2002

The ligase detection reaction: the evolution of a mutation detection strategy

Hannah D. Farquar
Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Part of the Chemistry Commons

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_dissertations/3855

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contactgradetd@lsu.edu.
THE LIGASE DETECTION REACTION: THE EVOLUTION OF A MUTATION DETECTION STRATEGY

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College
In partial fulfillment of the Requirements for the degree of Doctor of Philosophy

In

The Department of Chemistry

By
Hannah D. Farquar
B.S., Louisiana Tech University, 1996
December 2002
Acknowledgements

I would like to thank Dr. Robert P. Hammer for being patient enough to deal with my very steep learning curve and being an understanding colleague at the top of that curve. I would also like to thank Dr. Steven A. Soper for all of his teaching and support which has made me not only more successful but also a much better scientist. Dr. Kathleen Morden deserves much thanks for being a strong female scientist and a great role model.

Dr. Tod J. Miller, a postdoctoral researcher, has had an enormous impact on my life as a graduate student. He was an academic mentor as well as a great friend to me. He taught me how to think through problems yet never lose the big picture. This lesson has proven to be one of great importance as I have dug deeper into the research hole. Dr. Andrea S. Saurage was incredibly helpful in my first years as a graduate student teaching me every technique I needed to begin my graduate studies and continues to be an immense support sharing her experiences post-LSU.

I would also not have attained the level of satisfaction with my graduate studies had it not been for two very impressive undergraduate researchers with whom I have worked: Camille E. Tallini and Arsham Sheybani. They are not only two of the most intelligent people I have ever had the pleasure of meeting, but they are also two of the most open minded and well meaning human beings I believe I will ever come across. They are certain to be incredibly successful in every endeavor they attempt.

Finally, I must thank my husband, George R. Farquar, for being so immensely wonderful and understanding. George, you not only give me the best scientific
discussions and wonderfully patient answers to my not so well thought out questions, but you give me the security and sanity that is required to achieve the title of Ph.D. I can only hope that you understand how much of this dissertation as well as my life belong to you.
Table of Contents

Acknowledgments ......................................................................................................... ii

List of Tables ................................................................................................................ vii

List of Figures .............................................................................................................. viii

List of Schemes ............................................................................................................ xiv

List of Abbreviations ................................................................................................... xv

Abstract........................................................................................................................ xix

Chapter 1.  Introduction.................................................................................................1
  1.1 Mutations ............................................................................................................1
  1.2 PCR ...................................................................................................................6
  1.3 DNA Polymerases ..........................................................................................9
  1.4 Mutation Detection .....................................................................................15
    1.4.1 Scanning Methods ...........................................................................16
    1.4.2 Diagnostic Methods ......................................................................18
  1.5 References ....................................................................................................24

Chapter 2.  Nucleoside Base Analogs ..........................................................................29
  2.1 Introduction .................................................................................................29
    2.1.1 DNA Structure and Stability .................................................... 29
    2.1.2 DNA Polymerase Recognition .................................................. 36
    2.1.3 Nucleoside Base Analogs-The Literature ..................................37
      2.1.3.1 DNA Stability ....................................................................38
      2.1.3.2 DNA Polymerase Recognition ....................................... 43
  2.2 Pyrazole-Based Analogs .............................................................................46
    2.2.1 Introduction ............................................................................... 46
    2.2.2 Synthesis .................................................................................... 50
    2.2.3 Computational Analysis ........................................................... 54
    2.2.4 DNA Polymerase Recognition of Pyrazole Analogs ....... 60
  2.3 Thiazole and Thiazole N-oxide Analogs .................................................. 63
    2.3.1 Synthesis .................................................................................... 64
    2.3.2 Computational Analysis ........................................................... 70
    2.3.3 Thermodynamic Analysis .......................................................... 76
    2.3.4 DNA Polymerase Recognition ................................................ 78
  2.4 Conclusions ................................................................................................... 84
  2.5 References .................................................................................................... 85
Chapter 3. Nucleoside Backbone Analogs .................................................................92
  3.1 Introduction .........................................................................................................92
    3.1.1 Backbone Analogs-The Literature ..........................................................93
    3.1.2 Synthesis .....................................................................................................96
  3.2 Exonuclease Resistance of Phosphoramidate .................................................100
  3.3 Conclusions.......................................................................................................106
  3.4 References.........................................................................................................106

Chapter 4. Capillary and Microdevice Electrophoresis of LDR...............................110
  4.1 Introduction.......................................................................................................110
    4.1.1 LDR of Colorectal Cancer ......................................................................110
    4.1.2 Capillary and Microdevice Electrophoresis .........................................112
  4.2 Capillary Electrophoresis: Model Studies and LDR Reactions ......................114
    4.2.1 Introduction.............................................................................................114
    4.2.2 Model Studies .........................................................................................115
    4.2.3 LDR Reactions .......................................................................................116
  4.3 Microdevice Electrophoresis: Model Studies ..................................................122
  4.4 Conclusions......................................................................................................125
  4.5 References.......................................................................................................125

Chapter 5. Experimental Details ...............................................................................129
  5.1 Nucleoside Base Analogs-Chapter 2 ................................................................129
    5.1.1 Synthesis of 1-(β-D-2´-deoxyribofuranosyl)-4-(2-thiazolyl)pyrazole .......129
    5.1.2 Computational Analysis of Pyrazole-Based Analogs ..............................130
    5.1.3 DNA Polymerase Extension Reactions of Pyrazole Analogs ..........130
    5.1.4 Synthesis of Thiazole and Thiazole N-oxide Analogs .........................131
      5.1.4.1 Synthesis of 2-deoxy-3,5-di-O-p-toluoyl-β-D-erythro-pentofuranosyl cyanide (2) ..........................................................131
      5.1.4.2 Synthesis of 2,5-anhydro-3-deoxy-4,6-di-O-p-toluoyl-D ribo-hexothioamide (3) ..........................................................132
      5.1.4.3 Synthesis of 2-(2´-Deoxy-3´,5´-di-O-p-toluoyl-β-D-ribofuranosyl)thiazole (4) ..........................................................132
      5.1.4.4 Synthesis of 2-(2´-Deoxy-3´,5´-di-O-p-toluoyl-β-D-ribofuranosyl)thiazole-N-oxide (5) ..........................................................132
      5.1.4.5 Synthesis of 2-(2´-Deoxy-β-D-ribofuranosyl)thiazole(6) ..................134
      5.1.4.6 Synthesis of 2-(2´-Deoxy-5´-O-dimethoxytrityl-β-D-ribofuranosyl)thiazole (8) ..........................................................134
      5.1.4.7 Synthesis of 2-(2´-Deoxy-5´-O-dimethoxytrityl-β-D-ribofuranosyl)thiazole-3´-O-(2-cyanoethyl)-N,N-disopropylphosphoramidite (9) ..........................................................135
      5.1.4.8 Synthesis of 2-(2´-Deoxy-β-D-ribofuranosyl)thiazole-N-oxide (7) ..136
5.1.4.9 Synthesis of 2-(2’-Deoxy-5’-O-dimethoxytrityl-β-D-ribofuranosyl)thiazole-N-oxide (10) ..................................................136
5.1.4.10 Synthesis of 2-(2´-deoxy-5´-O-dimethoxytrityl-β-D-ribofuranosyl)thiazole-N-oxide-3´-O-(2-cyanoethyl)-N,N-diisopropylphosphoramidite (11) ...........................................137
5.1.4.11 Deprotection of DNA .....................................................137
5.1.5 Computational Analysis of Thiazole and Thiazole N-oxide Analogs .................................................................................138
  5.1.4.1 Semi-empirical Analysis ......................................................138
  5.1.4.2 Molecular Dynamics Analysis .............................................138
5.1.5 Thermodynamic Analysis of Thiazole and Thiazole N-oxide Analogs ..................................................................................139
5.1.6 DNA Polymerase Recognition of Thiazole and Thiazole N-oxide Analogs ...........................................................................140
5.2 Nucleoside Backbone Analogs-Chapter 3..................................................143
  5.2.1 Synthesis of N3’→P5’ Phosphoramidate Backbone Modification .................................................................143
  5.2.2 Exonuclease Resistance Assays ..............................................144
  5.2.3 Kinetic Exonuclease Resistance Assays ...............................146
5.3 Capillary and Microdevice Electrophoresis of LDR-Chapter 4.........146
  5.3.1 DNA sample preparation .......................................................146
  5.3.2 LDR Conditions ......................................................................146
  5.3.3 Instrumentation-CE ...............................................................147
  5.3.4 Instrumentation-Microdevice ................................................147
5.4 References ..................................................................................148

Appendix: Thermal Melting Raw Data and MALDI Mass Spectra ..........150

Vita .................................................................................................................177
List of Tables

Table 2.1  Thermodynamic values obtained from thermal melting analysis of the Dickerson dodecamer (5’-CGCXAATTYGCG-3’). PzI = iodopyrazole, PzN = nitropyrazole, PzP = propynylpyrazole, and PzT = thiazolylpyrazole. ........................................................................................................48

Table 2.2  Semiempirical (PM3) computational analysis (MM+ minimization) comparing the natural bases and pyrazole analogs. ......................57

Table 2.3  Semiempirical (PM3) computation analysis of thiazole and thiazole N-oxide nucleoside analogs. ..................................................71

Table 2.4  Thermodynamic properties of the Dickerson dodecamer (5’-CGCQAATTYGCG-3’) containing thiazole and thiazole N-oxide nucleoside analogs. ........................................................................................................77

Table 2.5  Kinetic parameters measured for single base insertion of dATP or dGTP opposite thymine, cytosine, and thiazole. .........................84

Table 3.1  Quantitation of percentages of extension and cleavage of each primer. .......................................................................................104

Table 5.1  Optimized parameters for kinetic incorporation analysis. ........142
List of Figures

Figure 1.1 The four natural bases found in DNA................................................2
Figure 1.2 Tautomerization of the natural bases and deamination of adenine and cytosine which lead to mutations........................................................4
Figure 1.3 Formation of T-T dimer and T-C dimer caused by UV irradiation and lead to mutations.................................................................6
Figure 1.4 Schematic representation of the polymerase chain reaction (PCR). The first step includes denaturing of the double stranded DNA, followed by annealing of primers in the second step, and finally extension of the primers by DNA polymerase in the final step. ......7
Figure 1.5 The nucleotidyl transfer mechanism catalyzed by DNA polymerase. The atoms in red indicate the carboxylate triad (carboxylate groups of Asp and Glu residues) which position the Mg ions for chemistry to occur.................................................................9
Figure 1.6 The conserved shape of all known DNA polymerases. Adapted from the file 1JXE of the Protein Data Bank (PDB). .....................................11
Figure 1.7 Space filling models of Taq polymerase in the open complex (top) and the closed complex (bottom). The cleft between the thumb and fingers decreases in size when the polymerase is ready to catalyze the nucleotidyl transfer reaction. Images adapted from files 2KTQ and 3KTQ of the Protein Data Bank (PDB). ..............................................13
Figure 1.8 The cleavage mechanism used by the exonuclease domain of DNA polymerases. The pink atoms indicate the hydroxide ion formed for nucleophilic attack of the phosphate phosphorous. The red atoms indicate the carboxylate oxygens of Asp and Glu residues located in the exonuclease active site. .................................................................14
Figure 1.9 Schematic representations of scanning methods.............................17
Figure 1.10 Schematic representations of diagnostic methods. .........................18
Figure 1.11 Schematic representation of the ligase detection reaction (LDR). PCR products are denatured and common and discriminating primers annealed. DNA ligase only ligates those duplexes which contain exact matches and thermocycling amplifies the amount of product formed linearly. ...............................................................................................21
Figure 1.12  Diagram of PCR step using nucleoside base analogs (Q) and the restriction endonuclease (RE) step of PCR/RE/LDR. PCR primers containing analogs are annealed to the DNA and extended. Natural primers are then annealed and extended forcing DNA polymerase to incorporate a natural base across from the analog. The incorporation forms an RE site which can then be cleaved leaving only mutant sequences to be detected by LDR. ....................................................23

Figure 2.1  The secondary structure of DNA indicating major and minor grooves on the left and helix width, pitch height, and distance between bases on the right.................................................................30

Figure 2.2  Syn and anti conformations of nucleosides. ..................................................31

Figure 2.3  Sugar puckering modes of furanose rings in DNA. The C2-endo conformation is preferred in DNA. .................................................................32

Figure 2.4  A. Systematic numbering of bases; B. The positions of hydrogen bond donors and acceptors positioned in the grooves of DNA; C. The geometry of natural base pairs: C1-C1 distances and C1-N bond angles. .................................................................................................33

Figure 2.5  A typical transition in absorbance seen as temperature is increased causing all double stranded DNA to form single strands........................................................................................................35

Figure 2.6  Structures of nucleoside base analogs discussed in Chapter 2 ......40

Figure 2.7  Structures of pyrazole based analogs................................................................47

Figure 2.8  Graphical representation of the stabilities of each of the pyrazole analogs reported earlier. All of the analogs are most stable when base paired with adenine (top) and of the analogs thiazolopyrazole is the most stable when paired with all of the natural bases (bottom). ..49

Figure 2.9  The possible syn/anti conformations of the pyrazole analogs and probable base pair geometries formed with adenine..........................51

Figure 2.10  Plots of -AA values obtained from thermodynamic analysis versus calculated surface area and dipole moment values of pyrazole analogs. ..................................................................................58

Figure 2.11  Plots of –ΔΔG values obtained from thermodynamic analysis versus calculated polarizability value of the pyrazole analogs (top) and of the natural bases (bottom)..........................................................59
Figure 2.12  The actions performed by DNA polymerase during template directed primer extension. X represents the base in the primer. X can be either a natural base or an analog. Y and Z represent bases that are being added to the primer strand by DNA polymerase. These bases can also be either natural or analog bases. The DNA polymerase makes specific and separate contacts with X, Y, and Z during replication ..........................................................60

Figure 2.13  Extension assays of primers on templates containing thymine (control), nitropyrazole (PzN), and thiazolylpyrazole (PzT). The primer is 18nt while the template is 25nt in length. Extension reactions were performed in 20 µL with 250 nM radiolabeled primer and template, and 200 µM solution of all four dNTPs, AmpliTaq® polymerase and the supplied buffer (100 mM Tris-HCl pH 8.3, 500 mM KCl). An additional 8 mM MgCl₂ was added to the reaction. ..........................................................62

Figure 2.14  Structures of tiazofurin, thiazole, and thiazole N-oxide base analogs. .................................................................63

Figure 2.15  ORTEP diagrams of p-toluoyl protected thiazole analog (A) and p-toluoyl protected thiazole N-oxide analog (B). .........................66

Figure 2.16  HPLC trace of Dickerson dodecamer containing thiazole (A) and thiazole N-oxide (B). HPLC was performed using a Hydrocell NS 1000 (anion-exchange) column (150 x 4.6 mm column) with Solvent A = 25 mM CHES pH 8, 30% methanol and Solvent B = 25 mM CHES, 1 M (NH₄)₂SO₄, pH 8, 30% methanol (0-100% B in 60 min with a flow rate = 1 mL/ min). .........................................................69

Figure 2.17  HPLC trace of the 28nt template containing thiazole N-oxide. HPLC was performed using a DNA Pac-100 (anion-exchange) column with Solvent A = 1 M NaCl, 25 mM NaOH (pH 12.4) and Solvent B = 25 mM NaOH (pH 12.4) (100-500 mM NaCl in 2 min, 500-900 mM NaCl in 18 min, and 900-1000 mM NaCl in 3 min with a flow rate = 6 mL/min). ..........................................................69

Figure 2.18  Calculated partial atomic charges and electrostatic potential maps of cytosine, thymine, thiazole, and thiazole N-oxide. In the electrostatic potential maps blue indicates a negative charge while red indicates a positive charge. SYBYL MOPAC calculations at the PM3 were performed and the electrostatic potential was mapped onto a Connolly surface using these charges.................................73

Figure 2.19  Molecular Dynamics structures of the Dickerson dodecamer containing a cytosine-guanine base pair at the arrow (A) and a
thiazole N-oxide-guanine base pair at the arrow (B). Molecular Dynamics simulations were performed by ramping to 300 K or 400 K for 1 ps and then stepping down to zero. These simulations used a 1 fs time step and constant pressure. SHAKE was used to constrain all hydrogens except those of the thiazole N-oxide-guanine base pair.

Figure 2.20 Cytosine-guanine (A) and thiazole N-oxide-guanine (B) base pairs extracted from the Molecular Dynamics structures in Figure 2.19.

Figure 2.21 Extension reactions performed with templates containing thymine (T), cytosine (C), thiazole (Z), and thiazole N-oxide (X). The percentages directly under the electrophoresis lanes correspond to the percentage of full-length product. The percentage under the full-length product percentage is the percentage of stall product (24nt for the 23nt primer and 23nt for the 21nt primer). The third percentage in the case of the 21nt primer is the percentage of stall product formed at 24nt. Extension reactions were performed in 20 µL with 200 nM radiolabeled primer and template, and 200 µM solution of all four dNTPs, polymerase and the supplied buffer. (Taq: 100 mM KCl, 100 mM (NH₄)₂SO₄, 200 mM Tris-Cl, 20 mM MgSO₄, 1% Triton X-100 (pH 8.8). An additional 8 mM MgCl₂ was added to the reaction. Kf(exo-): 100 mM Tris-Cl pH 7.5, 5 mM MgCl₂, 7.5 mM DTT.)

Figure 3.1 Structure of the phosphodiester (A), phosphorothioate (B), and phosphoramidate (C) backbone modifications.

Figure 3.2 The anomeric effect caused by the addition of the 3´-nitrogen of the phosphoramidate modification. The lone pair of the nitrogen interacts with the σ* antibonding orbital of the P-O5´ bond. The structure is stabilized by the chlorine ion and the C₃-endo furanose conformation.

Figure 3.3 Sequences of the primers as well as the match and mismatch templates used in the assay to test the exonuclease resistance of the phosphoramidate modification.

Figure 3.4 The assay designed to test the exonuclease resistance of the phosphoramidate modification.

Figure 3.5 Electrophoretic gels of the assay to test the exonuclease resistance of the phosphoramidate modification extended by A. Taq, B. Vent, or C. Pfu DNA polymerase. O=phosphodiester; S=phosphorothioate, N=phosphoramidate. Extension reactions were performed with 20 µL...
pM primer, 50 pM template, and 200 µM solution of all four dNTPs, AmpliTaq®, Vent, and Pfu polymerase and the supplied buffers. Extension of the match template was digested with BslI restriction endonuclease (supplied buffer: 1 M NaCl, 500 mM Tris-HCl, 100 mM MgCl2, 1 mM DTT, pH 7.9) at 55 ºC for 60 min. Extension of the mismatch template was digested with TaqI restriction endonuclease (supplied buffer: 1 M NaCl, 100 mM Tris-HCl, 100 mM MgCl2, pH 8.4) at 65 ºC for 60 min. 

Figure 3.6 Electrophoretic gels of the timed analysis of exonucleolytic cleavage of the phosphodiester (O), phosphorothioate (S), and phosphoramidate (N) primers. Reactions were performed as extension reactions as described in Figure 3.5 without the addition of dATP and dGTP. 

Figure 4.1 Locations of codons 12, 13, and 61 in the K-ras gene. 

Figure 4.2 Schematic representation of the ligase detection reaction (LDR). 

Figure 4.3 The twin T microdevice. A is the waste well, B is the sample well, C is the buffer reservoir, and D is the separation well. 

Figure 4.4 Analysis of 1:100 model sample using the POP5 matrix. The samples were run at 12 kV and capillary temperature was controlled at 50 ºC. 

Figure 4.5 Analysis of the 1:100 model sample using the eCap matrix. Samples were run at 12 kV at 30 ºC. 

Figure 4.6 5T5C analysis of the A) 1:100 model sample and B) 1:1000 model sample. Samples were run at 250 V and 25 ºC. 

Figure 4.7 LDR primer sequences used in reactions for CE analysis. 

Figure 4.8 LDR blank sample for CE analysis. Samples were run as described in Figure 4.6. 

Figure 4.9 A)1:20 and B) 1:100 mutant to wild type LDR reactions for CE analysis. Samples were run as described in Figure 4.6. 

Figure 4.10 Modified twin T design used in the current described analysis. A is the sample well, B the waste well, C the buffer reservoir and D the separation well. During injection a positive voltage was placed on waste well B while the sample well A was grounded. During separation a positive voltage was applied to well D while C was
grounded and positive voltages of 15 and 10% of the voltage applied to D was applied to the sample well A and waste well B. ...........122

Figure 4.11  Microdevice analysis of a 1:1 model sample of the 25-nucleotide and 51-nucleotide oligonucleotides. POP5 was the matrix used and no pullback voltages were applied during separation. Injection was at 500 V for 5 s and separation was at 3.5 kV. .................................124

Figure 4.12  Microdevice analysis of a 1:1 model sample of the 25-nucleotide and 51-nucleotide oligonucleotides. POP5 was the matrix used and pullback voltages were applied during separation. Injection was at 500 V for 5 s and separation was at 3.5 kV. Pullback voltages on the sample and waste wells during separation were 500 and 300 V, respectively. ..............................................................124

Figure 4.13  Microdevice analysis in POP5 of the 1:100 model sample. Samples were separated as described in Figure 4.12. .................................125
List of Schemes

Scheme 2.1  Synthesis of the pyrazole analogs. ..........................................................53

Scheme 2.2  Improved synthesis of thiazolylpyrazole analog. ..............................54

Scheme 2.3  DNA synthesis by the cyanoethyl phosphoramidite approach.  TCA =
             trichloroacetic acid, DMT = dimethoxytrityl, CPG = controlled pore
glass solid support. ..................................................................................55

Scheme 2.4  Synthesis of thiazole and thiazole N-oxide analogs. ...........65

Scheme 2.5  Synthesis of thiazole and thiazole N-oxide phosphoramidites......67

Scheme 2.6  Proposed mechanism for the deoxygenation of thiazole N-oxide
             containing oligonucleotides during DNA synthesis......................71

Scheme 3.1  Reduction of 3´-DMT azidothymine..................................................97

Scheme 3.2  Formation of the phosphoramidate linkage by solid phase
             synthesis. ......................................................................................98

Scheme 3.3  Solid phase DNA synthesis by the cyanoethyl phosphoramidite
             approach. ......................................................................................99
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>5T5C</td>
<td>5% crosslinked acrylamide capillary for CE</td>
</tr>
<tr>
<td>A</td>
<td>Deoxyadenosine</td>
</tr>
<tr>
<td>Å</td>
<td>Angstrom</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>Bz</td>
<td>Benzoyl</td>
</tr>
<tr>
<td>C</td>
<td>Deoxycytidine</td>
</tr>
<tr>
<td>CE</td>
<td>Capillary electrophoresis</td>
</tr>
<tr>
<td>CH$_2$Cl$_2$</td>
<td>Dichloromethane</td>
</tr>
<tr>
<td>CH$_3$OH</td>
<td>Methanol</td>
</tr>
<tr>
<td>CHCl$_3$</td>
<td>Chloroform</td>
</tr>
<tr>
<td>CHES</td>
<td>2-(N-cyclohexylamino)ethanesulfonic acid</td>
</tr>
<tr>
<td>CPG</td>
<td>Controlled pore glass</td>
</tr>
<tr>
<td>dATP</td>
<td>Deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>DCM</td>
<td>Dichloromethane</td>
</tr>
<tr>
<td>dCTP</td>
<td>Deoxycytidine triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>Deoxyguanosine triphosphate</td>
</tr>
<tr>
<td>DIPA</td>
<td>Diisopropylamine</td>
</tr>
<tr>
<td>DMAP</td>
<td>4,4-Dimethylaminopyridine</td>
</tr>
<tr>
<td>DME</td>
<td>Dimethoxyethane</td>
</tr>
<tr>
<td>DMT</td>
<td>Dimethoxytrityl</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>DMTCl</td>
<td>Dimethoxytrityl chloride</td>
</tr>
<tr>
<td>DMTr</td>
<td>Dimethoxytrityl</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleoside triphosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>Deoxythymidine triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>eCap</td>
<td>ssDNA 100-R replaceable gel matrix for CE</td>
</tr>
<tr>
<td>EDITH</td>
<td>3-Ethoxy-1,2,4-dithiazoline-5-one</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>equiv</td>
<td>Equivalent</td>
</tr>
<tr>
<td>Et₃N</td>
<td>Triethylamine</td>
</tr>
<tr>
<td>EtOAc</td>
<td>Ethyl acetate</td>
</tr>
<tr>
<td>EtOH</td>
<td>Ethanol</td>
</tr>
<tr>
<td>G</td>
<td>Deoxyguanosine</td>
</tr>
<tr>
<td>∆G</td>
<td>Gibbs free energy</td>
</tr>
<tr>
<td>h</td>
<td>Hour</td>
</tr>
<tr>
<td>∆H</td>
<td>Enthalpy</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>HDA</td>
<td>Heteroduplex analysis</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>Ib</td>
<td>Isobutyryl</td>
</tr>
<tr>
<td>KCl</td>
<td>Potassium Chloride</td>
</tr>
<tr>
<td>Kf</td>
<td>E. coli DNA polymerase I</td>
</tr>
</tbody>
</table>
Kf(exo-)

E. coli DNA polymerase I without the exonuclease domain

Km

Michaelis constant

LDR

Ligase detection reaction

m.p.

Melting point

MALDI

Matrix assisted laser desorption ionization mass spectrometry

mCPBA

m-Chloroperoxybenzoic acid

MeOH

Methanol

MgCl₂

Magnesium chloride

MgSO₄

Magnesium sulfate

min

Minute

mRNA

Messenger RNA

Na₂CO₃

Sodium carbonate

Na₂PO₄

Sodium phosphate

Na₂S₂O₃

Sodium thiosulfate

NaCl

Sodium chloride

NaHCO₃

Sodium bicarbonate

NaOH

Sodium hydroxide

NH₄OH

Ammonium hydroxide

(NH₄)₂SO₄

Ammonium sulfate

NMR

Nuclear Magnetic Resonance

nt

Nucleotide

PCR

Polymerase Chain Reaction

PDB

Protein Database
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pfu</em></td>
<td><em>Pyrococcus furiosus</em> DNA polymerase</td>
</tr>
<tr>
<td>POP5</td>
<td>Performance Optimized Polymer replaceable gel matrix for CE</td>
</tr>
<tr>
<td>RE</td>
<td>Restriction endonuclease</td>
</tr>
<tr>
<td><em>Rf</em></td>
<td>Retention factor</td>
</tr>
<tr>
<td>RMS</td>
<td>Root mean square fit</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td><em>s</em></td>
<td>Second</td>
</tr>
<tr>
<td>ΔS</td>
<td>Entropy</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td><em>T</em></td>
<td>Deoxythymidine</td>
</tr>
<tr>
<td><em>Taq</em></td>
<td><em>Thermus aquaticus</em> DNA polymerase</td>
</tr>
<tr>
<td>TCA</td>
<td>Trichloroacetic acid</td>
</tr>
<tr>
<td>TFAA</td>
<td>Trifluoroacetic anhydride</td>
</tr>
<tr>
<td>THF</td>
<td>Tetrahydrofuran</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin Layer Chromatography</td>
</tr>
<tr>
<td><em>Tm</em></td>
<td>Melting temperature</td>
</tr>
<tr>
<td>Tol</td>
<td><em>p</em>-Toluoyl</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>U</td>
<td>Unit</td>
</tr>
<tr>
<td>UHP</td>
<td>Urea-hydrogen peroxide complex</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>Vent</td>
<td><em>Thermococcus litoralis</em> DNA polymerase</td>
</tr>
<tr>
<td><em>V</em>&lt;sub&gt;max&lt;/sub&gt;</td>
<td>Velocity measured at <em>K</em>&lt;sub&gt;m&lt;/sub&gt;</td>
</tr>
</tbody>
</table>
Abstract

Early detection of genetic mutations is important for control of diseases such as cancer and Alzheimer’s. Early detection requires methods that detect small amounts of mutated DNA in very large amounts of normal or wild type DNA. One method to detect mutated DNA is the ligase detection reaction (LDR). Since its inception, LDR has evolved greatly from a simple detection reaction after PCR amplification to PCR/RE/LDR, a scheme that uses nucleoside base analogs in PCR to convert wild type sequences to sequences containing restriction endonuclease (RE) sites which can then be cleaved leaving only mutant sequences for detection by LDR. Analysis of LDR has also evolved from slab gel electrophoresis to microarray analysis.

Understanding the structure and DNA polymerase recognition of nucleoside base analogs used in PCR/RE/LDR is key to improving this detection scheme. The use of higher fidelity DNA polymerase containing 3’→5’ exonuclease domains for error correction is also important in early detection of genetic diseases. Pyrazole-based nucleoside analogs have been studied computationally and enzymatically. The stability a DNA containing these analogs depends largely on the dipole moment of the analogs, rather than polarizability or surface area. Reduced DNA polymerase recognition is due in part to altered base pair geometry, either inherent or created by DNA polymerase. Thiazole and thiazole N-oxide analogs to be used in the PCR/RE/LDR assay have been synthesized and characterized computationally, thermodynamically, and enzymatically. The N-oxide, a pyrimidine O2 mimic, enhances DNA stability and DNA polymerase recognition. The N-oxide increases electrostatic properties and solvation by the formation of a hydrogen bond when base paired with guanine. Enzymatic analysis
indicated a preference for the base pairing of thiazole \( N \)-oxide with guanine and thiazole with adenine. An N3→P5’ phosphoramidate backbone analog has shown to inhibit the exonuclease activity of higher fidelity DNA polymerases for use in PCR/RE/LDR.

The evolution of the analysis of LDR continues with the adaptation to capillary and microdevice electrophoresis. These formats were used to analyze model samples and LDR reactions mimicking low abundant mutations. These improved techniques greatly improve the resolution of LDR analysis.
Chapter 1: Introduction

1.1 Mutations

Deoxyribonucleic acid or DNA carries the genetic information of the cell. This information is processed within the cell during replication and transcription. During replication an enzyme, DNA polymerase, copies the information and during transcription an enzyme, RNA polymerase, transcribes the DNA into messenger ribonucleic acid, mRNA which is then further processed during translation into proteins which are the machines of the cell. Transcription only occurs on portions of DNA that are genes. The locations of transcription are governed by specific sequences termed promoters. Portions of genes that are not transcribed are termed introns while transcription of DNA occurs on other portions of DNA termed exons. The process, DNA → RNA → PROTEINS, is what governs the operation of most forms of life including humans. The information of DNA is stored within a sequence which is made up of four nucleotides, each containing a different aromatic base: adenine (A), thymine (T), guanine (G), and cytosine (C) (Figure 1.1). Any error, or mutation, that occurs in the strings of nucleic acids may cause the above-described process to fail. Therefore, the study and detection of mutations is very important to understanding the links between genetics and disease.

Mutations are classified into many different categories based on the protein products formed after transcription and translation. Some mutations that occur can be silent. During translation codons, groups of three bases, code for amino acids. If the amino acid that is encoded does not change, even though the codon sequence is changed, the mutation that occurred is termed a silent mutation. For example, the codons CGC, CGA, and AGG are degenerate, they all code for arginine. Therefore, if a mutation occurs forming any of these three codons arginine is still encoded and there is no change in the protein sequence that is synthesized. Another
example of a silent mutation is one in which the mutation causes a change in a codon that changes the amino acid coded for yet the changed amino acid within the sequence of a protein does not change the function of the protein. For instance, if the amino acid is on the outside of the protein and not contained within the active site the protein will still function normally.

![Image of DNA bases]

**Figure 1.1.** The four natural bases found in DNA.

The types of mutations that are not silent are classified into frameshift mutations and point mutations. Frameshift mutations occur when a portion of the DNA sequence is deleted or another sequence is inserted during replication or transcription. These types of mutations lead to a complete change of codon sequences throughout the length of the exon, consequently frequently leading to a complete loss of protein synthesis. Point mutations are characterized by the substitution of a single base for another at a specific location. If a purine is substituted for another purine or a pyrimidine for another pyrimidine (e.g. A for G or T for C) this mutation is termed a transition. If a pyrimidine is substituted for a purine or a purine for a pyrimidine (e.g. A for T or C for G) a transversion occurs. Point mutations are distinguished as either nonsense or missense in nature. Nonsense mutations change a codon to a stop codon while missense mutations cause amino acid changes that change the function of a protein.

Mutations that occur within the genome can be spontaneous or can be induced. Spontaneous mutations result from the natural process of the cells while induced mutations are
caused by the interaction of DNA with an outside agent or mutagen.\textsuperscript{1,2} The most important and frequent cause of spontaneous mutation is errant DNA replication. If a DNA polymerase makes a “mistake” and inserts an incorrect base, and this base is not corrected, the mutation is then conserved throughout further rounds of replication and transcription therefore affecting the sequence of the mRNA produced and the protein synthesized. The rate of this type of mutation and the mechanisms of correction will be discussed later. Spontaneous mutations are also seen during replication in locations within a DNA duplex that contain long “runs” or repeats of nucleotides. The DNA polymerase tends to “slip” and therefore repeat many bases during replication consequently causing gene expansion.\textsuperscript{1,3,1,4} Spontaneous mutations can also be caused by natural structural alterations of the bases and by damage to the bases caused by metabolites of other cellular processes.\textsuperscript{1,2} One structural change that has been proposed to cause mutations is the tautomerization of the bases (Figure 1.2). Guanine and thymine favor the keto form, but the enol form can occur. Adenine and cytosine bases favor an amino form, but can exist in an imino form. These changes lead to different pairing trends among the bases. For example, when guanine is in the enol form a G:T base pair is favorable rather the normal G:C base pair causing a point mutation. Other structural changes that can occur include the deamination of cytosine to uracil (Figure 1.2), the oxidative deamination of adenine to hypoxanthine (Figure 1.2), and the alkylation of bases. When cytosine is deaminated to uracil an A is coded for rather than a G. Hypoxanthine encodes for C rather than the T encoded by adenine. As a result deamination causes point mutations to occur within the genome. Alkylation can also cause pairing aberrations, but can also cause spontaneous breakdown of the DNA molecule. Free radical damage to DNA bases is a cause of spontaneous mutation that has
TAUTOMERIZATION OF THE NATURAL BASES

DEAMINATION OF ADENINE AND CYTOSINE

Figure 1.2. Tautomerization of the natural bases and deamination of adenine and cytosine which lead to mutations.
received much attention. Free radicals are produced as a result of oxidative metabolism and can oxidize the bases to cause mispairing.

Natural or man-made chemicals that alter DNA bases are termed mutagens. Some DNA base analogs that mimic the DNA natural bases and are used in research and for other purposes can be mutagens by being incorporated into the DNA duplex during replication causing the formation of point mutations during the following rounds of replication or transcription. These types of base analogs include bromouracil and aminopurine. Mutagens can also be chemicals that affect the structure of the natural bases therefore causing mispairing. For example, nitrous acid causes the deamination of cytosine and adenine while methyl methanesulfonate and ethyl methanesulfonate are alkylating agents. Intercalating dyes such as acridine orange, proflavin, and ethidium bromide are also mutagens due to their ability to insert between the bases of a DNA strand causing the DNA polymerase to insert an extra base across from the intercalating agent.

Ionizing radiation (X-rays and gamma rays) and ultraviolet (UV) light also cause mutations. Ionizing radiation is energetic enough to cause the formation of reactive ions when reacting with biological molecules. This type of radiation also causes the formation of free radicals, formed from water molecules, and peroxide formed during cellular respiration that can then act as mutagens. UV radiation is less energetic than ionizing radiation but the bases of DNA preferentially absorb these wavelengths. The most common and most lethal mutations caused by UV radiation are the pyrimidine-pyrimidine dimers (Figure 1.3). These dimers are formed by the covalent attachment of pyrimidines in the same strand of DNA. These dimers are very bulky therefore drastically changing the secondary structure of a DNA duplex causing the cessation of DNA replication and transcription completely.
Figure 1.3. Formation of T-T dimer and T-C dimer caused by UV irradiation and lead to mutations.

While there are repair pathways for most of these mutations, not all are repaired and are therefore left to replicate.¹² Mutations that are not repaired can then possibly lead to disease if they are located within the coding region of a gene. Many diseases, such as diabetes, cancer, asthma, myocardial infarction, atherosclerosis, cystic fibrosis, Alzheimer’s and sickle cell anemia, are associated with genetic alterations in genomic DNA (www.ncbi.nlm.nih.gov/disease).

1.2 PCR

The polymerase chain reaction (PCR) has become the most widely used technique for DNA analysis and is thus the centerpiece of most mutation detection strategies. PCR is used in most detection methods detecting mutations. Most methods used are not sensitive enough to detect genomic DNA therefore sections containing mutations to be analyzed must be amplified by PCR prior to analysis. The methods for mutation detection will be discussed later.
PCR: Polymerase Chain Reaction

**Denaturation**
"melting"
95°C

**Annealing**
54°C

**Extension**
72°C

The three steps are cycled for an exponential increase in the amount of DNA

**Figure 1.4.** Schematic representation of the polymerase chain reaction (PCR). The first step includes denaturing of the double stranded DNA, followed by annealing of primers in the second step, and finally extension of the primers by DNA polymerase in the final step.

PCR was invented in 1983 by Dr. Kary B. Mullis for which he was awarded the Nobel Prize in Chemistry ten years later. The three steps of PCR which are cycled can be seen in Figure 1.4. First, a double helical strand of DNA is separated or denatured with high temperature to form two single strands of DNA. Secondly, primers, or shorter lengths of DNA, are annealed to the two longer strands or templates. Two different primers are used to bracket the target sequence to be amplified. One primer is complementary to one strand at the beginning of the target sequence while the second primer is complementary to the other strand at the end of the sequence. Finally, a DNA polymerase enzyme and deoxyribonucleotide triphosphates
(dNTPs) extend the primers. These three steps are repeated or cycled and the target sequence amplified exponentially.

The DNA polymerase catalyzes the addition of the incoming deoxyribonucleotide triphosphate (dNTP) to the primer to form a phosphodiester linkage (Figure 1.5). For PCR to work a thermostable polymerase must be used in order to withstand the high temperatures required for the amplification cycles. Before PCR the DNA polymerases most widely used for research purposes were *E. coli* Polymerase I large fragment, otherwise known as Klenow fragment (Kf), and T7 DNA polymerase, which extend efficiently at 37 ºC. The first thermostable DNA polymerase discovered was *Thermus aquaticus* (Taq). Since *Taq* many more thermostable polymerases have been discovered: *Theratoga maritima* (Tma), *Pyrococcus furiosus* (Pfu), *Thermus thermophilus* (Tth), *Pyrococcus species* GB-D (Deep Vent), and *Thermococcus litoralis* (Tli or Vent). All of these enzymes have optimal activity at temperatures from 60 ºC-80 ºC, which makes them all useful in PCR. One distinguishing characteristic is the presence or absence of an exonuclease domain. The exonuclease domain of a polymerase allows correction or proofreading of extension mistakes or nucleoside mismatches. This characteristic contributes highly to the fidelity or efficiency of the enzyme. The existence of an exonuclease domain increases the fidelity therefore decreasing the number of mistakes that occur in the extension of a template.

The DNA polymerase is the most significant element of PCR and DNA replication, as mentioned in Section 1.1. The mechanism by which DNA polymerase replicates DNA is of utmost importance to understand the formation of most spontaneous mutations as well as to improve mutation detection methods as most rely on PCR.
Figure 1.5. The nucleotidyl transfer mechanism catalyzed by DNA polymerase. The atoms in red indicate the carboxylate triad (carboxylate groups of Asp and Glu residues) which position the Mg ions for chemistry to occur.

1.3 DNA Polymerases

DNA polymerases function in many ways within the cell. By catalyzing the nucleophilic attack of the free 3′-hydroxyl of the primer on the phosphorus of the α-phosphate of the incoming dNTP it is able to function in DNA replication, repair, and recombination (Figure 1.5).
The many different polymerases within an organism differ slightly in structure yet share the aforementioned functions. The known polymerases are classified according to structural differences, although all known polymerases are thought to have the same basic mechanism of action and similar overall structure, especially sharing the shape of the polymerase domain. The family of polymerases of interest in this document is the Pol A or Pol I family, which includes bacterial polymerases such as *E. coli* Pol I, *Taq*, and bacteriophage T7 DNA polymerase. Figure 1.6 is the crystal structure of *Taq* polymerase with the different subdomains of the structure labeled. The polymerase domain of such polymerases resembles an open right hand (Figure 1.6). The palm region is the subdomain that contains the active site. The thumb and fingers subdomains flank the palm and contain regions which interact with the template and dNTP and binds duplex DNA upstream of the 3’ end of the primer strand, respectively.

The DNA polymerase is an amazing “machine” when accomplishing template-directed polymerization of nucleic acids. The polymerase must be able to chemically recognize and bind a primed DNA strand, recognize and bind a correct dNTP dictated by the template, catalyze a nucleotidyl transfer reaction, and reposition the newly extended primer for further cycles of nucleotidyl transfer. The overall mechanism of DNA polymerase is seen below:\(^{1,17-20}\):

\[
\begin{align*}
\text{Pol} & \leftrightarrow \text{Pol} \cdot \text{DNA} \\
& \leftrightarrow \text{Pol} \cdot \text{DNA} \cdot \text{dNTP} \\
& \leftrightarrow \text{Pol}^* \cdot \text{DNA} \cdot \text{dNTP} \\
& \leftrightarrow \text{Pol}^* \cdot \text{DNA}_{+1} \cdot \text{PPi} \\
& \leftrightarrow \text{Pol} \cdot \text{DNA}_{+1} \cdot \text{PPi} \\
& \leftrightarrow \text{Pol} \cdot \text{DNA}_{+1} + \text{PPi} \\
& \rightarrow \text{Pol} \cdot \text{DNA}_{+1}
\end{align*}
\]

Pol = DNA polymerase; Pol* = conformational change; PPi = pyrophosphate

In the first step, the DNA polymerase binds the primed DNA template, which is then followed by binding of the dNTP. The next step is a conformational change of the polymerase, which is the rate-limiting step because this is dependent on the correct dNTP being bound. The
Conformational change consists of the fingers subdomain moving in towards the palm subdomain forming a “closed” complex.\textsuperscript{1,21} This change can be seen in Figure 1.7. Once the conformational change occurs the chemistry of nucleotidyl transfer occurs very quickly. After transfer is complete, the conformation of the enzyme is returned to its original form, “open”, which is also a very slow step, followed by release of pyrophosphate and a repositioning of the DNA for further rounds of polymerization.

\textbf{Figure 1.6.} The conserved shape of all known DNA polymerases. Adapted from the file 1JXE of the Protein Data Bank (PDB).
When selecting the dNTP for incorporation, the polymerase binds dNTPs in between the fingers and thumb subdomains. The 3’-OH of the primer is held close to this junction forming a tight fit. This restricted area limits the orientations of the incoming dNTP to certain geometries, ensuring that selection of the correct dNTP occurs through recognition of base pair geometry. Once the correct dNTP is chosen, the closed conformation is formed, signaling nucleotidyl transferase chemistry to commence.

The active site contains several acidic and polar amino acid residues as well as two metal cations (usually Mg$^{2+}$) that position the components of the reaction. One metal ion binds the 3’-OH of the primer and the $\alpha$-phosphate of the incoming dNTP. The positions are such that the pKa of the 3’-OH of the primer is lowered to favor the formation of a hydroxide anion to facilitate nucleophilic attack. The second ion binds all of the phosphate groups of the dNTP helping in alignment so the reaction can proceed and for stabilization of the pyrophosphate product formed.

Processivity is a characteristic that varies among polymerases. Some polymerases perform many nucleotidyl transfer reactions per binding event, therefore making them much more processive than those that only perform one or a few. Residues in the thumb subdomain make many contacts with duplex DNA blocking the DNA from completely dissociating from the Pol·DNA complex. Therefore, the more contacts the greater the processivity of the polymerase.

If an incorrect dNTP is incorporated, the exonuclease domain is able to correct the mistake. Not all polymerases have an exonuclease domain, but those that do exhibit much higher fidelities. The exonuclease domain can reduce the error rate of replication by approximately four orders of magnitude. The polymerase domain is approximately 30-40 Å
Figure 1.7. Space filling models of Taq polymerase in the open complex (top) and the closed complex (bottom). The cleft between the thumb and fingers decreases in size when the polymerase is ready to catalyze the nucleotidyl transfer reaction. Images adapted from files 2KTQ and 3KTQ of the Protein Data Bank (PDB).
away from the exonuclease domain: therefore, shuttling of the single stranded primer must take
place for exonucleolytic cleavage to occur.\textsuperscript{1,22,1,26} This happens when the fingers and thumb
subdomains that contact the final base pair in the primer-template duplex sense distortions
caused by mismatches. This causes a stall in the change to a closed conformation, which gives
the primer and template time to melt, or separate. Then, the primer strand is transferred to the
exonuclease domain where phosphodiester cleavage occurs through a two metal ion mechanism
similar to that described for polymerization only in reverse (Figure 1.8). One metal ion is bound
to the carboxylates of amino acid side chains and two oxygens of the phosphodiester linkage, one
bridging oxygen, and one non-bridging oxygen. The second metal ion is bound to carboxylates,
the same non-bridging oxygen, and a hydroxide ion formed from a water molecule. This
hydroxide ion is in position for nucleophilic attack onto the phosphorous atom of the backbone.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.jpg}
\caption{The cleavage mechanism used by the exonuclease domain of DNA polymerases. The pink atoms indicate the hydroxide ion formed for nucleophilic attack of the phosphate phosphorous. The red atoms indicate the carboxylate oxygens of Asp and Glu residues located in the exonuclease active site.}
\end{figure}
1.4 Mutation Detection

Identifying particular genetic mutation(s) is very important in order to utilize the benefits of presymptomatic diagnosis and to individualize the treatment. For example, cancer is the second leading cause of death in the United States, comprising 23% of all deaths. Many forms could be prevented with early detection of the genetic alterations that cause the disease, which are acquired early in the disease and can be conserved throughout its progression. Detection of genetic predispositions to, for example, heart disease and cancer, would revolutionize clinical diagnosis and permit significant progress in preventative medicine. Once the molecular basis of any disease is understood, analysis of the relevant genetic material requires robust tools to uncover its presence. Focus will be placed upon the detection of point mutations although some methods described can also be used for detecting frameshift mutations. An important problem in detecting point mutations is the occurrence of low abundant mutations or mutations that occur in low concentrations in the early development of a disease. Techniques that are sensitive enough to detect low concentrations of mutations in very high concentrations of normal (wild type) DNA are a requirement for the early detection of disease.

Point mutations that are associated with certain diseases may fall consistently in one position or several different positions along a large span of a gene. For this reason two basic types of methods for mutation detection have been developed, scanning and diagnostic techniques. Diagnostic methods are employed when the location of the mutation of interest is already characterized, while scanning or screening methods are used for an unknown location. While scanning methods potentially allow for the detection of all mutations in a specific region of DNA, the presence of the mutation(s) is the only information obtained and further analysis is required to obtain the specific genotype. Diagnostic methods detect an already characterized
mutation and give the specific sequence composition of the mutation. Although direct DNA sequencing (determining the primary structure of the DNA fragment under investigation) is the “gold standard” for diagnostics, the immense time and labor required to directly sequence DNA has lead to the development of alternative methodologies. In addition, direct sequencing cannot determine the presence of low abundant mutations and is generally ineffective at analyzing heterozygous samples containing both normal and mutant DNA. The detection of low abundant mutations has become a major goal of mutation detection for early detection of disease or the recurrence of disease. A brief synopsis of some detection techniques follows.

1.4.1 Scanning Methods

One of the first scanning methods developed was restriction fragment length polymorphism (RFLP). RFLPs consists of fragments of DNA produced after restriction enzyme sites found in the target are cut by restriction endonucleases that recognize short sequences. Point mutations, which must occur in the cutting site of the restriction endonuclease, change the pattern of fragments produced by these polymorphisms allowing separation and analysis via electrophoresis to uncover the presence of the mutation. The technique is limited in that the mutation must occur at an existing restriction site or an additional restriction site must be created by the mutation.

Many other scanning methods utilize the fact that changes in the properties of DNA occur when a mismatch is present in the duplex due to the presence of the mutation. These properties include modifications in electrophoretic mobilities and/or changes in duplex stability. Heteroduplex analysis (HDA) is a detection strategy that detects mutant DNA molecules by thermal denaturation and reannealing to form homo- and heteroduplexes of wild type and mutant alleles that have different electrophoretic mobilities due to bulges or thermodynamic instabilities
HDA can also be analyzed using denaturing high performance liquid chromatography (DHPLC). Denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis (TGGE), and constant denaturing gel electrophoresis (CDGE) are methods that enhance the unique electrophoretic mobilities of duplexed DNA containing mismatches through their reduced thermodynamic stability. Single strand conformation polymorphism (SSCP) utilizes assay conditions, such as the addition of denaturants that are conducive to the formation of different secondary structures of single stranded DNA. Differing conformations of wild type and mutant DNA, which can be sorted by electrophoresis, exhibit sequence specific electrophoretic mobilities. Figure 1.9 contains schematic representations of RFLP, HDA, and SSCP.

**Figure 1.9.** Schematic representations of scanning methods.
1.4.2 Diagnostic Methods

The majority of diagnostic techniques use, as a first step in the assay, amplification via the polymerase chain reaction (PCR). PCR can be done either non-specifically, in which primers flank the coding region containing the mutation(s) or specifically, in which allele specific primers that hybridize to the mutation site are used in the PCR step. If allele specific PCR is not used, the primary PCR reaction must be followed by another reaction that can discriminate between the wild type and mutant DNA. A major concern associated with mutation diagnostic methods is the ability to detect low abundant mutations. Diagnostic techniques must be sensitive enough to detect the mutant DNA without generating false positives or false negatives. Figure 1.10 contains schematic representations of the diagnostic methods described below.

**Allele-Specific PCR**

Each allele is amplified in separate reactions. PCR product is the indication of the presence of mutant alleles.

**Enzymatic Cleavage**

Fragments are formed after cleavage of the mismatched heteroduplex. The presence of smaller cleavage fragments indicates the presence of mutant alleles.

**Minisequencing**

The formation of an extension product one base longer than the primer indicates the presence of mutant alleles.

*Figure 1.10.* Schematic representations of diagnostic methods.
Allele specific PCR is accomplished with specially designed primers of defined lengths, which hybridize to the target at the mutation(s) potentially producing mismatches, with the premise being that mismatches in the primers impede or reduce the efficiency in the amplification step. Two common techniques are PCR-RFLP and amplification refractory mutation system (ARMS). In PCR-RFLP, only the mutant alleles are amplified by the addition of mismatches in the PCR primers before the restriction digests are performed. This allows much lower frequency mutations to be detected. In ARMS, two amplification reactions are performed each with an amplifying primer specific for mutant or wild type DNA, thereby amplifying only one allele in each reaction for subsequent electrophoretic analysis.

There are also many enzymatic and chemical cleavage assays developed that are used for diagnostic analysis of mutations. Different chemicals for each base (A, C, G, T) are used to cleave base mismatches formed when heteroduplexes are created. Enzymes such as bacteriophage T4 endonuclease VII, T7 endonuclease I, RNase A, or Cleavase I cleave mutant duplexed DNA according to changes in their secondary structure. Cleavage fragments can then be analyzed by electrophoresis. MutS is also an enzyme that binds to mismatched DNA duplexes and the complex that forms is visualized with atomic force microscopy (AFM).

“Minisequencing” or single nucleotide primer extension is another detection method that has been used for detecting mutations. A primer is extended by only one nucleotide if the 3’ end is complementary to the target being analyzed. The resultant fragments can be analyzed by electrophoresis, on DNA microarrays, or using matrix assisted laser desorption ionization (MALDI) mass spectrometry. The sensitivity of this technique is limited by the efficiency of the DNA polymerase extending the primer.
Another family of diagnostic methods utilizes a DNA ligase enzyme. A DNA ligase is an enzyme that will catalyze the formation of a phosphodiester bond preferentially in oligonucleotides hybridized to a target DNA strand. DNA ligase has very high selectivity to “seal” DNA in which there is perfect complementarity (match). A single base mismatch at the junction inhibits ligation. The ligase detection reaction is one such diagnostic method (Figure 1.11). LDR uses two adjacent primers and the thermostable ligase, \( Tth \) DNA ligase\(^{1,50-52} \), to distinguish all four bases potentially found at any position. Two allele specific primers are used which differ in the base at their 3'-end. This difference corresponds to the mutation difference of the gene. A common primer, which is fluorescently labeled, is also used that is selective for the allele specific primer that hybridizes to the mutant sequence. Since only the mutant discriminating primer ligates with the common primer this is all that is detected.

In low level mutation detection of genetic diseases such as cancer, PCR/LDR detection reactions have been used to detect mutations with a sensitivity of 1 mutant in 4000 normal alleles.\(^1,53\) In order to increase this sensitivity a restriction endonuclease (RE) digestion step has been added in order to eliminate all normal or wild type alleles.\(^1,54\) A restriction endonuclease is an enzyme that catalyzes the hydrolysis of phosphodiester bonds in double helical DNA. Restriction endonucleases recognize specific sequences of four to eight base pairs and cleave very specifically the sequence recognized. Therefore a specific sequence can be intentionally introduced into DNA and only these strands will be cleaved.

Figure 1.12 shows an example of how this is accomplished with a detection strategy termed PCR/RE/LDR. PCR/RFLP described above accomplishes the same goal but uses strategically placed mismatches in the PCR primer for conversion, yet mismatches that are read
Figure 1.11. Schematic representation of the ligase detection reaction (LDR). PCR products are denatured and common and discriminating primers annealed. DNA ligase only ligates those duplexes which contain exact matches and thermocycling amplifies the amount of product formed linearly.
by DNA polymerases leading to errors in the further rounds of PCR, that lead to false positive results in analysis. In PCR/RE/LDR nucleotide base analogs (indicated by Q in Figure 1.12) are used instead of mismatches. The analogs preferentially pair with one of the natural bases and therefore encode for that base specifically allowing for better conversion of the wild type sequence to one containing a restriction endonuclease recognition site. As will be explained in Chapter 2, the Hammer research group synthesized many nucleoside base analogs for use in the PCR/RE/LDR detection strategy. It was found that DNA polymerases tend to either “slip” when reading the analog or can flip out a base in the primer or template therefore extending the template by one less base. Due to the aforementioned problems, detection limits of mutation:wild type sequences were not increased from the original PCR/LDR strategy. It is the purpose of this document to discuss research performed in synthesizing new nucleoside analogs and in understanding the chemical factors that play a part in the mechanism by which the DNA polymerase extends a primer.

The pyrazole based nucleoside analogs previously synthesized were analyzed computationally and in polymerase recognition assays to better understand the structural consequences of altering a nucleoside as well as the DNA polymerase reaction to the alterations. New nucleoside base analogs were designed in an attempt to explore direct hydrogen bonding contacts between DNA polymerase and the base opposite which a nucleoside triphosphate would be incorporated. These analogs included a thiazole ring as the aromatic ring. The addition of an N-oxide allowed probing of the importance of the O2 atom of pyrimidines to DNA replication.

A backbone modification, the N3’→P5’ phosphoramidate, was also synthesized and tested in DNA polymerase mediated DNA synthesis in order to inhibit the 3’→5’ exonuclease
**Figure 1.12.** Diagram of the PCR step using nucleoside base analogs (Q) and the restriction endonuclease (RE) step of PCR/RE/LDR. PCR primers containing analogs are annealed to the DNA and extended. Natural primers are then annealed and extended forcing DNA polymerase to incorporate a natural base across from the analog. The incorporation forms a RE site which can then be cleaved leaving only mutant sequences to be detected by LDR.
activity of higher fidelity polymerases. The inhibition of the exonuclease activity of these higher fidelity polymerases would allow for their use in PCR/RE/LDR without the excision of the nucleoside base analogs. The use of higher fidelity polymerases is suggested to circumvent the problems associated with DNA polymerase slippage and flipping out of bases.

A final evolution of LDR is associated with better electrophoretic detection of products produced in the assay. The original assay made use of slab gel electrophoresis. Current methods of electrophoresis, such as capillary electrophoresis (CE) and microdevice electrophoresis increase the resolution of detection of DNA. Application of these methods to PCR/LDR is the current step in making LDR practical for mutation detection.

1.5 References

1.8 Longley, M. J.; Bennett, S. E.; Mosbaugh, D. W. Characterization of the 5' to 3' Exonuclease Associated with Thermus-Aquaticus DNA-Polymerase. Nucleic Acids Research 1990, 18, 7317-7322.


Chapter 2: Nucleoside Base Analogs

The mutation detection strategy PCR/RE/LDR depends on the use of nucleoside base analogs for conversion of natural sequences to restriction endonuclease recognition sites. These analogs should preferably be equally stable in DNA as the natural bases and be replicated by DNA polymerases equally efficiently as natural bases. In designing these analogs the factors contributing to DNA stability and DNA polymerase recognition need to be analyzed and understood for successful use in mutation detection strategies.

2.1 Introduction

Many chemical factors that affect DNA duplex stability and DNA polymerase recognition of DNA. Nucleoside base analogs have contributed much to the understanding of these factors. Base analogs have also been used to increase the degeneracy of the genetic code. Below are outlined the factors important to DNA duplex stability and recognition as well as a review of the literature concerning DNA nucleoside base analogs.

2.1.1 DNA Structure and Stability

DNA is composed of two strands of aromatic bases connected to a sugar-phosphate backbone (Figure 2.1). The bases are nitrogenous heterocyclic rings and the backbone is composed of furanose rings connected by phosphodiester bonds. The phosphodiester bonds have several preferred geometries, which leads to some limitation in sugar-phosphate backbone conformation. The two strands are wound into a right-handed double helix to form a DNA molecule.
Figure 2.1. The secondary structure of DNA indicating major and minor grooves on the left and helix width, pitch height, and distance between bases on the right.

Red = oxygen
Blue = nitrogen
Orange = phosphorus
The rings of the bases are connected to the sugar rings by a β-glycosidic linkage. This bond is described by the torsion angle $\chi$ (Figure 2.2).\textsuperscript{2,1,2,2} The bases can be in either the syn or anti conformation with the anti conformation being such that the bulk of the base points away from the sugar and the syn with the bulk over the sugar ring. The bases are planar due to the aromatic system of which the amino groups of cytosine and adenine are an integrated part. The amino groups as well as other functional groups on the bases are hydrogen bond donors and acceptors. Although the base rings are planar, the sugar ring exists in puckered conformations (Figure 2.3).\textsuperscript{2,1,2,2} It can be puckered in an envelope ($C_2$-endo or $C_2$-exo; $C_3$-endo or $C_3$-exo) or symmetrical ($C_2$-exo-$C_3$-endo) and unsymmetrical (major $C_3$-endo and minor $C_2$-exo) twist forms. Although these are the major puckering modes puckering can occur intermediate to these modes and is defined by the pseudorotation angle, $P$. DNA structure prefers the $C_2$-endo puckering mode ($P = 137^\circ$ to $194^\circ$) yet the syn/anti conformation dictated by the $\chi$ torsion angle can affect puckering mode. The more syn the position of the base the more the $C_2$-endo puckering mode is favored.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{nucleosides.png}
\caption{Syn and anti conformations of nucleosides.}
\end{figure}

The bases pair with each other in a complementary fashion to stabilize the double helix: the pyrimidine C with the purine G (forming three hydrogen bonds) and the pyrimidine T with the purine A (forming two hydrogen bonds). The two pyrimidine-purine base pairs form structures
similar in size and geometry (Figure 2.4). The C1'-C1' distance of each pair is approximately 11 Å and the C1'-N bond angles are approximately 50°. It is the isostructural nature of the G-C and A-T base pairs that allow the folding of DNA. In the helix the pitch is the distance along the helix required to complete one helix turn and the pitch height is the number of base pairs in this distance (Figure 2.1). In B-form DNA the pitch is approximately 34 Å with a pitch height of approximately 10 base pairs. The distance between a planar base pair and its neighbor is about 3.4 Å. The double helix also forms grooves in the overall structure. The minor groove, which is indicated by the O2 of pyrimidines and N3 of purines, has a width of 5.7 Å and a depth of 7.5 Å while the major groove is 11.7 Å wide and 8.5 Å deep.

![Sugar puckering modes of furanose rings in DNA](image)

**Figure 2.3.** Sugar puckering modes of furanose rings in DNA. The C2'-endo conformation is preferred in DNA.

The above-described structure of the DNA double helix is optimal for stability and there are many chemical factors that govern the structure and stability of the DNA molecule. These factors (discussed below) can be measured by measuring the melting temperature (T_m) of the DNA double helix. This is performed with optical thermal denaturation studies. Heating a solution of DNA causes the denaturation of the double helix into two single strands of polynucleotides. In the case of highly cooperative transitions, once the temperature reaches the
Figure 2.4. A. Systematic numbering of bases; B. The positions of hydrogen bond donors and acceptors positioned in the grooves of DNA; C. The geometry of natural base pairs: $C_1$-$C_1'$ distances and $C_1$-$N$ bond angles.
Tₘ even a slight increase in temperature causes the forces holding the two strands together to cease and all of the hydrogen bonds between the base pairs are broken. Therefore, the thermal denaturation of DNA can often be viewed as a simple two-state equilibrium (single stranded to duplex), which can be measured by UV absorbance. Figure 2.5 shows an example of such a transition. The UV absorbance of a DNA sample increases as the duplex is melted: therefore, the Tₘ is the temperature at which the UV absorbance is centered between the absorbance of double stranded DNA and the increased absorbance of single stranded DNA. The more stable the DNA double helix, the higher the Tₘ. From the Tₘ (and/or the melting curve) can be calculated thermodynamic properties by a van’t Hoff style analysis:

\[
\frac{1}{T_m} = \left( \frac{R}{\Delta H} \right) \cdot \ln \left( \frac{C_r}{4} \right) + \frac{\Delta S}{\Delta H} \quad \text{for non-self-complementary sequences, and}
\]

\[
\frac{1}{T_m} = \left( \frac{R}{\Delta H} \right) \cdot \ln(C_r) + \frac{\Delta S}{\Delta H} \quad \text{for self-complementary sequences, and}
\]

Gibb’s free energy, ΔG, can then be calculated by:

\[
\Delta G = \Delta H - T\Delta S
\]

DNA duplex formation is driven by enthalpy rather than entropy. The enthalpy term (ΔH) is normally a large and negative while the entropy term (-TΔS) is almost as large a positive value making the free energy favorable.

The enthalpy term is governed by the electrostatic properties that affect DNA duplex stability. These properties include hydrogen bonding, base stacking, and the repulsion of negative charges of the two backbone phosphates.\(^{2,2,5}\) Hydrogen bonding has been defined; base stacking is the ability of the planar aromatic bases to \(\pi\)-stack with their neighboring bases in
either a single or double stranded DNA molecule. In double stranded DNA there is not only base stacking within the strand (intrastrand stacking) but bases also stack with bases in the complementary strand (interstrand stacking) due to the turn of the double helix. Hydrogen bonding and repulsion of negative charges in the backbones enthalpically affect only the helical structure and not the single stranded DNA molecule, but base stacking is important enthalpically in both cases.

![Figure 2.5](image)

**Figure 2.5.** A typical transition in absorbance seen as temperature is increased causing all double stranded DNA to form single strands.

The entropy term is governed by solvation and restriction of bond rotations. Due to the many electronegative atoms of the DNA molecule it is easily dissolved in water. When the double helix is formed many of these atoms are unable to make hydrogen bonds with water because they are hydrogen bonding with other nucleosides: therefore, the process of solvation is more favorable entropically than formation of the double helix. Another entropically unfavorable factor is the restricted rotation of bonds in the backbone. When the helix is formed
these backbone torsion angles are stringently restricted as compared with that of a single stranded molecule. This is also seen with the restriction of the $\chi$ angle when a duplex is formed.

Although hydrogen bonding within a duplex is unfavorable entropically the enthalpic gain from the formation of many hydrogen bonds between the bases largely makes up for the entropic loss. The importance of hydrogen bonding to duplex stability will also be discussed later as this was studied using non-hydrogen bonding nucleoside analogs.

The electronic contributions that stabilize base stacking interactions in a DNA double helix include permanent dipoles, hydrophobic forces, dipole-induced dipole interactions (and therefore polarizability).\textsuperscript{2.2,2.5} The surface area of stacking also plays a role in base stacking. It has been shown that the purines stack much more efficiently than the pyrimidines.\textsuperscript{2.6} This is due in part to the larger surface area of the purines as well as to hydrophobic forces, which dictate that it is more favorable for the planar aromatic surfaces of the bases to stack with each other than interact with water.

2.1.2 DNA Polymerase Recognition

Some of the above-mentioned contributions to DNA duplex stability may also play a role in DNA polymerase recognition of DNA and high fidelity replication. DNA polymerases make specific hydrogen bonding contacts with DNA in the active site of the enzyme as well as making a stacking interaction (tyrosine residue) with the template base across from the incoming dNTP.\textsuperscript{2.7-10} Most of the hydrogen bonding contacts are between the negatively charged backbone and basic amino acid residues of the polymerase, but some of the bases are hydrogen bonded in the minor groove at O2 of pyrimidines and N3 of purines (Figure 2.4). While the hydrogen bonds forming base pairs in the double stranded portion of the DNA do not seem to be recognized by the polymerase, when incorporating a dNTP there does seem to be formation of
hydrogen bonds between the last template base of the duplex and the incoming dNTP as well as normal stacking interactions with the last primer base. Therefore, it can be said that hydrogen bonding and base stacking are important to correct nucleotide selection during replication.

Restrictions to size recognition during DNA replication are also of importance. Due to the tight fit of the incoming dNTP in the polymerase active site it is hypothesized that the shape of the base pair formed causes a steric restriction to inserting the correct dNTP. It has been hypothesized that the base pair, when the incoming dNTP is hydrogen bonding to its partner in the template, must be of the size of a natural Watson-Crick base pair.

DNA polymerases convert the duplex DNA into A-form DNA in the active site. The base pairs tend to show more A-character the closer to the point of insertion. This is seen in crystal structures by analyzing the widening of the minor groove. In A-form DNA the minor groove is wider than in B-form DNA. This widening in the active site may be due to hydrogen bonding contacts between the thumb subdomain and backbone atoms in the minor groove. In A-form DNA the distance between base pair faces is 2.6 Å as compared to 3.4 Å in B form DNA. This closer interaction in the active site may lead to an amplification of base stacking interactions.

2.1.3 Nucleoside Base Analogs-The Literature

Nucleoside base analogs have been used to test the import of the above mentioned factors to DNA duplex stability and DNA recognition by DNA polymerase. As well as base analogs that occur naturally, there are different classes of novel analogs that have been synthesized and analyzed via thermal denaturation studies and DNA polymerase extension assays. Figure 2.6 shows the structures of all of the analogs discussed herein. The first analogs to be studied were abasic sites in which the carbon-nitrogen glycosidic bond is hydrolyzed and no base is present,
and a nucleoside where the base is replaced with a phenyl ring. There is a large number of analogs that modify the natural bases altering hydrogen bonding and base stacking interactions as well as non-hydrogen bonding analogs that probe the importance of hydrogen bonds. A number of analogs were also designed as universal base analogs that would pair to all four of the natural bases with similar efficiency. A final class of analogs is azole based base analogs, which have also been used as universal base analogs.

2.1.3.1 DNA Stability

The abasic site and phenyl ring, when incorporated into DNA, were found to be very destabilizing to the duplex due to the lack of hydrogen bonds and base stacking in the case of the abasic site and hydrogen bonds alone in the case of the phenyl ring. The abasic site destabilized the duplex by almost 16 °C on average compared to the control sequence while the phenyl ring was almost 15 °C destabilizing on average. These examples started the probes of DNA duplex stability and have become standards by which destabilization is measured.

Examples of purines being modified are that of O6 methylguanine and N6 methoxyadenine. Both of these analogs were found to destabilize the duplex. Another adenine analog in which the ring nitrogen at the 3-position is replaced with a carbon has also been synthesized has been found to destabilize the duplex by only 1 °C. A pyrimidine modification includes an analog of thymine that lacks the O2 keto group which has been synthesized and tested by thermal denaturation. This analog was found to only destabilize the duplex by 1 °C. The above two modifications of adenine and thymine lack the minor groove binding atoms and were found to induce a change from B-form DNA to A-form DNA due to a lack of the spine of hydrations which has been thought to be most important for DNA duplex stability in water. Modifications of pyrimidines have also included removing the 4-keto
and 4-amino group of thymine and cytosine, respectively.²²² Both of these modifications are incredibly destabilizing to a DNA duplex to the point that in the thermal denaturation studies of the thymine mimic no random coil to helix transition was seen. There has been much more attention placed on the modification of pyrimidines at the 5-position since the discovery that the addition of a methyl group to the 5-position of cytosine greatly stabilizes a DNA-DNA and DNA-RNA duplex.²²³,²²⁴ Polarizability of the base was suggested to be the cause of increased stability: therefore, further modifications at the 5-position of pyrimidines have been made to increase this property. A group of alkyl, alkenyl, and alkynyl substituents was tested.²²⁵ Alkyl groups larger than methyl at the 5-position of thymine were shown to destabilize a DNA duplex while alkyl substituents on cytosine were stabilizing. The alkynyl groups propynyl, pentynyl, and hexynyl were stabilizing on thymine yet propynyl was destabilizing to cytosine. There have also been 5-position substitutions of halogen atoms. A series of thymine modifications were synthesized contained iodine, bromine, and fluorine.²²⁶ Bromine and fluorine modified thymines were stabilizing while iodine was only slightly destabilizing. A thymine mimic was also synthesized with a propynyl substituent at the 5-position.²²⁷ It was found to add approximately 1.5 °C stability to a DNA-RNA duplex. A final series of 5-substituted pyrimidines containing a pyridine, thiazole, imidazole, and thiophene was synthesized.²²⁸ All of the listed substituents were found to be slightly stabilizing, especially thiazole, to a DNA-RNA duplex (+0.3 °C, +1.7 °C, +0.7 °C, +0.4 °C, respectively).
Figure 2.6. Structures of nucleoside base analogs discussed in Chapter 2.
Other analogs that have been synthesized completely remove the hydrogen bonding capability of the nucleoside in order to probe the importance of base stacking. These analogs include difluorotoluene, trimethylbenzene, and dimethylindole.\textsuperscript{2.11,2.29} These analogs were designed as steric mimics of thymine and adenine, but without hydrophilic oxygens or nitrogen containing groups. All of these analogs destabilized a DNA duplex by almost 20 °C despite their structural similarity to the natural bases. It was determined that although stacking was maintained due to the lack of hydrogen bonds, base pair geometry was not preserved and caused destabilization. It is probable that the destabilization was caused in part by the hydrophobic nature of the analogs although this is not certain due the lack of entropic data in the report.

Universal base analogs differ from the already mentioned analogs in that they usually do not resemble the natural bases in any way. One of the first of the universal analogs was 5-nitroindole.\textsuperscript{2.30} It was found to be equally stable base paired with all of the natural bases and only slightly destabilizing (-5 °C) to the duplex as a whole. A group of isocarbostyril nucleoside analogs has been reported.\textsuperscript{2.31,2.32} These analogs are slightly destabilizing to the duplex but pair equally well with the four natural bases. Another recent addition to the family of universal bases is a pyrrolopyrimidine nucleoside analog of adenine in which in the sugar is attached to the N8 position of the base.\textsuperscript{2.33} This analog has been shown to be equally stable with each of the four natural bases and is as stabilizing to the duplex as an A-T base pair. The base pairs formed have different geometries than the natural base pairs but two hydrogen bonds are formed between each of the pairs. Further structural analysis of this analog is needed for a better understanding of the forces responsible for the stability this analog imparts.

Another group of analogs that were designed as universal analogs are the azole based nucleosides. These analogs are nitrogen containing five-membered rings which would decrease
stacking due to a decrease in size of the heterocycle but have added substituents to make up for this decrease. 3-nitropyrrrole was the first of these nucleoside reported.\textsuperscript{2,34} It was shown to equally pair with all four of the natural bases yet was found to destabilize the DNA duplex by approximately 10 °C. Experimental and computational structural analysis was performed and it was found that a 3-nitropyrrrole base pair varied only slightly from the natural base pairs. NMR analysis of a DNA duplex containing 3-nitropyrrrole indicated that the bulkiness of the nitro group protruded into the major groove and was the main cause of duplex destabilization.\textsuperscript{2,35}

Another report compared the stability of 3-nitopyrrrole with a number of other azole based nucleoside analogs.\textsuperscript{2,36} These included pyrazole, 4-nitropyrazole, 4-nitroimidazole and 5-nitroindole as well as comparing all of these to an abasic site. In increasing order the order of stability was found to be abasic site (\(\Delta G = -4.7\) kcal/mol), pyrazole (\(\Delta G = -4.8\) kcal/mol), nitropyrrrole (\(\Delta G = -6.3\) kcal/mol), 4-nitropyrazole (\(\Delta G = -6.5\) kcal/mol), 4-nitroimidazole (\(\Delta G = -7.0\) kcal/mol) and finally 5-nitroindole (\(\Delta G = -8.1\) kcal/mol). While the stability of 5-nitroindole is due to increased stacking area, the stability due to 4-nitroimidazole is due to a hydrogen bond formed when the base is in the syn conformation. When comparing the abasic site, pyrazole, and nitropyrazole it is seen that the enthalpy and entropy terms of pyrazole and nitropyrazole are consistently low in absolute value while when the abasic site is paired opposite adenine or cytosine these values are quite large. This indicates that electrostatic effects and solvation play a larger role in stabilizing an abasic site. Carboxamide groups were also added to azole rings in the hopes of developing bases capable of pairing with two or more of the natural bases termed universal bases.\textsuperscript{2,37} The carboxamide group was designed such that it could rotate and have different hydrogen bonding patterns to accommodate the four natural bases.\textsuperscript{2,38}

Thermodynamic analysis of pyrrole-3-carboxamide was very destabilizing to the duplex and
showed a preference for pairing with thymine. It was reported that the pyrrole ring stacked less efficiently and the amide group of the carboxamide did not obtain the position necessary for hydrogen bonding to maintain base pair geometry. It was also learned that there was an entropic price due to the restriction of rotation of the carboxamide when in the duplex and that if the base was in the anti conformation pointing the carboxamide into the major groove there was a distortion of base stacking.

2.1.3.2 DNA Polymerase Recognition

Few of the novel nucleoside analogs used to study DNA structure and stability have been used to study DNA polymerase mechanism and fidelity. Most analogs used to study DNA polymerase recognition of DNA have been naturally occurring analogs that are formed through mutagenic pathways.

Most mutations that occur in DNA change the structure of guanine. N2-methylguanine\(^2,39\), 8-methylguanine\(^2,40\), O6-methylguanine\(^2,41\), and N2-ethylguanine\(^2,42\) have all been used in polymerase assays in determining the chemical reasons for DNA polymerase mechanism and fidelity. N2-methylguanine and N2-ethylguanine are formed by the reaction of guanine with aldehydes (formaldehyde and acetaldehyde, respectively). It was found that N2-methylguanine encodes cytosine (84%) and thymine (9.4%) in primer extension reactions with exonuclease deficient Klenow fragment of \textit{E.coli} polymerase I, or Kf(exo-). This was mirrored in the kinetic analysis which indicated that the incorporation of thymine is only 1.2 times less efficient than cytosine. Therefore, in a small amount, \textit{G}→\textit{A} transition mutations are formed. With Kf(exo-) N2-ethylguanine coded for guanine (40%) and cytosine (39%) forming \textit{G}→\textit{C} transitions. These assays indicated that the 2-position amino group is critical for polymerase fidelity. 8-methylguanine was found to encode cytosine and only in small quantities adenine and
guanine (approximately 1% of each). This indicates that position 8 of guanine is not critical in correct dNTP incorporation by DNA polymerases. O6-methylguanine was found to encode thymine 76% of the time although kinetic analysis indicated that the efficiency of formation of this base pair was very low compared to a natural guanine-cytosine base pair and the rate of formation was lower than that of a thymine-guanine mismatch. The hydrogen bond formed with this keto group is of great importance to DNA polymerase fidelity. When studying O6-methylguanine a guanine analog lacking N3, 3-deazaguanine, was synthesized for comparison.\textsuperscript{2,7,2,41} The rate of incorporation of cytosine opposite 3-deazaguanine was reduced by 170-fold with \( K_f^{(exo-)} \) indicating that the hydrogen bond formed between N3 and an amino acid residue in the active site is of utmost importance for recognizing a correct base pair. A reduced efficiency was also reported indicating that the rate of the formation of the closed conformation of the polymerase was decreased.

The novel synthesized nucleoside base analogs used in polymerase assays include the isocarbostyril analogs.\textsuperscript{2,32,2,43} These analogs were found to encode for the natural bases with slightly similar efficiencies but with greatly reduced efficiencies compared with the natural bases. The isocarbostyril analog containing a propynyl functionality was use as a triphosphate in extension reaction and was found to be inserted less efficiently than the natural bases. Other base analogs synthesized as triphosphates include thymine and cytosine mimics that lack the O2 to test minor groove recognition by DNA polymerase.\textsuperscript{2,44} It was reported that in extension reactions with \( K_f(\text{exo-}) \) and \textit{Taq} polymerase activity was inhibited with increasing amounts of analog indicating the O2 of pyrimidines is necessary for base pair recognition during dNTP incorporation by DNA polymerases. Both of these analogs were C-nucleosides in which a carbon-carbon bond rather than the natural carbon-nitrogen bond connects the base to the sugar.
It was shown with a psuedo-thymine analog, which is also a C-nucleoside, that DNA polymerase activity is not altered by this modification.\textsuperscript{2,45} Incorporation of adenine opposite pseudo-thymine was similar in efficiency to the incorporation of adenine opposite thymine.

Much enzymatic analysis has been reported with the nonpolar, non-hydrogen bonding analogs difluorotoluene (thymine mimic), 4-methylbenzimidazole (adenine mimic lacking an N3 minor groove hydrogen bond acceptor), and 9-methyl-1-H-imidazopyridine (adenine mimic containing N3).\textsuperscript{2,46-51} Difluorotoluene was used as a dNTP and incorporation efficiency opposite adenine was found to be very similar in efficiency to the natural adenine-thymine base pair and with difluorotoluene in the template dATP was inserted only 3-fold less efficiently than opposite thymine. These findings indicate that geometry plays a crucial role in base pair recognition and formation by DNA polymerase. In extension reactions with K\textsubscript{f}(exo-) and Taq polymerases the three analogs were placed both in the primer and template strands in separate reactions. It was found that extension was more efficient with the natural base in the primer strand as compared to the analogs in the primer strand indicating that critical contacts between residues in the active site and the last base pair of the duplex being replicated are more important in the primer strand than in the template strand during the extension step of replication (Figure 2.10). It was also found that the adenine mimic lacking the minor groove hydrogen bond acceptor was replicated much less efficiently than the mimic containing the hydrogen bond acceptor. This indicates that this location of contact is important to DNA polymerase processivity and fidelity.

Enzymatic analysis was also performed with azole nucleobase analogs. 3-nitropyrrrole, pyrrole-3-carboxamide, pyrazole-4-carboxamide, pyrazole-3-carboxamide, and imidazole-4-carboxamide were all incorporated into PCR primers and used to generate PCR products which were sequenced to determine which bases were incorporated across from the analogs.\textsuperscript{2,52} It was
found that adenine was incorporated the most opposite 3-nitropyrrrole and pyrazole-4-carboxamide (75% and 63%, respectively). Thymine was incorporated opposite pyrrole-3-carboxamide and imidazole-4-carboxamide most often (65% and 77%, respectively). Pyrazole-3-carboxamide coded for cytosine 51% and thymine 45%. It is interesting that the abasic site tested coded for adenine 92% and all of the azole analogs coded for adenine most except pyrazole-3-carboxamide. The “A-rule” states that DNA polymerases tend to insert an adenine when the chemical rules of incorporation do not apply to an insertion situation. The “A-rule” is a possible explanation for the adenine incorporation preference.

2.2 Pyrazole Based Analogs

2.2.1 Introduction

In the Hammer research group a series of pyrazole based analogs have been synthesized, thermodynamically characterized, and incorporated into PCR primers for analysis of polymerase insertion opposite these analogs for use in the PCR/RE/LDR mutation detection reaction. These analogs have substituents at the 4-position of the pyrazole ring that vary in size, polarity, and stacking ability. These substituents include, in increasing order of stacking ability, iodo, nitro, propynyl, and thiazolyl (Figure 2.7). The complete characterization of these analogs will contribute to the understanding of DNA stability and polymerase recognition which will ultimately be useful for mutation detection.

It was reported that all of the pyrazole based analogs when incorporated into the self-complementary Dickerson dodecamer (5´-CGCXAATTYGCG-3´) were most stabilizing to a DNA duplex when paired with adenine although they were all quite destabilizing (-11 ºC to –31 ºC). The stability of the duplex increased with increasing stacking ability: iodo ($\Delta G = -6.3$ ºC) < nitro ($\Delta G = -6.8$ ºC) = propynyl ($\Delta G = -6.8$ ºC) < thiazolyl ($\Delta G = -7.9$ ºC). The measured
thermodynamic values as reported earlier are seen in Table 2.1 and bar graphs comparing the stability of each of the analogs can be seen in Figure 2.8. This decrease in stability was consistent with a decrease in enthalpy as well as entropy indicating that these analogs affect both solvation and steric interactions as well as electrostatic effects. Solvation is affected by the increased hydrophobicity of the analogs. The loss of groove hydration leads to the decreased stability of a 4-substituted pyrazole-adenine base pair. The loss of hydrogen bonding and stacking ability lead to the small enthalpy term. The small 5-membered pyrazole ring lacks the stacking surface area of the natural bases. Although the enthalpy and entropy values are consistently lower than the control sequences the values do not decrease with decreasing stacking ability. The trend for the decrease in thermodynamic values is: Nitropyrazole < propynylpyrazole < iodopyrazole < thiazolylpyrazole. It can be hypothesized that thiazolylpyrazole has the greatest enthalpy and entropy values due the increased size of the stacking surface and increased hydrophilicity of the thiazole ring compared to the other base analogs. Iodopyrazole and propynylpyrazole show a greater increase in enthalpy and entropy than nitropyrazole possibly due to rearrangement. As can be seen in Figure 2.9 if these analogs
Table 2.1. Thermodynamic values obtained from thermal melting analysis of the Dickerson dodecamer (5’-CGCXAATTYGCG-3’). PzI = iodopyrazole, PzN = nitropyrazole, PzP = propynylpyrazole, and PzT = thiazolylpyrazole.²,⁵⁶

<table>
<thead>
<tr>
<th>XY</th>
<th>(\Delta T_m(°C))/sub.</th>
<th>(\Delta H) (kcal/mol)</th>
<th>(\Delta S) (cal/K•mol)</th>
<th>(\Delta G) (kcal/mol) (298K)</th>
<th>(\Delta \Delta G) (kcal/mol) (298K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-A</td>
<td>-69.5</td>
<td>-189.3</td>
<td>-13.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G-C</td>
<td>-87.7</td>
<td>-235.9</td>
<td>-17.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-G</td>
<td>-69.1</td>
<td>-183.9</td>
<td>-14.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-T</td>
<td>-80.9</td>
<td>-219.3</td>
<td>-15.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PzI-A</td>
<td>-16.7</td>
<td>-39.0</td>
<td>-6.3</td>
<td>-6.8</td>
<td></td>
</tr>
<tr>
<td>PzI-C</td>
<td>-31.1</td>
<td>-28.8</td>
<td>-4.5</td>
<td>-12.9</td>
<td></td>
</tr>
<tr>
<td>PzI-G</td>
<td>-26.3</td>
<td>-26.8</td>
<td>-5.1</td>
<td>-9.2</td>
<td></td>
</tr>
<tr>
<td>PzI-T</td>
<td>-24.1</td>
<td>-26.5</td>
<td>-5.5</td>
<td>-10.0</td>
<td></td>
</tr>
<tr>
<td>PzN-A</td>
<td>-14.3</td>
<td>-31.9</td>
<td>-6.8</td>
<td>-6.3</td>
<td></td>
</tr>
<tr>
<td>PzN-C</td>
<td>-23.5</td>
<td>-28.9</td>
<td>-6.1</td>
<td>-11.3</td>
<td></td>
</tr>
<tr>
<td>PzN-G</td>
<td>-21.8</td>
<td>-41.6</td>
<td>-5.7</td>
<td>-8.6</td>
<td></td>
</tr>
<tr>
<td>PzN-T</td>
<td>-21.2</td>
<td>-31.6</td>
<td>-6.0</td>
<td>-9.5</td>
<td></td>
</tr>
<tr>
<td>PzP-A</td>
<td>-14.7</td>
<td>-35.8</td>
<td>-6.8</td>
<td>-6.3</td>
<td></td>
</tr>
<tr>
<td>PzP-C</td>
<td>-25.8</td>
<td>-41.3</td>
<td>-5.2</td>
<td>-12.2</td>
<td></td>
</tr>
<tr>
<td>PzP-G</td>
<td>-18.1</td>
<td>-43.0</td>
<td>-6.7</td>
<td>-7.6</td>
<td></td>
</tr>
<tr>
<td>PzP-T</td>
<td>-23.1</td>
<td>-25.3</td>
<td>-5.8</td>
<td>-9.7</td>
<td></td>
</tr>
<tr>
<td>PzT-A</td>
<td>-11.2</td>
<td>-47.6</td>
<td>-7.9</td>
<td>-5.2</td>
<td></td>
</tr>
<tr>
<td>PzT-C</td>
<td>-21.5</td>
<td>-32.7</td>
<td>-6.5</td>
<td>-10.9</td>
<td></td>
</tr>
<tr>
<td>PzT-G</td>
<td>-13.0</td>
<td>-42.9</td>
<td>-8.2</td>
<td>-6.1</td>
<td></td>
</tr>
<tr>
<td>PzT-T</td>
<td>-17.0</td>
<td>-33.8</td>
<td>-6.9</td>
<td>-8.6</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.8. Graphical representation of the stabilities of each of the pyrazole analogs reported earlier. All of the analogs are most stable when base paired with adenine (top) and of the analogs thiazolylpyrazole is the most stable when paired with all of the natural bases (bottom).
occur in the *anti* conformation the propynyl and iodo substituents can stick into the major groove relieving some steric stress of the duplex. Nitropyrazole is unable to relieve this stress in the same way due to the favored conformation in Figure 2.8. Therefore, although the stabilities of nitropyrazole and propynylpyrazole are the same these analogs accomplish this in different ways. More structural characterization is required to analyze the thermodynamic values fully.

Nitropyrazole and thiazolylpyrazole were also incorporated into PCR primers and used in PCR reactions. The resultant PCR products were sequenced to determine the incorporated bases opposite the analogs. It was found that with *Taq* polymerase nitropyrazole coded for adenine 94% of the time, and thiazolylpyrazole coded for adenine 93% of the time. This agrees with the thermodynamic data that these two analogs are most stable when paired with adenine. This is again probably due to the base pair geometry formed with adenine (Figure 2.8). Cytosine is the only other base that these interactions are possible with, but the small size of the pyrazole ring would increase the C$_1$-C$_1$' distance too much.

In order to affirm the above theories about pyrazole analogs, structural computational analysis of the factors leading to duplex stability were calculated and compared to the thermodynamic data. DNA polymerase extension reactions were also performed with nitropyrazole and thiazolylpyrazole in order to understand further a stall product seen in the PCR product analysis and understand better the mechanism by which DNA polymerase chooses a dNTP to insert opposite the pyrazole analogs.

### 2.2.2 Synthesis

The synthesis of the pyrazole analogs was reported elsewhere.$^{2,56}$ The reported synthesis is outlined in Scheme 2.1. The $p$-toluoyl protected furanose ring was used to form the 1-chloro-2-deoxy-3,5-di-$O$-tolouoyl-D-erythropentofuranose which was then reacted with the sodium salt
Figure 2.9. The possible syn/anti conformations of the pyrazole analogs and probable base pair geometries formed with adenine.
of the iodopyrazole or nitropyrazole. Palladium chemistry was used to convert iodopyrazole to propynylpyrazole and thiazolylpyrazole. The synthesis of thiazolylpyrazole required the formation of trimethyltin pyrazole before Pd coupling with 2-bromothiazole. The yield of the \( p \)-toluoyl protected thiazolylpyrazole nucleoside was reported as being only 60%. In an effort to increase yield and ease of synthesis a direct Pd catalyzed coupling of \( p \)-toluoyl protected iodopyrazole with thiazolyl zinc bromide was attempted (Scheme 2.2). It was previously shown that biaryls could be formed by Pd catalyzed reactions of aryl-Zn derivatives with aryl halides.\(^2\)\(^{57,58}\) These reactions were found to proceed successfully with 5 mol\% Pd catalyst. When the coupling of iodopyrazole and thiazolylzinc bromide was attempted 20 mol\% Pd was required. This led to problems with purification. Due to the large excess of (tetrakis)triphenylphosphine palladium used, a large amount of triphenylphosphine oxide was formed, lowering the solubility of the final products. The amount of triphenylphosphine oxide formation could be reduced when the reaction was performed under an inert atmosphere (dry box). Even though purification was greatly hindered a final yield of 90% was obtained. The incorporation of these analogs into DNA oligonucleotides is illustrated in Scheme 2.3. The synthesized analogs were \( p \)-toluoyl deprotected. The 5´-hydroxyl was dimethoxytrityl protected and the 3´-hydroxyl was converted to a 2-cyanoethyl \( N,N \)-diisopropylphosphoramidite for use in the solid phase synthesis of oligonucleotides containing the analogs nitropyrazole and thiazolylpyrazole.\(^2\)\(^{56}\)
Scheme 2.1. Synthesis of the pyrazole analogs.
Scheme 2.2. Improved synthesis of thiazolylpyrazole analog.

2.2.3 Computational Analysis

Computational analysis of the nucleoside analogs was performed using the program Hyperchem 6.0 (Hypercube, Inc.). Geometric optimization of the natural bases as well as the pyrazole analogs was performed with the AMBER and MM+ force fields. While the RMS fit of the natural base database entries was lower with an AMBER minimization the RMS fit of the pyrazole analogs with the crystal structures obtained for the $p$-toluoyl protected nucleosides was lower with an MM+ minimization. It was decided to continue analysis using the minimized structures from the MM+ geometrical optimization. Semiempirical (PM3) calculations were then performed on the minimized structures of the natural bases as well as the pyrazole analogs. The values for surface area, dipole moment, and polarizability are shown in Table 2.2. The RMS fit values for the natural bases are in comparison to the database values in Hyperchem 6.0 while those of the pyrazole analogs are in comparison to the crystal structures obtained as described above.
Scheme 2.3. DNA synthesis by the cyanoethyl phosphoramidite approach. TCA = trichloroacetic acid, DMT = dimethoxytrityl, and CPG = controlled pore glass solid support.
The values for surface area, dipole moment, and polarizability for the pyrazole analogs are within the range of the natural bases. Thiazolylpyrazole was found to have the largest values of the properties computed. Propynylpyrazole had a very large surface area, yet its dipole moment is comparable to that of iodopyrazole. Nitropyrazole is smaller yet has a large dipole moment.

In an attempt to correlate the calculated data with the thermodynamic data, plots of surface area, dipole moment, and polarizability versus \( -\Delta G_{37^\circ C} \) of the duplexes were made (Figures 2.10 and 2.11). When each of the calculated parameters of the analogs is plotted versus \( -\Delta G \) values of each analog paired with the natural bases there is good correlation with dipole moment only. Conversely, when the calculated values of the natural bases are plotted versus \( -\Delta G \) values good correlation is obtained with polarizability. This indicates that the dipole moments of the analogs are inducing electrostatic interactions or polarization of the natural bases. It is also seen that the surface area of the analogs shows intermediate correlation with stability. Surface area tends to affect base pairing with purines more than with pyrimidines. This would seem to indicate that the major contributor to DNA duplex stability when a pyrazole analog is incorporated is due to electrostatic interactions. It is also found that the three properties measured tend to show greatest correlation with stability when paired with the purines. This is consistent with the thermodynamic data indicating that the analogs are most stable when base paired with adenine.
Table 2.2. Semiempirical (PM3) computational analysis (MM+ minimization) comparing the natural bases and the pyrazole analogs.

<table>
<thead>
<tr>
<th>BASE</th>
<th>RMS fit (Å)</th>
<th>Surface Area (Å²)</th>
<th>Dipole Moment (Debyes)</th>
<th>Polarizability (Å³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenine</td>
<td>0.12</td>
<td>134.8</td>
<td>2.14</td>
<td>14.9</td>
</tr>
<tr>
<td>Cytosine</td>
<td>0.05</td>
<td>118.8</td>
<td>5.92</td>
<td>10.9</td>
</tr>
<tr>
<td>Guanine</td>
<td>0.07</td>
<td>142.9</td>
<td>7.02</td>
<td>14.2</td>
</tr>
<tr>
<td>Thymine</td>
<td>0.06</td>
<td>134.3</td>
<td>4.11</td>
<td>11.9</td>
</tr>
<tr>
<td>ANALOG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitropyrazole</td>
<td>0.06</td>
<td>112.9</td>
<td>5.08</td>
<td>9.4</td>
</tr>
<tr>
<td>Propynylpyrazole</td>
<td>0.05</td>
<td>134.4</td>
<td>2.52</td>
<td>12.0</td>
</tr>
<tr>
<td>Iodopyrazole</td>
<td>0.06</td>
<td>118.6</td>
<td>2.57</td>
<td>12.6</td>
</tr>
<tr>
<td>Thiazolylyrazole</td>
<td>0.04</td>
<td>147.0</td>
<td>9.06</td>
<td>15.9</td>
</tr>
</tbody>
</table>
**Figure 2.10.** Plots of $\Delta \Delta G$ values obtained from thermodynamic analysis versus calculated surface area and dipole moment values of pyrazole analogs.
Figure 2.11. Plots of -ΔΔG values obtained from thermodynamic analysis versus calculated polarizability value of the pyrazole analogs (top) and of the natural bases (bottom).
2.2.4 DNA Polymerase Recognition of Pyrazole Analogs

DNA polymerase extension reactions were performed using oligonucleotide templates containing the pyrazole analogs nitropyrazole and thiazolylpyrazole. Due to the evidence of shorter PCR products seen in the aforementioned PCR assays termination was hypothesized as a reason for these fragments. It was thought that DNA polymerase was “stalling” when inserting dNTPs opposite the analogs (insertion step) and at the next base following the analogs (extension step). (Figure 2.12) Extension reactions were used to probe the site of termination and quantitate the amount of termination in hopes of learning more about DNA polymerase fidelity mechanisms.

![DNA Polymerase Diagram](image-url)

**Figure 2.12.** The actions performed by DNA polymerase during template directed primer extension. X represents the base in the primer. X can be either a natural base or an analog. Y and Z represent bases that are being added to the primer strand by DNA polymerase. These bases can also be either natural or analog bases. The DNA polymerase makes specific and separate contacts with X, Y, and Z during replication.
The sequences of primer (18nt) and template (25nt) used in the DNA polymerase extension reactions can be seen below.

5´-CGC TCA ACT ATG TAA ACG  
3´-GCG AGT TGA TAC ATT TGC Q TAC TAG-5´

Q = thymine, nitropyrazole or thiazolylpyrazole

The 5´-33P radiolabeled primer was extended directed by the templates containing thymine as well as the pyrazole analogs with Taq and Kf(exo-) polymerases. The resultant extension products formed were then analyzed by electrophoresis. The extension reactions can be seen in Figure 2.13. The primer extended on the control template was fully extended after only one minute. The primer extended on the nitropyrazole template did form full-length product even after one minute but termination product is still seen after 30 minutes. With the thiazolyl pyrazole template less full-length extension product was seen than compared with either the control or nitropyrazole template. There is slightly increasing amounts of termination product formed as time increases. The major termination products seen with both the nitropyrazole and thiazolylpyrazole template correspond to the addition of a dNTP opposite the analog. This indicates that a dNTP can be incorporated opposite the analog but that the polymerase then stalls when incorporating the next dNTP. This indicates either an inability of the polymerase to form the closed complex or to form the phosphodiester bond. This may be due to the change in the overall duplex structure, lack of hydrogen bonds positioning the analog in the active site, or a lack of stacking interactions with the tyrosine residue in the active site. The change in overall duplex structure and lack of stacking interactions in the active site are less probable than the lack of hydrogen bonding interactions in incorporation of a dNTP opposite the analog since nitropyrazole was extended more efficiently than thiazolylpyrazole. Thiazolylpyrazole was found to form a more thermodynamically stable duplex than nitropyrazole and the introduction

61
of the thiazole ring to the pyrazole ring increases stacking interactions more than nitropyrazole. The lack of minor groove hydrogen bonds to position the template base to be “read” by the polymerase would cause the wrong positioning of the template base, which would not allow the incoming dNTP to be in the correct position for nucleophilic attack by the 3’-hydroxyl of the primer. In the extension step when binding of a dNTP to be inserted opposite the base to the 5’-side of the analog, a change in the geometry of the analog-natural base (presumably adenine) base pair probably moves the 3’-hydroxyl of the primer out of position to perform nucleophilic attack. In either case if the correct positioning is not attained a closed complex will not form and phosphodiester bond formation will not occur.

**Figure 2.13.** Extension assays of primers on templates containing thymine (control), nitropyrazole (PzN), and thiazolylpyrazole (PzT). The primer is 18nt and the template is 25nt in length. Extension reactions were performed in 20 µL with 250 nM radiolabeled primer and template, and 200 µM solution of all four dNTPs, AmpliTaq® polymerase and the supplied buffer (100 mM Tris-HCl pH 8.3, 500 mM KCl). An additional 8 mM MgCl₂ was added to the reaction.
The results indicate that nitropyrazole forms a base pair that more closely resembles the geometry of a natural base pair. Although thiazolylpyrazole has more stacking ability and forms a more stable duplex structure than nitropyrazole the overall geometry of the base pair is distorted to such a degree that polymerization is almost impossible even after extended amounts of time.

### 2.3 Thiazole and Thiazole N-oxide Analogs

Thiazole and thiazole N-oxide base analogs were designed as pyrimidine mimics in order to probe the importance of the O2 minor groove hydrogen bond acceptor atom. These analogs are modeled after the C-ribonucleoside tiazofurin, which has shown to be an antitumor agent.\(^2.59,2.60\) Tiazofurin has a conformation such that the partially positive sulfur of the thiazole ring is energetically favored to be located over the partially negatively charged oxygen of the furanose ring. This conformation allows the N-oxide to align in a DNA duplex in the same manner as the O2 of the pyrimidines. Figure 2.14 compares the structures of tiazofurin, thiazole, and thiazole N-oxide. Thiazole and thiazole N-oxide nucleoside analogs and oligonucleotides containing the analogs were synthesized.\(^2.61\) Computational, thermodynamic, and enzymatic analyses of thiazole and thiazole N-oxide and oligonucleotides were performed using the thiazole analog as a control in order to probe the effects of the N-oxide directly.

![Figure 2.14. Structures of tiazofurin, thiazole, and thiazole N-oxide base analogs.](image-url)
2.3.1 Synthesis

Many reports have provided synthetic details for preparation of the thiazole ring in tiazofurin and related analogs.\textsuperscript{2,62-64} The thiazole ring was constructed from ribofuranosylcyanide using the classical Hantzsch method, but with modifications to allow the preparation of 2-(2’-deoxy-β-D-ribofuranosyl)thiazole and its N-oxide.\textsuperscript{2,65} The formation of the thiazole and thiazole N-oxide nucleoside analogs can be seen in Scheme 2.4. Reaction of chlorosugar 1 with excess sodium cyanide in dry DME provided the nitrile 2 in a 3.2:1 β:α mixture of anomers in 83% yield. As previously reported separation of these anomers proved difficult.\textsuperscript{2,62} Quantitative conversion of 2 to the thioamide by treatment with hydrogen sulfide and a catalytic amount of DMAP allowed ready purification of the β-anomer to provide compound 3. The thioamide 3 was refluxed with 50% aqueous chloroacetaldehyde in 2:1 THF:ethanol\textsuperscript{2,66}, but the reaction was low-yielding (28%) and chromatographic separation of the thiazole nucleoside (4) from side-products proved difficult. However, reaction of 3 with bromoacetaldehyde diethyl acetal and a catalytic amount of 4 M HCl/dioxane in refluxing acetone\textsuperscript{2,67} provided compound 4 in improved yield (70%), with ready purification from side-products.

For formation of the N-oxide (5), compound 4 was treated with 2 equiv of \textit{m}CPBA\textsuperscript{2,68} in THF (72 h) to provide the N-oxide 5 in 48% yield. Longer reaction times provided little additional product, although unreacted thiazole 4 was readily recycled. However a more efficient method using 1.1 equivalents each of trifluoroacetic anhydride (TFAA) and hydrogen peroxide urea complex (UHP)\textsuperscript{2,69} was utilized, which provided a much higher yield (97%) after only 3 h at room temperature. Removal of the hydroxyl protecting groups with NH\textsubscript{4}OH proved destructive to the thiazole ring of 4. Therefore, the esters of 4 and 5 were dissolved in 50 mM
Scheme 2.4. Synthesis of thiazole and thiazole N-oxide analogs.

sodium methoxide/methanol to provide the free nucleoside analogs 6 and 7 in 74% and 83% yields, respectively.

The structure and absolute configuration of 4, 5 and 6 were confirmed by X-ray crystallography (Figure 2.15). Structural analysis indicates a C2- endo-sugar conformation (P = ~180). The thiazole ring is planar with the sulfur positioned over the oxygen of the furanose ring. The dihedral angles, \( \chi \) (O1\(^{\prime}\)-C1\(^{\prime}\)-C2-S1), for the toluoyl protected thiazole nucleoside, the toluoyl protected thiazole N-oxide nucleoside, and the free thiazole nucleoside are 21.6°, 23.0°, and 32.6°, which is comparable to other crystal structure data obtained for similar thiazole nucleosides.\(^{2,70}\)
For incorporation into oligonucleotides, nucleoside analogs 6 and 7 were converted to their protected nucleoside phosphoramidites (Scheme 2.5). The 5′-hydroxyl of each was dimethoxytritylated in anhydrous pyridine using 1.0 equiv of Et₃N and dimethoxytrityl chloride (DMTCl) to provide DMT ethers 8 and 10 in 70% and 90% yield, respectively. For complete conversion, compound 7 required treatment with a second equiv of Et₃N and DMTCl. Phosphitylation of thiazole 8 was accomplished by treatment with 2-cyanoethyl tetraisopropylphosphorodiamidite (3 equiv), diisopropylamine (DIPA; 1 equiv) and tetrazole (1 equiv) in dry CH₂Cl₂. After a 2 h reaction time, CH₃OH was added and the product purified to provide thiazole phosphoramidite 9 in 89% yield. These conditions were initially used for the phosphitylation of thiazole N-oxide nucleoside 10 and resulted in a complex reaction mixture containing significant amounts of the deoxygenated thiazole amidite 9 as indicated by TLC.
Attempted phosphitylation of the $N$-oxide 10 with the more reactive chlorophoramidite reagent also led to a mixture of products containing mostly starting material as well as thiazole phosphoramidite 9. Thus a modified protocol was developed whereby phosphordiamidite reagent is added portionwise to avoid large excesses. To a solution of DMT-$N$-oxide 10 in CH$_2$Cl$_2$ was added DIPA (1.1 equiv), tetrazole (1.1 equiv), 2-cyanoethyl tetraisopropyl-phosphorodiamidite (1.3 equiv) and was reacted for 30 min. A second portion of 2-cyanoethyl tetraisopropylphosphorodiamidite (1.3 equiv) was added. Following an additional 20 min of reaction, CH$_3$OH was added to quench the reaction and thiazole $N$-oxide phosphoramidite 11 was purified by flash chromatography in 67% yield.

![Scheme 2.5. Synthesis of thiazole and thiazole $N$-oxide phosphoramidites.](image)

For oligonucleotide synthesis, each nucleoside phosphoramidite was dissolved in dry acetonitrile to a 0.1 M concentration and used to prepare several self-complementary oligonucleotides based on the Dickerson dodecamer sequence, as well as oligonucleotides to examine the enzymatic properties of each analog (Section 2.3.4). Due to the observed ammonia
lability of thiazole, phenoxyacetyl (Pac), 4-isopropylphenoxyacetyl and acetyl N-protection for dA, dG and dC phosphoramidites, respectively, were used for oligonucleotide assembly. (Glen Research, Sterling, VA) The synthesis was performed on a 1.0 µmol scale, and all coupling steps except for the thiazole analogs were >98% by trityl assay. (vide infra) Base and phosphodiester deprotection was accomplished using 50 mM potassium carbonate in methanol (200 µl) at room temperature for 3 h. Acetic acid (1.5 eq.) in water (200 µl) was added and the resulting solutions evaporated to dryness.

The deprotected oligonucleotides were analyzed by anion exchange HPLC. Figure 2.16 shows the resulting HPLC profile for the oligonucleotide sequence 5´-(CGCQAATTGCG)-3´, where Q represents a nucleoside having either thiazole or thiazole N-oxide as heterocyclic base. Analysis of the products by MALDI spectroscopy indicates that the thiazole nucleoside is completely stable to the steps used for oligonucleotide synthesis and deprotection. However, three product peaks are evident for the thiazole-N-oxide oligonucleotide. The mass of the peaks corresponds to an octanucleotide (Figure 2.16 peak III) as well as intact dodecanucleotides containing either thiazole (Figure 2.16 peak I) or thiazole N-oxide (Figure 2.16 peak II). The column used for anion exchange chromatography of the above dodecamer was not sufficient for separating the peaks of the larger 28-nucleotide sequence for enzymatic analysis: 5´-ACTGQTCTCCTATAGTGAGTCGTATTA-3´ where Q represents the position containing thiazole or thiazole N-oxide. A quaternary ammonium functionalized polymeric anion exchange column was purchased from Dionex Corp. The HPLC trace of the 28-nucleotide sequence containing thiazole N-oxide can be seen in Figure 2.17. Five peaks were resolved in this analysis using 1 M NaCl, 25 mM NaOH (pH 12.4). MALDI analysis was performed on each
peak and the first two peaks are of the mass of a 23-nucleotide sequence (calculated: 7005, measured: 7000), the third peak corresponds to the 28-nucleotide sequence containing thiazole

**Figure 2.16.** HPLC trace of Dickerson dodecamer containing thiazole (A) and thiazole $N$-oxide (B). HPLC was performed using a Hydrocell NS 1000 (anion-exchange) column (150 x 4.6 mm column) with Solvent A = 25 mM CHES pH 8, 30% methanol and Solvent B = 25 mM CHES, 1 M (NH$_4$)$_2$SO$_4$, pH 8, 30% methanol (0-100% B in 60 min with a flow rate = 1 mL/min).

**Figure 2.17.** HPLC trace of the 28nt template containing thiazole $N$-oxide. HPLC was performed using a DNAPac-100 (anion-exchange) column with Solvent A = 1 M NaCl, 25 mM NaOH (pH 12.4) and Solvent B = 25 mM NaOH (pH 12.4) (100-500 mM NaCl in 2 min, 500-900 mM NaCl in 18 min, and 900-1000 mM NaCl in 3 min with a flow rate = 6 mL/min).
\(N\)-oxide (calculated: 8519, measured: 8513) and the fourth and fifth peaks correspond to a 28-nucleotide sequence containing thiazole (calculated: 8504, measured: 8496).

Both the problem in the synthesis of the thiazole \(N\)-oxide phosphoramidite 11 and the deoxygenation of thiazole \(N\)-oxide during solid phase oligonucleotide synthesis suggest that the activated P(III) reagents are causing the reduction. A common method for reduction of aromatic \(N\)-oxides is the reaction with phosphorous trichloride.\(^{2,71,72}\) Thus, it appears that for deoxygenation to occur the \(N\)-oxide must first be covalently attached to the P(III) moiety. In Scheme 2.6 a mechanism (A) is suggested whereby the thiazole \(N\)-oxide, whether resin bound or in solution reacts with activated phosphoramidite reagent to give a phosphitylated \(N\)-oxide species 12. Once attached to the oxygen the phosphorous can donate its electrons to reduce the thiazole and create a metaphosphate-like intermediate. This mechanism is favored over direct nucleophilic reduction by the P(III) reagents (B) as we find the thiazole \(N\)-oxide phosphoramidite to be stable to long term storage. This explains the decreased coupling efficiency of the thiazole \(N\)-oxide phosphoramidite and therefore large peak corresponding to the octanucleotide seen in the HPLC chromatogram. The tetrazole-activated \(N\)-oxide may be reacting with the phosphorous of other \(N\)-oxide phosphoramidite molecules, thus decreasing the concentration of the active phosphorous species.

### 2.3.2 Computational Analysis

Computational analysis of thiazole and thiazole \(N\)-oxide nucleoside base analogs was carried out using Hyperchem 6.0 (Hypercube, Inc.). The nucleosides were minimized using the AMBER force field since this force field is proven more correct for use with DNA and RNA and further molecular dynamics analysis of these analogs would warrant the use of this force field. Semiempirical (PM3) calculations were performed on the minimized structures and surface area,
Scheme 2.6. Proposed mechanism for the deoxygenation of thiazole N-oxide containing oligonucleotides during DNA synthesis.

Table 2.3. Semiempirical (PM3) computational analysis of thiazole and thiazole N-oxide nucleoside analogs.

<table>
<thead>
<tr>
<th>BASES</th>
<th>Surface Area (Å²)</th>
<th>Polarizability (Å³)</th>
<th>Dipole Moment (Debyes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytosine</td>
<td>120.7</td>
<td>10.9</td>
<td>6.58</td>
</tr>
<tr>
<td>Thymine</td>
<td>134.5</td>
<td>11.9</td>
<td>4.28</td>
</tr>
<tr>
<td>Thiazole</td>
<td>91.8</td>
<td>10.0</td>
<td>1.42</td>
</tr>
<tr>
<td>Thiazole N-oxide</td>
<td>100.9</td>
<td>10.6</td>
<td>4.57</td>
</tr>
</tbody>
</table>
polarizability, and dipole moment values obtained. The calculated values can be seen in Table 2.3.

While the thiazole analog falls terribly short of the pyrimidines in the value of each parameter, with the simple addition of the N-oxide the values of the calculated parameters become very close to those of the pyrimidines. The charge of the N-oxide adds a great deal to the dipole moment. In order to measure the charge difference between O2 atoms of cytosine and thymine compared with thiazole N-oxide partial atomic charges were calculated semiempirically (PM3) and the electrostatic potential mapped (Figure 2.18). It is obvious that the charge of the N-oxide is much larger than that of the O2 of cytosine and thymine. From the electrostatic potential maps (Figure 2.18) the direction of the dipole of thiazole N-oxide more closely mimics that of cytosine rather than thymine: therefore, a more favorable base pair should be expected with guanine. A more favorable base pair with guanine is also expected when inspecting the hydrogen bonding patterns of the natural base pairs. O2 of thymine does not form a hydrogen bond in an adenine-thymine base pair, but the O2 of cytosine does interact with the amino group of N2 of guanine.

Preliminary Molecular Dynamics analysis of a Dickerson dodecamer containing a thiazole N-oxide-guanine base pair was performed. The Molecular Dynamics structure of the Dickerson dodecamer with thiazole N-oxide incorporated can be seen in Figure 2.19 compared with the normal B-type Dickerson dodecamer. The double helix is distorted slightly with the incorporation of the nucleoside analog. Due to the thiazole ring and the amino group of guanine being moved out of the plane of the π stack the planarity of the base pairs directly above and below the thiazole N-oxide-guanine base pair are also distorted. The distance between the
Figure 2.18. Calculated partial atomic charges and electrostatic potential maps of cytosine, thymine, thiazole, and thiazole N-oxide. In the electrostatic potential maps blue indicates a negative charge while red indicates a positive charge. SYBYL MOPAC calculations at the PM3 level were performed and the electrostatic potential was mapped onto a Connolly surface using these charges.
Figure 2.19. Molecular Dynamics structures of the Dickerson dodecamer containing a cytosine-guanine base pair at the arrow (A) and a thiazole N-oxide-guanine base pair at the arrow (B). Molecular Dynamics simulations were performed by ramping to 300 K or 400 K for 1 ps and then stepping down to zero. These simulations used a 1 fs time step and constant pressure. SHAKE was used to constrain all hydrogens except those of the thiazole N-oxide-guanine base pair.
thiazole N-oxide-guanine base pair and the adenine-thymine base pair just below it is 3.9 Å, whereas this distance in the natural B-type DNA is 3.7 Å. A double helix shape aberration is also evident in the measurement of the C1'-C1' distances. The C1'-C1' distance of the cytosine-guanine base pair is 10.73 Å while this distance measures 11.20 Å in the thiazole-N-oxide-guanine base pair. An overall fit of the nucleotide monomers gives an RMS value of 1.72 showing that the double helix must change dramatically throughout the molecule to accommodate the thiazole N-oxide analog. These distortions lead to an overall reduced stability of the double helix.

The cytosine-guanine and thiazole N-oxide-guanine base pairs were extracted from the aforementioned DNA molecules and are shown in Figure 2.20. The first structural element that can be seen in these base pairs is the formation of a hydrogen bond between thiazole N-oxide and N1 of guanine rather than the expected N2 amino group. The measurements of the hydrogen bond distances vary dramatically. The hydrogen bond distance between the thiazole N-oxide and guanine is 2.56 Å, whereas the hydrogen bonds in the cytosine-guanine base pair measure 1.73 Å, 1.83 Å, and 1.73 Å. The longer length of the hydrogen bond formed between thiazole N-oxide and guanine means that this noncovalent interaction imparts much less stability energetically to the base pair. The increase in hydrogen bond length is due to the size of the thiazole ring. The thiazole ring is much smaller than a cytosine ring which forces the sugar to base torsional angle to be strained. It is apparent that although this analog perturbs the double helix, the hydrogen bond between the N-oxide oxygen and the hydrogen of N1 of guanine allows the two nucleotides to stay within close contact and allow the π system of the thiazole ring to base stack with the base pairs directly above and below it.
It has been shown through molecular modeling and Molecular Dynamics simulations that thiazole $N$-oxide pairs very strongly with guanine. The hydrogen bond formed between these two bases is of utmost importance in the stability of DNA.

![Diagram](image1)

**Figure 2.20.** Cytosine-guanine (A) and thiazole $N$-oxide-guanine (B) base pairs extracted from the Molecular Dynamics structures in Figure 2.19.

### 2.3.3 Thermodynamic Analysis

Thermodynamic analysis was performed with thiazole and thiazole $N$-oxide incorporated into a Dickerson dodecamer seen below:

$$5\prime-\text{CGCQAA}T\text{TYGC}G-3\prime$$
$$3\prime-\text{GCGYTTAA}Q\text{CG}G-5\prime$$

where $Q =$ cytosine, thymine, thiazole or thiazole $N$-oxide and $Y =$ adenine or guanine.
Table 2.4. Thermodynamic properties of the Dickerson dodecamer (5’-CGCQAATTTYGCG-3’) containing thiazole and thiazole N-oxide nucleoside analogs.

<table>
<thead>
<tr>
<th>Q-Y</th>
<th>Tm (100 µM)</th>
<th>-ΔTm/sub.</th>
<th>ΔH (kcal/mol)</th>
<th>ΔS (cal/K·mol)</th>
<th>ΔG37 °C (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-G</td>
<td>67.9</td>
<td>----</td>
<td>-55.43</td>
<td>-144.16</td>
<td>-10.72</td>
</tr>
<tr>
<td>T-A</td>
<td>67.1</td>
<td>----</td>
<td>-79.92</td>
<td>-216.87</td>
<td>-12.65</td>
</tr>
<tr>
<td>Z-G</td>
<td>16.5</td>
<td>-25.7</td>
<td>-53.18</td>
<td>-164.74</td>
<td>-2.08</td>
</tr>
<tr>
<td>Z-A</td>
<td>11.5</td>
<td>-27.8</td>
<td>-35.79</td>
<td>-107.01</td>
<td>-2.60</td>
</tr>
<tr>
<td>X-G</td>
<td>26.6</td>
<td>-20.7</td>
<td>-141.74</td>
<td>-442.97</td>
<td>-4.35</td>
</tr>
<tr>
<td>X-A</td>
<td>&lt;0</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The experimentally determined Tm and thermodynamic properties of enthalpy, entropy, and free energy can be seen in Table 2.4. Thiazole and thiazole N-oxide are very destabilizing to the duplex as can be seen from the value of -ΔTm/substitution. The loss of stabilization of the duplex is due mainly to reduction of size of the thiazole heterocycle as compared to the pyrimidine ring of thymine and cytosine.\textsuperscript{2.11} The decreased surface area of the ring leads to a loss of intra- and especially interstrand base stacking which distorts the base stack and destabilizes the duplex. A net increase of 11.4 °C is seen with the addition of the N-oxide to the thiazole ring in the dodecamer due largely to the formation of a hydrogen bond between the N-oxide oxygen and the N2 amino group or N1 of guanine (as indicated by molecular dynamics analysis). As shown through computational analysis the large charge and dipole moment of the N-oxide allows hydrogen bonding with guanine. This increase in stability is also due to the increase in solvation evident in the entropy increase of the duplex. The hydrogen bonding and solvation of thiazole N-oxide allows for increased stability of the duplex over thiazole. The melt
curve of the duplex containing a thiazole N-oxide-adenine base pair seems to show no duplex formation. This is presumably caused by the inability of adenine to form a hydrogen bond with thiazole N-oxide. Another possible reason for the lack of duplex formation is the ability of thiazole N-oxide to be highly solvated, which may possibly compete with the limited base stacking of the thiazole ring. This allows no formation of a thiazole N-oxide-adenine base pair which in turn does not allow the formation of the duplex resulting in an uncooperative melt of the duplex.

2.3.4 DNA Polymerase Recognition

In order to test the importance of O2 of pyrimidines to DNA polymerase fidelity an extension assay was developed. The sequences used are below:

21nt primer: 5´-TAA TAC GAC TCA CTA TAG GGA
23nt primer: 5´-TAA TAC GAC TCA CTA TAG GGA GA
24G primer: 5´-TAA TAC GAC TCA CTA TAG GGA GAG
24A primer: 5´-TAA TAC GAC TCA CTA TAG GGA GAA
template: 3´-ATT ATG CTG AGT GAT ATC CCT CTQ GTC A-5´
where Q = thymine in T-template, cytosine in C-template, thiazole in Z-template, or thiazole N-oxide in X-template.

The above primers were designed to test many proposed premises about DNA replication by DNA polymerase. The 21nt primer is three bases away from the analog in the template: therefore, the formation of the A-form DNA can occur before the polymerase encounters the analog. It has been noted that the last three base pairs of a primer-template complex are in A-form rather than the more usual B-form. The question probed was whether polymerase efficiency is greater if this transition is allowed to occur before the interruption of the duplex by the analog. The 23nt primer was designed so as to probe the insertion step (Figure 2.12). The next dNTP to be incorporated is opposite the analog: therefore, direct minor groove contacts in the first single stranded base of the template strand can be explored. The 24nt primers are
designed such that the extension step (Figure 2.12) can be investigated. The final base pair in the primer-template complex is thought to contain many important minor groove contacts with the polymerase active site: an asparagine with the primer strand and a glutamine with the template strand. The comparison of thiazole with no hydrogen bonding ability and thiazole N-oxide with considerable hydrogen bonding ability would explore these template contacts in the final base pair of the double stranded primer-template complex.

The 23nt primer was extended in the presence of all four dNTPs with Taq and Kf(exo-) DNA polymerases. Additionally polymerase extensions were performed in the absence of dATP or dGTP in order to probe which base was inserted opposite the analogs more efficiently. This was also investigated more quantitatively with kinetic single nucleotide insertion assays. These assays were designed such that with the 23nt primer only one dNTP is offered to the DNA polymerase at varying concentrations yet keeping time as well enzyme and DNA concentrations constant. The percentage extension product formed can then be divided by the time (in minutes) to obtain the velocity of insertion for each dNTP concentration. With this value a plot of dNTP concentration versus dNTP concentration divided by velocity can be obtained. This plot is termed a Hanes-Woolf plot which is a linear plot of a Michaelis-Menten enzyme kinetics plot. Michaelis-Menten kinetics is based on the following model:

\begin{equation}
E + S \overset{k_+}{\rightleftharpoons} ES \rightarrow E + P
\end{equation}

where E = enzyme, S = substrate, ES = an enzyme-substrate complex, and P = product formed.

The Michaelis-Menten equation is:

\begin{equation}
V = V_{\text{max}} \left( \frac{[S]}{[S] + K_m} \right)
\end{equation}
From the Hanes-Woolf plot $V_{\text{max}}$ and $K_m$ values are attained. The $K_m$ value is termed the Michaelis constant and is the concentration of substrate at which there is complete formation of the ES complex. The $V_{\text{max}}$ value is a measure of the maximal velocity at which $K_m$ is reached. $V_{\text{max}}$ divided by $K_m$ then defines the efficiency of the enzyme. These values have been obtained for single nucleotide extension of dATP and dGTP on the T, C, and Z templates.

The results of the extension reactions described above are seen in Figure 2.21. These extensions were carried out with both $K_f$(exo-) and Taq polymerases. When extending the 23nt primer, full extension is seen with both the natural base templates. Contrastingly the thiazole template gave only 45% extension with Taq polymerase and 68% extension with $K_f$(exo-). The lack of full-length extension product is caused by the stalling of the polymerase at the addition of a dNTP when thiazole is in the final base pair of the primer-template duplex. A dNTP is incorporated opposite thiazole such that there is a 40% stall product with Taq and 12% with $K_f$(exo-) formed at 24nt. This indicates that Taq is much more restrictive when reading the final thiazole containing base pair and cannot as efficiently incorporate the next dNTP opposite the next base after thiazole in the template. The thiazole N-oxide template gave only 22% full-length extension product with Taq and 23% with $K_f$(exo-) and the stall product (24nt) is formed in 31% yield with Taq and 26% yield with $K_f$(exo-). The main cause of loss of full-length extension product is the insertion of a dNTP opposite thiazole N-oxide. While the loss of full-length extension product with thiazole in the template can be attributed to incorporation of a dNTP opposite the next template base after thiazole, the loss of full-length extension product with the thiazole N-oxide template is due to the incorporation of a dNTP opposite the analog. This could possibly be due to the fact that with no hydrogen bonds made with the glutamine in the active site that the A-rule is adhered to whereas if a hydrogen bond is present there is a
continual search for a suitable dNTP that meets the other requirements for dNTP incorporation. Another possibility may be that there is simply more water molecules around the $N$-oxide which sterically inhibits a dNTP being placed opposite thiazole $N$-oxide.

When extending the 23nt primer in the absence of dATP or dGTP it is evident that extension is much more efficient with the thiazole template when dATP is present and conversely, extension is more efficient with the thiazole $N$-oxide template when dGTP is present. To further probe this the 24nt primers were extended. Extension of the 24G primer with the thiazole template was less efficient than with the 24A primer, yet extension of the 24G primer with the thiazole $N$-oxide template was much more efficient than extension with the 24A primer. The results of these assays possibly indicate that thiazole is encoding for adenine while thiazole $N$-oxide is encoding for guanine. A possibility for the reduced extension product seen with extension of the 23nt primer on the thiazole $N$-oxide template may be that an adenine is being inserted as per the A-rule when extension of the next base pair would be more efficient if a guanine was inserted. These hypotheses were further investigated with the kinetic analysis of single base insertion (vide infra).

The final extension assay is extension of the 21nt primer. With both polymerases assayed there is more product formed than with the 23nt primer. This indicates that the polymerases may have much more ability to convert the modified sequence structure to A-form if the modification does not begin in the active site. If the formation of A-form DNA does enhance the importance of stacking this explains the decreased amount of product compared with the natural sequences because the thiazole ring is much smaller: therefore, the area of stacking is much smaller.
Figure 2.21. Extension reactions performed with templates containing thymine (T), cytosine (C), thiazole (Z), and thiazole N-oxide (X). The percentages directly under the electrophoresis lanes correspond to the percentage of full-length product. The percentage under the full-length product percentage is the percentage of stall product (24nt for the 23nt primer and 23nt for the 21nt primer). The third percentage in the case of the 21nt primer is the percentage of stall product formed at 24nt. Extension reactions were performed in 20 µL with 200 nM radiolabeled primer and template, and 200 µM solution of all four dNTPs, polymerase and the supplied buffer. (Taq: 100 mM KCl, 100 mM (NH₄)₂SO₄, 200 mM Tris-HCl, 20 mM MgSO₄, 1% Triton X-100 pH 8.8). An additional 8 mM MgCl₂ was added to the reaction. Kf(exo-): 100 mM Tris-HCl pH 7.5, 5 mM MgCl₂, 7.5 mM DTT.)
Results of the kinetic single base insertion reactions are seen in Table 2.5. Insertion of adenine opposite thymine is almost 250-fold more efficient than insertion of guanine opposite adenine forming a thymine-guanine mismatch. The formation of a cytosine-guanine base pair is more than 150-fold more efficient than the formation of a cytosine-adenine mispair. The formation of a thiazole-adenine base pair is only 4-fold less efficient than formation of adenine-thymine base pair and insertion of guanine opposite thiazole is 34-fold less efficient than the formation of a cytosine-guanine base pair. It is interesting to note that the formation of base pairs containing thiazole is more efficient than the formation of mispairs. This seems to contradict the extension reactions lacking dATP or dGTP and the extensions with the 24G and 24A primers. More full-length extension product is formed with the templates containing thymine or cytosine than those containing thiazole even though extension of the natural templates would require the formation of mismatches or already contain a mismatch at the 3´-end of the primer strand. This can be explained by the values of K_m and V_max for thiazole. The efficiency of the incorporation of adenine and guanine opposite thiazole is mainly gained from the velocity of base pair formation. The K_m value for adenine incorporation opposite thiazole is similar to that of a cytosine-adenine base pair while the K_m for guanine incorporation opposite thiazole is much larger than any of the other values measured. These values indicate that the polymerase must go through a lengthy trial and error process of binding many dNTPs before catalyzing the formation of the phosphodiester bond very quickly, especially in the case of a thiazole-adenine base pair. Therefore, in the single nucleotide insertion reactions, where only one dNTP is available to the DNA polymerase, it is perceivable that the efficiency of a thiazole-adenine or thiazole-guanine base pair would be greater than a natural mismatch but in the
Table 2.5. Kinetic parameters measured for single base insertion of dATP or dGTP opposite thymine, cytosine, and thiazole.

<table>
<thead>
<tr>
<th>Template Base</th>
<th>dNTP</th>
<th>Kₘ (µM)</th>
<th>Vₘₐₓ (%/min)</th>
<th>Efficiency (Vₘₐₓ/Kₘ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T dATP</td>
<td>2.3(±0.7)</td>
<td>19.9(±0.9)</td>
<td>8.8 x 10⁶</td>
<td></td>
</tr>
<tr>
<td>T dGTP</td>
<td>85.7(±16.4)</td>
<td>3.1(±0.4)</td>
<td>3.6 x 10⁴</td>
<td></td>
</tr>
<tr>
<td>C dATP</td>
<td>25.8(±11.7)</td>
<td>5.2(±1.8)</td>
<td>2.0 x 10⁴</td>
<td></td>
</tr>
<tr>
<td>C dGTP</td>
<td>2.8(±0.3)</td>
<td>9.4(±0.6)</td>
<td>3.3 x 10⁶</td>
<td></td>
</tr>
<tr>
<td>Z dATP</td>
<td>20.1(±5.4)</td>
<td>42.0(±15.4)</td>
<td>2.1 x 10⁶</td>
<td></td>
</tr>
<tr>
<td>Z dGTP</td>
<td>150.0(±2.0)</td>
<td>14.6(±4.8)</td>
<td>9.7 x 10⁴</td>
<td></td>
</tr>
</tbody>
</table>

extension reactions when multiple dNTPs are available the polymerase may statistically be able to insert a more energetically stable base.

2.4 Conclusions

By studying pyrazole based analogs and thiazole and thiazole N-oxide analogs much has been learned about designing new base analogs for the use in mutation detection strategies. A lack of base pair geometry while not as important to DNA stability as electrostatic interactions and stacking area does seem to play a role in DNA polymerase recognition. Therefore, when designing base analogs for mutation detection strategies more emphasis should be placed on base pair geometry (C₁'-C₁' distances and C₁' bond angles). Therefore, the size of the analog and accessible atoms for hydrogen bonding should be designed carefully. From the analysis of thiazole and thiazole N-oxide the minor groove hydrogen bond sites are important for dNTP selection in DNA insertion and extension so these sites should be mimicked in designing analogs for mutation detection.
2.5 References


2.2 Saenger, W. Principles of Nucleic Acid Structure; Springer-Verlag: New York, 1984; 556.


2.7 Spratt, T. E. Identification of hydrogen bonds between Escherichia coli DNA polymerase I (Klenow fragment) and the minor groove of DNA by amino acid substitution of the polymerase and atomic substitution of the DNA. Biochemistry 2001, 40, 2647-2652.


<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
<th>Journal</th>
<th>Year</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.37</td>
<td>Johnson, W. T.; Zhang, P. M.; Bergstrom, D. E.</td>
<td>The synthesis and stability of oligodeoxyribonucleotides containing the deoxyadenosine mimic 1-(2'-deoxy-beta-D-ribofuranosyl)imidazole-4-carboxamide.</td>
<td>1997</td>
<td>559-567</td>
</tr>
<tr>
<td>2.38</td>
<td>Klewer, D. A.; Zhang, P. M.; Bergstrom, D. E.; Davisson, V. J.; LiWang, A. C.</td>
<td>Conformations of nucleoside analogue 1-(2'-deoxy-beta-D-ribofuranosyl)-1,2,4-triazole-3-carboxamide in different DNA sequence contexts.</td>
<td>2001</td>
<td>1518-1527</td>
</tr>
<tr>
<td>2.41</td>
<td>Spratt, T. E.</td>
<td>Klenow fragment DNA interaction required for the incorporation of nucleotides opposite guanine and O-6-methylguanine.</td>
<td>1997</td>
<td>13292-13297</td>
</tr>
<tr>
<td>2.42</td>
<td>Terashima, I.; Matsuda, T.; Fang, T. W.; Suzuki, N.; Kobayashi, J. et al.</td>
<td>Miscoding potential of the N-2-ethyl-2'-deoxyguanosine DNA adduct by the exonuclease-free Klenow fragment of Escherichia coli DNA polymerase I.</td>
<td>2001</td>
<td>4106-4114</td>
</tr>
<tr>
<td>2.44</td>
<td>Guo, M. J.; Hildbrand, S.; Leumann, C. J.; McLaughlin, L. W.; Waring, M. J.</td>
<td>Inhibition of DNA polymerase reactions by pyrimidine nucleotide analogues lacking the 2-keto group.</td>
<td>1998</td>
<td>1863-1869</td>
</tr>
<tr>
<td>2.45</td>
<td>Lutz, S.; Burgstaller, P.; Benner, S. A.</td>
<td>An in vitro screening technique for DNA polymerases that can incorporate modified nucleotides. Pseudothymidine as a substrate for thermostable polymerases.</td>
<td>1999</td>
<td>2792-2798</td>
</tr>
<tr>
<td>2.46</td>
<td>Moran, S.; Ren, R. X. F.; Kool, E. T.</td>
<td>A thymidine triphosphate shape analog lacking Watson-Crick pairing ability is replicated with high sequence selectivity.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the National Academy of Sciences of the United States of America 1997, 94, 10506-10511.


Chapter 3: Nucleoside Backbone Analogs

3.1 Introduction

In the point mutation detection assay PCR/RE/LDR, nucleoside analogs were shown to increase the efficiency of base conversion for the formation of restriction enzyme recognition sites, but the desired sensitivity of the assay was still not achieved. This was due to the poor base pairing of the analog which causes “primer slippage” as well as the analogs may convert the base in question to multiple other bases which causes mismatches during PCR with the natural base primers after conversion.\(^3\)\(^1\) In other point mutation detection strategies, mismatches are strategically placed and must be retained to ensure that false positives are not generated. To increase the sensitivity of the PCR/RE/LDR assay and other mutation detection strategies, higher fidelity DNA polymerases that have exonuclease (proofreading) activity must be used. The use of backbone analogs would be a solution to the problem of removal of nucleoside analogs or strategically placed mismatches by the exonuclease domain of higher fidelity enzymes. There are a number of backbone analogs which have been synthesized and proven resistant to nucleases.\(^3\)\(^2\)\(^,\)\(^3\)\(^3\) One such backbone modification is the N3´→ P5´ phosphoramidate in which nitrogen replaces the 3´-bridging oxygen of the phosphate backbone (3´-NHP(O)(O⁻)O-5´) (Figure 3.1).

![Figure 3.1. Structure of the phosphodiester (A), phosphorothioate (B), and phosphoramidate (C) backbone modifications.](image)

\(^3\)\(^1\) In other point mutation detection strategies, mismatches are strategically placed and must be retained to ensure that false positives are not generated.
3.1.1 Backbone Analogs-The Literature

Many of the nucleoside backbone analogs were synthesized and tested for nuclease resistance due the advent of their use in antisense and antigene technology which is currently being tested for its use in the treatment of diseases such as cancer, AIDS, and inflammation. This new technology is based on the concept of inhibition of protein synthesis through the blockage of information transfer by binding oligonucleotides to either mRNA (antisense) or genomic DNA (antigene). Backbone analogs have since been used in the study of DNA structure and function, and in site-directed mutagenesis.

The backbone analogs synthesized contain from a simple single atom replacement to complete replacement of the backbone such as in peptide nucleic acids (PNAs). PNAs are comprised of a peptide backbone with DNA bases attached at the \( \alpha \)-carbon. The most studied of the backbone analogs though are the phosphorothioates in which a non-bridging oxygen of the backbone phosphate is replaced by a sulfur atom (Figure 3.1). This atom substitution changes the overall charge of the modified DNA molecule yet does not change the hybridization properties drastically. Phosphorothioates were also found to inhibit the cleavage activity of nuclease. The cleavage of a phosphorothioate containing oligonucleotide was found to be approximately ten times slower than the cleavage of a natural phosphodiester containing oligonucleotide using snake venom phosphodiesterase. The synthesis of phosphorothioates also leads to the formation of diastereomers, \( R_P \) and \( S_P \), which adds an extra step of purification if they are to be used because most nuclease can degrade one or other of the diastereomers. For instance, snake venom phosphodiesterase was found to cleave only the \( R_P \) isomer while nuclease S1 and P1 cleave only the \( S_P \) isomer. Phosphorothioate linkages were also found to be resistant to the exonuclease activity of \( E. coli \) DNA polymerase I, but T4 DNA
polymerase cleaves these modified linkages at similar rates as the natural phosphodiester backbone.\textsuperscript{3,9} Due to this limited exonuclease resistance, phosphorothioate backbone modifications have only been used in site-directed mutagenesis experiments and PCR.\textsuperscript{3,13-17} It was found that the use of phosphorothioates did inhibit the exonuclease activity of higher fidelity enzymes yet the polymerase activity of the enzymes was decreased. Further study of this modification by capillary electrophoresis found that a single 3´-modification afforded more exonucleolytic protection than multiple modifications.\textsuperscript{3,18}

The N3´→P5´ phosphoramidate backbone modification has been discussed as the second generation of antisense therapy due to its unique structural properties, increased hybridization, and increased nuclease resistance. Oligonucleotides with every linkage replaced by a phosphoramidate modification have been studied intensely. Circular dichroism (CD), Molecular Dynamics, NMR, and X-ray crystallography all indicate that these oligonucleotides exhibit a helical conformation more consistent with A-form DNA than the more natural B-form.\textsuperscript{3,19-22} This is enough to explain the increased hybridization to mRNA as RNA displays an A-form conformation. Further structural analysis indicates that this A-form conformation is adopted due the preferred C3´-endo furanose conformation.\textsuperscript{3,20} Increased hybridization can also been seen in the increased melting temperatures (T\textsubscript{m}) of oligonucleotides containing phosphoramidate linkages. These modifications were found to increase the T\textsubscript{m} by as much as 27.5 °C.\textsuperscript{3,19} Also indicated by structural analysis is the increased hydration of the grooves due to the more hydrogen bond donating character of the phosphoramidate as compared with the more hydrogen bond accepting character of the natural phosphodiester backbone. The increased hydration causes a strong anomeric effect between the 3´-nitrogen lone pair and the antibonding orbital of the adjacent P-O5´ bond (Figure 3.2). The phosphoramidate modification has also shown
increased nuclease resistance.\textsuperscript{3,23,3,24} A fully modified oligonucleotide resisted cleavage by snake venom phosphodiesterase a little more than an order of magnitude compared with the natural phosphodiester backbone (26 h compared with 20 min). The only explanation for this resistance has been the anomeric effect stated above.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{anomeric_effect.png}
\caption{The anomeric effect caused by the addition of the 3′-nitrogen of the phosphoramidate modification. The lone pair of the nitrogen interacts with the $\sigma^*$ antibonding orbital of the P-O5′ bond. The structure is stabilized by the chlorine ion and the C$_3$-endo furanose conformation.}
\end{figure}

Due to the increased protection to nucleases afforded oligonucleotides by the N3′→P5′ phosphoramidate backbone modification and the abovementioned finding that a single 3′-phosphorothioate modification affords more protection than multiple modifications the phosphoramidate was proposed for use in point mutation detection strategies. Also of interest was the structural finding that the phosphoramidate-modified oligonucleotides were of the A-form conformation. As stated in Chapter 2 the catalysis site of DNA polymerase converts the final few base pairs to A-form. Due to the conformation of a primer-template duplex containing a 3′-phosphoramidate modification in the primer strand already being A-form the problem of decreased DNA polymerase fidelity found with phosphorothioate modified oligonucleotides would be circumvented.
3.1.2 Synthesis

N3’→ P5’ phosphoramidate backbone modification containing oligonucleotides have been synthesized in a number of ways. One method was through chemical ligation in aqueous media. An amine-exchange procedure was also reported for the formation of a phosphoramidate linkage directly on a solid support. The 3’-aminonucleoside was attached to the solid support and was then reacted with the next phosphoramidite in the presence of tetrazole. Another approach coupled preformed phosphoramidate dimers on a solid support. One method of dimer formation included the coupling of a phosphotriester with a 3’-azidonucleoside in a Staudinger reaction forming an imine intermediate which is followed by a Michaelis-Arbuzov transformation enhanced by LiCl transforming the imine to an amide. The resulting dimer product was then converted to a phosphoramidite through standard phosphitilation for use in solid phase oligonucleotide synthesis. Another method of dimer formation coupled a 3’-aminonucleoside with a 3’-O-Levulinyl protected nucleoside converted to a phosphonate ester in the presence of carbon tetrachloride and triethylamine in an Atherton-Todd procedure. Conventional phosphitilation following 3’-deprotection yielded the dimer phosphoramidite for use in solid phase oligonucleotide synthesis. This method was later converted to a synthesis performed directly on the solid support. The attached nucleoside was converted to a phosphonate ester which was then coupled with a 3’-aminonucleoside in the presence of carbon tetrachloride/triethylamine to form the phosphoramidate linkage. Repeated couplings were performed to yield the all phosphoramidate oligonucleotide. This solid phase procedure was later performed using an oxidative phosphorylation reaction in which the coupling of the 3’-aminonucleoside with the phosphonate ester is activated by iodine.
As the stepwise yields for the above reported oligonucleotides were similar the solid phase methods using carbon tetrachloride/triethylamine and iodine as activation agents of coupling were used for the synthesis of the oligonucleotides used in the experiments described below. The 3′-aminothymidine was synthesized by reduction of the azido group with sodium borohydride of a 5′-dimethoxytrityl-protected 3′-azidothymidine (Scheme 3.1). The 3′-aminothymidine was then coupled to a solid supported nucleoside phosphoramidite which had been converted to a 5′-H-phosphonate-2-cyanoethyl diester (Scheme 3.2). Conventional steps for solid phase oligonucleotide synthesis (Scheme 3.3) were then performed to yield an oligonucleotide containing only one penultimate 3′ N3′→ P5′ phosphoramidate backbone modification.

**Scheme 3.1.** Reduction of 3′-DMT azidothymine.
Scheme 3.2. Formation of the phosphoramide linkage by solid phase synthesis.
Scheme 3.3. Solid phase DNA synthesis by the cyanoethyl phosphoramidite approach.
3.2 Exonuclease Resistance of Phosphoramidate

In order to test the extension efficiency and exonuclease resistance of 3′-phosphoramidate modifications and compare it to that of a 3′-phosphorothioate modification and the natural phosphodiester backbone, three primers containing a natural phosphodiester bond, a phosphorothioate bond, or a phosphoramidate bond in the penultimate 3′-position were extended on a template with either a C-G match (Match template) or a C-A mismatch (Mismatch template). Template and primer sequences are shown in Figure 3.3. The extensions were performed with Taq DNA polymerase which has no exonuclease activity and Vent and Pfu DNA polymerases which do contain 3′-5′ exonuclease domains. From these extension reactions could be quantitated the amount of extension product for each primer extended with each enzyme to determine the efficiency by which the polymerase extends the phosphoramidate and compare this with the efficiency of the phosphodiester and phosphorothioate primers. A time-dependent cleavage assay was also developed in which Vent and Pfu DNA polymerases were provided dNTPs that would not form Watson-Crick base pairs with the next two bases in the template strand. This assay allowed the qualitative analysis of the rate of phosphorothioate and phosphoramidate modification cleavage.

5′-GCT CCC CTA CCA TGA CTT TAT TxC-3′
  x= phosphodiester (O), phosphorothioate (S), or  phosphoramidate(N)

**Match template:**
5′-GGA AGA CCG ATC TTC GAA ATA A GG TCA TGG TAG GGG AGC-3′

**Mismatch template:**
5′-GGA AGA CCG ATC TTC GAA ATA AAG TCA TGG TAG GGG AGC-3′

*The two templates have a G-A base difference at position 17.*

**Figure 3.3.** Sequences of the primers as well as the match and mismatch templates used in the assay to test the exonuclease resistance of the phosphoramidate modification.
When the primer is annealed and extended to the match template there is a BslI restriction endonuclease site which can be utilized to check if the enzyme correctly extended the primer strand (Figure 3.4). The annealed and extended primer/mismatch duplex contained a TaqI restriction endonuclease site if and only if the exonuclease domain of the DNA polymerase removed the mismatched base, which is linked via the modification, from the primer. By measuring the extent of TaqI cleavage the extent to which the mismatched base was removed can be measured. The amount of cleavage is equivalent to the exonuclease resistance of the phosphoramidate and phosphorothioate modifications which can be compared to that of the natural phosphodiester backbone.

The extended “match” template has a BslI restriction endonuclease cleavage site:

5’-GCT CCC CTA CCA TGA CTT TAT T CC GAA GAT CGG TCT TCC-3’ 39nt
3’-CGA GGG GAT GGT ACT GAA ATAA GG CTT CTA GCC AGA AGG-5’

The extended “mismatch” template has a TaqI restriction endonuclease cleavage site only if the exo+ polymerase can correct the mismatch:

5’-GCT CCC CTA CCA TGA CTT TAT T’TC GAA GAT CGG TCT TCC-3’ 39nt
3’-CGA GGG GAT GGT ACT GAA ATAA AG CTT CTA GCC AGA AGG-5’

Figure 3.4. The assay designed to test the exonuclease resistance of the phosphoramidate modification.
Figure 3.5 shows the radiolabeled primers and resulting products formed in the assay separated on a polyacrylamide electrophoresis gel. The determination of extension efficiency of the primers by the polymerases indicated that both the phosphorothioate and phosphoramidate modifications were recognized by all of the polymerases tested. The extension efficiencies of all three primers annealed to the match template are quite similar indicating that the incorporation of these backbone modifications does not affect the catalysis performed by the polymerase. Although it was expected that even with a mismatch the phosphoramidate containing primer would be extended with greater efficiency due to the A-form conformation adopted, the extension efficiency of the phosphoramidate containing primer was greatly reduced with all three DNA polymerases. This is possibly due to the fact that the A-form conformation is adopted because the sugar pucker is changed to C3'-endo which may disrupt the formation of the closed conformation of the polymerase catalytic active site. Due to the observation that the phosphorothioate containing primer also reduced extension efficiency of the primer/mismatch duplex another possible explanation for the reduced efficiency may be that the primer spends more time in the exonuclease domain since the mismatch is not corrected. The phosphoramidate primer has more hydrogen bonding capability than the phosphorothioate primer which may make the binding of the former more favorable than the latter which indicates that the phosphoramidate may spend more time in the exonuclease domain than either the phosphodiester or the phosphorothioate primers.

In the evaluation of the exonuclease resistance of the backbone modifications the Bs/I restriction digest interestingly shows reduced cleavage of the phosphorothioate and phosphoramidate primers (Table 3.1). The modifications are contained within the reading frame of the restriction endonuclease which may decrease binding of the enzyme to these modified
Figure 3.5. Electrophoretic gels of the assay to test the exonuclease resistance of the phosphoramidate modification extended by A. Taq, B. Vent, or C. Pfu DNA polymerase. O=phosphodiester; S=phosphorothioate, N=phosphoramidate. Extension reactions were performed with 20 pM primer, 50 pM template, and 200 µM solution of all four dNTPs, AmpliTaq®, Vent, and Pfu polymerases and the supplied buffers. Extension of the match template was digested with BsuI restriction endonuclease (supplied buffer: 1 M NaCl, 500 mM Tris-HCl, 100 mM MgCl₂, 1 mM DTT, pH 7.9) at 55 ºC for 60 min. Extension of the mismatch template was digested with TaqαI restriction endonuclease (supplied buffer: 1 M NaCl, 100 mM Tris-HCl, 100 mM MgCl₂, pH 8.4) at 65 ºC for 60 min.
primers. The amount of *Taq*α*I* restriction endonuclease cleavage of the mismatched template extension product is equivalent to the extent of correction and therefore gives an indication of modification stability. With Vent DNA polymerase nearly complete cleavage of the phosphodiester primer can be compared with 75% cleavage of the phosphorothioate primer, and only 16% cleavage of the phosphoramidate primer. Although mismatch extension of the phosphorothioate and phosphoramidate primers was not very efficient with *Pfu* polymerase there was nearly complete cleavage of the phosphodiester primer, and only small amounts of cleavage of the phosphorothioate and phosphoramidate primers. These results indicate that the phosphoramidate primer affords the greatest stability and therefore the greatest protection against exonuclease activity.

**Table 3.1. Quantitation of percentages of extension and cleavage of each primer.**

### Match Template:

<table>
<thead>
<tr>
<th></th>
<th><em>Taq</em> Polymerase</th>
<th>Vent Polymerase</th>
<th><em>Pfu</em> Polymerase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% extension</td>
<td>% cleavage</td>
<td>% extension</td>
</tr>
<tr>
<td>O</td>
<td>99%</td>
<td>59%</td>
<td>99%</td>
</tr>
<tr>
<td>S</td>
<td>93%</td>
<td>41%</td>
<td>93%</td>
</tr>
<tr>
<td>N</td>
<td>96%</td>
<td>1%</td>
<td>85%</td>
</tr>
</tbody>
</table>

### Mismatch Template:

<table>
<thead>
<tr>
<th></th>
<th>% extension</th>
<th>% cleavage</th>
<th>% extension</th>
<th>% cleavage</th>
<th>% extension</th>
<th>% cleavage</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>53%</td>
<td>&lt;1%</td>
<td>99%</td>
<td>99%</td>
<td>80%</td>
<td>58%</td>
</tr>
<tr>
<td>S</td>
<td>67%</td>
<td>&lt;1%</td>
<td>72%</td>
<td>75%</td>
<td>10%</td>
<td>7%</td>
</tr>
<tr>
<td>N</td>
<td>37%</td>
<td>&lt;1%</td>
<td>39%</td>
<td>16%</td>
<td>3%</td>
<td>1%</td>
</tr>
</tbody>
</table>
Figure 3.6. Electrophoretic gels of the timed analysis of exonucleolytic cleavage of the phosphodiester (O), phosphorothioate (S), and phosphoramidate (N) primers. Reactions were performed as extension reactions as described in Figure 3.5 without the addition of dATP and dGTP.

Analysis of the time-dependent cleavage assay shows cleavage of the phosphodiester primer after 5 s, cleavage of the phosphorothioate primer after 5 min and no cleavage of the phosphoramidate primer even after 30 min with *Pfu* DNA polymerase (Figure 3.6). Once again it was shown that the phosphoramidate modification exhibits greater exonuclease resistance when compared with the natural phosphodiester or the phosphorothioate modification.

There are a number of possibilities as to the reason for the increased exonuclease resistance of the phosphoramidate backbone modification. In the mechanism of exonucleolytic cleavage introduced in Chapter 1 a metal ion binds to the 3′-oxygen. The metal ion provided in the assays described above was Mg$^{2+}$, which may have a higher affinity for binding oxygen rather than nitrogen due to the lower electronegativity of nitrogen. As was mentioned earlier in
respect to the structural studies done on phosphoramidate containing oligonucleotides it was found that there is more hydration in the grooves due to the greater hydrogen bonding donating properties of nitrogen as compared with oxygen. Also in the structural studies the anomeric effect between the antibonding orbital of the P-O5′ bond and the lone pair of nitrogen caused a configuration that allowed for the binding of chlorine ions to the nitrogen. Due to binding of chlorine ions and water there may be steric hindrance to the binding of the Mg^{2+} metal ion which is crucial for exonucleolytic cleavage to occur. A final possibility for the increased exonuclease resistance of the phosphoramidate modification may be due to the inability of the nitrogen to stabilize the negative charge generated after nucleophilic attack of the phosphorous by a hydroxide ion due to the decreased electronegativity of nitrogen compared with oxygen.

3.3 Conclusions

The N3′→ P5′ phosphoramidate backbone modification is resistant to the exonucleolytic cleavage of Vent and Pfu DNA polymerases. It has been shown that this backbone modification is more exonuclease resistant than the natural phosphodiester backbone and the much used phosphorothioate modification. This modification would be of great use in PCR primers, the PCR/RE/LDR mutation detection strategy, and other point mutation detection strategies utilizing strategically placed mismatches.

3.4 References


3.16 Nikiforov, T. T.; Rendle, R. B.; Kotewicz, M. L.; Rogers, Y. H. The Use of Phosphorothioate Primers and Exonuclease Hydrolysis for the Preparation of Single-


Chapter 4: Capillary and Microdevice Electrophoresis of LDR

4.1 Introduction

4.1.1 LDR of Colorectal Cancer

Point mutations occurring in the K-ras gene have been shown to be good biomarkers for colorectal cancers.\textsuperscript{4,1-4} 30-50\% of all colorectal cancers contain a point mutation in codons 12, 13, and 61 in the K-ras gene (Figure 4.1). These point mutations are conserved throughout the history of the disease which makes these mutations suitable for detection of these cancers. In the early stages of colorectal cancer there are low copy numbers of mutated DNA as compared with normal or wild type DNA. Early detection of colorectal cancer requires the detection of these low level mutations which is made difficult by the large amount of normal DNA present.

![Diagram showing locations of codons 12, 13, and 61 in the K-ras gene.]

**Figure 4.1.** Locations of codons 12, 13, and 61 in the K-ras gene.
The ligase detection reaction (LDR) was designed as a point mutation detection strategy in order to identify such low abundant mutations.\textsuperscript{4,5-8} In LDR (Figure 4.2) oligonucleotide primers (a common and all possible discriminating primers) are annealed adjacently to one strand of the target DNA. The adjoining discriminating and common primers are covalently joined by a thermostable DNA ligase to form an LDR product if the target nucleotide at the mutation site is complementary to the 3'-end of the discriminating primer. All possible mutations at a specific site can be analyzed by including all possible discriminating primers (of differing lengths) corresponding to the possible nucleotide substitutions. Generation of an LDR

\textbf{Figure 4.2.} Schematic representation of the ligase detection reaction (LDR).
product indicates the presence of a mutation while the size of the product indicates which specific substitution is present. LDR products were first analyzed by slab gel electrophoresis. This method made use of fluorescein labeled discriminating primers and a phosphorylated common primer, both in excess of the target DNA to which the primers were annealed. This excess of discriminating primers and common primer is required to ensure complete annealing of the target DNA and efficient ligation. This surplus of primers causes annealing competition when the assay is used for analysis of multiple mutations or multiplexing. Using this analysis method a 1:1000 ratio of mutant to wild type sequences could be discriminated.

In an effort to better analyze products of the LDR strategy capillary and microdevice electrophoresis was employed. These electrophoresis formats not only decrease time of analysis but also increase resolution.

### 4.1.2 Capillary and Microdevice Electrophoresis

Capillary electrophoresis (CE) consists of two buffer reservoirs with a separation capillary filled with a carrier electrolyte making electrical contact between them. Samples are loaded onto the capillary either electrokinetically, by applying an electric field from the sample reservoir to the waste reservoir, or hydrodynamically, by applying an external pressure to the sample container. The sample is then separated by the application of an electric field and detected by an online detector close to the outlet of the capillary. In order to separate DNA a sieving matrix must be used inside the capillary tube. These sieving polymers can either be cross-linked or linear, with the choice mandated by the required resolving power and the ease of use. While the cross-linked gels typically provide better resolving power than the linear gels, the linear gels can be replaced in the capillary while cross-linked gels cannot. The replacement of the linear gels is required due to the deleterious effects of the high electric field
on the gel. Detection can be accomplished by either ultraviolet absorbance or by fluorescence using intercalating dyes or covalently linked fluorophores for laser induced fluorescence (LIF) detection. Since most CE analyses use only nanoliter volumes of sample, highly sensitive detection is required; therefore in most cases LIF is used in CE applications.

CE has been used in the analysis of a multitude of mutation detection strategies. Some more recent examples include the detection of mutations in the *BRCA1* and *BRCA2* breast cancer susceptibility genes by the single strand conformation polymorphism (SSCP) and heteroduplex analysis (HDA) methods which employed the use of CE.\(^4,21\) The percentage of mutations detected was increased to 100% efficiency. CE was also used to analyze the products of a minisequencing assay to detect mutations in the *CYP21* gene, which causes a deficiency in the 21-hydroxylase enzyme.\(^4,22\)

The first microdevices developed for electrophoretic purposes were constructed on glass microscope slides and produced comparable resolution to conventional CE but at a fraction of the time.\(^4,23,4,24\) Electrophoresis on a planar chip builds on the inherent advantages associated with CE, including high surface-to-volume ratios produced by using very small channels for the separation, which permit the application of higher electric fields for faster separations. Microdevices have been applied to the separation of DNA in a variety of genetic applications ranging from the simple sizing of PCR products to mutation detection strategies such as SSCP and HDA.\(^4,25\) The simplest device format for performing electrophoresis is the twin T design, which contains a cross-like structure.\(^4,26,4,27\) (Figure 4.3) The “arms” of the T are used for electrokinetic injection of sample and are offset to define a fixed injection volume, while the longer channel is used for separation. A sieving matrix is introduced into the device as in CE. The switching of voltages from well to well allows the loading and separation of samples.
4.2 Capillary Electrophoresis: Model Studies and LDR Reactions

4.2.1 Introduction

CE has been applied to a number of different mutation detection strategies. A problem caused in the CE analysis of LDR reactions is the electroosmotic flow (EOF) caused by severe injection biases. These injection biases are due to the large excess of LDR primers in the reactions in comparison with LDR products produced. In order to detect the low concentrations of LDR products higher injection voltages and times were required which in effect caused a reduction in resolution. In order to choose a suitable matrix for LDR analysis and to prove that an LDR product could be resolved in a large excess of primers, model studies using synthetic oligonucleotides were performed. LDR reactions containing varying ratios of mutant and wild type templates were also analyzed by CE.

Figure 4.3. The twin T microdevice. A is the waste well, B is the sample well, C is the buffer reservoir and D is the separation well.
4.2.2 Model Studies

A 25 nucleotide oligonucleotide which was fluorescein labeled on its 5’-end was used to simulate the labeled discriminating primers while 44 and 51 nucleotide fluorescein labeled oligonucleotides were used to simulate the LDR products which were to be resolved. In an LDR reaction, which contains a 1:20 molar ratio of mutant DNA template to wild type template, the ratio of LDR product to excess discriminating primers becomes 1:2 if DNA ligase ligates annealed primers at 100% efficiency. In an LDR reaction that has a 1:1000 ratio of mutant to wild type templates, the ratio of LDR products to discriminating primers becomes 1:100 (if the reaction is 100% efficient). If the ligation efficiency falls below 100% then the ratio of LDR product to discriminating primer becomes much higher. Therefore, samples containing 1:10, 1:100, and 1:1000 molar ratios of 25 to 44 and 51 nucleotide model oligonucleotides were analyzed for optimization of separation conditions and testing of sieving matrices.

The matrices tried were the POP5 performance optimized polymer separation matrix (herein referred to as POP5), the eCap ssDNA 100-R Gel matrix (herein referred to as eCap), and the µ-PAGE-5 (5%T, 5%C) (indicating the amount of crosslinking present in the acrylamide capillary) 5T5C capillary (herein referred to as 5T5C). The POP5 and eCap gels are replaceable linear entangled polymers while the 5T5C capillary is a multiple run cross-linked capillary. Conditions for injection and separation were optimized separately for each matrix. While the 1:10 and 1:100 molar ratios of mock LDR product to mock discriminating primer were able to be separated using all of the matrices only the 5T5C was able to separate the 1:1000 ratio due to the large injection biases caused by extended times and voltages required during injection.

The use of the POP5 gel not only was unable to separate the 1:1000 model sample but there was severe loss of resolution if the gel was used for more than one injection. This required
refilling the capillary after each run. The total run time to separate the model samples with POP5 was found to be just over 30 minutes. Separation of the 1:100 model sample using POP5 can be seen in Figure 4.4. The eCap gel while also unable to separate the 1:1000 model sample did not show a loss in resolution until after five runs and the separation of the oligonucleotides was performed in approximately 28 minutes. Figure 4.5 is an electropherogram of the 1:100 model sample using the eCap matrix. The shorter separation time required combined with the ability of eCap to sustain up to five runs decreases the time requirement for analysis. The 5T5C capillary was able to separate all three of the model samples tested and with better resolution than the linear gels although with longer migration times (Figure 4.6). The migration times did vary among runs (The 44 fragment migrated at 37 min in the 1:100 sample while at 40 minutes in the 1:1000.) due to the required clipping of the capillary which is necessary due to bubbles and other abnormalities formed from the higher voltages and times required for injection. It can be seen in the electropherograms using 5T5C that the peak height of the 51-nucleotide fragment is not equivalent to that of the 44-nucleotide fragment although they were in equimolar concentrations in the sample. This is possibly a reflection of the different labeling efficiencies with fluorescein during synthesis or possibly another artifact of the severe injection biases.

4.2.3 LDR Reactions

LDR reactions were performed using DNA templates containing a G12V mutation in codon 12 of the K-ras gene. Three labeled discriminating primers were included to probe for all of the possible point mutations (G12V, G12A, G12D) at this location of codon 12. The discriminating primers were of 24-, 25-, and 26-nucleotides and the phosphorylated common primer was of 20-nucleotides in length. Sequences of the LDR primers used are shown in Figure
Figure 4.4. Analysis of 1:100 model sample using the POP5 matrix. The samples were run at 12 kV and capillary temperature was controlled at 50 °C.

Figure 4.5. Analysis of the 1:100 model sample using the eCap matrix. Samples were run at 12 kV at 30 °C.
Figure 4.6. 5T5C analysis of the A) 1:100 model sample and B) 1:1000 model sample. Samples were run at 250 V at 25 °C.
Discriminating Primers:

c12.2V  5'-Fluor CAA AAA CTT GTG GTA GTT GGA GCT GT-3’  26nt

c12.2A  5'-Fluor CAA AAC TTG TGG TAG TTG GAG CTG C-3’  25nt

c12.2D  5'-Fluor AAA ACT TGT GGT AGT TGG AGC TCA-3’  24nt

Common Primer:
c12 com-2  5’-pTGG CGT AGG CAA GAG TGC CT-3’  20nt

Figure 4.7. LDR primer sequences used in reactions for CE analysis.

4.7. Since only the G12V template was used in the reaction only a 46-nucleotide LDR product was expected. After cycling the LDR reactions they were desalted so as to decrease any further injection biases caused by the salts required in the reactions.

As the 5T5C capillary was found to perform the highest resolution separations as compared to the POP5 and eCap linear gels, this separation matrix was used in analysis of the LDR samples. An LDR blank sample was prepared by performing an LDR reaction containing no template (Figure 4.8). This sample was used as a negative control. All that is evident in the electropherogram is the large discriminating primer peak. LDR samples were prepared using both a 1:20 and 1:100 molar ratios of mutant to wild type DNA template. When the electropherograms of these samples are compared with that of the LDR blank sample the generation of the 46-nucleotide LDR product is apparent (Figure 4.9). As the amount of mutant template is decreased resolution between the discriminating primer peak and the LDR product peak is greatly reduced. As in the model studies, inconsistent migration rates were observed.
Figure 4.8. LDR blank sample for CE analysis.\textsuperscript{4,29} Samples were run as described in Figure 4.6.
Figure 4.9. A) 1:20 and B) 1:100 mutant to wild type LDR reactions for CE analysis.\textsuperscript{4.29} Samples were run as described in Figure 4.6.
4.3 Microdevice Electrophoresis: Model Studies

The microdevice used in these studies was of a modified twin T design (Figure 4.10) in which the separation channel was lengthened and contained two turns. The microchannel was hot-embossed in poly (methyl methacrylate) (PMMA) using a mold master produced by a LIGA process. The final device used was assembled by annealing a PMMA cover plate to the open face of the device at elevated temperatures in a circulating air oven. The separation channel was 10 cm in length and the side channels were 0.5 cm. There are four wells labeled A, B, C, and D. Well A is the sample well, B the waste well, C the buffer reservoir, and D the separation well. During injection a positive voltage was placed on waste well B while the sample well A was grounded. During separation a positive voltage was applied to well D while C was grounded and positive voltages of 15 and 10% of the voltage applied to D was applied to the sample well A and waste well B to prevent diffusion of remaining sample in these wells into the separation channel. A detector designed in the research lab of Dr. Steven Soper accomplished the LIF detection.

Figure 4.10. Modified twin T design used in the current described analysis. A is the sample well, B the waste well, C the buffer reservoir and D the separation well. During injection a positive voltage was placed on waste well B while the sample well A was grounded. During separation a positive voltage was applied to well D while C was grounded and positive voltages of 15 and 10% of the voltage applied to D was applied to the sample well A and waste well B.
POP5 and eCap matrices were not used in the CE analysis of LDR reactions due to the severe injection biases. These injection problems are alleviated in the microdevice format due to the ability to switch voltages during injection and separation and the volume based injection as compared with the electrokinetic injection of CE. POP5 was chosen as a matrix for analyzing model samples, as described above, in the microdevice electrophoresis format. The model samples were used to optimize the electrophoretic injection as well as separation conditions for analysis. A separation of the labeled 25-nucleotide and 51-nucleotide model sample was performed with no voltage pullbacks on wells A and B during separation (Figure 4.11). Although the entire microchannel was 10 cm, detection was performed at 3.5 cm before the first turn in the modified twin T. The lack of pullback voltages leads to severe band broadening in the form of tailing due to the leakage of residual sample from injection into the separation channel during separation. When pullback voltages are applied and detection is performed at 9.5 cm tailing is eliminated (Figure 4.12). A 1:100 model sample of 44- and 51-nucleotide LDR product mimics to 25-nucleotide discriminating primer mimic was prepared and analyzed by microdevice electrophoresis with the optimized conditions described above. Very good resolution was obtained with this sample as can be seen in Figure 4.13. The major problem with microdevice electrophoresis still to be studied is the lack of complete reproducibility due to the manual loading of the microchannel. The lack of suitable technology to fill the microchannel with separation matrix is the foremost concern of applying microdevice electrophoresis to mutation detection strategies. Further model studies and analysis of LDR products need to be performed to prove the efficiency of microdevice electrophoresis.
Figure 4.11. Microdevice analysis of a 1:1 model sample of the 25-nucleotide and 51-nucleotide oligonucleotides. POP5 was the matrix used and no pullback voltages were applied during separation. Injection was at 500 V for 5 s and separation was at 3.5 kV.

Figure 4.12. Microdevice analysis of a 1:1 model sample of the 25-nucleotide and 51-nucleotide oligonucleotides. POP5 was the matrix used and pullback voltages were applied during separation. Injection was at 500 V for 5 s and separation was at 3.5 kV. Pullback voltages on the sample and waste wells during separation were 500 and 300 V, respectively.
Figure 4.13. Microdevice analysis in POP5 of the 1:100 model sample.\textsuperscript{4.29} Samples were separated as described in Figure 4.12.

4.4 Conclusions

Although not completely studied, microdevice electrophoresis of LDR reactions is more efficient and superior in resolution to CE analysis, while both CE and microdevice electrophoresis are superior to slab gel electrophoresis in time and resolution. The use of this technology for analyzing mutation detection strategies increases resolution and decreases time which is essential for the ultimate goal of microdevice electrophoresis: doctor’s office evaluation of disease. These small devices, when completely optimized, will be a breakthrough for diagnostic care.

4.5 References


Chapter 5: Experimental Details

5.1 Nucleoside Base Analogs-Chapter 2

5.1.1 Synthesis of 1-((β-D-2´-deoxyribofuranosyl)-4-(2-thiazolyl)-pyrazole

In an inert atmosphere Pd(PPh3)4 (0.05 g, 0.05 mmol) was dissolved in dry THF (10 mL). To this was added thiazolyl zinc bromide (2.7 mL of a 0.5 M solution, 1.4 mmol) after which 1-(β-D-2´-deoxyribofuranosyl)-3´,5´-bis-O-(p-toluoyl)-4-iodo-pyrazole (0.5 g, 0.9 mmol) was added to the Pd-thiazole mixture. The reaction was stirred for 1 h after which another 0.05 equiv of Pd(PPh3)4 was added to the reaction. This was repeated three more times to reach a final 25 mol% of palladium catalyst. THF was evaporated under reduced pressure and the reaction was washed with NH4Cl (20 mL), ether (3 x 20 mL), and dried over CaCl2. The product was purified by column chromatography (2:1 hexanes:ethyl acetate) to yield a yellow oil in 90% (0.4 g). 1-(β-D-2´-deoxyribofuranosyl)-3´,5´-bis-O-(p-toluoyl)-4-(2-thiazolyl)-pyrazole (0.3 g, 0.6 mmol) was dissolved in methanol (20 mL) to which was added sodium methoxide (0.2 M, 20 mL). The reaction was stirred for 2 h after which acetic acid (5 mL) was added to quench. The solvent was evaporated under reduced pressure and the white solid dissolved in chloroform in which sodium acetate precipitated and was filtered. The product was purified by radial chromatography (9:1 CHCl3:CH3OH) to yield a white foam in 63%.

1-(β-D-2´-Deoxyribofuranosyl)-4-(2-thiazolyl)-pyrazole was converted to the 5´-dimethoxytrityl protected 3´-O-2-cyanoethyl-N,N-diisopropylphosphoramidite as previously described for incorporation into an oligodeoxynucleotide.5,1 The above-mentioned compounds were characterized by 1H and 31P NMR as previously described.5,1
5.1.2 Computational Analysis of Pyrazole-Based Analogs

Bases with a hydrogen atom at the sugar attachment point were drawn in Hyperchem 6.0 (Hypercube, Inc.) and minimized with a Polak-Ribiere conjugate gradient minimization technique (leaving all values as default) using the MM+ force field. RMS fits with crystal structure data were found to be 0.06 or less for each analog. Semiempirical PM3 calculations produced the QSAR properties of surface area, dipole moment and polarizability values for each analog.

5.1.3 DNA Polymerase Extension Reactions of Pyrazole Analogs

Templates and primers were synthesized by cyanoethyl phosphoramidite chemistry at a 1.0 µmol scale on an Applied Biosystems 394 DNA/RNA Synthesizer using a standard 500 Å CPG column. Coupling times were doubled for the incorporation of thiazolylpyrazole and nitropyrazole analog phosphoramidites.

All oligonucleotides were purified by electrophoresis on a 20% polyacrylamide/7 M urea gel. The bands were visualized by UV shadowing, excised from the gel, and eluted overnight at 37 ºC in a 1X elution buffer (0.1% SDS, 0.5 M ammonium acetate, and 10 mM magnesium acetate OR 0.5 M NaCl, 5 mM EDTA, pH 8.0). The eluant was purified with a C18 Sep Pak (Waters). Finally, the oligonucleotides were quantified by UV absorbance at 260 nm.

Primers were 5’-end 33P-labeled using T4 Polynucleotide Kinase (New England Biolabs) (10 U/50 pmol primer) and (γ-33P)ATP (Amersham) (150 µCi/50 pmol primer) and purity analyzed by electrophoresis on a 12% polyacrylamide/ 7 M urea analytical gel.

Extension reactions were performed in 20 µL with 250 nM radiolabeled primer and template, and 200 µM solution of all four dNTPs, polymerase and the supplied buffer. The following polymerases and buffers were used:
(a) 5 U AmpliTaq® DNA polymerase with supplied 10X Reaction Buffer (100 mM Tris-HCl pH 8.3, 500 mM KCl) (Perkin-Elmer). An additional 8 mM MgCl₂ was added to the reaction.

(b) 2.5 U Klenow fragment (exo-) DNA polymerase with supplied 10X Reaction Buffer (100 mM Tris-HCl pH 7.5, 5 mM MgCl₂, 7.5 mM dithiothrietol) (New England Biolabs)

Extension reactions were stopped at the appropriate time by the addition of 20 µL gel loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol FF) and the products electrophoresed on a 12% polyacrylamide/7 M urea analytical polyacrylamide gel. After drying for 2 h at 80 ºC the gel was imaged on a Molecular Dynamics phosphorscreen imaging plate then read on a Storm 840 scanner and analyzed using Molecular Dynamics Image-Quant software. The electrophoresis lanes were subjected to area analysis which created an electropherogram based on the intensities of the imaged bands. The area of the product peak was divided by the sum of the areas of all of the peaks in the lane. This value was multiplied by 100 to obtain the percentage of extension.

5.1.4 Synthesis of Thiazole and Thiazole N-oxide Analogs

5.1.4.1 Synthesis of 2-deoxy-3,5-di-O-p-toluoyl-β-D-erythro-pentofuranosyl cyanide (2)

Chlorosugar 1 (4.23 g, 10.9 mmol) was dissolved in dry DME (100 mL), sodium cyanide (2.14 g, 43.60 mmol) was added, and the mixture was stirred for 24 h at room temperature under argon. The solvent was evaporated and the resulting yellow oil was dissolved in 200 mL EtOAc then extracted with sat. aq. NaHCO₃ (1 x 100 mL) and sat. aq. NaCl (1 x 100 mL). The organic layer was dried over Na₂SO₄, filtered and evaporated to a white solid. Purification by silica gel flash chromatography using hexanes/EtOAc as eluant provided cyanosugar 2 as a 3.2:1 ratio of β/α anomers in 83% yield (3.18 g). For ease of purification, this anomeric mixture was used
directly in the following reaction. \( R = 0.63 \) (50% EtOAc/Hexanes); \(^1\)H NMR (250 MHz, CDCl\(_3\), \( \beta \) anomer) \( \delta \) 7.92 (m, 4H), (m, 4H), (m, 1H), (dd, 0.75H, \( J = 2.6, 6.6 \) Hz, \( \beta \)-anomer), 4.56 (m, 3H), 2.70 (m, 2H), 2.41 (s, 6H); \(^{13}\)C NMR (62.5 MHz, CDCl\(_3\)) 166.1, 166.0, 144.6, 144.1, 129.9, 129.7, 129.6, 129.2, 126.6, 126.2, 117.9, 83.8, 75.4, 74.8, 66.5, 65.8, 63.7, 63.6, 37.8, 21.7.

5.1.4.2 Synthesis of 2,5-anhydro-3-deoxy-4,6-di-\( \text{O-} \)-p-toluoyl-D-ribo-hexonthioamide (3)

Hydrogen sulfide was condensed into a pressure vial cooled to -77 °C containing a 3.2:1 \( \beta/\alpha \) anomeric mixture of cyanosugar 2 (3.10 g, 8.17 mmol) and dimethylaminopyridine (50 mg, 0.041 mmol). The vessel was sealed and allowed to warm to room temperature. After 24 h, the pressure vial was cooled to -77 °C and opened, then allowed to slowly warm to room temperature. The resulting yellow foam was coevaporated with 2 x 50 mL of dichloromethane, then purified by silica gel chromatography using hexanes/ethyl acetate as eluant to provide thioamide 3 in 61% yield (2.06 g). \( R = 0.57 \) (50%EtOAc:Hexanes); \(^1\)H NMR (250 MHz, CDCl\(_3\)) \( \delta \) 8.37 (s, 1H), 7.92 (m, 4H), 7.63 (s, 1H), 7.27 (m, 4H), 5.48 (d, 1H, \( J = 5.6 \) Hz), 5.11 (dd, 1H, \( J = 5.9, 4.7 \)), 4.63 (m, 3H), 2.86 (dd, 1H, \( J = 6, 8.2 \) Hz), 2.26 (m, 1H).; \(^{13}\)C NMR (62.5 MHz, CDCl\(_3\)) 205.3, 166.7, 165.9, 144.2, 129.6, 129.5, 129.2, 129.1, 126.3, 84.9, 84.8, 75.7, 64.3, 40.0, 21.5; FAB(+)HRMS calcd for \( \text{C}_{22}\text{H}_{24}\text{NO}_{5}\text{S} \) (M+H)+ 414.1392, found 414.1365.

5.1.4.3 Synthesis of 2-(2´-Deoxy-3´,5´-di-\( \text{O-} \)-p-toluoyl-\( \beta \)-D-ribofuranosyl)thiazole (4)

A. Thioamide 3 (1.16g, 3.0 mmol) was dissolved in THF (33 mL) and EtOH (16 mL). 50% chloroacetaldehyde solution (3.18 mL, 25 mmol) was added and the reaction refluxed for 24 h. Another equiv of chloroacetaldehyde solution (3.18 mL, 25 mmol) was added and the reaction refluxed for another 24 h period. The solvents were evaporated and the residue dissolved in ethyl acetate, extracted with saturated aq. NaHCO\(_3\) (1 x 100 mL) and saturated aq.
NaCl (1 x 100 mL), and dried over MgSO₄. Silica gel column chromatography with ethyl acetate/hexanes yielded 42% as a light yellow solid.

B. Thioamide 3 (0.50 g, 1.21 mmol) was dissolved in 7.1 mL of dry acetone, and 0.9 mL (0.69 g, 3.5 mmol) of bromoacetaldehyde diethyl acetal, followed by 13 µl of a 4 M HCl/dioxane solution, was added. The resulting clear solution was refluxed under argon for 24 h. The brown colored solution was evaporated to a dark oil, then dissolved in 100 mL EA and extracted with saturated aq. NaHCO₃ (1 x 100 mL) and saturated aq. NaCl (1 x 100 mL). The organic layer was dried over Na₂SO₄, filtered and evaporated to a dark brown liquid. Purification by silica gel chromatography with hexanes/EA provided tol-thiazole 4, a light brown gum, in 70% yield (0.370 g). Crystallization from ethanol provided the title compound as a solid. ¹H NMR (250 MHz, CDCl₃) δ 7.93 (m, 4H), 7.74 (d, 1H, J=3.4 Hz), 7.24 (m, 5H), 5.62 (m, 2H), 4.61 (m, 3H), 2.78 (m, 1H), 2.47 (m, 1H), 2.43 (s, 3H), 2.39 (s, 3H); ¹³C (62.5 MHz, CDCl₃) 171.4, 166.2, 166.0, 147.0, 143.8, 142.6, 129.7, 129.2, 129.1, 126.9, 126.7, 119.1, 83.5, 78.8, 76.7, 64.4, 40.2, 21.6; FAB(+)HRMS calcd for C₂₄H₂₄NO₅S (M+H)+ 438.1392, found 438.1382.

5.1.4.4 Synthesis of 2-(2´-Deoxy-3´,5´-di-O-p-toluoyl-β-D-ribofuranosyl)thiazole-N-oxide (5)

A. Tol-thiazole 4 (380 mg, 0.87 mmol) was dissolved in ethyl acetate (5 mL) and m-chloroperoxybenzoic acid (320 mg, 1.85 mmol) was added. The stirred reaction was allowed to proceed for 72 h, then purified by silica gel column chromatography using ethyl acetate/methanol as eluant to provide tol-thiazole N-oxide 5, a colorless oil, in 46% yield (180 mg). Repeated evaporation from ethyl acetate provided the title compound as a crystalline solid.

B. Tol-thiazole 4 (100 mg, 0.23 mmol) was dissolved in dichloromethane (5 mL) after which hydrogen peroxide urea complex (43.27 mg, 0.46 mmol) was added and allowed to stir for 5 min at room temperature. Trifluoroacetic anhydride (96.61 mg, 0.46 mmol) was then slowly
added to the reaction and stirred for 1 h. Na$_2$S$_2$O$_3$ (5 mL) was then added to the reaction mixture and allowed to stir for 15 min destroying any residual peroxides. The organic layer was extracted with NaCl (1 X 50 mL) and dried using Na$_2$SO$_4$. The mixture was purified by silica gel column chromatography using ethyl acetate/hexanes, which resulted in tol-thiazole $N$-oxide 5, a colorless oil that solidifies at room temperature, resulting in a 97 % yield (101 mg).

$R_f$ = 0.37 (12.5% MeOH:EtOAc); $^1$H NMR (250 MHz, CDCl$_3$) $\delta$ 7.94 (m, 4H), 7.70 (d, 1H, $J$=4.0 Hz), 7.27 (m, 5H), 5.77 (dd, 1H, $J$=5.6, 4.7 Hz), 5.68 (d, 1H, $J$=5.9 Hz), 4.63 (m, 3H) 3.14 (dd, 1H, $J$=8.2, 5.7), 2.45 (s, 3H), 2.42 (s, 3H), 2.31 (m, 1H); $^{13}$C NMR (62.5 MHz, CDCl$_3$) 166.1, 149.3, 144.3, 144.0, 137.6, 129.8, 129.6, 129.2, 126.7, 126.5, 117.1, 83.6, 76.6, 75.3, 64.2, 37.8, 21.7; FAB(+)HRMS calcd for C$_{24}$H$_{24}$NO$_6$S (M+H)$^+$ 454.1341, found 454.1318.

5.1.4.5 Synthesis of 2-(2´-Deoxy-$\beta$-D-ribofuranosyl)thiazole (6)

Tol-thiazole 4 (220 mg, 0.51 mmol) was dissolved in 16.2 mL of dry methanol, and 0.20 M sodium methoxide in methanol (5.4 mL) was added. After 2 h, glacial acetic acid (129 mg, 2.14 mmol) was added and the solution was evaporated to an oil. Flash chromatography on silica gel with ethyl acetate/methanol as eluant provided 2-(2´-deoxy-$\beta$-D-ribofuranosyl)thiazole as a white solid in 74% yield (75 mg). Thiazole 6 was crystallized from benzene, mp 129-131 °C. $R_f$ = 0.46 (12.5% MeOH:EtOAc); $^1$H NMR (250 MHz, CD$_3$OD) $\delta$ 7.73 (d, 1H, $J$=3.3 Hz), 7.52 (d, 1H, 3.3 Hz), 5.42 (dd, 1H, $J$=6.2, 3.2 Hz), 4.35 (m, 1H), 4.00 (m, 1H), 3.63 (m, 2H), 2.39 (m, 1H), 2.13 (m, 1H); $^{13}$C NMR (62.5 MHz, CDCl$_3$) 175.3, 143.1, 120.7, 89.6, 78.9, 74.0, 63.9, 43.6; FAB(+)HRMS calcd for C$_{8}$H$_{12}$NO$_3$S (M+H)$^+$ 202.0555, found 202.0542.

5.1.4.6 Synthesis of 2-(2´-Deoxy-5´-O-dimethoxytrityl-$\beta$-D-ribofuranosyl)thiazole (8)

Thiazole 6 (100 mg, 0.50 mmol) was dissolved in 4 mL dry pyridine. To this solution was added 70 µl of triethylamine (51 mg, 0.50 mmol) and dimethoxytrityl chloride (170 mg,
0.50 mmol). Following 3 h of stirring at room temperature, 0.20 mL of methanol was added and the solvents were evaporated. Flash chromatography on silica gel with hexanes/ethyl acetate provided the title compound in 70% yield (170 mg) as a white foam from dry dichloromethane. $R_f=0.34$ (75% EtOAc:Hexanes); $^1H$ NMR (250 MHz, CDCl$_3$) $\delta$ 7.70 (d, 1H, $J$=3.3 Hz), 7.46-7.23 (m, 10H), 6.80 (d, 4H, $J$=8.8 Hz), 5.50 (dd, 1H, $J$=6.2, 2.8 Hz), 4.42 (m, 1H), 4.12 (m, 1H), 3.76 (s, 6H), 3.30 (m, 2H), 3.04 (bs, 1H), 2.44 (m, 1H), 2.26 (m, 1H); $^{13}C$ NMR (62.5 MHz, CDCl$_3$) 173.2, 158.4, 144.7, 142.3, 135.9, 135.8, 130.0, 128.0, 127.8, 126.7, 118.9, 113.0, 86.6, 86.1, 77.9, 73.7, 64.0, 55.1, 42.7; FAB(+)HRMS calcd for C$_{29}$H$_{29}$NO$_5$S (M+Na)$^+$ 526.1681, found 526.1656.

5.1.4.7 Synthesis of 2-(2´-deoxy-5´-O-dimethoxytrityl-β-D-ribofuranosyl)thiazole-3´-O-(2-cyanoethyl)-N,N-diisopropylphosphoramidite (9)

DMT-thiazole 8 (80 mg, 0.16 mmol) was coevaporated with dry dichloromethane (3 x 5 mL), and then dissolved in dichloromethane anhydrous (2 mL). To this solution was added 2-cyanoethyl tetraisopropylphosphorodiamidite (152 µl, 0.48 mol), diisopropylamine (22 mL, 0.16 mmol), and tetrazole (11.2 mg, 0.16 mmol). The reaction was stirred for 2 h, methanol was added (100 µl), followed by dichloromethane (50 mL), the reaction was extracted with 10% Na$_2$CO$_3$ (2 x 50 mL) and dried over Na$_2$SO$_4$. Flash chromatography on silica gel using ethyl acetate/hexanes as eluant provided thiazole phosphoramidite 9 as a white foam from dry diethyl ether in 89% yield (10 mg). Mixture of two diastereomers, $R_f=0.43$, 0.53 (50% EtOAc:Hexanes); $^1H$ NMR (250 MHz, CDCl$_3$) $\delta$ 7.66 (d, 1H, $J$=4.1), 7.33 (m, 11H), 6.85 (m, 4H), 5.53 (dd, 1H, $J$=5.4, 4.2), 4.54 (m, 1H), 4.29 (m, 1H), 3.81 (s, 6H), 3.63 (m, 3H), 3.38 (m, 1H), 3.24 (m, 1H), 2.63 (t, 2H, $J$=6.3), 2.44 (t, 1H, $J$=6.6), 2.30 (m, 1H), 1.16 (m, 12H); $^{31}P$ NMR (101.2 MHz, CDCl$_3$) (mixture of two diasteromers) $\delta$ 150.07, 149.44; FAB(+)HRMS calcd for C$_{8}$H$_{12}$NO$_4$S (M)$^+$ 703.2862, found 703.2827.
5.1.4.8 Synthesis of 2-(2´-Deoxy-β-d-ribofuranosyl)thiazole-N-oxide (7)

To tol-thiazole N-oxide 5 (150 mg, 0.33 mmol) dissolved in methanol (10.4 mL) was added a 0.20 M solution of sodium methoxide in methanol (3.45 mL, 0.69 mmol). After 2 h, glacial acetic acid was added (42 mg, 0.69 mmol) and the solvents were evaporated to provide a thick oil. Flash chromatography on silica gel using EtOAc/methanol as eluant provided thiazole N-oxide 7 as a viscous oil in 83% yield (60 mg). \( R_f = 0.33 \) (50% MeOH:EtOAc); \(^1\)H NMR (250 MHz, CD\(_3\)OD) \( \delta \) 7.92 (m, 4H), 7.68 (d, 1H, \( J = 4.0 \)), 7.25 (m, 5H), 5.76 (dd, 1H, \( J = 6.6, 4.7 \)), 5.66 (d, 1H, \( J = 5.9 \)), 4.62 (m, 3H), 3.11 (dd, 1H, \( J = 5.6, 8.3 \)), 2.43 (s, 3H), 2.40 (s, 3H), 2.32 (m, 1H); \(^13\)C NMR (62.5 MHz, CD\(_3\)OD) 155.7, 137.9, 120.36, 89.4, 75.5, 73.9, 63.6, 41.0; FAB(+)HRMS calcld for C\(_8\)H\(_{12}\)NO\(_4\)S (M+ H\(^+\)) 218.0504, found 218.0481.

5.1.4.9 Synthesis of 2-(2´-Deoxy-5´-O-dimethoxytrityl-β-d-ribofuranosyl)thiazole-N-oxide (10)

Thiazole N-oxide 7 (50 mg, 0.23 mmol) was coevaporated with dry pyridine (2 x 5 mL), then dissolved in anhydrous pyridine (2 mL). To this solution was added 33 \( \mu \)l of triethylamine (24 mg, 0.23 mmol) and dimethoxytrityl chloride (78 mg, 0.23 mmol). The mixture was stirred for 2 h, then a second equivalent of triethylamine (24 mg, 33 \( \mu \)l) and dimethoxytrityl chloride (78 mg, 0.23 mmol) was added. After an additional 1 h, methanol (0.2 mL) was added, the reaction was evaporated to dryness, dissolved in ethyl acetate (50 mL) and extracted with saturated aq. NaHCO\(_3\) (1 x 50 mL), saturated aq. NaCl (1 x 20 mL) and the organic layer was dried over Na\(_2\)SO\(_4\). The solution was filtered, evaporated to dryness, and purified by flash chromatography on silica gel using methanol/dichloromethane as eluant to provide thiazole phsphoramidite 9 as a white foam from diethyl ether in 90% yield (110 mg). \( R_f = 0.25 \) (12.5% MeOH:EtOAc); \(^1\)H NMR (250 MHz, CDCl\(_3\)) \( \delta \) 7.61 (d, 1H, \( J = 4.1 \)), 7.30 (m, 11H), 6.82 (m, 4H), 5.73 (dd, 1H, \( J = 5.7, 4.1 \)), 4.49 (bs, 1H), 4.15 (m, 2H), 3.79 (s, 6H), 3.20 (m, 2H), 2.89 (m,
$^1$H, 2.14 (m, 1H); $^{13}$C NMR (62.5 MHz, CDCl$_3$) 158.9, 145.2, 137.7, 136.4, 136.3, 130.5, 128.6, 128.2, 127.2, 117.3, 113.5, 87.3, 86.6, 75.2, 74.2, 64.3, 60.8, 55.6, 41.0; FAB(+)HRMS calcd for C$_8$H$_{12}$NO$_4$S (M)$^+$ 519.1732, found 519.1709.

5.1.4.10 Synthesis of 2-(2´-deoxy-5´-O-dimethoxytrityl-β-D-ribofuranosyl)thiazole-N-oxide-3´-O-(2-cyanoethyl)-N,N-diisopropylphosphoramidite (11)

Thiazole phosphoramidite 9 (110 mg, 0.21 mmol) was dried by evaporation from anhydrous dichloromethane (2 x 5 mL), and dissolved in 2 mL dry dichloromethane. To this solution was added diisopropylamine (32 µl, 0.23 mmol), 2-cyanoethyl tetraisopropylphosphorodiamidite (89 µl, 0.28 mmol) and tetrazole (16 mg, 0.23 mmol). The solution was stirred for 30 minutes and a second portion of 2-cyanoethyl tetraisopropylphosphorodiamidite (57 µl, 0.18 mmol) was added. The reaction was allowed to proceed an additional 20 min., then methanol (0.5 mL), followed by 50 mL of dichloromethane, was added. The organic layer was extracted with 10% aq. Na$_2$CO$_3$ (1 x 50 mL), sat. aq. NaCl (1 x 20 mL), dried over Na$_2$SO$_4$, filtered and the solvents removed by evaporation. Flash chromatography on silica gel with methanol/ethyl acetate containing 0.1% triethylamine provided thiazole N-oxide phosphoramidite 11 as a white foam from diethyl ether in 67% yield (100 mg). Mixture of two diastereomers, $R_f$=0.49, 0.53 (10% MeOH:EtOAc); $^1$H NMR (250 MHz, CDCl$_3$) δ 7.66 (d, 1H, $J$=4.1), 7.32 (m, 12H), 6.86 (m, 4H), 5.65 (dd, 1H, $J$=6.0, 3.5), 4.54 (m, 1H), 4.24 (m, 1H), 3.81 (s, 6H), 3.61 (m, 3H), 3.38 (s, 1H), 3.17 (m, 1H), 2.93 (m, 1H) 2.67 (t, 1H, $J$=6.3), 2.44 (t, 1H, $J$=6.3), 2.19 (m, 1H), 1.10 (m, 12H); $^{31}$P NMR (101.2 MHz, CDCl$_3$) δ 149.2, 148.6; FAB(+)HRMS calcd for C$_8$H$_{12}$NO$_4$S (M)$^+$ 720.2889, found 720.2875

5.1.4.11 Deprotection of DNA

On a small scale approximately 10% of the column material was taken from the cartridge and treated with 50 mM K$_2$CO$_3$ (100 µL) for 3 h. 1.5 eq. of acetic acid in water was added to
quench. The solvent was removed, avoiding the solid support, and evaporated to dryness. HPLC purification was performed using either a Hydrocell NS 1000 (anion-exchange) column (150 x 4.6 mm column) with Solvent A = 25 mM CHES pH 8, 30% methanol and Solvent B = 25 mM CHES, 1 M (NH₄)₂SO₄, pH 8, 30% methanol (0-100% B in 60 min with a flow rate = 1 mL/min) or a DNAPac-100 (anion-exchange) column with Solvent A = 1 M NaCl, 25 mM NaOH (pH 12.4) and Solvent B = 25 mM NaOH (pH 12.4) (100-500 mM NaCl in 2 min, 500-900 mM NaCl in 18 min, and 900-1000 mM NaCl in 3 min with a flow rate = 6 mL/min).

5.1.5 Computational Analysis of Thiazole and Thiazole N-oxide Analogs

5.1.5.1 Semi-empirical Analysis

Bases with a hydrogen atom at the sugar attachment point were drawn in Hyperchem 6.0 (Hypercube, Inc.) and minimized with a Polak-Ribiere conjugate gradient minimization technique (leaving all values as default) using the AMBER force field. Semiempirical PM3 calculations produced the QSAR properties of surface area, dipole moment and polarizability values for each analog.

In order to produce the electrostatic potential maps, in SYBYL MOPAC calculations at the PM3 level were performed and the electrostatic potential was mapped onto a Connolly surface using these charges.

5.1.5.2 Molecular Dynamics Analysis

The thiazole N-oxide nucleoside analog was built by modifying the cytidine nucleoside from the AMBER95dna dictionary. The nitrogen of the N-oxide was defined as NA which is an sp² nitrogen in aromatic rings with a hydrogen attached. The oxygen of the N-oxide was defined as O2 which is an sp² oxygen in anionic acids. The sulfur was defined as S which is sulfur in methionine. With these definitions a 1000 iteration minimization with Kollman charges using
the Amber 4.1 force field produced a structure which has an RMS value with the crystal structure of 0.07. A further systematic search was performed to insure that the lowest energy conformation was being used. This search produced a structure which had a slightly lower energy (about 2 kcal/mol) and an RMS fit with the crystal structure of 0.05. This structure was merged into a B-type Dickerson dodecamer built using the AMBER95dna dictionary. The modified DNA double helix structure was minimized using the Amber 4.1 force field and Kollman charges. This minimized structure was used directly in the Molecular Dynamics simulations.

Molecular Dynamics simulations were performed by ramping to 300 K or 400 K for 1 ps and then stepping down to zero. These simulations used a 1 fs time step and constant pressure. SHAKE was used to constrain all hydrogens except those of the thiazole N-oxide-guanine base pair.

5.1.6 Thermodynamic Analysis of Thiazole and Thiazole N-oxide Analogs

Oligonucleotides were synthesized by cyanoethyl phosphoramidite chemistry at a 1.0 µmol scale using phenoxyacetyl (Pac) phosphoramidites (Glen Research) on an Applied Biosystems 394 DNA/RNA Synthesizer using a standard 500 Å CPG column. Coupling times were doubled for the incorporation of thiazole and thiazole N-oxide analog phosphoramidites. Oligonucleotides were purified by anion exchange HPLC using a Hydrocell NS 1000 (anion exchange) 150 x 4.6 mm column. Solvent A was 25 mM CHES pH 8, 30% MeOH while Solvent B was 25 mM CHES/ 1 M (NH₄)₂SO₄ pH 8, 30% MeOH. The program consisted of ramping 0-100% B over 60 min and the flow rate = 1 mL/min.

Samples were prepared at multiple concentrations in degassed 1X buffer (1 M NaCl, 1 mM EDTA, 10 mM Na₂PO₄). The path lengths of the cells used for analysis ranged from 0.01
cm to 1 cm. The samples were heated to 95 °C for 2 min to drive the strands to full separation. They were then cooled to 5 °C for 5 min to ensure full duplex formation. The temperature was then ramped from 5 °C to 95 °C in 0.5 °C increments. Absorbance readings were recorded at each increment at 260 nm by a Gilford Response II UV/Vis Spectrophotometer equipped with a Peltier temperature controlling device and thermal software. Following the melt, absorbance readings were taken at 5 °C to ensure that evaporation did not take place during the melt.

Extinction coefficients for the modified nucleosides were measured at 260 nm (4.97x10^-4 M for thiazole and 9.2x10^-8 M for thiazole N-oxide). These values were then used to determine the extinction coefficient of the DNA strand by the nearest neighbor estimate using the straight sum method.\textsuperscript{5,3,5,4}

\( T_m, \Delta H, \Delta S, \) and \( \Delta G \) were obtained for each sample using \textit{Meltwin}\textsuperscript{©} provided by Jeffrey McDowell which fits a curve to the raw data points.\textsuperscript{5,5} Van’t Hoff analysis of three or more concentrations was performed to obtain the thermodynamic values presented.

\textbf{5.1.7 DNA Polymerase Recognition of Thiazole and Thiazole N-oxide Analogs}

Templates and primers were synthesized by cyanoethyl phosphoramidite chemistry at a 1.0 µmol scale using phenoxyacetyl (Pac) phosphoramidites (Glen Research) on an Applied Biosystems 394 DNA/RNA Synthesizer using a standard 500 Å CPG column. Coupling times were doubled for the incorporation of thiazole and thiazole N-oxide analog phosphoramidites.

Oligonucleotides containing thiazole were purified by electrophoresis on a 20% polyacrylamide/7 M urea gel while those containing thiazole N-oxide were purified by HPLC. The bands of the electrophoresis gel were visualized by UV shadowing, excised from the gel, and eluted overnight at 37 °C in a 1X elution buffer (0.1% SDS, 0.5 M ammonium acetate, and 10 mM magnesium acetate OR 0.5 M NaCl, 5 mM EDTA, pH 8.0). The eluant was purified.
with a C18 Sep Pak (Waters). The oligonucleotides containing thiazole N-oxide were purified on a quaternary ammonium functionalized polymeric anion exchange column purchased from Dionex Corp, the DNAPac PA-100 column. using Buffer A (1 M NaCl, 25 mM NaOH pH 12.4) and Buffer B (25 mM NaOH pH 12.4). The program consisted of ramping from 100-500 mM NaCl in 2 min, 500-900 mM NaCl in 18 min, and 900-1000 mM NaCl in 3 min and the flow rate = 6 mL/min. The fractions were collected and desalted using a C18 Sep Pak (Waters) and analyzed by MALDI. Finally, the oligonucleotides were quantified by UV absorbance at 260 nm.

Primers were 5’-end 33P-labeled using T4 Polynucleotide Kinase (New England Biolabs) (20 U/2 nmol primer) and (γ-33P)ATP (Amersham) (250 µCi/2 nmol primer) and purity analyzed by electrophoresis on a 12% polyacrylamide/7 M urea analytical gel.

Extension reactions were performed in 20 µL with 200 nM radiolabeled primer and template, and 200 µM solution of all four dNTPs, polymerase and the supplied buffer. The following polymerases and buffers were used:

(a) 5 U Taq DNA polymerase with supplied 10X Reaction Buffer (100 mM KCl, 100 mM (NH4)2SO4, 200 mM Tris-HCl, 20 mM MgSO4, 1% Triton X-100 pH 8.8) (New England Biolabs). An additional 8 mM MgCl2 was added to the reaction.

(b) 5 U Klenow fragment (exo-) DNA polymerase with supplied 10X Reaction Buffer (100 mM Tris-HCl pH 7.5, 5 mM MgCl2, 7.5 mM dithiothrietol) (New England Biolabs)

Extension reactions were stopped at the appropriate time by the addition of 20 µL gel loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol FF) and the products elecrophoresed on a 12% polyacrylamide/7 M urea analytical polyacrylamide gel. After drying for 2 h at 80 ºC the gel was imaged on a Molecular Dynamics phosphorscreen.
imaging plate then read on a Storm 840 scanner and analyzed using Molecular Dynamics Image-Quant software. The electrophoresis lanes were subjected to area analysis which created an electropherogram based on the intensities of the imaged bands. The area of the product peak was divided by the sum of the areas of the all of the peaks in the lane. This value was multiplied by 100 to obtain the percentage of extension.

To analyze the kinetic incorporation of dNTPs opposite thiazole and thiazole N-oxide the concentration of K_f(exo-), concentrations of dNTP, and time of incorporation were optimized to keep the amount of extension at or below 20%. The concentration of DNA in the assay was kept constant at 2.5 µM. Table 5.1 shows these optimized values.

**Table 5.1. Optimized parameters for kinetic incorporation analysis.**

<table>
<thead>
<tr>
<th></th>
<th>K_f(exo-) (U/µl)</th>
<th>Time</th>
<th>dNTP (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T-template</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>0.005</td>
<td>30 s</td>
<td>2.5, 5, 7.5, 10, 20 µM</td>
</tr>
<tr>
<td>dGTP</td>
<td>0.005</td>
<td>5 min</td>
<td>50, 75, 100, 150, 200 µM</td>
</tr>
<tr>
<td><strong>C-template</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>0.02</td>
<td>5 min</td>
<td>25, 50, 75, 100, 150 µM</td>
</tr>
<tr>
<td>dGTP</td>
<td>0.005</td>
<td>1 min</td>
<td>2.5, 5, 7.5, 10, 15 µM</td>
</tr>
<tr>
<td><strong>Z-template</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>0.01</td>
<td>30 s</td>
<td>5, 10, 15, 20, 25 µM</td>
</tr>
<tr>
<td>dGTP</td>
<td>0.02</td>
<td>3 min</td>
<td>25, 50, 75, 100, 150 µM</td>
</tr>
</tbody>
</table>
The radiolabeled primer and template were incubated in 4X reaction buffer at 95 ºC for 1 min. The appropriate amount of this annealing mix was then added to Solution A containing 2X reaction buffer, 2X BSA, and the appropriate amount of Kf(exo-). Solution B was prepared by adding stock dNTP solution to 2X reaction buffer to obtain the varying dNTP concentrations. Solution A was incubated at 37 ºC for 2 min. The incorporation was initiated by the addition of an equal volume of Solution B to Solution A and the reaction incubated at 37 ºC for the appropriate length of time. Reactions were stopped by the addition of 10 µL gel loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol FF) and the products elecrophoresed on a 12% polyacrylamide/7 M urea analytical polyacrylamide gel. After drying for 2 h at 80 ºC the gel was imaged on a Molecular Dynamics phosphorscreen imaging plate then read on a Storm 840 scanner and analyzed using Molecular Dynamics Image-Quant software. The electrophoresis lanes were subjected to area analysis which created an electropherogram based on the intensities of the imaged bands. The area of the product peak was divided by the sum of the areas of the product and primer peaks. This value was multiplied by 100 to obtain the percentage of extension. The velocity of incorporation was calculated by dividing the percentage extension by the time in minutes. Linear plots of dNTP concentration divided by velocity versus dNTP concentration (Hanes-Wolf plots) were obtained using Microsoft Excel 2000. $K_m$ and $V_{max}$ values were then calculated from the slope and y-intercept values.

5.2 Nucleoside Backbone Analogs-Chapter 3

5.2.1 Synthesis of N3’→P5´ Phosphoramidate Backbone Modification

The 5´-O-DMT-2´,3´-dideoxy-3´-aminothymidine was synthesized as previously reported and then was used in a modified synthetic cycle for the automated synthesis of the primer used in
the extension reactions at a 1.0 µmol scale on an Applied Biosystems 394 DNA/RNA Synthesizer using a standard 500 Å CPG column.  

5.2.2 Exonuclease Resistance Assays

Templates and primers containing only phosphodiester linkages were synthesized by cyanoethyl phosphoramidite chemistry at a 1.0 µmol scale on an Applied Biosystems 394 DNA/RNA Synthesizer using a standard 500 Å CPG column. For synthesis of the primer containing a 3´-phosphorothioate linkage standard cyanoethyl phosphoramidite chemistry was used with an additional sulferization step utilizing 3-ethoxy-1,2,4-dithiazoline-5-one (EDITH) after the first phosphoramidite coupling. This primer was also synthesized at a 1.0 µmol scale on an Applied Biosystems 394 DNA/RNA Synthesizer using a standard 500 Å CPG column. All oligonucleotides were purified by electrophoresis on a 19% polyacrylamide/7 M urea gel. The bands were visualized by UV shadowing, excised from the gel, and eluted overnight at 37 ºC in a 1X elution buffer (0.1% SDS, 0.5 M ammonium acetate, and 10 mM magnesium acetate). The eluant was purified with a C_{18} Sep Pak (Waters). Finally, the oligonucleotides were quantified by UV absorbance at 260 nm.

Primers were 5´-end $^{33}$P-labeled using T4 Polynucleotide Kinase (New England Biolabs) (5 U/4 pmol) and ($\gamma^{33}$P)ATP (Amersham) (12 μCi/4 pmol) and purity analyzed by electrophoresis on an 8% polyacrylamide/7 M urea analytical gel.

Extension reactions were performed with 20 pM primer, 50 pM template, and 200 µM solution of all four dNTPs, polymerase and the supplied buffer. The following polymerases and buffers were used:

(a) 2.5 U AmpliTaq® DNA polymerase with supplied 10X Reaction Buffer (100 mM Tris-HCl pH 8.3, 500 mM KCl) (Perkin-Elmer)
(b) 1 U Vent(exo+) DNA polymerase with supplied 10X Reaction Buffer (100 mM KCl, 100 mM (NH₄)₂SO₄, 20 mM Tris-HCl pH 8.8, 20 mM MgSO₄, 1% Triton X-100) (New England Biolabs)

(c) 1.75 U Pfu(exo+) DNA polymerase with supplied 10X Reaction Buffer (200 mM Tris-HCl (pH 8.8), 100 mM KCl, 100 mM (NH₄)₂SO₄, 20 mM MgSO₄, 1% Triton X-100, 1mg/mL nuclease free bovine serum albumin (BSA)) (Stratagene)

Extension reactions were then subjected to restriction endonuclease digests. Extension of the template with a 3´-match base pair was digested with 10 U BslI restriction endonuclease (New England Biolabs) (supplied 10X reaction buffer: 1 M NaCl, 500 mM Tris-HCl, 100 mM MgCl₂, 1 mM DTT, pH 7.9) at 55 ºC for 60 min. Extension of the template with a 3´-mismatch base pair was digested with 1 U TaqαI restriction endonuclease (New England Biolabs)(supplied 10X reaction buffer: 1 M NaCl, 100 mM Tris-HCl, 100 mM MgCl₂, pH 8.4)) at 65 ºC for 60 min. Reactions were stopped by the addition of 4 µl gel loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol FF) and the products electrophoresed on an 8% polyacrylamide/7 M urea analytical polyacrylamide gel. After drying for 2 h at 80 ºC the gel was imaged on a Molecular Dynamics phosphorscreen imaging plate then read on a Storm 840 scanner and analyzed using Molecular Dynamics Image-Quant software. The electrophoresis lanes were subjected to area analysis which created an electropherogram based on the intensities of the imaged bands. The area of the product peak was divided by the sum of the areas of the product and primer peaks. This value was multiplied by 100 to obtain the percentage of extension.
5.2.3 Kinetic Exonuclease Resistance Assays

Primers were 5’-end \(^{33}\)P-labeled as stated above. The reactions were mixed as an extension reaction. Only Vent and \(Pfu\) polymerases were used in the above buffers. Aliquots of the reactions were removed and stopped at \(t = 0, 10\) s, \(30\) s, \(1\) min, \(5\) min, \(10\) min, \(30\) min, and \(1\) h. The reactions were stopped with gel loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol FF) and cooled on ice (0 °C). The aliquots were analyzed by electrophoresis on an 8% polyacrylamide/7 M urea analytical polyacrylamide gel. After drying for 2 h at 80 ºC the gel was imaged on a Molecular Dynamics phosphorscreen imaging plate then read on a Storm 840 scanner and analyzed using Molecular Dynamics Image-Quant software.

5.3 Capillary and Microdevice Electrophoresis of LDR-Chapter 4

5.3.1 DNA sample preparation

All LDR primers and model sample oligonucleotides were purchased from Integrated DNA Technologies (IDT, Coralville, IA) with PAGE purification and 5’-fluorescein labeling of LDR discriminating primers and model sample oligonucleotides. Model samples were prepared in 1:10, 1:100, and 1:1000 molar ratios of 44nt and 51nt oligonucleotides to 25nt oligonucleotide. The 25nt was 1 \(\mu\)M and the 44nt and 51nt were 100 nM, 10 nM, and 1 nM for the 1:10, 1:100, and 1:1000 samples, respectively.

5.3.2 LDR Conditions

PCR products from the amplification of exon 1 of the K-\textit{ras} gene were obtained from the laboratory of Dr. Francis Barany and used as templates in LDR reactions as described elsewhere.\(^{5,12,5,13}\) Briefly, the reactions were performed in a final volume of 20 \(\mu\)L. 500 nM of each discriminating primer and the appropriate amount of mutant and wild type templates for
the ratio being tested were used and ligated with Tth ligase (1 nM) in a 1X reaction buffer (10 mM Tris-HCl pH 7.6, 10 mM MgCl₂, 0.1 M KCl, and 20 µg/mL bovine serum albumin) as well as 10 mM dithiothreitol and 1.25 mM NAD⁺ (nicotinic adenine dinucleotide, a cofactor of Tth ligase). The reaction mixtures without the ligase were heated to 94 °C for 1.5 min. The ligase was then added to the reaction and the reaction cycled 20 times at 94 °C for 15 s and 65 °C for 4 min. Reactions were stopped by the addition of 0.5 µL of 0.5 mM EDTA in an ethanol-dry ice bath. Reactions were desalted on a C₁₈ Sep Pak (Waters) before electrophoretic analysis.

5.3.3 Instrumentation-CE

The POP5 Performance Optimized Polymers were obtained from PE Biosystems (Foster City, PA). A capillary (75 µM i.d., 50 cm effective length, 56 cm total length) was filled at approximately 1000 psi before each analysis. The optimized run voltage was 12 kV while the capillary temperature was controlled at 50 °C. The eCap ssDNA 100-R Gel, buffer, and neutrally coated capillary (75 µM i.d.) were purchased from Beckman Coulter (Fullerton, CA). A capillary (30 cm effective length, 36 cm total length) was filled at approximately 800 psi and was used for 5 separations at 308 V/cm at 30 °C. The µPAGE-5 (5%T, 5%C) polyacrylamide gel-filled capillary and buffer were purchased from Agilent Technologies (Palo Alto, CA). The capillary length was 56 cm while the effective length was 50 cm. Injection was typically at 10 kV for 10 s and separation at 250 V at 25 °C.

Separations were performed using a Beckman P/ACE CE fitted with an LIF module containing an Argon ion laser (excitation at 488 nm) and a 510 nm emission filter.

5.3.4 Instrumentation-Microdevice

The microdevices, LIF detection system, and switch box were all designed in-house as described earlier. The microdevice was manually filled with POP5 matrix then the injection,
waste, and buffer reservoirs were flushed with buffer. The buffer was removed from the sample well and the sample added just prior to analysis. Injection was typically at 500 V for 5 s and separation was at 3.5 kV. Pullback voltages on the sample and waste wells during separation were 500 and 300 V, respectively.

### 5.4 References


Appendix: Thermal Melting Raw Data and MALDI Mass Spectra

Contained in the following pages is the raw data for the thermodynamic analysis of the Dickerson dodecamer containing thiazole and thiazole N-oxide nucleoside analogs and well as MALDI mass spectrum of the Dickerson dodecamer as well as 28nt oligonucleotide synthesized containing thiazole and thiazole N-oxide analog. The thermodynamic values of $T_m$, $\Delta H$, $\Delta S$, and $\Delta G$ were calculated by the program Meltwin® for each melt curve shown. From the $T_m$ plots of $1/T_m$ versus $\ln(C)$ are obtained in a Van’t Hoff analysis giving averaged values of $\Delta H$, $\Delta S$, and $\Delta G$.

**C-G Dickerson Dodecamer:**

5’-CGCCAATTGCGC-3’

3’-GCGGTTAACGCG-5’

Van’t Hoff Analysis:

$\Delta H = -55.43$ kcal/mol

$\Delta S = -144.16$ cal/K·mol

$\Delta G = -10.72$ kcal/mol

$T_m = 67.9 \degree C$

1/Tm vs ln(C)
$T_m=53.97°C$

**Melt Curve 5.17E-06M**

$b = 1.0 \text{ cm}; \Delta H = -83.3; \Delta S = -230.47; \Delta G = -11.82$

$T_m=62.15°C$

**Melt Curve 1.61E-05M**

$b = 1.0 \text{ cm}; \Delta H = -46.07; \Delta S = -115.46; \Delta G = -10.26$
Melt Curve 1.30E-04M

\( T_m = 69.20^\circ C \)

\[ b = 0.1 \text{ cm}; \Delta H = -62.7; \Delta S = -165.37; \Delta G = -11.41 \]

Melt Curve 3.42E-05M

\( T_m = 64.97^\circ C \)

\[ b = 0.5 \text{ cm}; \Delta H = -54.88; \Delta S = -141.87; \Delta G = -10.88 \]
T_m=75.39°C
Melt Curve 8.96E-04M
b = 0.01 cm; ΔH = -79.34; ΔS = -213.7; ΔG = -13.06

T_m=75.55°C
Melt Curve 4.29E-04M
b = 0.02 cm; ΔH = -70.01; ΔS = -185.35; ΔG = -12.5
T-A Dickerson Dodecamer: 5’-CGCAAATTTCGC-3’
3’-GCGTTTAAAGCG-5’

Van’t Hoff Analysis:
$\Delta H = -79.92 \text{ kcal/mol}$
$\Delta S = -216.87 \text{ cal/K·mol}$
$\Delta G = -12.65 \text{ kcal/mol}$
$T_m = 67.1 ^\circ \text{C}$

$r^2 = 0.980$

1/T<sub>m</sub> vs ln(C)
Melt Curve 1.010E-05

\[ b = 1.0; \Delta H = -61.99; \Delta S = -162.81; \Delta G = -11.49 \]

Melt Curve 1.580E-05

\[ b = 0.5; \Delta H = -65.73; \Delta S = -174.07; \Delta G = -11.74 \]
Melt Curve 7.170E-05

\[ b = 0.1; \Delta H = -72.54; \Delta S = -196.41; \Delta G = -11.63 \]

Melt Curve 6.350E-04

\[ b = 0.02; \Delta H = -81.21; \Delta S = -220.53; \Delta G = -12.81 \]
$b = 0.1; \Delta H = -74.14; \Delta S = -199.67; \Delta G = -12.21$

$T_m = 73.38$
Z-A Dickerson Dodecamer: $5'$-CGCZAATTACGC-3'
3'-GCGATTAZGCG-5'

Van’t Hoff Analysis:
$\Delta H = -35.79$ kcal/mol
$\Delta S = -107.01$ cal/K·mol
$\Delta G = -2.60$ kcal/mol
$T_m = 11.5$ °C

$r^2 = 0.998$
**Melt Curve 6.590E-05**

\[ b = 0.1; \Delta H = -29.51; \Delta S = -84.94; \Delta G = -3.7 \]

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Absorbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.42</td>
</tr>
<tr>
<td>20</td>
<td>0.44</td>
</tr>
<tr>
<td>40</td>
<td>0.46</td>
</tr>
<tr>
<td>60</td>
<td>0.48</td>
</tr>
<tr>
<td>80</td>
<td>0.50</td>
</tr>
<tr>
<td>100</td>
<td>0.52</td>
</tr>
</tbody>
</table>

\[ T_m = 10.45 \]

**Melt Curve 4.226E-04**

\[ b = 0.02; \Delta H = -36.24; \Delta S = -108.35; \Delta G = -2.64 \]

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Absorbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.55</td>
</tr>
<tr>
<td>20</td>
<td>0.60</td>
</tr>
<tr>
<td>40</td>
<td>0.65</td>
</tr>
<tr>
<td>60</td>
<td>0.70</td>
</tr>
<tr>
<td>80</td>
<td>0.75</td>
</tr>
<tr>
<td>100</td>
<td>0.80</td>
</tr>
</tbody>
</table>

\[ T_m = 19.63 \]
Melt Curve 1.317E-03

\[ b = 0.01; \Delta H = -26.21; \Delta S = -74.92; \Delta G = -2.97 \]
Z-G Dickerson Dodecamer: 5’-CGCZAATTGC-3’
3’-GCGTAAZGC-5’

Van’t Hoff Analysis:
ΔH = -53.18 kcal/mol
ΔS = -164.74 cal/K·mol
ΔG = -2.08 kcal/mol
T_m = 16.5 °C

$r^2 = 0.956$

1/T_m vs ln(C)
Melt Curve 1.856E-05
b = 0.5; ΔH = -23.19; ΔS = -59.28; ΔG = -4.81

Tm = 13.46

Melt Curve 4.88E-05
b = 0.2; ΔH = -33.29; ΔS = -95.93; ΔG = -3.53

Tm = 14.65
Melt Curve 9.514E-05
b = 0.1; ΔH = -18.65; ΔS = -46.3; ΔG = -4.29

Melt Curve 3.189E-04
b = 0.02; ΔH = -21.19; ΔS = -55.76; ΔG = -3.89
Melt Curve 7.128E-04

\[ b = 0.01; \Delta H = -21.86; \Delta S = -59.15; \Delta G = -3.57 \]

\[ T_m = 24.06 \]
X-A Dickerson Dodecamer: 5’-CGCXAATTACGC-3’
3’-GCGATTAAXGCG-5’

Van’t Hoff Analysis:

$T_m = <0$
X-G Dickerson Dodecamer: 5’-CGCXAATTGCGC-3’
3’-GCGGTTAAXGCG-5’

Van’t Hoff Analysis:
ΔH = -141.74 kcal/mol
ΔS = -442.97 cal/K·mol
ΔG = -4.35 kcal/mol
T_m = 26.6 °C
Melt Curve 8.190E-04
b = 0.02; ΔH = -43.63; ΔS = -126.71; ΔG = -4.33

Melt Curve 1.890E-03
b = 0.01; ΔH = -41.68; ΔS = -121.45; ΔG = -4.01
Melt Curve 4.036E-04

\[ b = 0.02; \Delta H = -29.03; \Delta S = -78.38; \Delta G = -4.72 \]

\[ T_m = 35.98 \]
MALDI mass spectrum of 5'-CGCTAATTTCG-3'
oxide-T-Dickerson peak 1
MALDI mass spectrum of 5'-CGCoxAATTTGGC-3'

oxide-T-Dickerson peak 2
MALDI mass spectrum of 5’- ACTGThTCTCCCTATAGTAGCGTATTA-3’
MALDI mass spectrum of 5’- ACTGOxTCTCCCTATAGTGAGTCGTATTA-3’

8512.59
MALDI mass spectrum of 5’- TCTCCCTATAGTGAGTCGTATTA-3’
Vita

Hannah D. Farquar was born in Lafayette, Louisiana, on September 24, 1974. She was raised in Sunset, Louisiana, where she attended school. She earned a bachelor of science degree in chemistry with a biochemistry option from Louisiana Tech University in Ruston, Louisiana, in May of 1996. After graduation she worked as a laboratory technician at Southern Petroleum Laboratories in Lafayette, Louisiana, until 1997 when she began pursuing a doctoral degree from Louisiana State University Agricultural and Mechanical College in Baton Rouge, Louisiana, under the direction of Dr. Robert P. Hammer.

She has accepted a postdoctoral appointment at Oak Ridge National Laboratories in Oak Ridge, Tennessee, under the direction of Dr. Winston Chen developing new hybridization techniques for DNA mutation detection. The degree of Doctor of Philosophy will be conferred at the 2002 December Commencement.