



Publicly Accessible Penn Dissertations

Summer 8-14-2009

The Synthesis and Characterization of Glycol Nucleic Acids

Mark K. Schlegel

University of Pennsylvania, mschleg@gmail.com

Follow this and additional works at: <http://repository.upenn.edu/edissertations>

 Part of the [Organic Chemistry Commons](#)

Recommended Citation

Schlegel, Mark K., "The Synthesis and Characterization of Glycol Nucleic Acids" (2009). *Publicly Accessible Penn Dissertations*. 13.
<http://repository.upenn.edu/edissertations/13>

This paper is posted at Scholarly Commons. <http://repository.upenn.edu/edissertations/13>
For more information, please contact libraryrepository@pobox.upenn.edu.

The Synthesis and Characterization of Glycol Nucleic Acids

Abstract

A project was undertaken to elucidate the properties of the simplified glycol nucleic acid (GNA). GNA is novel in the fact that nucleic acid duplexes composed entirely of GNA show thermal stabilities that are superior to those of analogous DNA or RNA duplexes. Furthermore, GNA has been shown to pair with complementary sequences of RNA, but not with DNA.

The first step towards understanding the thermal stabilities of GNA duplexes is the development of a straightforward synthesis of the phosphoramidites for solid phase oligonucleotide synthesis. Chapter 2 describes work towards a new set of exocyclic amino protection groups which could be removed in less time and under milder conditions than those previously reported. This new scheme results in a vastly improved synthesis of the individual phosphoramidites and allows for quicker access to the subsequent oligonucleotides.

With easy access to GNA oligonucleotides in hand, the next goal was to use spectroscopy to compare the duplex formation properties with that of DNA as outlined in Chapter 3. These studies pointed to the preorganization of the single strands and increased stacking interactions as the main factors that increase the stability of GNA duplexes.

Although the studies in Chapter 3 present a basic understanding of GNA duplex formation, it did not provide the direct structural insight that was desired. Chapter 4 presents the synthesis and pairing properties of three artificial metal-mediated base pairs in GNA duplexes; one of which was used as a handle in an 8-mer duplex for phasing the crystallographic data, thereby providing the initial structural insight that was desired.

Uncertain as to the extent by which the artificial metal-mediated base pair perturbs the overall structure of GNA, several other structures of GNA duplexes containing exclusively Watson-Crick base pairs are presented in Chapter 5. The most impressive structure of a 6-mer GNA duplex displays many common structural features to that of the 8-mer duplex containing artificial base pairs.

Overall this research has provided insight into the exceptional duplex formation properties of glycol nucleic acids and should provide the basis for future research on the application of GNA duplexes for various biological or technical purposes.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Chemistry

First Advisor

Dr. Eric Meggers

Keywords

Nucleic Acid Analogs, Biomimetic Structures, Duplex Stability, Nucleic Acid Crystallography, Thermodynamics, Supramolecular Chemistry

Subject Categories

Organic Chemistry

THE SYNTHESIS AND CHARACTERIZATION
OF GLYCOL NUCLEIC ACIDS

Mark K. Schlegel

A DISSERTATION

in

Chemistry

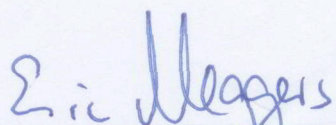
Presented to the Faculties of the University of Pennsylvania

in

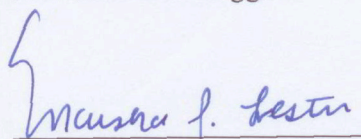
Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2009



Supervisor of Dissertation
Professor Eric Meggers



Graduate Group Chairperson
Professor Marsha Lester

Professor Ronen Marmorstein

Professor Barry Cooperman

Professor Jeffery Saven

Dedicated to my family

First and foremost, I would like to acknowledge my research advisor, Dr. Eric Meggers. Over the years Eric's presence in lab was very helpful in learning how to approach scientific problems and conduct research in a thorough but efficient manner. Eric also maintains a passion for science that seems to be infectious to those around him. This passion and all the things I learned working with Eric are important lessons that will no doubt be helpful in my future career. I want to thank him for his dedication and hard work for the benefit of his students and I wish him continued success in the future.

Another person who deserves a great amount of my gratitude is Dr. Lars-Oliver Essen. Much of my later research was centered around crystallography and Prof. Essen unselfishly dedicated a great amount of effort into helping me with this aspect of my project. I can honestly say that much of my success would not have been possible without his help and enthusiasm for the project at hand. Furthermore, I also need to thank the members of his research group, each of whom was very willing to lend a hand with different aspects of the process. I wish him the best of luck in all of his future endeavors.

I would also like to thank the members of my committee, Professors Ronen Marmorstein, Barry Cooperman, and Jeffery Saven. These three professors were very helpful over the years in discussing my research and offering opinions and suggestions as outsiders to the project. I must specifically thank Prof. Marmorstein for the extra effort he and his lab provided in helping me with the beginning parts of the crystallography part of my project.

I was in the unique situation of being a part of two different lab environments; first at Penn and most recently at the University of Marburg, which has allowed me the pleasure of getting to know many different people over the course of my research. Initially at Penn, I must thank Howie Bregman and Doug Williams as they were tremendous resources upon joining the Meggers lab, and later became good friends. Although I had little real life synthetic experience at the time, they were both more than willing to help and make suggestions during the first couple of months in lab. Additional new members of the lab, Nick Pagano and Craig Streu, also became good friends over the years and I will miss the good times we had in and out of lab in Philadelphia and in Europe. I should specifically thank Nick as we made the trip to Marburg together and got to experience many different and exciting things together. Finally, I must thank Ekin Atilla-Gokcumen and Jasna Maksimoska who were very helpful in the beginning on the biological side of things in lab. I think having them in the lab made joining a group that is more dedicated to synthetic chemistry less of a daunting experience. Also I need to thank our secretary, Andrea Carrol, as she was always happy to help and chat with us during our lunchtime breaks.

During my stay in Marburg, I also had the privilege of meeting many new people. I would first like to thank Martin Bohn and Johannes Beerlink as both of these guys were very welcoming to us Americans when we arrived at the department. These guys not only made me feel welcome at lab, but also exposed me to many different things in Germany, most particularly soccer and beer. I also had the opportunity to welcome many new lab mates into our group; Nick Bierman, Matthis Bischof, Sebastian Blanck, Sandra

Dieckmann, Li Feng, Lei Gong, Stefan Mollin, Friedrich Wendel, Alexander Wilbuer, and Gunther Zimmermann. Over the years I have grown to be friends with all of these guys and I will miss the good times we had in and out of lab. The list would not be complete without our lab technician, Katja Kraeling, and our secretary, Ina Pinnschmidt who have both helped me tremendously with many issues in and pertaining to lab.

I also would like to thank the facilities directors at both Penn and Marburg as they all maintained excellent facilities. In particular, George Furst and Rakesh Kohli who maintained the NMR facility and MS facility at Penn, respectively. At Marburg, Xiulan Xie and Uwe Linne also oversee the NMR and MS facilities, respectively. I can say that without the hard work and dedication of these people it would have made our research very difficult.

Not to be forgotten are the members of my extended family who have provided support in many different ways throughout the years. Most especially, I should thank my parents, Brian and Stephanie, who always encouraged me to work hard, believe in myself, and make the best out of what I choose to do with my life. These values will always be a part of who I am. I also need to thank my big brother, Matt, for his encouragement, support, and just being a friend whenever I needed one. I wish him the best of luck with all of the new experiences that come with having a new family and continued success with his career.

Last, but certainly not least, I must thank my girlfriend, Julia Fischer. It would be hard to completely explain how much she helped me in and out of lab over the last few years. Julia provided me with an endless amount of support and encouragement and her

dedication and hard work towards becoming a teacher (even teaching a little German to me against a slight amount of resistance) have inspired me. The experiences that we shared together were what made my time in Marburg truly special. I can say that the time I spent here would not have been complete without her. I am excited to continue our journey together as we move to Berlin.

ABSTRACT

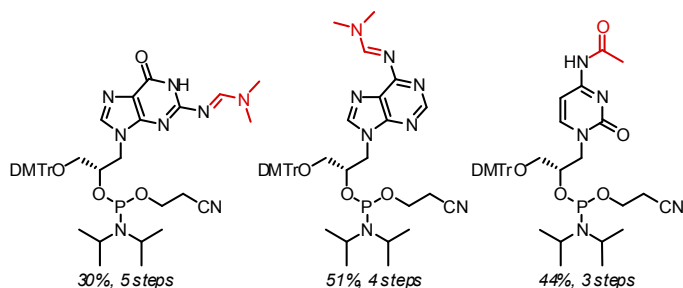
THE SYNTHESIS AND CHARACTERIZATION OF GLYCOL NUCLEIC ACIDS

Mark K. Schlegel

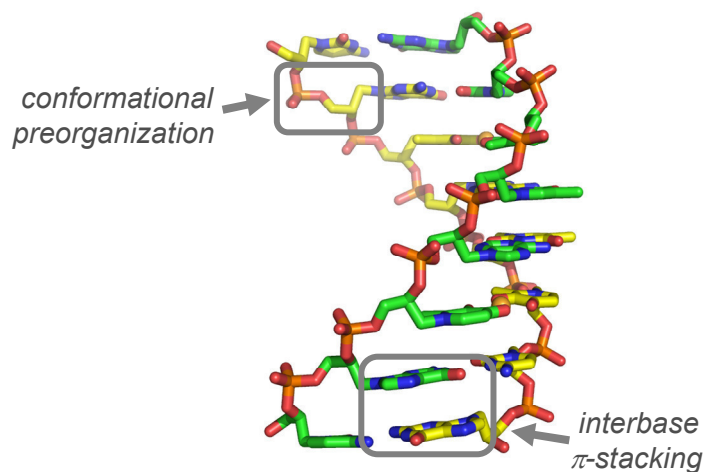
Advisor: Dr. Eric Meggers

A project was undertaken to elucidate the properties of the simplified glycol nucleic acid (GNA). GNA is novel in the fact that nucleic acid duplexes composed entirely of GNA show thermal stabilities that are superior to those of analogous DNA or RNA duplexes. Furthermore, GNA has been shown to pair with complementary sequences of RNA, but not with DNA.

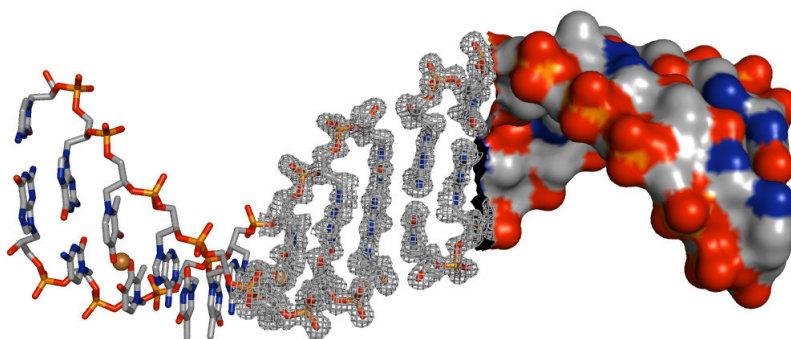
The first step towards understanding the thermal stabilities of GNA duplexes is the development of a straightforward synthesis of the phosphoramidites for solid phase oligonucleotide synthesis. Chapter 2 describes work towards a new set of exocyclic amino protection groups which could be removed in less time and under milder conditions than those previously reported. This new scheme results in a vastly improved synthesis of the individual phosphoramidites and allows for quicker access to the subsequent oligonucleotides.



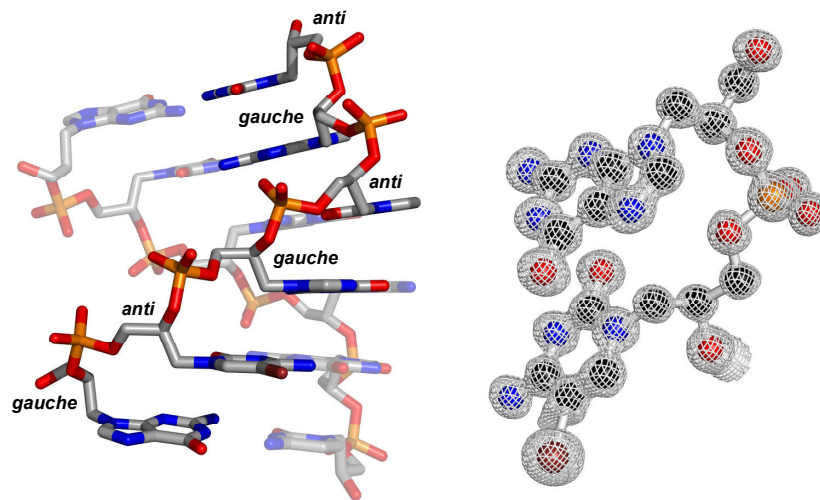
With easy access to GNA oligonucleotides in hand, the next goal was to use spectroscopy to compare the duplex formation properties with that of DNA as outlined in Chapter 3. These studies pointed to the preorganization of the single strands and increased stacking interactions as the main factors that increase the stability of GNA duplexes.



Although the studies in Chapter 3 present a basic understanding of GNA duplex formation, it did not provide the direct structural insight that was desired. Chapter 4 presents the synthesis and pairing properties of three artificial metal-mediated base pairs in GNA duplexes; one of which was used as a handle in an 8-mer duplex for phasing the crystallographic data; thereby providing the initial structural insight that was desired.



Uncertain as to the extent by which the artificial metal-mediated base pair perturbs the overall structure of GNA, several other structures of GNA duplexes containing exclusively Watson-Crick base pairs are presented in Chapter 5. The most impressive structure of a 6-mer GNA duplex displays many common structural features to that of the 8-mer duplex containing artificial base pairs.



Overall this research has provided insight into the exceptional duplex formation properties of glycol nucleic acids and should provide the basis for future research on the application of GNA duplexes for various biological or technical purposes.

Table of Contents

Dedication	ii
Acknowledgements	iii
Abstract.....	vii
List of Schemes.....	xiii
List of Figures.....	xv
List of Tables	xxvii
Chapter 1: Introduction	1
Chapter 1.1. Background information.....	2
Chapter 1.2. Acyclic analogs of nucleic acids containing phosphodiester linkages	6
Chapter 1.3. Previous work with glycol nucleic acids (GNA).....	8
Chapter 1.4. References.....	12
Chapter 2: Synthesis of glycol nucleoside phosphoramidites and their incorporation into GNA oligonucleotides.....	14
Chapter 2.1. Initial synthesis of glycol nucleoside phosphoramidites.....	15
Chapter 2.2. Initial protocol for GNA oligonucleotide synthesis and purification.	22
Chapter 2.3. Development of new routes towards (S)-G phosphoramidite.....	31
Chapter 2.4. Modified synthesis of (S)-A and (S)-C phosphoramidites.....	37

Chapter 2.5. Second generation GNA oligonucleotide synthesis and purification .	40
Chapter 2.6. Conclusions.....	46
Chapter 2.7. Experimental procedures.....	48
Chapter 2.8. References.....	79
Appendix to chapter 2: ^1H , ^{13}C , ^{31}P NMR spectra, IR spectra; Oligonucleotide synthesizer protocols.....	81
Chapter 3: Thermal and thermodynamic properties of glycol nucleic acid duplexes...	136
Chapter 3.1. From single incorporation to completely acyclic oligonucleotides.....	137
Chapter 3.2. Watson-Crick pairing, salt dependence, CD spectroscopy	142
Chapter 3.3. Thermal and thermodynamic stabilities of GNA duplexes.....	149
Chapter 3.3.1. Thermodynamics of GNA duplex formation	149
Chapter 3.3.2. Probing GNA duplex formation using circular dichroism ..	160
Chapter 3.3.3. Intercalators and NMR analysis of single glycol nucleosides	165
Chapter 3.4. Crosspairing of GNA single strands with DNA and RNA.....	168
Chapter 3.4.1. Initial crosspairing results of (<i>S</i>)- and (<i>R</i>)-GNA with DNA and RNA	168
Chapter 3.4.2. Crosspairing of mixed sequences of (<i>S</i>)-GNA with RNA.....	174
Chapter 3.5. Conclusions.....	187

Chapter 3.6. Experimental procedures.....	189
Chapter 3.7. References.....	205
Appendix to chapter 3: ^1H and ^{31}P NMR spectra	208
Chapter 4: Metal-mediated base pairing in GNA.....	227
Chapter 4.1. Synthesis of unnatural metal-mediated base pairs	228
Chapter 4.2. Incorporation of (<i>S</i>)-H and (<i>S</i>)-P into GNA oligonucleotides.....	237
Chapter 4.3. Thermal stabilities of metal-mediated base pairs in GNA.....	238
Chapter 4.4. Metal-mediated base pairing as a tool for crystallography	250
Chapter 4.4.1. Duplex structure of a (<i>S</i>)-GNA duplex containing H-Cu-H	250
Chapter 4.4.2. New insight into the high duplex stability of GNA.....	259
Chapter 4.5. Conclusions.....	262
Chapter 4.6. Experimental procedures.....	264
Chapter 4.7. References.....	282
Appendix to chapter 4: ^1H , ^{13}C , ^{31}P NMR spectra, IR spectra, Crystallographic tables.....	286
Chapter 5: Crystallography of GNA duplexes.....	322
Chapter 5.1. Derivatizing glycol nucleic acids for phasing crystallographic data..	323
Chapter 5.1.1. Synthesis of brominated nucleoside phosphoramidite derivatives	325
Chapter 5.1.2. Phosphoroselenoate derivatives of GNA oligonucleotides ...	330

Chapter 5.2. Crystallographic conditions.....	337
Chapter 5.3. Structure of the GNA duplex 3'-GCGCGC-2'	340
Chapter 5.4. Structure of the GNA duplex 3'-CTCTAGAG-2'	350
Chapter 5.5. Structure of the GNA duplex 3'-ATGCGCAT-2'	357
Chapter 5.6. Conclusions.....	360
Chapter 5.7. Experimental procedures.....	362
Chapter 5.8. References.....	374
Appendix to chapter 5: ¹ H, ¹³ C, ³¹ P NMR spectra, IR spectra, Crystallographic tables.....	376

List of Schemes

Scheme 2.1. Synthesis of (<i>S</i>)-G phosphoramidite.....	17
Scheme 2.2. Synthesis of (<i>R</i>)-4,4'-dimethoxytritylglycidol (2.11).....	18
Scheme 2.3. Synthesis of (<i>S</i>)-A phosphoramidite.....	19
Scheme 2.4. Synthesis of phosphoramidites of the pyrimidine bases: a) (<i>S</i>)-C, b) (<i>S</i>)-T.....	21
Scheme 2.5. Synthesis of succinyl esters: a) guanine glycol nucleoside, b) adenine glycol nucleoside.....	23
Scheme 2.6. Synthesis of succinyl esters: a) cytosine glycol nucleoside, b) thymine glycol nucleoside.....	24
Scheme 2.7. General scheme for the synthesis of glycol nucleoside functionalized solid supports using compound 2.24 as an example.	25

Scheme 2.8. Overview of the procedure for initial GNA oligonucleotide purification.	27
Scheme 2.9. Attempted ring-opening using protected guanine derivative compound 2.31.	32
Scheme 2.10. Synthesis of <i>O</i>-6-benzyloxyguanine (2.36).	34
Scheme 2.11. Synthesis of (<i>S</i>)-G^{DPC} phosphoramidite.	34
Scheme 2.12. Synthesis of (<i>S</i>)-G* phosphoramidite.	36
Scheme 2.13. Synthesis of (<i>S</i>)-A* phosphoramidite.	38
Scheme 2.14. Synthesis of (<i>S</i>)-C* phosphoramidite.	39
Scheme 2.15. Synthesis of the succinyl ester of (<i>S</i>)-C* glycol nucleoside.	40
Scheme 2.16. Synthesis of succinyl esters: a) (<i>S</i>)-G* glycol nucleoside, b) (<i>S</i>)-A* glycol nucleoside.	41
Scheme 2.17. Overview of the procedure for streamlined GNA oligonucleotide purification.	42
Scheme 3.1. Synthesis of (<i>S</i>)-N phosphoramidite.	178
Scheme 3.2. Synthesis of (<i>S</i>)-D phosphoramidite.	179
Scheme 3.3. Attempted synthesis of (<i>S</i>)-I phosphoramidite starting with: a) inosine, and b) 6-chloropurine.	180
Scheme 3.4. Synthesis of benzylglycidol.	181
Scheme 3.5. Attempted synthesis of (<i>S</i>)-I phosphoramidite using compound 3.14 for ring opening of 6-chloropurine.	182
Scheme 3.6. Successful synthesis of (<i>S</i>)-I phosphoramidite.	183

Scheme 4.1. Synthesis of 3-benzyloxy-2-methyl-4-pyridone (4.3).....	230
Scheme 4.2. Synthesis of (S)-H ^{Piv} phosphoramidite.....	231
Scheme 4.3. Synthesis of (S)-H phosphoramidite.....	232
Scheme 4.4. Attempted synthesis of compound 4.12.....	233
Scheme 4.5. Synthesis of (S)-H ^{NB} phosphoramidite.....	234
Scheme 4.6. Synthesis of (S)-P phosphoramidite.	236
Scheme 5.1. Synthesis of (S)- ^{Br} U phosphoramidite.....	326
Scheme 5.2. Attempted synthesis of (S)- ^{Br} C phosphoramidite	327
Scheme 5.3. Synthesis of (S)- ^{Br} C phosphoramidite.....	328
Scheme 5.4. Synthesis of (S)- ^{Br} C* phosphoramidite.....	329
Scheme 5.5. Overview of the procedure for synthesizing GNA oligonucleotides containing phosphoroselenoate linkages.....	331

List of Figures

Figure 1.1. Schematic representation of oligonucleotide directed inhibition of gene expression: a) the antisense approach and b) the antigene approach.....	3
Figure 1.2. Several examples of sugar modified DNA/RNA analogs.	4
Figure 1.3. Structures of different acyclic nucleosides: a) Constructed from deleting one or more atoms of the ribose sugar. b) Constructed from simple alkyl chains.....	7
Figure 1.4. A completely artificial base pair in DNA.....	8

Figure 1.5. Comparison of the constitution of DNA and RNA with that of the two enantiomers of GNA.	9
Figure 2.1. HPLC traces of the GNA oligonucleotide 3'-TTTTTTTTTT-2'.....	28
Figure 2.2. HPLC trace of the purified GNA oligo 3'-CACATTATTGTTGTA-2'.	29
Figure 2.3. MALDI spectrum of the GNA oligonucleotide 3'-CACATTATTGTTGTA-2' (calc for C ₁₁₈ H ₁₅₇ N ₅₀ O ₇₆ P ₁₄ [M-H] ⁻ 3925.5).	30
Figure 2.4. Crude "Trityl-OFF" HPLC trace of the GNA oligonucleotide 3'-CATGTCGTGCGTA-2'.....	44
Figure A2.1.1. ¹ H NMR spectrum of compound 2.3 (500 MHz, DMSO- <i>d</i> ₆).	82
Figure A2.2.1. ¹ H NMR spectrum of compound 2.4 (500 MHz, DMSO- <i>d</i> ₆).	83
Figure A2.3.1. ¹ H NMR spectrum of compound 2.7 (360 MHz, DMSO- <i>d</i> ₆).	84
Figure A2.4.1. ¹ H NMR spectrum of compound 2.9a (300 MHz, CDCl ₃).....	85
Figure A2.5.1. ³¹ P NMR spectrum of compound (<i>S</i>)-G (121 MHz, CDCl ₃).....	86
Figure A2.6.1. ¹ H NMR spectrum of compound 2.11 (500 MHz, CDCl ₃).....	87
Figure A2.7.1. ¹ H NMR spectrum of compound 2.13 (300 MHz, CDCl ₃).....	88
Figure A2.8.1. ¹ H NMR spectrum of compound 2.15 (300 MHz, CDCl ₃).....	89
Figure A2.9.1. ³¹ P NMR spectrum of compound (<i>S</i>)-A (121 MHz, CDCl ₃).....	90
Figure A2.10.1. ¹ H NMR spectrum of compound 2.17 (300 MHz, CDCl ₃).....	91
Figure A2.11.1. ³¹ P NMR spectrum of compound (<i>S</i>)-C (121 MHz, CDCl ₃).....	92
Figure A2.12.1. ¹ H NMR spectrum of compound 2.19 (500 MHz, CDCl ₃).....	93
Figure A2.13.1. ³¹ P NMR spectrum of compound (<i>S</i>)-T (121 MHz, CDCl ₃).....	94

Figure A2.14.1. ^1H NMR spectrum of compound 2.21 (500 MHz, CDCl_3).....	95
Figure A2.15.1. ^1H NMR spectrum of compound 2.22 (500 MHz, CDCl_3).....	96
Figure A2.16.1. ^1H NMR spectrum of compound 2.23 (500 MHz, CDCl_3).....	97
Figure A2.17.1. ^1H NMR spectrum of compound 2.24 (500 MHz, CDCl_3).....	98
Figure A2.18.1. ^1H NMR spectrum of compound 2.29 in (360 MHz, $\text{DMSO-}d_6$).....	99
Figure A2.19.1. ^1H NMR spectrum of compound 2.31 (360 MHz, $\text{DMSO-}d_6$).....	100
Figure A2.20.1. ^1H NMR spectrum of compound 2.34 (300 MHz, D_2O).	101
Figure A2.21.1. ^1H NMR spectrum of compound 2.36 (300 MHz, $\text{DMSO-}d_6$).....	102
Figure A2.22.1. ^1H NMR spectrum of compound 2.37 (500 MHz, CDCl_3).....	103
Figure A2.22.2. ^{13}C NMR spectrum of compound 2.37 (125 MHz, CDCl_3).	104
Figure A2.22.3. IR spectrum of compound 2.37 (film).....	105
Figure A2.23.1. ^1H NMR spectrum of compound 2.38 (300 MHz, $\text{DMSO-}d_6$).....	106
Figure A2.23.2. ^{13}C NMR spectrum of compound 2.38 (75 MHz, $\text{DMSO-}d_6$).....	107
Figure A2.23.3. IR spectrum of compound 2.38 (film).....	108
Figure A2.24.1. ^1H NMR spectrum of compound 2.39 (500 MHz, CDCl_3).....	109
Figure A2.24.2. ^{13}C NMR spectrum of compound 2.39 (125 MHz, CDCl_3).	110
Figure A2.24.3. IR spectrum of compound 2.39 (film).....	111
Figure A2.25.1. ^{31}P NMR spectrum of phosphoramidite (<i>S</i>)- G^{DPC} (121 MHz, CDCl_3) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).....	112
Figure A2.26.1. ^1H NMR spectrum of compound 2.41 (300 MHz, CDCl_3).....	113
Figure A2.26.2. ^{13}C NMR spectrum of compound 2.41 (75 MHz, CDCl_3).	114
Figure A2.26.3. IR spectrum of compound 2.41 (solid).....	115

Figure A2.27.1. ³¹P NMR spectrum of phosphoramidite (S)-G* (162 MHz, CDCl₃) with trimethyl phosphate as an internal standard (δ = 3.06 ppm).	116
Figure A2.28.1. ¹H NMR spectrum of compound 2.44 (300 MHz, CDCl₃).	117
Figure A2.28.2. ¹³C NMR spectrum of compound 2.44 (75 MHz, CDCl₃).	118
Figure A2.28.3. IR spectrum of compound 2.44 (solid).	119
Figure A2.29.1. ³¹P NMR spectrum of phosphoramidite (S)-A* (162 MHz, CDCl₃) with trimethyl phosphate as an internal standard (δ = 3.06 ppm).	120
Figure A2.30.1. ¹H NMR spectrum of compound 2.46 (300 MHz, CDCl₃).	121
Figure A2.30.2. ¹³C NMR spectrum of compound 2.46 (75 MHz, CDCl₃).	122
Figure A2.30.3. IR spectrum of compound 2.46 (solid).	123
Figure A2.31.1. ³¹P NMR spectrum of phosphoramidite (S)-C* (162 MHz, CDCl₃) with trimethyl phosphate as an internal standard (δ = 3.06 ppm).	124
Figure A2.32.1. ¹H NMR spectrum of compound 2.47 (300 MHz, CDCl₃).	125
Figure A2.33.1. ¹H NMR spectrum of compound 2.48 (300 MHz, CDCl₃).	126
Figure A2.34.1. ¹H NMR spectrum of compound 2.49 (300 MHz, CDCl₃).	127
Figure 3.1. A representative thermal melting curve as monitored by UV spectroscopy of the GNA duplex 3'-AATATTATTATTTTA-2' and its complement.	140
Figure 3.2. Comparison of the effect of sodium concentration on the thermal stability of a GNA duplex and DNA duplex of the sequence CACATTATTGTTGTA and its complement.	144

Figure 3.3. Thermal melting temperatures (T_m) of GNA and DNA duplexes of the sequence CACATTATTGTTGTA and its complement in the presence of DMF....	145
Figure 3.4. CD spectra of a (<i>S</i>)-GNA and DNA duplexes of the sequence CACATTATTGTTGTA at a duplex concentration of 10 μM.	146
Figure 3.5. Temperature-dependent CD measurement of the (<i>S</i>)-GNA duplex 3'-CACATTATTGTTGTA-2' and its complement at a duplex concentration of 2 μM in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride.	147
Figure 3.6. Thermal melting temperatures (T_m) of self-complementary GNA duplexes of the sequence 3'-CGAATTCG-2' (open circles) and 3'-CGAATTCGA-2' (filled squares).	158
Figure 3.7. Thermal melting temperatures (T_m) of self-complementary DNA duplexes of the sequence 5'-CGAATTCG-3' (open circles) and 5'-CGAATTCGA-3' (filled squares).	159
Figure 3.8. Melting of the GNA single strand 3'-CACATTATTGTTGTA-2' as monitored by circular dichroism (CD).	160
Figure 3.9. Melting of the DNA single strand 5'-CACATTATTGTTGTA-3' as monitored by circular dichroism (CD).	161
Figure 3.10. CD spectra of GNA single strands with varying length and sequence at concentrations of 20 μM (15mers), 30 μM (10mer), and 60 μM (5mers).....	163
Figure 3.11. Effect of salt concentration on the CD spectra of GNA single strands with the sequence 3'-CACATTATTGTTGTA-2' at a concentration of 20 μM.	164

Figure 3.12. Structure of (<i>S</i>)-1-(2,3-dihydroxypropyl)cytosine determined by NMR.	165
Figure 3.13. CD spectra of GNA, DNA, and RNA homoduplexes of the sequence A₁₅:T₁₅ at a duplex concentration of 4 μM.	170
Figure 3.14. CD spectra of a (<i>S</i>)-GNA T₁₅:RNA A₁₅ heteroduplex at a duplex concentration of 4 μM and the corresponding single strands at a concentration of 4 μM.	171
Figure 3.15. CD spectra of a (<i>S</i>)-GNA T₁₅:DNA A₁₅ heteroduplex at a duplex concentration of 4 μM and the corresponding single strands at a concentration of 4 μM.	172
Figure 3.16. CD and Job plot analysis of the (<i>S</i>)-GNA T₁₅:RNA A₁₅ heteroduplex.	173
Figure 3.17. Crosspairing of mixed sequences of (<i>S</i>)-GNA with RNA.	175
Figure 3.18. Structure of the three unnatural base pairs used in this study and their possible hydrogen bonding patterns.	177
Figure A3.1.1. ¹H NMR spectrum of the N-9 regisomer of compound 3.2 (300 MHz, CDCl₃).	209
Figure A3.1.2. ¹H NMR spectrum of the N-7 regisomer of compound 3.2 (300 MHz, CDCl₃).	210
Figure A3.2.1. ¹H NMR spectrum of compound 3.3 (300 MHz, CDCl₃).	211
Figure A3.3.1. ¹H NMR spectrum of phosphoramidite (<i>S</i>)-N (121 MHz, CDCl₃).	212
Figure A3.4.1. ¹H NMR spectrum of compound 3.5 (300 MHz, CDCl₃).	213

Figure A3.5.1. ^1H NMR spectrum of compound 3.6 (300 MHz, CDCl_3).....	214
Figure A3.6.1. ^1H NMR spectrum of phosphoramidite (<i>S</i>)-D (121 MHz, CDCl_3). ..	215
Figure A3.7.1. ^1H NMR spectrum of the two tautomeric forms (?) of compound 3.8a (360 MHz, CDCl_3).....	216
Figure A3.8.1. ^1H NMR spectrum of compound 3.11 (500 MHz, acetone- d_6).....	217
Figure A3.9.1. ^1H NMR spectrum of compound 3.14 (500 MHz, CDCl_3).....	218
Figure A3.10.1. ^1H NMR spectrum of compound 3.15 (500 MHz, CDCl_3).....	219
Figure A3.11.1. ^1H NMR spectrum of compound 3.16 (360 MHz, MeOD).....	220
Figure A3.12.1. ^1H NMR spectrum of compound 3.17 (360 MHz, MeOD).....	221
Figure A3.13.1. ^1H NMR spectrum of compound 3.18 (360 MHz, MeOD).....	222
Figure A3.14.1. ^1H NMR spectrum of compound 3.20 (500 MHz, CDCl_3).....	223
Figure A3.15.1. ^1H NMR spectrum of compound 3.8c (500 MHz, CDCl_3).....	224
Figure A3.16.1. ^1H NMR spectrum of compound 3.9c (360 MHz, CDCl_3).....	225
Figure A3.17.1. ^1H NMR spectrum of phosphoramidite (<i>S</i>)-I (162 MHz, CDCl_3).....	226
Figure 4.1. Constitution of the hydroxypyridone (H) and pyridylpurine (P) homobase pairs.....	229
Figure 4.2. Copper(II)-dependent stability of a GNA duplex containing a hydroxypyridone (H) homo-base pair with the reference curve shown for an A•T base pair.....	240

Figure 4.3. Nickel(II)-dependent stability of a GNA duplex containing a pyridylpurine (P) homo-base pair with the reference curve shown for an A•T base pair.242

Figure 4.4. Copper(II)-dependent stability of a GNA duplex containing a hydroxypyridone (H):pyridylpurine (P) hetero-base pair with the reference curve shown for an A•T base pair.....244

Figure 4.5. Thermal melting curves monitored by UV spectroscopy of 1:1 mixtures (2 μ M each strand) of 3'-AATATTATTATTTTA-2':3'-TAAAATAATAATATT-2' ($T_m = 51$ °C), 3'-AATATTATTATTTTA-2':3'-TAAAATATTAATATT-2' ($T_m = 36$ °C), and 3'-AATATTATTATTTTA-2':3'-TAAAATAATAATATT-2' with 4 μ M HgCl₂ ($T_m = 41$ °C).....245

Figure 4.6. HPLC traces of the GNA oligonucleotide 3'-AATAPTATTAH^{NB}TTTA-2' before (dotted line) and after (solid line) exposure to UV light for three hours.247

Figure 4.7. Thermal melting curves monitored by UV spectroscopy of 3'-AAT APT ATT AHT TTA-2' and its complement at 2 μM duplex concentration (solid line)......248
Figure 4.8. Thermal melting curves monitored by UV spectroscopy of 3'-AATATTAPTATTTTA-2' and its complement at 2 μM duplex concentration with two equivalents of nickel. Also shown is the same duplex containing two equivalents of nickel followed by the addition of 2 equivalents of copper.....249
Figure 4.9. Thermal melting curves monitored by UV spectroscopy of (S)-GNA oligos containing H:H metallo-base pairs for crystallography at a duplex concentration of 2 μM and 8 μM added CuSO₄......252
Figure 4.10. Overall GNA duplex structure. The structure has been deposited in the Protein Data Bank under PDB code 2JJA.254
Figure 4.11. Electron density of the H-Cu-H base pair [a] and the terminal G-C base pair [b]......256
Figure 4.12. Details of the GNA duplex structure.....257
Figure 4.13. (S)-1-(2,3-dihydroxypropyl)cytosine (colored balls and sticks) superimposed with the conformation of a cytosine nucleotide within a GNA duplex (green sticks)......260
Figure A4.1.1. ¹H NMR spectrum of compound 4.2 (500 MHz, CDCl₃).....287
Figure A4.2.1. ¹H NMR spectrum of compound 4.3 (500 MHz, CDCl₃).....288
Figure A4.3.1. ¹H NMR spectrum of compound 4.4 (500 MHz, CDCl₃).....289
Figure A4.3.2. ¹³C NMR spectrum of compound 4.4 (125 MHz, CDCl₃).290

Figure A4.3.3. IR spectrum of compound 4.4 (film).	291
Figure A4.4.1. ^1H NMR spectrum of compound 4.5 (500 MHz, CDCl_3).....	292
Figure A4.4.2. ^{13}C NMR spectrum of compound 4.5 (90 MHz, CDCl_3).	293
Figure A4.4.3. IR spectrum of compound 4.5 (film).	294
Figure A4.5.1. ^1H NMR spectrum of compound 4.7 (500 MHz, CDCl_3).....	295
Figure A4.6.1. ^{31}P NMR spectrum of phosphoramidite (<i>S</i>)- H^{Piv} (121 MHz, CDCl_3)....	296
.....	
Figure A4.7.1. ^1H NMR spectrum of compound 4.8 (500 MHz, 373K, $\text{DMSO-}d_6$).....	297
.....	
Figure A4.7.2. ^{13}C NMR spectrum of compound 4.8 (125 MHz, 373K, $\text{DMSO-}d_6$).....	298
.....	
Figure A4.7.3. IR spectrum of compound 4.8 (film).	299
Figure A4.8.1. ^{31}P NMR spectrum of phosphoramidite (<i>S</i>)- H^{DPC} (121 MHz, CDCl_3)..	300
.....	
Figure A4.9.1. ^1H NMR spectrum of compound 4.10 (500 MHz, CDCl_3).....	301
Figure A4.10.1. ^1H NMR spectrum of compound 4.11 (360 MHz, MeOD).	302
Figure A4.11.1. ^1H NMR spectrum of compound 4.16 (500 MHz, CDCl_3).....	303
Figure A4.12.1. ^1H NMR spectrum of compound 4.17 (500 MHz, CDCl_3).....	304
Figure A4.13.1. ^1H NMR spectrum of compound 4.18 (500 MHz, CDCl_3).....	305
Figure A4.14.1. ^1H NMR spectrum of compound 4.19 (500 MHz, CDCl_3).....	306
Figure A4.15.1. ^{31}P NMR spectrum of phosphoramidite (<i>S</i>)- H^{NB} (121 MHz, CDCl_3)..	307
.....	

Figure A4.16.1. ^1H NMR spectrum of compound 4.21 (360 MHz, CDCl_3).....	308
Figure A4.16.2. ^{13}C NMR spectrum of compound 4.21 (90 MHz, CDCl_3).	309
Figure A4.16.3. IR spectrum of compound 4.21 (film).	310
Figure A4.17.1. ^1H NMR spectrum of compound 4.22 (360 MHz, CDCl_3).....	311
Figure A4.17.2. ^{13}C NMR spectrum of compound 4.22 (75 MHz, CDCl_3).	312
Figure A4.17.3. IR spectrum of compound 4.22 (film).	313
Figure A4.18.1. ^1H NMR spectrum of compound 4.25 (400 MHz, CDCl_3).....	314
Figure A4.18.2. ^{13}C NMR spectrum of compound 4.25 (125 MHz, CDCl_3).	315
Figure A4.18.3. IR spectrum of compound 4.25 (film).	316
Figure A4.19.1. ^1H NMR spectrum of compound 4.26 (400 MHz, CDCl_3).....	317
Figure A4.19.2. ^{13}C NMR spectrum of compound 4.26 (100 MHz, CDCl_3).	318
Figure A4.19.3. IR spectrum of compound 4.26 (film).	319
Figure A4.20.1. ^{31}P NMR spectrum of phosphoramidite (<i>S</i>)-P (162 MHz, CDCl_3).	320
.....	
Figure 5.1. Comparison of the T•A and C•G Watson-Crick base pairs with the $^{\text{Br}}\text{U}\cdot\text{A}$ and $^{\text{Br}}\text{C}\cdot\text{G}$ base pair used in this study.	325
Figure 5.2. Crude HPLC trace of the GNA oligonucleotide 3'- $\text{A}_{\text{PSe}}\text{TGC GCAT}$ -2'.	332
.....	
Figure 5.3. Crude HPLC trace of the GNA oligonucleotide 3'- $\text{CTC}_{\text{PSe}}\text{TAGAG}$ -2'.	334
.....	

Figure 5.4. Crude HPLC trace of GNA oligonucleotides functionalized using a phosphoroselenoate linkage demonstrating the differences in separation ability of the two diastereomers.	335
Figure 5.5. Packing contacts between duplexes in the crystal structure of 3'-G^{Br}CGCGC-2'.	342
Figure 5.6. Overall structures of Type N (a) and Type M (b)^x GNA oligonucleotide helices.	343
Figure 5.7. Structure of a single duplex of 3'-G^{Br}CGCGC-2' (a) and electron density of the 3'-terminal glycol nucleotides (b) at 1.5 sigma.	346
Figure 5.8. CD spectra of the self-complementary GNA duplexes 3'-GCGCGC-2' (MKS5) and 3'-G^{Br}CGCGC-2' (MKS132) at a duplex concentration of 20 μM.	349
Figure 5.9. Structure of a single duplex of 3'-CTC^{Br}UAGAG-2' (a) and electron density of the center two glycol nucleotides (b) at 1.5 sigma.	352
Figure 5.10. Packing contacts between two duplexes in the crystal structure of 3'-CTC^{Br}UAGAG-2'.	353
Figure 5.11. Anomalous difference density map at 5.0 sigma for peak data of 3'-CTC_{PSe}TAGAG-2' using phases from the brominated derivative (MKS80).	354
Figure 5.12. Superimposed models of: a) The structure of MKS80 (red) and MKS22 (blue).....	356
Figure 5.13. Structure of a single duplex of 3'-ATGCGCAT-2' missing one terminal adenine nucleotide (a) and electron density of the terminal two nucleotides (b) at 1.5 sigma.....	359

Figure A5.1.1. ^1H NMR spectrum of compound 5.2 (300 MHz, CDCl_3).....	377
Figure A5.1.2. ^{13}C NMR spectrum of compound 5.2 (75 MHz, CDCl_3).	378
Figure A5.1.3. IR spectrum of compound 5.2 (solid).....	379
Figure A5.2.1. ^{31}P NMR spectrum of phosphoramidite (<i>S</i>)- $^{\text{Br}}\text{U}$ (162 MHz, CDCl_3) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).	380
Figure A5.3.1. ^1H NMR spectrum of compound 5.4 (300 MHz, CDCl_3).....	381
Figure A5.3.2. ^{13}C NMR spectrum of compound 5.4 (75 MHz, CDCl_3).	382
Figure A5.3.3. IR spectrum of compound 5.4 (solid).....	383
Figure A5.4.1. ^1H NMR spectrum of compound 5.6 (300 MHz, CDCl_3).....	384
Figure A5.5.1. ^1H NMR spectrum of compound 5.7 (300 MHz, CDCl_3).....	385
Figure A5.6.1. ^{31}P NMR spectrum of phosphoramidite (<i>S</i>)- $^{\text{Br}}\text{C}$ (162 MHz, CDCl_3).....	386
Figure A5.7.1. ^1H NMR spectrum of compound 5.8 (300 MHz, CDCl_3).....	387
Figure A5.7.2. ^{13}C NMR spectrum of compound 5.8 (75 MHz, CDCl_3).	388
Figure A5.7.3. IR spectrum of compound 5.8 (solid).....	389
Figure A5.8.1. ^{31}P NMR spectrum of phosphoramidite (<i>S</i>)- $^{\text{Br}}\text{C}^*$ (162 MHz, CDCl_3) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).	390

List of Tables

Table 1.1. Thermal stabilities of DNA, RNA and GNA duplexes.....	10
Table A.2.1.1. Table of reagents for oligonucleotide synthesis.	128
Table A2.2.1. DNA oligonucleotide synthesis protocol: DNA (1 μM).	128

Table A2.3.1. Initial GNA oligonucleotide synthesis protocol with changes highlighted in red: 1000GNA (1 μM).	130
Table A2.4.1. Final GNA oligonucleotide synthesis protocol with changes highlighted in red: 1000GNA3 (1 μM).	133
Table 3.1. Thermal stabilities of duplexes of DNA, GNA, and DNA/GNA hybrids.	138
Table 3.2. Thermal stabilities of (<i>S</i>)-GNA and DNA duplexes.	141
Table 3.3. Thermal stabilities (T_m [$^{\circ}$C]) of matched and mismatched Watson-Crick base pairs in duplex GNA.	143
Table 3.4. Thermal and thermodynamic stability of (<i>S</i>)-GNA and DNA duplexes.	150
Table 3.5. Thermal and thermodynamic stabilities of (<i>S</i>)-GNA duplexes with overhanging nucleotides.	152
Table 3.6. Thermal and thermodynamic stabilities of DNA duplexes with overhanging nucleotides.	155
Table 3.7. Thermal and thermodynamic stabilities of (<i>S</i>)-GNA and DNA duplexes with overhanging nucleotides.	156
Table 3.8. Effects of intercalators on GNA and DNA duplexes.	167
Table 3.9. Thermal stabilities (T_m [$^{\circ}$C]) for the cross-pairing between strands of (<i>S</i>)-GNA, (<i>R</i>)-GNA, RNA, and DNA.	169
Table 3.10. Thermal stabilities (T_m [$^{\circ}$C]) of matched and mismatched base pairs in GNA/RNA heteroduplexes.	185

Table 4.1. Thermal stabilities of GNA duplexes in the presence of metal ions.	239
Table 4.2. Thermal stabilities of GNA duplexes in the presence of metal ions.	241
Table 4.3. Thermal stabilities of GNA duplexes in the presence of metal ions.	243
Table 4.4. GNA oligonucleotides for crystallography utilizing metal-mediated base pairing.	251
Table 4.5. Comparison of Average Helical Parameters for (S)-GNA, B-DNA, and A-DNA.	255
Table A4.1.1. Data collection, phasing, and refinement statistics of MKS64.	321
Table 5.1. Formulation of the Nucleic Acid Mini Screen (Hampton Res.)	338
Table 5.2. GNA oligos synthesized to solve the structure of 3'-GCCGCGC-2'	341
Table 5.3. Comparison of the average helical parameters for (S)-GNA, B-DNA, and A-DNA.	344
Table 5.4. Backbone torsional angles.	347
Table 5.5. Phosphate bond angles.	348
Table 5.6. GNA oligos synthesized to solve the structure of 3'-CTCTAGAG-2'	351
Table 5.7. GNA oligos synthesized to solve the structure of 3'-ATGCGCAT-2'	357
Table A5.1.1. Data collection statistics of unsolved structures.	391
Table A5.2.1. Data collection, phasing, and refinement statistics of MKS132.	392
Table A5.3.1. Data collection, phasing, and refinement statistics of MKS80.	393
Table A5.4.1. Data collection, phasing, and refinement statistics of MKS74-1.	394
Table A5.5.1. Data collection and refinement statistics of MKS22.	395
Table A5.6.1. Data collection and refinement statistics of MKS42.	396

Chapter 1

Introduction

Chapter 1.1. Background Information

Analogues of DNA and RNA have become extremely popular in the medicinal field for their application towards the inhibition of gene expression via the antisense or antigene approach.^{1,2} In principle, both of these approaches function via the simple rules of Watson-Crick base pairing introduced over 50 years ago with the first structure of a B-form duplex composed of DNA; those are adenine pairs with thymine, and guanine pairs with cytosine.³ The advantage of using this method over designing drugs that use small molecules to target protein inhibition is that only the sequence of the target DNA or RNA needs to be known. In the antisense approach, an oligonucleotide is introduced which is complementary to a mRNA of interest (Figure 1.1a).¹ Duplex formation between these two oligonucleotides thereby prevents transcription of this mRNA by direct physical blockage or by the recruitment of cellular enzymes that break down the complex. The antigene approach functions on the level of genomic, double-stranded DNA instead of the transcribed mRNA. Oligonucleotides are introduced which interfere with the transcription of the gene of interest by forming a parallel or antiparallel triple helix with the double-stranded DNA (Figure 1.1b).² One obvious advantage of the antigene approach over the antisense approach is that much less of the inhibitory oligonucleotide would be needed since there are only two copies of the gene in each cell versus possibly thousands of mRNA copies.

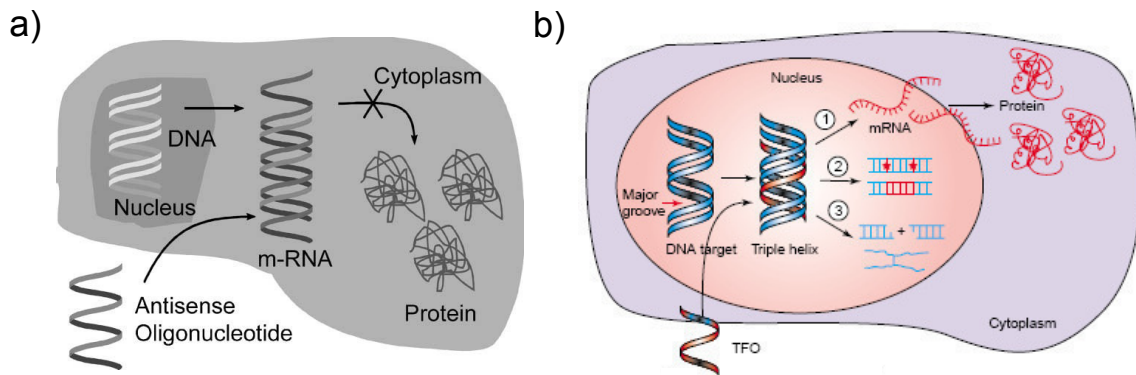


Figure 1.1. Schematic representation of oligonucleotide directed inhibition of gene expression: a) the antisense approach¹ and b) the antigene approach.²

Although natural oligonucleotides composed of DNA or RNA possess the ability to inhibit gene expression in this manner, they have one major drawback. Since both of these natural nucleotides are degraded biologically (and in the case of RNA, chemically), they are not ideal to develop as antisense or antigene drugs because of their low bioavailability.⁴ This has led researchers to investigate modifications of the DNA or RNA scaffold to produce nucleic acids analogs with increased affinities towards their targets and a greater stability which is usually accompanied by a lack of recognition by nucleases in the cellular environment. The three ways of altering nucleic acid structure are by modifying the nucleobases,^{5,6} the phosphodiester linkage between nucleosides,⁷⁻¹² or the sugar of the backbone.^{1,13,14} Although sugar alternatives are probably the most common modification of individual nucleotides (see Figure 1.2), there is still little understanding as to how structural changes of individual nucleotides will affect the overall structure when incorporated into oligonucleotides. Along these lines

Eschenmoser set out to investigate sugar alternatives as part of ongoing research into the etiology of nucleic acid structure.¹³ Although Eschenmoser was more interested in why nature chose a deoxyribose and ribose backbone for information storage and carriage, respectively, these studies pointed to the importance of conformational restriction of nucleosides to preorganize oligonucleotides for duplex formation, thereby leading to novel analogs, especially LNA¹⁵ with very promising RNA binding affinities.

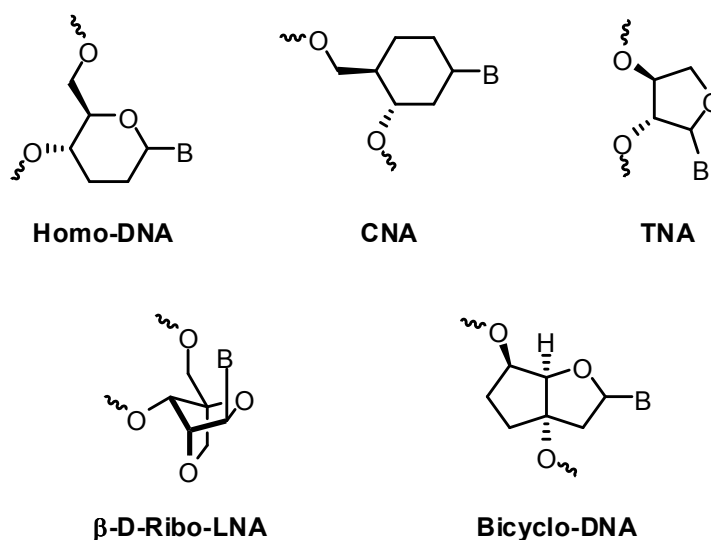


Figure 1.2. Several examples of sugar modified DNA/RNA analogs.

Not to be lost in the field of medicine is the application of oligonucleotide analogs towards the fields of structural DNA nanotechnology and nanoelectronics.^{16,17} DNA has become extremely interesting outside of the biological context since it is a molecule which has properties such as the high affinity and specificity of complementary sequences, a largely regular backbone regardless of sequence, and the ability to store

information by the arrangement of base pairs. Based on these principles, DNA has already been used to create higher order three-dimensional structures and arrays,^{16,17} nanomechanical devices,¹⁶⁻¹⁸ and DNA duplexes containing arrays of metal ions.^{19,20}

Chapter 1.2. Acyclic analogs of nucleic acids containing phosphodiester linkages

Acyclic glyceronucleotide containing oligos have been proposed as simpler precursors to RNA on the timescale of evolution since the corresponding nucleosides are presumably simpler to synthesize from simple precursors under prebiotic conditions than are those from the ribose and deoxyribose sugars.⁴ Furthermore, oligonucleotides constructed from these acyclic analogs may also possess greater stability towards nucleases and therefore have a greater bioavailability in a biological environment compared to using natural oligonucleotides which are degraded easily under these conditions. Based on these two proposals, researchers became highly interested in synthesizing acyclic (“flexible”) nucleoside analogs for the purpose of antisense therapy (see Figure 1.3). Several analogs of DNA and RNA in which an atom or bond has been removed from the (deoxy)ribose ring have been synthesized and tested for their hybridization properties with DNA.^{4,21,22} Other analogs containing dihydroxypropyl,²³ dihydroxybutyl,²⁴ and dihydroxypentyl²⁵ nucleosides have also been synthesized and subsequently tested for their hybridization and nuclease resistance properties. In each case, incorporation of one or several of these “flexible” nucleotides into a DNA duplex resulted in significant destabilizations compared to DNA nucleosides. On the other hand, the substitution of an acyclic nucleotide into an oligonucleotide composed of DNA generally conferred greater stability towards nuclease degradation compared to an oligonucleotide composed entirely of deoxyribonucleotides. Although these acyclic nucleosides were interesting based on their nuclease resistance properties, the

disappointing thermal stabilities were thought to be the result of the increase in the loss of entropy upon the transition from single oligonucleotides to a duplex in solution. Furthermore, it was proposed that oligonucleotides composed entirely of “flexible” nucleosides would not have the conformational preorganization required to form stable duplexes with natural oligonucleotides in solution; effectively diminishing the likelihood that these oligonucleotides could be interesting for antisense applications or as a possible precursor to natural oligonucleotides.^{4,23}

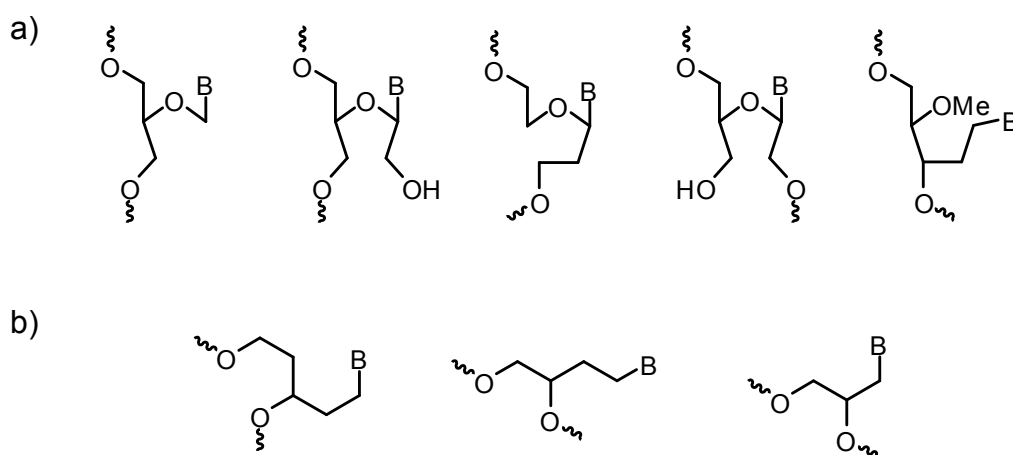


Figure 1.3. Structures of different acyclic nucleosides.²³ a) Constructed from deleting one or more atoms of the ribose sugar. b) Constructed from simple alkyl chains.

Chapter 1.3. Previous work with glycol nucleic acids (GNA)

Several years later, the Meggers group became interested in using an acyclic backbone to incorporate a metal-mediated base pair into the middle of a DNA duplex.²⁶ Although the earlier reports on acyclic nucleosides demonstrated large decreases in thermal stability in DNA duplexes, Meggers suggested that the use of a flexible linker might be overcome by extraordinary base-pairing strength. In this way, it was envisioned one could obtain structurally simplified, functional nucleotides that could be accessed in greater quantities than the corresponding deoxyribonucleotides. Surprisingly, a comparison of the 8-hydroxyquinoline nucleotides **HQ** and **C₃HQ** (Figure 1.4), containing the 2'-deoxyribose and dihydroxypropyl linkers, respectively, demonstrated that the Cu⁺² mediated homobase pair composed of the two flexible nucleotides resulted in a DNA duplex with greater thermal stability than that composed of the **HQ** nucleotides.²⁶ Meggers hypothesized that the greater stability could be attributed to less strain with the slimmer acyclic backbone (with better accommodation of the expanded C_{1'}-C_{1'} distance) versus the normal 2'-deoxyribose backbone.

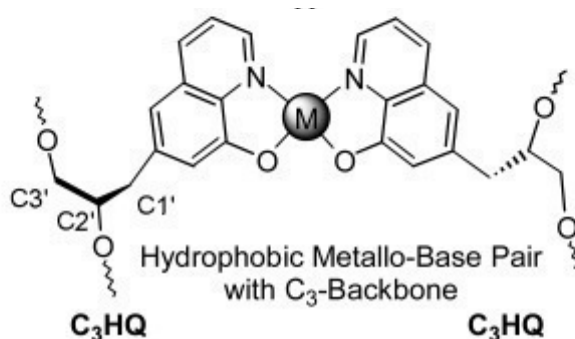


Figure 1.4. A completely artificial base pair in DNA.²⁶

A subsequent report by the Meggers group went one step further in synthesizing oligonucleotides composed entirely of the “flexible” dihydroxypropyl backbone. This was inspired by the previous work of Eschenmoser and others who demonstrated that stable duplexes with Watson-Crick base pairing could be supported by sugars different from those in DNA or RNA. Accordingly, the individual nucleoside phosphoramidites of both the (*S*) and (*R*)-enantiomers containing the adenine and thymine nucleobases were synthesized and successfully incorporated into oligonucleotides which were called glycol nucleic acids (GNA, Figure 1.5).²⁷

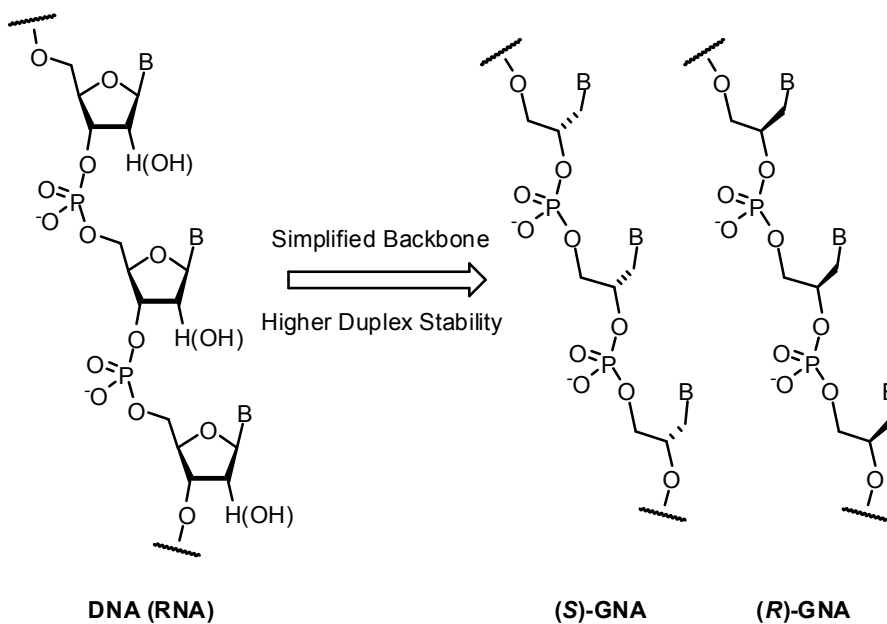


Figure 1.5. Comparison of the constitution of DNA and RNA with that of the two enantiomers of GNA.

Both temperature dependent UV-spectroscopy and CD measurements pointed to the formation of stable GNA duplexes composed of single strands pairing in a strictly antiparallel fashion. However, most surprising was that the thermal stability of the GNA duplex *exceeded* that of the analogous DNA and RNA duplexes by 22.5 and 22.0 °C, respectively (Table 1.1). This was astonishing considering all previous reports of oligonucleotides containing flexible nucleotides came to the conclusion that an acyclic backbone does not possess the required conformational preorganization required for duplex formation, although it was not eliminated as a possibility.⁴ Furthermore, it was shown that the (*S*)-enantiomer of GNA could form stable heteroduplexes with that of RNA, rendering GNA as a possible predecessor of RNA as a genetic material.²⁷ A later communication detailed the synthesis of the glycol nucleoside phosphoramidites of all five Watson-Crick nucleobases²⁸ (adenine, guanine, cytosine, thymine, and uracil) which allows one to access GNA duplexes composed of both Watson-Crick base pairs, and with superior stabilities to analogous DNA or RNA duplexes.

Table 1.1. Thermal stabilities of DNA, RNA and GNA duplexes.²⁷

System	Sequence	T _m (°C)
GNA Antiparallel	3'-TAAAATTTATATTATTAA-2' 2'-ATTTTAAATATAATAATT-3'	63
GNA Parallel	3'-TAAAATTTATATTATTAA-2' 3'-ATTTTAAATATAATAATT-2'	No T _m
DNA	5'-TAAAATTTATATTATTAA-3' 3'-ATTTTAAATATAATAATT-5'	40.5
RNA	5'-TAAAATTTATATTATTAA-3' 3'-ATTTTAAATATAATAATT-5'	42.5

With these results in mind, a study was proposed whose main goal was the elucidation of the properties of GNA. Since all previous reports detailed the disappointing stabilities of “flexible” acyclic nucleosides, we wanted to understand how a backbone composed of an acyclic, three carbon, phosphodiester backbone could support the formation of antiparallel duplexes with superior thermal stabilities than corresponding duplexes of DNA or RNA. Furthermore, it was desirable to further investigate the crosspairing ability of (*S*)-GNA and RNA. As mentioned previously, this is a focal point of research in medicinal chemistry with nucleic acid analogs; mainly designing a nucleic acid analog with specific base pairing properties that can evade nuclease degradation within the cell. Finally, based on the synthetic accessibility of glycol nucleosides, we were interested in the synthesis and evaluation of novel base-pairs as a means to create functionalized duplexes composed entirely of GNA.

Chapter 1.4. References

- (1) Leumann, C. J. *Bioorganic & Medicinal Chemistry* **2002**, *10*, 841.
- (2) Buchini, S.; Leumann, C. J. *Current Opinion in Chemical Biology* **2003**, *7*, 717.
- (3) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.
- (4) Schneider, K. C.; Benner, S. A. *Journal of the American Chemical Society* **1990**, *112*, 453-455.
- (5) Henry, A. A.; Romesberg, F. E. *Current Opinion in Chemical Biology* **2003**, *7*, 727.
- (6) Clever, G. H.; Kaul, C.; Carell, T. *Angewandte Chemie International Edition* **2007**, *46*, 6226-6236.
- (7) Mori, K.; Boiziau, C.; Cazenave, C.; Matsukura, M.; Subasinghe, C.; Cohen, J. S.; Broder, S.; Toulme, J. J.; Stein, C. A. *Nucl. Acids Res.* **1989**, *17*, 8207-8219.
- (8) Uhlmann, E.; Peyman, A. *Chemical Reviews* **1990**, *90*, 543-584.
- (9) Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 6123.
- (10) Marshall, W. S.; Caruthers, M. H. *Science* **1993**, *259*, 1564-1570.
- (11) Ulf, D. *Angewandte Chemie International Edition* **1997**, *36*, 1886-1889.
- (12) Nielsen, P. E. *Accounts of Chemical Research* **1999**, *32*, 624-630.
- (13) Eschenmoser, A. *Science* **1999**, *284*, 2118.
- (14) Herdewijn, P. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1999**, *1489*, 167.
- (15) Braasch, D. A.; Corey, D. R. *Chemistry & Biology* **2001**, *8*, 1.

- (16) Seeman, N. C. *Chemistry & Biology* **2003**, *10*, 1151.
- (17) Seeman, N. *Molecular Biotechnology* **2007**, *37*, 246.
- (18) Modi, S.; Swetha, M. G.; Goswami, D.; Gupta, G. D.; Mayor, S.; Krishnan, Y. *Nat Nano* **2009**, *4*, 325.
- (19) Tanaka, K.; Clever, G. H.; Takezawa, Y.; Yamada, Y.; Kaul, C.; Shionoya, M.; Carell, T. *Nat Nano* **2006**, *1*, 190.
- (20) Clever, G. H.; Carell, T. *Angewandte Chemie International Edition* **2007**, *46*, 250-253.
- (21) Azymah, M.; Chavis, C.; Lucas, M.; Morvan, F. o.; Imbach, J.-L. *Nucleosides, Nucleotides and Nucleic Acids* **1992**, *11*, 1241 - 1255.
- (22) Nielsen, P.; Kirpekar, F.; Wengel, J. *Nucl. Acids Res.* **1994**, *22*, 703-710.
- (23) Nielsen, P.; Dreißøe, L. H.; Wengel, J. *Bioorganic & Medicinal Chemistry* **1995**, *3*, 19.
- (24) Augustyns, K.; Van Aerschot, A.; Van Schepdael, A.; Urbanke, C.; Herdewijn, P. *Nucl. Acids Res.* **1991**, *19*, 2587-2593.
- (25) Vandendriessche, F.; Augustyns, K.; Van Aerschot, A.; Busson, R.; Hoogmartens, J.; Herdewijn, P. *Tetrahedron* **1993**, *49*, 7223.
- (26) Zhang, L.; Meggers, E. *Journal of the American Chemical Society* **2005**, *127*, 74-75.
- (27) Zhang, L.; Peritz, A.; Meggers, E. *Journal of the American Chemical Society* **2005**, *127*, 4174-4175.
- (28) Zhang, L.; Peritz, A.; Carroll, P. J.; Meggers, E. *Synthesis* **2006**, 645-653.

Chapter 2

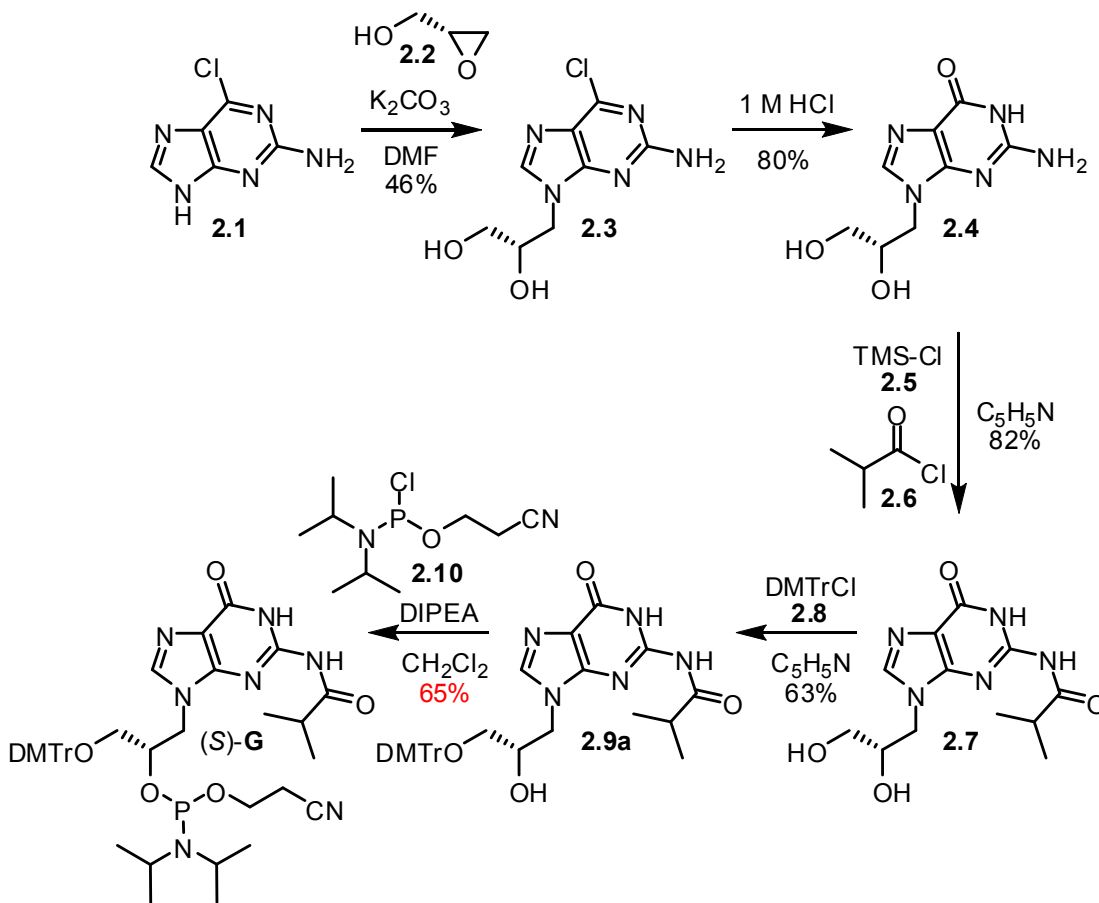
Synthesis of glycol nucleoside phosphoramidites and their incorporation into GNA oligonucleotides

Chapter 2.1. Initial synthesis of glycol nucleoside phosphoramidites

The synthesis of glycol nucleoside phosphoramidites¹ is a straight-forward process starting from enantiopure (*R*)- or (*S*)-glycidol. As previously reported, nucleophilic ring-opening of enantiopure glycidol produces the glycol nucleoside derivatives with good regioselectivity for the *N*-9 and *N*-1 regioisomers of purines and pyrimidines, respectively.¹ The initial yields ranged from 39-59%, with higher yields generally obtained for the pyrimidines. The pyrimidine nucleosides could then be converted to the nucleoside phosphoramidites directly while the purine nucleosides required 1-3 additional protection steps before conversion to the phosphoramidites. The synthesis of these glycol nucleoside phosphoramidites represents a large savings of both time and material as they proceed in fewer steps and without the separation of the two anomeric forms as is required for the synthesis of DNA phosphoramidites.²

For the synthesis of guanine phosphoramidite (*S*)-**G**, the first step is the ring opening of (*R*)-glycidol (**2.2**) using 2-amino-6-chloropurine (**2.1**) with 0.17 equivalents of potassium carbonate to produce compound **2.3** (Scheme 2.1) in 46% yield. The somewhat low yield is due to the observed formation of both the *N*-7 and *N*-9 regioisomers in this reaction. Compound **2.3** is then hydrolyzed using 1 M HCl to give compound **2.4** in 80% yield. Transient protection of the exocyclic amino group with trimethylsilyl chloride (**2.5**) and further reaction with isobutyryl chloride (**2.6**) in pyridine yields the protected diol **2.7** in 82% yield. This can then be further selectively protected at the 3'-hydroxyl group with dimethoxytriyl chloride (**2.8**) in pyridine to yield

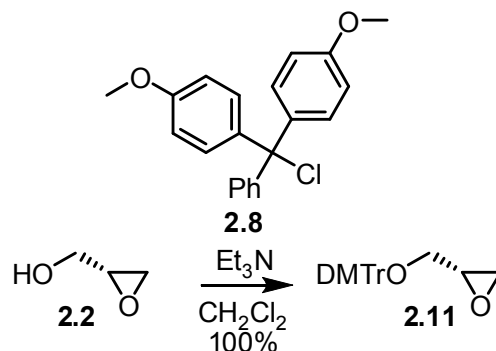
compound **2.9a** in 63% yield. Compound **2.9a** can then be converted to phosphoramidite (*S*)-**G** using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) in the presence of excess *N,N*-diisopropylethylamine (product was usually not completely pure as determined by ³¹P NMR). Unfortunately, the product (*S*)-**G** is not stable to either silica gel or aluminum oxide flash chromatography and must be purified via a precipitation protocol.¹ The product is precipitated from rapidly stirred pentane by the slow addition of a solution of the crude product in methylene chloride, however, most times it was not as pure as one ideally would desire for oligonucleotide synthesis. Furthermore, the product is unstable and slowly decomposes even when stored under argon at -20 °C.



Scheme 2.1. Synthesis of (*S*)-**G** phosphoramidite.

For the synthesis of the adenine, cytosine, and thymine glycol nucleoside phosphoramidites, nucleophilic ring-opening is not performed on (*R*)-glycidol (**2.2**). Instead, the 3'-hydroxyl group of (*R*)-glycidol (**2.2**) is first protected using 4,4'-dimethoxytrityl chloride (**2.8**) and excess triethylamine in methylene chloride to yield compound **2.11** in quantitative yield (Scheme 2.2). The advantage of using compound **2.11** in the subsequent ring-opening reactions is that the desired product is more non-

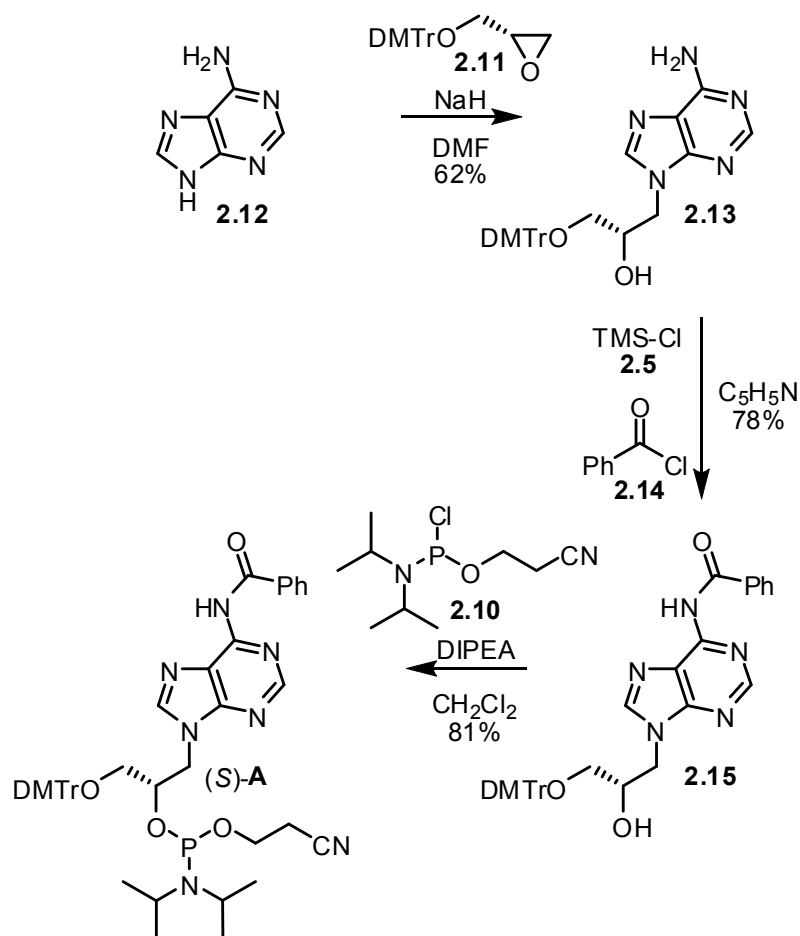
polar as opposed to its non-protected counterpart (**2.2**), thereby facilitating product purification.



Scheme 2.2. Synthesis of (*S*)-glycidyl 4,4'-dimethoxytrityl ether (**2.11**).

The synthesis of the adenine phosphoramidite (*S*)-**A** can be accomplished in four steps starting with ring opening of compound **2.11** using adenine (**2.12**) and 0.2 equivalents of sodium hydride to produce compound **2.13** in 62% yield (Scheme 2.3). The exocyclic amino group is then protected using a transient protection procedure in which the 2'-hydroxyl group of compound **2.13** is first protected with trimethylsilyl chloride (**2.5**) and then reacted with benzoyl chloride (**2.14**). Afterwards, during the work-up of the reaction, the 2'-trimethylsilyloxy group is cleaved by the addition of aqueous ammonia to produce compound **2.15** in 78% yield. It should be noted that yields for this protection step were highly variable (from 40-78%) based on how well the removal of the “transient” TMS group was accomplished during ammonia treatment. It was normally observed after ammonia treatment that a significant spot on TLC remained corresponding to the TMS protected product indicating this intermediate to be more stable than expected. An improved procedure which utilizes TBAF to remove the TMS

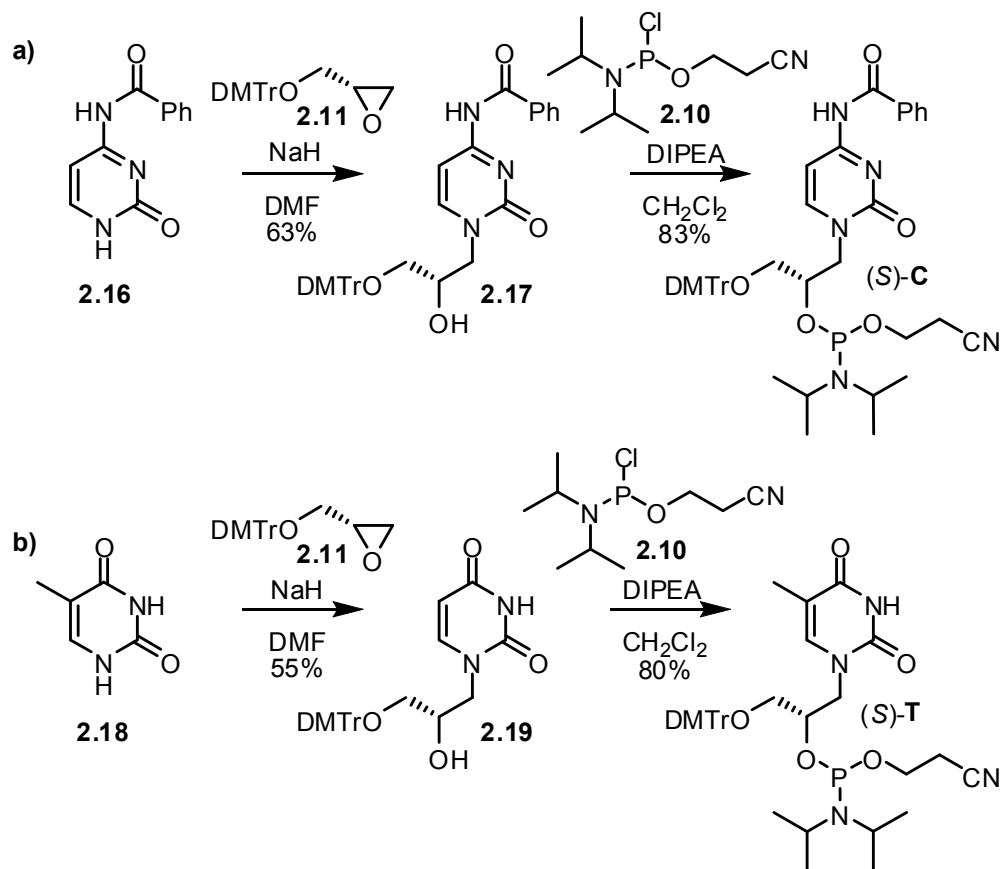
group was developed with much more consistent results. After treatment with aqueous ammonia and concentration of the reaction mixture, the crude product was redissolved in THF and treated with two equivalents of TBAF to produce compound **2.15** with consistently higher yields. Compound **2.15** is then converted to phosphoramidite (*S*)-**A** in 81% yield using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess *N,N*-diisopropylethylamine.



Scheme 2.3. Synthesis of (*S*)-**A** phosphoramidite.

The phosphoramidite (*S*)-**C** is produced starting again with the ring opening of compound **2.11** with *N*-4-benzoylcytosine (**2.16**) using 0.2 equivalents of sodium hydride to produce compound **2.17** in 63% yield (Scheme 2.4). Since the exocyclic amine of compound **2.17** is already protected, it can be directly converted to the phosphoramidite (*S*)-**C** using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess *N,N*-diisopropylethylamine in 83% yield. Similarly, the phosphoramidite (*S*)-**T** is synthesized starting with ring opening of compound **2.11** with thymine (**2.18**) using 0.2 equivalents of sodium hydride to produce compound **2.19** in 55% yield (Scheme 2.4). Again, since no further protection is required, compound **2.19** is then converted to the phosphoramidite (*S*)-**T** using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite and excess *N,N*-diisopropylethylamine in 80% yield.

The synthesis of the four glycol nucleoside phosphoramidites (*S*)-**G**, (*S*)-**A**, (*S*)-**C**, and (*S*)-**T** proceed in five, four, three, and three steps, respectively. Overall yields are 39% for (*S*)-**A**, 12% for (*S*)-**G**, 52% for (*S*)-**C** and 44% for (*S*)-**T** compared to the published overall yields¹ of 26%, 8%, 39%, and 38%, respectively. The (*R*)-enantiomer of each of these phosphoramidites is also easily accessible simply by starting with (*S*)-glycidol during the ring-opening step.¹

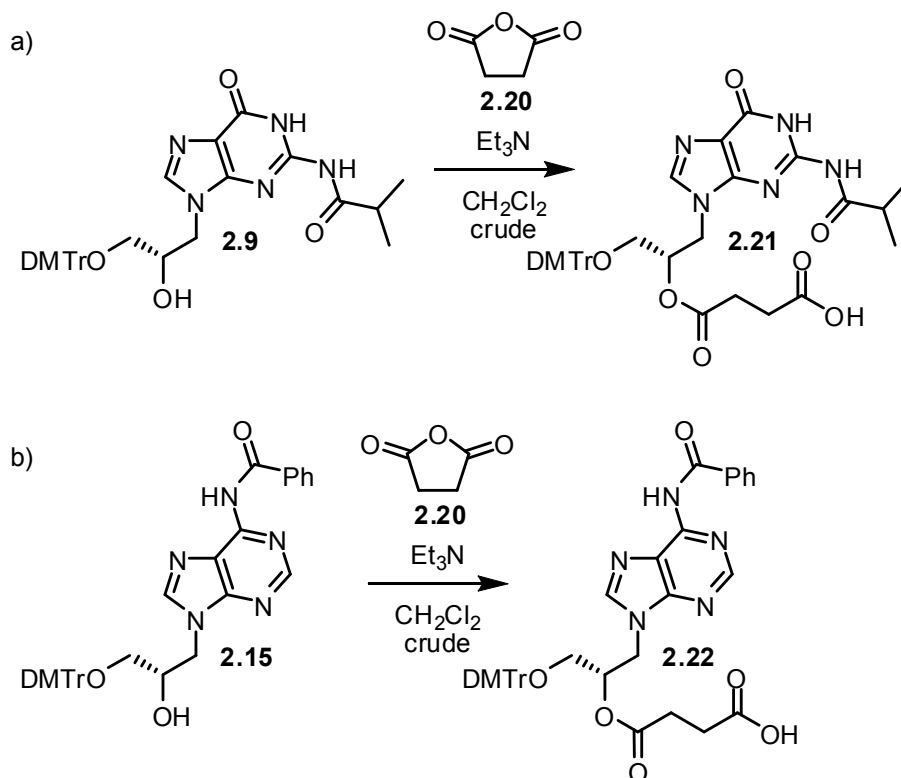


Chapter 2.2. Initial protocol for GNA oligonucleotide synthesis and purification

The synthesis and purification of phosphoramidites was followed by repeated coevaporation with toluene to remove trace amounts of water. Afterwards, the phosphoramidite was left under high vacuum overnight and then repeatedly dissolved and repeatedly evaporated with methylene chloride to remove traces of toluene. Again, the phosphoramidite was left overnight under high vacuum. The phosphoramidites could then be transferred as a solid to separate vials and stored under nitrogen or argon at -20 °C. All phosphoramidites were stable stored under these conditions except for the (*S*)-**G** phosphoramidite, as mentioned previously. Generally, on the evening prior to oligonucleotide synthesis the phosphoramidites [(*S*)-**A**, (*S*)-**G**, (*S*)-**C**, and (*S*)-**T**] were then transferred to flasks and left under high vacuum overnight to ensure dryness.

Solid supports were synthesized from the dimethoxytrityl protected nucleosides as previously reported.³ Briefly, nucleosides **2.9**, **2.15**, **2.17**, and **2.19** were functionalized using succinic anhydride (**2.20**) in the presence of triethylamine to produce compounds **2.21-2.24** (Scheme 2.5 and 2.6). These succinimidyl esters are then linked via an amide linkage to long chain alkylamine controlled pore glass (Scheme 2.7) using 1-hydroxybenzotriazole (**2.25**) and *N,N'*-diisopropylcarbodiimide (**2.26**) to provide the solid support linked nucleosides. After capping of unreacted hydroxyl groups using acetic anhydride, the solid support could be washed, dried, and used for solid phase synthesis of GNA oligonucleotides. Loadings were generally higher than that of normal DNA solid supports and ranged from 55-70 μ moles of nucleoside per gram of solid support.

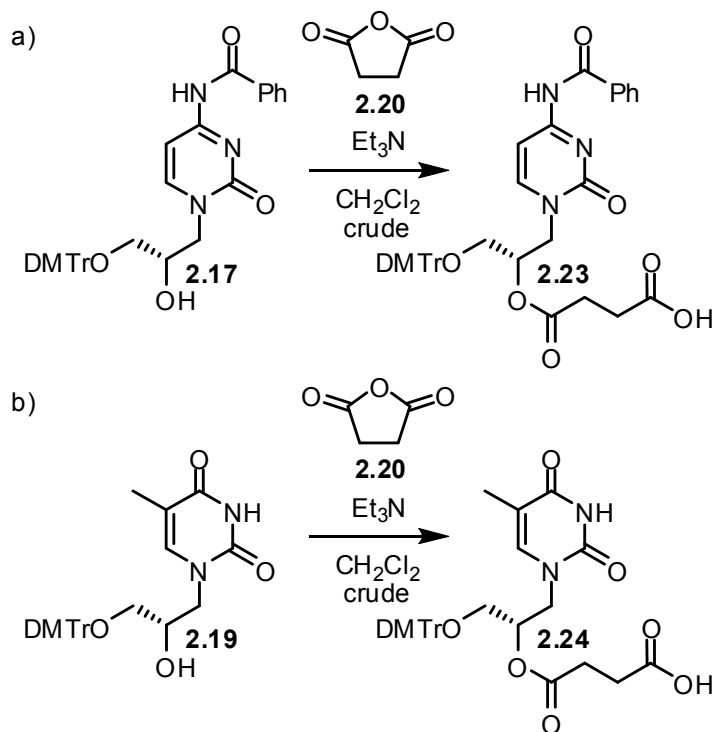
Loadings were measured by the absorbance of the released trityl cation³ when dissolving 3-5 mg of the solid support in 10% dichloroacetic acid in methylene chloride.



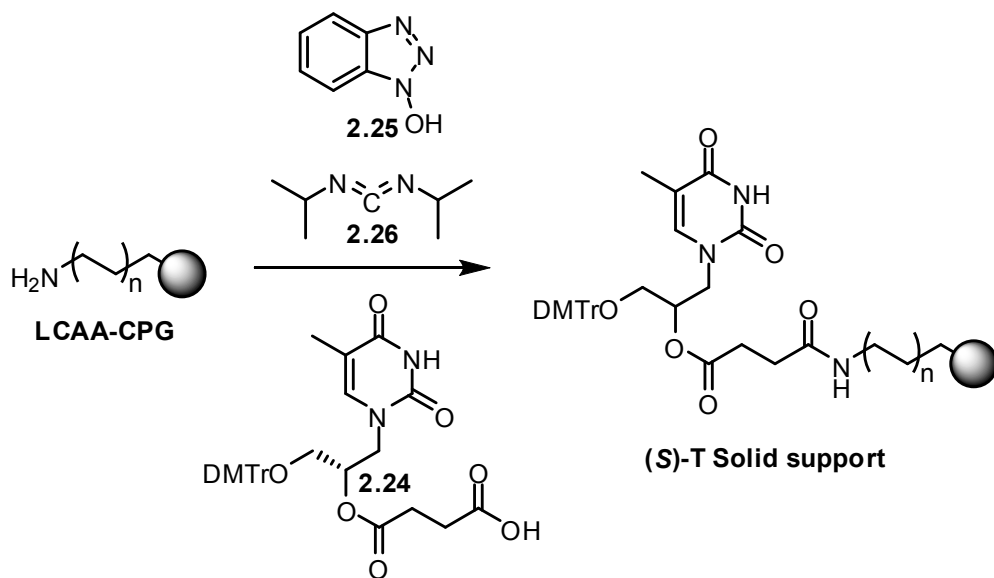
Scheme 2.5. Synthesis of succinyl esters from: a) guanine glycol nucleoside, b) adenine glycol nucleoside.

The next morning, the phosphoramidites were dissolved using anhydrous acetonitrile to a final concentration of 100 mM. Initial syntheses were attempted using 50 mM concentrations of phosphoramidites, however, much better coupling yields, and therefore oligonucleotides, were obtained when the concentration was doubled to 100 mM. These solutions were transferred to oven dried vials specially designed for use with the automated oligonucleotide synthesizer.⁴ Care was taken to avoid exposure to air

during all manipulations. Solid supports were weighed out and transferred to polypropylene synthesis columns sealed with column frits (Glen Research). Prior to synthesis, the solid support was dried for five minutes directly on the synthesizer using argon.



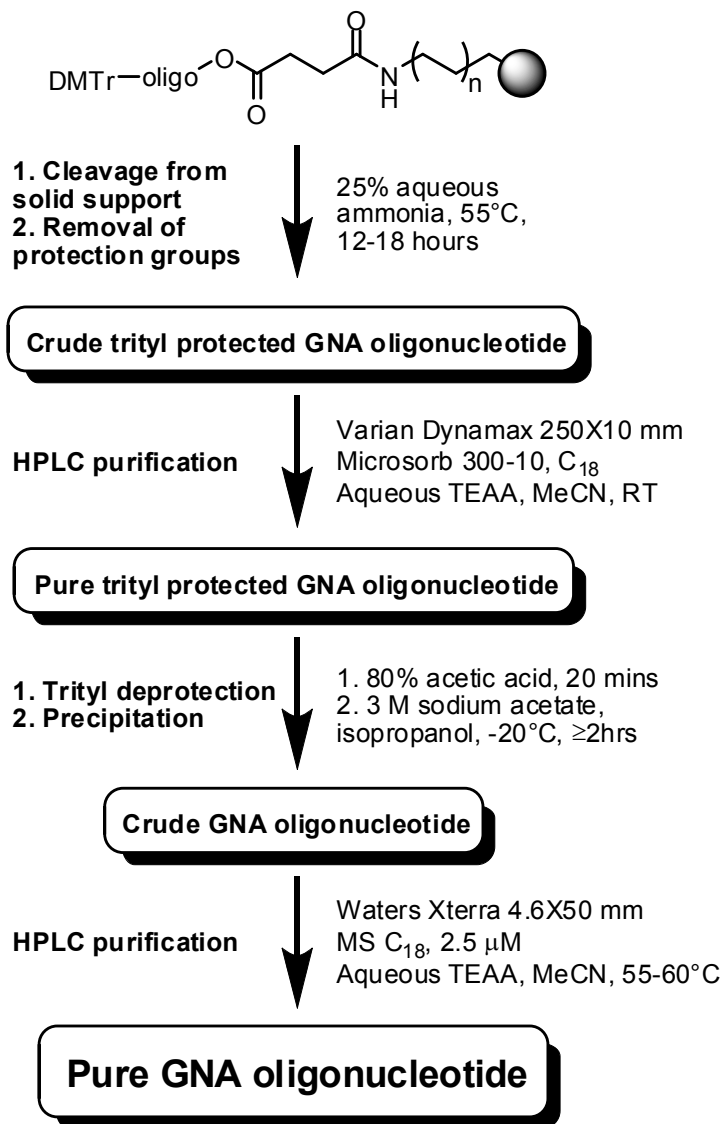
Scheme 2.6. Synthesis of succinyl esters from: a) cytosine glycol nucleoside, b) thymine glycol nucleoside.



Scheme 2.7. General scheme for the synthesis of glycol nucleoside functionalized solid supports using compound **2.24** as an example.

Oligonucleotides were synthesized on an Applied Biosystems ABI 394 Automated DNA/RNA synthesizer on a 0.2 or 1.0 micromole scale; however, best purities and yields were obtained when using a 1.0 micromole scale. Only minor changes were needed from standard DNA synthetic cycles⁵ (outlined specifically in the appendix to chapter 2), cycle: 1000GNA, steps 45-47, 53, 60). Most significantly, the coupling times were increased from 25 seconds for DNA phosphoramidites to 180 seconds from GNA phosphoramidites. Best results were obtained when the last 4,4'-dimethoxytrityl group was left attached to the 3'-hydroxyl group after synthesis as a purification handle. The solid support containing oligonucleotide was dried under argon on the synthesizer and then transferred to a screw cap vial capable of withstanding increased pressure during the deprotection step. Deprotection and cleavage from the solid support was carried out using 25% NH_4OH at 55-60 °C for 12-18 hours (overnight).

After cooling and concentration, GNA oligonucleotides could then be purified by reverse phase HPLC (column) using the 3'-dimethoxytrityloxy group as a purification handle (Scheme 2.8). This allows for the separation of shorter, capped "failures sequences" from the full length product. A representative "Trityl-ON" HPLC trace is shown below (Figure 2.1a) in which the main peak at 7.2 mins represents the full length, DMTr-protected product. The resulting fractions are combined and the 4,4'-dimethoxytrityl group cleaved by dissolving the DMTr-oligo in 80% acetic acid (200 μ L) for 20 minutes. Afterwards, an equal volume of 3M sodium acetate was added and the oligo precipitated by the addition of isopropanol (1.1 mL). The solution was cooled to -20 °C for at least two hours and then centrifuged at 14000 rpm at 0-4 °C for 10 minutes. After decanting the liquid, the pellet was resuspended in isopropanol (1.5 mL) and centrifuged once more. The resulting pellet was finally dried after decanting the isopropanol. Subsequent purification was performed using reverse phase HPLC (Waters Xterra) to separate the desired GNA oligo from both failure sequences and post-synthetic modifications (see figure 2.1b for representative HPLC trace of crude "Trityl-OFF" oligo). Oligonucleotides were at least 98% pure, as determined by HPLC (see representative trace of a pure GNA oligonucleotide in Figure 2.2).



Scheme 2.8. Overview of the initial procedure for GNA oligonucleotide purification.

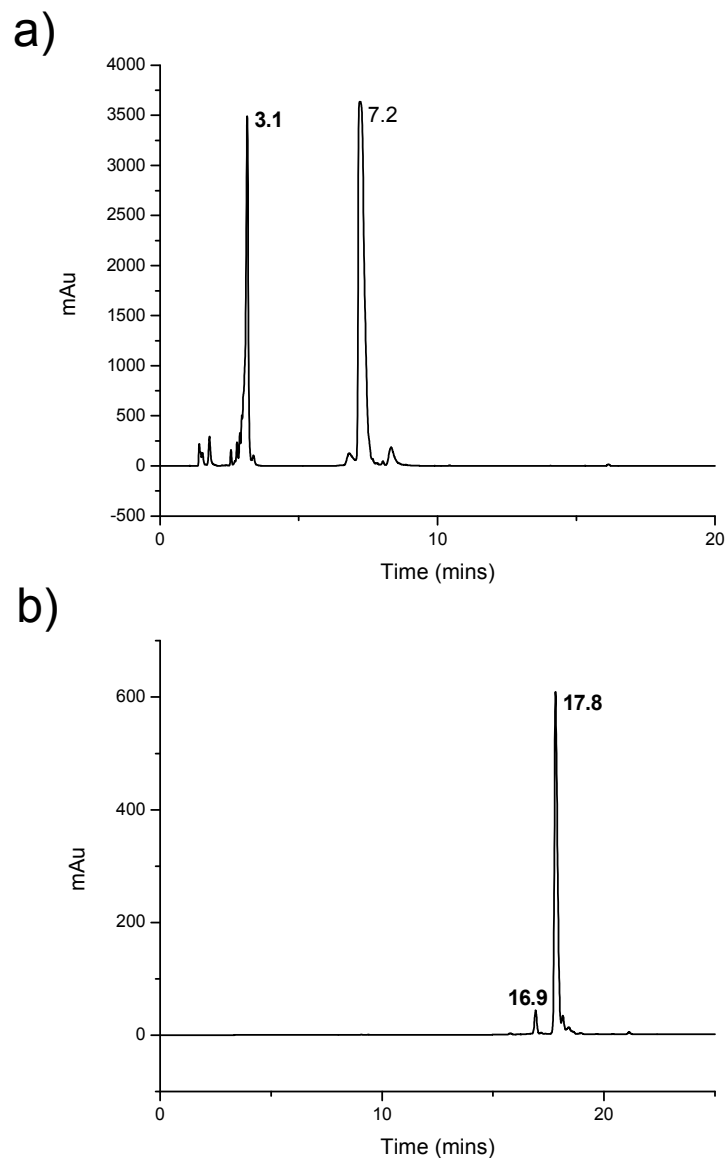


Figure 2.1. HPLC traces of the GNA oligonucleotide 3'-TTTTTTTTTT-2'. a) Crude “Trityl-ON” HPLC trace in which the peak at 7.2 minutes represents the dimethoxytrityl protected full length product. The crude oligo was eluted using a Microsorb 300-10 C₁₈ column with a linear gradient (flow = 3.0 mL/min) from 5-80% acetonitrile in 20 minutes and 95-20% aqueous tritethylammonium acetate buffer (50 mM, pH=7.0). b) Crude “Trityl-OFF” HPLC trace in which the peak at 17.8 minutes represents the full length product. The crude oligo was eluted using a Waters Xterra column (MS C₁₈, 4.6 x 50 mm, 2.5 μM particle size) at 60 °C with a linear gradient (flow = 1.0 mL/min) from 3-13% acetonitrile in 30 minutes and 97-87% aqueous tritethylammonium acetate buffer (50 mM, pH=7.0).

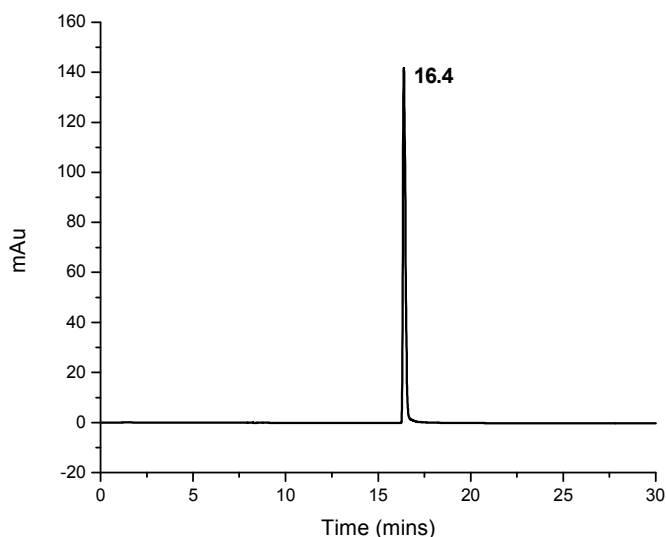


Figure 2.2. HPLC trace of the purified GNA oligo 3'-CACATTATTGTTGTA-2'. The oligo was eluted using a Waters Xterra column (MS C₁₈, 4.6 x 50 mm, 2.5 μ M particle size) at 60 °C with a linear gradient (flow = 1.0 mL/min) from 3-13% acetonitrile in 30 minutes and 97-87% aqueous tritethylammonium acetate buffer (50 mM, pH=7.0).

MALDI was used as a means to confirm the identity of synthesized GNA oligonucleotides. Samples were prepared at a concentration of approximately 10 μ M. One microliter of a saturated solution of trihydroxyacetophenone in 50% aqueous acetonitrile was mixed with one microliter of 100 mM ammonium tartrate and then one microliter of the oligo sample. The dried sample was then analyzed in the negative mode for detection of the sample mass. Best results were obtained when the oligo solution was desalted using DOWEX 50WX8 cation exchange beads⁶ (hydrogen form) prior to mixing with the matrix and ammonium tartrate solutions. A representative MALDI spectrum is shown in Figure 2.3. After identification, extinction coefficients of pure samples were calculated from deoxynucleotide increments.

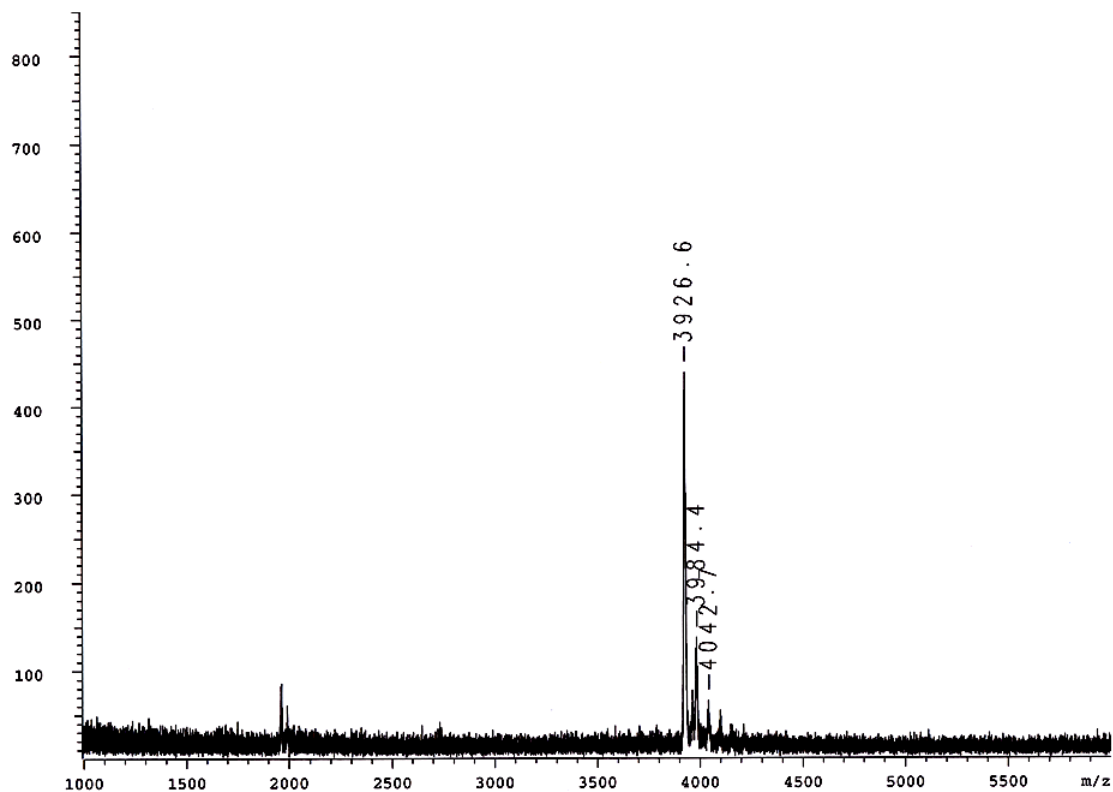


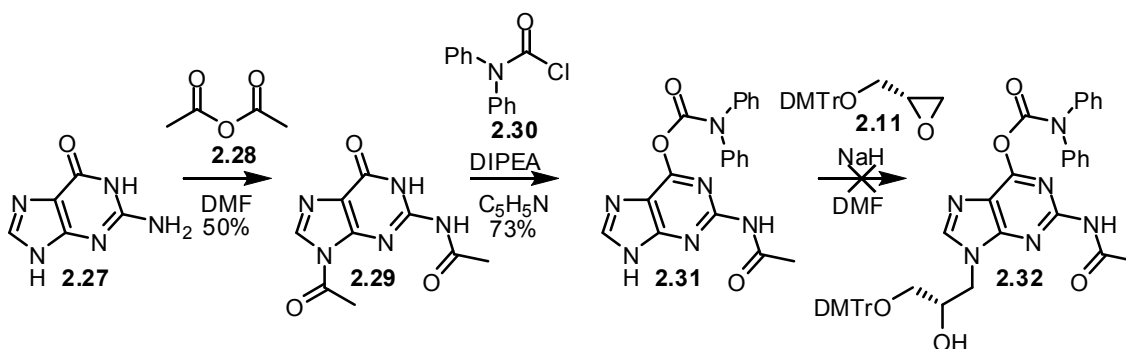
Figure 2.3. MALDI spectrum of the purified GNA oligonucleotide 3'-CACATTATTGTTGTA-2' (calc for $C_{118}H_{157}N_{50}O_{76}P_{14}$ $[M-H]^-$ 3925.5). The sample ($\sim 10 \mu\text{M}$ in water) was analyzed in negative mode after desalting over DOWEX beads.

Chapter 2.3. Development of new routes towards (*S*)-**G** phosphoramidite

The biggest initial problem with the synthesis of glycol nucleoside phosphoramidites was the lack of stability of the (*S*)-**G** phosphoramidite. It is known for DNA that the *N*-2-isobutyryl protected guanosine phosphoramidite is slightly unstable and prone to oxidation, even when one uses extreme cautionary measures to exclude air and water during its manipulation.⁷ This phenomenon is apparently more extreme for the (*S*)-**G** phosphoramidite in which the product seems to have a half life of approximately only one month when stored at -20 °C under argon. Therefore, it was desirable to synthesize a more stable derivative of (*S*)-**G** which could be purified more easily and produce a cleaner product compatible with longer storage times.

Initially, it was desired to develop a new synthetic route towards (*S*)-**G** that was simpler and more economical. Previous experience had shown difficulties in the conversion of compound **2.3** to **2.4** with the main problem being salt-contaminated product that was hard to purify. Moreover, the presence of salt in the product seemed to hinder the subsequent conversion to compound **2.9a** when carried forward without further purification (see Scheme 2.1). Therefore, it was envisioned that one could use guanine, or some derivative thereof, to perform the ring opening reaction as a more direct route towards the product (*S*)-**G**. Previous experience in our lab had shown that ring opening of compound **2.11** with unprotected guanine (**2.27**) does not result in the desired product. Attempts thereafter involved the use of a protected derivative of guanine (scheme 2.9), compound **2.29**, which was synthesized by first reacting guanine (**2.27**) with acetic anhydride (**2.28**) to produce compound **2.29** in 50% yield.⁸ This could then be converted

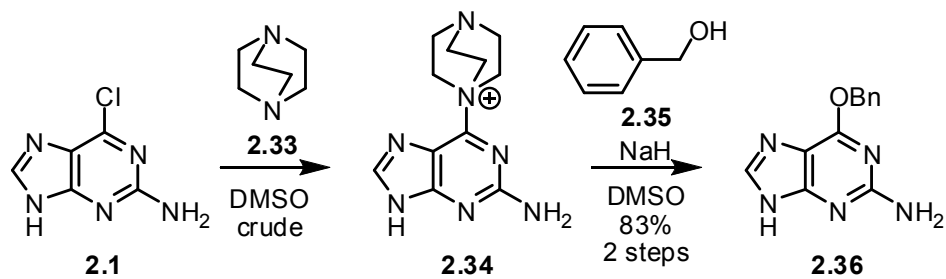
to compound **2.31** by the reaction with diphenylcarbamoyl chloride (**2.30**) and *N,N*-diisopropylethylamine in 73% yield. Unfortunately, an attempt at the ring opening of compound **2.11** with compound **2.31** and 0.2 equivalents of sodium hydride did not produce the desired product **2.32**.



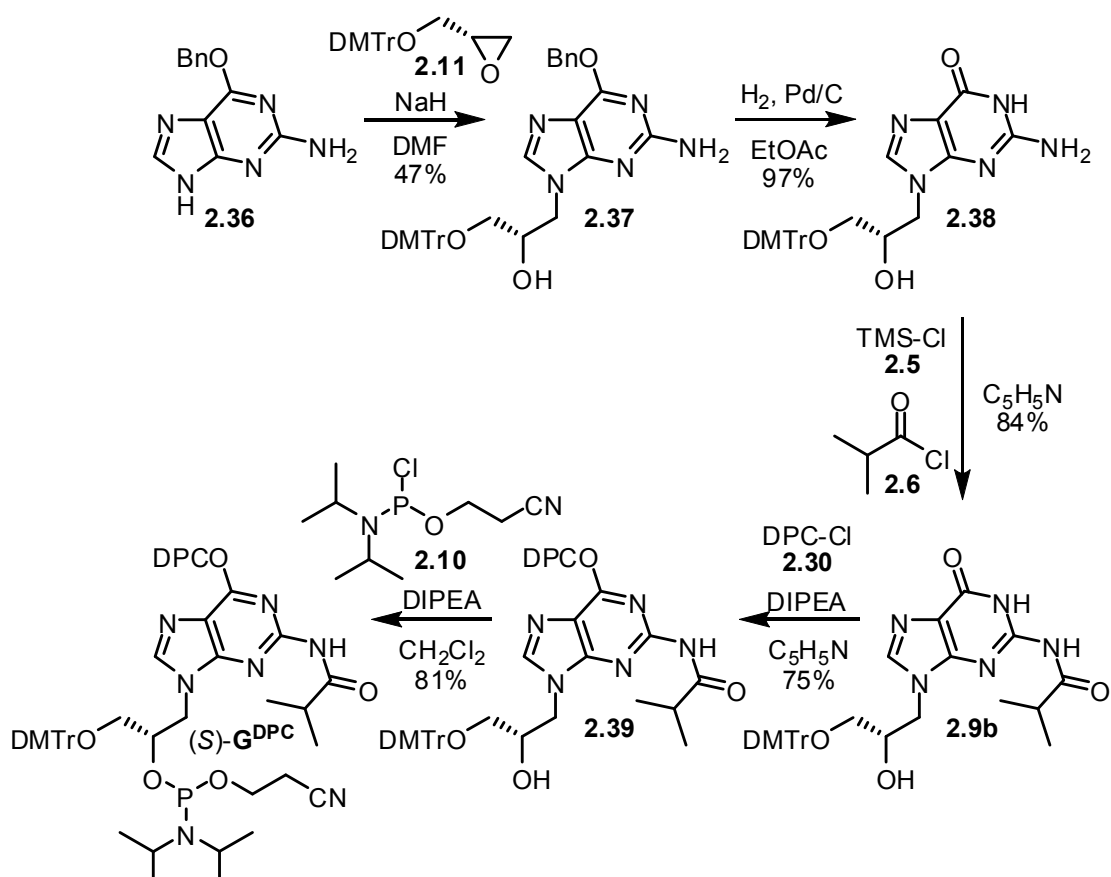
Scheme 2.9. Attempted ring-opening using protected guanine derivative compound **2.31**.

The next attempt involved the use of *O*-6-benzyloxyguanine (**2.36**) for the ring opening reaction of **2.11**. The synthesis of compound **2.36** starts with the reaction of 2-amino-6-chloropurine (**2.1**) with 1,4-diazabicyclo[2.2.2]octane (**2.33**) to produce the quaternary amine compound **2.34**. Afterwards, benzyl alcohol (**2.35**) is deprotonated using sodium hydride and reacted with compound **2.34** to produce *O*-6-benzyloxyguanine (**2.36**) in 83% over two steps⁹ (Scheme 2.10). Afterwards, *O*-6-benzyloxyguanine (**2.36**) and 0.2 equivalents of sodium hydride were used successfully in the ring opening of compound **2.11** to afford compound **2.37** in 47% yield (Scheme 2.11). The benzyl group was subsequently removed using catalytic hydrogenation to produce compound **2.38** in 97% yield. Reaction of compound **2.38** with isobutyryl chloride (**2.6**) in pyridine after transient protection with trimethylsilyl chloride (**2.5**) produces compound **2.9b** in 84%

yield. Again, as observed for the formation of compound **2.15**, best yields were obtained when two equivalents of TBAF were added to the product of this reaction after redissolving in THF to completely remove the “transient” trimethylsilyl protection group. This route starting with the ring opening of compound **2.11** with *O*-6-benzyloxyguanine (**2.36**) proceeds to the formation of compound **2.9b** in six steps with 32% overall yield versus the route in Scheme 2.1 to compound **2.9a** which requires four steps with 17% overall yield. It should be noted that *O*-6-benzyloxyguanine (**2.36**) is a commercially available compound, but the decision was made to synthesize compound **2.36** based on cost and availability. Compound **2.9b** could then be further protected using diphenylcarbonyl chloride (**2.30**) and *N,N*-diisopropylethylamine to produce compound **2.39** in 75% yield. Diphenylcarbonyl protection was chosen to further protect the guanine derivative **2.9b** in the *O*-6 position based on previous reports of increased stability of similar derivatives.¹⁰⁻¹¹ Subsequent conversion to the phosphoramidite (*S*)-**G^{DPC}** proceeds using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess *N,N*-diisopropylethylamine in methylene chloride in 81% yield. Gratifyingly, (*S*)-**G^{DPC}** was stable to flash chromatography over silica gel and a more pure phosphoramidite could be obtained. Moreover, the phosphoramidite (*S*)-**G^{DPC}** was stable when stored under argon at -20 °C. This new synthesis has an overall yield of 23% in six steps versus 12% in five steps for the previous synthesis of (*S*)-**G** in scheme 2.1.

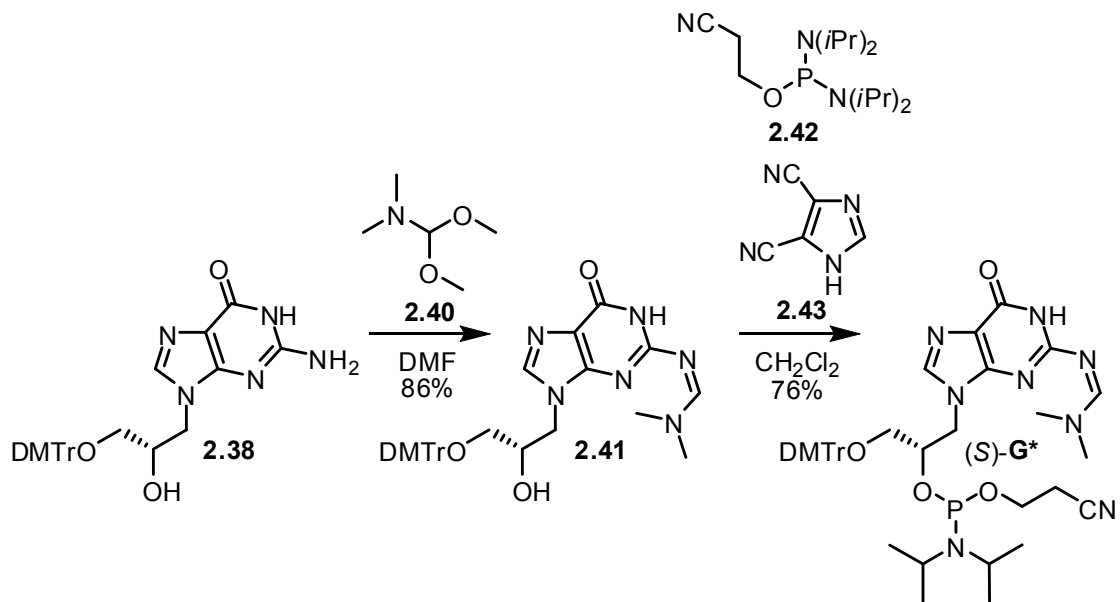


Scheme 2.10. Synthesis of *O*-6-benzyloxyguanine (2.36).



Scheme 2.11. Synthesis of (*S*)- G^{DPC} phosphoramidite.

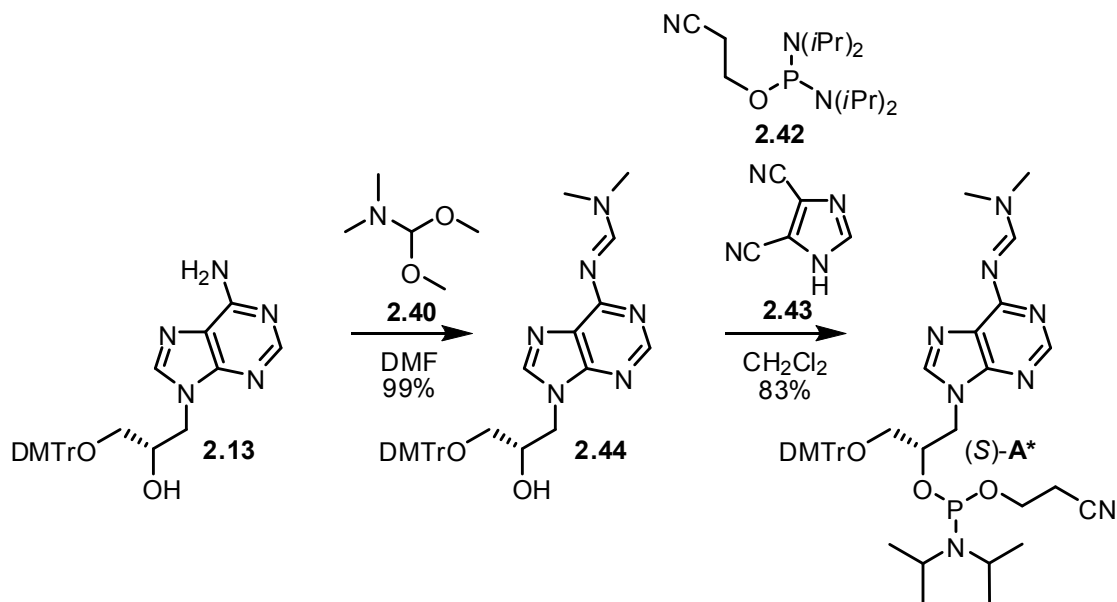
The major shortcoming of (*S*)-**G**^{DPC} phosphoramidite is that it suffers from the incomplete removal of protection groups after solid phase oligonucleotide synthesis, especially with the incorporation of several guanine nucleotides into a GNA oligonucleotide. Furthermore, the conditions required to remove the DPC protection group are harsh (overnight in 25% aqueous ammonia at 55 °C) and not completely compatible with syntheses requiring milder deprotection. Therefore, although (*S*)-**G**^{DPC} was more stable than (*S*)-**G**, we were still interested in further improving the protection group scheme. The original amide protection scheme of the exocyclic amines of guanine and adenine derivatives¹²⁻¹³ suffers from the formation of side products, long work-up times, and sub-optimal yields. Encouraged by other reports¹⁴⁻¹⁶ of success using amidine protection, we investigated protecting the exocyclic amino group of compound **2.38** as a dimethylformamidine.¹⁷ This was accomplished quite easily by heating a mixture of compound **2.38** and dimethylformamide-dimethylacetal (**2.40**) in DMF at 60 °C for one hour yielding compound **2.41** in 86% yield (Scheme 2.12). Conversion to the phosphoramidite (*S*)-**G**^{*} was accomplished in 76% yield by the reaction with *N,N,N',N'*-tetrakisopropylphosphordiamidite (**2.42**) and 0.7 equivalents of 4,5-dicyanoimidazole (**2.43**) in methylene chloride.¹⁸ Gratifyingly, this new phosphoramidite was also stable to chromatography over silica gel unlike its *N*-2-isobutyryl counterpart. The new phosphoramidite (*S*)-**G**^{*} is also stable for extended periods of time when stored under nitrogen at -20 °C. This new synthesis using amidine protection has an overall yield of 30% in five steps compared to 12% in five steps for the old synthesis in Scheme 2.1.



Scheme 2.12. Synthesis of (S)-G* phosphoramidite.

Chapter 2.4. Modified synthesis of (S)-A and (S)-C

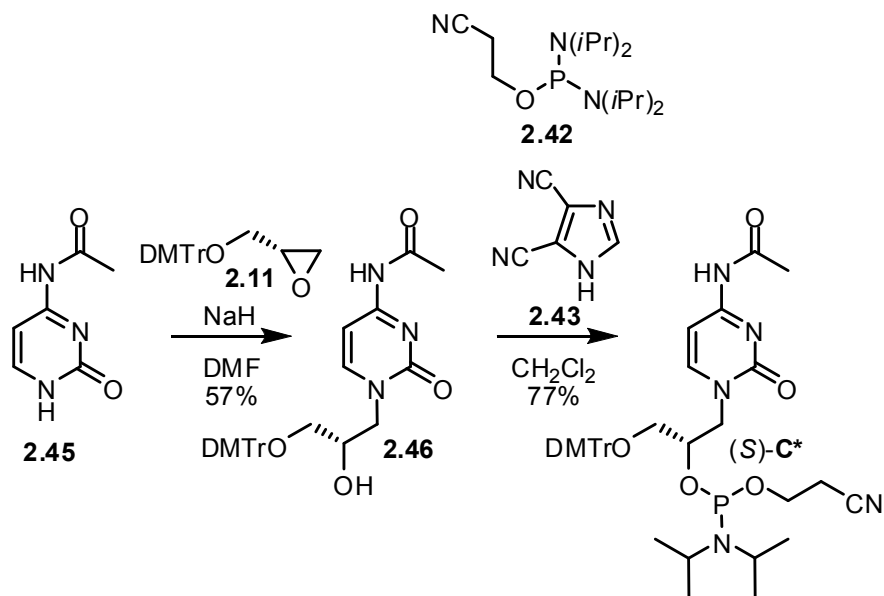
Due to the observed simplicity of amidine protection of the guanine derivative (S)-G* and the inherent problems with amide protection of the exocyclic amine of adenine derivatives, we were also interested in investigating *N*-6-dimethylformamidine protection¹⁷ of compound **2.13**. Amidine protection has been previously investigated in the context of adenosine nucleosides in an attempt to minimize acid catalyzed depurination during oligonucleotide synthesis.¹⁵ Accordingly, the reaction of compound **2.13** with dimethylformamide-dimethylacetal (**2.40**) in DMF afforded compound **2.44** in 99% yield (Scheme 2.13). The following conversion to phosphoramidite (S)-A* using *N,N,N',N'*-tetraisopropylphosphordiamidite (**2.42**) and 0.7 equivalents of 4,5-dicyanoimidazole (**2.43**) could be accomplished in 83% yield. This new synthesis has an improved overall yield of 51% over four steps for (S)-A* compared to 39% over four steps for the previous phosphoramidite (S)-A.



Scheme 2.13. Synthesis of (S)-A* phosphoramidite.

Furthermore, we were also interested in rendering phosphoramidite (S)-C more amenable towards milder deprotection.¹⁷ Therefore, we replaced the benzoyl protection group of (S)-C against an acetyl group in (S)-C*. N-4-acetylcytosine DNA phosphoramidites are widely used in mild and ultramild oligonucleotide synthesis.¹⁹ Accordingly, epoxide ring opening of compound **2.11** using N-4-acetylcytosine (**2.45**) and 0.2 equivalents of sodium hydride in DMF afforded compound **2.46** in 57% yield (Scheme 2.14). Subsequent conversion to the phosphoramidite (S)-C* using N,N,N',N'-tetraisopropylphosphordiamidite (**2.42**) and 0.7 equivalents of 4,5-dicyanoimidazole (**2.43**) proceeded in 77%. It should be noted that attempts to purify compound (S)-C* using flash chromatography over silica gel were unsuccessful and that the pure phosphoramidite could only be isolated using basic alumina (Brockmann Type II). This

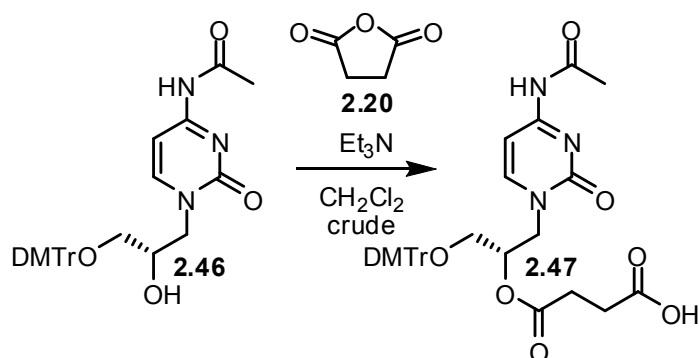
new synthesis has an overall yield of 44% over three steps for (*S*)-**C*** which is slightly lower than the overall yield of 52% over three steps for phosphoramidite (*S*)-**C**.



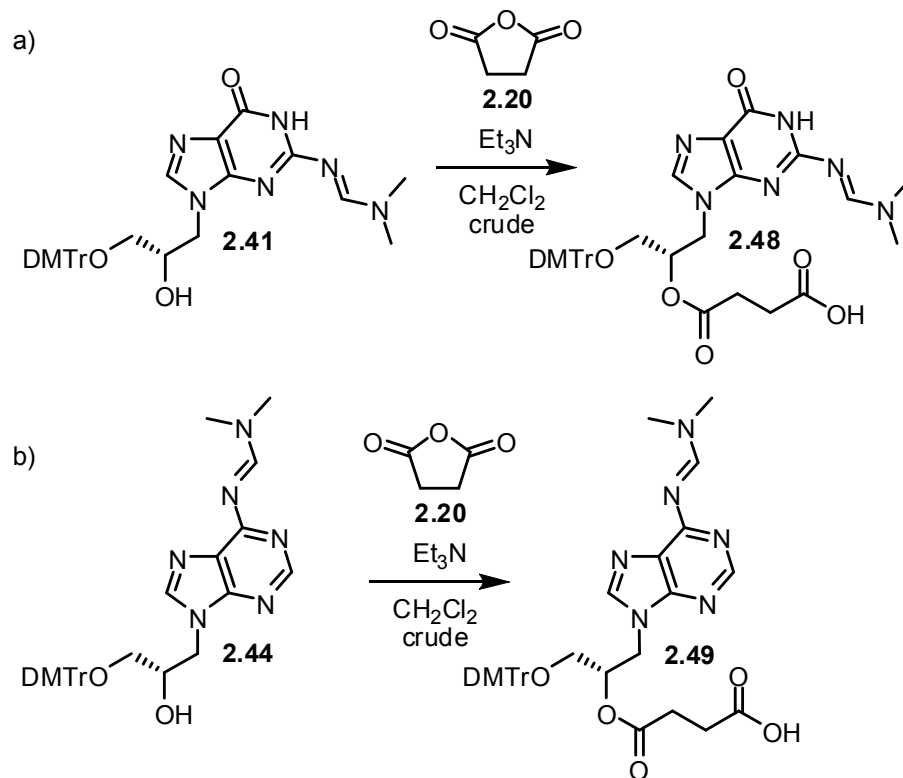
Scheme 2.14. Synthesis of (*S*)-**C*** phosphoramidite.

Chapter 2.5. Second generation GNA oligonucleotide synthesis and purification

As before, the synthesis of all phosphoramidites was followed by the storage procedure as stated in chapter 2.2. Solid supports were also synthesized as previously described in chapter 2.2 (see schemes 2.15 and 2.16 for synthesis of derivatized nucleosides and scheme 2.7 for the linkage to the CPG). The phosphoramidites [(*S*)-A*, (*S*)-G*, (*S*)-C*, and (*S*)-T] were left under high vacuum overnight and the next morning they were dissolved using anhydrous acetonitrile to a final concentration of 100 mM. These solutions were transferred to oven dried vials specially designed for use with the automated oligonucleotide synthesizer. Care was taken to avoid exposure to air during all manipulations. Solid supports were weighed out and transferred to polypropylene synthesis columns sealed with column frits. Prior to synthesis, the solid support was dried for five minutes directly on the synthesizer using argon.



Scheme 2.15. Synthesis of the succinyl ester from (*S*)-C* glycol nucleoside.



Scheme 2.16. Synthesis of succinyl esters from: a) (*S*)-**G*** glycol nucleoside, b) (*S*)-**A*** glycol nucleoside.

Oligonucleotides were synthesized on an Applied Biosystems ABI 394 Automated DNA/RNA synthesizer. In addition to the changes stated in chapter 2.2, the cycle was altered so that upon initial introduction of phosphoramidite and activator to the column it was left to react for 60 seconds before the introduction of the second portion of these reagents (outlined specifically in the appendix, cycle 1000GNA3, steps 12, 23, 34, 45, 49-50, 55-58, 65, 80-86). Best results were obtained when the last 4,4'-dimethoxytrityl group protecting the 3'-hydroxy group was left attached after synthesis as a purification handle. The solid support containing oligonucleotide was dried under argon on the synthesizer and then transferred to a screw cap vial capable of withstanding

product.²⁰⁻²¹ Since the DMTr group is removed directly on the column, this allows for a “Trityl-ON” purification of the target oligonucleotide in a much shorter time frame than HPLC. The Sep-Pak column was first pre-equilibrated using 10 mL of acetonitrile followed by 10 mL of 5 mM triethylammonium acetate buffer (pH=7.0). The solution (including solid support) was then applied three times via syringe to the column to ensure complete binding of the oligonucleotide. The column was then washed via syringe with the following solvents in the order stated: 15 mL of 3% NH₄OH, 10 mL of water, 10 mL of 1.5% aqueous TFA, and finally with 10 mL of water. The 3% aqueous ammonia serves to wash failure sequences from the column before cleaving the dimethoxytrityl group using 1.5% aqueous TFA. The crude “Trityl-OFF” oligonucleotide can then be eluted using 2.5 mL of 20% aqueous acetonitrile and purified using a Waters-Xterra column. Identity was confirmed by MALDI as in chapter 2.2. A representative HPLC trace of a crude “Trityl-OFF” GNA oligonucleotide solution is shown in Figure 2.4.

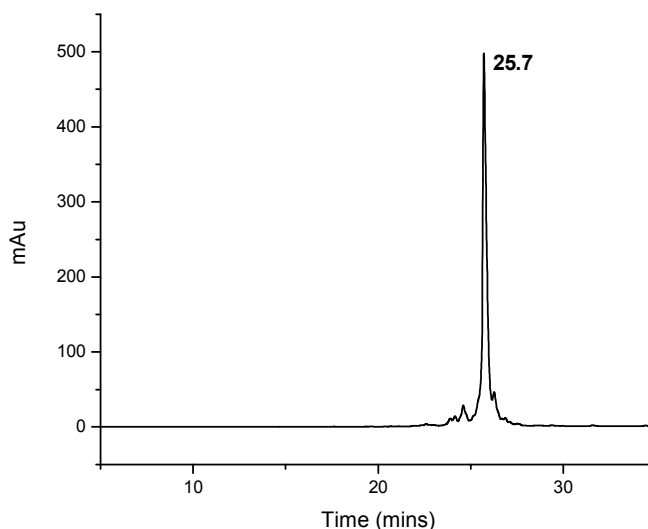


Figure 2.4. Crude “Trityl-OFF” HPLC trace of the GNA oligonucleotide 3’-CATGTCGTGCGT A-2’. The peak at 25.7 minutes represents the full length product. The crude oligo was eluted using a Waters Xterra column (MS C₁₈, 4.6 x 50 mm, 2.5 μM particle size) at 60 °C with a linear gradient (flow = 1.0 mL/min) from 2-8% acetonitrile in 30 minutes and 98-92% aqueous tritethylammonium acetate buffer (50 mM, pH=7.0).

The major advantage of this solid phase extraction method of purification using a Sep-Pak C₁₈ column versus the initial purification protocol (Scheme 2.8) is the amount of time saved. A “Trityl-ON” purification of a GNA oligonucleotide using a Sep-Pak column proceeds in approximately 15 minutes versus approximately one day for a reverse phase HPLC purification. Using the Sep-Pak column allows for the purification and cleavage of the remaining trityl group directly on the column, while the initial procedure involves precipitation of the crude oligonucleotide following cleavage of the remaining trityl group after HPLC purification. Furthermore, use of the dimethylformamide (adenine, guanine) and the acetyl protection groups (cytosine) allows for a much quicker deprotection procedure than the original phosphoramidites synthesized in chapter 2.1.

Removal of these new protection groups can be achieved in just 15 minutes at 55 °C versus 12-18 hours at 55 °C for the original protection groups. This is also advantageous considering the crude oligonucleotide is exposed to highly basic conditions at elevated temperatures for a shorter amount of time; conditions known to degrade and modify oligonucleotides. This results in a total time savings of approximately two days without sacrificing the purity or the integrity of the crude oligonucleotide product.

Chapter 2.6. Conclusions

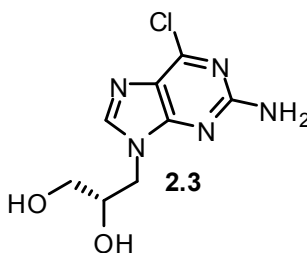
Although the initial published synthesis of GNA phosphoramidites¹ is much simpler, less time consuming, and more efficient than the corresponding DNA phosphoramidites, there was room for improvement. The reactions employed to protect the exocyclic amines of adenine and guanine involved highly air or moisture sensitive reagents. Furthermore, these reactions suffered from sub-optimal yields and the formation of side products that are not easily removed. The synthesis of (*S*)-**G** phosphoramidite is particularly problematic, especially considering the lack of stability inherent to the final compound.

The development of the phosphoramidites (*S*)-**A***, (*S*)-**G***, and (*S*)-**C*** have not only streamlined the synthesis of the phosphoramidites, but also the subsequent oligonucleotides.¹⁷ The use of dimethylformamide protection for the exocyclic amines of adenine and guanine provide stable phosphoramidites with high yields and simple purifications. The reaction employed to protect the exocyclic amines of the adenine and guanine derivatives **2.13** and **2.38** as dimethylformamides, respectively, proceeds with the formation of no side products and is complete in just one hour. This renders formamide protection a superior protection scheme to amide protection for GNA. Furthermore, this new protection scheme utilizing dimethylformamide [(*S*)-**A***, (*S*)-**G***] and acetyl [(*S*)-**C***] protection groups allows for a quicker post-synthetic work up of the resulting GNA oligonucleotides. One can access solutions of crude “Trityl-OFF” GNA

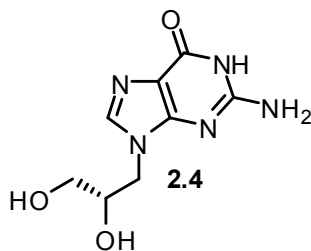
oligonucleotides in just 30 minutes as opposed to approximately two days for the initial synthetic conditions involving an extra HPLC purification step.

Chapter 2.7. Experimental procedures

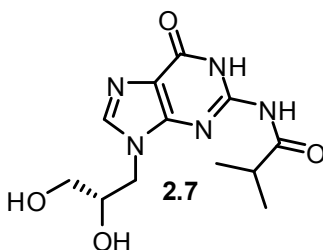
General procedures and reagents. NMR spectra were recorded on a Bruker DRX-500 (500 MHz), DMX-360 (360 MHz), or DMX-300 (300 MHz) spectrometer. High-resolution mass spectra were obtained with a Micromass AutoSpec or Thermo LTQ-FT instrument using ES ionization. Infrared spectra were recorded either on a Perkin Elmer 1600, Nicolet 510, or Bruker alpha series FTIR spectrometer. Solvents and reagents were used as supplied from Aldrich, Acros, Fluka, or TCI. Reactions were performed under an atmosphere of argon or nitrogen unless otherwise specified.



Compound 2.3. To a suspension of 2-amino-6-chloropurine (2.50 g, 14.8 mmol) and potassium carbonate (350 mg, 2.5 mmol) in DMF (50 mL) was added (*R*)-glycidol (1.00 mL, 15.1 mmol) and the suspension heated to 90 °C overnight. The solution was concentrated, redissolved in methanol, and then dry loaded onto silica gel. The crude product was purified via flash chromatography over silica gel eluting with 10:1 EtOAc:MeOH to afford compound **2.3** as a white solid (1.64 g, 46%). ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm) 8.02 (s, 1H), 6.87 (s, 2H), 5.08 (d, *J* = 5.4 Hz, 1H), 4.79 (t, *J* = 5.6 Hz, 1H), 4.19 (dd, *J* = 13.9, 3.5 Hz, 1H), 3.91 (dd, *J* = 13.9, 8.6, 1H), 3.81 (m, 1H), 3.40 (m, 1H), 3.32 (m, 1H). Data matches that of previously published data.¹

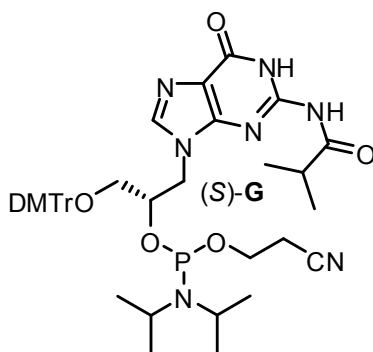


Compound 2.4. A suspension of compound **2.3** (1.90 g, 7.8 mmol) in 1 M HCl (68 mL) was heated to 85 °C for three hours. After cooling to room temperature, the pH was adjusted to approximately 9 using concentrated aqueous ammonia and the solvent was removed until the start of the formation of a white precipitate. This solution was then cooled to 4 °C overnight. The next morning, the product was filtered to afford compound **2.4** as a light blue solid (1.40 g, 80%). ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm) 7.59 (s, 1H), 6.43 (s, 2H), 5.05 (br, 1H), 4.77 (br, 1H), 4.08 (dd, *J* = 13.6, 3.4 Hz, 1H), 3.81 (dd, *J* = 13.6, 8.1 Hz, 1H), 3.75 (m, 1H), 3.36 (m, 1H), 3.28 (m, 1H). Data matches that of previously published data.¹



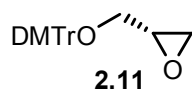
Compound 2.7. To a nitrogen purged suspension of Compound **2.4** (1.38 g, 6.1 mmol) in anhydrous pyridine (41 mL) was added trimethylsilyl chloride (5.8 mL, 46.0 mmol) and the resulting solution was allowed to stir at room temperature for two hours. After cooling to 0 °C, isobutyryl chloride (3.2 mL, 30.6 mmol) was added dropwise and the resulting solution was allowed to warm up gradually to room temperature and stir

2H), 7.37-7.16 (m, 7H), 6.81 (m, 4H), 5.71 (br, 1H), 4.46 (br, 1H), 4.28 (dd, $J = 14.0, 2.3$ Hz, 1H), 4.01 (dd, $J = 14.0, 8.7$ Hz, 1H), 3.77 (s, 6H), 3.29 (dd, $J = 9.6, 4.6$ Hz, 1H), 3.19 (dd, $J = 9.4, 6.1$ Hz, 1H), 2.69 (p, $J = 6.9$ Hz, 1H), 1.25 (dd, $J = 6.3, 3.2, 6H$). Data matches that of previously published data.¹

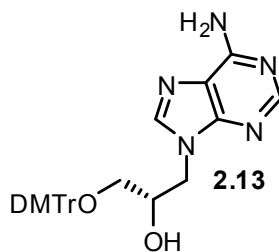


Compound (S)-G. To an argon purged solution of Compound **2.9a** (475 mg, 0.80 mmol) and *N,N*-diisopropylethylamine (0.20 mL, 1.1 mmol) in tetrahydrofuran (7.0 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.19 mL, 0.85 mmol) dropwise and the solution stirred overnight at room temperature. The solution was washed once with saturated aqueous NaHCO₃ extracted into CH₂Cl₂, dried over Na₂SO₄, and finally concentrated by rotary evaporation. A solution of the crude product in methylene chloride (4.0 mL) was precipitated by the addition to rapidly stirring pentane (350 mL). After filtration of the white precipitate, the solid was redissolved in methylene chloride (20 mL) and pentane added to the solution until the first sign of precipitation (~50 mL). The solution was stored at -20 °C for two hours after which the solution was decanted and the remaining oil redissolved in methylene chloride and concentrated to afford compound (S)-G as a white foam with a purity of ~60% based on NMR integration

(410 mg, 65%). ^{31}P NMR (121 MHz, CDCl_3) δ (ppm) 147.2, 147.1. Data matches that of previously published data.¹

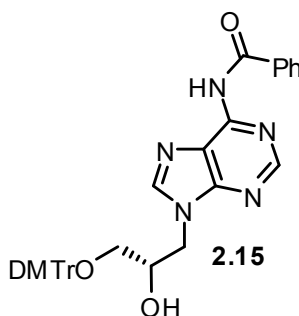


Compound 2.11. To a solution of (*R*)-glycidol (2.00 mL, 30.1 mmol) and triethylamine (10.5 mL, 75.3 mmol) in methylene chloride (68 mL) was added dimethoxytrityl chloride (12.8 g, 37.7 mmol) and allowed to stir overnight under nitrogen. The next morning, the solution was diluted with methylene chloride and washed once with saturated aqueous NaHCO_3 , dried over Na_2SO_4 , and concentrated by rotary evaporation. The crude product was purified via flash chromatography over silica gel eluting with 10:1:0.01 Hexanes:EtOAc:Et₃N to afford compound **2.11** as a colorless oil (11.6 g, 102%). ^1H NMR (500 MHz, CDCl_3) δ (ppm) 7.47 (m, 2H), 7.35 (m, 4H), 7.29 (m, 2H), 7.21 (m, 1H), 6.84 (m, 4H), 3.79 (s, 6H), 3.32 (dd, $J = 10.0, 2.4$ Hz, 1H), 3.17-3.10 (m, 2H), 2.78 (dd, $J = 4.9, 4.2$ Hz, 1H), 2.63 (dd, $J = 5.0, 2.5$ Hz, 1H). Data matches that of previously published data.¹



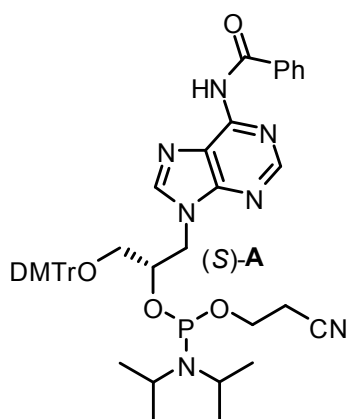
Compound 2.13. To a suspension of adenine (1.10 g, 8.1 mmol) in 16 mL of anhydrous DMF was added NaH (60% in mineral oil, 65 mg, 1.6 mmol) and allowed to stir under nitrogen for one hour. A solution of compound **2.11** (2.90 g, 7.7 mmol) in 16 mL of anhydrous DMF was added to the above solution and the reaction was heated to 110 °C

overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column chromatography starting with 100:1 EtOAc:Et₃N, then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound **2.13** as a white foam (2.60 g, 62%). ¹H NMR (300 MHz, CDCl₃) δ (ppm) 8.25 (s, 1H), 7.72 (s, 1H), 7.39 (m, 2H), 7.31-7.17 (m, 7H), 6.81 (m, 4H), 5.85 (s, 2H), 4.41 (dd, *J* = 14.2, 2.5 Hz, 1H), 4.29 (dd, *J* = 14.2, 6.7 Hz, 1H), 4.19 (m, 1H), 3.78 (s, 6H), 3.75 (s, 1H), 3.26 (dd, *J* = 9.6, 5.5 Hz, 1H), 3.04 (dd, *J* = 9.6, 6.4 Hz, 1H). Data matches that of previously published data.¹



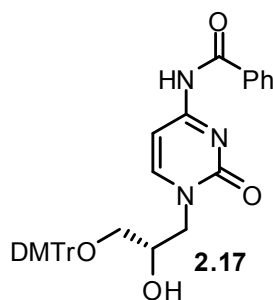
Compound 2.15. To an argon purged solution of compound **2.13** (5.30 g, 10.4 mmol) in anhydrous pyridine (83 mL) was added trimethylsilyl chloride (5.3 mL, 41.4 mmol) and the solution allowed to stir at room temperature for two hours. The solution was then cooled to 0 °C and benzoyl chloride (1.81 mL, 15.5 mmol) added dropwise. After allowing the solution to warm up gradually and stir for an additional two hours, it was cooled once more to 0 °C and 13 mL of water added to quench the reaction. After stirring for 15 minutes, 30 mL of concentrated aqueous ammonia was added and allowed to stir for another 30 minutes after warming to room temperature. The solution was diluted with methylene chloride, washed with water, dried over Na₂SO₄, and

concentrated. The crude product was then redissolved in 200 mL of methylene chloride, TBAF was added (1M solution in THF, 20 mL), and then allowed to stir at room temperature for one hour. After concentration, the crude product was purified via flash chromatography over silica gel starting with 100:1 EtOAc:Et₃N, then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound **2.15** as a light yellow foam (4.95 g, 78%). ¹H NMR (300 MHz, CDCl₃) δ (ppm) 9.39 (br, 1H), 8.69 (s, 1H), 8.02 (d, *J* = 7.2 Hz, 2H), 7.95 (s, 1H), 7.58 (m, 1H), 7.51-7.39 (m, 4H), 7.32-7.18 (m, 7H), 6.82 (m, 4H), 4.49 (dd, *J* = 13.9, 2.3 Hz, 1H), 4.29 (dd, *J* = 13.9, 7.4 Hz, 1H), 4.21 (m, 1H), 3.78 (s, 6H), 3.67 (s, 1H), 3.25 (dd, *J* = 9.8, 5.4 Hz, 1H), 3.17 (dd, *J* = 9.7, 5.6 Hz, 1H). Data matches that of previously published data.¹

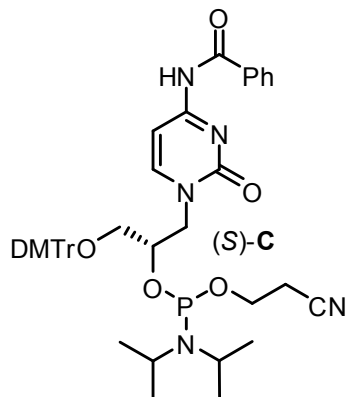


Compound (S)-A. To an argon purged solution of Compound **2.15** (2.15 g, 3.5 mmol) and *N,N*-diisopropylethylamine (3.65 mL, 21.0 mmol) in methylene chloride (58 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (1.15 mL, 5.2 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The solution was washed one time with saturated aqueous NaHCO₃ extracted into CH₂Cl₂, dried over Na₂SO₄, and finally concentrated by rotary evaporation. The crude product

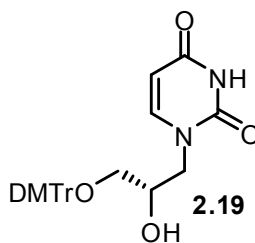
was purified by column chromatography starting with 1:1:0.01 Hexanes:EtOAc:Et₃N, then with 1:2:0.01 Hexanes:EtOAc:Et₃N to afford compound (*S*)-**A** as a white foam (2.30 g, 81%). ³¹P NMR (121 MHz, CDCl₃) δ (ppm) 150.4, 149.8. Data matches that of previously published data.¹



Compound 2.17. To a suspension of *N*⁴-benzoylcytosine (2.59 g, 12.0) in 24.0 mL of anhydrous DMF was added NaH (60% in mineral oil, 96 mg, 2.4 mmol) and allowed to stir under nitrogen for one hour. A solution of compound **2.11** (4.30 g, 11.4 mmol) in 24.0 mL of anhydrous DMF was added to the above solution and heated to 110 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column chromatography starting with 3:2:0.01 Hexanes:Acetone:Et₃N, then eluting with 1:1:0.01 Hexanes:Acetone:Et₃N to afford compound **2.17** as a light yellow foam (4.50 g, 63%). ¹H NMR (300 MHz, CDCl₃) δ (ppm) 8.80 (br, 1H), 7.90 (m, 2H), 7.61 (m, 2H), 7.50 (m, 2H), 7.42 (m, 3H), 7.33-7.17 (m, 7H), 6.82 (m, 4H), 4.36 (dd, *J* = 13.6, 2.8 Hz, 1H), 4.23 (m, 1H), 3.83 (m, 2H), 3.77 (s, 6H), 3.26 (dd, *J* = 9.7, 5.3 Hz, 1H), 3.12 (dd, *J* = 9.6, 5.9 Hz, 1H). Data matches that of previously published data.¹

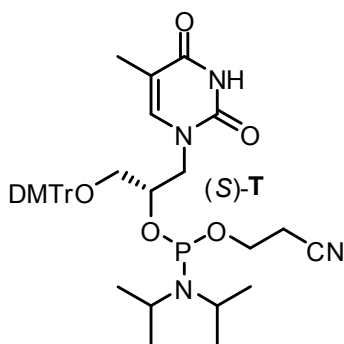


Compound (S)-C. To a nitrogen purged solution of Compound **2.17** (1.22 g, 2.1 mmol) and *N,N*-diisopropylethylamine (2.16 mL, 12.4 mmol) in methylene chloride (34 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.69 mL, 3.1 mmol) dropwise and the solution stirred for two hours at room temperature under nitrogen. The solution was washed one time with saturated aqueous NaHCO₃ extracted into CH₂Cl₂, dried over Na₂SO₄, and finally concentrated by rotary evaporation. The crude product was purified by column chromatography starting with 1:2:0.01 Hexanes:EtOAc:Et₃N, then with 1:4:0.01 Hexanes:EtOAc:Et₃N to afford compound (S)-C as a white foam (1.35 g, 83%). ³¹P NMR (121 MHz, CDCl₃) δ (ppm) 150.0, 149.9. Data matches that of previously published data.¹



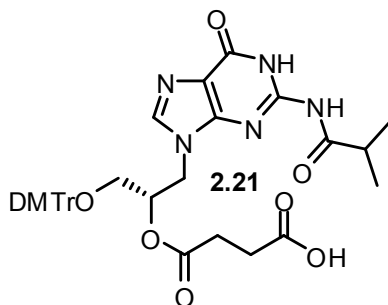
Compound 2.19. To a suspension of thymine (2.00 g, 15.9 mmol) in 33.0 mL of anhydrous DMF was added NaH (60% in mineral oil, 130 mg, 3.2 mmol) and allowed to

stir under argon for one hour. A solution of compound **2.11** (5.70 g, 15.1 mmol) in 30.0 mL of anhydrous DMF was added to the above solution and heated to 110 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column chromatography eluting with 1:2:0.01 Hexanes:EtOAc:Et₃N to afford compound **2.19** as a white foam (4.20 g, 55%). ¹H NMR (500 MHz, CDCl₃) δ (ppm) 8.26 (s, 1H), 7.41 (m, 2H), 7.33-7.21 (m, 7H), 7.05 (m, 1H), 6.84 (m, 4H), 4.07 (m, 1H), 4.02 (dd, *J* = 14.1, 3.0 Hz, 1H), 3.80 (s, 6H), 3.66 (dd, *J* = 14.1, 7.2 Hz, 1H), 3.19 (d, *J* = 5.4 Hz, 2H), 2.85 (d, *J* = 4.6 Hz, 1H), 1.85 (s, 3H). Data matches that of previously published data.¹

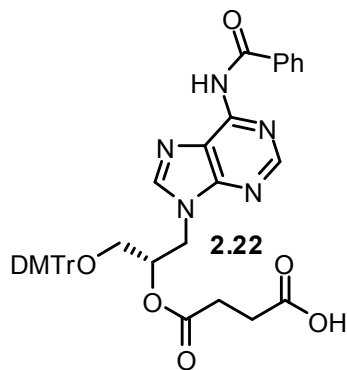


Compound (S)-T. To a nitrogen purged solution of Compound **2.19** (2.10 g, 4.2 mmol) and *N,N*-diisopropylethylamine (4.35 mL, 25.1 mmol) in methylene chloride (70 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (1.40 mL, 6.3 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The solution was washed one time with saturated aqueous NaHCO₃ extracted into CH₂Cl₂, dried over Na₂SO₄, and finally concentrated by rotary evaporation. The crude product was purified by column chromatography eluting with 1:1:0.01 Hexanes:EtOAc:Et₃N to

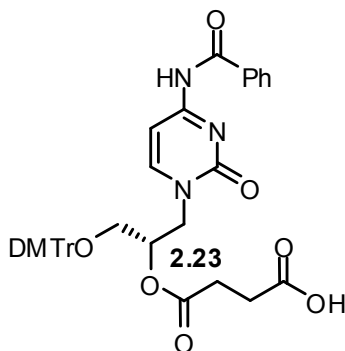
afford compound (*S*)-**T** as a white foam (2.35 g, 80%). ^{31}P NMR (121 MHz, CDCl_3) δ (ppm) 150.1, 150.0. Data matches that of previously published data.¹



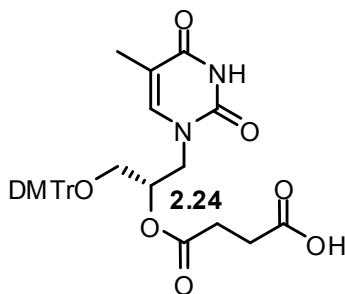
Compound 2.21. To a solution of compound **2.9** (150 mg, 0.25 mmol) and succinic anhydride (38 mg, 0.38 mmol) in methylene chloride (2.5 mL) was added triethylamine (0.11 mL, 0.75 mmol) and the solution allowed to stir under argon for four hours. The solution was diluted with methylene chloride, washed once with a solution of 4% aqueous citric acid, dried over Na_2SO_4 , and finally concentrated to afford compound **2.21** as a white foam (168 mg). The product was used in the next step without further purification. ^1H NMR (500 MHz, CDCl_3) δ (ppm) 10.10 (s, 1H), 7.54 (s, 1H), 7.39 (d, $J = 7.5$ Hz, 2H), 7.31-7.16 (m, 7H), 6.83 (d, $J = 8.1$ Hz, 4H), 5.22 (m, 1H), 4.50 (m, 1H), 4.32 (dd, $J = 14.8, 3.1$ Hz, 1H), 3.78 (s, 6H), 3.25 (q, $J = 4.9$ Hz, 1H), 2.89-2.50 (m, 5H), 2.42 (m, 1H), 1.11 (t, $J = 7.7$ Hz, 6H).



Compound 2.22. To a solution of compound **2.15** (155 mg, 0.25 mmol) and succinic anhydride (38 mg, 0.38 mmol) in methylene chloride (2.5 mL) was added triethylamine (0.11 mL, 0.75 mmol) and the solution allowed to stir under argon for four hours. The solution was diluted with methylene chloride, washed once with a solution of 4% aqueous citric acid, dried over Na₂SO₄, and finally concentrated to afford compound **2.22** as a tan foam (173 mg). The product was used in the next step without further purification. ¹H NMR (500 MHz, CDCl₃) δ (ppm) 8.68 (s, 1H), 8.08 (s, 1H), 8.01 (d, *J* = 7.4 Hz, 2H), 7.54 (m, 1H), 7.48-7.41 (m, 4H), 7.33-7.19 (m, 7H), 6.84 (d, *J* = 8.3 Hz, 4H), 5.39 (m, 1H), 4.57 (m, 2H), 3.78 (s, 6H), 3.29 (d, *J* = 4.8 Hz, 2H), 2.63-2.48 (m, 4H).

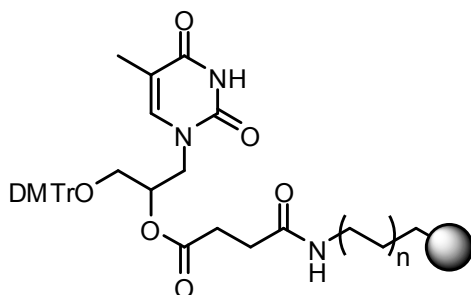


Compound 2.23. To a solution of compound **2.17** (151 mg, 0.26 mmol) and succinic anhydride (38 mg, 0.38 mmol) in methylene chloride (2.5 mL) was added triethylamine (0.11 mL, 0.75 mmol) and the solution allowed to stir under argon for four hours. The solution was diluted with methylene chloride, washed once with a solution of 4% aqueous citric acid, dried over Na₂SO₄, and finally concentrated to afford compound **2.23** as a white foam (200 mg). The product was used in the next step without further purification. ¹H NMR (500 MHz, CDCl₃) δ (ppm) 7.94 (d, 2H), 7.56 (t, *J* = 7.5 Hz, 1H), 7.50-7.41 (m, 6H), 7.35-7.20 (m, 7H), 6.85 (d, *J* = 7.8 Hz, 4H), 5.30 (m, 1H), 4.23 (m, 1H), 4.11 (dd, *J* = 13.9, 8.1 Hz, 1H), 3.79 (s, 6H), 3.33-3.24 (m, 2H), 2.79-2.55 (m, 4H).



Compound 2.24. To a solution of compound **2.19** (141 mg, 0.28 mmol) and succinic anhydride (42 mg, 0.42 mmol) in methylene chloride (2.8 mL) was added triethylamine (0.12 mL, 0.84 mmol) and the solution allowed to stir under argon for four hours. The

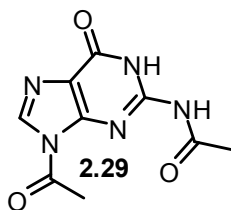
solution was diluted with methylene chloride, washed once with a solution of 4% aqueous citric acid, dried over Na₂SO₄, and finally concentrated to afford compound **2.24** as a white foam (173 mg). The product was used in the next step without further purification. ¹H NMR (500 MHz, CDCl₃) δ (ppm) 7.41 (m, 2H), 7.33-7.20 (m, 7H), 7.01 (m, 1H), 6.84 (d, *J* = 8.8 Hz, 4H), 5.28 (m, 1H), 4.09 (dd, *J* = 14.5, 3.8 Hz, 1H), 3.93 (dd, *J* = 14.5, 8.0 Hz, 1H), 3.79 (s, 6H), 3.30-3.23 (m, 2H), 2.71-2.52 (m, 4H), 1.84 (s, 3H).



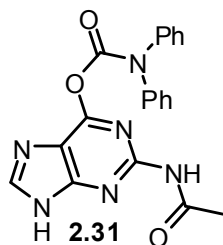
(S)-T Solid support

Synthesis of solid supports [(S)-T as example]. To a suspension of long-chain alkylamine controlled pore glass (300 mg, mesh 120-200, 174 μmole/gram loading) and 1-hydroxybenzotriazole (2 mg, 0.02 mmol) in anhydrous acetonitrile (2 mL) and pyridine (0.10 mL) was added *N,N'*-diisopropylcarbodiimide (24 μL, 0.15 mmol) in a sealed glass vial. The suspension was shaken at room temperature for 30 minutes. Afterwards, compound **2.24** (33 mg, 0.055 mmol) was added to the vial and then shaken overnight at room temperature. The next morning, the solid support was filtered and washed three times with both methanol and then chloroform. After drying, the solid support was transferred to a new vial and suspended in 8:1:1 THF:2,6-lutidine:acetic anhydride (5 mL) and 9:1 THF:*N*-methylimidazole (5 mL) and shaken for another hour. The solid support was filtered once more and washed three times with both methanol and then

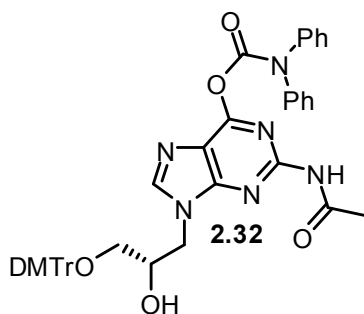
chloroform. The solid supports were transferred to new glass vials and dried under high vacuum to ensure all the excess solvent was removed. The loading was measured by accurately weighing approximately 5 mg of the solid support and suspending it in 10% wt:wt trichloroacetic acid in methylene chloride. The maximum absorbance was measured at 504 nm and the loading calculated based on the extinction coefficient of 76 mL*cm/ μ mole for the dimethoxytrityl group. Typical values were in the range from 55-70 μ mole/gram.



Compound 2.29. To a suspension of guanine (1.03 g, 6.8 mmol) in DMF (8.5 mL) was added acetic anhydride (1.60 mL, 17.0 mmol). The suspension was heated to reflux for three hours after which it was cooled, filtered, and the solid washed with ethanol. NMR showed only protection of the exocyclic amino group, so the product was resuspended in DMF (8.5 mL) and acetic anhydride (2.58 mL, 27.3 mmol) added. The solution was again heated to reflux at which point the solid dissolved. After two hours the solution was allowed to cool slowly to room temperature and then to 0 °C, the precipitate filtered, and the solid washed with ethanol to afford compound **2.29** as a fluffy tan solid (800 mg, 50%). ¹H NMR (360 MHz, DMSO-*d*₆) δ (ppm) 8.44 (s, 1H), 2.82 (s, 3H), 2.22 (s, 3H). Data matches that of previously published data.⁸



Compound 2.31. To a suspension of compound **2.29** (790 mg, 3.4 mmol) and diphenylcarbonyl chloride (855 mg, 3.7 mmol) in pyridine (16 mL) was added *N,N*-diisopropylethylamine (1.17 mL, 6.7 mmol) and the solution allowed to stir at room temperature for one hour. The reaction was quenched by the addition of water (2 mL) and allowed to stir for 10 minutes after which it was concentrated and coevaporated with toluene. The solid was redissolved in a 1:1 mixture of ethanol:water (40 mL) and heated over a steam bath for 1.5 hours after which the solution was cooled, filtered, and washed with ethanol to afford compound **2.31** as a purple solid (950 mg, 73%). ^1H NMR (360 MHz, $\text{DMSO-}d_6$) δ (ppm) 13.53 (br, 1H), 10.57 (s, 1H), 8.44 (s, 1H), 7.54-7.28 (m, 10H), 2.17 (s, 3H). Data matches that of previously published data.⁸

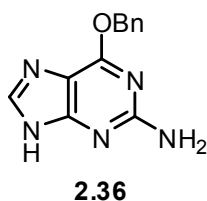


Attempted synthesis of compound 2.32. To a suspension of compound **2.31** (200 mg, 0.51 mmol) in 1.0 mL of anhydrous DMF was added NaH (60% in mineral oil, 4 mg, 0.10 mmol) and allowed to stir under nitrogen for one hour. A solution of compound **2.11** (185 mg, 0.49 mmol) in 1.0 mL of anhydrous DMF was added to the above solution

and the reaction was heated to 90 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column chromatography starting with 50:1:0.01 EtOAc:MeOH:Et₃N, the eluting with 40:3:0.01 EtOAc:MeOH:Et₃N. ¹H NMR did not indicate the desired product.

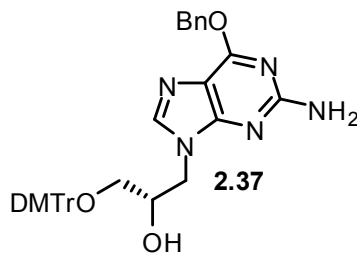


Compound 2.34. A solution of 2-amino-6-chloropurine (2.50 g, 14.7 mmol) and 1,4-diazabicyclo[2.2.2]octane (9.10 g, 81.1 mmol) in DMSO (15 mL) was allowed to stir for three hours at room temperature. To the white suspension was added 150 mL of ethyl acetate and allowed to stir for another three hours. The white precipitate was then filtered and washed with ethyl acetate and diethyl ether to afford compound **2.34** as a white solid which was carried crude to the next step. ¹H NMR (300 MHz, D₂O) δ (ppm) 8.21 (s, 1H), 4.16 (t, *J* = 7.6 Hz, 3H), 3.40 (t, *J* = 7.6 Hz, 3H). Data matches that of previously published data.⁹



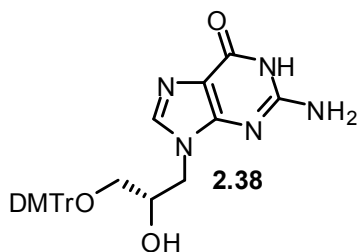
Compound 2.36. To a suspension of sodium hydride (60% in mineral oil, 1.18 g, 29.5 mmol) in anhydrous DMSO (7.5 mL) was slowly added benzyl alcohol (8.40 mL, 81.1

mmol). After reacting for two hours, crude compound **2.34** (14.8 mmol) and DMSO (7.5 mL) were added to the above solution and allowed to stir at room temperature under nitrogen overnight. The next morning, the solution was dumped into ice water and the pH of the solution adjusted to 8 with concentrated acetic acid. After removing the solvent from the resulting suspension, the solid was redissolved in a mixture of methylene chloride and methanol and dry loaded onto silica gel. The crude product was purified via flash chromatography over silica gel starting with 20:1 CH₂Cl₂:MeOH, then eluting with 15:1 CH₂Cl₂:MeOH to afford compound **2.36** as a white solid (3.10 g, 83% two steps). ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm) 12.45 (s, 1H), 7.82 (s, 1H), 7.51 (m, 2H), 7.43-7.31 (m, 3H), 6.32 (s, 2H), 5.48 (s, 2H). Data matches that of previously published data.⁹



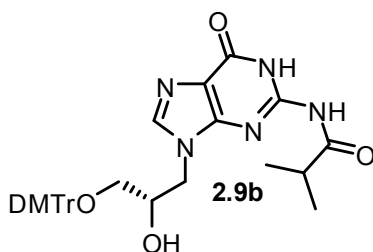
Compound 2.37. Compound **2.36** (3.1 g, 12.8 mmol) was partially dissolved in anhydrous DMF (25 mL) under a nitrogen atmosphere. NaH was added (105 mg, 2.6 mmol, 60% in mineral oil) and the solution was allowed to stir under nitrogen for one hour. In a separate flask, compound **2.11** was dissolved in 26 mL of DMF, added to the first solution, and then heated to 90 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column

chromatography starting with 2:1:0.01 Hexanes:Acetone:Et₃N, then eluting with 3:2:0.01 Hexanes:Acetone:Et₃N to afford compound **2.37** as a light yellow foam (3.7 g, 47%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 7.53 (s, 1H), 7.51 (d, *J* = 7.1 Hz, 2H), 7.42 (d, *J* = 7.5 Hz, 2H), 7.35-7.25 (m, 9H), 7.21 (t, *J* = 7.3 Hz, 1H), 6.82 (d, *J* = 8.8 Hz, 4H), 5.52 (s, 2H), 5.11 (b, 1H), 4.85 (s, 2H), 4.28 (m, 1H), 4.16 (m, 2H), 3.78 (s, 6H), 3.21 (dd, *J* = 9.5, 4.3 Hz, 1H), 3.03 (dd, *J* = 9.4, 5.6 Hz, 1H). ¹³C-NMR (125 MHz, CDCl₃) δ (ppm) 161.0, 158.8, 158.7, 154.0, 144.8, 140.8, 136.5, 136.0, 135.9, 130.1, 128.5, 128.4, 128.2, 128.1, 128.0, 127.0, 115.4, 113.3, 86.4, 69.5, 68.3, 64.7, 55.3, 48.5. IR (film) ν (cm⁻¹) = 3515, 1401, 3341, 3212, 3065, 3034, 2934, 2834, 1616, 1589, 1510, 1456, 1410, 1385, 1356, 1333, 1300, 1252, 1175, 1154, 1101, 1061, 1028, 907, 828, 789, 756, 727, 698, 633, 581. HRMS calcd for C₃₆H₃₅N₅O₅ (M+Na)⁺ 640.2536, found (M+Na)⁺ 640.2530.



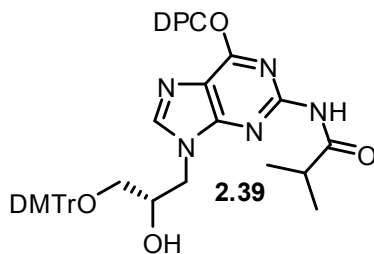
Compound 2.38. Compound **2.37** (3.7g, 6 mmol) and Pd/C (3.7 g, 5% on Carbon) were suspended in EtOAc (150 mL) and the solution was purged with nitrogen, then H₂, and allowed to stir under a hydrogen atmosphere for 3 hours after which TLC showed completion of the reaction. The mixture was filtered through celite and washed with 2:1:0.01 EtOAc:MeOH:Et₃N to afford crude compound **2.38** as a tan solid. The crude product was used in the next step further without purification. ¹H-NMR (300 MHz,

DMSO-*d*₆) δ (ppm) 10.76 (b, 1H), 7.57 (s, 1H), 7.41 (d, $J = 7.4$ Hz, 2H), 7.33-7.18 (m, 7H), 6.88 (dd, $J = 8.9, 3.4$ Hz, 4H), 6.50 (b, 2H), 5.39 (b, 1H), 4.08-3.90 (m, 3H), 3.73 (s, 6H), 2.95 (dd, $J = 9.5, 5.2$ Hz, 1H), 2.89 (dd, $J = 9.4, 4.8$ Hz, 1H). ¹³C-NMR (75 MHz, DMSO-*d*₆) δ (ppm) 158.0, 157.1, 153.6, 151.3, 144.9, 138.0, 135.6, 129.7, 127.8, 126.6, 116.4, 113.1, 85.3, 68.0, 65.7, 55.0, 49.3. IR (film) ν (cm⁻¹) =3412, 3127, 1699, 1607, 1578, 1541, 1508, 1481, 1462, 1443, 1412, 1379, 1302, 1250, 1173, 1152, 1113, 1074, 1024, 986, 899, 828, 789, 777, 752, 725, 692, 629, 579, 548, 519, 432, 422, 414, 407. HRMS calcd for C₂₉H₂₉N₅O₅ (M+Na)⁺ 550.2066, found (M+Na)⁺ 550.2061.



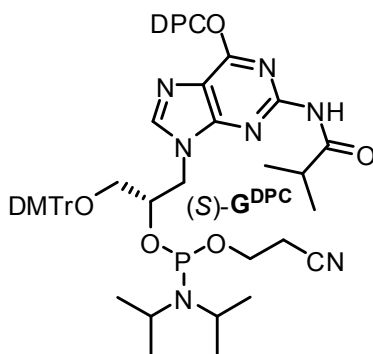
Compound 2.9b. To a nitrogen purged suspension of Compound **2.38** (3.0 g, 5.7 mmol) in anhydrous pyridine (38 mL) was added trimethylsilyl chloride (2.9 mL, 22.7 mmol) and the resulting solution was allowed to stir at room temperature for two hours. After cooling to 0 °C, isobutyryl chloride (1.2 mL, 11.4 mmol) was added dropwise and the resulting solution allowed to warm up gradually to room temperature and stir for an additional 2 hours. After cooling again to 0 °C, 20 mL of water was added to quench the reaction and allowed to warm up and stir for 15 minutes at room temperature. The reaction was then washed with water and extracted two times into methylene chloride. After drying over sodium sulfate and concentration, the resulting oil was redissolved in THF (120 mL) and TBAF added under nitrogen (6 mmol, 1 M solution in THF). The

solution was washed again with water and extracted two times into ethyl acetate, dried over sodium sulfate, and concentrated. The resulting crude foam was purified via column chromatography starting with 100:1 EtOAc:Et₃N, then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound **2.9b** as a white foam (2.85 g, 84%). ¹H NMR data matched that of compound **2.9a** and reported data.¹

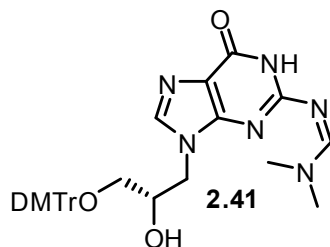


Compound 2.39. Compound **2.9b** (2.25 g, 3.8 mmol) and diphenylcarbonyl chloride (960 mg, 4.1 mmol) were dissolved in anhydrous pyridine (30 mL) under nitrogen. To this solution was added *N,N*-diisopropylethylamine (0.73 mL, 4.1 mmol) and then allowed to stir for 1 hour at room temperature. The resulting dark red solution was diluted with CHCl₂, washed with saturated aqueous NaHCO₃, dried over Na₂SO₄, concentrated, and finally coevaporated with toluene. The crude product was purified by column chromatography starting with 3:1:0.01 Hexanes:Acetone:Et₃N, then with 2:1:0.01 Hexanes:Acetone:Et₃N, and finally eluting with 3:2:0.01 Hexanes:Acetone:Et₃N to afford compound **2.39** as a light orange foam (2.4 g, 80%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 8.06 (s, 1H), 7.89 (s, 1H), 7.49-7.33 (m, 10H), 7.26 (m, 8H), 7.18 (t, *J* = 7.3 Hz, 1H), 6.79 (d, *J* = 8.8 Hz, 4H), 5.33 (b, 1H), 4.50 (dd, *J* = 14.3, 1.5 Hz, 1H), 4.28 (dd, *J* = 14.4 6.7 Hz, 1H), 4.18 (m, 1H), 3.75 (s, 6H), 3.39 (dd, *J* = 9.6, 5.2 Hz, 1H), 2.95 (m, 1H), 2.65 (m, 1H), 1.23 (d, *J* = 3.4 Hz, 3H), 1.22 (d,

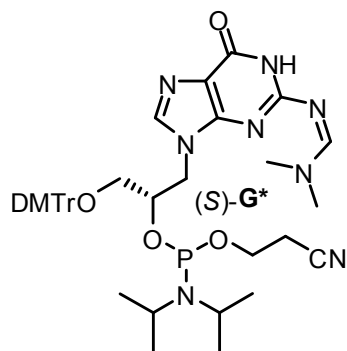
$J = 3.4$ Hz, 3H). $^{13}\text{C-NMR}$ (125 MHz, CDCl_3) δ (ppm) 158.7, 156.2, 155.3, 151.5, 150.4, 145.3, 144.7, 141.9, 135.9, 135.8, 130.09, 129.97, 129.3, 128.13, 128.02, 127.0, 121.3, 113.3, 86.6, 69.8, 64.3, 55.3, 50.0, 36.7, 19.34, 19.30. IR (film) ν (cm^{-1}) = 3413, 3065, 2971, 2933, 1756, 1625, 1591, 1511, 1493, 1447, 1408, 1385, 1337, 1301, 1251, 1220, 1177, 1154, 1073, 1051, 1029, 1002, 978, 695. HRMS calcd for $\text{C}_{46}\text{H}_{44}\text{N}_6\text{O}_7$ ($\text{M}+\text{Na}$) $^+$ 815.3164, found ($\text{M}+\text{Na}$) $^+$ 815.3164.



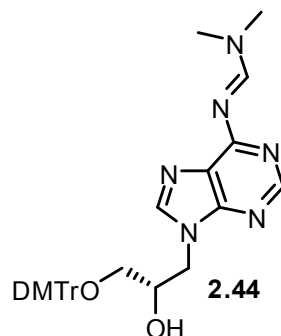
Compound (S)-G^{DPC}. To a nitrogen purged solution of Compound **2.39** (2.85 g, 3.6 mmol) and *N,N*-diisopropylethylamine (3.8 mL, 21.6 mmol) in methylene chloride (60 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (1.6 mL, 7.2 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The solution was diluted with CH_2Cl_2 , washed one time with saturated aqueous NaHCO_3 dried over Na_2SO_4 , and finally concentrated by rotary evaporation. The crude product was purified by column chromatography starting with 2:1:0.01 Hexanes:EtOAc: Et_3N , then with 3:2:0.01 Hexanes:EtOAc: Et_3N to afford compound (S)-G^{DPC} as a white foam (2.9 g, 81%). ^{31}P NMR (121 MHz, CDCl_3) δ (ppm) 150.4, 149.9. HRMS calcd for $\text{C}_{55}\text{H}_{61}\text{N}_8\text{O}_8$ ($\text{M}+\text{Na}$) $^+$ 1015.4248, found ($\text{M}+\text{Na}$) $^+$ 1015.4242.



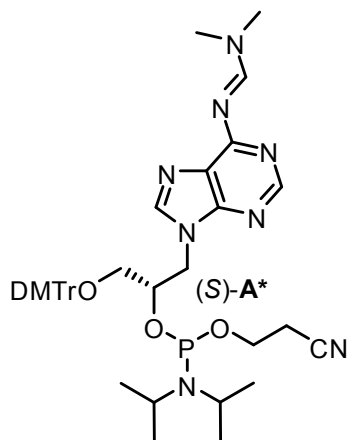
Compound 2.41. To a solution of compound **2.38** (2.63 g, 4.99 mmol) in anhydrous DMF (16.0 mL) was added dimethylformamide-dimethylacetal (2.35 mL, 17.4 mmol) and heated to 60 °C for one hour. After cooling and removal of the DMF, the residue was redissolved in methylene chloride, washed once with saturated aqueous NaHCO₃, dried over Na₂SO₄, and finally concentrated. The product was purified via column chromatography starting with 100:1 EtOAc:Et₃N, then eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **2.41** as a white foam (2.50 g, 86%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 8.73 (s, 1H), 8.45 (s, 1H), 7.49 (s, 1H), 7.45 (m, 2H), 7.36-7.17 (m, 7H), 6.82 (m, 4H), 4.38 (dd, *J* = 14.0, 2.4 Hz, 1H), 4.31 (m, 1H), 4.08 (q, *J* = 7.0 Hz, 1H), 3.78 (s, 6H), 3.32 (dd, *J* = 9.5, 5.0 Hz, 1H), 3.05 (m, 4H), 2.97 (s, 3H). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm) 158.7, 158.2, 157.5, 156.5, 150.3, 145.1, 139.5, 136.2, 136.1, 130.1, 128.2, 128.0, 126.9, 120.0, 113.3, 86.4, 69.3, 65.0, 55.4, 48.6, 41.3, 35.2. IR (solid) ν (cm⁻¹) = 2929, 2836, 1630, 1558, 1506, 1444, 1416, 1399, 1345, 1326, 1300, 1245, 1174, 1110, 1066, 1024, 981, 827, 755, 726, 701, 644, 581. HRMS calcd for C₃₂H₃₄N₆O₅Na (M+Na)⁺ 605.2483, found (M+Na)⁺ 605.2477.



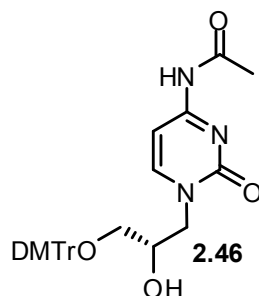
Compound (S)-G*. To a solution of compound **2.41** (1.80 g, 3.09 mmol) in 15.5 mL of anhydrous methylene chloride under nitrogen was added a 1 M solution of 4,5-dicyanoimidazole (2.20 mL in acetonitrile). 2-cyanoethyl *N,N,N',N'*-tetraisopropylphosphordiamidite (1.03 mL, 3.24 mmol) was then added dropwise and the solution stirred at room temperature. After two hours, the reaction mixture was diluted with methylene chloride, washed twice with saturated aqueous NaHCO₃, dried over Na₂SO₄, and then concentrated. The product was purified via column chromatography starting with 1:1:0.01 Hexanes:Acetone:Et₃N, then eluting with 1:2:0.01 Hexanes:Acetone:Et₃N to afford compound (S)-G* as a white foam (1.85 g, 76%). ³¹P-NMR (162 MHz, CDCl₃) δ (ppm) 150.3, 150.0. HRMS calcd for C₄₁H₅₂N₈O₆P (M+H)⁺ 783.3742, found (M+H)⁺ 783.3736.



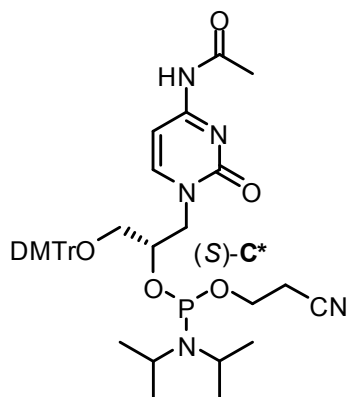
Compound 2.44. To a solution of compound **2.13** (1.32 g, 2.58 mmol) in 7.5 mL of anhydrous DMF was added dimethylformamide-dimethylacetal (1.21 mL, 9.04 mmol) and heated to 60 °C for one hour. After cooling and removal of the DMF, the residue was redissolved in methylene chloride, washed once with saturated aqueous NaHCO₃, dried over Na₂SO₄, and finally concentrated. The product was purified via column chromatography starting with 100:1 EtOAc:Et₃N, then eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **2.44** as a white foam (1.45 g, 99%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 8.89 (s, 1H), 8.44 (s, 1H), 7.84 (s, 1H), 7.43 (m, 2H), 7.35-7.17 (m, 7H), 6.82 (m, 4H), 4.48 (m, 1H), 4.27 (m, 1H), 3.79 (s, 6H), 3.29 (dd, *J* = 9.4, 4.7 Hz, 1H), 3.23 (s, 3H), 3.21 (s, 3H), 3.10 (dd, *J* = 9.4, 5.7 Hz, 1H). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm) 159.5, 158.7, 158.4, 152.1, 151.8, 144.8, 143.0, 135.91, 135.83, 130.1, 128.11, 128.00, 127.0, 125.8, 113.3, 86.5, 69.6, 64.8, 55.3, 48.7, 41.4, 35.2. IR (solid) ν (cm⁻¹) = 2929, 2836, 1630, 1558, 1506, 1444, 1416, 1399, 1345, 1326, 1300, 1245, 1174, 1110, 1066, 1024, 981, 827, 755, 726, 701, 644, 581. HRMS calcd for C₃₂H₃₅N₆O₄ (M+H)⁺ 567.2714, found (M+H)⁺ 567.2709.



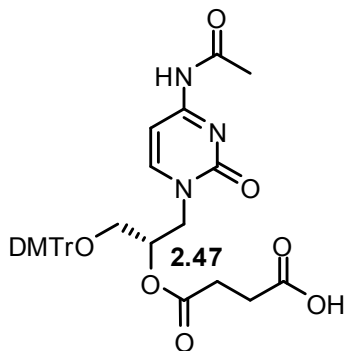
Compound (S)-A*. To a solution of **2.44** (1.82 g, 3.21 mmol) in 16.0 mL of anhydrous methylene chloride under nitrogen was added a 1 M solution of 4,5-dicyanoimidazole (2.20 mL in acetonitrile). 2-cyanoethyl *N,N,N',N'*-tetraisopropylphosphordiamidite (1.07 mL, 3.37 mmol) was then added dropwise and the solution stirred at room temperature. After two hours, the reaction mixture was diluted with methylene chloride, washed twice with saturated aqueous NaHCO₃, dried over Na₂SO₄, and then concentrated. The product was purified via column chromatography starting with 3:2:0.01 Hexanes:Acetone:Et₃N, then eluting with 1:1:0.01 Hexanes:Acetone:Et₃N to afford compound (S)-A* as a white foam (2.05 g, 83%). ³¹P-NMR (162 MHz, CDCl₃) δ (ppm) 150.4, 149.9. HRMS calcd for C₄₁H₅₂N₈O₅P (M+H)⁺ 767.3793, found (M+H)⁺ 767.3781.



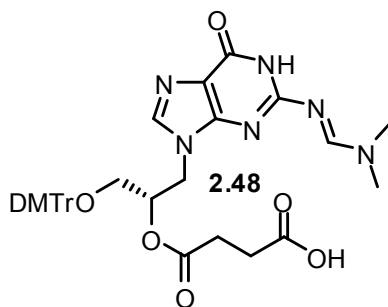
Compound 2.46. To a suspension of *N*⁴-acetylcytosine (1.16 g, 7.57 mmol) in 15.0 mL of anhydrous DMF was added NaH (60 mg, 1.5 mmol, 60% in mineral oil) and allowed to stir under nitrogen for one hour. A solution of **2.11** (2.71 g, 7.20 mmol) in 15.0 mL of anhydrous DMF was added to the above solution and the reaction was heated to 110 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column chromatography starting with 3:2:0.01 Hexanes:Acetone:Et₃N, then eluting with 1:1:0.01 Hexanes:Acetone:Et₃N to afford compound **2.46** as a light yellow foam (2.30 g, 57%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 9.36 (b, 1H), 7.57 (d, *J* = 7.3 Hz, 1H), 7.41 (m, 2H), 7.34-7.17 (m, 8H), 6.82 (m, 4H), 4.33 (dd, *J* = 13.5, 2.6 Hz, 1H), 4.21 (b, 1H), 3.91 (b, 1H), 3.78 (m, 7H), 3.24 (dd, *J* = 9.6, 5.1 Hz, 1H), 3.08 (dd, *J* = 9.6, 6.0 Hz, 1H), 2.22 (s, 3H). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm) 171.1, 162.9, 158.7, 157.4, 150.5, 144.7, 135.75, 135.71, 130.1, 128.09, 128.03, 127.0, 113.3, 96.7, 86.4, 68.9, 64.5, 55.3, 54.6, 24.9. IR (solid) ν (cm⁻¹) = 3256, 2963, 2929, 2836, 1630, 1558, 1507, 1445, 1417, 1347, 1299, 1245, 1174, 1112, 1069, 1030, 827, 789, 755, 727, 701, 645, 582. HRMS calcd for C₃₀H₃₁N₃O₆Na (M+Na)⁺ 552.2105, found (M+Na)⁺ 552.2104.



Compound (S)-C*. To a solution of **2.46** (1.06 g, 2.00 mmol) in 10.0 mL of anhydrous methylene chloride under nitrogen was added a 1 M solution of 4,5-dicyanoimidazole (1.40 mL in acetonitrile). 2-cyanoethyl *N,N,N',N'*-tetraisopropylphosphordiamidite (0.67 mL, 2.1 mmol) was then added dropwise and the solution stirred at room temperature. After two hours, the reaction mixture was diluted with methylene chloride, washed once with saturated aqueous NaHCO₃, dried over Na₂SO₄, and then concentrated. The product was purified via column chromatography (basic alumina, Brockmann Type II) starting with 100:1 EtOAc:Et₃N, then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound (S)-C* as a white foam (1.12 g, 77%). ³¹P-NMR (162 MHz, CDCl₃) δ (ppm) 150.3, 150.1. HRMS calcd for C₃₉H₄₉N₅O₇P (M+H)⁺ 730.3364, found (M+H)⁺ 730.3358.

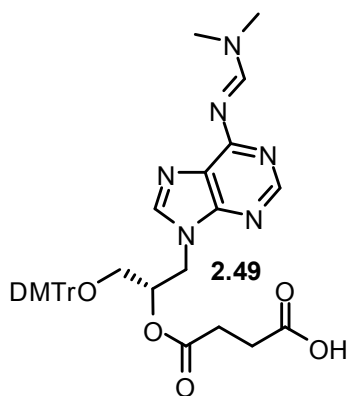


Compound 2.47. To a solution of compound **2.46** (106 mg, 0.20 mmol) and succinic anhydride (30 mg, 0.30 mmol) in methylene chloride (2.0 mL) was added triethylamine (0.09 mL, 0.60 mmol) and the solution allowed to stir under nitrogen for four hours. The solution was diluted with methylene chloride, washed once with a solution of 4% aqueous citric acid, dried over Na₂SO₄, and finally concentrated to afford compound **2.47** as a white foam (150 mg). The product was used in the next step without further purification. ¹H NMR (300 MHz, CDCl₃) δ (ppm) 7.57 (d, J = 7.5 Hz, 1H), 7.43 (m, 2H), 7.35-7.18 (m, 8H), 6.84 (d, J = 8.9 Hz, 4H), 5.37 (m, 1H), 4.39 (dd, J = 14.0, 3.4 Hz, 1H), 3.98 (dd, J = 14.0, 7.9 Hz, 1H), 3.79 (s, 6H), 3.32 (dd, J = 10.5, 4.6 Hz, 1H), 3.23 (dd, J = 10.4, 4.4 Hz, 1H), 2.73-2.60 (m, 4H), 2.14 (s, 3H).



Compound 2.48. To a solution of compound **2.41** (110 mg, 0.19 mmol) and succinic anhydride (28 mg, 0.28 mmol) in methylene chloride (1.9 mL) was added triethylamine

(0.08 mL, 0.57 mmol) and the solution allowed to stir under nitrogen for four hours. The solution was diluted with methylene chloride, washed once with a solution of 4% aqueous citric acid, dried over Na₂SO₄, and finally concentrated to afford compound **2.48** as a white foam (130 mg). The product was used in the next step without further purification. ¹H NMR (300 MHz, CDCl₃) δ (ppm) 8.40 (s, 1H), 7.70 (s, 1H), 7.41 (m, 2H), 7.32-7.16 (m, 7H), 6.81 (d, J = 8.2 Hz, 4H), 5.40 (m, 1H), 4.45 (dd, J = 14.6, 4.1 Hz, 1H), 4.31 (dd, J = 14.6, 6.4 Hz, 1H), 3.78 (s, 6H), 3.21 (d, J = 5.1 Hz, 2H), 3.02 (s, 3H), 2.96 (s, 3H), 2.72-2.56 (m, 4H).



Compound 2.49. To a solution of compound **2.44** (104 mg, 0.18 mmol) and succinic anhydride (28 mg, 0.28 mmol) in methylene chloride (1.8 mL) was added triethylamine (0.08 mL, 0.55 mmol) and the solution allowed to stir under nitrogen for four hours. The solution was diluted with methylene chloride, washed once with a solution of 4% aqueous citric acid, dried over Na₂SO₄, and finally concentrated to afford compound **2.49** as a tan foam (165 mg). The product was used in the next step without further purification. ¹H NMR (300 MHz, CDCl₃) δ (ppm) 8.86 (s, 1H), 8.51 (s, 1H), 8.09 (s, 1H), 7.42 (m, 2H), 7.34-7.17 (m, 7H), 6.83 (d, J = 8.8 Hz, 4H), 5.35 (m, 1H), 4.62-4.47

(m, 2H), 3.78 (s, 6H), 3.31 (dd, J = 10.5, 4.6 Hz, 1H), 3.25-2.19 (m, 4H), 3.17 (s, 3H), 2.71-2.56 (m, 4H).

MALDI-TOF MS: GNA oligonucleotide samples were prepared at a concentration of approximately 10-20 μM . One microliter of a saturated solution of trihydroxyacetophenone in 50% aqueous acetonitrile was mixed with one microliter of 100 mM ammonium tartrate and then one microliter of the oligo sample. The dried sample was then analyzed in negative mode for detection of the sample mass on an ABI Voyager 6030 or a Bruker Biflex III MALDI-TOF instrument. Best results were obtained when the oligo solution was desalted using DOWEX 50WX8 cation exchange beads (hydrogen form) prior to mixing with the matrix and ammonium tartrate solutions. The oligonucleotide was incubated at the concentration stated above with approximately 1 microliter of DOWEX beads for approximately 5-10 minutes, mixed thoroughly by pipetting the solution up and down, and incubated for another 5-10 minutes after which it was mixed with the matrix for spotting.

Chapter 2.8. References

- (1) Zhang, L.; Peritz, A. E.; Carroll, P. J.; Meggers, E. *Synthesis* **2006**, 645-653.
- (2) Zhang, L.; Meggers, E. *J. Am. Chem. Soc.* **2005**, 127, 74-75.
- (3) Damha, M. J.; Giannaris, P. A.; Zabarylo, S. V. *Nucleic Acids Research* **1990**, 18, 3813-3821.
- (4) Vials from commercially purchased DNA phosphoramidites (Glen Research) specially designed for ABI synthesizers were cleaned and reused for this purpose
- (5) Applied Biosystems Instruments: ABI 394 DNA/RNA Synthesizer Manual
- (6) Langley, G. J.; Herniman, J. M.; Davies, N. L.; Brown, T. *Rapid Commun. Mass Spectrom.* **1999**, 13, 1717-1723.
- (7) Krotz, A. H.; Rentel, C.; Gorman, D.; Olsen, P.; Gaus, H. J.; McArdle, J. V.; Scozzari, A. N. *Nucleosides, Nucleotides and Nucleic Acids* **2004**, 23, 767-775.
- (8) Zou, R.; Robins, M. J. *Can. J. Chem.* **1987**, 65, 1436-1437.
- (9) Lembicz, N. K.; Grant, S.; William, C.; Griffin, R. J.; Heath, S. L.; Golding, B. T. *J. Chem. Soc. Perkin Trans. I* **1997**, 185-186.
- (10) Kamimura, T.; Tsuchiya, M.; Koura, K.; Sekine, M.; Hata, T. *Tet. Lett.* **1983**, 24, 2775-2778.
- (11) Seela, F.; Kroeschel, R. *Bioconjugate Chem.* **2001**, 12, 1043-1050.
- (12) Schaller, H.; Weimann, G.; Lerch B.; Khorana H. G. *J. Am. Chem. Soc.* **1963**, 85, 3821-3827.
- (13) Ti, G. S.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1982**, 104, 1316-1319.

- (14) McBride, L. J.; Kierzek, R.; Beaucage, S. L.; Caruthers, M. H. *J Am. Chem. Soc.* **1986**, 108, 2040-2048.
- (15) Vu, H.; McCollum, C.; Jacobsen, K.; Theisen, P.; Vinayak, R.; Spiess, E.; Andrus, A. *Tetrahedron Lett.* **1990**, 31, 7269-7272.
- (16) Vinayak, R.; Anderson, P.; McCollum, C.; Hampel, A. *Nucleic Acids Res.* **1992**, 20, 1265-1269.
- (17) Schlegel, M. K.; Meggers, E. *J. Org. Chem.* **2009**, 74, 4615-4618.
- (18) Rosenbohm, C.; Christensen, S. M.; Sørensen, M. D.; Pedersen, D. S.; Larsen, L.-E.; Wengel, J.; Koch, T. *Org. Biomol. Chem.* **2003**, 1, 655-663.
- (19) Reddy, M. P.; Hanna, N. B.; Farooqui, F. *Tetrahedron Letters* **1994**, 23, 4311-4314.
- (20) Reddy, M. P.; Farooqui, F. U.S. Patent 5,770,723, **1998**.
- (21) Gilar, M.; Bouvier, E. S. P. *J. of Chromatography A* **2000**, 167-177.

Appendix to Chapter 2

^1H , ^{13}C , and ^{31}P NMR spectra

IR spectra

Oligonucleotide Synthesizer Protocols

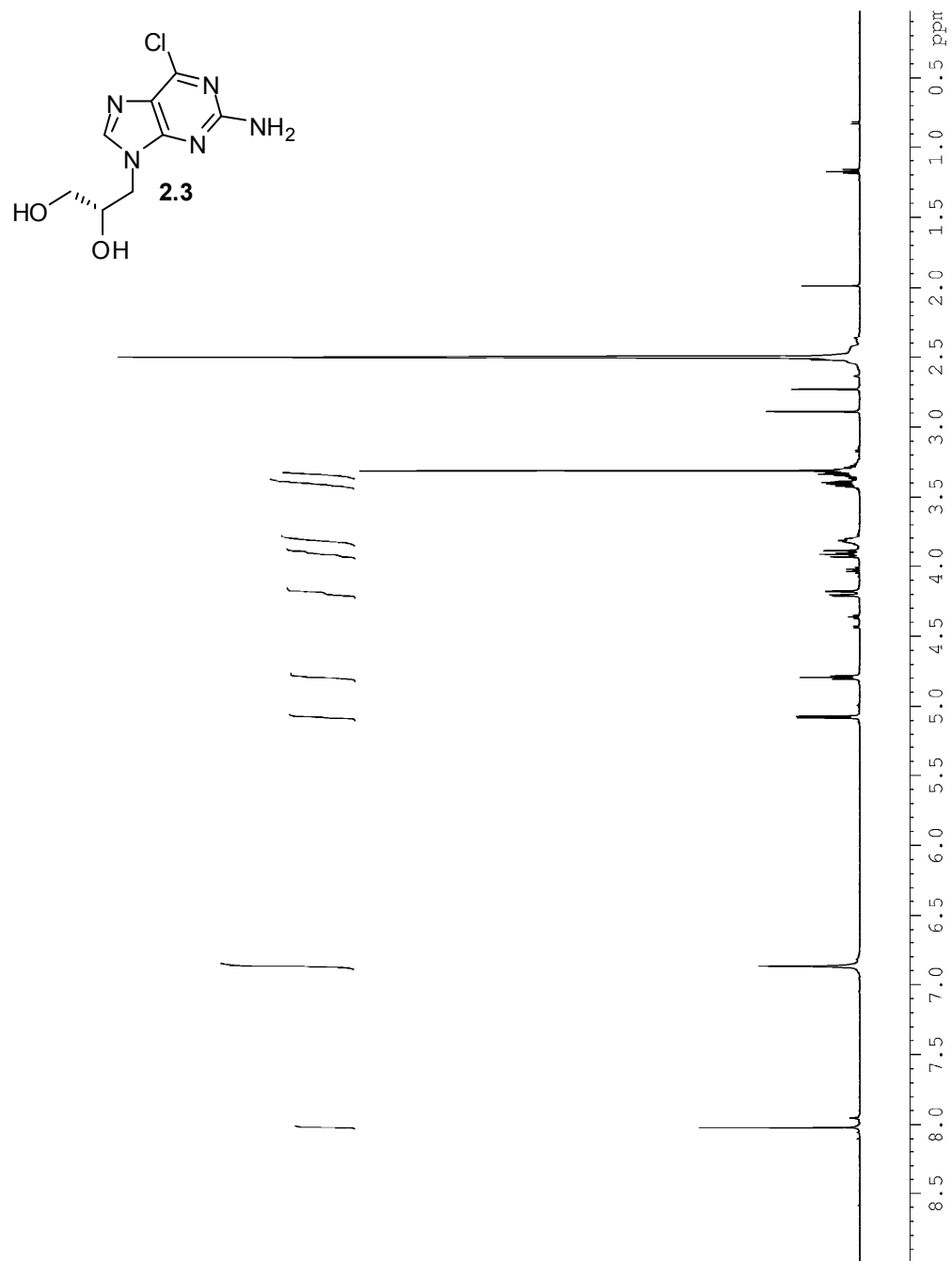


Figure A2.1.1. ^1H NMR spectrum of compound **2.3** (500 MHz, $\text{DMSO}-d_6$).

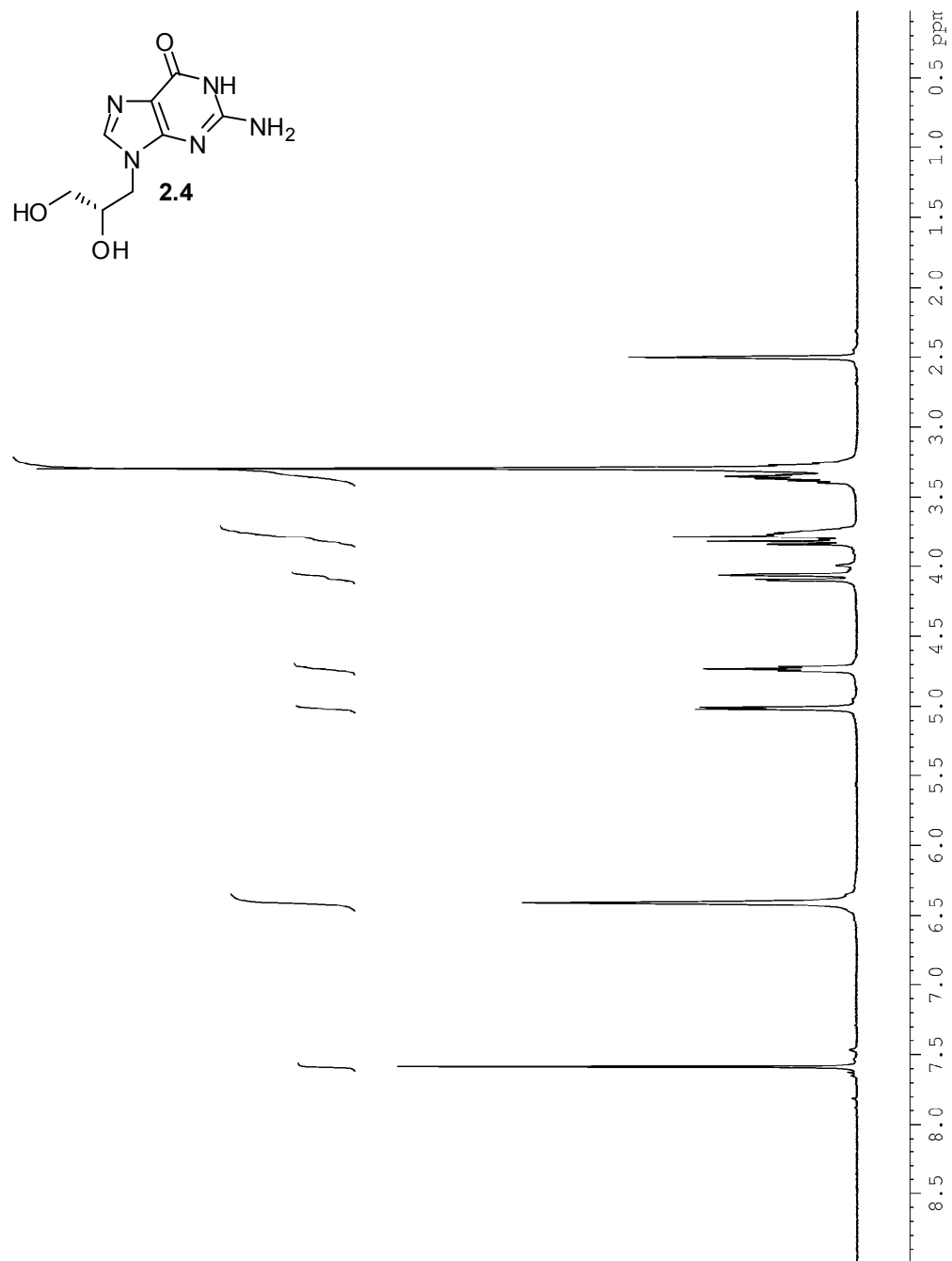


Figure A2.2.1. ¹H NMR spectrum of compound **2.4** (500 MHz, DMSO-*d*₆).

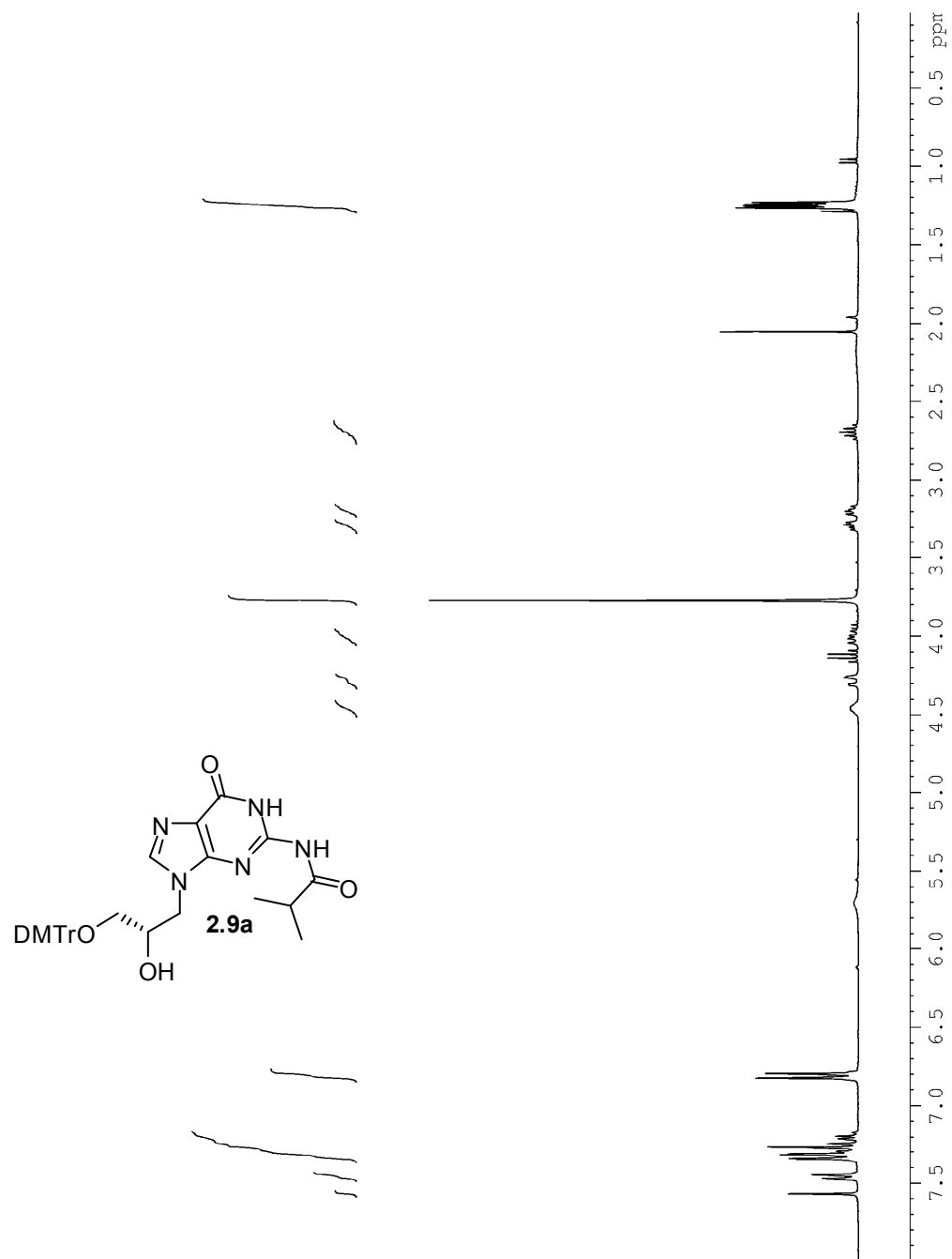


Figure A2.4.1. ^1H NMR spectrum of compound **2.9a** (300 MHz, CDCl_3).

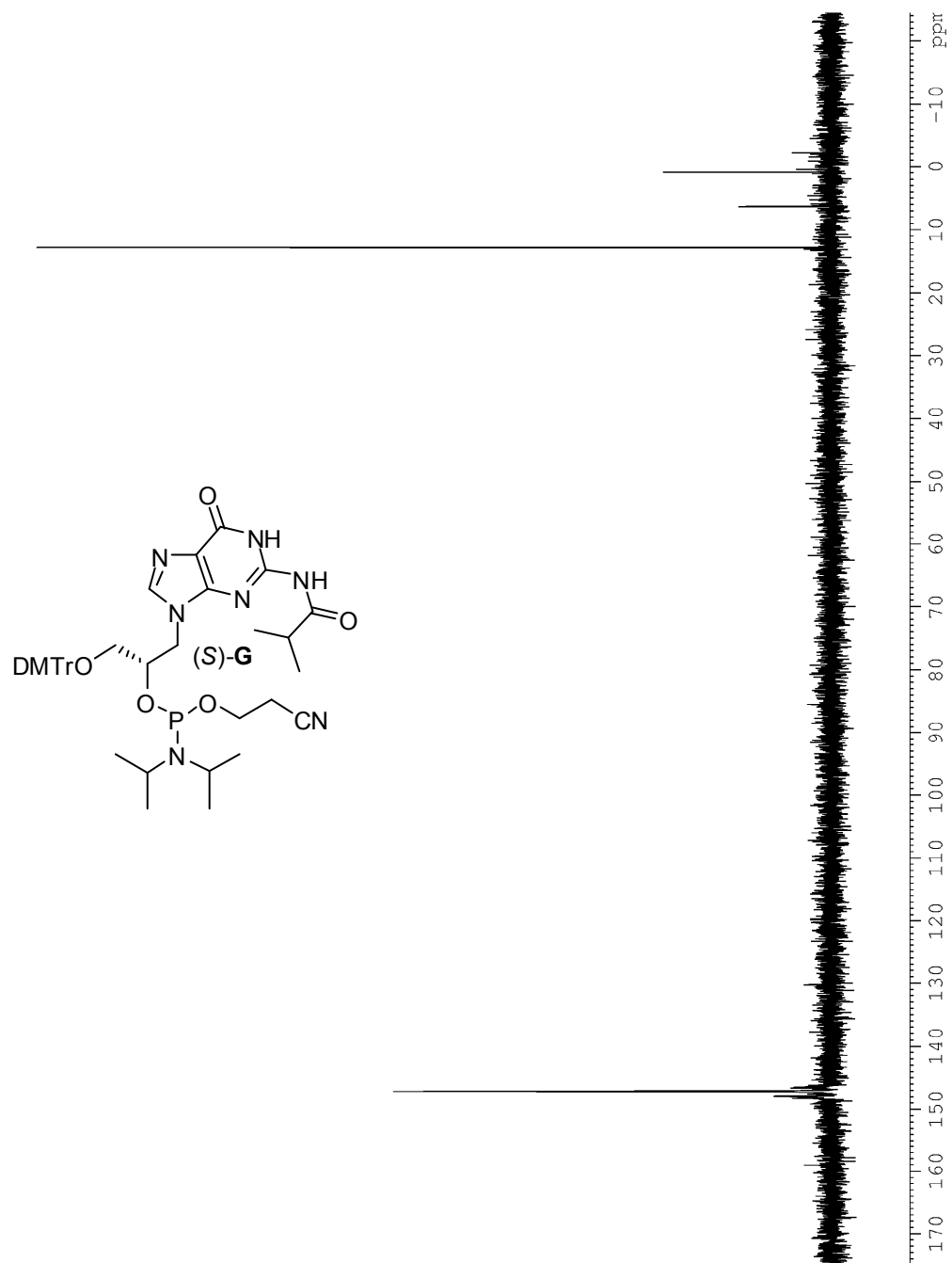


Figure A2.5.1. ^{31}P NMR spectrum of compound (S)-G (121 MHz, CDCl_3).

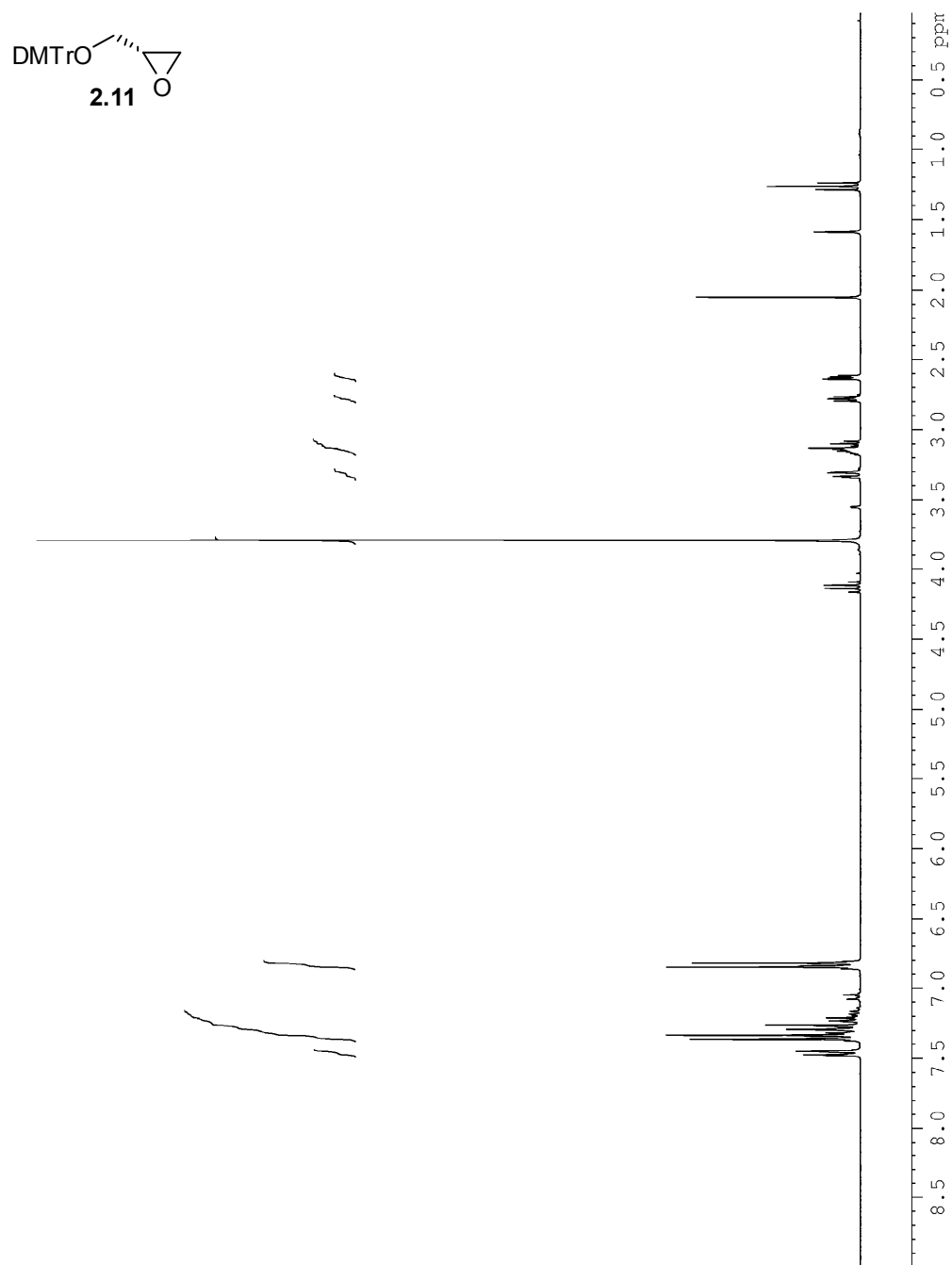


Figure A2.6.1. ^1H NMR spectrum of compound **2.11** (500 MHz, CDCl_3).

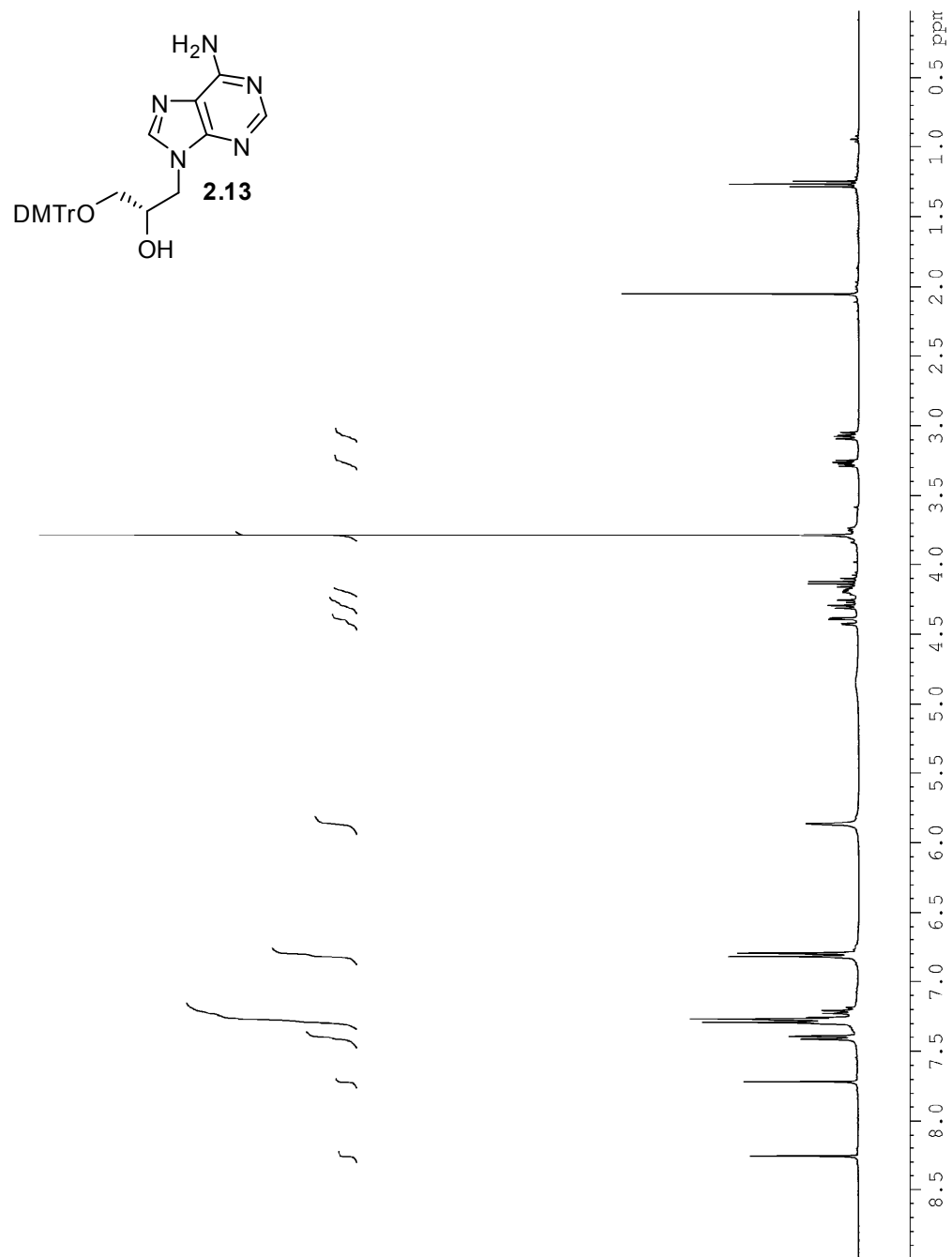


Figure A2.7.1. ^1H NMR spectrum of compound **2.13** (300 MHz, CDCl_3).

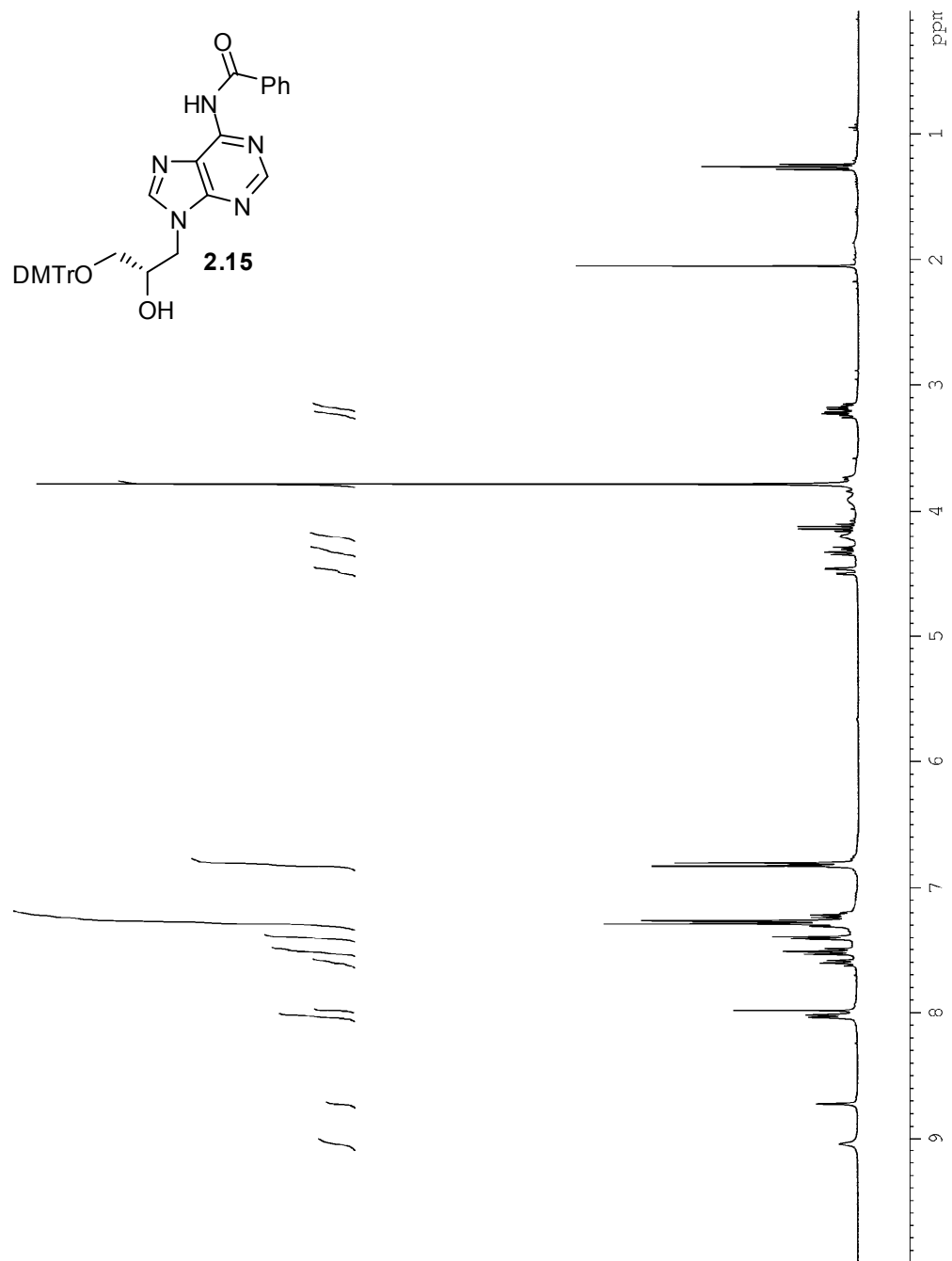


Figure A2.8.1. ^1H NMR spectrum of compound **2.15** (300 MHz, CDCl_3).

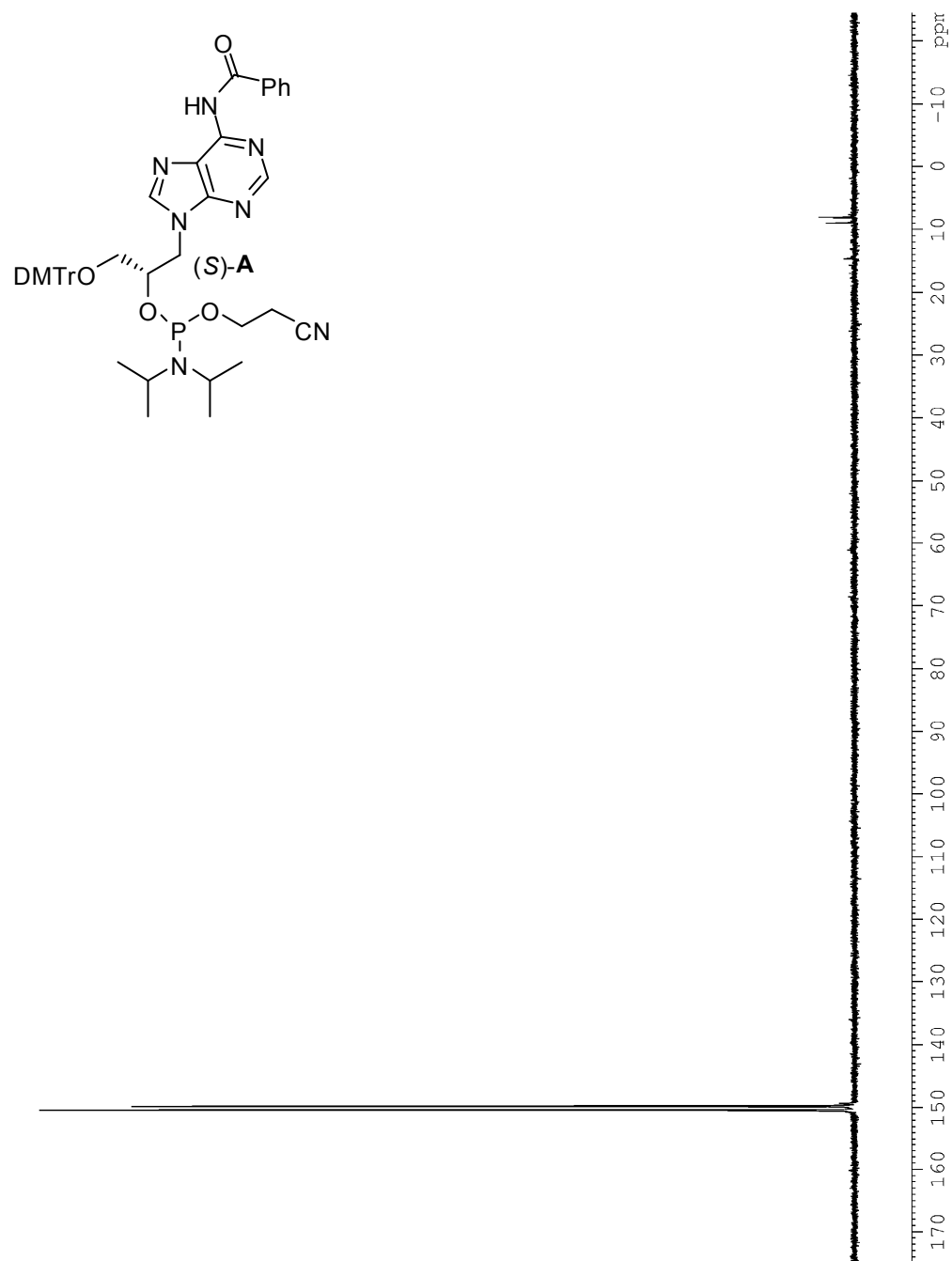


Figure A2.9.1. ^{31}P NMR spectrum of compound (S)-A (121 MHz, CDCl_3).

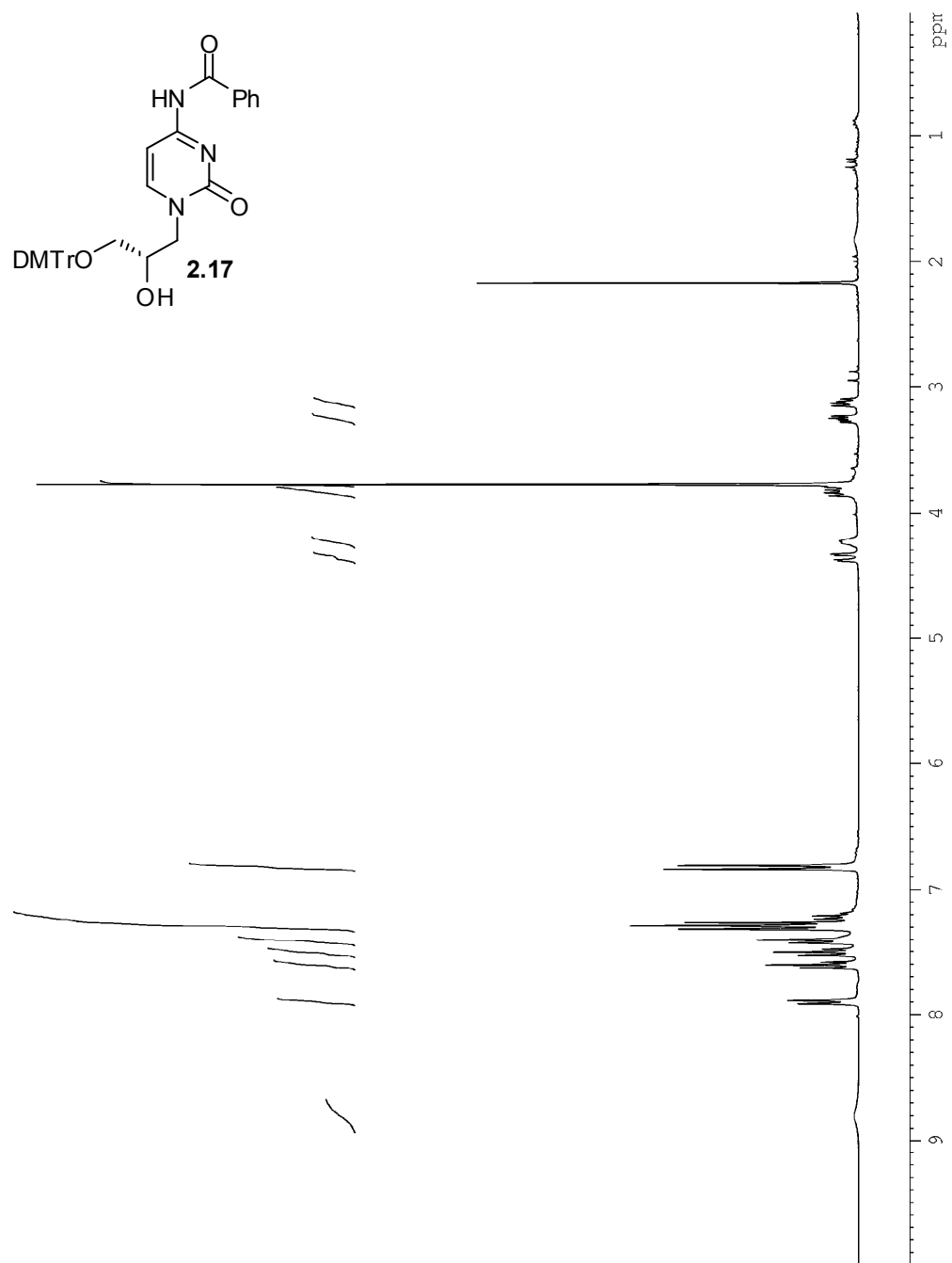


Figure A2.10.1. ^1H NMR spectrum of compound **2.17** (300 MHz, CDCl_3).

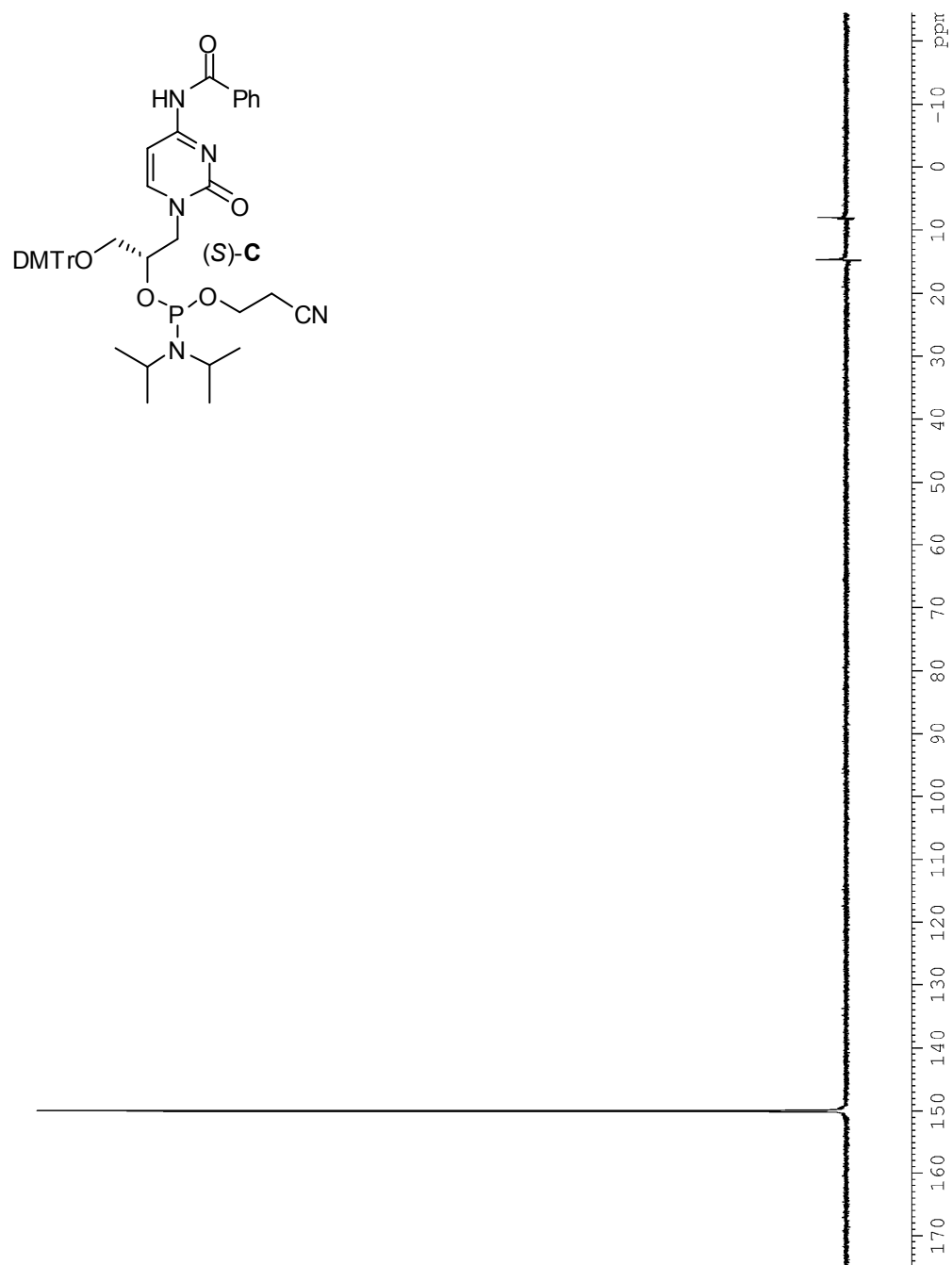


Figure A2.11.1. ^{31}P NMR spectrum of compound (S)-C (121 MHz, CDCl_3).

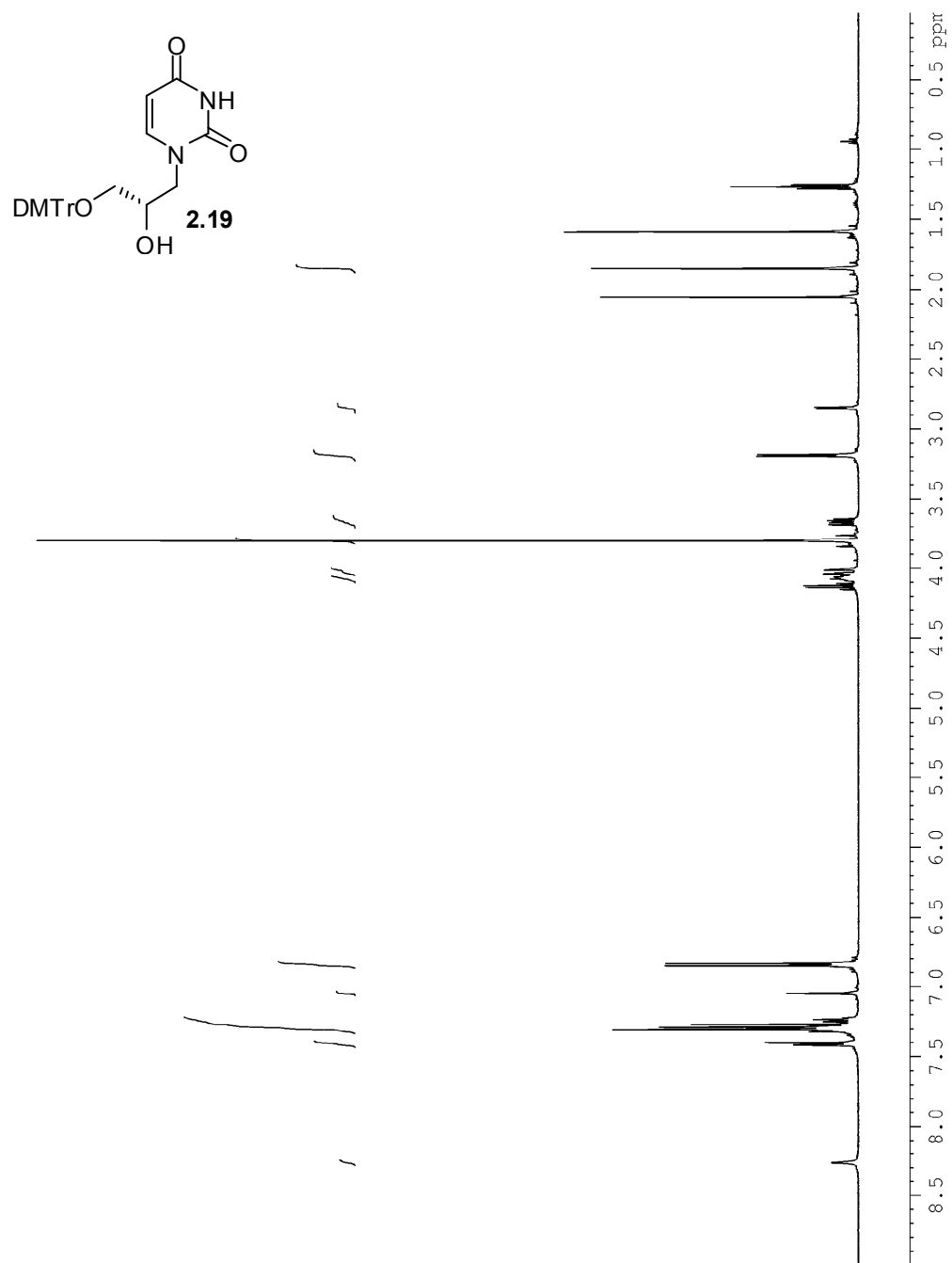


Figure A2.12.1. ^1H NMR spectrum of compound **2.19** (500 MHz, CDCl_3).

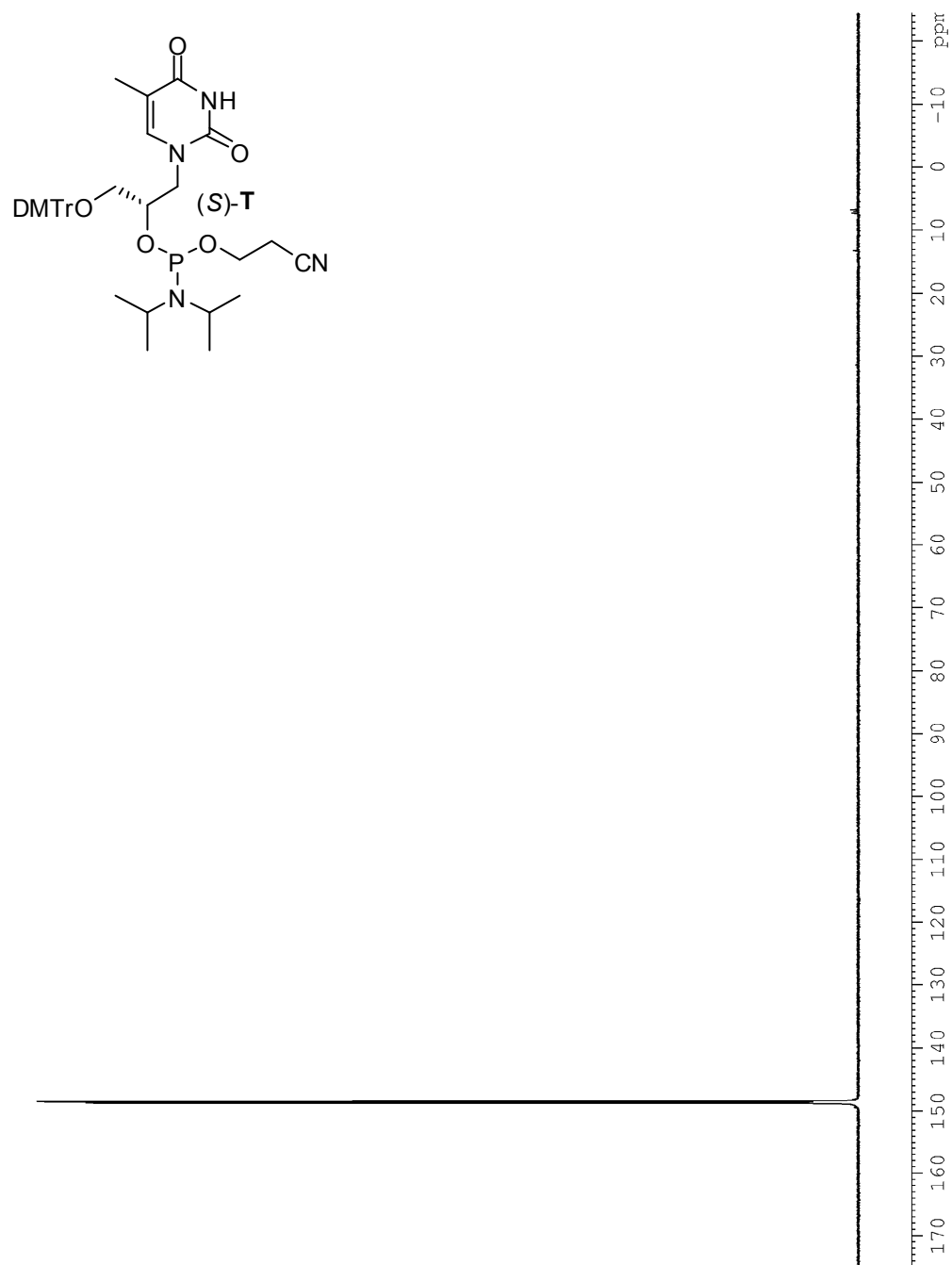


Figure A2.13.1. ^{31}P NMR spectrum of compound (S)-T (121 MHz, CDCl_3).

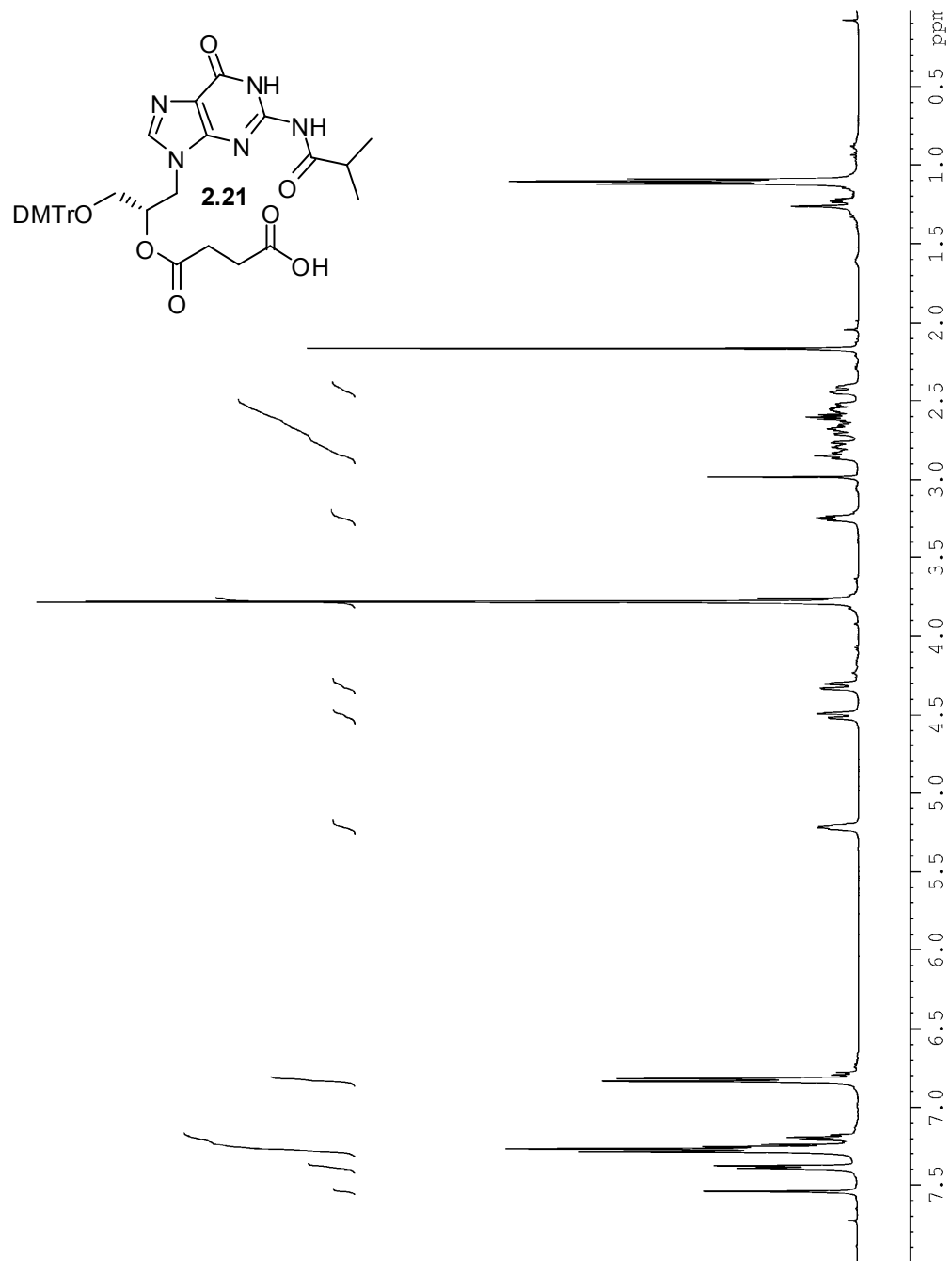


Figure A2.14.1. ^1H NMR spectrum of compound **2.21** (500 MHz, CDCl_3).

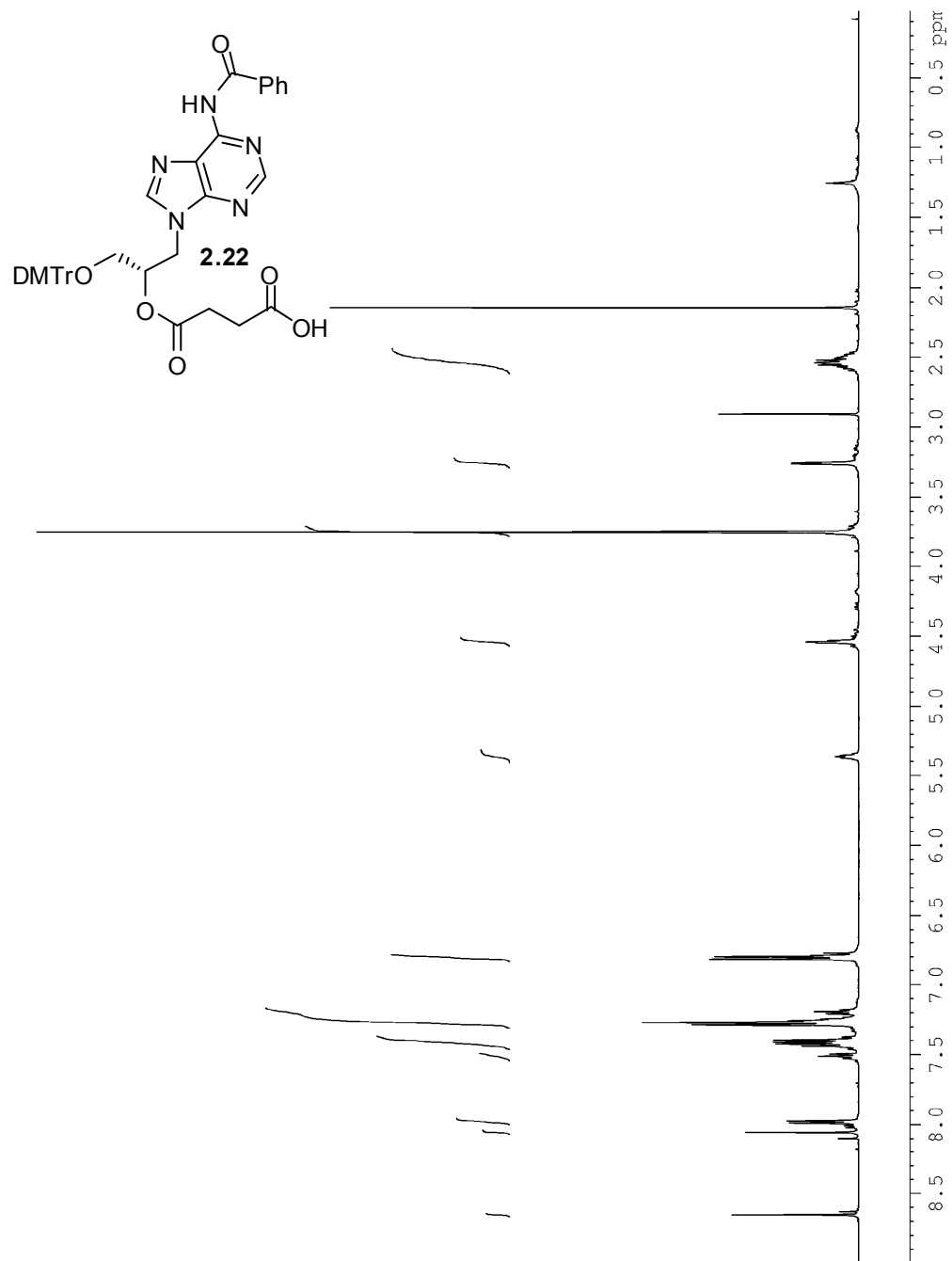


Figure A2.15.1. ^1H NMR spectrum of compound **2.22** (500 MHz, CDCl_3).

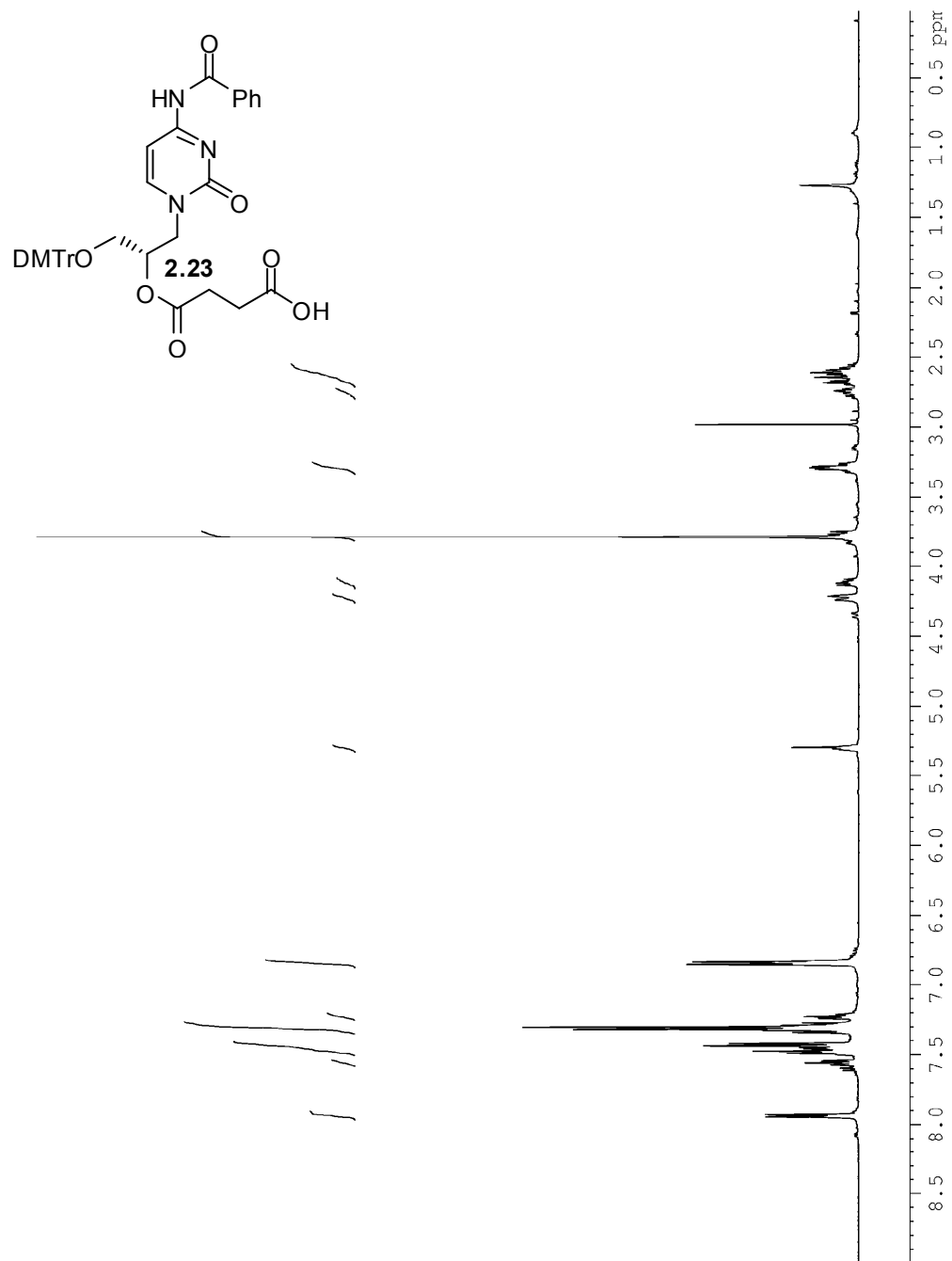


Figure A2.16.1. ^1H NMR spectrum of compound **2.23** (500 MHz, CDCl_3).

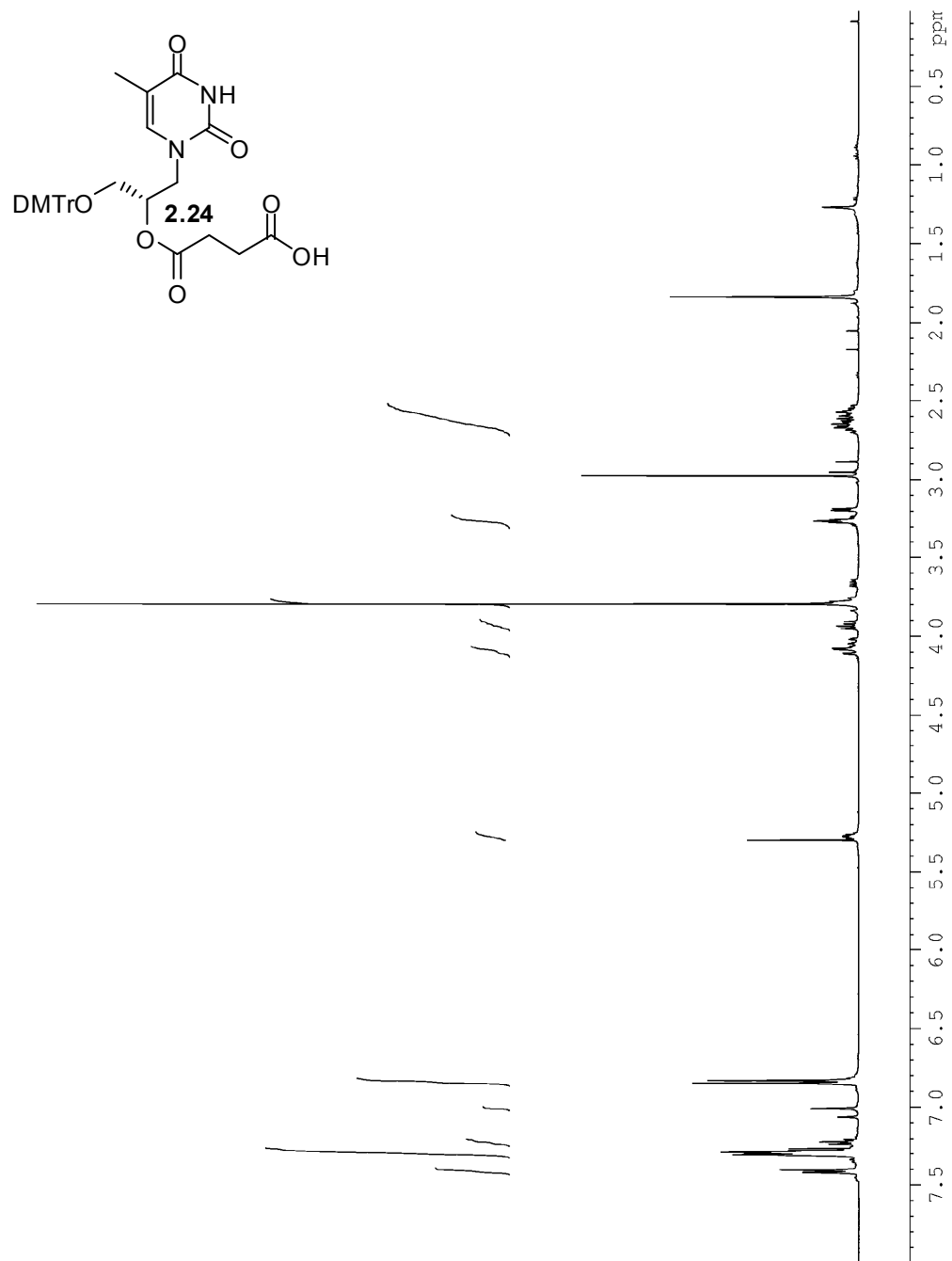


Figure A2.17.1. ^1H NMR spectrum of compound **2.24** (500 MHz, CDCl_3).

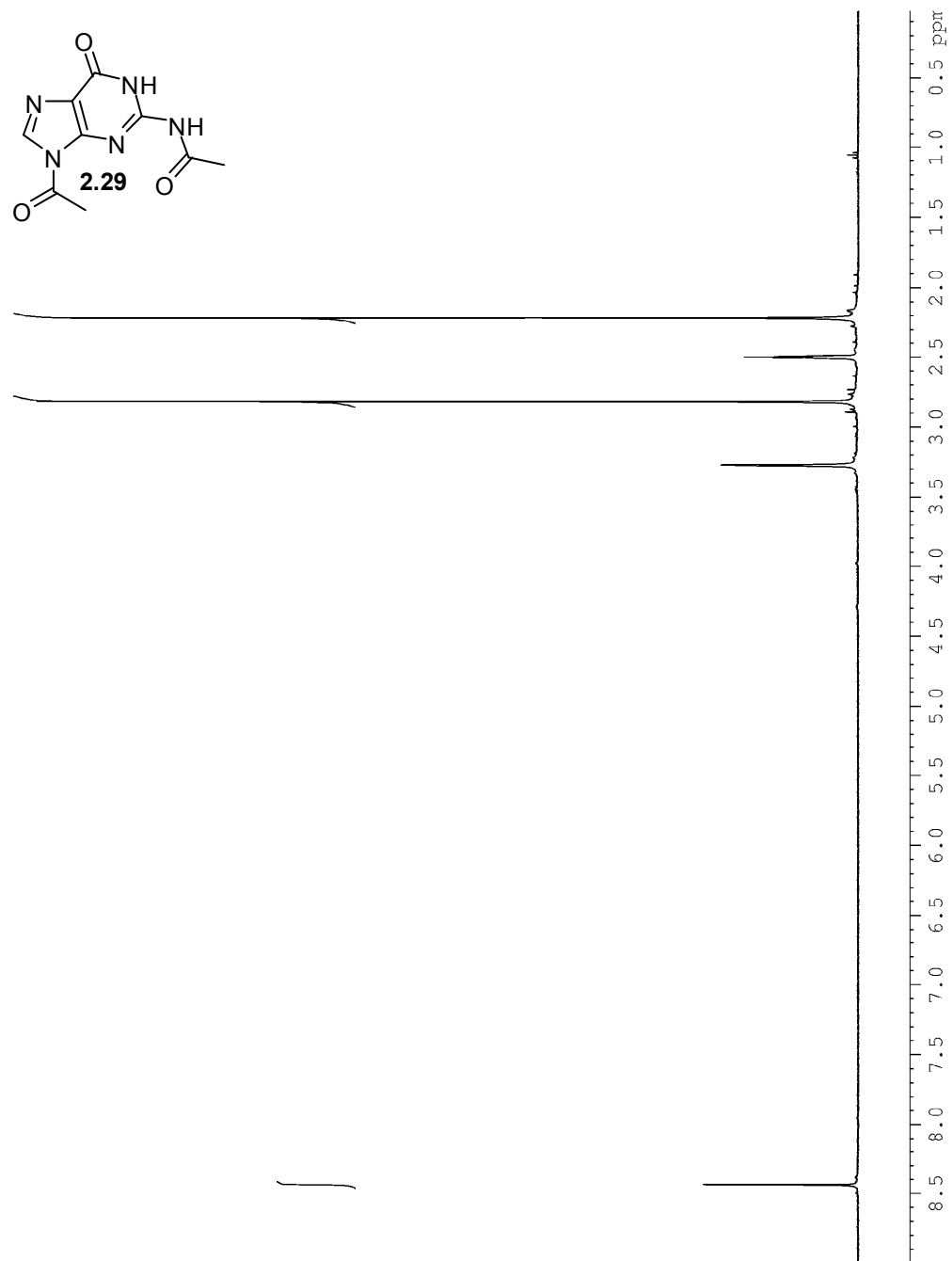


Figure A2.18.1. ^1H NMR spectrum of compound **2.29** in (360 MHz, $\text{DMSO-}d_6$).

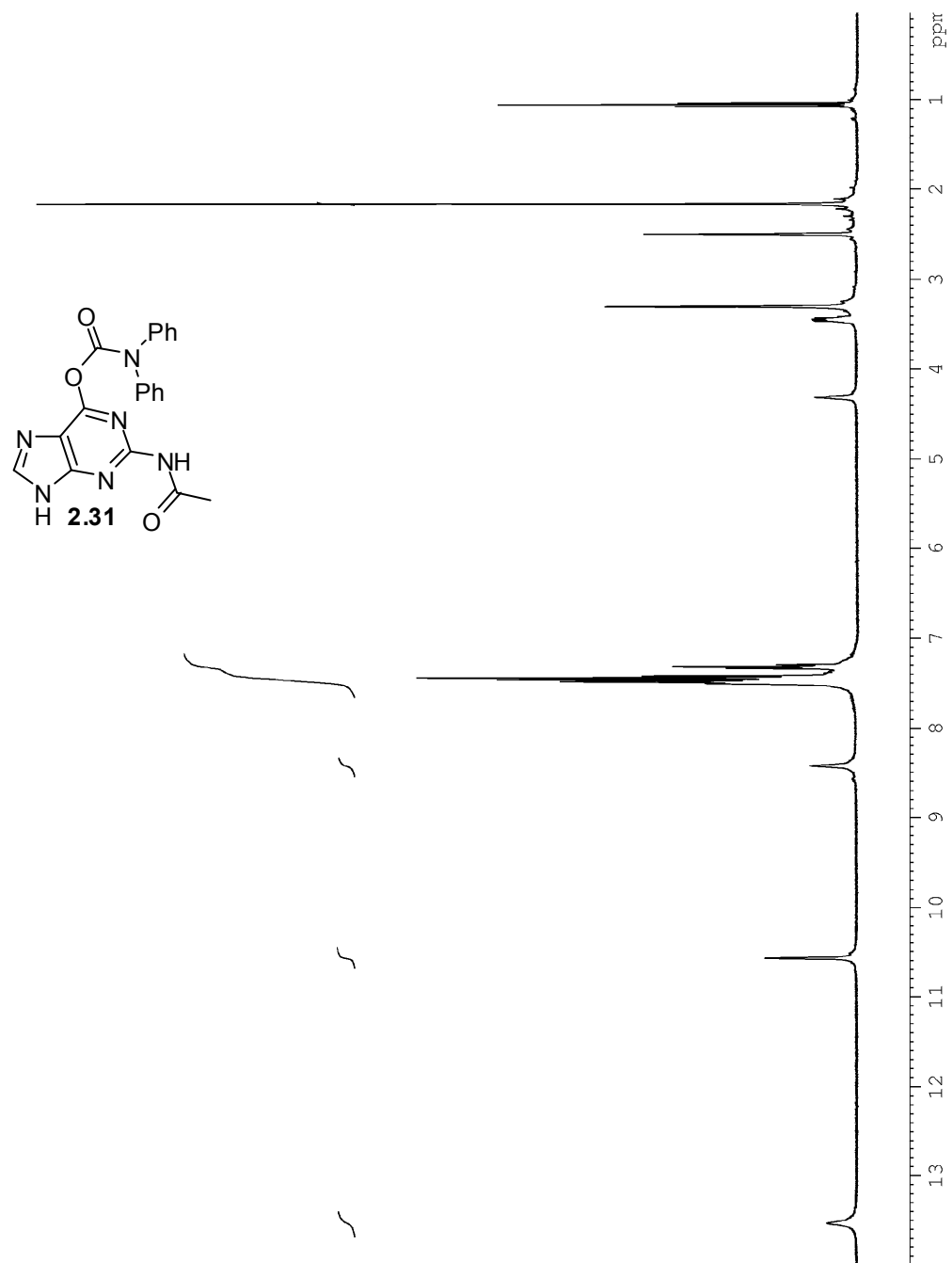


Figure A2.19.1. ^1H NMR spectrum of compound 2.31 (360 MHz, $\text{DMSO-}d_6$).

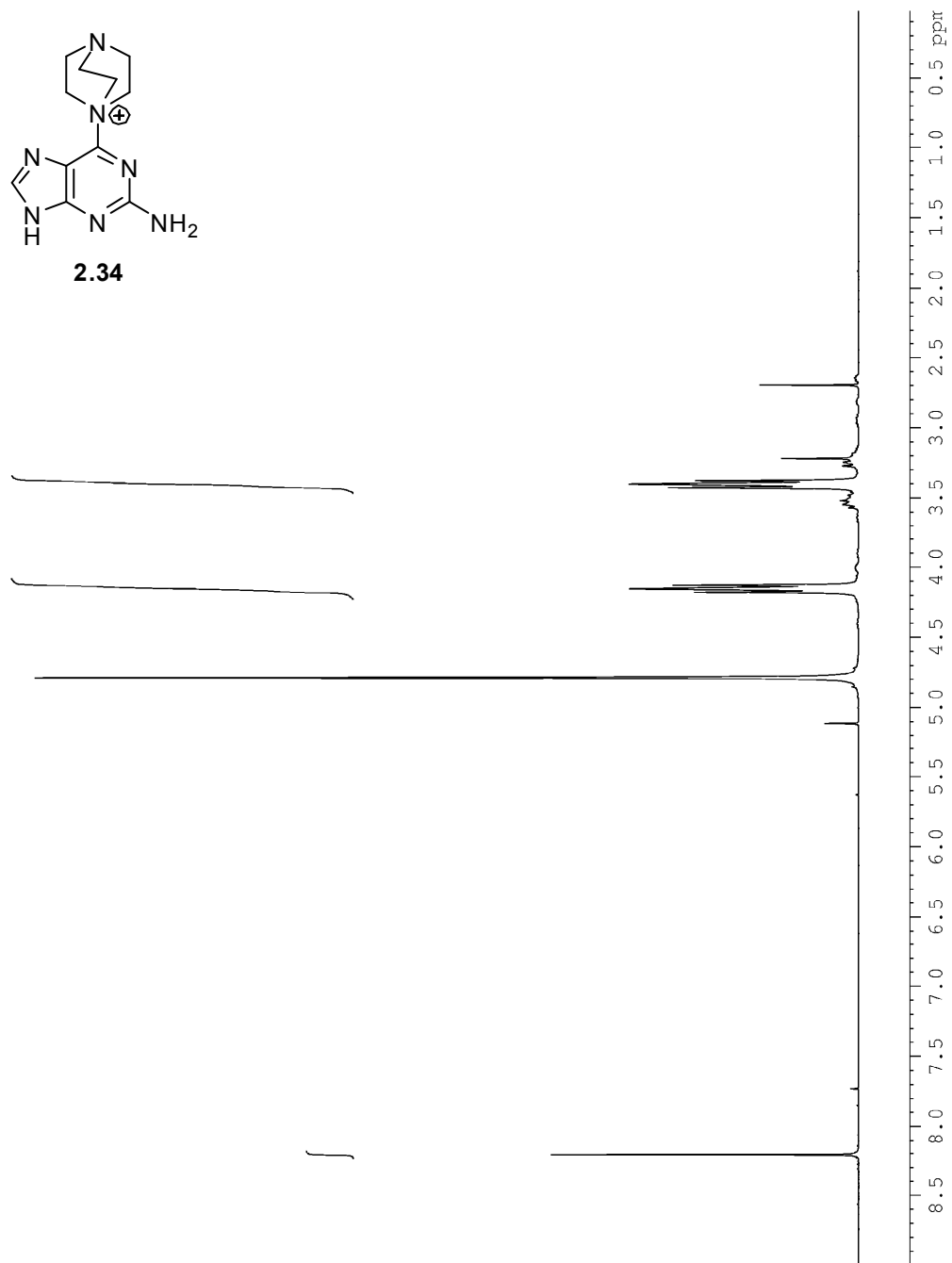


Figure A2.20.1. ¹H NMR spectrum of compound **2.34** (300 MHz, D₂O).

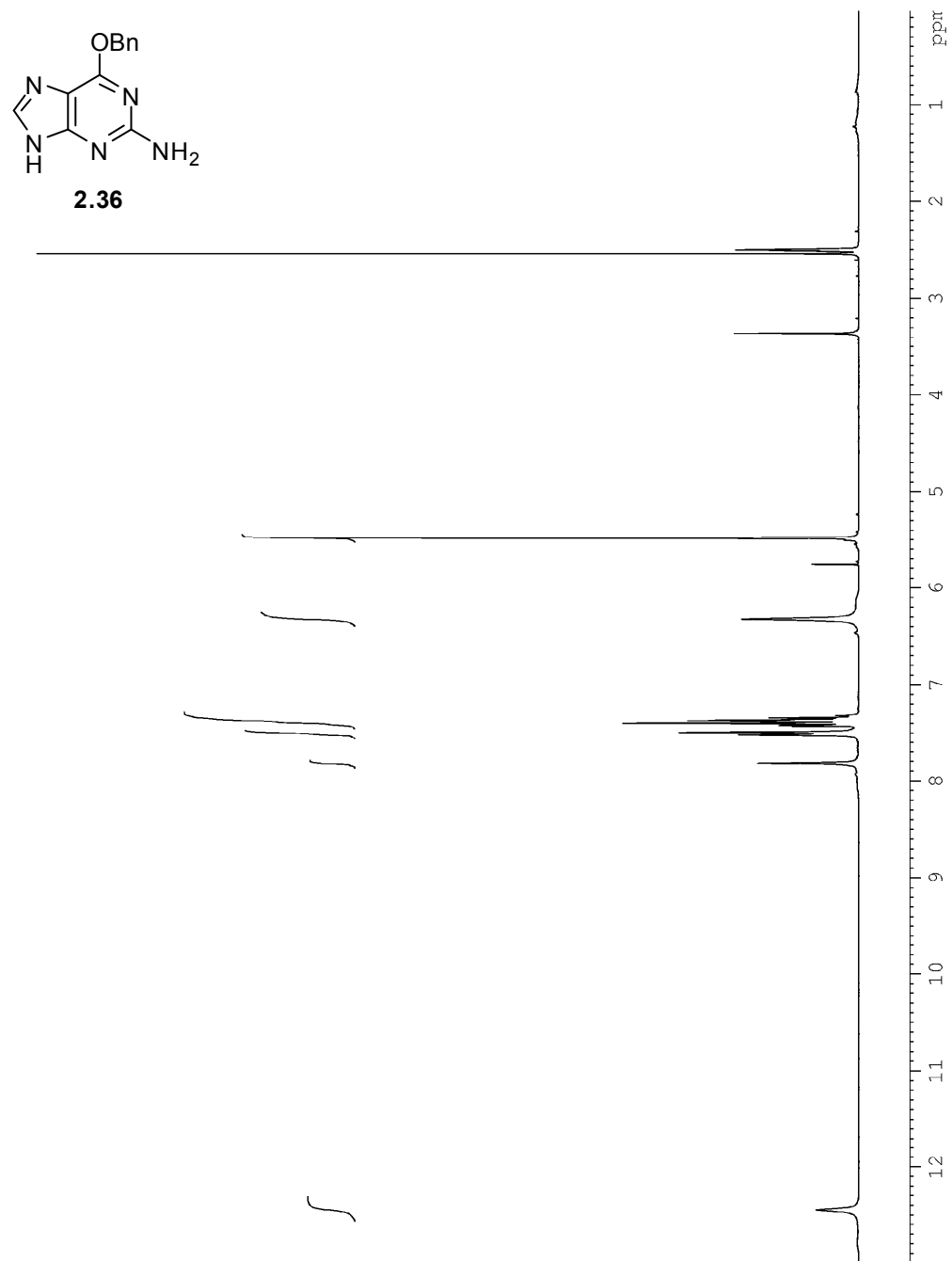


Figure A2.21.1. ^1H NMR spectrum of compound **2.36** (300 MHz, $\text{DMSO-}d_6$).

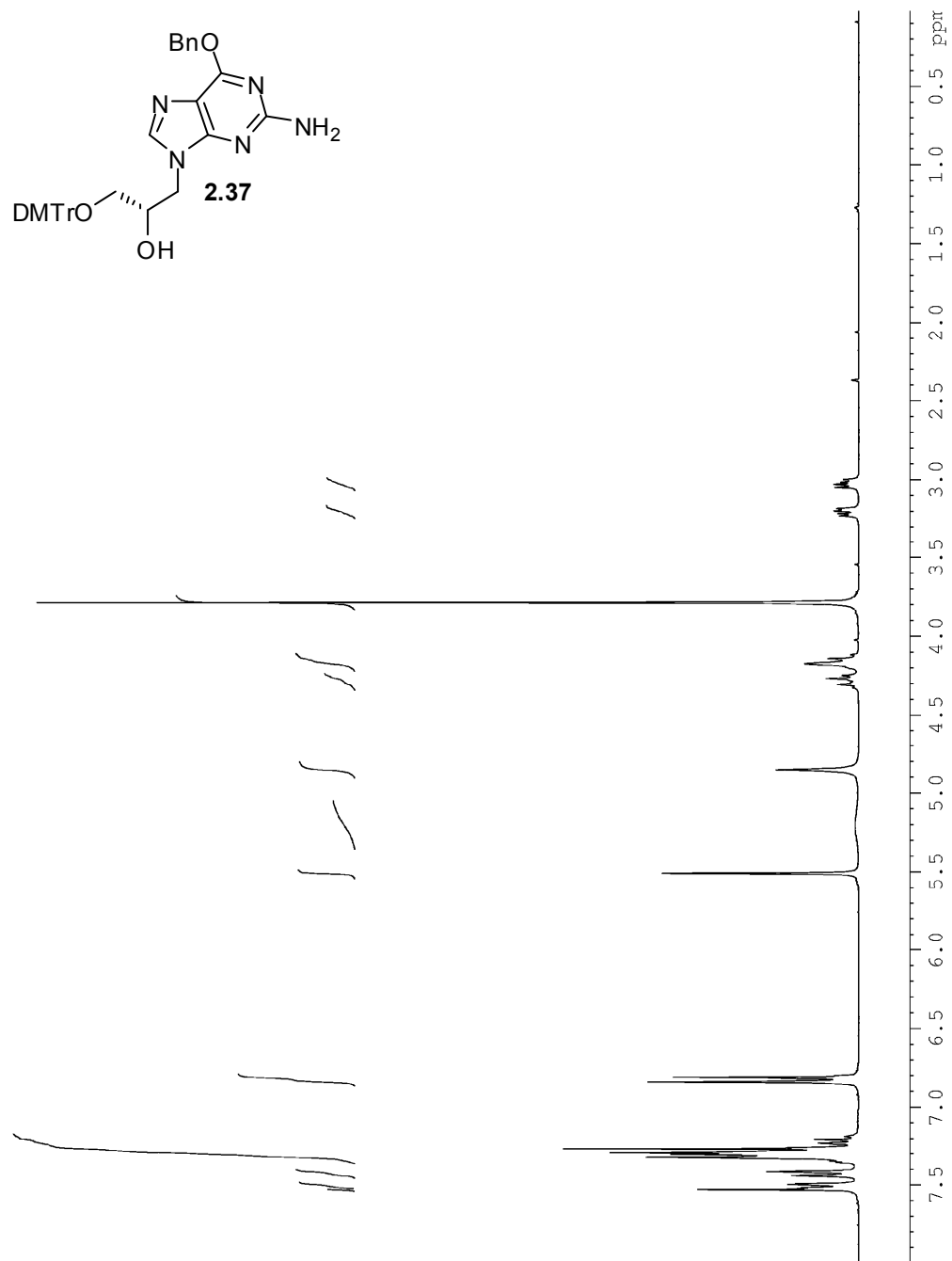


Figure A2.22.1. ^1H NMR spectrum of compound **2.37** (500 MHz, CDCl_3).

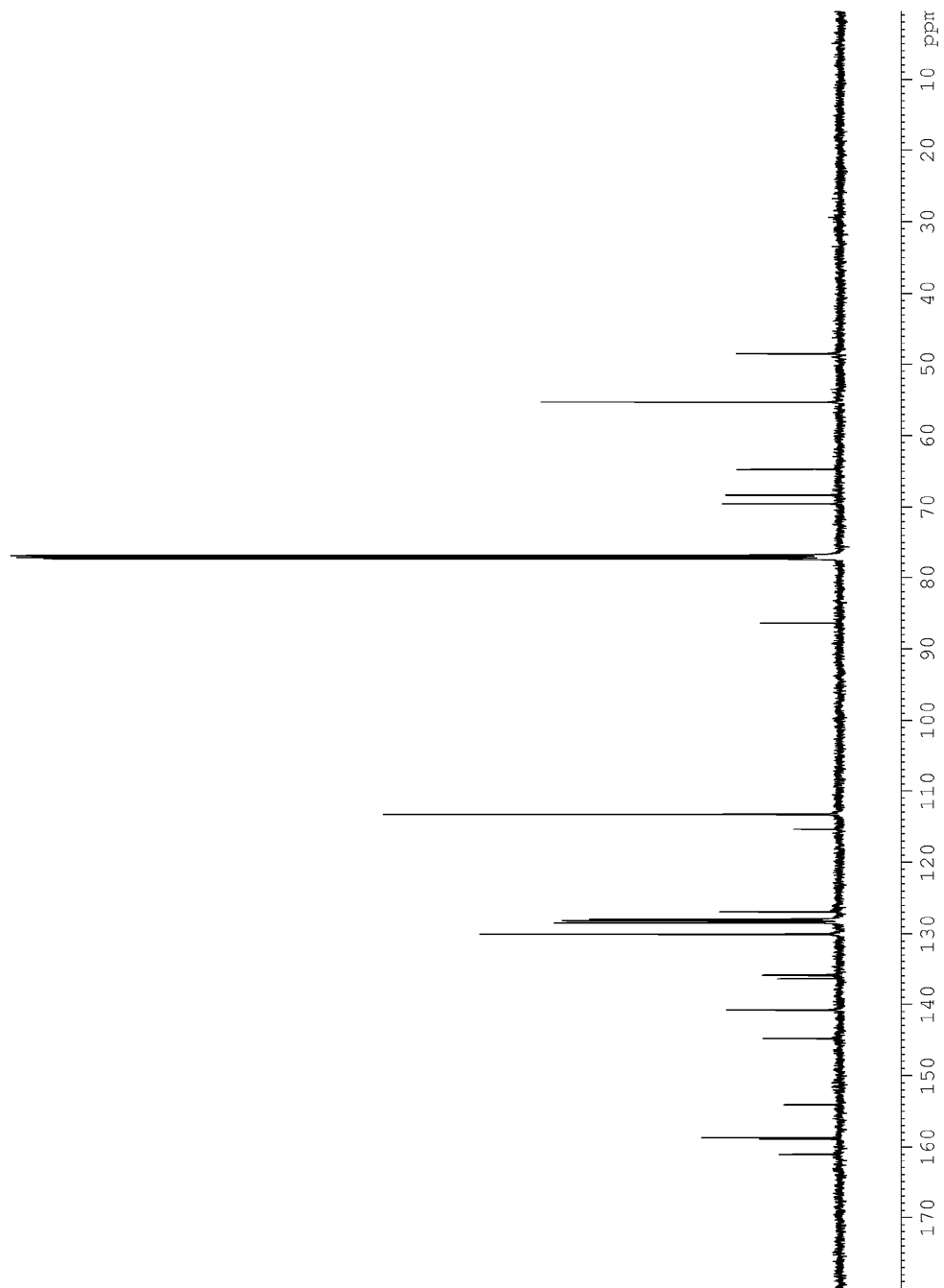


Figure A2.22.2. ^{13}C NMR spectrum of compound **2.37** (125 MHz, CDCl_3).

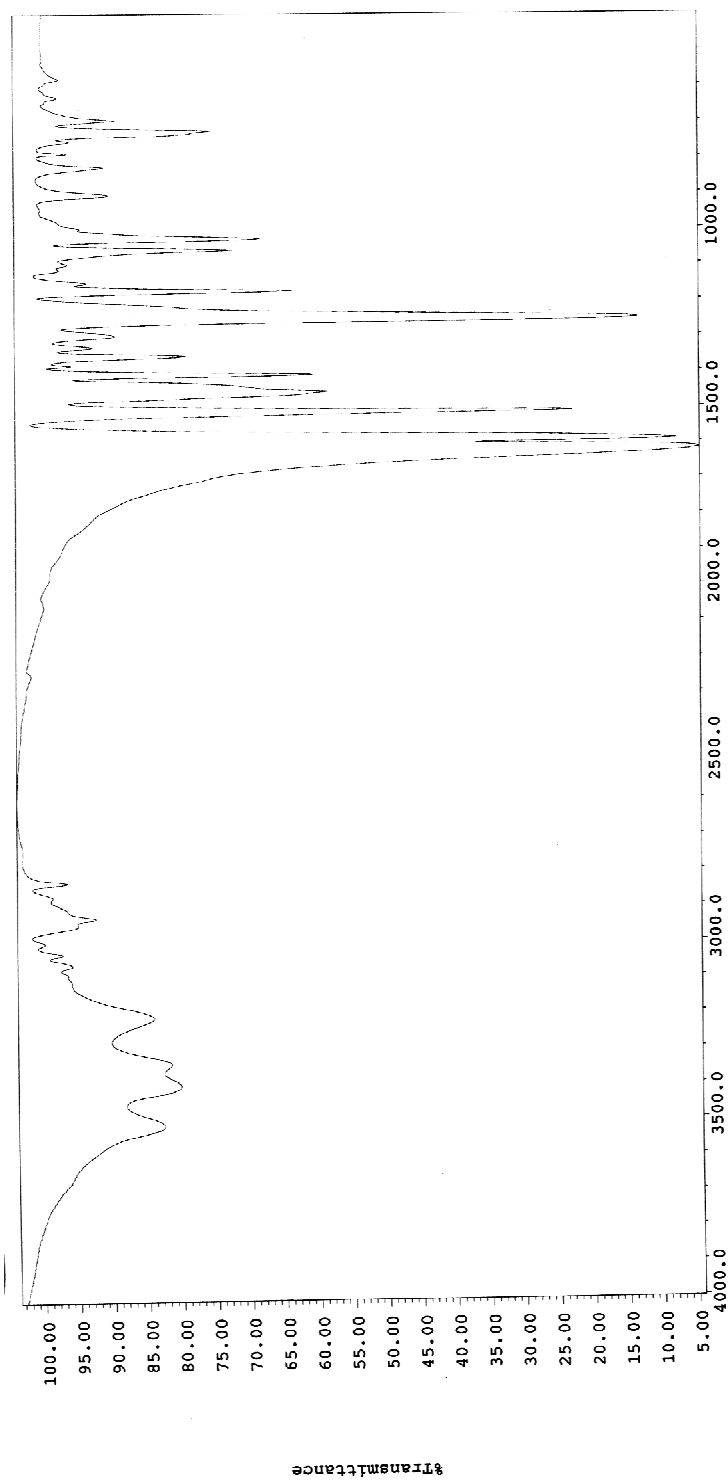


Figure A2.22.3. IR spectrum of compound 2.37 (film).

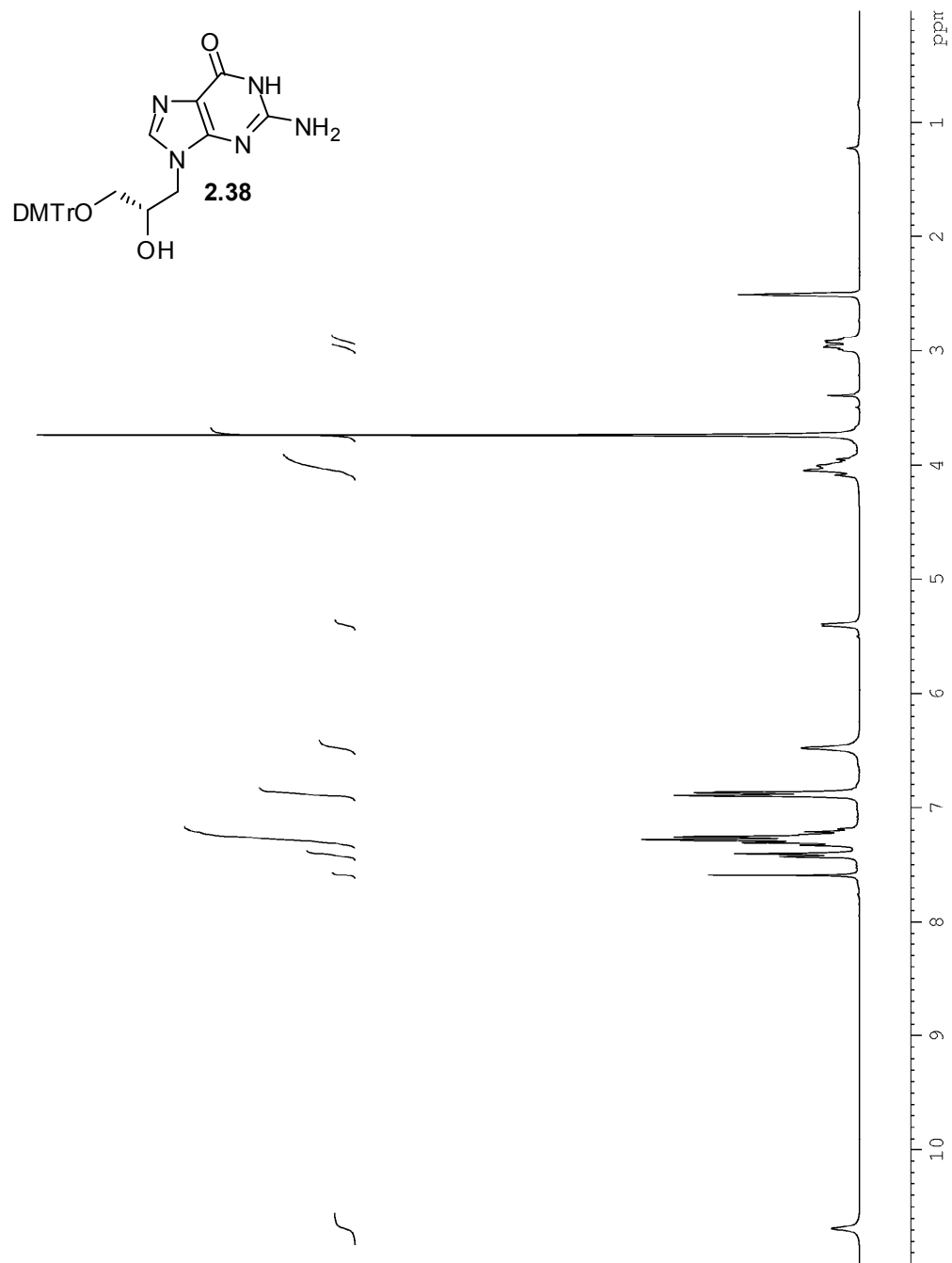


Figure A2.23.1. ^1H NMR spectrum of compound **2.38** (300 MHz, $\text{DMSO}-d_6$).

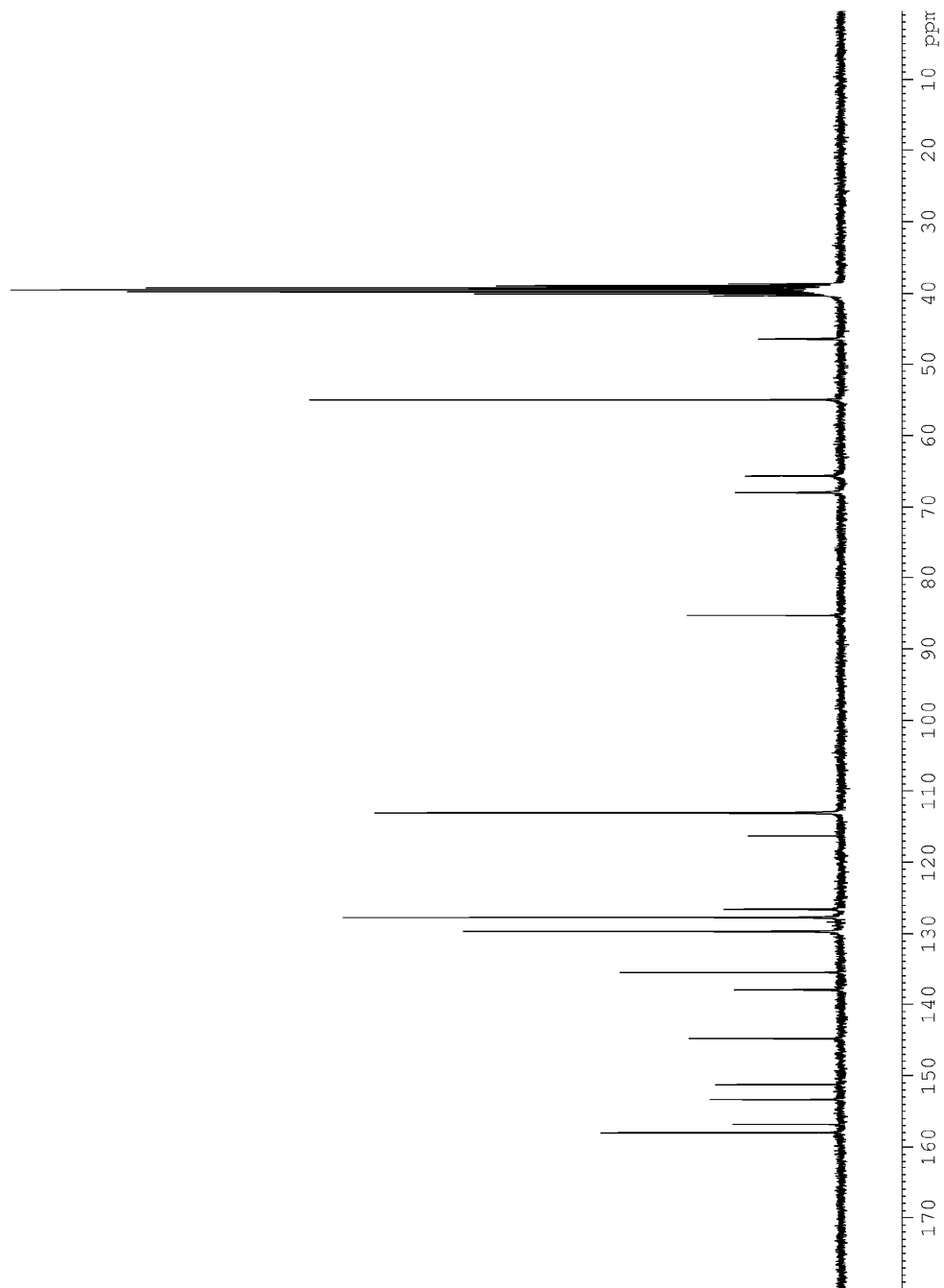


Figure A2.23.2. ^{13}C NMR spectrum of compound **2.38** (75 MHz, $\text{DMSO-}d_6$).

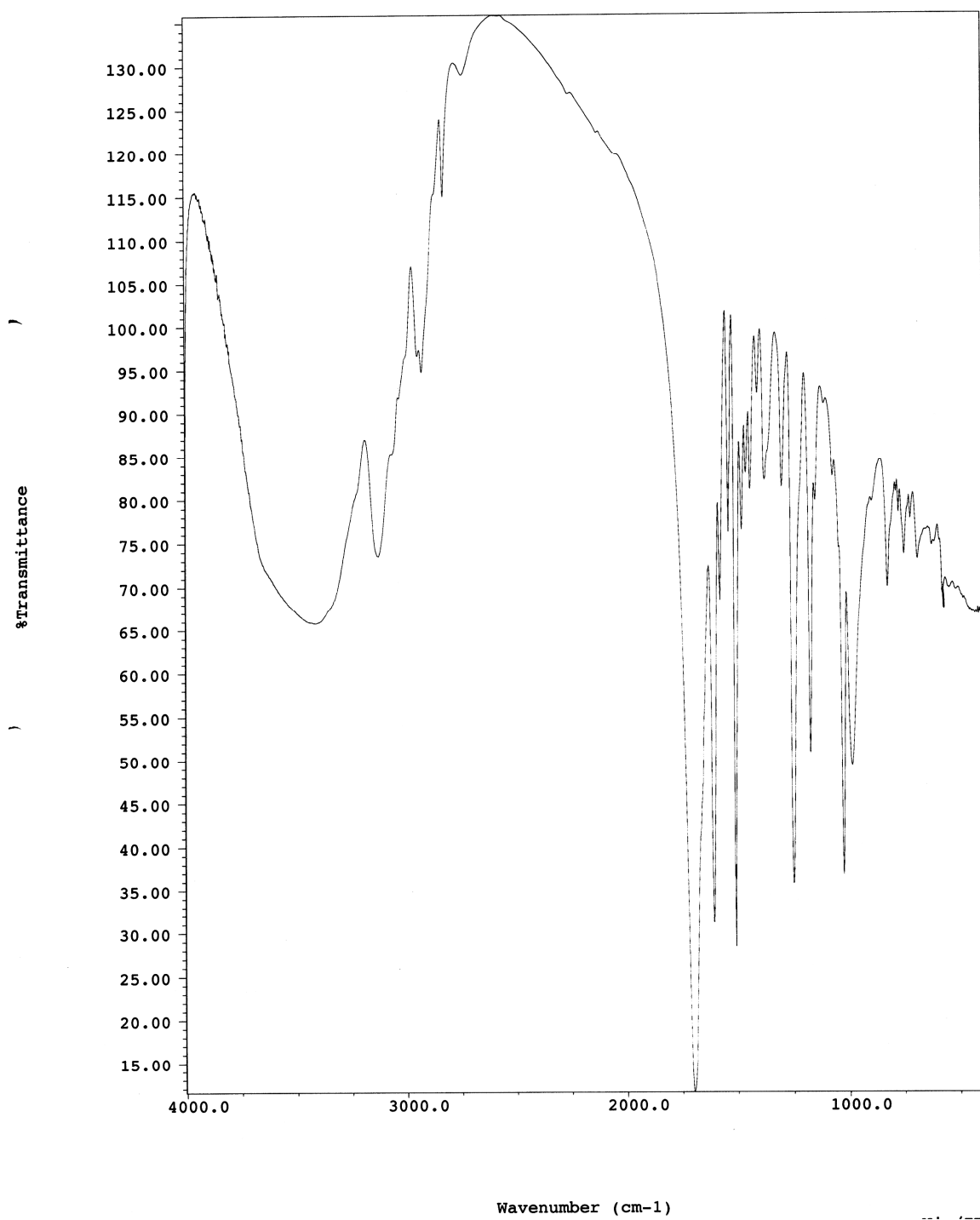


Figure A2.23.3. IR spectrum of compound **2.38** (film).

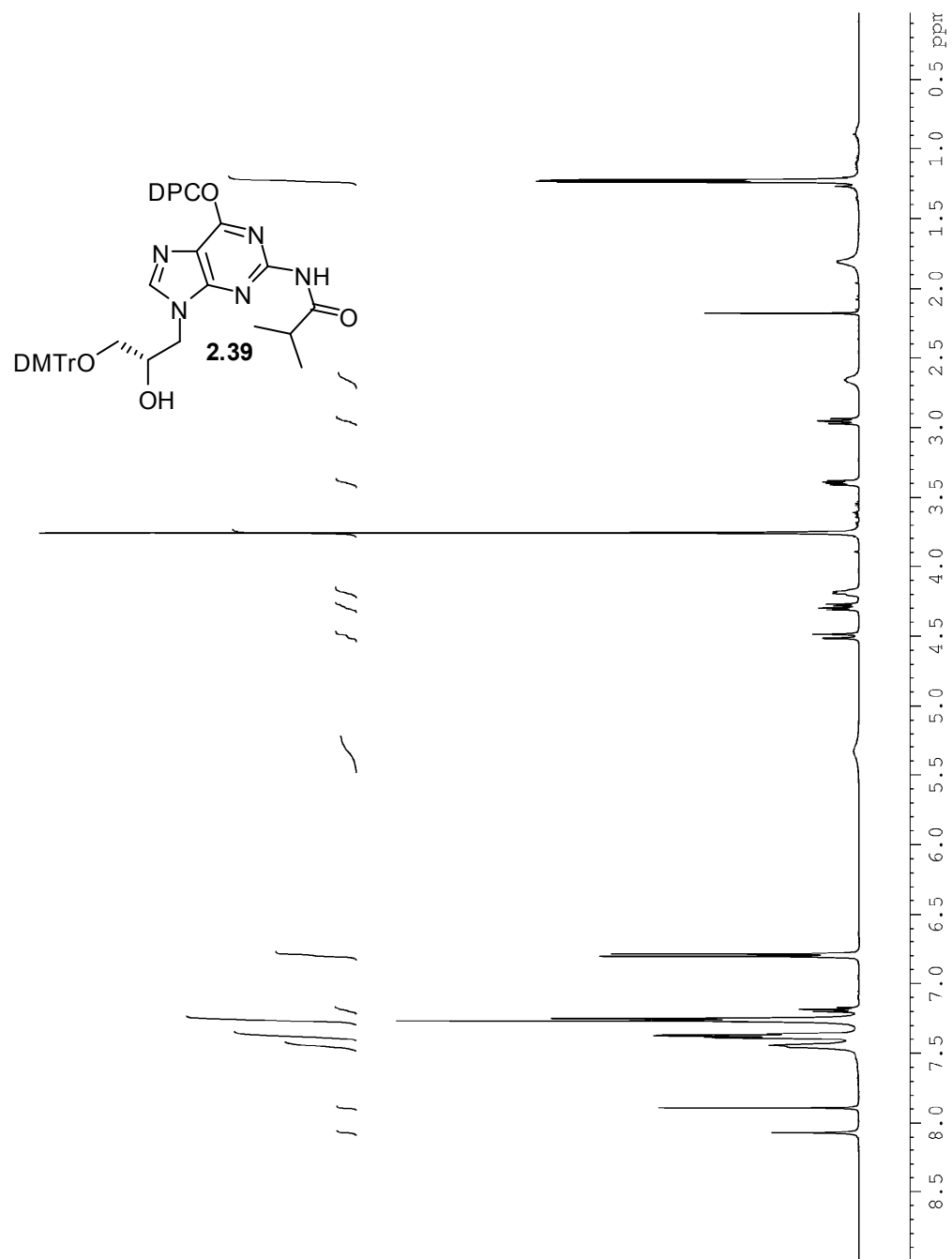


Figure A2.24.1. ^1H NMR spectrum of compound **2.39** (500 MHz, CDCl_3).

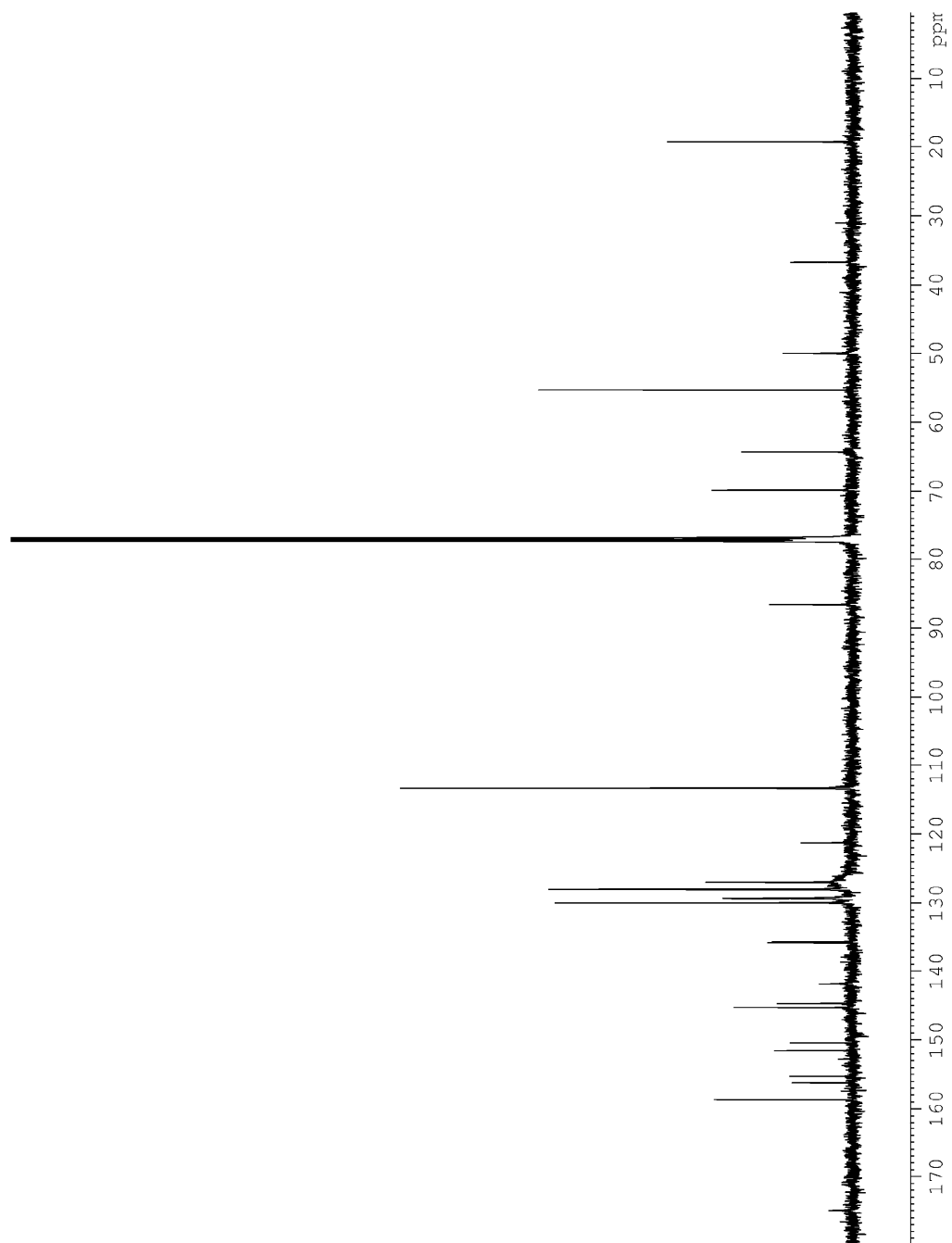


Figure A2.24.2. ^{13}C NMR spectrum of compound **2.39** (125 MHz, CDCl_3).

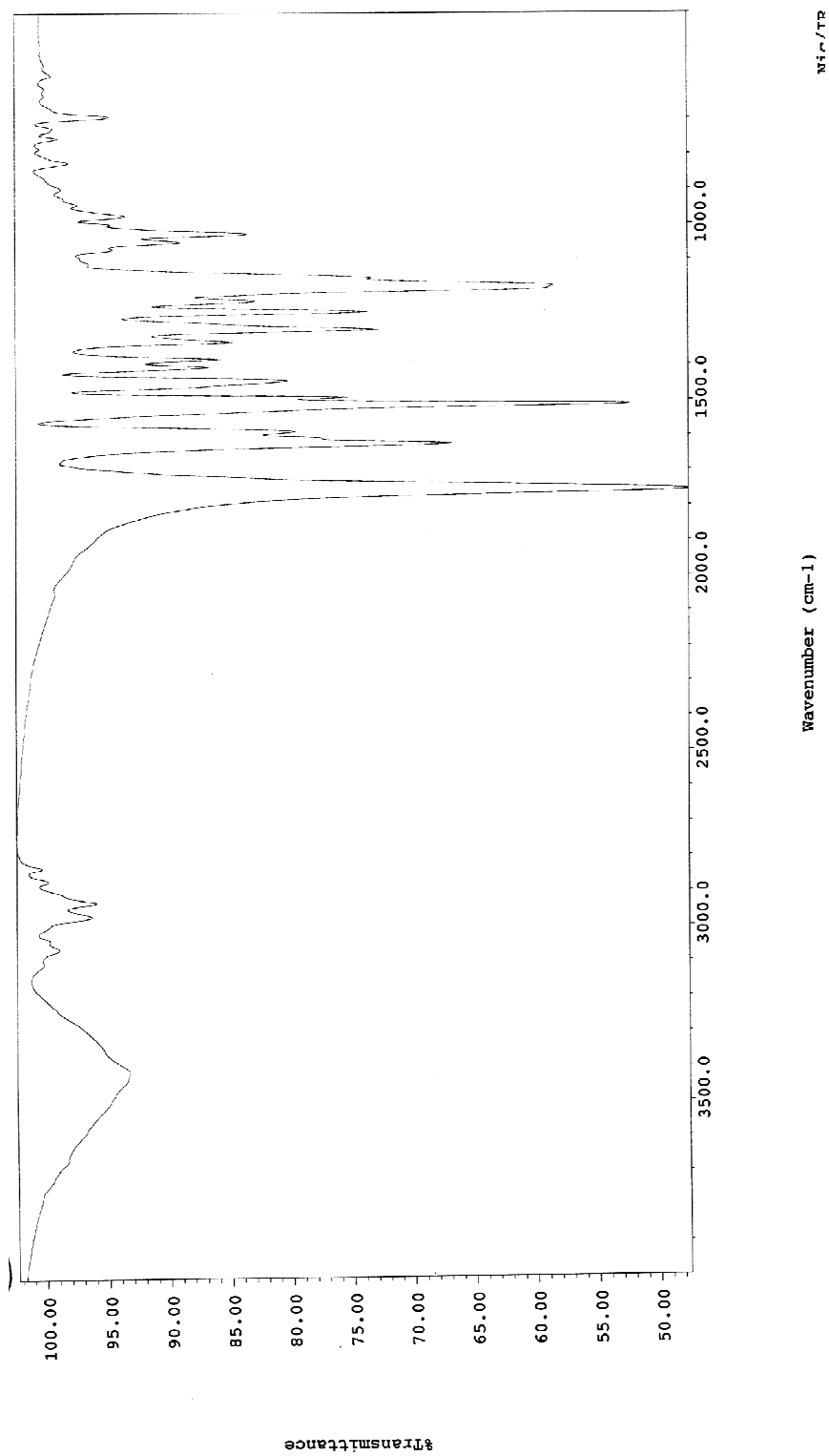


Figure A2.24.3. IR spectrum of compound 2.39 (film).

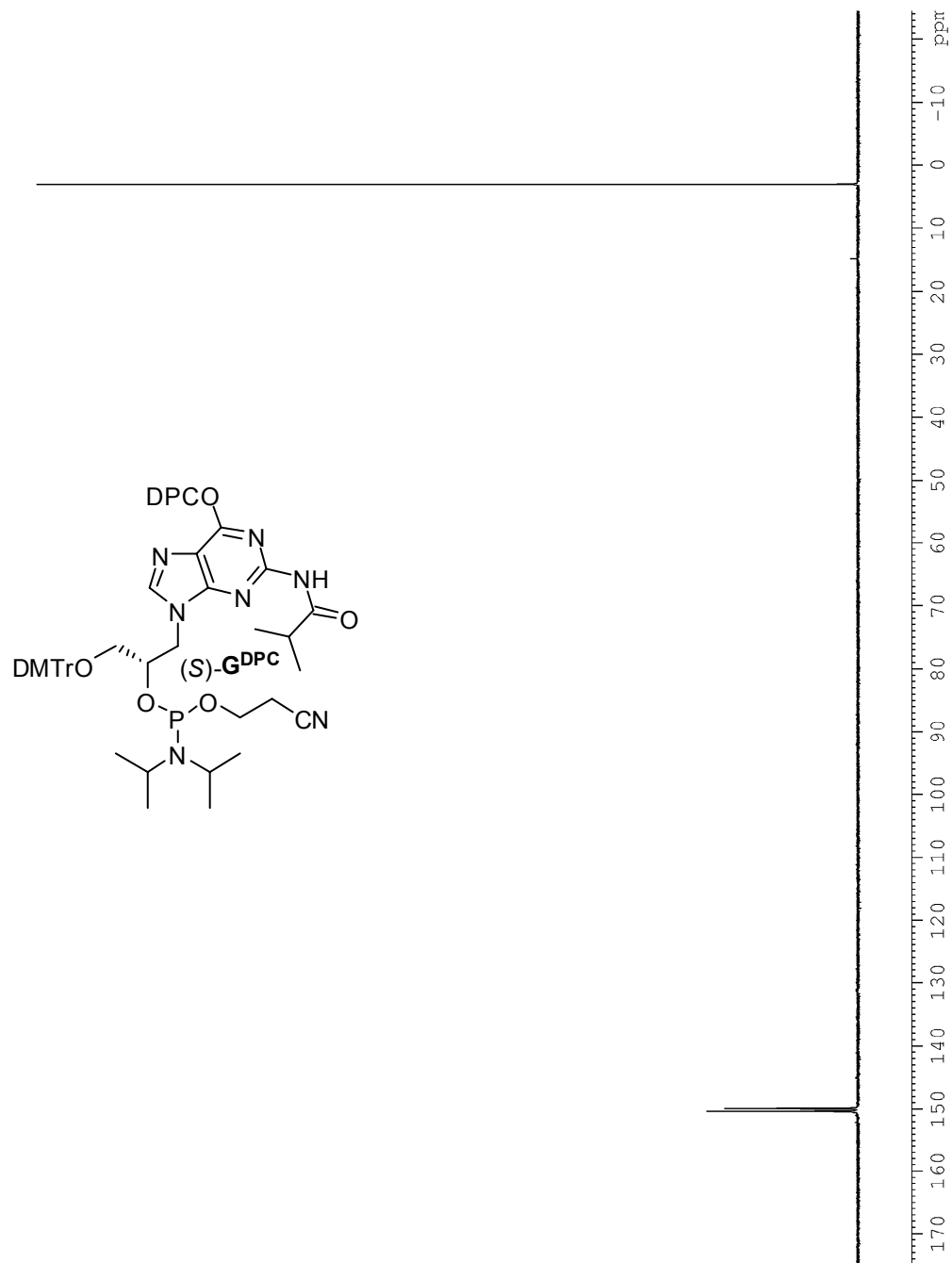


Figure A2.25.1. ^{31}P NMR spectrum of phosphoramidite (*S*)- G^{DPC} (121 MHz, CDCl_3) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).

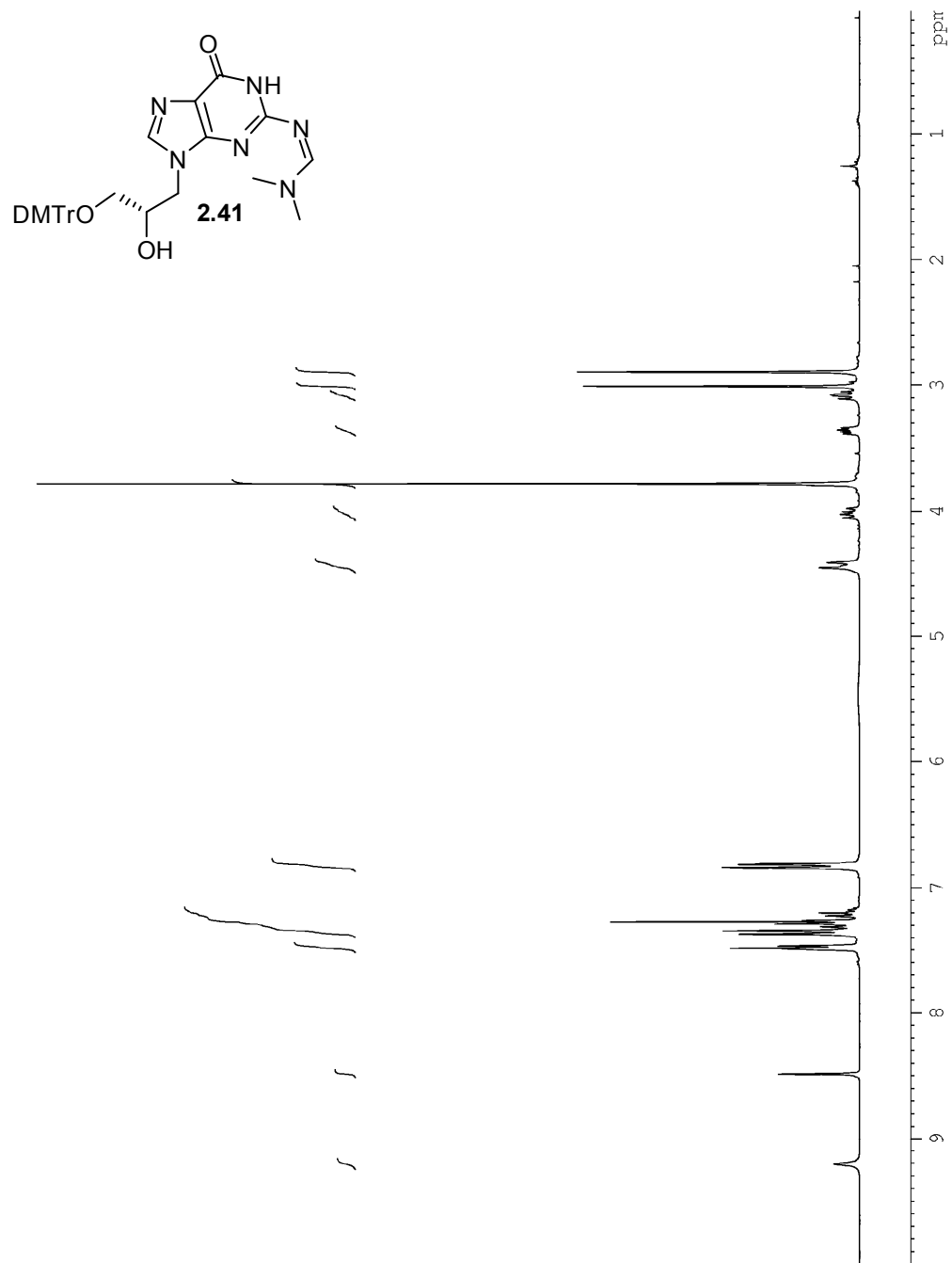


Figure A2.26.1. ^1H NMR spectrum of compound **2.41** (300 MHz, CDCl_3).

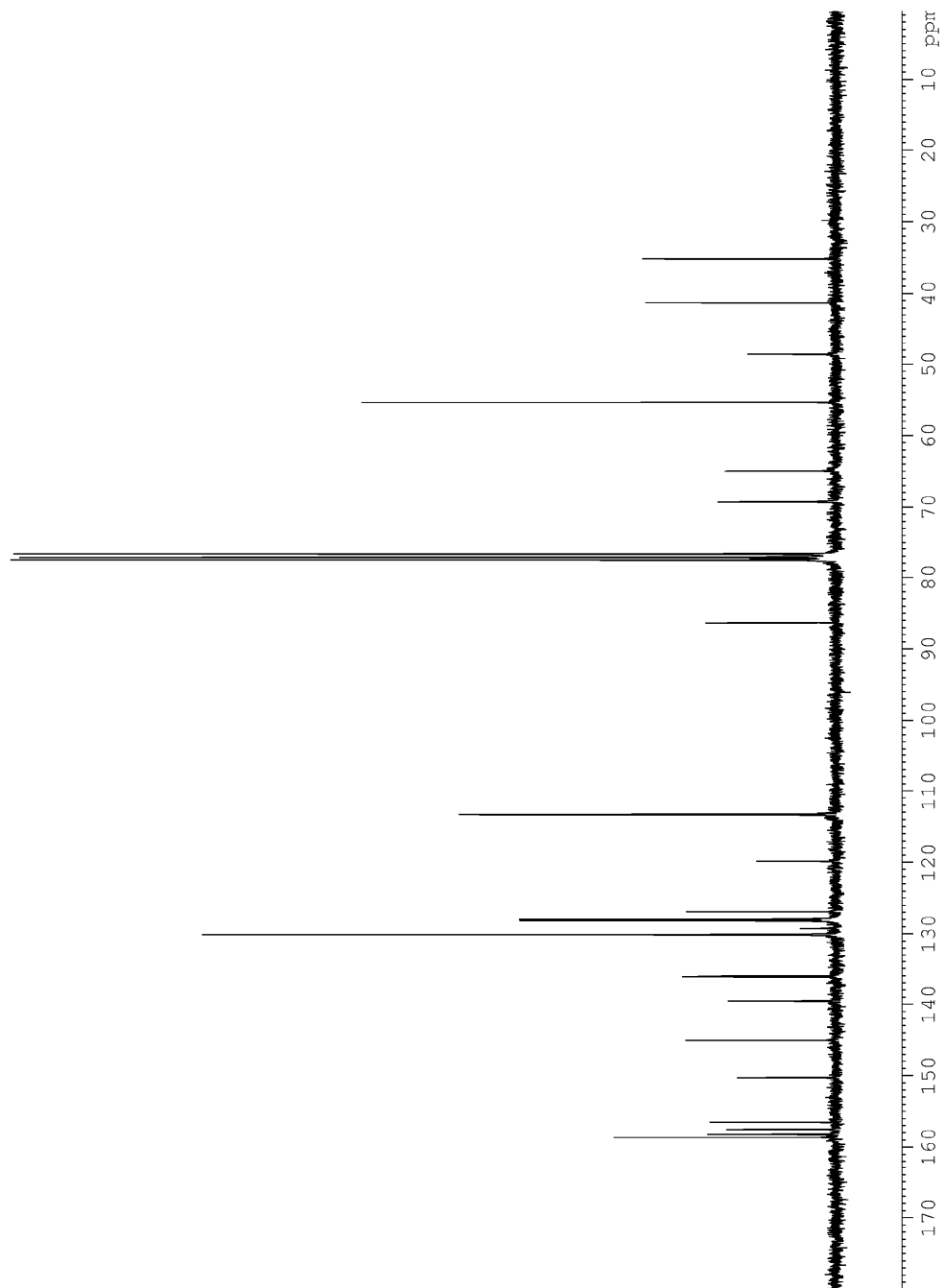


Figure A2.26.2. ^{13}C NMR spectrum of compound **2.41** (75 MHz, CDCl_3).

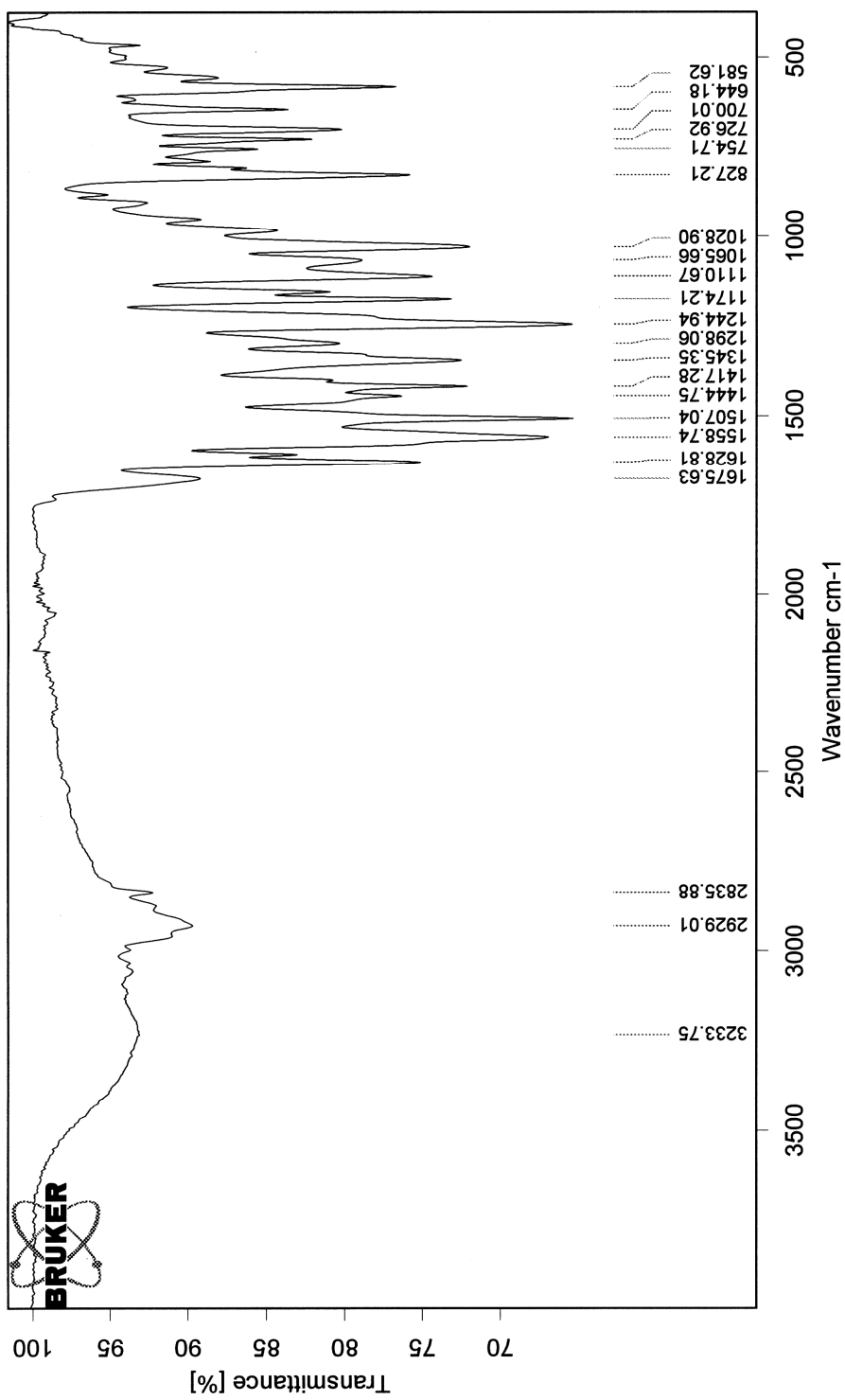


Figure A2.26.3. IR spectrum of compound 2.41 (solid).

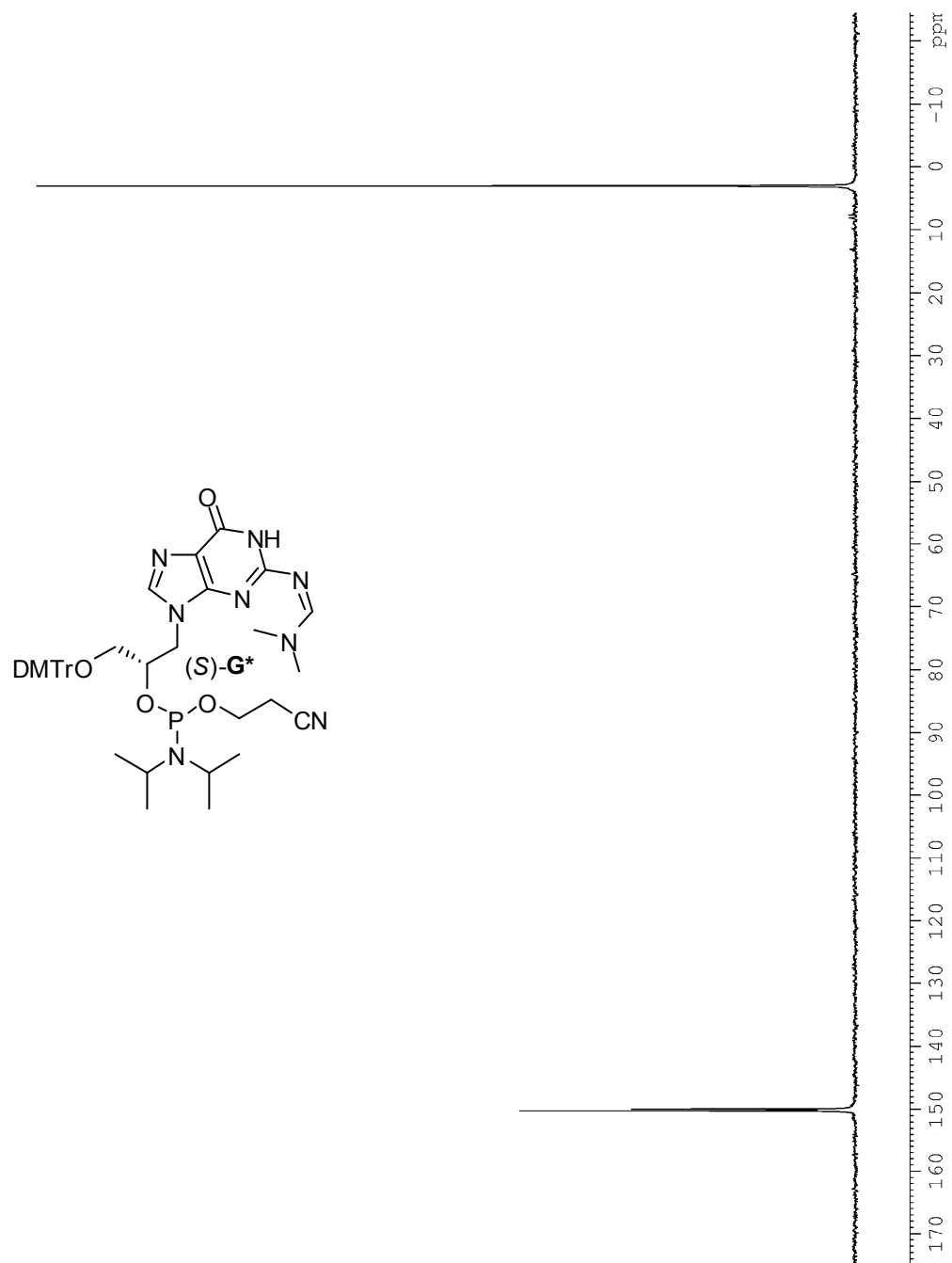


Figure A2.27.1. ³¹P NMR spectrum of phosphoramidite (S)-G* (162 MHz, CDCl₃) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).

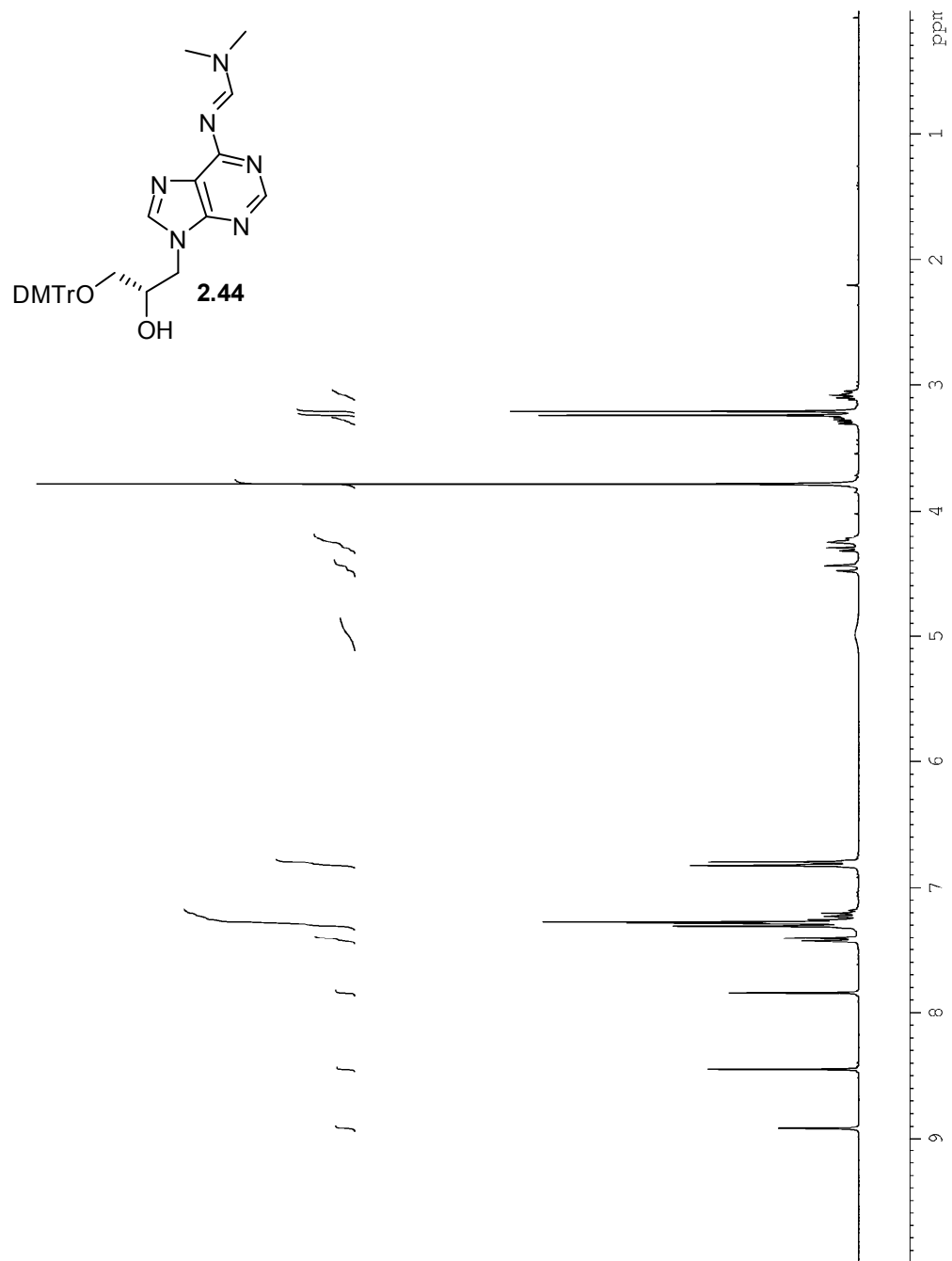


Figure A2.28.1. ^1H NMR spectrum of compound **2.44** (300 MHz, CDCl_3).

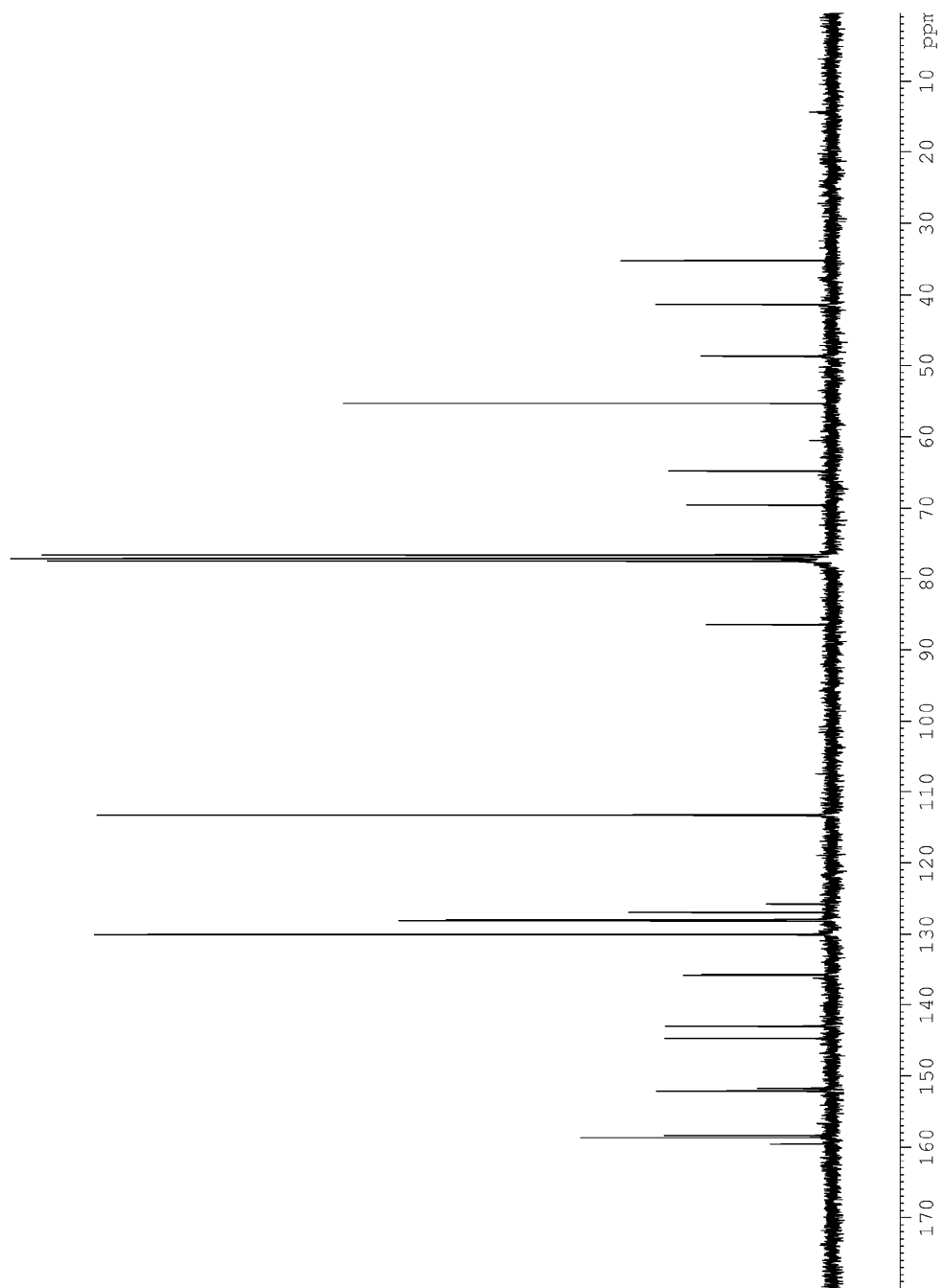


Figure A2.28.2. ^{13}C NMR spectrum of compound **2.44** (75 MHz, CDCl_3).

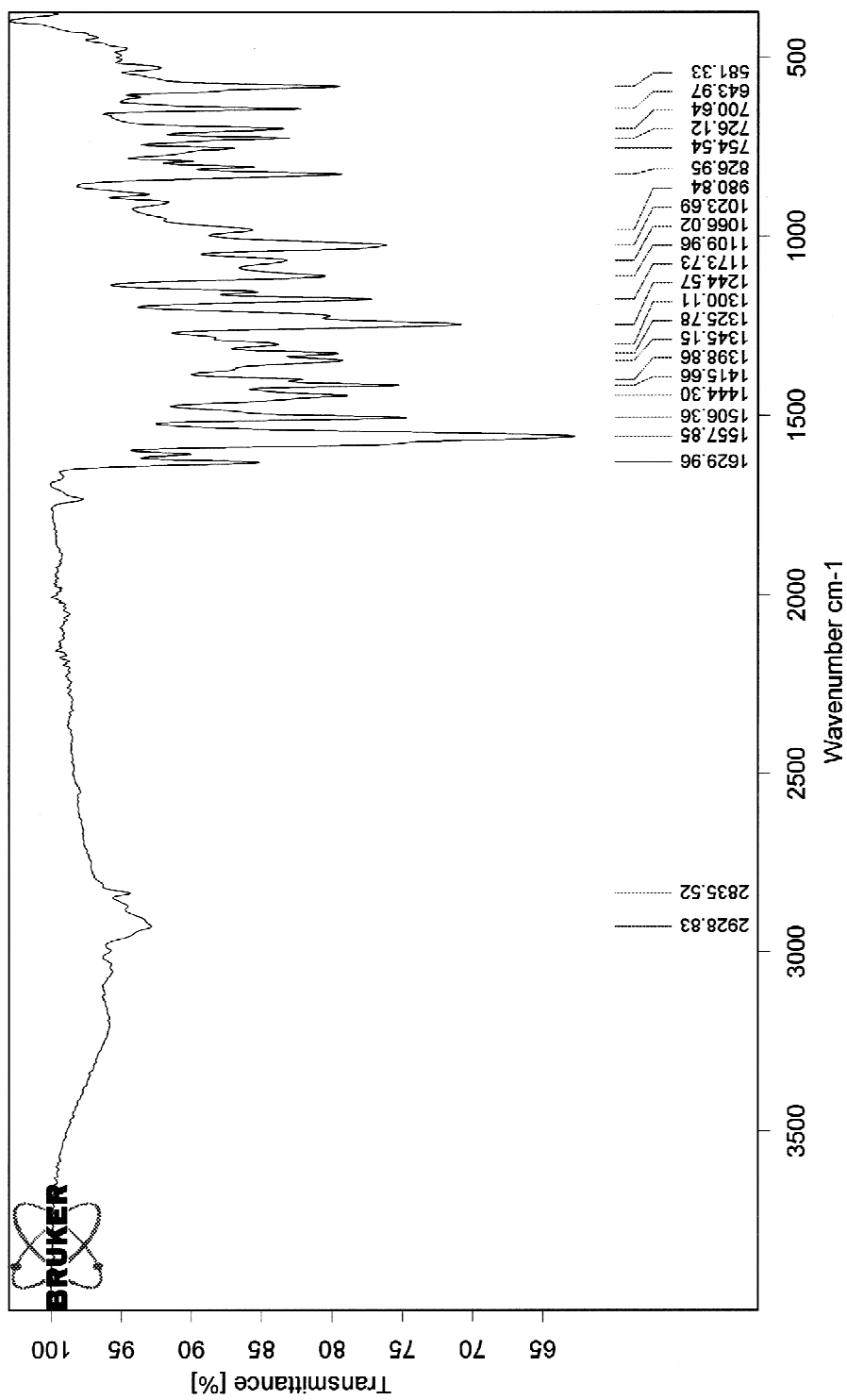


Figure A2.28.3. IR spectrum of compound 2.44 (solid).

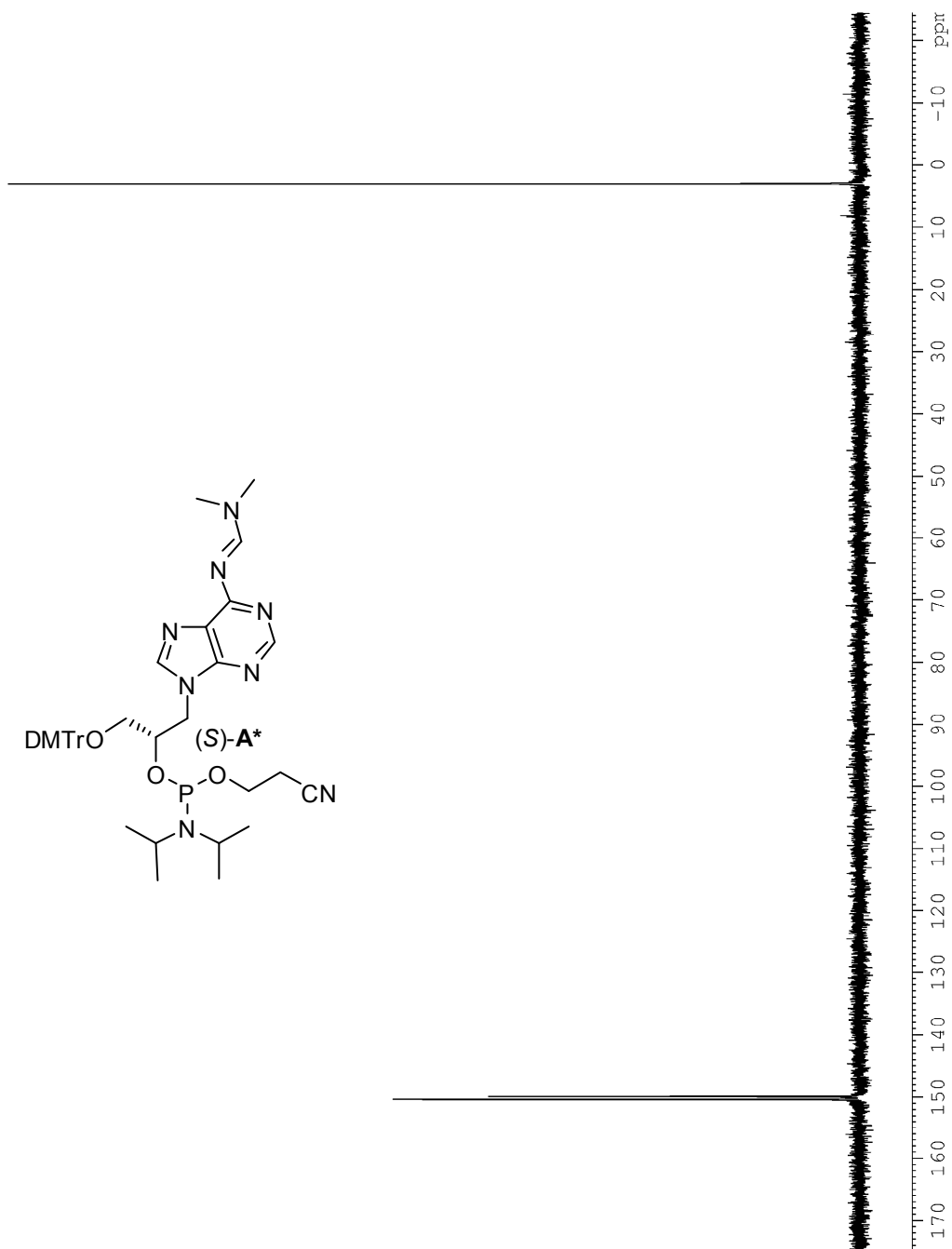


Figure A2.29.1. ³¹P NMR spectrum of phosphoramidite (S)-A* (162 MHz, CDCl₃) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).

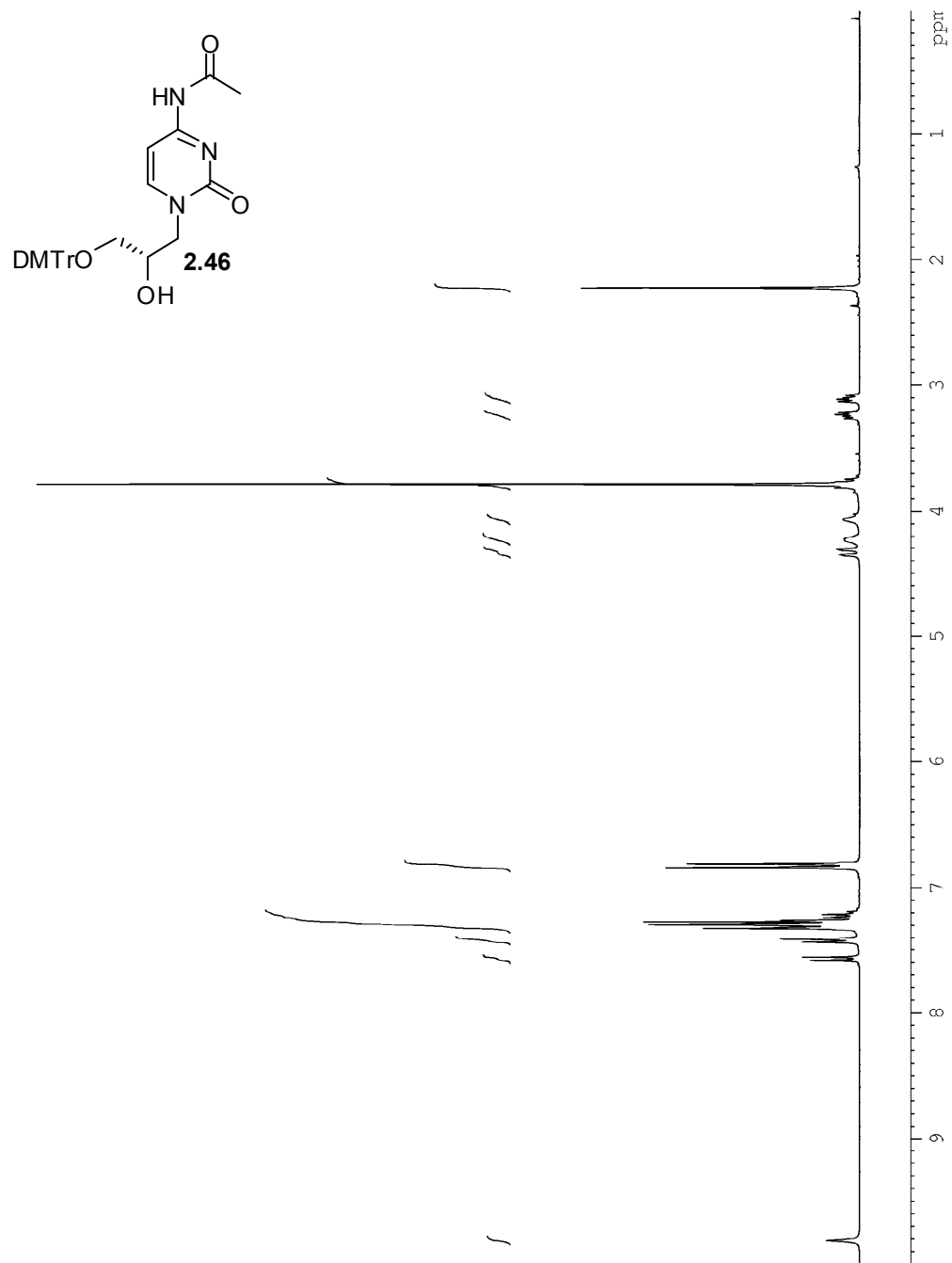


Figure A2.30.1. ^1H NMR spectrum of compound **2.46** (300 MHz, CDCl_3).

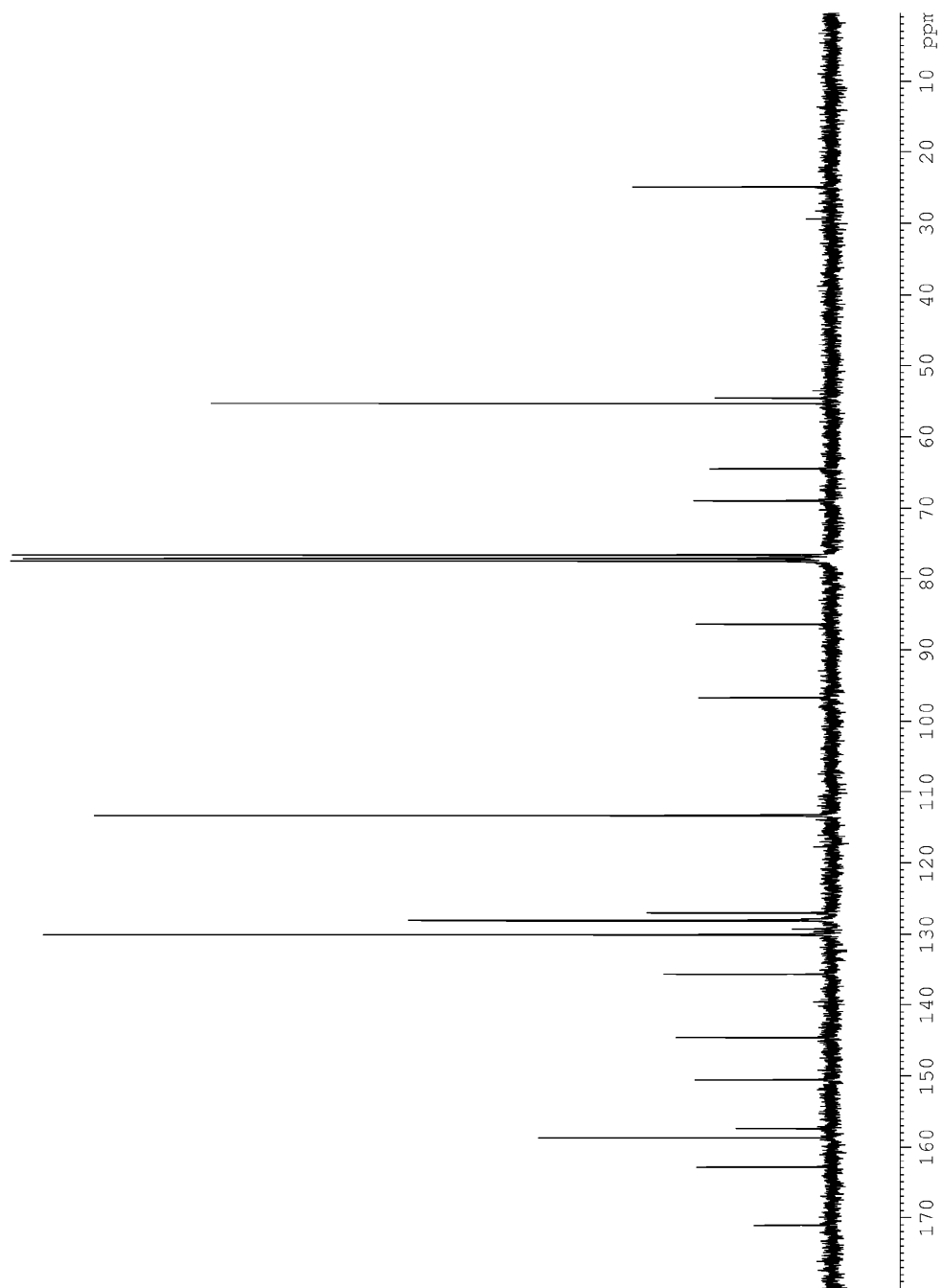


Figure A2.30.2. ^{13}C NMR spectrum of compound **2.46** (75 MHz, CDCl_3).

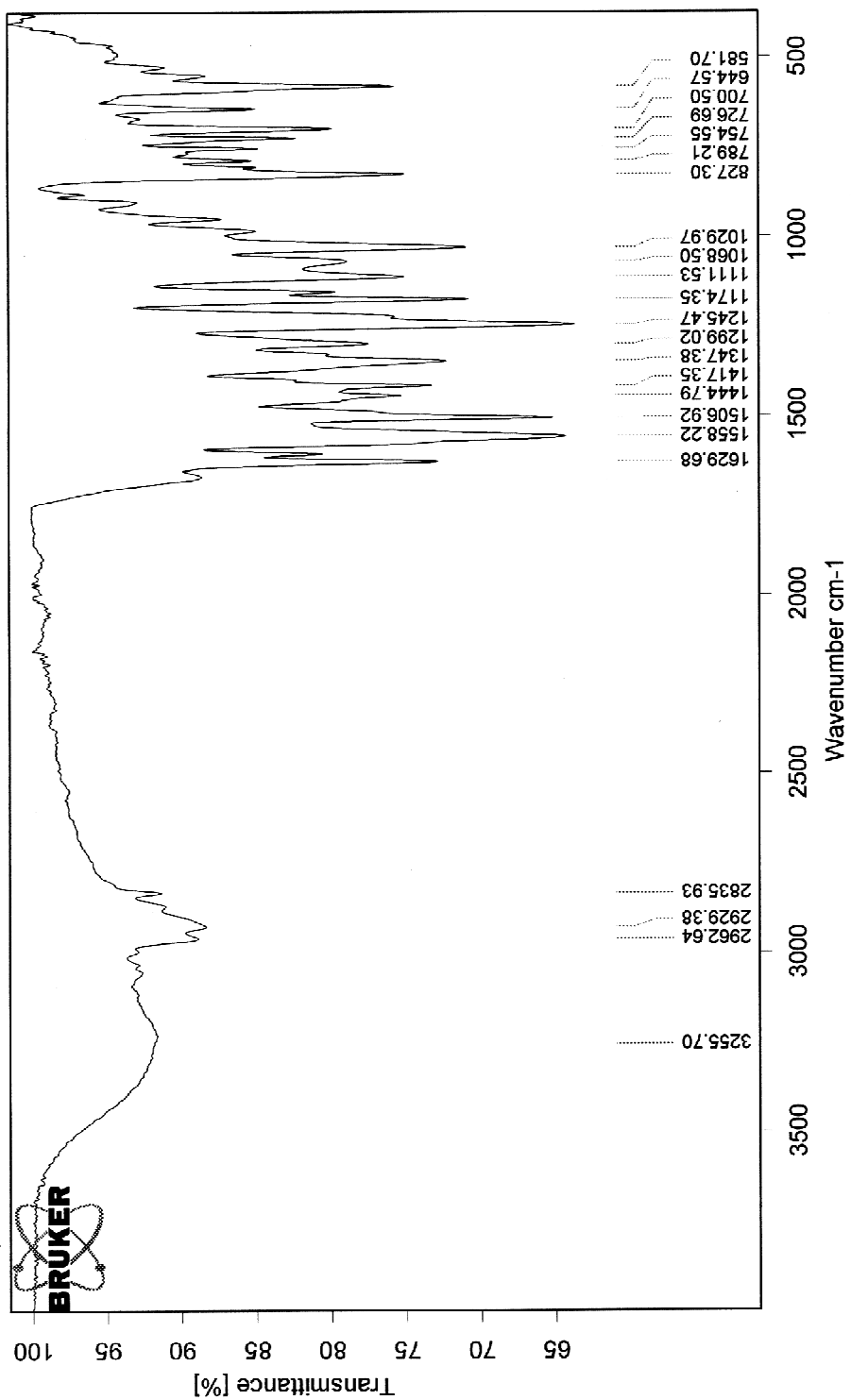


Figure A2.30.3. IR spectrum of compound 2.46 (solid).

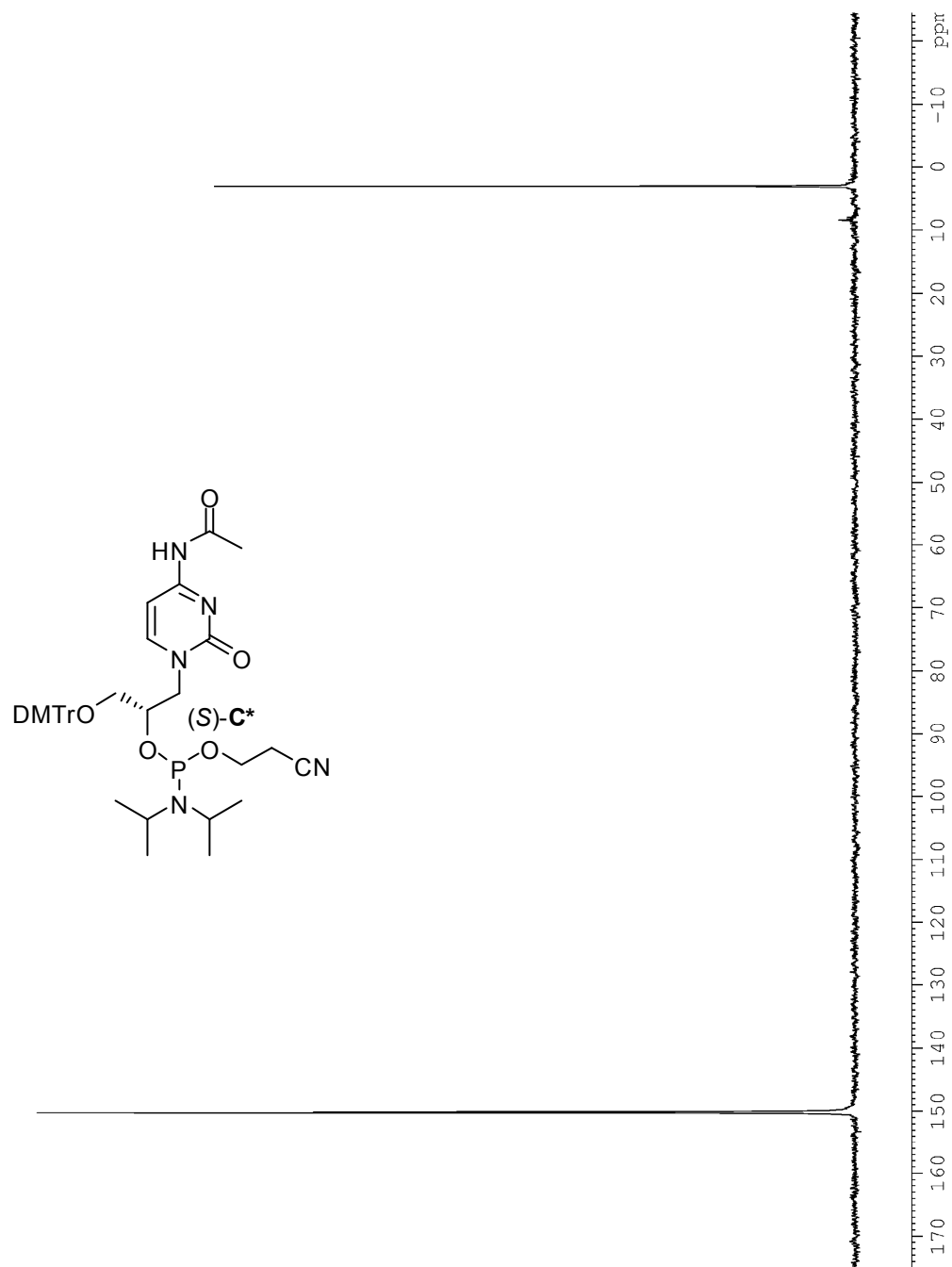


Figure A2.31.1. ³¹P NMR spectrum of phosphoramidite (*S*)-**C*** (162 MHz, CDCl₃) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).

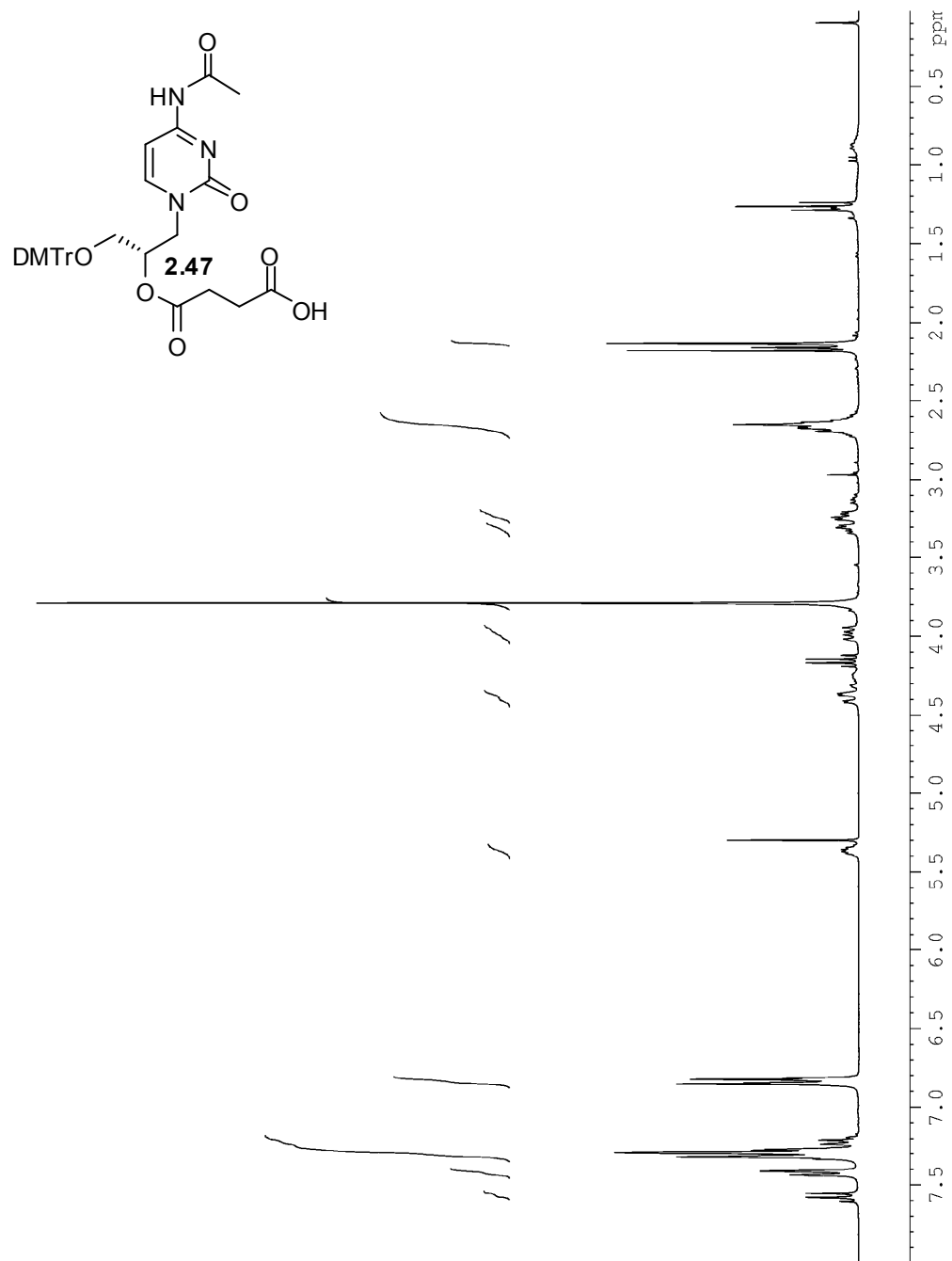


Figure A2.32.1. ^1H NMR spectrum of compound **2.47** (300 MHz, CDCl_3).

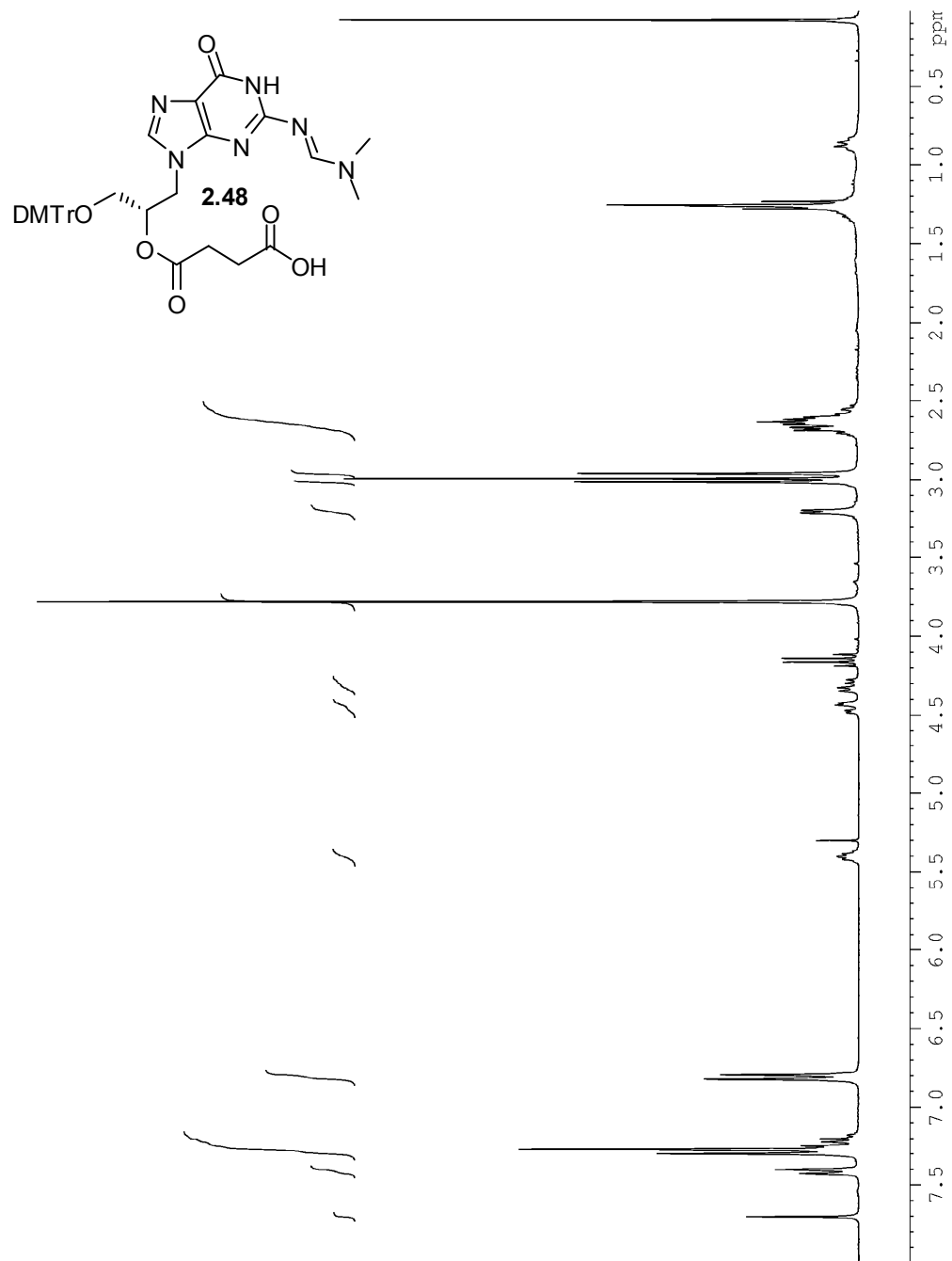


Figure A2.33.1. ^1H NMR spectrum of compound **2.48** (300 MHz, CDCl_3).

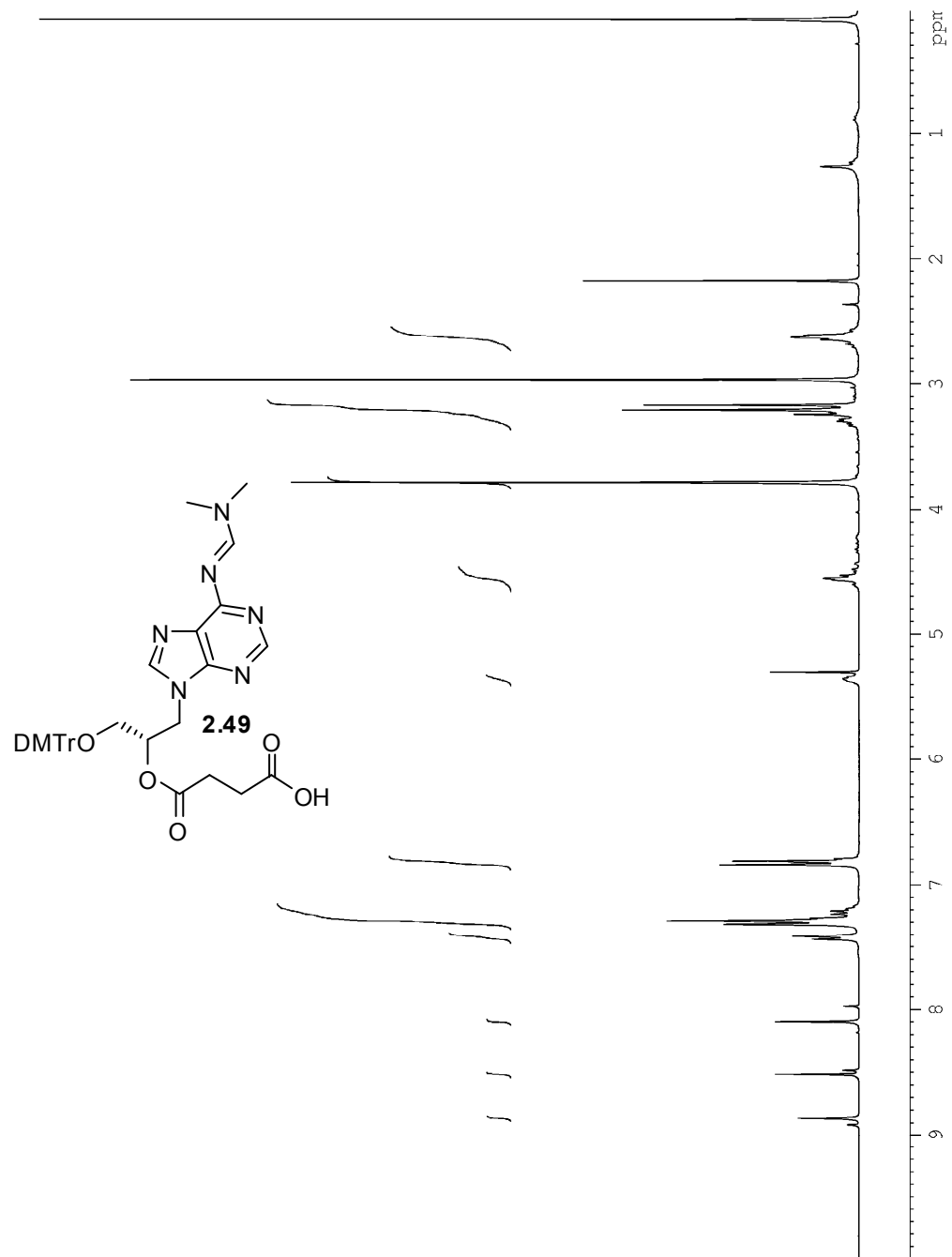


Figure A2.34.1. ¹H NMR spectrum of compound **2.49** (300 MHz, CDCl₃).

Table A.2.1.1. Table of reagents for oligonucleotide synthesis

Name	Composition
Reagent 18	100% anhydrous acetonitrile
Reagent Tet	0.25 M 5-ethylthiotetrazole in acetonitrile
Reagent Cap	Acetic anhydride in THF/Pyridine 16% 1-methylimidazole in THF
Reagent 15	0.02 M I ₂ in THF/Pyridine/H ₂ O
Reagent 19	Argon
Reagent 14	3% dichloroacetic acid in CH ₂ Cl ₂

Table A2.2.1. DNA oligonucleotide synthesis protocol: DNA (1 μM)

Step	Function	Name	Time	A	G	C	T	5	6	7	8
1	106	Begin									
2	64	18 to waste	3.0								
3	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5	1	Block flush	4.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6	101	Phos prep	3.0								
7	140	Column 1 ON									
8	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9	58	Tet to waste	1.7								
10	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
11	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
13	43	Push to column									
14	141	Column 1 OFF									
15	142	Column 2 ON									
16	64	18 to waste	4.0								
17	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
18	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
19	58	Tet to waste	1.7								
20	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
21	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
22	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
23	43	Push to column									

24	143	Column 2 OFF											
25	144	Column 3 ON											
26	64	18 to waste	4.0										
27	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
28	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
29	58	Tet to waste	1.7										
30	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
31	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
32	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
33	43	Push to column											
34	145	Column 3 OFF											
35	146	Column 4 ON											
36	64	18 to waste	4.0										
37	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
38	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
39	58	Tet to waste	1.7										
40	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
41	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
42	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
43	43	Push to column											
44	147	Column 4 OFF											
45	103	Wait	25.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
46	102	Cap prep	3.0										
47	64	18 to waste	4.0										
48	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
49	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
50	39	Cap to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
51	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
52	64	18 to waste	4.0										
53	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
54	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
55	41	15 to column	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
56	64	18 to Waste	4.0										
57	1	Block Flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
58	103	Wait	15.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
59	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
60	4	Flush to waste	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
61	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
62	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
63	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
64	105	Start detrityl											
65	64	18 to waste	4.0										
66	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
67	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

68	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
69	167	If Monitoring									
70	44	19 to column	35.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
71	40	14 to column	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
72	135	Monitor trityls									
73	40	14 to column	85.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
74	136	Monitor noise									
75	40	14 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
76	137	Stop Monitor									
77	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
78	2	Reverse flush	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
79	168	If not monitoring									
80	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
81	3	Trityl flush	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
82	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
83	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
84	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
85	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
86	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
87	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
88	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
89	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
90	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
91	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
92	3	Trityl flush	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
93	106	End monitoring									
94	42	18 to column	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
95	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
96	1	Block flush	4.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
97	107	End									

Table A2.3.1. Initial GNA oligonucleotide synthesis protocol with changes highlighted in red: 1000GNA (1 μ M)

Step	Function	Name	Time	A	G	C	T	5	6	7	8
1	106	Begin									
2	64	18 to waste	3.0								
3	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5	1	Block flush	4.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6	101	Phos prep	3.0								
7	140	Column 1 ON									
8	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

9	58	Tet to waste	1.7										
10	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
11	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
13	43	Push to column											
14	141	Column 1 OFF											
15	142	Column 2 ON											
16	64	18 to waste	4.0										
17	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
18	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
19	58	Tet to waste	1.7										
20	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
21	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
22	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
23	43	Push to column											
24	143	Column 2 OFF											
25	144	Column 3 ON											
26	64	18 to waste	4.0										
27	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
28	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
29	58	Tet to waste	1.7										
30	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
31	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
32	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
33	43	Push to column											
34	145	Column 3 OFF											
35	146	Column 4 ON											
36	64	18 to waste	4.0										
37	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
38	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
39	58	Tet to waste	1.7										
40	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
41	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
42	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
43	43	Push to column											
44	147	Column 4 OFF											
45	103	Wait	180.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
46	103	Wait	60.0	No	Yes	No	No	No	No	No	No	No	No
47	103	Wait	300.0	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
48	102	Cap prep	3.0										
49	64	18 to waste	4.0										
50	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
51	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
52	39	Cap to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

53	103	Wait	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
54	64	18 to waste	4.0								
55	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
56	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
57	41	15 to column	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
58	64	18 to Waste	4.0								
59	1	Block Flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
60	103	Wait	20.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
61	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
62	4	Flush to waste	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
63	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
64	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
65	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
66	105	Start detrityl									
67	64	18 to waste	4.0								
68	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
69	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
70	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
71	167	If Monitoring									
72	44	19 to column	35.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
73	40	14 to column	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
74	135	Monitor trityls									
75	40	14 to column	85.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
76	136	Monitor noise									
77	40	14 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
78	137	Stop Monitor									
79	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
80	2	Reverse flush	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
81	168	If not monitoring									
82	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
83	3	Trityl flush	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
84	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
85	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
86	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
87	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
88	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
89	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
90	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
91	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
92	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
93	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
94	3	Trityl flush	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
95	106	End monitoring									
96	42	18 to column	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

97	2	Reverse flush	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
98	1	Block flush	4.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
99	107	End										

Table A.2.4.1. Final GNA oligonucleotide synthesis protocol with changes highlighted in red: 1000GNA3 (1 μ M)

Step	Function	Name	Time	A	G	C	T	5	6	7	8
1	106	Begin									
2	64	18 to waste	3.0								
3	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5	1	Block flush	4.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6	101	Phos prep	3.0								
7	140	Column 1 ON									
8	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9	58	Tet to waste	1.7								
10	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
11	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12	103	Wait	60.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
13	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
14	43	Push to column									
15	141	Column 1 OFF									
16	142	Column 2 ON									
17	64	18 to waste	4.0								
18	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
19	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
20	58	Tet to waste	1.7								
21	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
22	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
23	103	Wait	60.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
24	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
25	43	Push to column									
26	143	Column 2 OFF									
27	144	Column 3 ON									
28	64	18 to waste	4.0								
29	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
31	58	Tet to waste	1.7								
32	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
33	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
34	103	Wait	60.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
35	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

36	43	Push to column										
37	145	Column 3 OFF										
38	146	Column 4 ON										
39	64	18 to waste	4.0									
40	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
41	111	Block vent	2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
42	58	Tet to waste	1.7									
43	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
44	34	Tet to column	1.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
45	103	Wait	60.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
46	33	B+Tet to column	2.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
47	43	Push to column										
48	147	Column 4 OFF										
49	103	Wait	180.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
50	103	Wait	300.0	No	No	No	No	Yes	Yes	Yes	Yes	Yes
51	102	Cap prep	3.0									
52	64	18 to waste	4.0									
53	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
54	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
55	39	Cap to column	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
56	103	Wait	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
57	39	Cap to column	7.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
58	103	Wait	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
59	64	18 to waste	4.0									
60	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
61	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
62	41	15 to column	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
63	64	18 to Waste	4.0									
64	1	Block Flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
65	103	Wait	20.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
66	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
67	4	Flush to waste	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
68	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
69	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
70	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
71	105	Start detrityl										
72	64	18 to waste	4.0									
73	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
74	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
75	1	Block flush	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
76	167	If Monitoring										
77	44	19 to column	35.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
78	40	14 to column	3.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
79	135	Monitor trityls										

80	40	14 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
81	103	Wait	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
82	40	14 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
83	103	Wait	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
84	40	14 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
85	103	Wait	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
86	40	14 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
87	136	Monitor noise									
88	40	14 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
89	137	Stop Monitor									
90	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
91	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
92	168	If not monitoring									
93	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
94	3	Trityl flush	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
95	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
96	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
97	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
98	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
99	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
100	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
101	40	14 to column	6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
102	103	Wait	5.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
103	3	Trityl flush	5.0	No	No	Yes	Yes	Yes	Yes	Yes	Yes
104	42	18 to column	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
105	3	Trityl flush	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
106	106	End monitoring									
107	42	18 to column	8.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
108	2	Reverse flush	10.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
109	1	Block flush	4.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
110	107	End									

Chapter 3

Thermal and thermodynamic properties of glycol nucleic acid duplexes

Chapter 3.1. From single incorporation to completely acyclic oligonucleotides

The initial discovery that GNA oligonucleotides could form stable duplexes in an antiparallel fashion was surprising.¹⁻² In fact, experiments within the last two decades had shown that incorporation of one or two acyclic nucleotides into DNA oligonucleotides resulted in a large thermal destabilization of the resulting duplexes.³⁻⁹ Furthermore, it was long thought that nucleic acid analogs containing phosphodiester backbones needed to be cyclic in order to confer the conformational preorganization required for duplex formation.³⁻¹² However, this seems not to be the case with GNA, especially considering duplexes consisting of the acyclic GNA backbone display a higher thermal stability than DNA and RNA duplexes of the same sequence. Therefore, we were interested in developing an understanding as to the reasons for the stability of GNA duplexes, as well as their pairing behavior with DNA and RNA.

We first wanted to understand the relationship between the incorporation of single GNA nucleotides into DNA and completely acyclic GNA oligonucleotides.¹ Investigation of the temperature dependent UV melting at 260 nm of DNA/GNA hybrid systems by then post doctoral coworker Dr. Lili Zhang confirmed previous findings in which single incorporation of a GNA nucleotide resulted in a large thermal destabilization (Table 3.1). For example, incorporation of one glycol thymine nucleotide into the middle of a 15-mer DNA duplex resulted in a reduction of thermal stability to 34 °C from 47 °C for the pure DNA duplex. Incorporation of a total of three glycol nucleotides in the middle of the duplex resulted in a further destabilization to 29 °C.

Finally, substitution of one entire 15-mer single strand with GNA nucleotides results in no observable duplex formation with the complementary DNA single strand by UV dependent melting.

Table 3.1. Thermal stabilities of duplexes of DNA, GNA, and DNA/GNA hybrids^[a]

Entry	Duplex ^[b]	T _m (°C) ^[c]
1	5'-CACATTATTGTTGTA-3' 3'-GTGTAATAACAACAT-5'	47
2	5'-CACATT A TTGTTGTA-3' 3'-GTGTAATAACAACAT-5'	34 (40)
3	5'-CACATT AT TGTTGTA-3' 3'-GTGTAATAACAACAT-5'	29 (28)
4	3'- CACATTATTGTTGTA -2' 3'-GTGTAATAACAACAT-5'	no T _m (no T _m)
5	5'-CACATT A TTGTTGTA-3' 3'-GTGTA T AACAACAT-5'	33 (33)
6	5'-CACATT AT TGTTGTA-3' 3'-GTGTA TA ACAACAT-5'	39 (39)
7	3'- CAC ATT ATT GTT GTA -2' 2'- GTG TAA TAA CAA CAT -3'	71

[a] Measured in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM NaCl and 2 μM of each strand. [b] GNA nucleotides shown in red. [c] Shown for *S* nucleotides with the values for *R* nucleotides given in parentheses

Next, the substitution of an entire single base pair into the middle of the DNA duplex results again in a large destabilization to 33 °C. Further substitution of three base pairs in the middle of the duplex with glycol nucleotides results in a duplex with a

stability of 39 °C. Although this duplex is still destabilized relative to the pure DNA duplex, it is only destabilized by 1.3 °C per glycol nucleotide versus 6 °C per nucleotide when substituting single nucleotides. As expected from previous studies, substitution of all base pairs with GNA nucleotides results in an extremely stable duplex with a thermal melting point of 71 °C. This surpasses the stability of the pure DNA duplex by 24 °C. These results, in combination with previous results¹⁻² from our group, demonstrate that the assumption about the importance of a cyclic backbone for conformational preorganization does not hold for GNA. Instead, the results from Table 3.1 suggest the destabilization is caused by a structural incompatibility between the 2'-deoxyribonucleotide and glycol oligonucleotide backbones since GNA is able to pair strongly with itself, but not with DNA nucleotides.

Additionally, we were interested in determining how general the duplex formation was for GNA, and how comparable it is to that of DNA. Over the years, we have measured the thermal melting points (T_m) of many GNA duplexes. The T_m is defined as the temperature at which a nucleic acid is halfway between its fully annealed and fully denatured states (i.e. – when 50% of the base pairs are dissociated). Measurements of T_m values can be made by monitoring the change in absorbance as a function of temperature at a wavelength corresponding to the maximal absorbance of the nucleobases, normally 260 nm. The T_m value can then be determined from the resulting sigmoidal curve (Figure 3.1) by either a nonlinear fit or the first derivative. Table 3.2 shows the T_m values for several GNA duplexes of varying sequence and length.¹ Comparing these values to those of DNA in parentheses shows that GNA duplexes not only display higher thermal

melting points in each case, but follow the stability trends shown for DNA duplexes; that is, melting temperature increases with length and increasing G:C base pair content.

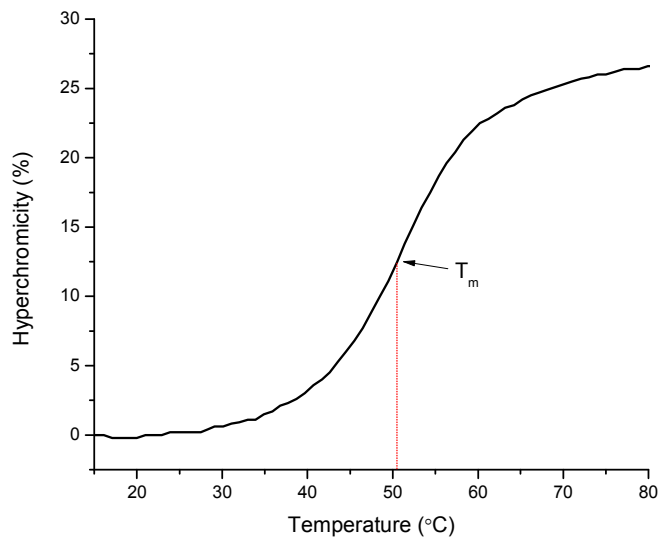


Figure 3.1. A representative thermal melting curve as monitored by UV spectroscopy of the GNA duplex 3'-AATATTATTATTTA-2' and its complement. Measurement was made in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride. The melting temperature (T_m) is indicated by the arrow and represents the inflection point of the curve. The T_m can be calculated by the first derivative or nonlinear fit of the data.

Table 3.2. Thermal stabilities of (*S*)-GNA and DNA duplexes ^[a]

Entry	[NaCl]	Duplex	T _m (°C)
1	1 M	3'-ATTAAT-2' 2'-TAATTA-3'	< 20
2	100 mM	3'-GCGCGC-2' 2'-CGC GCG-3'	53 ^[b]
3	1 M	3'-AAATATTT-2' 2'-TTTATAAA-3'	25
4	150 mM	3'-AACTAGTT-2' 2'-TTGATCAA-3'	38
5	500 mM	3'-CGAATTCG-2' 2'-GCTTAAGC-3'	54 (36)
6	100 mM	3'-TACGCACGACAT G-2' 2'-ATGCGTGCTGTA C-2'	83
7	500 mM	3'-AATATTATTATTTTA-2' 2'-TTATAATAATAAAAAT-3'	59 (41)
8	150 mM	3'-AAAAAAAAAAAAAAAAA-2' 2'-TTTTTTTTTTTTTTTT-3'	65 (40)
9	100 mM	3'-CACATTATTGTTGTA-2' 2'-GTGTAATAACAACAT-3'	71 (47)
10	200 mM	3'-TAAAATTTATATTATTAA-2' 2'-ATTTTAAATATAATAATT-3'	63 (41)

[a] Measured in 10 mM sodium phosphate buffer (pH = 7.0) with the indicated amount of NaCl at 2 μM duplex concentration. Values for DNA shown in parentheses. [b] Measured using CD spectroscopy.

Chapter 3.2. Watson-Crick pairing, salt dependence, CD spectroscopy

At this point, it was clear that GNA was forming duplexes in an antiparallel fashion with thermal stabilities greater than those of the corresponding DNA duplexes. Considering we were somewhat astonished that this was possible with a completely acyclic backbone, we were interested in understanding GNA duplex formation and comparing its properties with that of DNA. It was hoped that this would allow insight into the reason for the higher thermal stability. First, we wanted to determine the rules for nucleobase pairing in the context of GNA duplexes. Therefore, co-worker Lili Zhang measured the stabilities of matched and mismatched Watson-Crick base pairs in GNA duplexes and compared it with those in DNA.¹ Thermal melting points were determined for a 15-mer duplex in which the center base pair was modified to give every possible combination of pairing (Table 3.3). In GNA duplexes, base pairs of A:T and G:C form the most stable duplexes with T_m values of 71 °C and 72 °C for A:T pairs and 75 °C and 76 °C for G:C pairs. Similarly for DNA, the most stable combinations of A:T and G:C provide T_m values of 47 °C and 48 °C, respectively. All other combinations are destabilizing in both the GNA and DNA systems displaying destabilizations of 6-17 °C for GNA and 9-23 °C for DNA duplexes. This suggests that hydrogen bonding functions in both GNA and DNA duplexes as a means to determine the highly specific interaction between the nucleobases on opposite strands, discriminating against other combinations of nucleobases based on slight energy differences.

Table 3.3. Thermal stabilities (T_m [$^{\circ}\text{C}$]) of matched and mismatched Watson-Crick base pairs in duplex GNA^[a]

		3'-CACATTAXTGTTGTA-2' 2'-GTGTAATYACAACAT-3'			
		X			
		A	G	C	T
Y	A	65 (34)	65 (38)	63 (34)	71 (47)
	G	66 (36)	62 (38)	76 (48)	62 (36)
	C	63 (31)	75 (48)	58 (25)	61 (33)
	T	72 (47)	62 (35)	59 (34)	61 (36)

[a] Measured in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM NaCl at 2 μM duplex concentration. Values for DNA shown in parentheses.

Moreover, it was interesting to understand the effect ionic strength of the solvent has on the formation of GNA duplexes. Metal cations such as Na^+ and Mg^{+2} are important for the formation of nucleic acid duplexes because they help to diminish the electrostatic repulsion of negative charges on phosphodiester groups of opposing strands during duplex formation.¹³ Metal cations have also been shown to bind to the G:C base pairs of DNA duplexes.¹⁴⁻¹⁵ Figure 3.2 shows the effect of increasing sodium ion concentration on the T_m value of a 15-mer GNA duplex and compares it to a DNA duplex of the same sequence.¹ The data show a similar trend for both duplexes in which the melting point increases sharply as the concentration of sodium ions is increased from 0 to 0.2 M. Thereafter, the stabilization of each duplex proceeds in almost a linear fashion as the concentration of sodium ions is increased, albeit it in a much slower fashion. This data would suggest that cations play an equally important role in diminishing the electrostatic repulsion during GNA and DNA duplex formation.

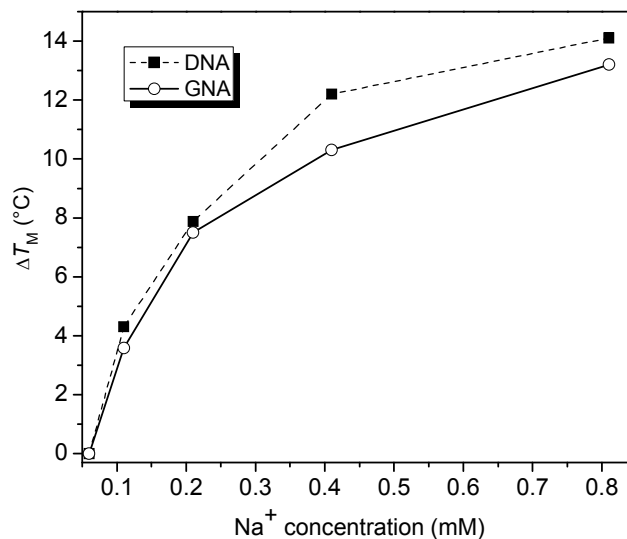


Figure 3.2. Comparison of the effect of sodium concentration on the thermal stability of a GNA duplex and DNA duplex of the sequence CACATTATTGTTGTA and its complement. The sequence is in the 3'→2' direction for GNA and 5'→3' direction of for DNA. Measurements were made in 10 mM sodium phosphate buffer (pH=7.0) with the indicated salt concentration and 2 μM of each strand. Each data point is the mean of two measurements.

Furthermore, we wanted to understand the effect of solvent content on GNA duplex formation compared to DNA duplex formation. Information about duplex formation can be gained by observing the melting behavior of duplexes in the presence of added organic co-solvent. The increase in organic content, and therefore decrease in aqueous content, of the solvent can have large effects on the hydrogen bonding and stacking interactions occurring during duplex formation.¹⁶⁻¹⁸ One would expect that the hydrogen bonding and stacking contributions towards duplex formation would increase and decrease, respectively, as the dielectric constant of the solvent is decreased. Measurement of the melting temperature of a 15-mer duplex in the presence of differing

amounts of *N,N'*-dimethylformamide showed similar results for both GNA and DNA duplexes (Figure 3.3).¹⁹ For GNA, every 10% increase in DMF content results in a destabilization of approximately 5 °C. A DNA duplex of the same sequence displays a destabilization of approximately 6 °C for every 10% increase in DMF content. This suggests that duplex formation for both GNA and DNA is affected equally as the dielectric constant of the solution is changed. Therefore, the solvation of the backbone and nucleobases is similar in both GNA and DNA.

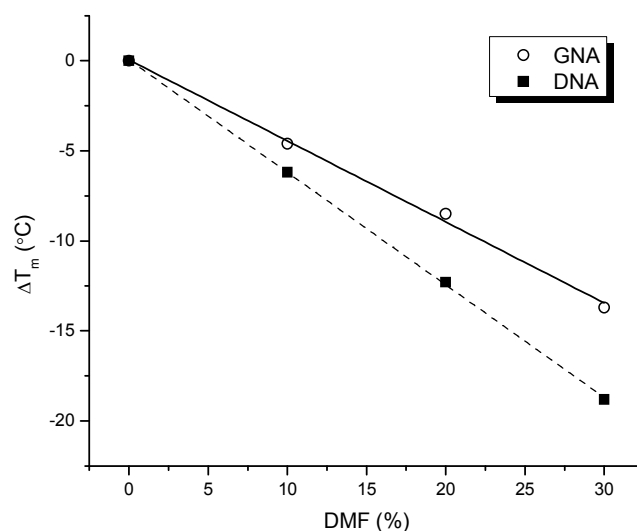


Figure 3.3. Thermal melting temperatures (T_m) of GNA and DNA duplexes of the sequence CAC ATT ATT GTT GTA and its complement in the presence of DMF. The shown sequence is in the direction 3'→2' for GNA and 5'→3' for DNA. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride and the indicated DMF percentage at 2 μ M duplex concentration. Each data point is the mean of two measurements.

Finally, we wanted to measure the CD spectrum of a GNA duplex and compare it to that of DNA. Circular dichroism is used extensively to determine the three dimensional configuration of proteins in solution. Although CD has also been used in the

determination of the secondary structure of DNA and RNA duplexes in solution, it is not as well understood as determining protein structure via CD spectroscopy.²⁰ In any case, it can be very indicative of slight structural changes and provide insight into the duplex stability. Accordingly, the CD spectrum of a GNA duplex (Figure 3.4) shows strong Cotton effects characteristic of the formation of a secondary structure.¹ The peaks at 270 nm, 220 nm, and 205 nm are indicative for the formation of a duplex of GNA. However, a comparison to the CD spectrum of a DNA duplex of the same sequence shows significant differences. Notably, the maxima and minima are inverted around 270, 250, 220, and 210 nm, and the intensities are smaller compared to those of GNA.

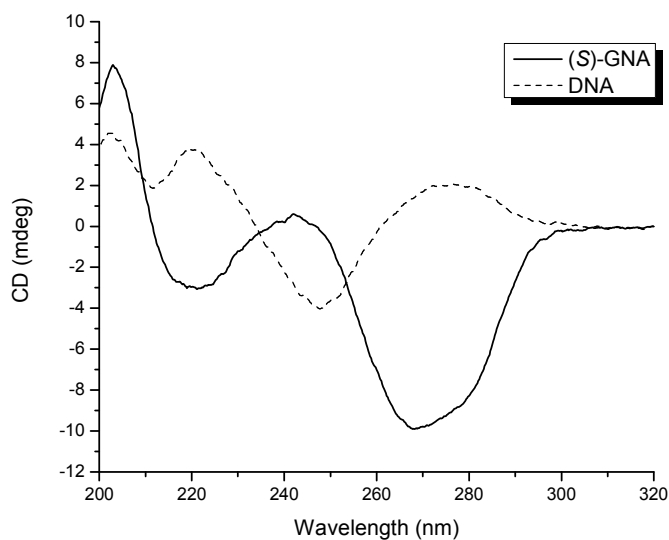


Figure 3.4. CD spectra of a (*S*)-GNA and DNA duplexes of the sequence CAC ATT ATT GTT GTA at a duplex concentration of 10 μ M. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride at 15 $^{\circ}$ C. Each curve is the average of ten measurements.

Moreover, quantitative analysis of the CD signal for the GNA duplex at 220 nm from 20 to 90 °C is shown in Figure 3.5. A sigmoidal fit to the data provides a T_m of 72 °C which is consistent with the melting point of 71 °C determined by UV-melting. A similar result is obtained by monitoring the CD signal at 270 nm, providing a melting point of 73 °C.¹

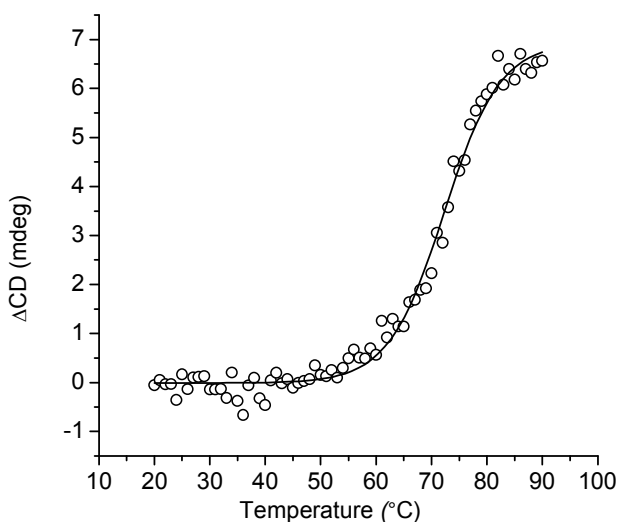


Figure 3.5. Temperature-dependent CD measurement of the (*S*)-GNA duplex 3'-CAC ATT ATT GTT GTA-2' and its complement at a duplex concentration of 2 μ M in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride. The CD signal (red dots) was measured at 220 nm as a function of temperature with a heating rate of 1 °C/min. A sigmoidal fit to the data (black line) yields a melting temperature of 72 °C based on the inflection point of the curve.

The data from the mismatch, ionic strength, and solvent content experiments suggest that the duplex formation properties of GNA and DNA are not significantly different. Both systems display similar thermal sensitivities to non Watson-Crick base

pairs, the amount of added metal cations in the solution, and the dielectric constant of the solution. The mismatch experiments also show that both GNA and DNA duplexes assemble based strictly on the Watson-Crick rules for pairing. The CD spectrum of the GNA duplex suggests the formation of a stable secondary structure which has a thermal melting temperature (T_m) almost identical to the melting temperature obtained using UV spectroscopy. Furthermore, this CD spectrum bears no resemblance to the CD spectrum of a corresponding DNA duplex under the same conditions.

Chapter 3.3. Thermal and thermodynamic stabilities of GNA duplexes

Chapter 3.3.1. Thermodynamics of GNA duplex formation

The next step towards an understanding of the higher thermal stabilities of GNA duplexes was to measure the thermodynamic parameters of duplex formation for GNA duplexes as compared to DNA duplexes.^{1,19} Over the course of several years, we had the opportunity to measure the thermodynamics of duplex formation for many different duplexes, the first of them shown in Table 3.4. The thermodynamics of duplex formation can be measured using the concentration method in which the T_m^{-1} is plotted against the $\ln[\text{oligo}]$.¹⁸ From this so called van't Hoff plot, the values of ΔH and ΔS can be extracted which allow for the subsequent calculation of the free energy of duplex formation, ΔG . From the four duplexes shown (Table 3.4), there is clearly a positive linear correlation between the T_m value and ΔG of duplex formation. Furthermore, GNA duplex formation is less exothermic than in DNA, but surprisingly, more entropically favorable in every case. Again, this was unexpected based on the assumption that the acyclic backbone of GNA is less conformationally preorganized to support duplex formation. Therefore, the data would suggest that the GNA backbone is adjusted to support duplex formation in a less entropically unfavorable manner.

Table 3.4. Thermal and thermodynamic stability of (S)-GNA and DNA duplexes^[a]

Entry	Duplex	T _m ^[b] (°C)	ΔG ^[c] (298K, kcal/mol)	ΔH ^[c] (kcal/mol)	-TΔS ^[c] (298K, kcal/mol)
1 ^[d]	CGA ATT CG GCT TAA GC	54 (36)	-12.2 (-9.4)	-54.0 (-59.4)	41.8 (49.9)
2 ^[e]	A ₁₅ T ₁₅	61 (37)	-19.5 (-12.6)	-102.3 (-104.2)	82.8 (91.6)
3 ^[d]	AAT ATT ATT ATT TTA TTA TAA TAA TAA AAT	59 (41)	-16.2 (-12.4)	-77.0 (-86.5)	60.8 (74.1)
4 ^[e]	CAC ATT ATT GTT GTA GTG TAA TAA CAA CAT	70 (47)	-21.1 (-15.2)	-95.6 (-103.3)	74.4 (88.1)

[a] Upper strand: 3'→2' direction for GNA and 5'→3' direction for DNA. Bottom strand: 2'→3' direction for GNA and 3'→5' direction for DNA. Values for DNA shown in parentheses. [b] Melting temperature indicated for duplex at a concentration of 2 μM. [c] ΔG, ΔH, and ΔS were estimated from van't Hoff plots. Each data point was measured in triplicate, and the mean value was taken. [d] Measured in 10 mM sodium phosphate buffer (pH=7.0) with 500 mM NaCl. [e] Measured in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM NaCl

These results led us to investigate the effect of overhanging nucleotides in GNA duplexes and compare it to DNA.^{19, 21-23} It was hoped that this would provide an indication of the stacking ability of GNA nucleotides during duplex formation compared to DNA. Since the terminal nucleotide has no pairing partner, the use of overhanging nucleotides also provides a measure of stacking ability that is separate from hydrogen bonding interactions. Therefore, we started with the self-complementary target sequence 5'(3')-CGA ATT CG-3'(2') and investigated both the melting temperatures and thermodynamic parameters for dangling nucleotides on both ends of duplexes of GNA and DNA. In GNA, the core sequence forms a duplex with a ΔH value of -54.0 kcal/mol

and a $-T\Delta S$ value of 41.8 kcal/mol (Entry 1, Table 3.5). The addition of a dangling base on the 3' end of GNA (Entries 2-5) has no significant stabilizing effect on duplex formation when compared to the ΔG and T_m of the core duplex (Entry 1). Although the change in ΔG is not significant for the addition of dangling bases, duplex formation is accompanied by a more unfavorable enthalpy (by 0.5 to 8.2 kcal/mol), but a more favorable entropy (by 0.6 to 7.5 kcal/mol) when compared to the core sequence. This is in stark contrast to the effect of dangling bases on the 2' end of the GNA duplex as shown in Entries 6-9 of Table 3.5. The addition of the dangling bases results in both higher thermal (T_m increases of 8.3 to 14.6 °C) and thermodynamic (ΔG increases of 2.2 to 4.4 kcal/mol) stabilities for these duplexes. Also in contrast to the dangling nucleotides on the 3' end, the dangling nucleotides on the 2' end results in more favorable enthalpic (by 6.8 to 14.9 kcal/mol) and less favorable entropic (by 4.4 to 13.3 kcal/mol) contributions to duplex formation. The ability of overhanging nucleobases to stabilize the duplex by stacking interactions on the 2' end of GNA is consistent with previous reports for overhanging nucleotides in both DNA and RNA. In the experiments with DNA and RNA it was determined that the stabilizing ability of the nucleobases increased with both decreasing polarity and increasing surface area (the polarity of the nucleobases increases in the order adenine<guanine<thymine<cytosine), consistent with what is found for the 2' overhanging nucleotides in GNA.

Table 3.5. Thermal and thermodynamic stabilities of (*S*)-GNA duplexes with overhanging nucleotides

Entry	Duplex ^[a]	T _m ^[c] (°C)	ΔG ^[d] (298K, kcal/mol)	ΔH ^[d] (kcal/mol)	-TΔS ^[d] (298K, kcal/mol)
1	3'-CGAATTCG-2' 2'-GCTTAAGC-3'	54.3	-12.2	-54.0	41.8
2	3'-ACGAATTCG-2' 2'-GCTTAAGCA-3'	55.2	-11.9	-50.0	38.1
3	3'-GCGAATTCG-2' 2'-GCTTAAGCG-3'	55.5	-11.6	-46.6	35.0
4	3'-CCGAATTCG-2' 2'-GCTTAAGCC-3'	54.8	-11.6	-45.8	34.3
5	3'-TCGAATTCG-2' 2'-GCTTAAGCT-3'	55.4	-12.3	-53.5	41.2
6	3'-CGAATTCGA-2' 2'-AGCTTAAGC-3'	68.9	-16.2	-68.9	52.7
7	3'-CGAATTCGG-2' 2'-GGCTTAAGC-3'	66.8	-15.8	-68.9	53.1
8	3'-CGAATTCGC-2' 2'-CGCTTAAGC-3'	62.6	-14.4	-63.1	48.6
9	3'-CGAATTCGT-2' 2'-TGCTTAAGC-3'	64.9	-14.6	-60.8	46.2

[a] Measured in 10 mM sodium phosphate buffer (pH=7.0) with 500 mM NaCl. [b] Melting temperature indicated for duplex at a concentration of 2 μM. [c] ΔG, ΔH, and ΔS were estimated from van't Hoff plots. Each data point was measured in triplicate, and the mean value was taken.

For DNA the core sequence forms a duplex with a ΔH value of -59.4 kcal/mol and a -TΔS value of 49.9 kcal/mol (Entry 1, Table 3.6). Adding an overhanging nucleotide to the 5' end (corresponding to the 3' end of GNA) results in both thermal

(T_m increases from 4.6 to 9.1 °C) and thermodynamic (ΔG increases from 0.6 to 1.8 kcal/mol) stabilization of the duplexes (Entries 2-5, Table 3.6). In three of the four cases, the duplex formation is accompanied by less favorable enthalpic (by 2.6 to 11 kcal/mol) and more favorable entropic (by 0.6 to 5.2 kcal/mol) contributions. This is in contrast to a similar experiment using the core sequence 5'-CGCGCG-3' in which a more favorable enthalpic and less favorable entropic contribution towards duplex formation was observed.²² However, the amount of stabilization provided by each of the stacking nucleotides is similar to that using the 6-mer DNA sequence. Similar to what is observed on the 5' end, dangling nucleotides on the 3' end of DNA also have a stabilizing effect, albeit to a lesser extent than on the 5' end. In contrast, dangling nucleotides on the 2' end of DNA have a large stabilizing effect on the duplex with an increase in T_m values from 9-15 °C based on the polarity of overhanging base. Dangling nucleotides on the corresponding 3' end of DNA also generally show a stabilizing effect (ΔG increase from 0-0.7 kcal/mol), albeit to a lesser degree than on the 5' end (Entries 6-9, Table 3.6). Again, this stabilization is accompanied by less favorable enthalpic (by 0.2 to 7.9 kcal/mol) and more favorable entropic (by 0.8 to 7.7 kcal/mol) contributions towards duplex formation. The stabilization trend on the 3' end is similar in which the purines (adenine and guanine) stabilize the duplex to a greater extent than the pyrimidines (cytosine and thymine). However, there is very little difference between the pairs (i.e.- adenine is similar to guanine).

Table 3.7 summarizes Tables 3.5 and 3.6 to emphasize the trends in the thermodynamic data. Averaging the thermodynamic effect of all dangling nucleotides

shows an average stabilization of 0.69 kcal/mol per nucleotide in GNA compared with 0.36 kcal/mol per nucleotide in DNA.¹⁹ This suggests that stacking is thermodynamically more favorable for GNA duplex formation versus DNA duplex formation. The potentially important factors for stabilizing face-to-face base interactions are electrostatic interactions (dipole-dipole and dipole-induced dipole), dispersion effects (momentary dipole-induced dipole), and solvation effects.²²

Table 3.6. Thermal and thermodynamic stabilities of DNA duplexes with overhanging nucleotides

Entry	Duplex ^[a]	T _m ^[c] (°C)	ΔG ^[d] (298K, kcal/mol)	ΔH ^[d] (kcal/mol)	-TΔS ^[d] (298K, kcal/mol)
1	5'-CGAATTCG-3' 3'-GCTTAAGC-5'	35.5	-9.4	-59.4	49.9
2	5'-ACGAATTCG-3' 3'-GCTTAAGCA-5'	45.0	-11.2	-60.4	49.3
3	5'-GCCGAATTCG-3' 3'-GCTTAAGCG-5'	42.6	-10.7	-59.1	48.4
4	5'-CCGAATTCG-3' 3'-GCTTAAGCC-5'	40.1	-10.0	-54.7	44.7
5	5'-TCGAATTCG-3' 3'-GCTTAAGCT-5'	41.1	-10.3	-56.8	46.5
6	5'-CGAATTCGA-3' 3'-AGCTTAAGC-5'	39.1	-10.0	-57.6	47.7
7	5'-CGAATTCGG-3' 3'-GGCTTAAGC-5'	39.4	-10.1	-59.2	49.1
8	5'-CGAATTCGC-3' 3'-CGCTTAAGC-5'	36.8	-9.3	-51.5	42.2
9	5'-CGAATTCGT-3' 3'-TGCTTAAGC-5'	36.9	-9.4	-54.1	44.7

[a] Measured in 10 mM sodium phosphate buffer (pH=7.0) with 500 mM NaCl. [b] Melting temperature indicated for duplex at a concentration of 2 μM. [c] ΔG, ΔH, and ΔS were estimated from van't Hoff plots. Each data point was measured in triplicate, and the mean value was taken.

Table 3.7. Thermal and thermodynamic stabilities of (*S*)-GNA and DNA duplexes with overhanging nucleotides^[a]

Entry	Overhanging nucleotide X ^[b]	XCGAATTCG GCTTAAGCX			CGAATTCGX XGCTTAAGC		
		T _m (°C)	ΔT _m (°C)	ΔΔG (298K, kcal/mol)	T _m (°C)	ΔT _m (°C)	ΔΔG (298K, kcal/mol)
<i>GNA duplexes</i>							
1	None	54.3	0	0	54.3	0	0
2	A	55.2	0.9	-0.2	68.9	14.6	4.0
3	G	54.8	0.5	-0.6	66.8	12.5	3.6
4	C	55.4	1.1	-0.6	62.6	8.3	2.3
5	T	55.5	1.2	0.1	64.9	10.6	2.4
<i>DNA Duplexes</i>							
6	None	35.5	0	0	35.5	0	0
7	A	45.0	9.5	1.8	39.1	3.6	0.6
8	G	42.6	7.1	1.3	39.4	3.9	0.7
9	C	40.1	4.6	0.6	36.9	1.4	-0.1
10	T	41.1	5.6	0.9	36.8	1.3	0.0

[a] Measured in 10 mM sodium phosphate buffer (pH = 7.0) with 500 mM NaCl and 2 μM of each duplex. [b] Sequence for the upper strand in the direction 3'→2' for GNA and 5'→3' for DNA.

Therefore, to further probe the nature of the stacking interactions in DNA and GNA, experiments determining the effect of the added organic cosolvent DMF to the stability of overhanging bases are shown in Figures 3.6 and 3.7. Since solvophobic effects are correlated with the extent of surface area excluded from solvent, one would

expect that the greater the extent of solvophobic contribution to molecular interactions, the more that an added organic cosolvent will affect such interactions.²² For both the GNA and DNA duplexes with an overhanging adenine nucleotide, the stabilities of the duplexes containing overhanging bases is more sensitive towards the addition of organic cosolvent than the core sequence. For GNA, the slope of the adenine overhang sequence is -0.63 compared to -0.47 for the core sequence. In DNA, the slope of the adenine overhang sequence is -0.72 versus -0.52 for the core sequence. The more negative slope for the DNA duplexes shows that the stacking in DNA is more sensitive towards solvophobic effects than in GNA, suggesting that stabilization by overhanging bases in DNA is more driven by solvent effects.²² However, if one considers the thermodynamic data presented in Tables 3.5 and 3.6, the increased thermodynamic stability of overhanging bases in GNA is accompanied by a more favorable enthalpic and less favorable entropic term towards duplex formation. Furthermore, the trend in which GNA duplex formation is less exothermic and more entropically favorable than in DNA is actually reversed with the addition of overhanging nucleotides. However, we are hesitant to over-interpret this data because of known effects of entropy-enthalpy compensation.²⁴ Thus, it can be concluded that the addition of overhanging bases in GNA results in larger thermodynamic stabilities than in DNA. The data presented here suggests that the increased stability is the result of slightly different interactions between nucleobases.

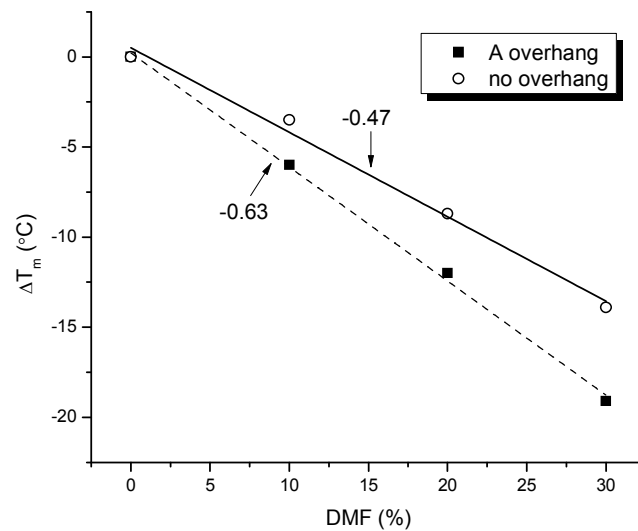


Figure 3.6. Thermal melting temperatures (T_m) of self-complementary GNA duplexes of the sequence 3'-CGAATTCG-2' (open circles) and 3'-CGAATTCGA-2' (filled squares). Also shown are the linear fits to the data with the corresponding slope of the fit. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride and the indicated DMF percentage at 2 μ M duplex concentration. Each data point is the mean of two measurements.

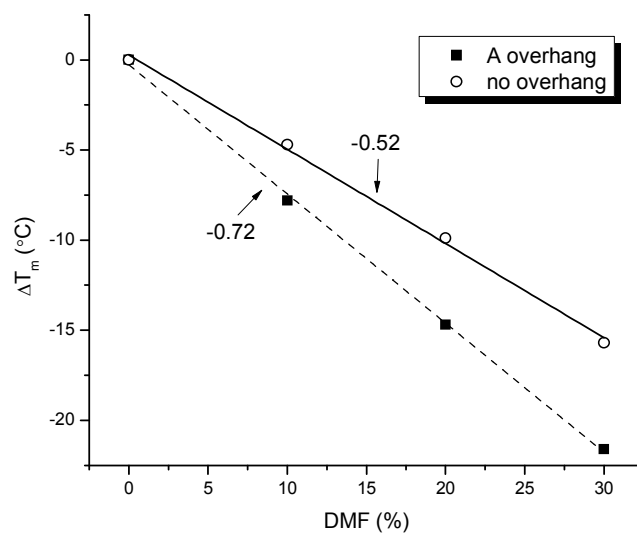


Figure 3.7. Thermal melting temperatures (T_m) of self-complementary DNA duplexes of the sequence 5'-CGAATTCG-3' (open circles) and 5'-CGAATTCGA-3' (filled squares). Also shown are the linear fits to the data with the corresponding slope of the fit. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride and the indicated DMF percentage at 2 μ M duplex concentration. Each data point is the mean of two measurements.

Chapter 3.3.2. Probing GNA duplex formation using circular dichroism

A more favorable entropic contribution to GNA duplex formation may indicate a strongly preorganized conformation in GNA single strands.²⁵ Therefore, we investigated the temperature-dependent conformation of GNA single strands in solution and compared it to DNA single strands of the same sequence.¹⁹ Figure 3.8 shows that the GNA single strand 3'-CACATTATTGTTGTA-2' displays significant Cotton effects that are only slightly less pronounced than those of the corresponding duplex at the same overall concentration of single strands. Upon heating from 15-90 °C, the CD signals gradually disappear indicating the GNA single strands are adopting a defined helical structure.

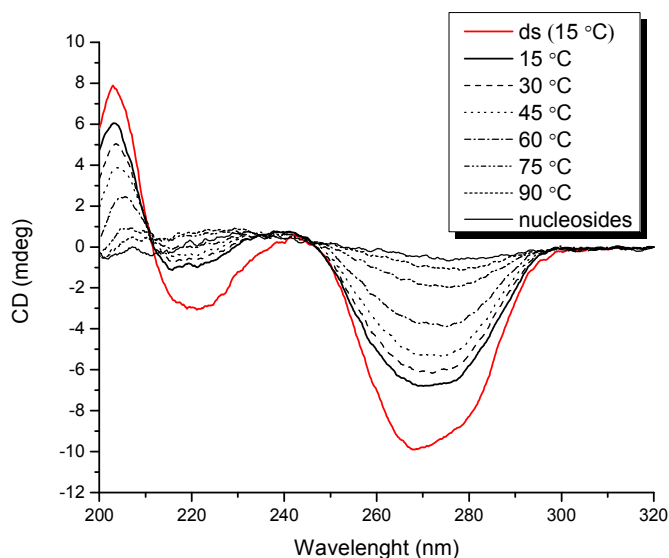


Figure 3.8. Melting of the GNA single strand 3'-CAC ATT ATT GTT GTA-2' as monitored by circular dichroism (CD). The measurements were made in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride at a concentration of 20 μ M. The CD spectra of the corresponding duplex (blue, 10 μ M per single strand) and mixture of nucleosides (red, 300 μ M total concentration) are shown for reference. Each curve is the average of at least five measurements.

Furthermore, upon cooling of the same solution from 90-15 °C, the CD signals at 205, 220, 240, and 270 nm reappear with equal intensity indicating reformation of this defined helical structure. From this it can be concluded that the conformation of the GNA single strand in solution resembles that of the individual strands in the duplex. As a comparison, analogous CD experiments with DNA single strands are not as conclusive and show less pronounced temperature dependent behavior (Figure 3.9). Thus, although acyclic, the GNA backbone can apparently lead to an ideal preorganization of the single strands for duplex formation.

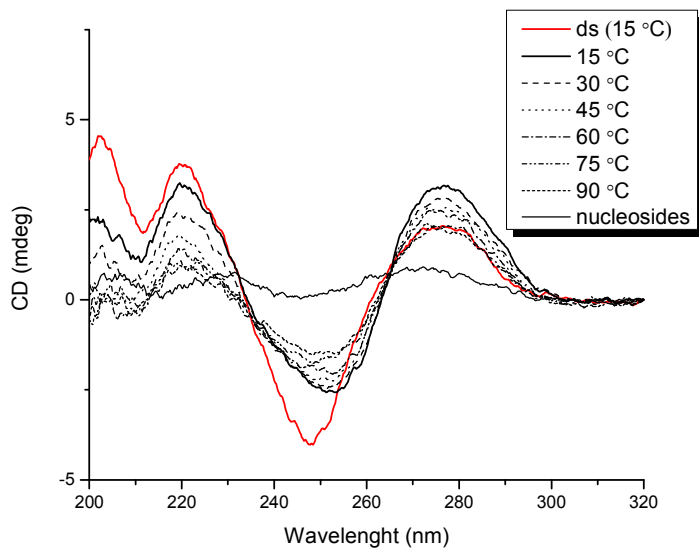


Figure 3.9. Melting of the DNA single strand 5'-CAC ATT ATT GTT GTA-3' as monitored by circular dichroism (CD). The measurements were made in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride at a concentration of 20 μ M. The CD spectra of the corresponding duplex (blue, 10 μ M per single strand) and mixture of nucleosides (red, 300 μ M total concentration) are shown for reference. Each curve is the average of at least five measurements.

To test the generality of this single strand preorganization, we investigated the CD spectra of GNA single strands of different sequences and lengths. Figure 3.10 demonstrates that not all GNA single strands adopt a preorganized conformation in solution. For example, a 5-mer of the sequence 3'-TTGTA-2', a 10-mer of the sequence 3'-TATTGTTGT A-2', and a 15-mer of the sequence 3'-T₁₅-2' show CD spectra with no significant Cotton effects compared to the 15-mer 3'-CATATTATTGTTGTA-2'. Single strands of the sequence 3'-A₅-2' and 3'-A₁₅-2' do show significant Cotton effects, but with maxima and minima that differ from that of duplex GNA indicating that these single strands are adopting a different conformation than is found in the duplex. Therefore, it is clear that not all GNA single strands demonstrate a strongly preorganized conformation in solution and that it is highly sequence and length dependent.

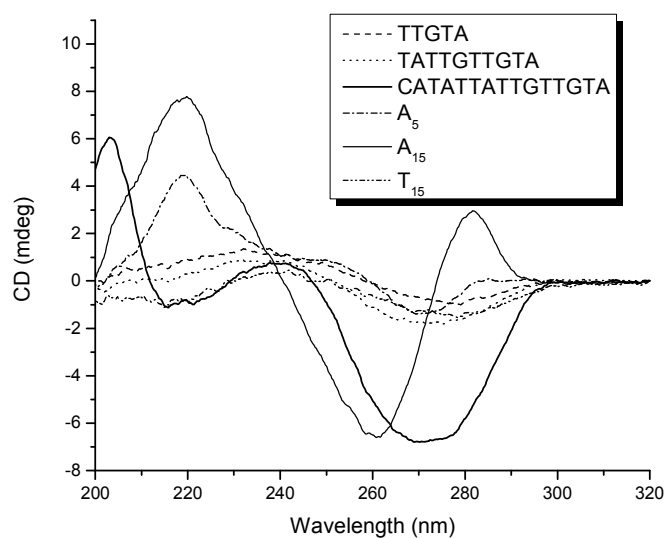


Figure 3.10. CD spectra of GNA single strands with varying length and sequence at concentrations of 20 μM (15mers), 30 μM (10mer), and 60 μM (5mers). Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride. Each curve is the average of ten measurements.

Proceeding further, we wanted to investigate the effect that the ionic strength has on preorganizing single strands of GNA. Since the presence of metal cations in solution help in the shielding of the negatively charged phosphates from each other, it was assumed that changes in the salt concentration should have a large effect on the conformational preorganization of the single strands. Figure 3.11 demonstrates that this is not the case for GNA. CD spectra of the sequence 3'-CATATTATTGTTGTA-2' in the presence of 0, 100, and 500 mM added sodium chloride do not differ significantly from each other in shape or intensity of the minimum peak at 270 nm. Only slight differences are observed in the maxima of 205 and 240 nm, and the minimum at 220 nm.

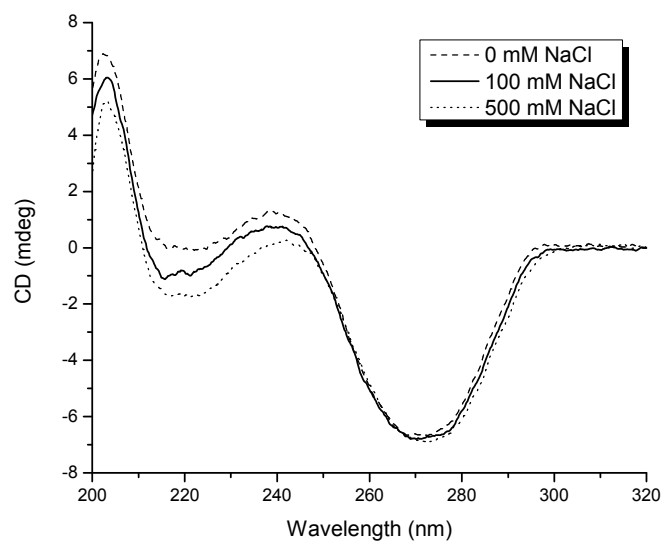


Figure 3.11. Effect of salt concentration on the CD spectra of GNA single strands at 15 °C with the sequence 3'-CACATTATTGTTGTA-2' at a concentration of 20 μ M. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with the stated sodium chloride concentration. Each curve is the average of ten measurements.

Chapter 3.3.3. Intercalators and NMR analysis of single glycol nucleosides

Additional support for the helical preorganization of GNA single strands was obtained from a comparison of the crystallographic data for single GNA nucleosides and NMR data of single nucleosides in organic and aqueous solutions.¹⁹ A comparison of the conformation of crystallized nucleosides with those in solution reveals that a preferred conformation also exists in solution. Curiously, all four nucleosides display sharp NMR signals of the protons in the dihydroxypropyl backbone, consistent with a somewhat defined structure and therefore preferred conformation (See ¹H NMR spectra in the Appendix to Chapter 2). A thorough NMR analysis of (*S*)-1-(2,3-dihydroxypropyl)cytosine reveals a similar conformation of the nucleoside in both the crystallized¹ form and in aqueous and organic solutions, as shown in Figure 3.12. Both the crystallized and NMR determined nucleosides adopt the same staggered conformation with an *anti* orientation of the nucleobase and a *gauche* conformation with respect to the vicinal C-O bonds.

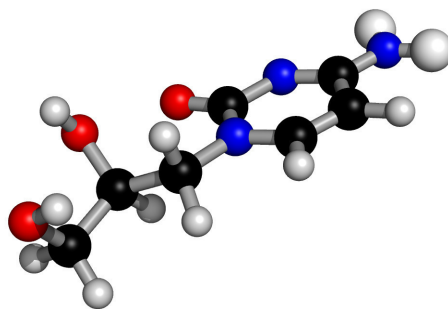


Figure 3.12. Structure of (*S*)-1-(2,3-dihydroxypropyl)cytosine determined by NMR.

Finally, the presence of an ideally preorganized backbone would imply that the duplex structure is quite rigid. To probe this assumption, we investigated the interaction of GNA duplexes with the well known intercalators ethidium bromide and proflavine.²⁶ Intercalation between neighboring base pairs requires the unwinding of the backbone, resulting in a modified backbone conformation.²⁷ Table 3.8 shows the results for intercalation of ethidium and proflavine into GNA and DNA duplexes. For the GNA duplex, addition of either intercalator has essentially no effect on the melting temperature. However, in DNA, the addition of the two intercalators results in a stabilization of the duplex in the amount of 6 °C and 4 °C for ethidium and proflavine, respectively. In addition to an increase in duplex melting temperature, a red-shift of the longest-wavelength absorption band of each intercalator would indicate binding by intercalation.²⁶ For ethidium, the absorption band at 477 nm shows no change in the GNA solution while it is shifted to 497 nm in the DNA solution. Furthermore, for proflavine, the absorption band at 442 nm shows no change in the GNA solution while it is shifted to 450 nm in the DNA solution. Finally, intercalation of ethidium and proflavine is accompanied by an increase and decrease in fluorescence intensity, respectively (Table 3.8).²⁶ Accordingly, the fluorescence intensity of ethidium is increased by 164% in the DNA solution while it remains essentially unchanged for the GNA solution. The fluorescence intensity of proflavine is decreased by 48% in the presence of DNA and by 15% in the presence of GNA. Although the decrease in fluorescence intensity of proflavine in the presence of GNA might suggest intercalation, the UV melting and absorbance data would argue the opposite way. It is quite possible

that the slight decrease in fluorescence is due to some other interaction of the proflavine with the GNA duplex (i.e. – groove binding or outside stacking). Identical experiments with concentrations of intercalators at 150 μM instead of 15 μM show similar results. Therefore, it can be concluded that the duplex structure of GNA does not allow for the unwinding of the backbone and insertion of intercalators between base pairs.

Table 3.8. Effects of intercalators on GNA and DNA duplexes^[a]

	T_m	ΔT_m	Max. Abs. ^[b] (nm)	Fluorescence ^[b]	Δ Fluorescence (%)
<i>GNA duplexes</i>					
No Intercalator	70.3	-	-	-	-
Ethidium	-	-	477	232	-
Duplex + Ethidium	69.8	-0.5	477	227	-2
Proflavine	-	-	442	13060	-
Duplex + Proflavine	70.2	-0.1	442	11083	-15
<i>DNA duplexes</i>					
No Intercalator	47.3	-	-	-	-
Ethidium	-	-	477	232	-
Duplex + Ethidium	53.3	6.0	497	613	164
Proflavine	-	-	442	13060	-
Duplex + Proflavine	51.3	4.0	450	6734	-48

[a] Measured in 100 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium chloride and 2 μM duplex with the sequence 5'(3')-CAC ATT ATT GTT GTA-3'(2') and its complement. Where stated is also included 15 μM of intercalator. [b] Measurements performed at 25 °C.

Chapter 3.4. Crosspairing of GNA single strands with DNA and RNA

Chapter 3.4.1. Initial crosspairing results of (S)- and (R)-GNA with DNA and RNA

Upon the introduction of a new nucleic acid analog with self-pairing ability usually comes interest as to how this new analog interacts with DNA and RNA. Much of this interest is aimed towards using the nucleic acid analog as a molecule to interfere with gene expression inside the cell. This can come on two different levels; first by interacting with mRNA and causing degradation of the mRNA, and second by interacting directly with the duplex DNA copy of the gene. The second method can prevent gene expression simply by not allowing the cell's machinery to access the gene in question. Regulating gene expression in this manner can be very powerful considering the rules that govern interaction of nucleic acids are very simple (Watson-Crick base pairing rules) and specific. This is compared to using small molecules to regulate gene expression in which one needs to know much more information about the protein to produce an effect (protein sequence, active site structure, mechanism of action, etc.).²⁸

Along these lines, we wanted to investigate how GNA interacts with both DNA and RNA. Since GNA has the attractive feature of easy synthetic access to both (*R*)- and (*S*)- enantiomers, we wanted to test the crosspairing ability of both enantiomers and determine whether any difference exists between their interactions with DNA and RNA. As shown in Table 3.9, 15-mer oligonucleotides comprised of poly-adenine or poly-thymine were used as an initial test of crosspairing ability.¹ It was hoped that using symmetrical oligonucleotides would allow for the observation of both parallel and

antiparallel pairing. Table 3.9 shows that (*S*)-GNA T₁₅ pairs strongly with (*S*)-GNA A₁₅ as expected ($T_m = 65$ °C). Furthermore (*S*)-GNA T₁₅ pairs weakly with (*R*)-GNA A₁₅ ($T_m = 38$ °C) and RNA A₁₅ ($T_m = 33$ °C). No sigmoidal melting is observed with DNA A₁₅. Besides pairing with (*S*)-GNA T₁₅, (*S*)-GNA A₁₅ also pairs weakly with (*R*)-GNA T₁₅ ($T_m = 35$ °C), DNA T₁₅ ($T_m = 21$ °C), and RNA U₁₅ ($T_m = 28$ °C). The (*R*)- enantiomers display weak crosspairing with the opposite enantiomeric form of GNA, but not with either DNA or RNA. The DNA and RNA oligonucleotides form stable duplexes with themselves ($T_m = 40$ °C for DNA, 29 °C for RNA) and each other ($T_m = 18$ °C and 35 °C), as expected. Interestingly, the two (*S*)-GNA:RNA hybrids show thermal stabilities similar to that of the RNA homoduplex (28 °C and 33 °C for (*S*)-GNA:RNA versus 29 °C for RNA:RNA).

Table 3.9. Thermal stabilities (T_m [°C]) for the cross-pairing between strands of (*S*)-GNA, (*R*)-GNA, RNA, and DNA^[a]

	(<i>S</i>)-GNA A ₁₅	(<i>R</i>)-GNA A ₁₅	DNA A ₁₅	RNA A ₁₅
(<i>S</i>)-GNA T ₁₅	65	38	no T_m	33
(<i>R</i>)-GNA T ₁₅	35	64	no T_m	no T_m
DNA T ₁₅	21	no T_m	40	35
RNA U ₁₅	28	no T_m	18	29

[a] Measured in 10 mM sodium phosphate buffer (pH=7.0) with 150 mM NaCl and 1 mM EDTA at 2 μ M duplex concentration.

To further test this crosspairing ability, we decided to investigate duplex formation by CD spectroscopy. Shown in Figure 3.13 are the CD spectra of

homoduplexes of (*S*)-GNA, RNA, and DNA using the same poly-adenine, poly-thymine pairing system. The (*S*)- enantiomer of GNA was chosen for this study based on the results from Table 3.9 which only show interaction of this enantiomer with the DNA and RNA oligonucleotides. The CD spectrum of the GNA duplex displays very few similarities to the spectra of either DNA or RNA. One apparent observation is that the minima and maxima of the GNA spectrum are the opposite sign in the DNA and RNA spectra; what might actually suggest the formation of an opposite handed duplex (i.e. – a left-handed duplex for GNA).

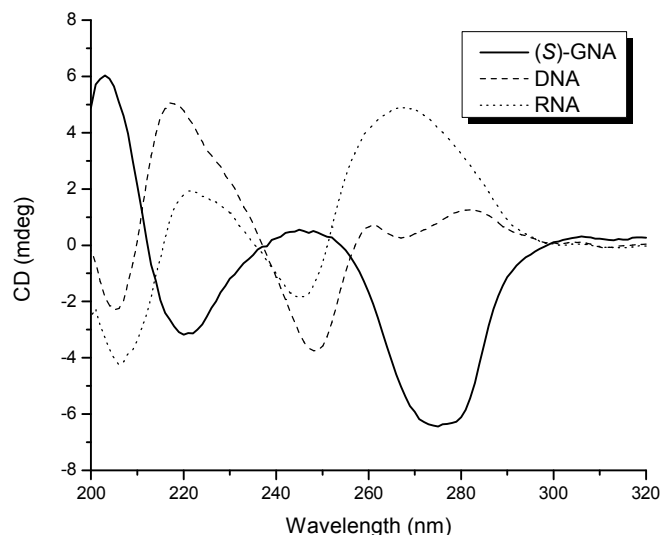


Figure 3.13. CD spectra of GNA, DNA, and RNA homoduplexes of the sequence A₁₅:T₁₅ at a duplex concentration of 4 μ M. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 150 mM sodium chloride and 1 mM EDTA at 15 $^{\circ}$ C. Each curve is the average of five measurements.

Proceeding further, the CD spectrum of the most stable (*S*)-GNA:RNA hybrid, (*S*)-GNA T₁₅:RNA A₁₅, was measured along with the corresponding single strands.

Figure 3.14 clearly shows that the CD spectrum is not simply the sum of the individual oligonucleotides; indicating the formation of a (*S*)-GNA:RNA heteroduplex. A similar CD spectrum was observed for the (*S*)-GNA A₁₅:RNA T₁₅ heteroduplex. In both spectra, the biggest indication of duplex formation is the peak at 283 nm which only appears in the curve for the heteroduplex. Analysis of the corresponding (*S*)-GNA T₁₅:DNA A₁₅ duplex in Figure 3.15 shows that the CD spectrum closely resembles that of the sum of the individual oligonucleotides. A similar result was obtained for the (*S*)-GNA A₁₅:DNA T₁₅ duplex indicating no formation of a duplex between GNA and DNA. This is consistent with previous results of our lab and the melting experiments in Table 3.9.

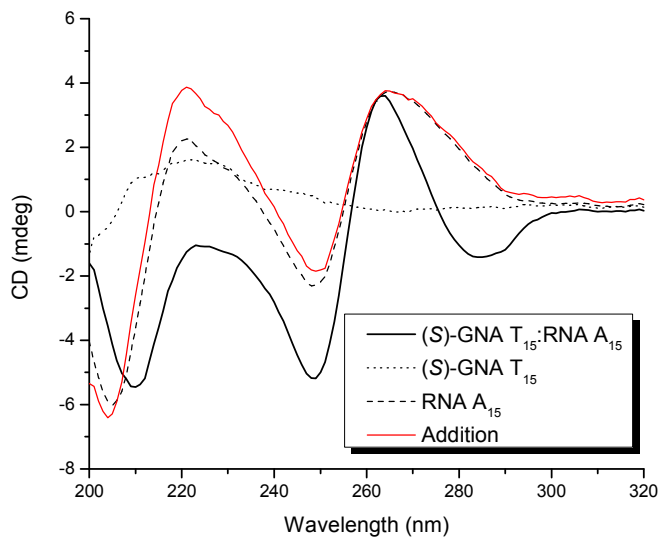


Figure 3.14. CD spectra of a (*S*)-GNA T₁₅:RNA A₁₅ heteroduplex at a duplex concentration of 4 μ M and the corresponding single strands at a concentration of 4 μ M. Also shown is the curve representing the addition of the spectra of the two single strands. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 150 mM sodium chloride and 1 mM EDTA at 15°C. Each curve is the average of five measurements.

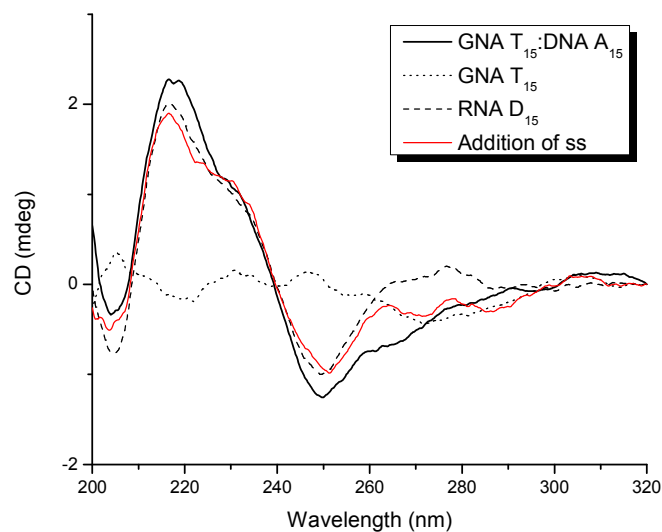


Figure 3.15. CD spectra of a (*S*)-GNA T₁₅:DNA A₁₅ heteroduplex at a duplex concentration of 4 μM and the corresponding single strands at a concentration of 4 μM. Also shown is the curve representing the addition of the spectra of the two single strands. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 150 mM sodium chloride and 1 mM EDTA at 15°C. Each curve is the average of five measurements.

Moreover, temperature dependent analysis of the CD signal at 283 nm from 5 to 60 °C, shown in Figure 3.16a, provides a T_m of 34 °C after a sigmoidal fit to the data. This is consistent with the melting point determined by UV melting ($T_m=33$ °C). A similar analysis at 283 nm of the (*S*)-GNA A₁₅:RNA T₁₅ duplex provides a T_m of 27 °C which is also consistent with the value from UV melting ($T_m=28$ °C). Finally, as shown in Figure 3.16b, Job Plot analysis of the (*S*)-GNA T₁₅:RNA A₁₅ heteroduplex indicates the formation of a 1:1 complex of these two oligonucleotides.²⁸

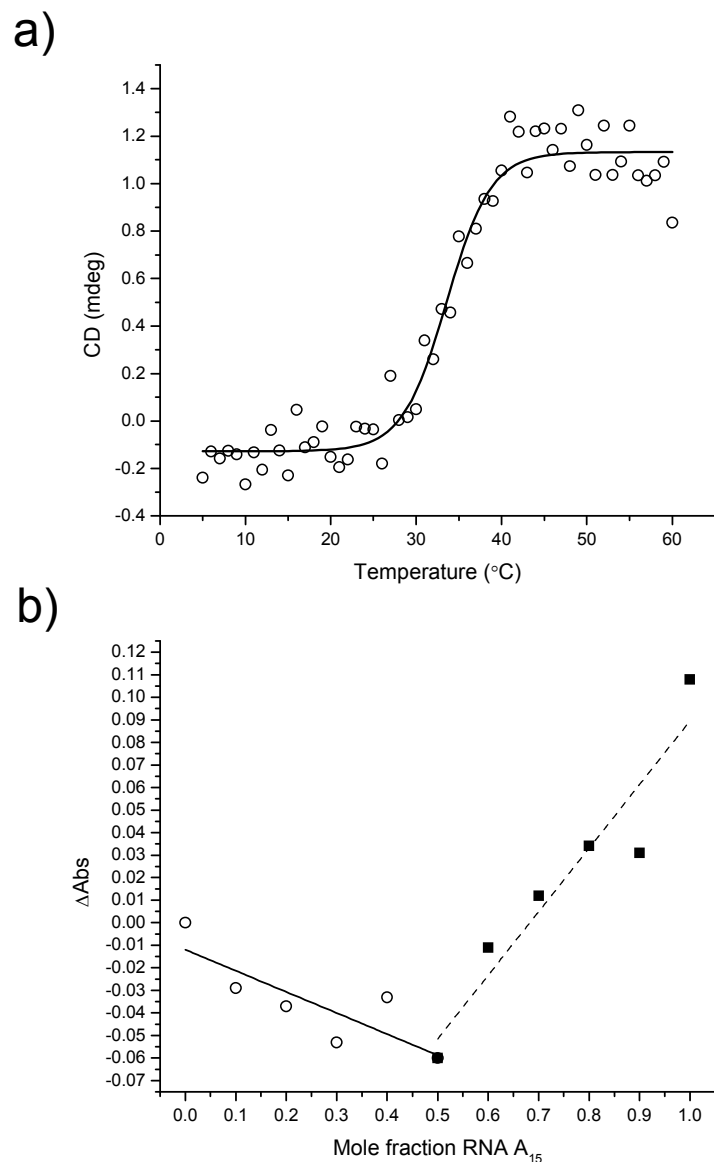


Figure 3.16. CD and Job plot analysis of the (S)-GNA T₁₅:RNA A₁₅ heteroduplex. a) Temperature-dependent CD measurement at a duplex concentration of 4 μ M in 10 mM sodium phosphate buffer (pH=7.0) with 150 mM sodium chloride and 1 mM EDTA. The CD signal (red dots) was measured at 283 nm as a function of temperature with a heating rate of 1 $^{\circ}$ C/min. A sigmoidal fit to the data (black line) yields a melting temperature of 34 $^{\circ}$ C based on the inflection point of the curve. b) Job plot analysis of (S)-GNA T₁₅ pairing with RNA A₁₅. Measurements were taken by mixing the appropriate amounts of a 5 μ M stock solution of each oligonucleotide in 10 mM sodium phosphate buffer (pH=7.0) with 150 mM sodium chloride and 1 mM EDTA. Linear fits to the data produces two lines which intersect at a value of 0.48.

Chaper 3.4.2. Crosspairing of mixed sequences of (S)-GNA and RNA

Once it was established that there is a real interaction between (S)-GNA and RNA using strands composed purely of adenine and thymine, we wanted to test how G:C base pairs affect the duplex formation properties.¹ Contrary to what one would expect, previous experience in our lab had shown that G:C base pairs destabilize these heteroduplexes. To further test this, both UV melting and CD spectroscopy were used to analyze (S)-GNA:RNA T₁₅:A₁₅ heteroduplexes containing varying numbers of G:C base pairs. Accordingly, Figure 3.17a shows the thermal melting curves of these heteroduplexes containing zero, one, and three G:C base pairs. Again, consistent with our previous results, the incorporation of one G:C base pair ($T_m = 24\text{ }^\circ\text{C}$) and three G:C base pairs (no observable T_m) lead to significant destabilizations compared to the heteroduplex containing no G:C base pairs ($T_m = 33\text{ }^\circ\text{C}$). As shown in Figure 3.17b, CD spectroscopy also points to the destabilization of these heteroduplexes based on the slow disappearance of the peaks at 283 and 210 nm as the number of G:C base pairs is increased from zero to three.

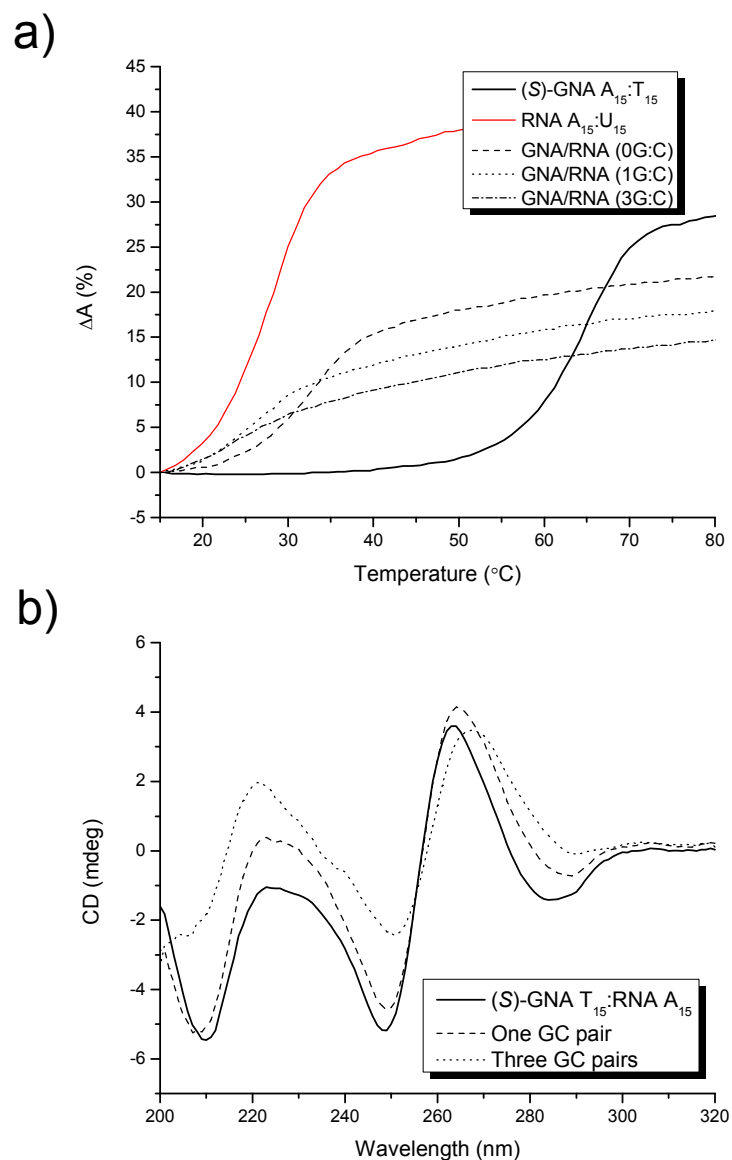


Figure 3.17. Crosspairing of mixed sequences of (*S*)-GNA with RNA. a) Thermal melting curves monitored by UV spectroscopy of 1:1 mixtures (2 μM each strand) of (*S*)-GNA $A_{15}:T_{15}$ ($T_m = 65^{\circ}\text{C}$), RNA $A_{15}:U_{15}$ ($T_m = 29^{\circ}\text{C}$), RNA:(*S*)-GNA $A_{15}:T_{15}$ ($T_m = 33^{\circ}\text{C}$), RNA:(*S*)-GNA $A_7CA_7:T_7GT_7$ ($T_m = 24^{\circ}\text{C}$), and RNA:(*S*)-GNA $A_3CA_3CA_3CA_3:T_3GT_3GT_3GT_3$ (no T_m). b) CD spectra of a (*S*)-GNA:RNA heteroduplexes with no G:C base pairs, one G:C pair [(*S*)-GNA $T_7GT_7:\text{RNA } A_7CA_7$], and three G:C base pairs [(*S*)-GNA $T_3GT_3GT_3GT_3:\text{RNA } A_3CA_3CA_3CA_3$] at a duplex concentration of 4 μM and a temperature of 15°C . Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 150 mM sodium chloride and 1 mM EDTA. Each curve is the average of five measurements.

Based on these surprising results, we were interested in performing a mismatch experiment of the five natural bases (A, G, C, T, U) and three synthetic bases (N, D, I) in GNA:RNA heteroduplexes. One initial theory was that the 2-amino group of guanine might be involved in a steric clash in the minor groove, thereby destabilizing heteroduplex formation. The incorporation of the three glycol nucleotides (*S*)-**N** (2-aminopurine), (*S*)-**D** (diaminopurine), and (*S*)-**I** (inosine) into GNA oligonucleotides was aimed towards testing this theory. The structures of these three nucleotides, compared to the natural ones, would allow us to directly test how the 2-amino group functions in GNA:RNA heteroduplex formation. For example, 2-aminopurine (N) provides a base pair with thymine of approximately equal stability to that of an A:T pair (Figure 3.18).³⁰ Furthermore, inosine (I) should pair with cytosine to form a base pair that is similar in strength to an A:T pair, but slightly destabilized compared to a G:C pair in the same position.³¹ Finally, 2,6-diaminopurine (D) is capable of base pairing with thymine and this base pair displays a stability greater than that of an A:T base pair, but slightly lower than that of a G:C base pair.³²

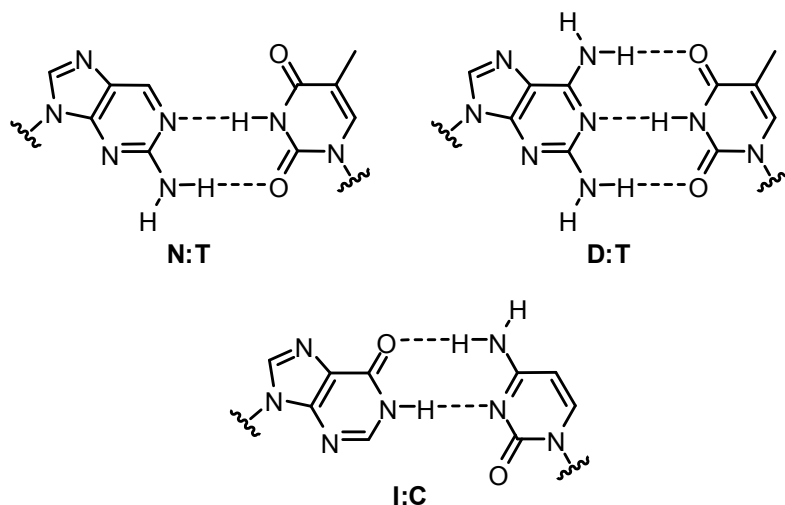
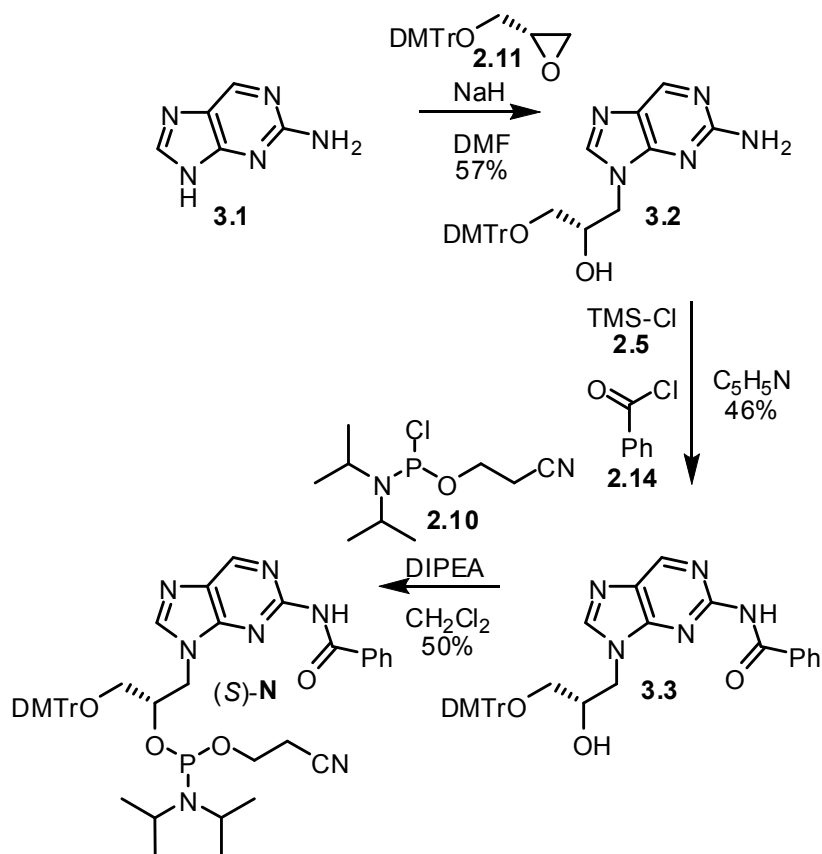


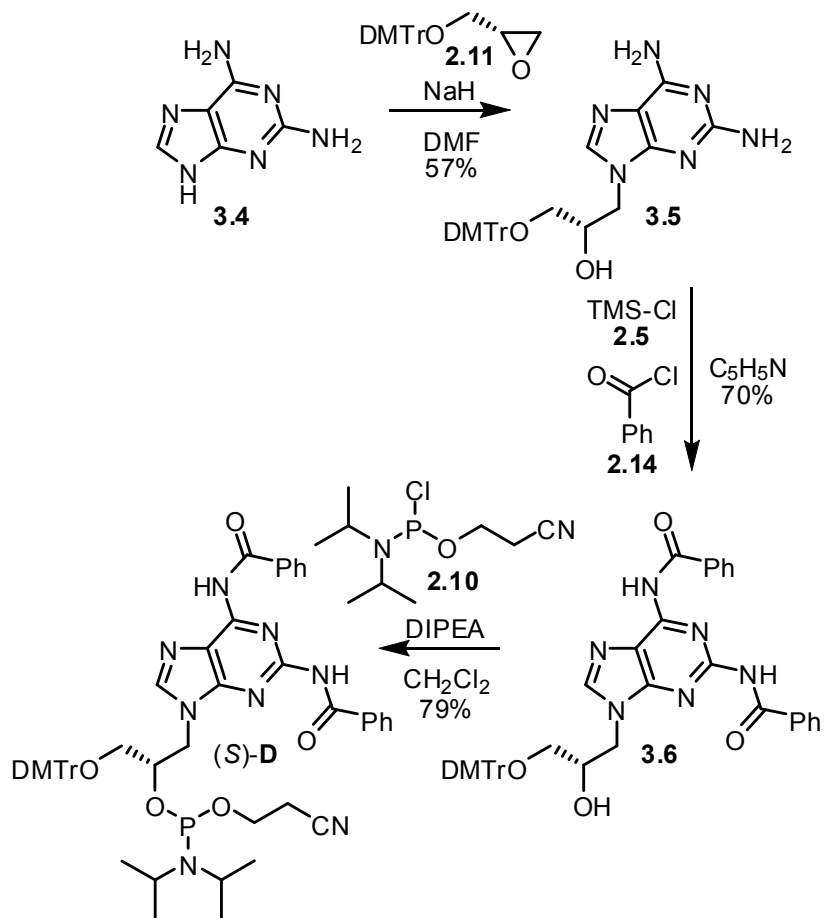
Figure 3.18. Structure of the three unnatural base pairs used in this study and their possible hydrogen bonding patterns.

For the synthesis of (*S*)-**N** glycol nucleoside phosphoramidite (Scheme 3.1), compound **2.11** was first ring-opened using 2-aminopurine (**3.1**) and 0.2 equivalents of sodium hydride to produce compound **3.2** in 57% yield. It should be noted that a significant amount (27%) of the *N*-7 ring opened product was also observed and that the correct *N*-9 regioisomer was determined using NMR. Protection of the exocyclic amine of compound **3.2** was accomplished by first transiently protecting the 2'-hydroxyl group using trimethylsilyl chloride (**2.5**) and then reacting with benzoyl chloride (**2.14**) in pyridine to produce compound **3.3** in 46% yield. Compound **3.3** was finally converted to the phosphoramidite (*S*)-**N** with 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.9**) and excess *N,N*-diisopropylethylamine in 50% yield.



Scheme 3.1. Synthesis of (S)-N phosphoramidite.

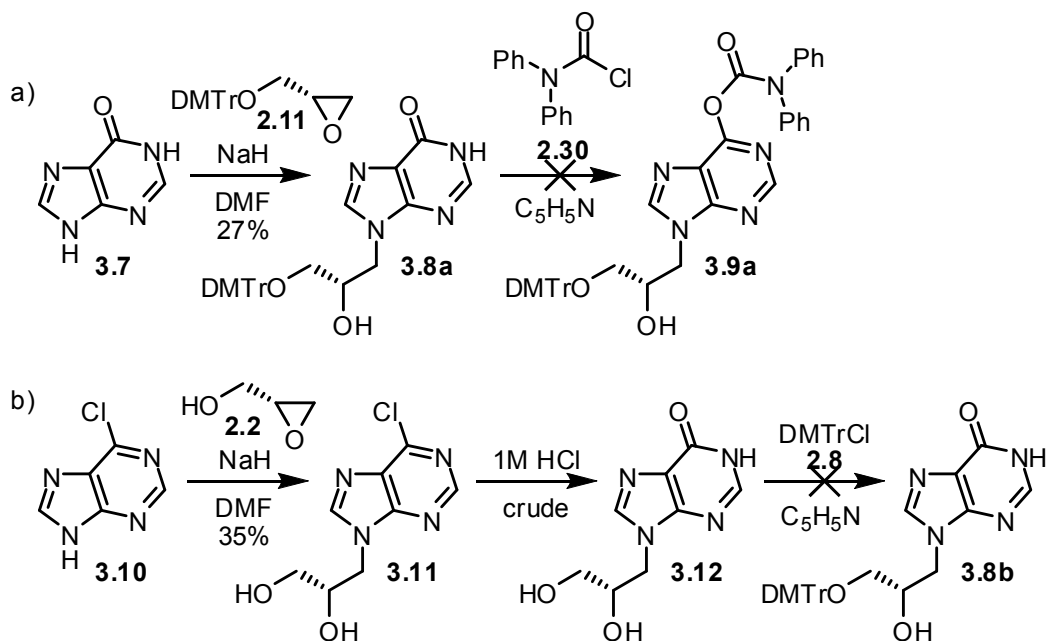
Similarly, the synthesis of (S)-D phosphoramidite starts with ring opening of compound **2.11** using diaminopurine (**3.4**) and 0.2 equivalents of sodium hydride to produce compound **3.5** in 57% yield (Scheme 3.2). Protection of the two exocyclic amines of compound **3.5** was accomplished by first reacting with trimethylsilyl chloride (**2.5**) and then with benzoyl chloride (**2.14**) in pyridine to afford compound **3.6** in 70% yield. Further conversion to the phosphoramidite (S)-D was afforded in 79% yield by the reaction of compound **3.6** with 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.9**) and excess *N,N*-diisopropylethylamine.



Scheme 3.2. Synthesis of (*S*)-**D** phosphoramidite.

The synthesis of (*S*)-**I** phosphoramidite proved to be more problematic. Initial attempts of ring-opening compound **2.11** using inosine (**3.7**) and 0.2 equivalents of sodium hydride yielded product **3.8a** which appeared to be a mixture of two tautomeric forms (Scheme 3.3a). However, subsequent protection attempts of compound **3.8a** using diphenylcarbonyl chloride (**2.30**) and *N,N*-diisopropylethylamine proved unsuccessful suggesting that compound **3.9a** was not the desired product. A second route which started with the ring opening of (*R*)-glycidol (**2.1**) with 6-chloropurine (**3.10**) and 0.17

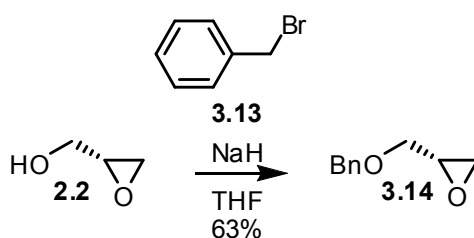
equivalents of potassium carbonate yielded the desired product **3.11** in 35% yield (Scheme 3.3b). Subsequent hydrolysis of compound **3.11** using 1 M HCl yielded compound **3.12** based on TLC. Unfortunately, the product could not be purified and was taken crude to the next step. Subsequent conversion to compound **3.8b** using 4,4'-dimethoxytrityl chloride (**2.7**) in pyridine was unsuccessful.



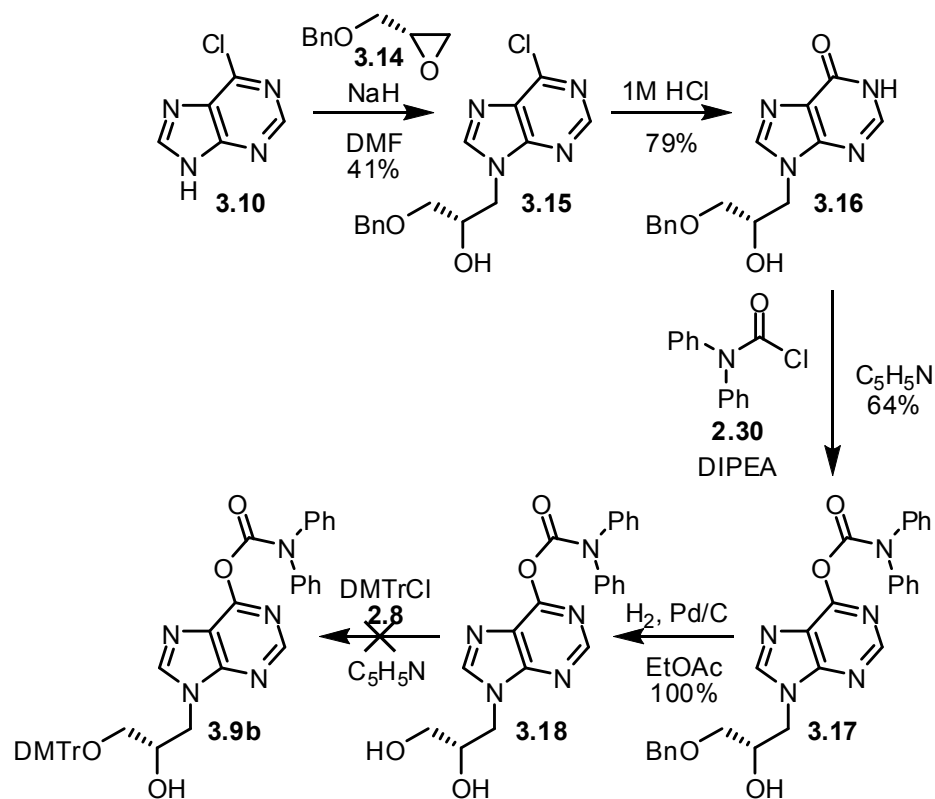
Scheme 3.3. Attempted synthesis of (*S*)-**I** phosphoramidite starting with: a) inosine, and b) 6-chloropurine.

Next, it was envisioned that the use of benzyl protection on the 3'-hydroxyl group would allow for the easy purification of the polar synthetic intermediates. Accordingly, (*R*)-glycidol was protected by the reaction with benzyl bromide (**3.13**) and sodium hydride to yield compound **3.14** in 63% yield (Scheme 3.4).³³ Subsequent ring-opening

of compound **3.14** with 6-chloropurine (**3.10**) and 0.2 equivalents of sodium hydride afforded the desired compound **3.15** in 41% yield (Scheme 3.5). Compound **3.15** was then hydrolyzed with 1 M HCl to provide compound **3.16** in 79% yield. The benzyl group allows for a simple purification of compound **3.16** to remove any salt byproducts. Reaction of compound **3.16** with diphenylcarbamoyl chloride and *N,N*-diisopropylethylamine in pyridine yields compound **3.17** in 64% yield. Catalytic hydrogenation was then used to produce compound **3.18** in quantitative yield. Unfortunately, an attempt to reprotect the 3'-hydroxyl group using 4,4'-dimethoxytrityl chloride in pyridine did not yield the desired compound **3.9b**.



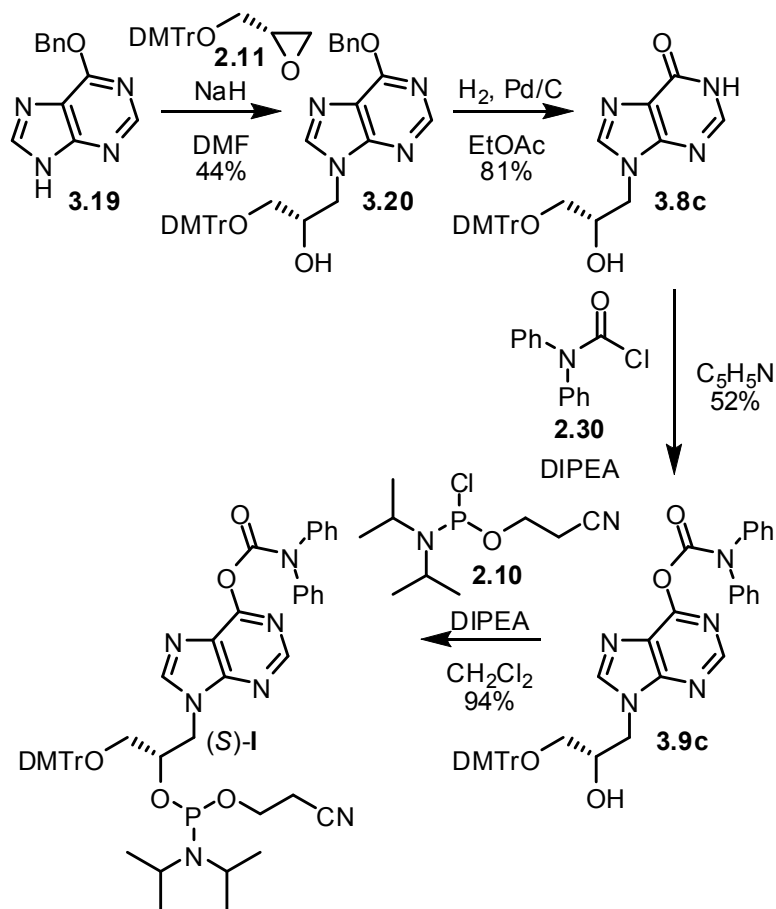
Scheme 3.4. Synthesis of benzylglycidol.³²



Scheme 3.5. Attempted synthesis of (*S*)-I phosphoramidite using compound **3.14** for ring opening of 6-chloropurine.

Finally, based on the success of ring-opening using *O*-6-benzyloxyguanine (**2.36**) for the synthesis of (*S*)-G, (*S*)-G^{DPC}, and (*S*)-G*, *O*-6-benzyloxyuracil (**3.19**) was chosen for the ring-opening of compound **2.11** in the presence of 0.2 equivalents of sodium hydride (Scheme 3.6). Compound **3.20**, which was formed in 44% yield, could then be converted to compound **3.8c** using catalytic hydrogenation in 81% yield. Gratifyingly, compound **3.8c** was then converted to compound **3.9c** by the reaction with diphenylcarbamoyl chloride and *N,N*-diisopropylethylamine in 52% yield. Finally,

compound **3.9c** was reacted with 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite and excess *N,N*-diisopropylethylamine to produce phosphoramidite (*S*)-**I** in 94% yield.



Scheme 3.6. Successful synthesis of (*S*)-**I** phosphoramidite.

With these three new phosphoramidites in hand [(*S*)-**N**, (*S*)-**D**, (*S*)-**I**], it was possible to perform the mismatch experiment in which the middle base pair of the 15-mer heteroduplex is substituted for the four Watson-Crick bases in GNA and RNA and the three unnatural bases in GNA as shown in Table 3.10. As expected, the most stable

combinations are those where adenine and thymine (or uracil) are at the middle position of the two 15-mers yielding melting temperatures (T_m) of 28 °C and 32 °C. Interestingly, A:A (27 °C) and A:G (25 °C) also form base pairs with stabilities similar to an A:U match (28 °C). A guanine nucleotide in the middle of the GNA 15-mer pairs with cytosine and guanine in the RNA oligo with a T_m of 24 °C, but pairs weakly with both A and U. The cytosine nucleotide in the GNA oligo forms a duplex with a T_m of 26 °C and 25 °C when pairing with A and G, respectively. In this case, it is interesting that the cytosine is forming a slightly more stable base pair with adenine, rather than guanine. The other two combinations C:C and C:U are highly destabilizing to the duplex. Besides strongly pairing with adenine (32 °C), thymine also pairs relatively well with guanine (27 °C), but is highly destabilizing when paired with C or U. Although the 2-aminopurine glycol nucleotide pairs strongest with uracil (23 °C), it results in largely destabilized duplexes when paired with all four RNA nucleotides. As expected, the most stable pair for diaminopurine is the D:U pair ($T_m = 28$ °C), however, all other combinations are only moderately stable (24 °C for D:A, 23 °C for D:G), or unstable (D:C). Finally, contrary to what one would expect, the inosine glycol nucleotide pairs strongest with guanine ($T_m = 25$ °C), followed by adenine ($T_m = 24$ °C), then uracil ($T_m = 22$ °C), and finally cytosine ($T_m = 21$ °C).

Table 3.10. Thermal stabilities (T_m [°C]) of matched and mismatched base pairs in GNA/RNA heteroduplexes^[a]

		5'-AAA AAA A X A AAA AAA-3'				
		2'-TTT TTT T Y T TTT TTT-3'				
		X				
		A	G	C	U	
	A	27	25	21	28	
	G	≤ 20	24	24	22	
	C	26	25	≤ 20	≤ 20	
Y	T	32	27	≤ 20	≤ 20	
	N	22	21	≤ 20	23	
	D	24	23	≤ 20	28	
	I	24	25	21	22	

[a] Measured in 10 mM sodium phosphate buffer (pH=7.0) with 150 mM NaCl and 1 mM EDTA at 2 μM duplex concentration. Each data point is the mean of two measurements.

It was hoped that the mismatch experiment would shed light into the destabilization of (*S*)-GNA:RNA heteroduplexes by G:C base pairs, but unfortunately the data does not help with understanding this phenomenon. A close analysis of the mismatch data does not reveal a clear trend of how the nucleobases are interacting with each other inside the duplex. It is entirely possible that the nucleobases are interacting in a fashion different from that of Watson-Crick base pairs in which the purine edge comprised of C-6, N-1, and C-2 interacts with the pyrimidine edge comprised of C-4, N-3, and C-2 (i.e. – alternative hydrogen bonding). However, this does confirm the observation that G:C base pairs are unstable in the heteroduplexes used for this study.

More interesting is there are base pairs (A:A, C:A, and T:G) that are more stable than either of the G:C base pairs. Although there is a large amount of interest in using GNA to interact with DNA and RNA,³⁴⁻³⁵ the crosspairing with RNA is still poorly understood and an in-depth structural analysis (crystallography or NMR) is needed to sort out the interesting behavior. However, from the results presented above, it is clear that (*S*)-GNA is forming heteroduplex structures with stabilities similar to DNA:RNA heteroduplexes, albeit in duplexes composed of sequences rich in adenine and thymine.

Chapter 3.5. Conclusions

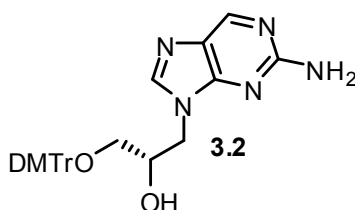
The initial discovery that GNA is capable of duplex formation that was thermally more stable than duplexes of DNA or RNA with the same sequence was surprising. Based on previous experiments in which single acyclic nucleosides were incorporated into DNA duplexes with decreased thermal stabilities, it was thought that two oligonucleotides composed entirely of acyclic nucleotides would not possess the conformational preorganization to form a duplex. The CD data presented here would suggest the opposite in which the GNA single strands are preorganized for duplex formation in solution, at least in some cases. An experiment probing the interactions of matched and mismatched base pairs in a GNA duplex shows that the hydrogen bonding interactions in GNA duplexes are, to a similar degree as in DNA, selective for the Watson-Crick base pairs. Experiments investigating the counter ion and solvent dependence of duplex formation show no significant differences between GNA and DNA duplexes. In contrast, measurements of the thermodynamic parameters ΔG , ΔH , and ΔS suggest that duplex formation is entropically favored for GNA, consistent with the CD results. Furthermore, the experiments with dangling nucleosides suggests that stacking interactions in GNA duplexes are thermodynamically more favorable than in DNA.

Unfortunately, the crosspairing of GNA single strands with DNA and RNA single strands is still poorly understood. Although it seems clear that only (*S*)-GNA pairs only with RNA, it is quite puzzling that this pairing is limited to sequences rich in adenine and thymine. Both UV melting and CD experiments shown evidence of duplex formation in

A:T rich heteroduplexes which gradually disappears as the number of G:C base pairs is increased in the sequence. A mismatch experiment, including the three unnatural bases N, D, and I does not help to increase our understanding of this crosspairing system. Although the synthesis of three unnatural glycol nucleoside phosphoramidites is presented [(*S*)-**D**, (*S*)-**I**, and (*S*)-**N**], there is much room for improvement in these synthesis. For other applications of (*S*)-**D** and (*S*)-**N**, formamidinium protection of the exocyclic amines should be investigated based on the ease of incorporation and stability of these protection groups in the (*S*)-**A*** and (*S*)-**G*** phosphoramidites.³⁶

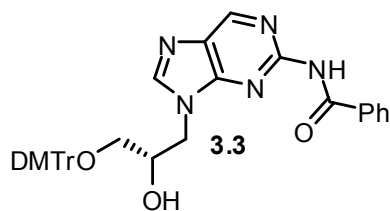
Chapter 3.6. Experimental procedures

General procedures and reagents. NMR spectra were recorded on a Bruker DRX-500 (500 MHz), DMX-360 (360 MHz), or DMX-300 (300 MHz) spectrometer. High-resolution mass spectra were obtained with a Micromass AutoSpec or Thermo LTQ-FT instrument using ES ionization. Infrared spectra were recorded either on a Perkin Elmer 1600, Nicolet 510, or Bruker alpha series FTIR spectrometer. Solvents and reagents were used as supplied from Aldrich, Acros, Fluka, or TCI. Reactions were performed under an atmosphere of argon or nitrogen unless otherwise specified.

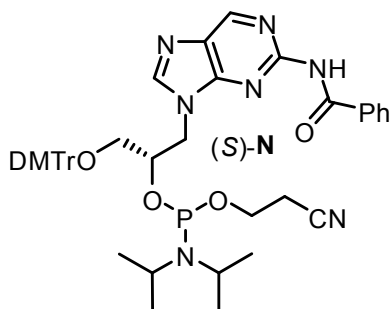


Compound 3.2. To a suspension of 2-aminopurine (510 mg, 3.8 mmol) in anhydrous DMF (8 mL) under nitrogen was added NaH (30 mg, 1.1 mmol, 60% in mineral oil) and the solution allowed to stir under nitrogen for one hour. A solution of compound **2.11** (1.35 g, 3.6 mmol) in DMF (7 mL) was added to the first solution and then heated to 90 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column chromatography starting with 100:1 EtOAc:Et₃N, then 50:1:0.01 EtOAc:MeOH:Et₃N, and finally eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **3.2** as a light yellow foam (1.10 g, 57%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 8.52 (s, 1H), 7.67 (s, 1H), 7.43 (m, 2H), 7.34-7.18 (m,

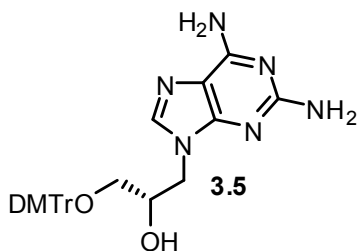
7H), 6.82 (d, $J = 8.9$ Hz, 4H), 5.20 (s, 2H), 4.57 (br, 1H), 4.33 (m, 1H), 4.16 (m, 2H), 3.78 (s, 6H), 3.26 (dd, $J = 9.5, 4.9$ Hz, 1H), 3.10 (dd, $J = 9.5, 5.4$ Hz, 1H).



Compound 3.3. To a solution of compound **3.2** (765 mg, 1.5 mmol) in anhydrous pyridine (12 mL) under nitrogen was added trimethylsilyl chloride (0.76 mL, 6.0 mmol) and allowed to stir for two hours at room temperature. The solution was then cooled to 0 °C followed by the dropwise addition of benzoyl chloride (0.26 mL, 2.2 mmol). The solution was allowed to gradually warm up to room temperature and stir for another two hours after which the solution was cooled back to 0 °C. Water (5 mL) was added, followed 10 minutes later by the addition of 25% aqueous ammonia (5 mL). After stirring for another 10 minutes, the solution was diluted with water and extracted two times with methylene chloride. The organic phase was dried over Na₂SO₄ and concentrated. The product was purified via column chromatography starting with 100:1 EtOAc:Et₃N then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound **3.3** as a light yellow foam (420 mg, 46%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 8.96 (br, 1H), 8.90 (s, 1H), 7.96 (s, 1H), 7.93 (s, 1H), 7.57 (m, 1H), 7.48 (m, 2H), 7.40 (m, 2H), 7.31-7.16 (m, 7H), 6.80 (m, 4H), 4.57 (m, 1H), 4.36 (dd, $J = 14.1, 6.5$ Hz, 1H), 4.28 (m, 1H), 3.77 (s, 6H), 3.42 (dd, $J = 9.6, 5.0$ Hz, 1H), 2.99 (dd, $J = 9.4, 7.9$ Hz, 1H).

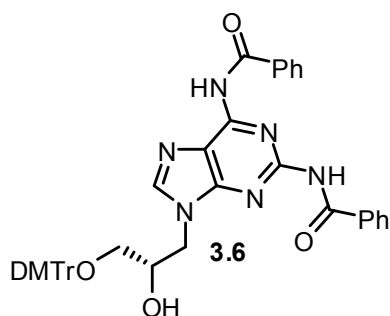


Compound (S)-N. To a nitrogen purged solution of compound **3.3** (420 mg, 0.68 mmol) and *N,N*-diisopropylethylamine (0.71 mL, 4.1 mmol) in methylene chloride (11 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.23 mL, 1.0 mmol) dropwise and the solution stirred for two hours at room temperature under nitrogen. The solution was diluted with methylene chloride and washed once with saturated aqueous NaHCO_3 , dried over Na_2SO_4 , and finally concentrated by rotoevaporation. The crude product was purified by column chromatography starting with 1:2:0.01 Hexanes:EtOAc: Et_3N then eluting with 100:1 EtOAc: Et_3N to afford compound (S)-N as a light yellow foam (275 mg, 50%). ^{31}P -NMR (121 MHz, CDCl_3) δ (ppm) 150.0, 149.8.



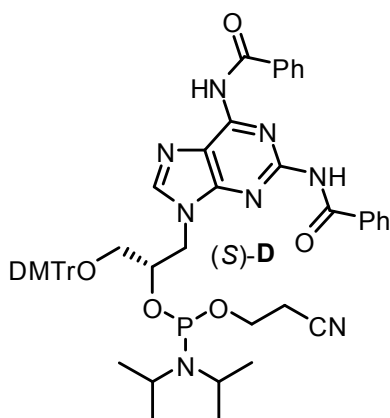
Compound 3.5. To a suspension of 2,6-diaminopurine (1.02 g, 6.8 mmol) in anhydrous DMF (14 mL) under argon was added NaH (55 mg, 1.4 mmol, 60% in mineral oil) and allowed to stir under argon for one hour. A solution of compound **2.11** (2.43 g, 6.5 mmol) in DMF (13 mL) was added to the first solution and then heated to 110 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting

oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column chromatography starting with 50:1:0.01 EtOAc:MeOH:Et₃N then eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **3.5** as a yellow foam (2.04 g, 57%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 7.43 (m, 2H), 7.34-7.16 (m, 8H), 6.81 (m, 4H), 6.02 (s, 2H), 5.04 (s, 2H), 4.26-4.06 (m, 3H), 3.76 (s, 6H), 3.29 (dd, J = 9.6, 4.7 Hz, 1H), 3.05 (dd, J = 9.5, 6.1 Hz, 1H).

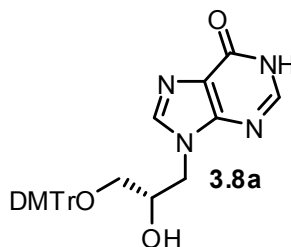


Compound 3.6. To a solution of compound **3.5** (810 mg, 1.5 mmol) in anhydrous pyridine (12 mL) under nitrogen was added trimethylsilyl chloride (0.78 mL, 6.2 mmol) and allowed to stir for two hours at room temperature. After two hours, the solution was cooled to 0 °C followed by the dropwise addition of benzoyl chloride (0.54 mL, 4.6 mmol). The solution was allowed to gradually warm up to room temperature and stir for another two hours after which the solution was cooled back to 0 °C. Water (5 mL) was added, followed 10 minutes later by the addition of 25% aqueous ammonia (5 mL). After stirring for another 10 minutes, the solution was diluted with water and extracted two times with methylene chloride. The organic phase was dried over Na₂SO₄ and concentrated. The product was purified via column chromatography starting with 100:1 EtOAc:Et₃N then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound **3.6** as a

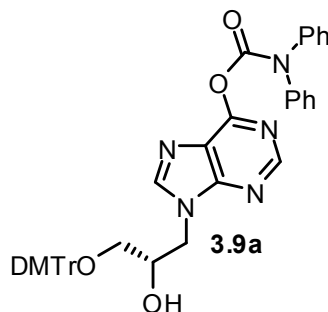
light yellow foam (790 mg, 70%). $^1\text{H-NMR}$ (300 MHz, CDCl_3) δ (ppm) 9.48 (s, 1H), 9.31 (br, 1H), 7.99 (d, $J = 7.5$ Hz, 2H), 7.92 (d, $J = 7.7$ Hz, 2H), 7.79 (s, 1H), 7.60 (m, 1H), 7.54-7.40 (m, 5H), 7.36-7.18 (m, 9H), 6.83 (d, $J = 8.8$ Hz, 4H), 4.55-4.33 (m, 3H), 3.77 (s, 6H), 3.49 (m, 1H), 3.09 (m, 1H).



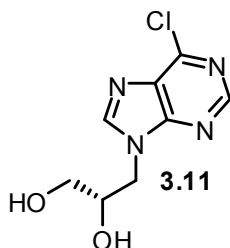
Compound (S)-D. To a nitrogen purged solution of compound **3.6** (610 mg, 0.83 mmol) and *N,N*-diisopropylethylamine (0.87 mL, 5.0 mmol) in methylene chloride (13.5 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.28 mL, 1.3 mmol) dropwise and the solution stirred for two hours at room temperature under nitrogen. The solution was diluted with methylene chloride and washed once with saturated aqueous NaHCO_3 , dried over Na_2SO_4 , and finally concentrated by rotoevaporation. The crude product was purified by column chromatography starting with 1:2:0.01 Hexanes:EtOAc:Et₃N then eluting with 1:4:0.01 Hexanes:EtOAc:Et₃N to afford compound (S)-D as a light yellow foam (610 mg, 79%). $^{31}\text{P-NMR}$ (121 MHz, CDCl_3) δ (ppm) 150.1, 149.6.



Attempted synthesis of Compound 3.8a. To a solution of inosine (246 mg, 1.81 mmol) in 3.8 mL of anhydrous DMF was added sodium hydride (60% in mineral oil, 15 mg, 0.38 mmol) and the solution allowed to stir under argon for one hour. A solution of compound **2.11** (650 mg, 1.72 mmol) in 3.5 mL of anhydrous DMF was added to the above solution and heated to 110 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via flash chromatography eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **3.8a** as a light yellow foam (250 mg, 27%). The ¹H NMR appears to be consistent with a mixture of two tautomeric forms. ¹H-NMR (360 MHz, CDCl₃) δ (ppm) 7.76 (d, J = 5.8 Hz, 2H), 7.68 (s, 1H), 7.61-7.50 (m, 5H), 7.48-7.37 (m, 8H), 7.32 (q, J = 7.3 Hz, 4H), 7.22 (m, 2H), 6.87 (m, 8H), 4.99 (m, 1H), 4.79 (br, 1H), 4.58-4.47 (m, 2H), 4.10-4.00 (m, 2H), 3.80 (s, 6H), 3.77 (s, 6H), 3.52 (dd, J = 9.1, 4.6 Hz, 1H), 3.43 (dd, J = 9.5, 4.4 Hz, 1H), 3.32 (m, 2H).

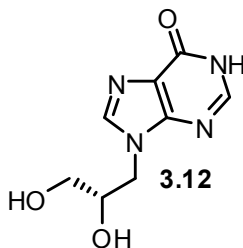


Attempted synthesis of Compound 3.9a. Compound **3.8a** (51 mg, 0.10 mmol) and diphenylcarbonyl chloride (28 mg, 0.12 mmol) were dissolved in anhydrous pyridine (0.80 mL) under argon. To this solution was added *N,N*-diisopropylethylamine (21 μ L, 0.12 mmol) and then allowed to stir for one hour at room temperature. The solution was diluted with methylene chloride, washed with saturated aqueous NaHCO_3 , dried over Na_2SO_4 , and finally concentrated. The crude product was purified by flash chromatography starting with 1:2:0.01 Hexanes:EtOAc: Et_3N , then eluting with 100:1 EtOAc: Et_3N . ^1H NMR was inconclusive.

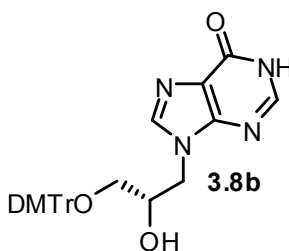


Compound 3.11. To a solution of 6-chloropurine (520 mg, 3.4 mmol) and potassium carbonate (79 mg, 0.57 mmol) in 13.4 mL of anhydrous DMF was added (*R*)-glycidol (0.23 mL, 3.4 mmol) (60% in mineral oil, 15 mg, 0.38 mmol) and the solution heated to 90 $^\circ\text{C}$ overnight. The next morning, the solution was cooled, all solvent removed, and the product dry loaded onto silica gel after redissolving in methanol. The product was purified via flash chromatography eluting with 10:1 EtOAc:MeOH to afford compound

3.11 as a white solid (268 mg, 35%). $^1\text{H-NMR}$ (500 MHz, d_6 -acetone) δ (ppm) 8.70 (s, 1H), 8.47 (s, 1H), 4.60 (dd, $J = 14.2, 3.6$ Hz, 1H), 4.37 (dd, $J = 14.2, 7.9$ Hz, 1H), 4.13 (m, 1H), 3.61 (m, 2H).

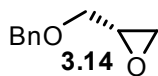


Compound 3.12. A suspension of compound **3.11** (253 mg, 1.1 mmol) in 1 M HCl (11 mL) was heated to 85 °C for three hours. After cooling to room temperature, the pH was adjusted to approximately 9 using concentrated aqueous ammonia and then all solvent was removed. The product was redissolved in methanol, dry loaded onto silica, and purified via flash chromatography over silica gel eluting with 8:1 MeCN:H₂O to afford compound **3.12** as a white solid. The $^1\text{H-NMR}$ was inconclusive and the product was taken crude to the next step.

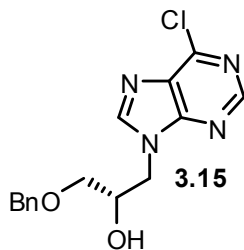


Attempted synthesis of Compound 3.8b. A solution of crude compound **3.12** (~1.1 mmol) and 4,4'-dimethoxytrityl chloride (450 mg, 1.3 mmol) in pyridine (4.4 mL) was allowed to stir at room temperature for five hours. TLC indicated no conversion to

product. An attempt was made to recover compound **3.12** by flash chromatography eluting with 1:1 EtOAc:MeOH but failed as nothing was eluted from the column.

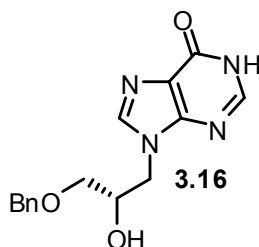


Compound 3.14. To a solution of NaH (725 mg, 18.1 mmol, 60% in mineral oil) in anhydrous THF (10.5 mL) under nitrogen at 0 °C was added (*R*)-glycidol (1.00 mL, 15.1 mmol) and allowed to warm up slowly to room temperature and stir for one hour. The solution was then cooled back to 0 °C, benzyl bromide (2.15 mL, 18.1 mmol) added dropwise, and the solution allowed to warm again to room temperature and stir overnight. The next morning, the solution was quenched with 15 mL of saturated aqueous NH₄Cl, extracted once into ethyl acetate, dried over MgSO₄, and concentrated by rotary evaporation. The crude product was purified via column chromatography using 8:1 Hexanes:EtOAc to afford compound **3.14** as a light yellow oil (1.55 g, 63%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 7.38-7.28 (m, 5H), 4.60 (m, 2H), 3.78 (dd, J = 11.4, 3.1 Hz, 1H), 3.47 (dd, J = 11.5, 5.8 Hz, 1H), 3.20 (m, 1H), 2.81 (t, J = 4.6 Hz, 1H), 2.63 (dd, J = 5.0, 2.7 Hz, 1H). ¹H NMR was consistent with published data.³³

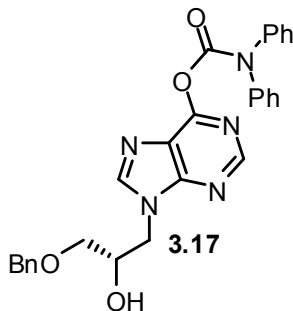


Compound 3.15. A solution of 6-chloropurine (362 mg, 2.3 mmol) and K₂CO₃ (49 mg, 0.35 mmol) in anhydrous DMF (4 mL) was allowed to stir under argon at room temperature for one hour. A solution of compound **3.14** (385 mg, 2.3 mmol) in

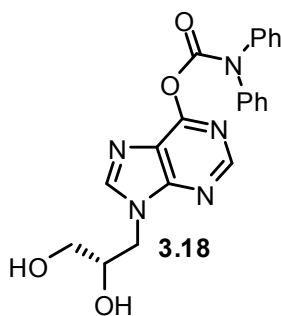
anhydrous DMF (3.5 mL) was added to the above solution and heated to 90 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via flash chromatography over silica gel starting with 1:2 Hexanes:EtOAc, then eluting with 100% EtOAc to afford compound **3.15** as a yellow oil (308 mg, 41%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 8.70 (s, 1H), 8.18 (s, 1H), 7.37-7.26 (m, 5H), 4.53 (d, J = 3.9 Hz, 2H), 4.49 (dd, J = 14.2, 3.2 Hz, 1H), 4.32 (dd, J = 10.2, 7.3 Hz, 1H), 4.25 (m, 1H), 3.58-3.52 (m, 2H), 3.48 (dd, J = 9.7, 5.6 Hz, 1H).



Compound 3.16. A suspension of compound **3.15** (308 mg, 0.97 mmol) in 1 M HCl (9.7 mL) was heated to 85 °C for three hours. After cooling to room temperature, the pH was adjusted to approximately 9 using concentrated aqueous ammonia and then all solvent was removed. The product was redissolved in methanol, filtered through celite, dry loaded onto silica, and purified via flash chromatography over silica gel loading with 100% EtOAc, then eluting with 20:3 EtOAc:MeOH to afford compound **3.16** as a white solid (230 mg, 79%). ¹H-NMR (360 MHz, MeOD) δ (ppm) 8.01 (m, 2H), 7.36-7.24 (m, 5H), 4.52 (s, 2H), 4.43 (dd, J = 14.0, 3.7 Hz, 1H), 4.25 (dd, J = 14.0, 7.7 Hz, 1H), 4.14 (m, 1H), 3.54-3.46 (m, 2H).

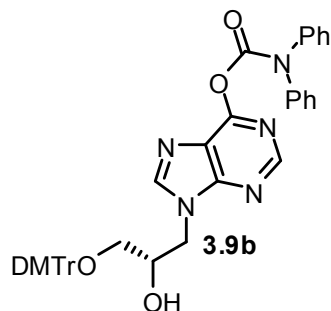


Compound 3.17. To a solution of compound **3.16** (50 mg, 0.17 mmol) and diphenylcarbamoyl chloride (46 mg, 0.20 mmol) in anhydrous pyridine (1.3 mL) was added *N,N*-diisopropylethylamine (35 μ L, 0.20 mmol) and allowed to stir under argon at room temperature for one hour. The resulting dark red solution was diluted with methylene chloride, washed with saturated aqueous NaHCO_3 , dried over Na_2SO_4 , and finally concentrated. The crude product was purified via flash chromatography over silica gel starting with 1:1 Hexanes:EtOAc, the eluting with 100% EtOAc to afford compound **3.17** as a white foam (53 mg, 64%). $^1\text{H-NMR}$ (360 MHz, MeOD) δ (ppm) 8.52 (s, 1H), 7.93 (s, 1H), 7.55-7.10 (m, 15H), 4.48 (s, 2H), 4.31 (dd, $J = 13.9, 3.5$ Hz, 1H), 4.16 (dd, $J = 14.0, 7.3$ Hz, 1H), 4.05 (br, 1H), 3.47-3.34 (m, 2H).

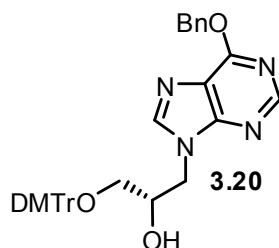


Compound 3.18. Compound **3.17** (53 mg, 0.12 mmol) and Pd/C (26 mg, 10% on carbon) were suspended in EtOAc (3.2 mL) and the solution was purged with argon, then hydrogen, and allowed to stir under a hydrogen atmosphere overnight. The mixture was

filtered through celite and washed with ethyl acetate to afford compound **3.18** as a colorless oil (43 mg, 100%). The product was used in the next step without further purification. $^1\text{H-NMR}$ (360 MHz, MeOD) δ (ppm) 8.53 (s, 1H), 7.97 (s, 1H), 7.54-7.14 (m, 10H), 4.34 (dd, $J = 14.1, 3.5$ Hz, 1H), 4.10 (m, 1H), 3.90 (br, 1H), 3.49 (d, $J = 5.3$, 2H).

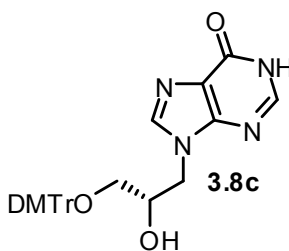


Attempted synthesis of Compound 3.9b. An argon purged solution of compound **3.18** (43 mg, 0.11 mmol) and 4,4'-dimethoxytrityl chloride (43 mg, 0.13 mmol) in anhydrous pyridine (0.50 mL) was allowed to stir at room temperature overnight. TLC indicated no conversion to product and no attempt was made at recovering the starting material.

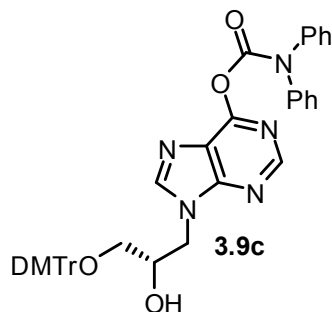


Compound 3.20. To a suspension of 6-benzyloxypurine (514 mg, 2.3 mmol) in anhydrous DMF (4.5 mL) under argon was added NaH (18 mg, 0.45 mmol, 60% in mineral oil) and the solution allowed to stir under argon for one hour. A solution of compound **2.11** (815 mg, 2.2 mmol) in DMF (4.5 mL) was added to the first solution and

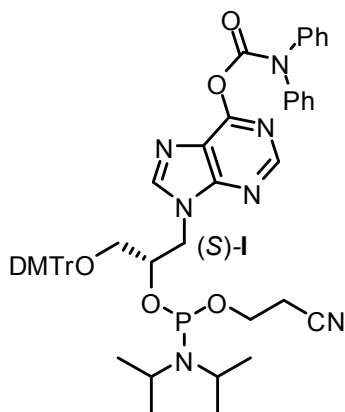
then heated to 100 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via column chromatography starting with 1:1:0.01 Hexanes:EtOAc:Et₃N, then with 1:2:0.01 Hexanes:EtOAc:Et₃N, and finally eluting with 100:1 EtOAc:Et₃N to afford compound **3.20** as a white foam (575 mg, 44%).
¹H-NMR (500 MHz, CDCl₃) δ (ppm) 8.44 (s, 1H), 7.88 (s, 1H), 7.54 (d, J = 7.0 Hz, 2H), 7.44 (d, J = 7.3 Hz, 2H), 7.34-7.20 (m, 10H), 6.83 (m, 4H), 5.57 (m, 2H), 4.44 (d, J = 11.8 Hz, 1H), 4.29-4.21 (m, 2H), 3.79 (s, 6H), 3.16 (dd, J = 9.7, 4.6 Hz, 1H), 3.12 (dd, J = 9.6, 5.1 Hz, 1H).



Compound 3.8c. Compound **3.20** (575 mg, 0.95 mmol) and Pd/C (287 mg, 10% on Carbon) were suspended in EtOAc (3 mL) and the solution was purged with argon, then hydrogen, and allowed to stir under a hydrogen atmosphere for one hour after which TLC showed completion of the reaction. The mixture was filtered through celite and washed with EtOAc:MeOH:Et₃N 40:3:0.01 to afford crude compound **3.8c** as a light brown solid (395 mg, 81%). The crude product was used in the next step further without purification.
¹H-NMR (500 MHz, CDCl₃) δ (ppm) 8.05 (br, 1H), 7.76 (br, 1H), 7.42 (d, J = 7.5 Hz, 2H), 7.35-7.17 (m, 7H), 6.83 (d, J = 8.8 Hz, 4H), 4.40 (m, 1H), 4.22 (br, 2H), 3.78 (s, 6H), 3.25 (m, 1H), 3.18 (m, 1H).



Compound 3.9c. (MKS347) To a solution of compound **3.8c** (350 mg, 0.68 mmol) and diphenylcarbonyl chloride (190 mg, 0.82 mmol) in anhydrous pyridine (5.5 mL) was added *N,N*-diisopropylethylamine (145 μ L, 0.83 mmol) and the solution allowed to stir at room temperature for one hour. The solution was diluted with methylene chloride and washed once with saturated aqueous NaHCO_3 , dried over Na_2SO_4 , concentrated, and coevaporated with toluene. The product was purified via column chromatography starting with 1:2:0.01 Hexanes:EtOAc: Et_3N , and finally eluting with 100:1 EtOAc: Et_3N to afford compound **3.9c** as a white foam (250 mg, 52%). $^1\text{H-NMR}$ (360 MHz, CDCl_3) δ (ppm) 8.06 (s, 1H), 7.63 (s, 1H), 7.50-7.07 (m, 19H), 6.84 (d, $J = 8.9$ Hz, 4H), 4.30 (br, 2H), 4.07 (br, 1H), 3.80 (s, 6H), 3.24 (br, 2H).



Compound (S)-I. (MKS349) To a nitrogen purged solution of compound **3.9c** (250 mg, 0.35 mmol) and *N,N*-diisopropylethylamine (370 μ L, 2.1 mmol) in methylene chloride (5.9 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (120 μ L, 0.53 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The solution was diluted with methylene chloride and washed once with saturated aqueous NaHCO_3 , dried over Na_2SO_4 , and finally concentrated by rotoevaporation. The crude product was purified by column chromatography starting with 1:1:0.01 Hexanes:EtOAc: Et_3N then eluting with 1:2:0.01 Hexanes:EtOAc: Et_3N to afford compound (S)-I as a white foam (300 mg, 94%). ^{31}P -NMR (162 MHz, CDCl_3) δ (ppm) 151.1, 150.6, 150.2, 150.0.

Thermal stability of nucleic acids monitored using UV spectroscopy: Samples of oligonucleotides were prepared at the stated concentrations and buffer concentrations. Experiments involving RNA were performed using RNase free buffers and water to ensure the integrity of the RNA. Melting studies were carried out in 1 cm path length quartz cuvettes with 200 μ L of sample covered by 125 μ L of mineral oil, or 150 μ L of

sample covered by 150 μL of mineral oil on a Beckmann DU800 spectrophotometer equipped with a thermoprogrammer. Melting curves were generally measured at 260 nm with a heating rate of 1 $^{\circ}\text{C}/\text{min}$. Melting temperatures (T_m) were calculated from either the first derivative or nonlinear fit to the heating curves. Experiments were performed at least in duplicate and the average value taken.

Circular dichroism of nucleic acids: Samples of oligonucleotides were prepared at the stated concentrations and buffer concentrations. Experiments involving RNA were performed using RNase free buffers and water to ensure the integrity of the RNA. Measurements were performed with a Jasco J-810 spectrometer or Aviv 62A DS spectrometer in 1 mm path length quartz cuvettes. Each spectrum was the result of at least five measurements, averaged in real time.

Chapter 3.7. References

- (1) Zhang, L.; Peritz, A. E.; Carroll, P. J.; Meggers, E. *Synthesis* **2006**, 645-653.
- (2) Zhang, L.; Meggers, E. *J. Am. Chem. Soc.* **2005**, 127, 74-75.
- (3) Nielson, P.; Dreißø L. H.; Wengel, J. *Bioorg. Med. Chem.* **1995**, 3, 19-28.
- (4) Schneider, K. C.; Benner, S. A. *J. Am. Chem. Soc.* **1990**, 112, 453-455.
- (5) Azymah, M.; Chavis, C.; Lucas, M.; Morvan, F.; Imbach, J.-L. *Nucleosides & Nucleotides* **1992**, 11, 1241-1255.
- (6) Vandendriessche, F.; Augustyns, K.; Van Aerschot, A.; Busson, R.; Hoogmartens, J.; Herdewijn, P. *Tetrahedron* **1993**, 49, 7223-7238.
- (7) Nielsen, P.; Kirpekar, F.; Wengel, J. *Nucleic Acid Res.* **1994**, 22, 703-710.
- (8) Peng, L.; Roth, H.-J. *Helv. Chim. Acta* **1997**, 80, 1494-1512.
- (9) Zhou, D.; Lagoja, I. M.; Rozenski, J.; Busson, R.; Van Aerschot, A.; Herdewijn, P. *ChemBioChem* **2005**, 6, 2298-2304.
- (10) Tarköy, M.; Leumann, C. *Angew. Chem. Int. Ed.* **1993**, 32, 1432-1434.
- (11) Steffens, R.; Leumann, C. J. *J. Am. Chem. Soc.* **1999**, 121, 3249-3255.
- (12) Wengel, J. *Acc. Chem. Res.* **1999**, 32, 301-310.
- (13) Nakano, S.; Fujimoto, M; Hara, H.; Sugimoto, N. *Nucleic Acids Res.* **1999**, 27, 2957-2965.
- (14) Burda, J. V.; Spopner, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B.* **1997**, 101, 9670-9677.

- (15) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341-8355.
- (16) Herskovits, T. T. *Archives of Biochemistry and Biophysics* **1962**, *97*, 474-484.
- (17) Sen, A.; Nielsen, P. E.; *Biophysical Journal* **2006**, *90*, 1329-1337.
- (18) Sen, A.; Nielsen, P. E.; *Nucleic Acids Res.* **2007**, *35*, 3367-3374.
- (19) Schlegel, M. K.; Xie, X.; Zhang, L.; Meggers, E. *Angew. Chem. Int. Ed.* **2009**, *48*, 960-963.
- (20) Johnson, C. J. *Circular Dichroism: Principles and Applications, Second Edition*; Berova, N., Nakanishi, K., Woody, R. W., Eds.; John Wiley & Sons, Inc.: New York, NY, 2000; 703-717.
- (21) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X.-F.; Sheils, C. J.; Paris, P. L.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **1996**, *118*, 8182-8183.
- (22) Guckian, K. M.; Ren, R. X.-F.; Chaudhuri, N. C.; Tahmassebi, D. T.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 2213-2222.
- (23) Kim, T. W.; Kool, E. T. *J. Org. Chem.* **2005**, *70*, 2048-2053.
- (24) Breslauer, K. J.; Remeta, D. P.; Chou, W.-Y.; Ferrante, R.; Curry, J.; Zaunczkowski, D.; Snyder, J. G.; Marky, L. A. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8922-8926.
- (25) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H.; *J. Am. Chem. Soc.* **2006**, *128*, 10258-10267.
- (26) Horowitz, E. D.; Hud, N. V. *J. Am. Chem. Soc.* **2006**, *128*, 15380-15381.
- (27) Chaires, J. B. *Biopolymers* **1997**, *44*, 201-215.

- (28) Leumann, C. J. *Bioorg. Med. Chem.* **2002**, 10, 841-854.
- (29) Wei, C.; Jia, G.; Yuan, J.; Feng, Z.; Li, C. *Biochemistry* **2006**, 45, 6681-6691.
- (30) McLaughlin, L.; Leong, T.; Benseler, F.; Piel, N. *Nucleic Acids Res.* **1988**, 16, 5631-5644.
- (31) Gaffney, B. L.; Marky, L. A.; Jones, R. A. *Tetrahedron* **1984**, 40, 3-13.
- (32) Martin, F. H.; Castro, M. M. *Nucleic Acids Res.* **1985**, 13, 8927-8938.
- (33) Walker, L. F.; Bourghida, A.; Connolly, S.; Wills, M. *J. Chem. Soc., Perkin Trans. 1*, **2002**, 965-981.
- (34) Horhota, A. T.; Szostak, J. W.; McLaughlin, L. W. *Org. Lett.* **2006**, 8, 5345-5347.
- (35) Tsai, C.-H.; Chen, J.; Szostak, J. W. *Proc. Natl. Acad. Sci. USA* **2007**, 104, 14598-14603.
- (36) Schlegel, M. K.; Meggers, E. *J. Org. Chem.* **2009**, 74, 4615-4618.

Appendix to Chapter 3

**^1H , ^{13}C , and ^{31}P NMR spectra
IR spectra**

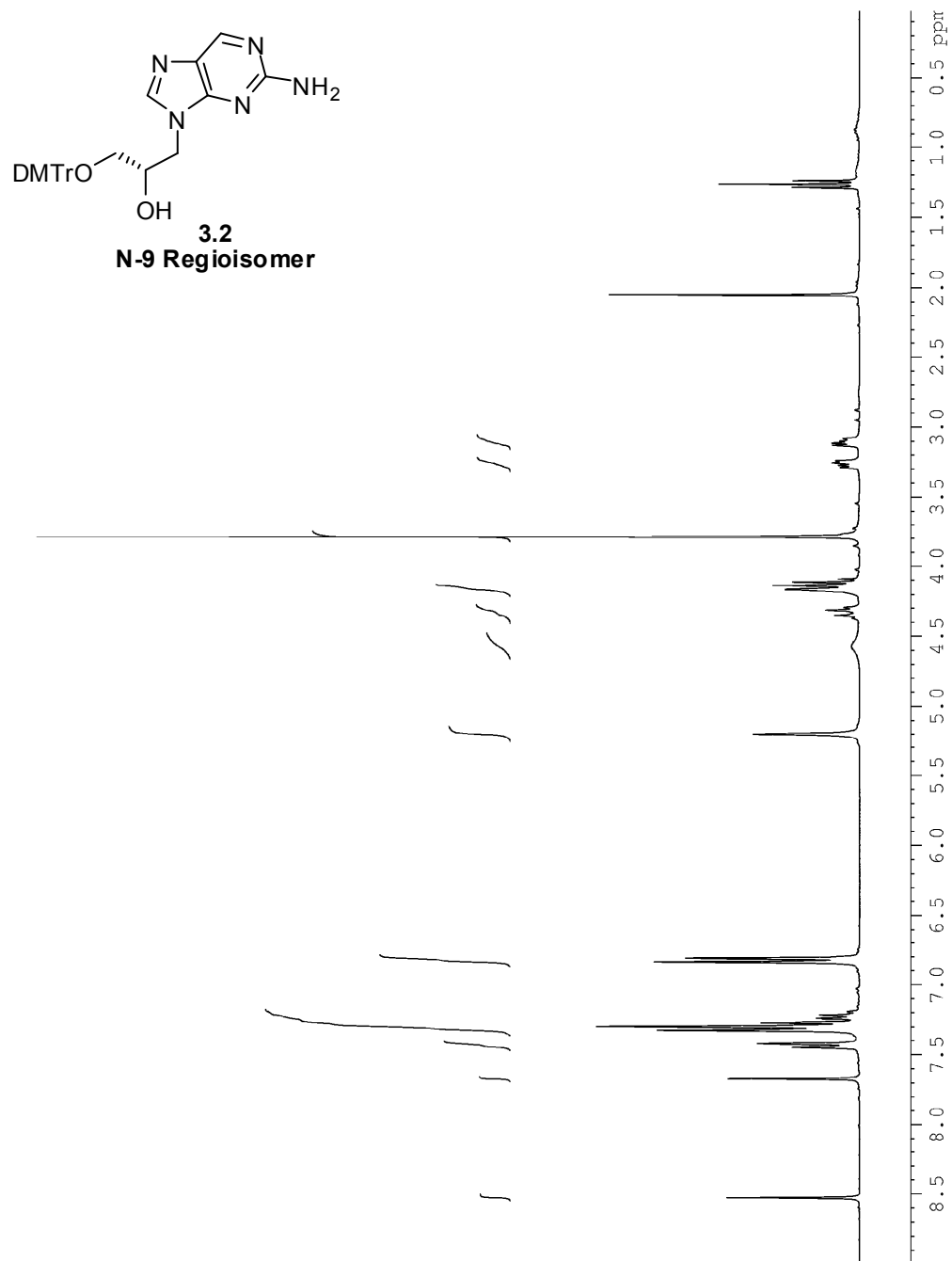


Figure A3.1.1. ¹H NMR spectrum of the N-9 regioisomer of compound 3.2 (300 MHz, CDCl₃).

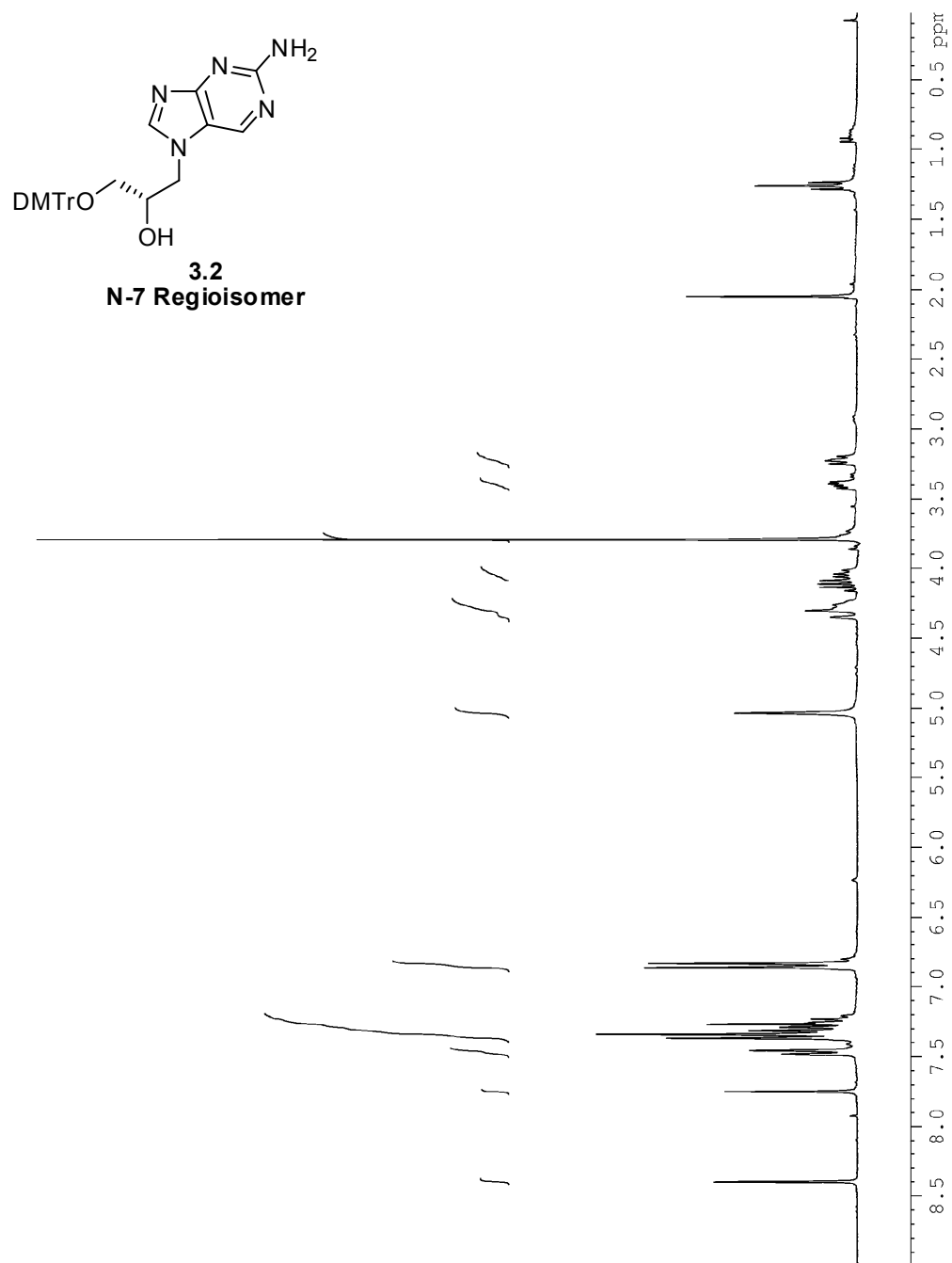


Figure A3.1.2. ^1H NMR spectrum of the N-7 regioisomer of compound **3.2** (300 MHz, CDCl_3).

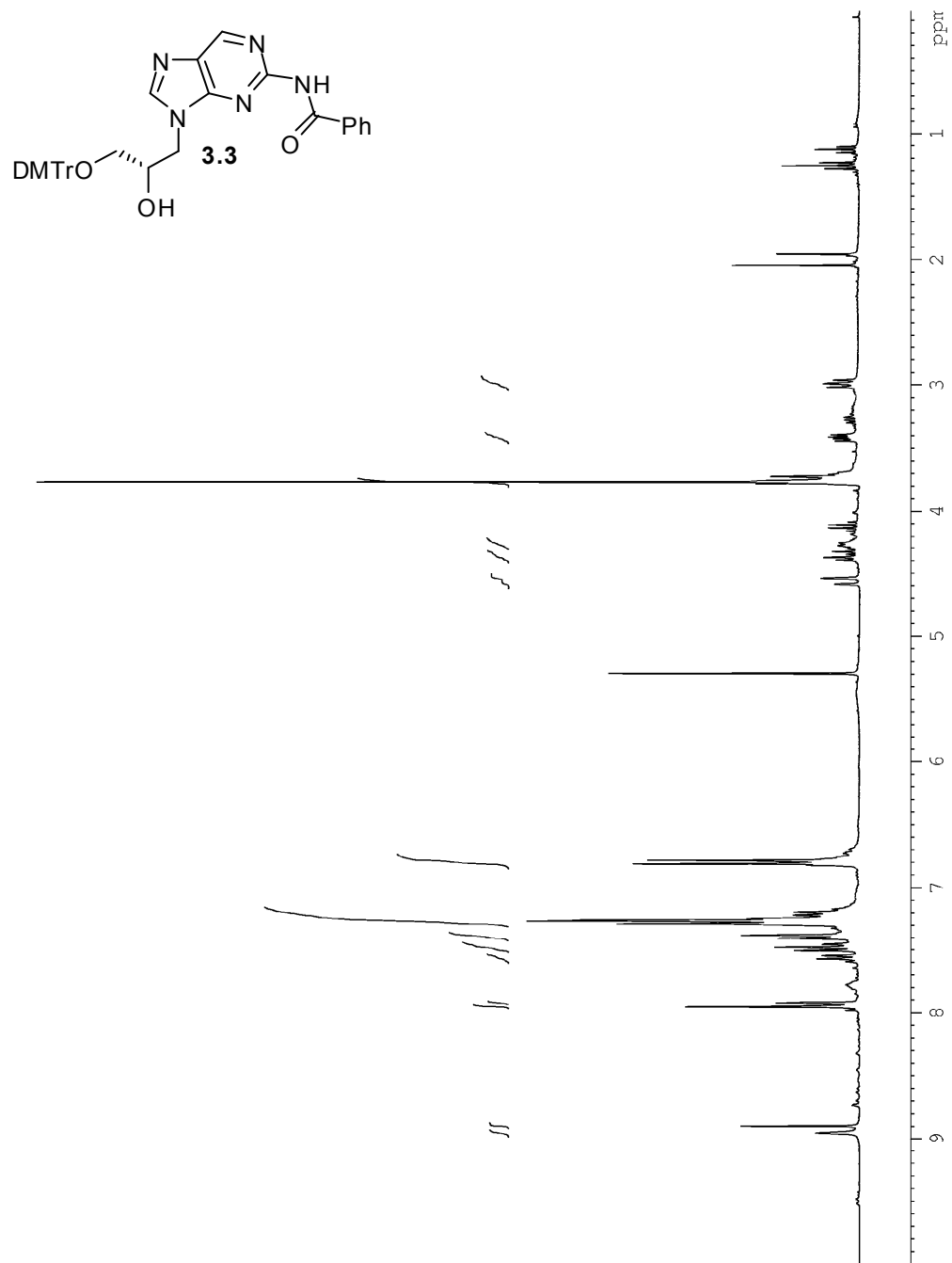


Figure A3.2.1. ¹H NMR spectrum of compound **3.3** (300 MHz, CDCl₃).

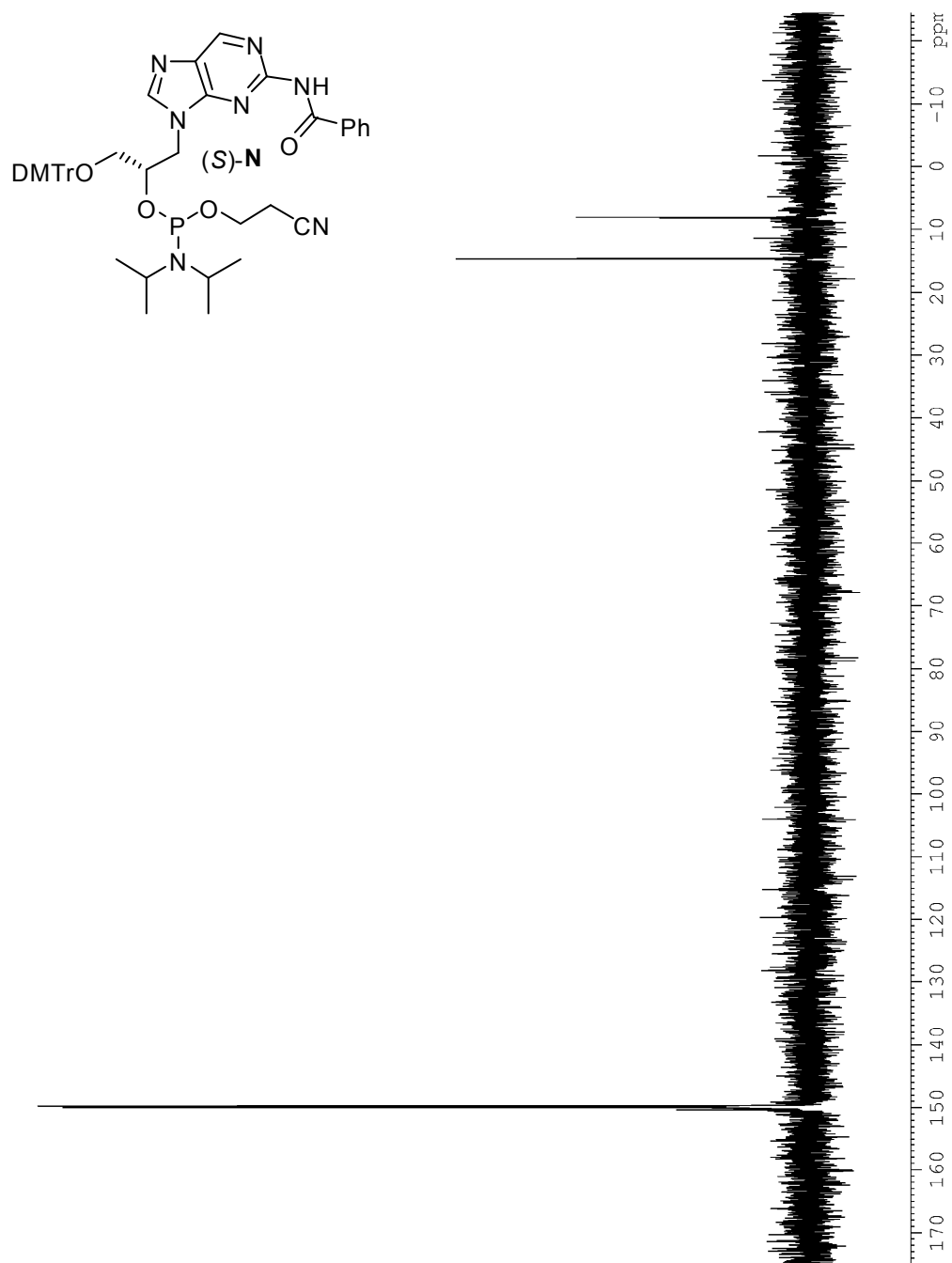


Figure A3.3.1. ¹H NMR spectrum of phosphoramidite (S)-N (121 MHz, CDCl₃).

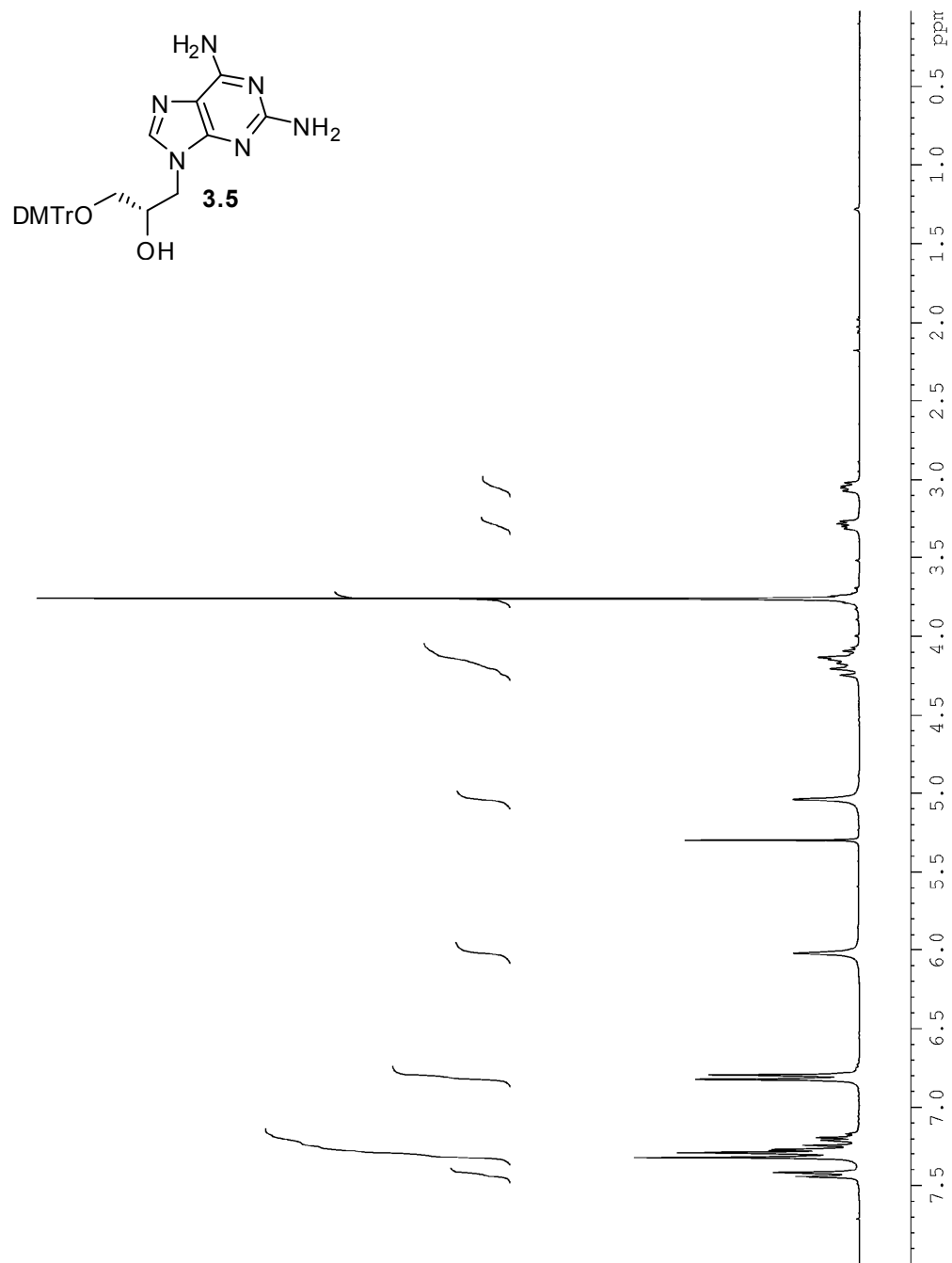


Figure A3.4.1. ¹H NMR spectrum of compound **3.5** (300 MHz, CDCl₃).

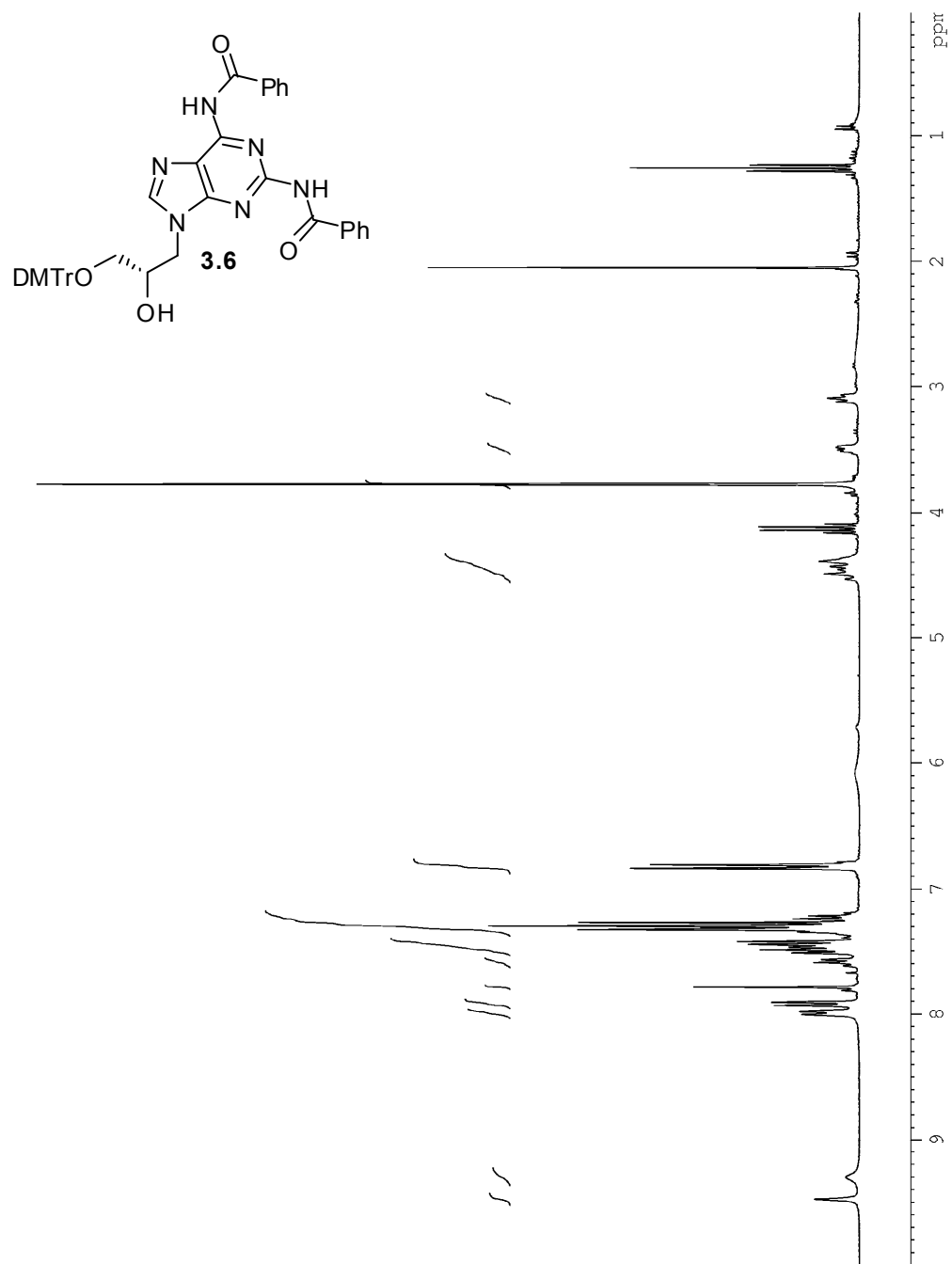


Figure A3.5.1. ^1H NMR spectrum of compound **3.6** (300 MHz, CDCl_3).

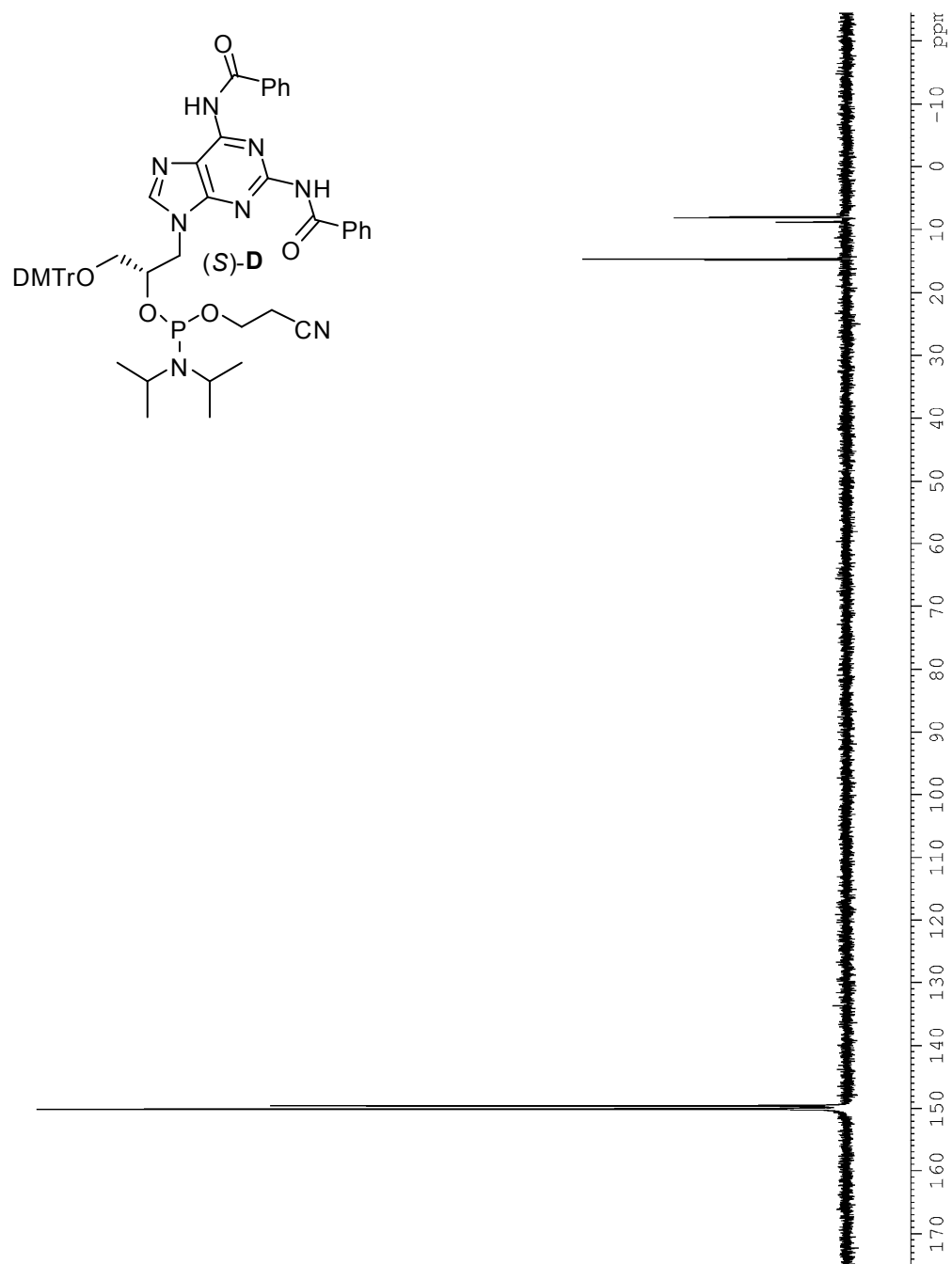


Figure A3.6.1. ^1H NMR spectrum of phosphoramidite (S)-D (121 MHz, CDCl_3).

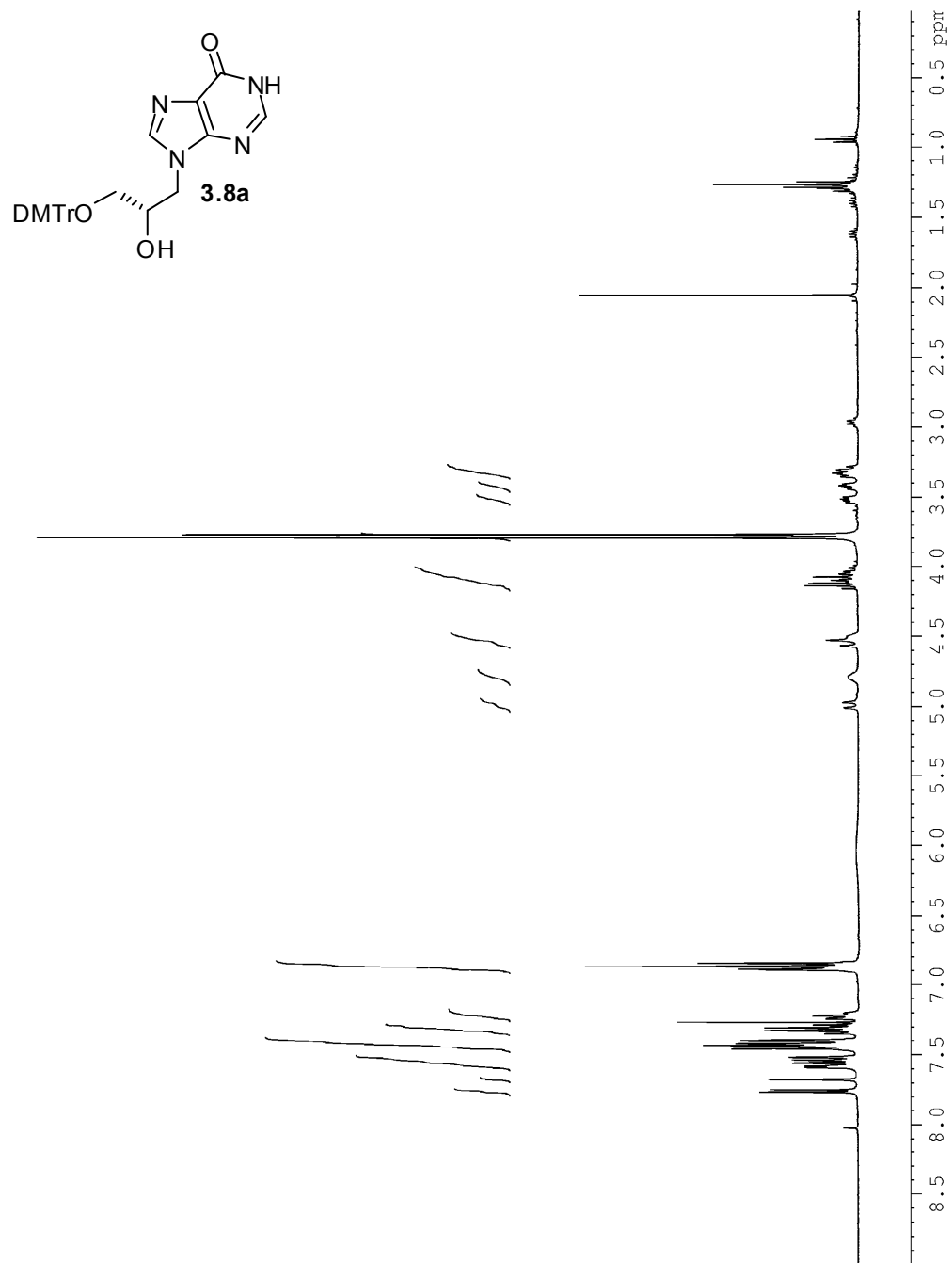


Figure A3.7.1. ^1H NMR spectrum of the two tautomeric forms (?) of compound **3.8a** (360 MHz, CDCl_3).

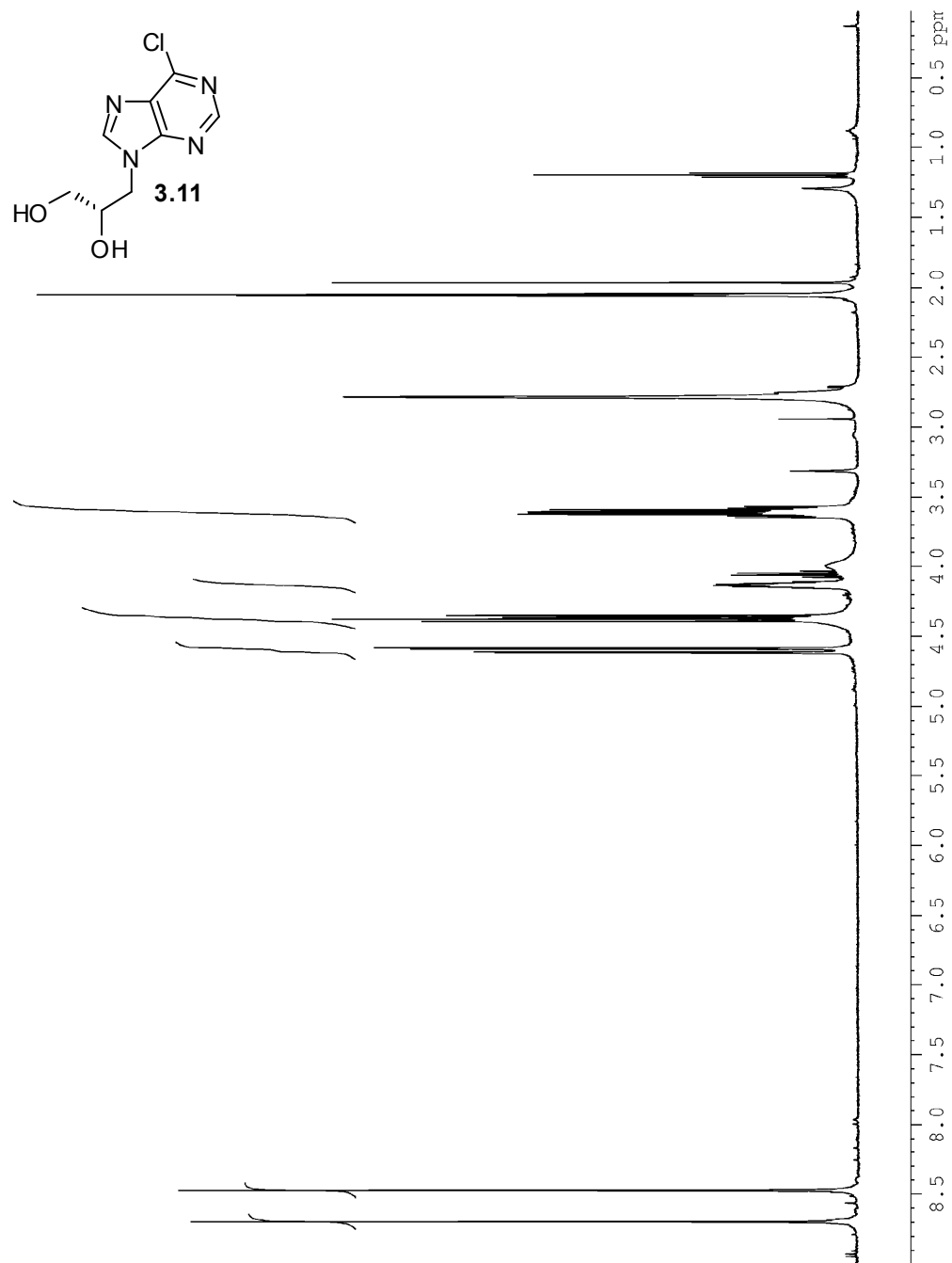


Figure A3.8.1. ^1H NMR spectrum of compound **3.11** (500 MHz, acetone- d_6).

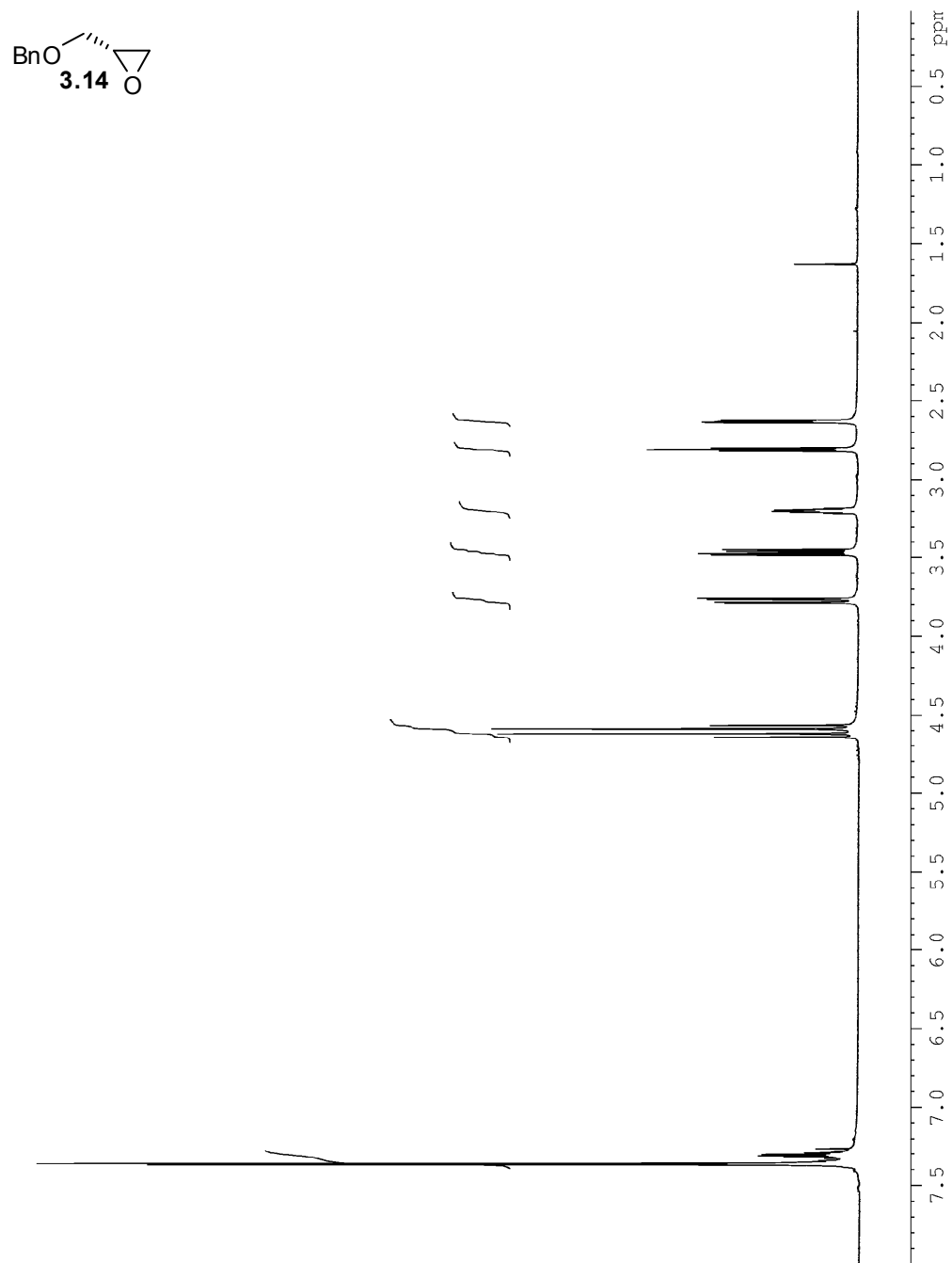


Figure A3.9.1. ^1H NMR spectrum of compound **3.14** (500 MHz, CDCl_3).

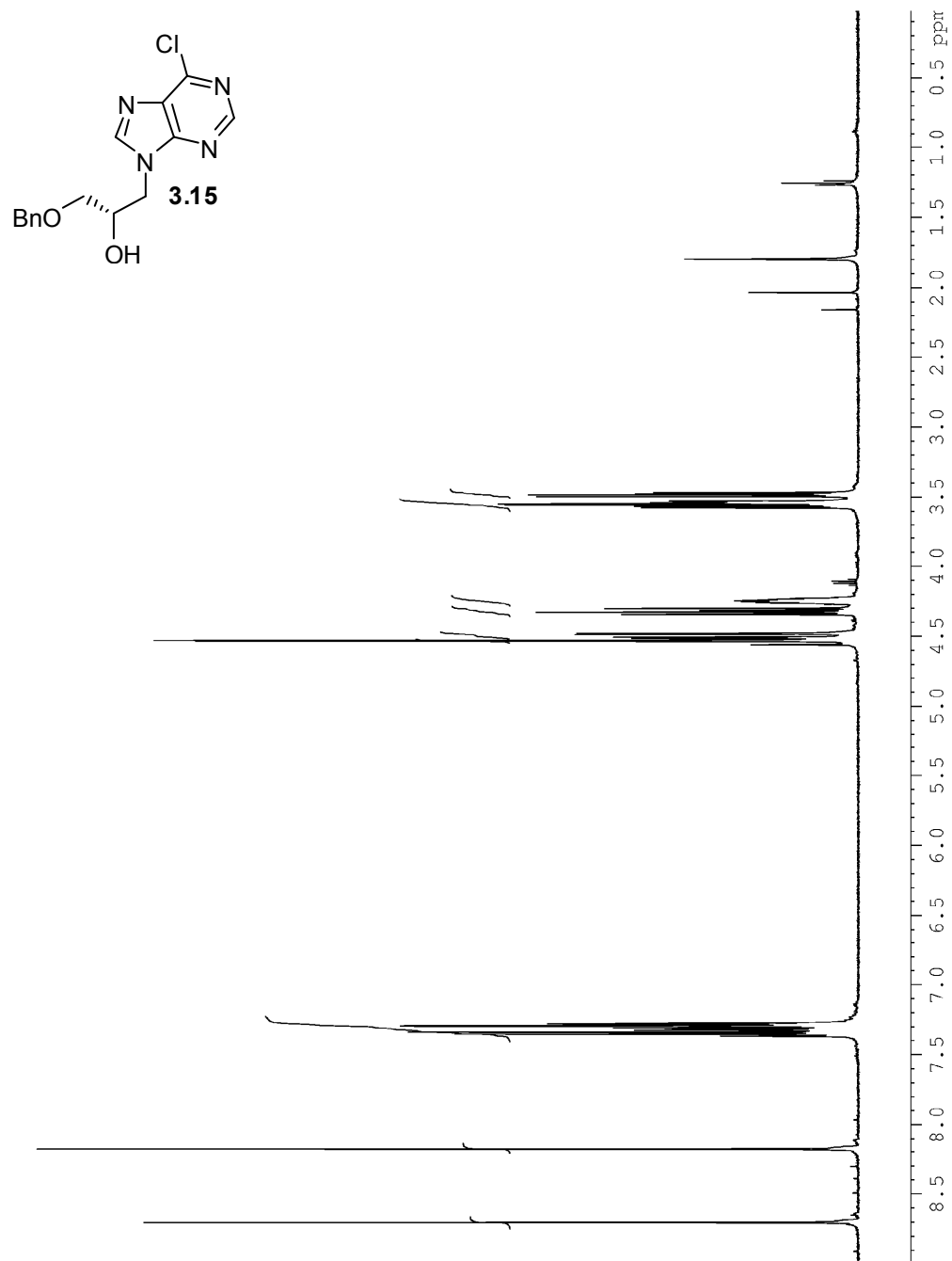


Figure A3.10.1. ^1H NMR spectrum of compound **3.15** (500 MHz, CDCl_3).

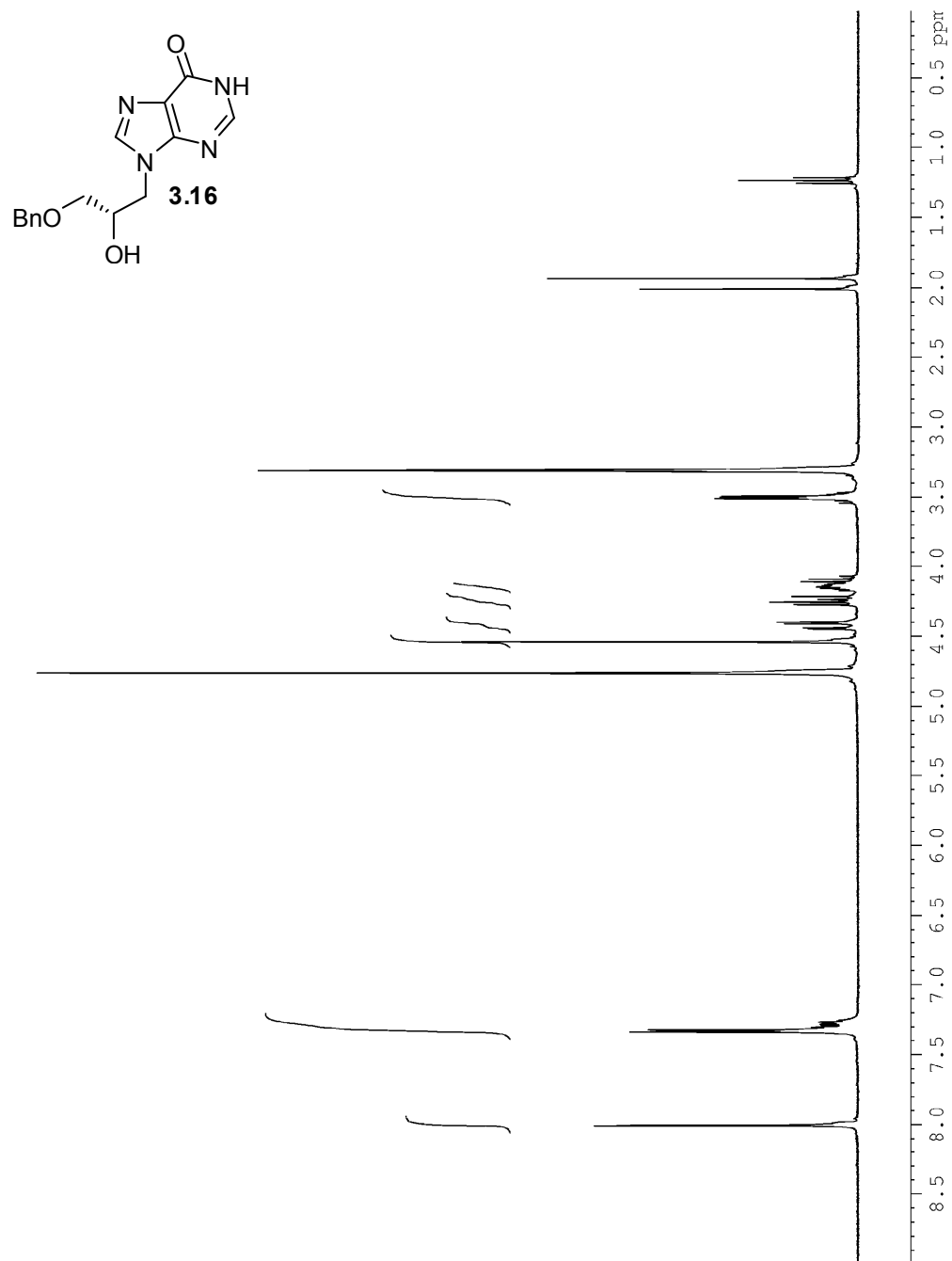


Figure A3.11.1. ^1H NMR spectrum of compound **3.16** (360 MHz, MeOD).

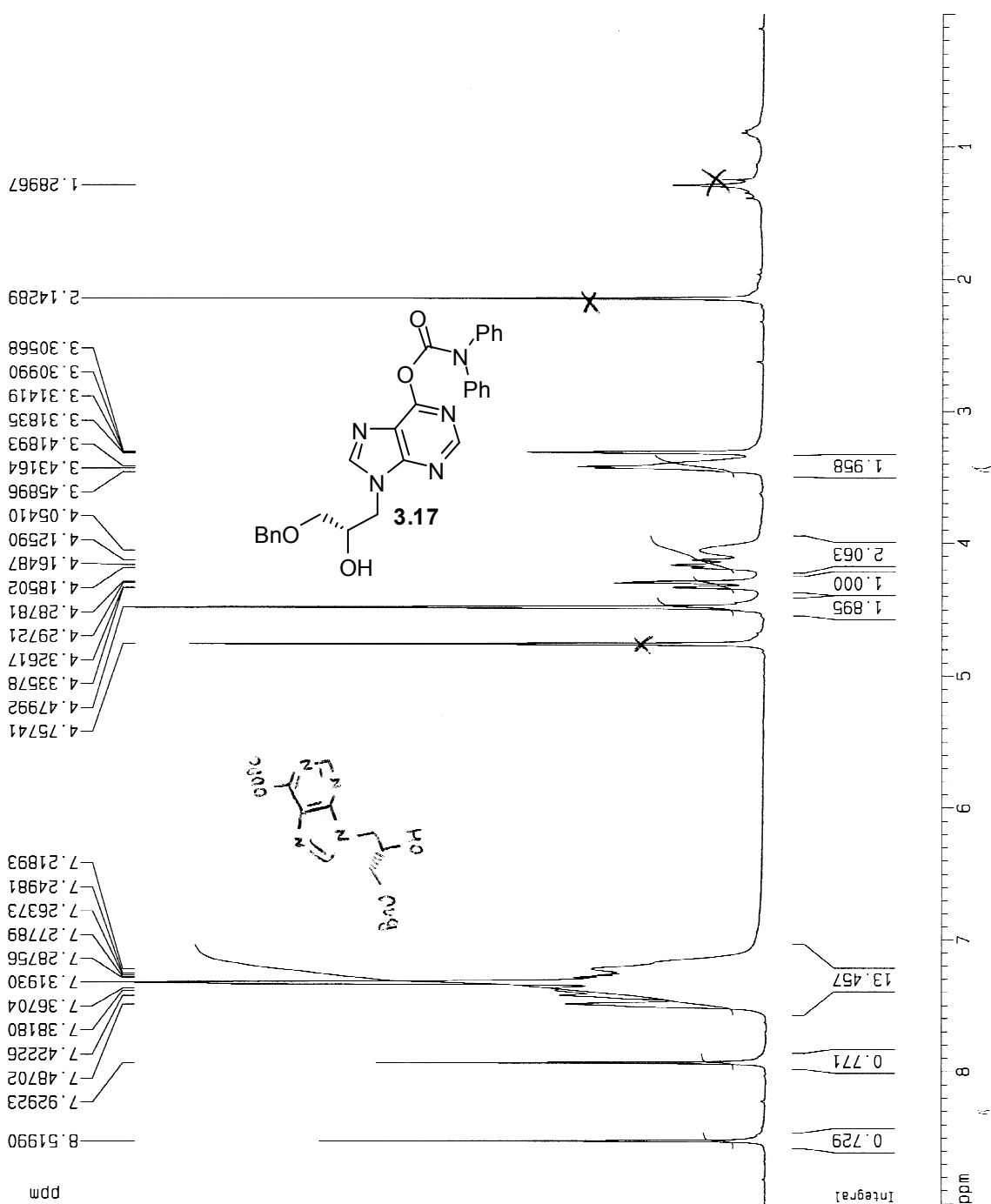


Figure A3.12.1. ^1H NMR spectrum of compound **3.17** (360 MHz, MeOD).

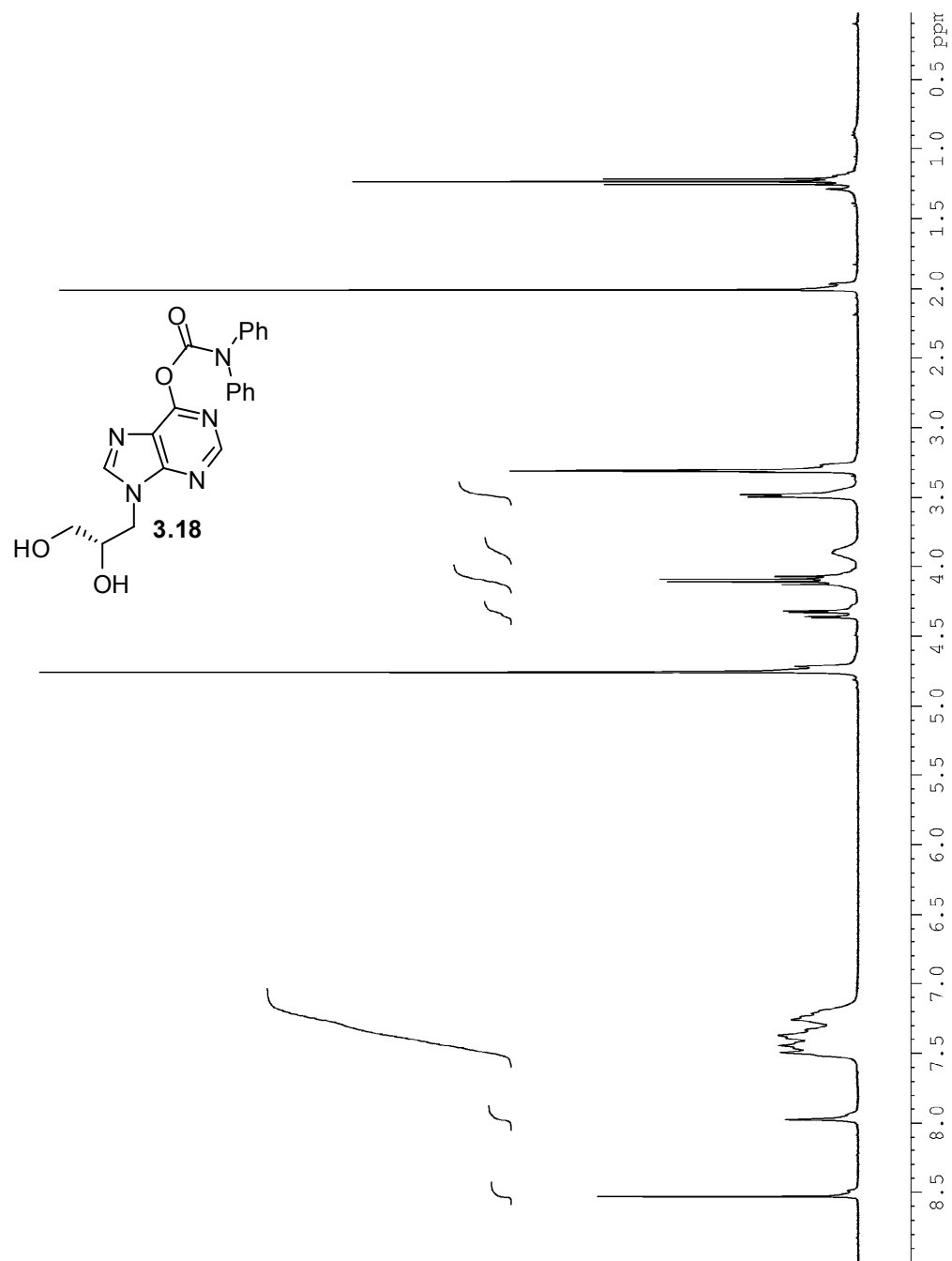


Figure A3.13.1. ^1H NMR spectrum of compound **3.18** (360 MHz, MeOD).

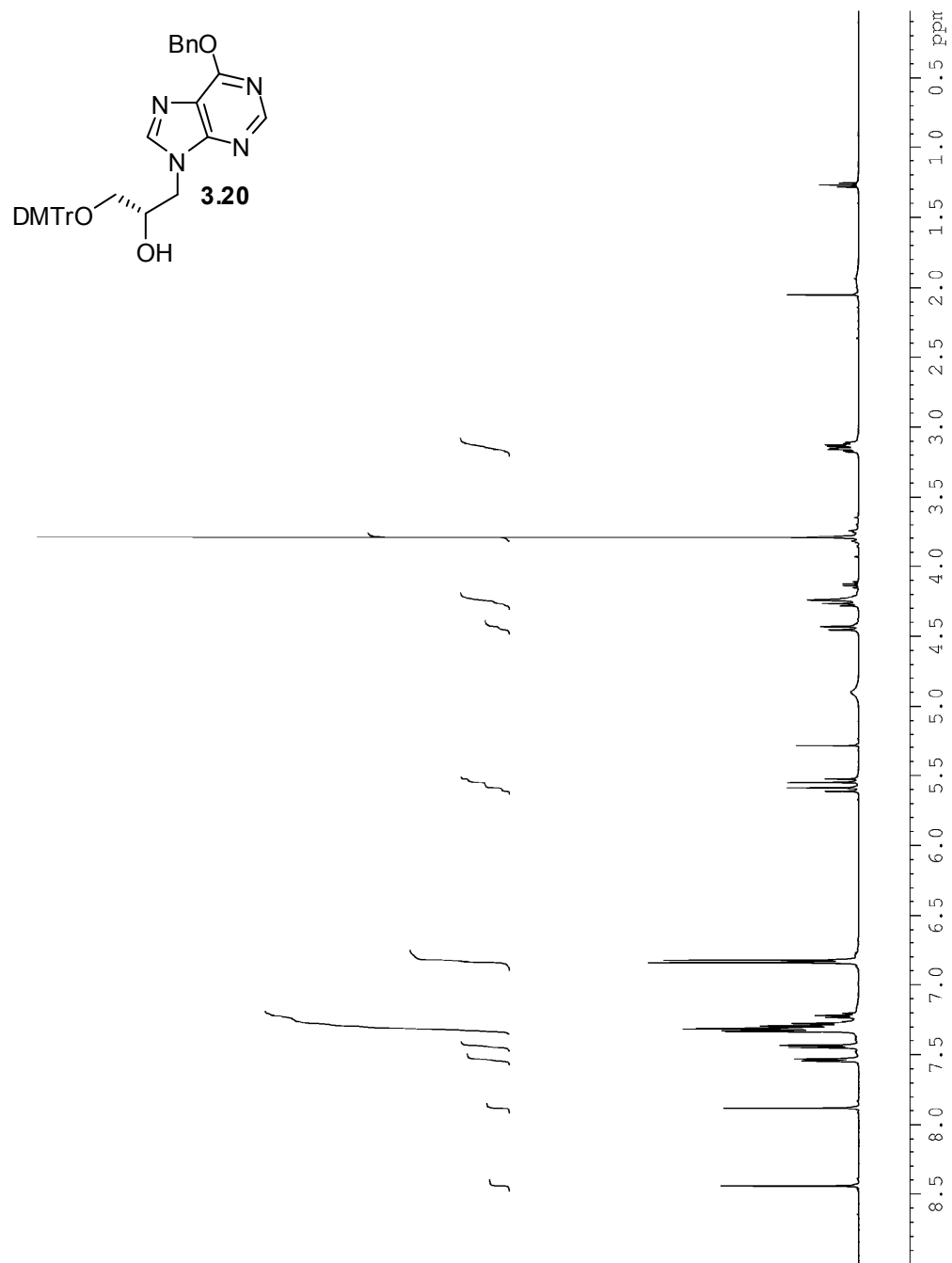


Figure A3.14.1. ¹H NMR spectrum of compound **3.20** (500 MHz, CDCl₃).

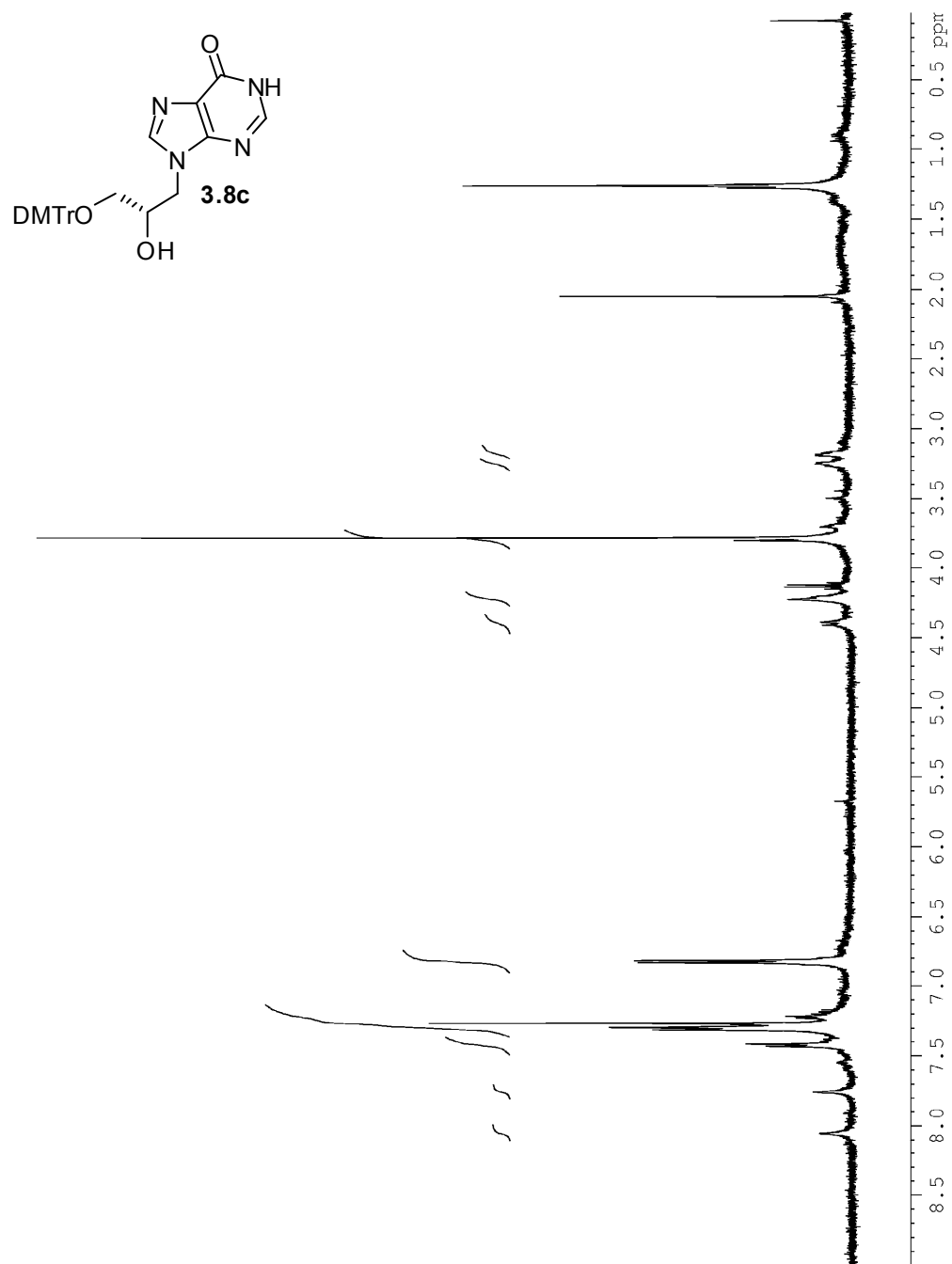


Figure A3.15.1. ^1H NMR spectrum of compound **3.8c** (500 MHz, CDCl_3).

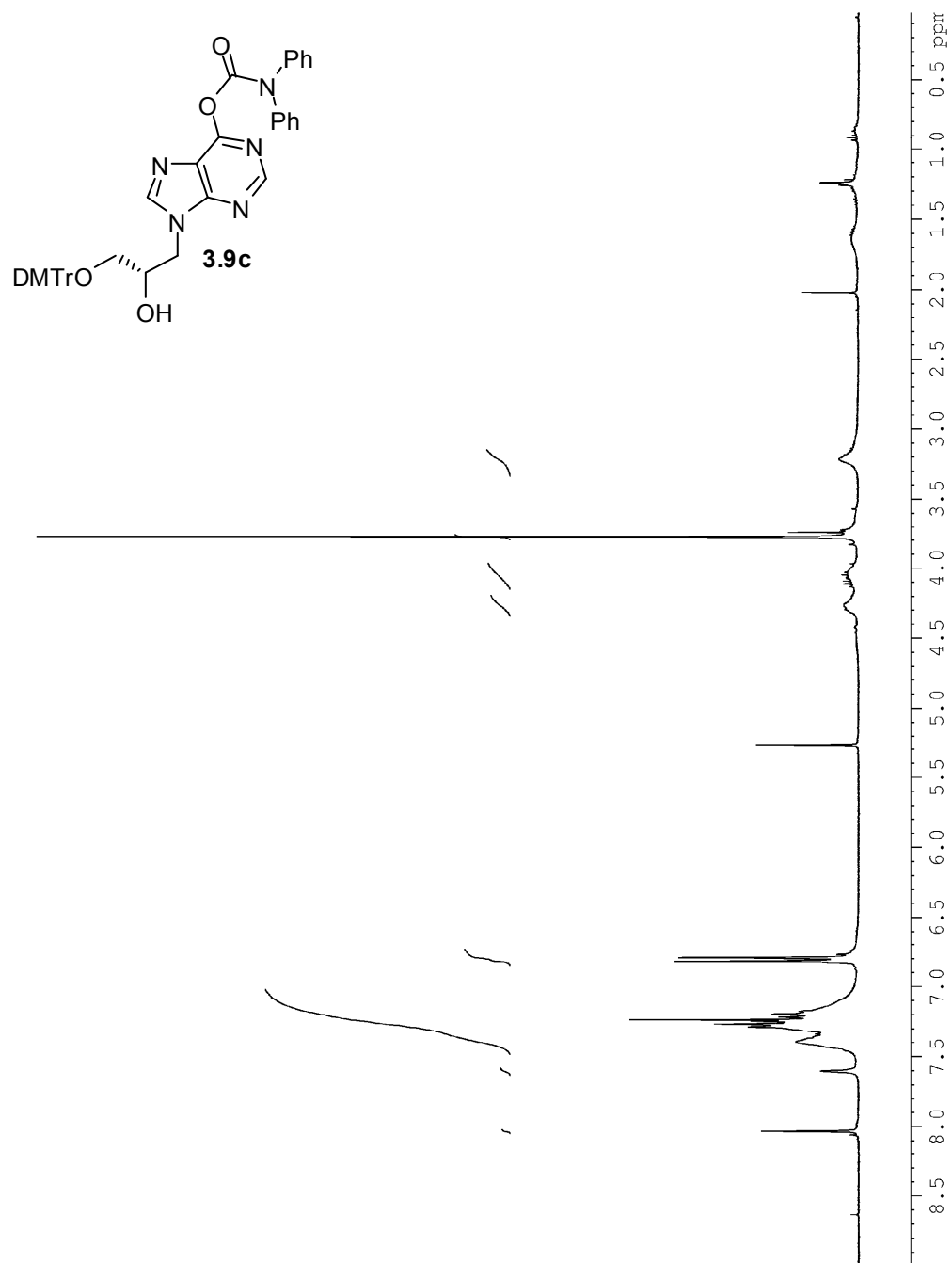


Figure A3.16.1. ^1H NMR spectrum of compound **3.9c** (360 MHz, CDCl_3).

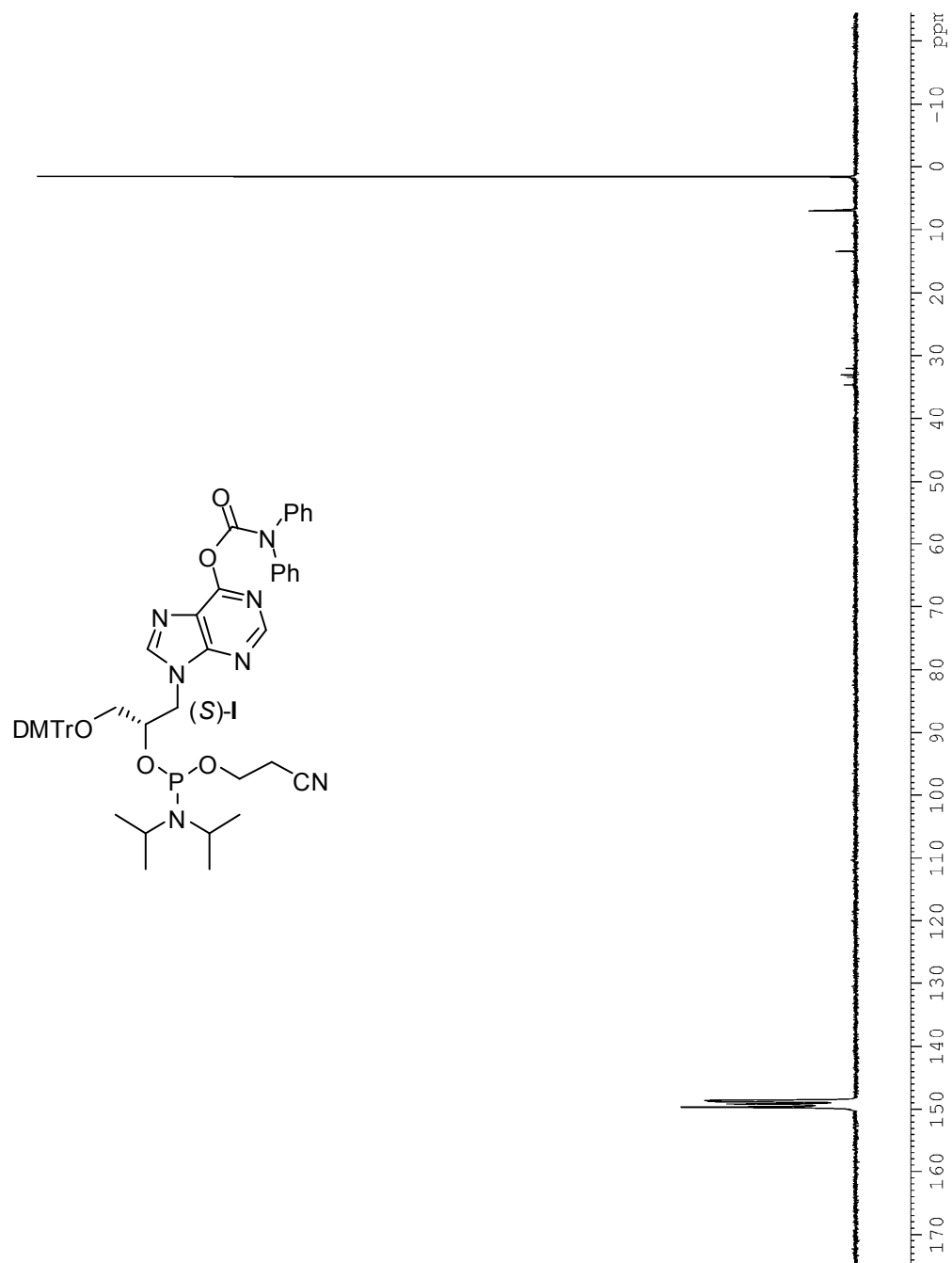


Figure A3.17.1. ^1H NMR spectrum of phosphoramidite (S)-I (162 MHz, CDCl_3).

Chapter 4

Metal-mediated base pairing in GNA

Chapter 4.1. Synthesis of unnatural metal-mediated base pairs

Due to its programmable self assembly, DNA is a promising biomaterial for the construction of nanoscale architectures (structural DNA nanotechnology) and has already been used for the assembly of a variety of periodic arrays in two dimensions and the construction of materials with predictable three-dimensional structures such as polyhedra and catenanes.^{1,2} In this respect, an important current goal is the functionalization of DNA in order to create nanoscale devices with novel properties.³⁻⁷ One recently developed strategy is the incorporation of metal-mediated base pairs into DNA which enables one to place metal ions within the nucleobase π -stacking thus providing an angstrom-scale control of the patterning of metal ions along a helix axis.⁸⁻¹⁰ It can easily be envisioned that a one dimensional array of redox-active metal ions positioned along a DNA double helix by metal-mediated base pairing may lead to nucleic acids with interesting magnetic and electronic properties.¹¹⁻¹³ However, despite such exciting prospects for metal-mediated base pairing, the multistep synthesis of such artificial nucleotides, including the chromatographical separation of anomeric mixtures, is often painstaking and thus makes their future large scale utilization in nanoscale devices questionable.

Based on these synthetic challenges for DNA, we were interested in the incorporation of metal-mediated base pairs into the GNA scaffold since it combines highly duplex stability, high base pairing fidelity, and easy synthetic access to individual nucleotide building blocks.¹⁴⁻¹⁹ Interestingly, it has previously been shown by our group

that a Cu^{+2} mediated, 8-hydroxyquinoline homobase pair containing the GNA backbone is capable of supporting duplex formation of a DNA duplex.²⁰ Unfortunately, from our experience the 8-hydroxyquinoline is oxidatively unstable, rendering it somewhat impractical for its incorporation into functionalized structures. Therefore, we investigated the synthesis, incorporation, and base pairing properties of the two metal-mediated base pairs composed of a hydroxypyridone and pyridylpurine, selective for Cu^{+2} and Ni^{+2} in DNA, respectively (Figure 4.1).^{19,21,22}

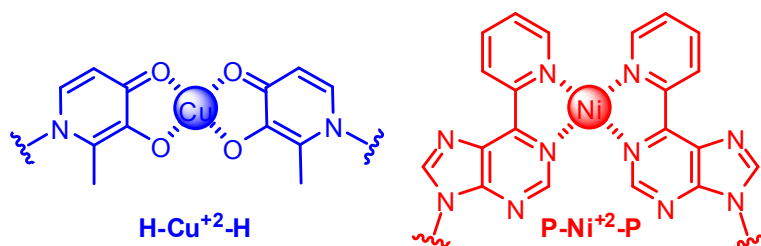
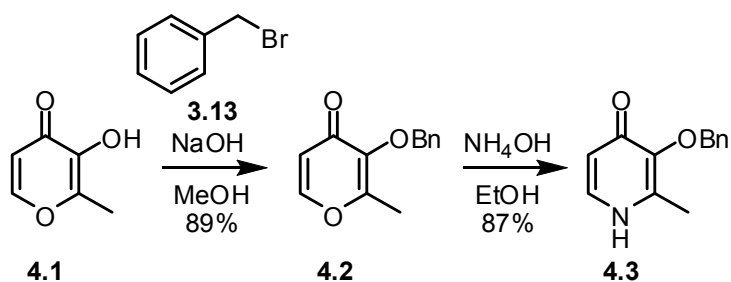


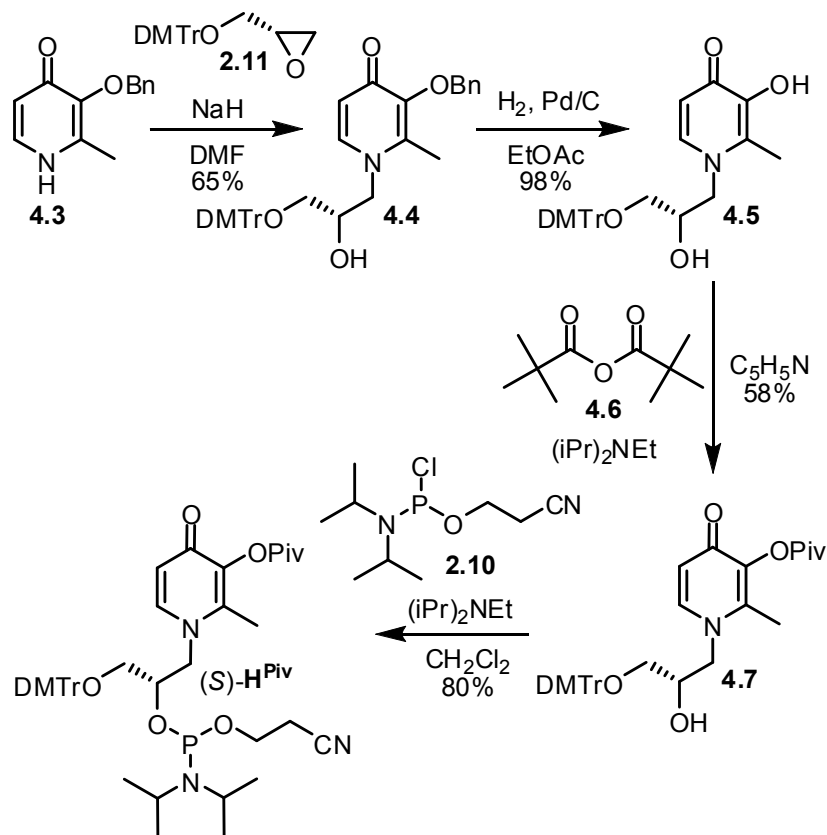
Figure 4.1. Constitution of the hydroxypyridone (H) and pyridylpurine (P) homobase pairs.

In order to obtain the nucleobase for the synthesis of (*S*)-**H** phosphoramidite, 3-hydroxy-2-methyl-4-pyrone (**4.1**) must first be protected by the reaction with benzyl bromide (**3.13**) and aqueous sodium hydroxide to produce compound **4.2** in 89% yield (Scheme 4.1).²³ Compound **4.2** can then be converted to pyridone **4.3** via a Michael addition using 25% aqueous ammonium hydroxide in 87% yield. Compound **4.3** and 0.2 equivalents of sodium hydride were then used in the ring opening of compound **2.11**, affording compound **4.4** in 65% yield (Scheme 4.2).¹⁹ The benzyl group of compound

4.4 was subsequently removed using catalytic hydrogenation to produce compound **4.5** in 98% yield. The initial synthesis involved re-protection of the 3'-hydroxyl group of compound **4.5** using pivalic anhydride (**4.6**) and *N,N*-diisopropylethylamine to produce compound **4.7** in 58% yield. Finally, compound **4.7** was converted to (*S*)-**H^{Piv}** using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess *N,N*-diisopropylethylamine in 80% yield. Unfortunately, initial attempts at the incorporation of (*S*)-**H^{Piv}** into GNA oligonucleotides failed, presumably due to its instability towards the reagents used during oligonucleotide synthesis.



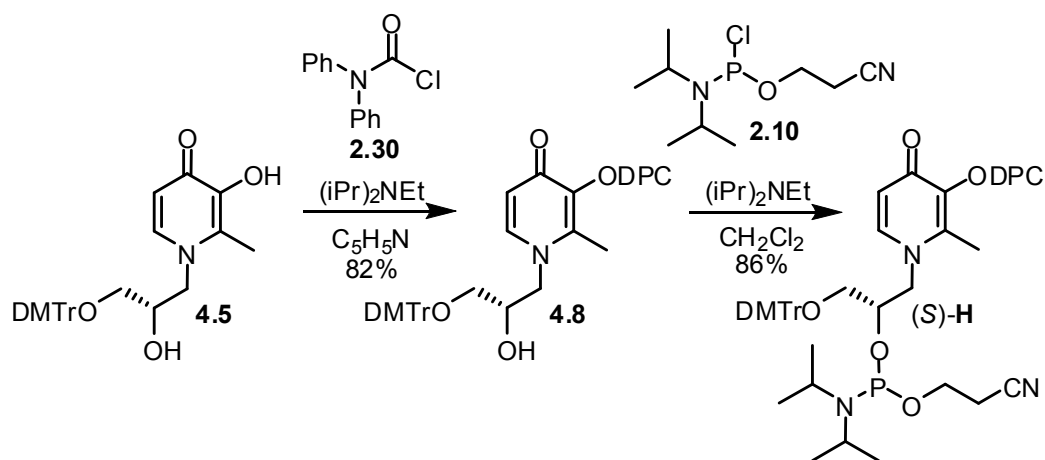
Scheme 4.1. Synthesis of 3-benzyloxy-2-methyl-4-pyridone (**4.3**).



Scheme 4.2. Synthesis of (S)-H^{Piv} phosphoramidite.

A subsequent attempt at protection of the 3'-hydroxyl group of compound **4.5** involved the use of diphenylcarbamoyl chloride. We envisioned that this protection group would provide more stability to the final phosphoramidite, especially towards the reagents during oligonucleotide synthesis, based on its successful use in protecting guanosine nucleosides.^{24,25} Accordingly, compound **4.5** was reacted with diphenylcarbamoyl chloride (**2.30**) and *N,N*-diisopropylethylamine to afford compound **4.8** in 82% yield (Scheme 4.3).¹⁹ Again, conversion to the phosphoramidite (S)-H proceeds using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess

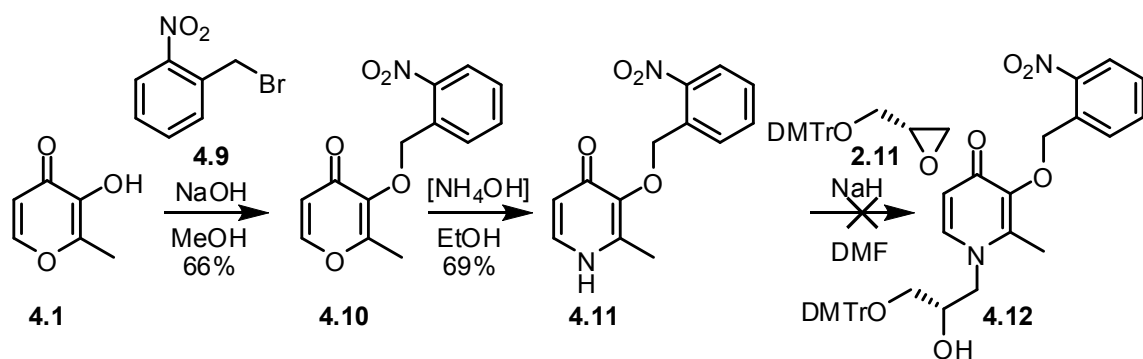
N,N-diisopropylethylamine in 86% yield. Overall, the synthesis of (*S*)-**H** phosphoramidite proceeds in five steps with a yield of 34% compared to the corresponding 2'-deoxynucleotide in which a nine step synthesis with an overall yield of 8% was reported.^{21,26} This protection scheme was shown to be compatible with the conditions during oligonucleotide synthesis and was used for the successful incorporation of (*S*)-**H** into GNA oligonucleotides.



Scheme 4.3. Synthesis of (*S*)-**H** phosphoramidite.

Furthermore, we were interested in developing a derivative of (*S*)-**H** with yet another protection scheme that is stable to both the conditions for oligonucleotide synthesis and the deprotection conditions after synthesis. In this way, it would be possible to obtain a purified GNA oligonucleotide with the hydroxypyridone nucleobase in a protected form. When desired, this oligonucleotide could then be deprotected and release the metal chelating functionality. Therefore, we attempted to synthesize a derivative of (*S*)-**H** in which the hydroxyl group of the hydroxypyridone nucleobase was protected as an *o*-nitrobenzyl derivative. It has been previously shown that *o*-nitrobenzyl

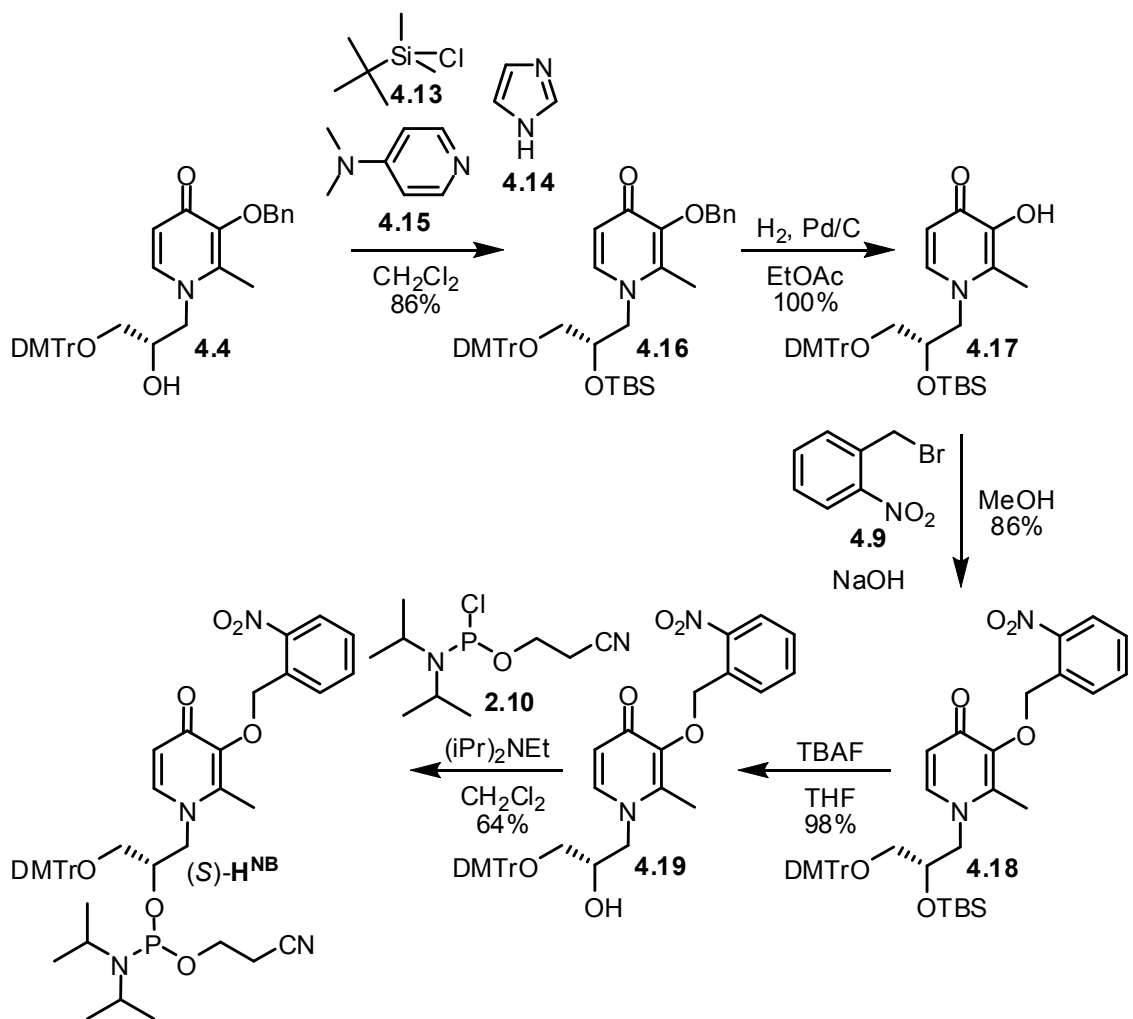
protection groups could be removed from oligonucleotides using ultraviolet light with a long enough wavelength to avoid the formation of thymine dimers.²⁷⁻³⁰ Protection of 3-hydroxy-2-methyl-4-pyrone (**4.1**) using *o*-nitrobenzyl bromide (**4.9**) and aqueous sodium hydroxide yielded compound **4.10** in 66% yield (Scheme 4.4). Compound **4.10** was then further converted to the pyridone **4.11** using 25% aqueous ammonium hydroxide in 69% yield. Unfortunately, attempts at ring opening of compound **2.11** with **4.11** did not produce the desired product.



Scheme 4.4. Attempted synthesis of compound **4.12**.

Therefore, an alternate route was developed in which the 2'-hydroxyl group of compound **4.4** was protected using *t*-butyldimethylsilyl chloride (**4.13**), imidazole (**4.14**), and DMAP (**4.15**) to produce compound **4.16** in 48% yield (Scheme 4.5). The benzyl group of compound **4.16** was subsequently removed using catalytic hydrogenation to produce compound **4.17** in quantitative yield. Protection of the nucleobase hydroxyl group was accomplished by the reaction of compound **4.17** with *o*-nitrobenzyl bromide (**4.9**) and aqueous sodium hydroxide to provide compound **4.18** in 86% yield. After TBAF deprotection to afford compound **4.19** in 98% yield, conversion to (*S*)-**H^{NB}**

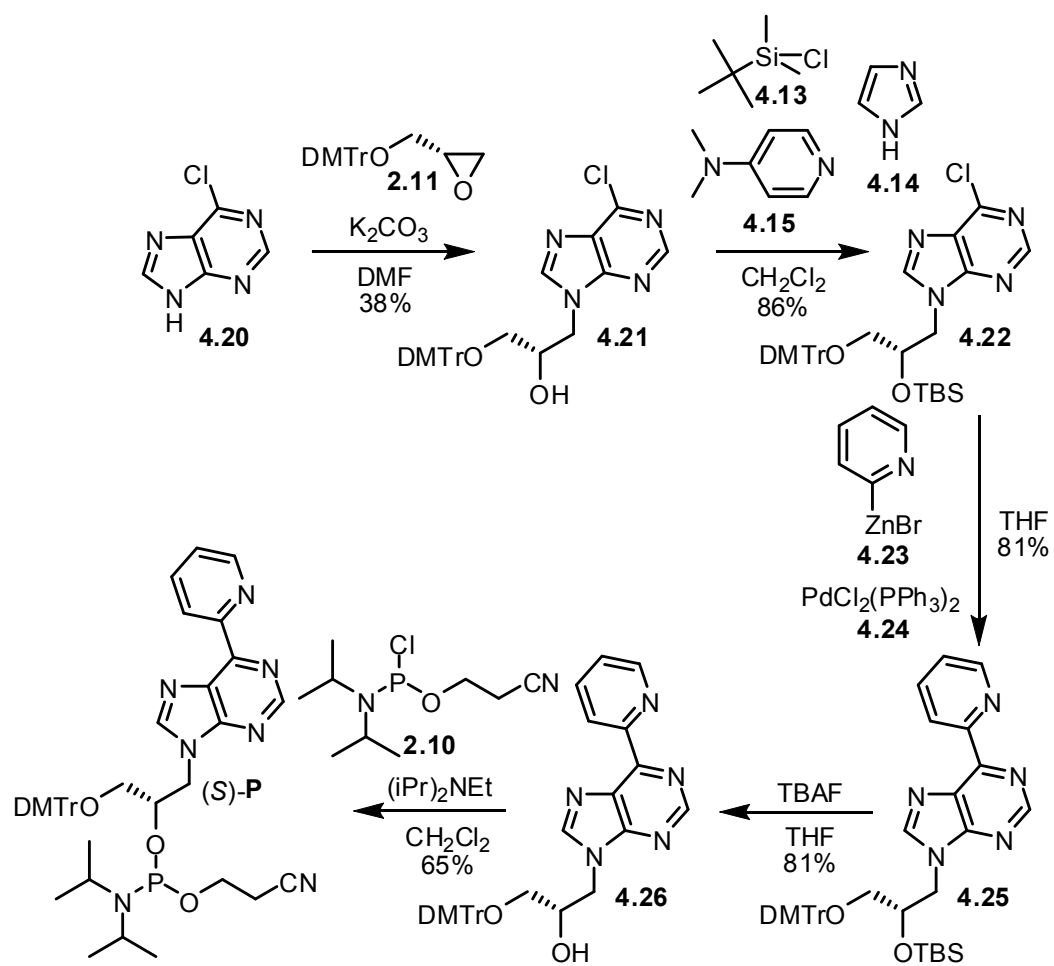
phosphoramidite was accomplished in 64% yield using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess *N,N*-diisopropylethylamine.



Scheme 4.5. Synthesis of (*S*)-**H^{NB}** phosphoramidite.

For the synthesis of (*S*)-**P** phosphoramidite (Scheme 4.6), 6-chloropurine (**4.20**) and 0.14 equivalents of potassium carbonate were used in the ring opening of compound **2.11** to produce compound **4.21** in 38% yield.¹⁹ The 2'-hydroxyl group of compound **4.21** was further protected using *t*-butyldimethylsilyl chloride (**4.13**), imidazole (**4.14**)

and catalytic dimethylaminopyridine (**4.15**) to produce compound **4.22** in 86% yield. This then allows for the conversion to compound **4.25** via a Negishi coupling with 2-pyridylzinc bromide (**4.23**) and PdCl₂(PPh₃)₃ (**4.24**) in 81% yield. After removal of the TBS group from compound **4.25** with TBAF in 81% yield, compound **4.26** could be reacted with 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess *N,N*-diisopropylethylamine to produce phosphoramidite (*S*)-**P** in 65% yield. This synthesis proceeds in five steps with an overall yield of 17% versus the synthesis of the analogous 2'-deoxynucleotide which has a reported overall yield of 26% over six steps.^{22,31,32} Although this synthesis proceeds with a lower overall yield than the corresponding 2'-deoxynucleotide, it proceeds without the separation of an anomeric mixture and can be easily accomplished with the opposite enantiomer of GNA rendering it a more attractive overall synthesis.



Scheme 4.6. Synthesis of (S)-P phosphoramidite.

Chapter 4.2. Incorporation of (S)-H and (S)-P into GNA oligonucleotides

As with the other phosphoramidites of GNA, the synthesis of phosphoramidites (S)-H and (S)-P was also followed by the general storage procedure in which trace water was removed by repeated coevaporation with toluene. The phosphoramidites [(S)-A, (S)-T, (S)-H, and (S)-P] were left under high vacuum overnight and the next morning they were dissolved using anhydrous acetonitrile to a final concentration of 100 mM. The solutions were transferred to oven dried vials specially designed for use with the automated oligonucleotide synthesizer. Care was taken to avoid exposure to air during all manipulations. Solid supports (A, T) were weighed out and transferred to polypropylene synthesis columns sealed with column frits. Prior to synthesis, the solid support was dried for five minutes directly on the synthesizer using argon.

Oligonucleotides were synthesized on an Applied Biosystems ABI 394 automated DNA/RNA synthesizer. In addition to the changes stated in Chapter 2.2, the coupling time for (S)-H and (S)-P phosphoramidites was extended five minutes to give a total coupling time of eight minutes for these bases. Best results were obtained when the last 4,4'-dimethoxytrityl group protecting the 3'-hydroxyl group was left attached after synthesis as a purification handle. The solid supports containing oligonucleotides were deprotected and purified as stated in Chapter 2.2. A representative HPLC trace is shown in Figure 4.1. Extinction coefficients were calculated based on deoxynucleotide increments and the reported $\epsilon = 5340^{21}$ for H and an estimated $\epsilon = 15400$ for P.

Chapter 4.3. Thermal stabilities of metal-mediated base pairs in GNA

The nucleotides **H** and **P** were incorporated into the middle of a 15-mer duplex in order to investigate the influence of metal ions on the thermal stabilities of H•H and P•P homo-base pairs in GNA.¹⁹ It was important that trace amounts of metal were removed from both the oligonucleotide and buffer solutions using Chelex 100 (Sigma) in order to obtain reproducible results. Chelex 100 is a resin functionalized with iminodiacetic acid groups in order to bind transition metals in solution. In the absence of transition metals, the H•H base pair destabilizes the 15-mer GNA duplex by 13.5 °C compared to an A•T base pair in the same position (Table 4.1). Similar to the H•H base pair in DNA, the GNA duplex is also stabilized the strongest by the addition of Cu⁺² ions, with an increase in T_m of 33.2 °C (Figure 4.2). This duplex containing one H•H pair is 19.7 °C more stable compared to an A•T base pair in the same position. Of the other metals tested, only the addition of Co(NO₃)₂ and ZnCl₂ result in significant stabilizations of 5.1 °C and 15.9 °C compared to the H•H pair in the absence of metal. The Cu⁺² stabilization of the H•H pair in a GNA duplex significantly exceeds the reported stabilization for the analogous 2'-deoxyribonucleotide H•H pair in a 15-mer DNA duplex in which a stabilization of 13.1 °C reported.

Table 4.1. Thermal stabilities of GNA duplexes in the presence of metal ions

3'-AAT ATT AXT ATT TTA-2'				
2'-TTA TAA TYA TAA AAT-3'				
Entry	X•Y	Metal ion	T _m (°C)	ΔT _m (°C)
1	A•T	None	50.5	-
2	A•T	NiCl ₂	51.0	+0.5
3	A•T	CuSO ₄	50.8	+0.3
4	A•T	ZnCl ₂	50.8	+0.3
5	H•H	None	37.0	-
6	H•H	CuSO ₄	70.2	+33.2
7	H•H	ZnCl ₂	52.9	+15.9
8	H•H	Co(NO ₃) ₂	42.1	+5.1
9	H•H	Ni(NO ₃) ₂	38.5	+1.5
10	H•H	Cd(NO ₃) ₂	37.9	+0.9
11	H•H	AuCl ₃	37.1	+0.1
12	H•H	AgNO ₃	37.0	0.0
13	H•H	Pd(NO ₃) ₂	36.6	-0.4

Measurements were performed in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate. Each sample contained 2 μM of each strand and 4 μM of the specified metal salt. Each data point is the average of at least two measurements.

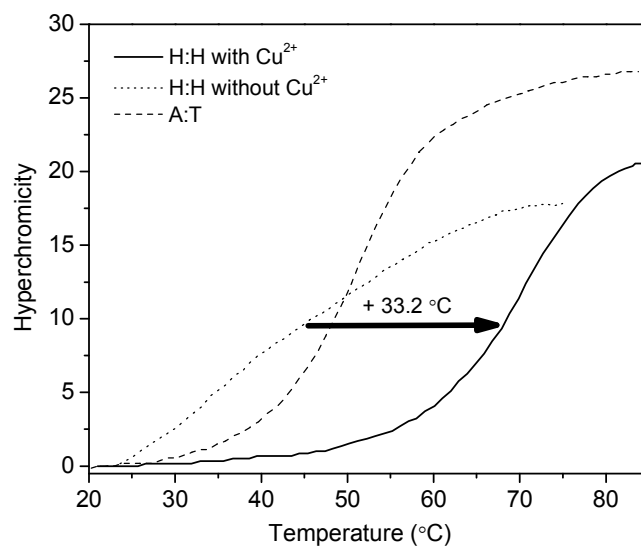


Figure 4.2. Copper(II)-dependent stability of a GNA duplex containing a hydroxypyridone (H) homo-base pair with the reference curve shown for an A•T base pair. Measurements were performed with a 2 μM duplex concentration and 4 μM CuSO_4 in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate.

Similarly, in the absence of transition metals the P•P base pair leads to a destabilization of 15.5 $^{\circ}\text{C}$ compared to an A•T base pair in the same position (Table 4.2).¹⁹ Upon the addition of Ni^{+2} ions, the duplex is stabilized with an increase in the T_m value of 17.9 $^{\circ}\text{C}$ (Figure 4.3). Thus, the duplex with a Ni^{+2} -containing P•P pair is 2.4 $^{\circ}\text{C}$ more stable compared to the GNA duplex containing an A•T base pair in the same position. Of the other metals tested, the addition of $\text{Co}(\text{NO}_3)_2$, CuCl_2 , and AgNO_3 show stabilizations of 7.0 $^{\circ}\text{C}$, 7.3 $^{\circ}\text{C}$, and 12.5 $^{\circ}\text{C}$, respectively. None of the other metals display any significant stabilizing effect on the GNA duplex. The stabilizing effect of Ni^{+2} on the P•P pair in GNA is similar to the reported stabilization of 17.6 $^{\circ}\text{C}$ (NiCl_2) and 18.1 $^{\circ}\text{C}$ ($\text{Ni}(\text{NO}_3)_2$) for the analogous 2'-deoxyribonucleotide P•P pair in DNA.

Similarly, in DNA CoCl_2 leads to a stabilization of the P•P base pair of 10.3 °C, whereas AgNO_3 and CuCl_2 , in contrast to what is observed in GNA, do not provide significant stabilization. Thus, the pyridylpurine base pair displays modulated metal ion selectivity within the GNA duplex with a preference order of $\text{Ni}^{+2} > \text{Ag}^+ > \text{Cu}^{+2} \approx \text{Co}^{+2}$, compared to $\text{Ni}^{+2} > \text{Co}^{+2} \gg \text{Cu}^{+2} \approx \text{Ag}^+$ in DNA.^{19,22}

Table 4.2. Thermal stabilities of GNA duplexes in the presence of metal ions

3'-AAT ATT AXT ATT TTA-2' 2'-TTA TAA TYA TAA AAT-3'				
Entry	X•Y	Metal ion	T _m (°C)	ΔT _m (°C)
1	P•P	None	35.0	-
2	P•P	NiCl_2	52.9	+17.9
3	P•P	AgNO_3	47.5	+12.5
4	P•P	CuCl_2	42.3	+7.3
5	P•P	$\text{Co}(\text{NO}_3)_2$	42.0	+7.0
6	P•P	CdSO_4	36.5	+1.5
7	P•P	ZnCl_2	36.0	+1.0
8	P•P	AuCl_3	34.6	-0.4
9	P•P	$\text{Pd}(\text{NO}_3)_2$	32.8	-2.2

Measurements were performed in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate. Each sample contained 2 μM of each strand and 4 μM of the specified metal salt. Each data point is the average of at least two measurements.

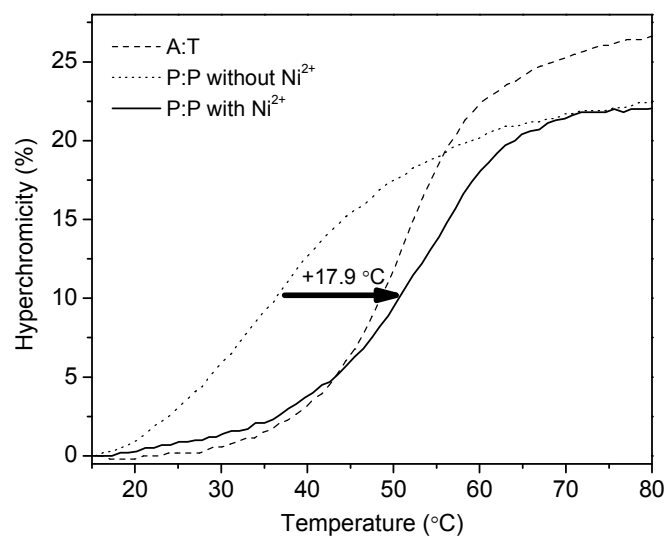


Figure 4.3. Nickel(II)-dependent stability of a GNA duplex containing a pyridylpurine (P) homo-base pair with the reference curve shown for an A•T base pair. Measurements were performed with a 2 μM duplex concentration and 4 μM NiCl_2 in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate.

In addition, we investigated the metal dependent crosspairing behavior of the H•P hetero-base pair.¹⁹ Again, in the absence of transition metals, the H•P pair shows a destabilization of 13.3 °C compared to an A•T base pair in the same position (Table 4.3). Interestingly, the addition of Cu^{+2} leads to a large stabilization of 37.1 °C resulting in a duplex with a stability that is 23.9 °C more stable than an A•T base pair in the same position (Figure 4.4). Also interesting is that the behavior of this hetero-base pair is similar to the H•H homo-base pair in that both ZnCl_2 and $\text{Co}(\text{NO}_3)_2$ provide significant stabilizations of 13.6 °C and 8.9 °C, respectively. Furthermore, this H•P hetero-base pair is significantly stabilized by the addition of Ni^{+2} , suggesting this base pair displays a metal ion selectivity that is a mixture of what is observed for the H•H and P•P base pairs.

Table 4.3. Thermal stabilities of GNA duplexes in the presence of metal ions

Entry	X•Y	Metal ion	T _m (°C)	ΔT _m (°C)
1	H•P	None	37.3	-
2	H•P	CuSO ₄	74.4	+37.1
3	H•P	ZnCl ₂	50.9	+13.6
4	H•P	Ni(NO ₃) ₂	50.1	+12.8
5	H•P	Co(NO ₃) ₂	46.2	+8.9
6	H•P	CdSO ₄	38.3	+1.0
7	H•P	AuCl ₃	37.4	+0.1
8	H•P	AgNO ₃	37.0	-0.3
9	H•P	Pd(NO ₃) ₂	36.4	-0.9

Measurements were performed in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate. Each sample contained 2 μM of each strand and 4 μM of the specified metal salt. Each data point is the average of at least two measurements.

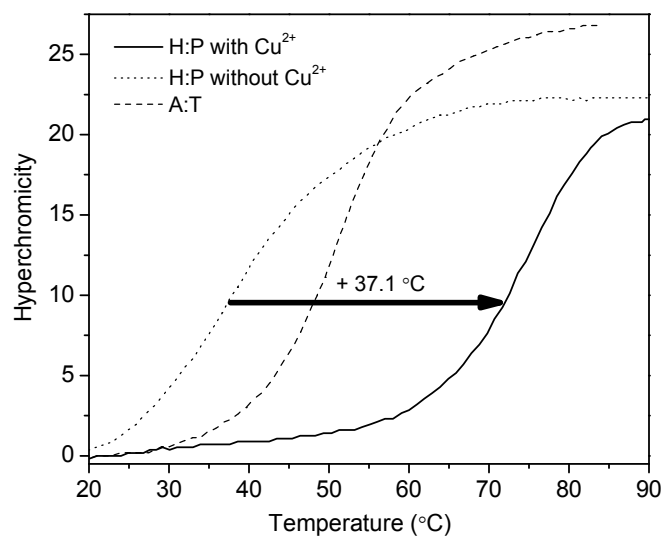


Figure 4.4. Copper(II)-dependent stability of a GNA duplex containing a hydroxypyridone (H):pyridylpurine (P) hetero-base pair with the reference curve shown for an A•T base pair. Measurements were performed with a 2 μM duplex concentration and 4 μM CuSO_4 in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate.

We were also interested in determining whether T•T mismatches could be stabilized in the context of a GNA duplex. Previous experiments with DNA have shown that thymine mismatches are capable of being stabilized by the addition of mercury(II), thus providing a metal-mediated base pair consisting of only natural nucleobases.³³ We therefore tested whether a T•T mismatch in the middle of a 15-mer GNA duplex could be stabilized by the addition of Hg^{+2} ions. Accordingly, the incorporation of one T•T mismatch into the middle of the 15-mer GNA duplex results in a destabilization of 14.5 $^\circ\text{C}$ compared to an A•T base pair in the same position (Figure 4.5). However, upon the addition of two equivalents of Hg^{+2} ions, the duplex is only destabilized by 9.8 $^\circ\text{C}$ versus an A•T base pair in the same position. This is in contrast to a similar experiment using

DNA in which a stabilization of 3 °C was observed for the addition of Hg⁺² to a 21-mer duplex compared to an A•T base pair in the same position.³³ Although this stabilization was 3 °C over the duplex, it should be noted that the overall stabilization to the T•T mismatch was only 10 °C, 5 °C higher than in GNA. Also, the duplex used in the study for DNA was composed of poly-purine and poly-pyrimidine oligonucleotides which may further explain the different results to the GNA study using a mixed sequence.

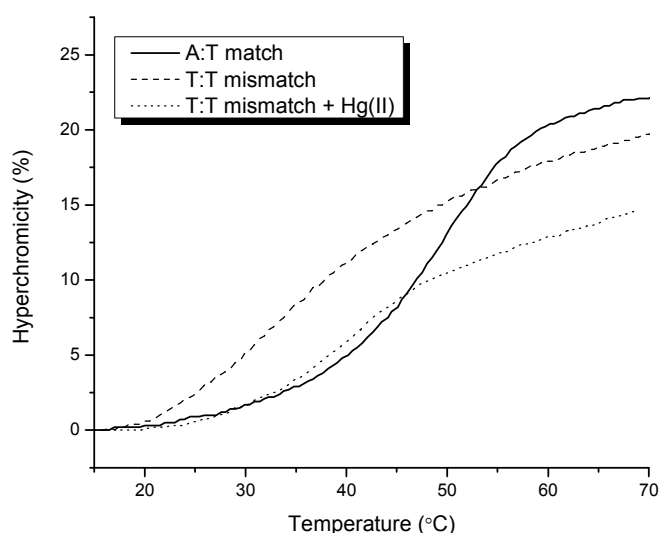


Figure 4.5. Thermal melting curves monitored by UV spectroscopy of 1:1 mixtures (2 μ M each strand) of 3'-AATATTATTATTTTA-2':3'-TAAAATAATAATATT-2' (T_m = 51 °C), 3'-AATATTATTATTTTA-2':3'-TAAAATATTAATATT-2' (T_m = 36 °C), and 3'-AATATTATTATTTTA-2':3'-TAAAATAATAATATT-2' with 4 μ M HgCl₂ (T_m = 41 °C). Measurements were made in 10 mM MOPS buffer (pH=7.1) with 100 mM sodium nitrate.

Finally, we wanted to determine whether it was possible to selectively incorporate two different metals into one GNA duplex, thereby accomplishing the site-specific incorporation of two different redox-active metals into a nucleic acid duplex. Initially,

we attempted to incorporate the two bases **H** and **P** (using phosphoramidites (*S*)-**H** and (*S*)-**P**) into a 15-mer duplex of the following sequence 3'-AATAPTATTAHTTTA-2' and its complement. Although the synthesis of the oligonucleotides proceeded normally, problems were encountered during the purification of the individual strands. The HPLC showed several broad peaks with little, if any, separation making this method of incorporation impractical. Therefore, it was envisioned that the use of the (*S*)-**H**^{NB} phosphoramidite would solve the purification problems because it would allow for a simpler purification of the oligonucleotides. Accordingly, the phosphoramidites (*S*)-**H**^{NB} and (*S*)-**P** were used to synthesize the two sequences 3'- AATAPTATTA**H**^{NB}TTTA-2' and 3'-TAAAH^{NB}TAATAPTATT-2'. As expected, the purification of these oligonucleotides proceeded smoothly using the *o*-nitrobenzyl group to protect the hydroxyl group of the hydroxypyridone nucleobase. Once these oligonucleotides were purified, irradiation with ultraviolet light was required to remove the *o*-nitrobenzyl protection group²⁸ rendering the hydroxypyridone nucleobase capable of Cu⁺² mediated homobase pairing. Afterwards, the oligonucleotides could be easily separated from other impurities by HPLC (Figure 4.6).

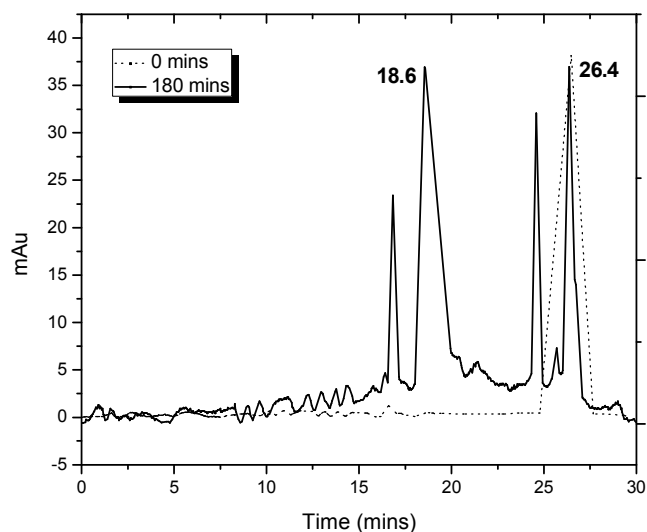


Figure 4.6. HPLC traces of the GNA oligonucleotide 3'-AAT APT ATT AH^{NB}T TTA-2' before (dotted line) and after (solid line) exposure to UV light for three hours. The peak at 26.4 minutes represents the full length product containing a nitrobenzyl protection group and the peak at 18.6 minutes represents the GNA oligonucleotide after removal of the nitrobenzyl protection group (3'-AAT APT ATT AHT TTA-2'). The crude oligos were eluted using a Waters Xterra column (MS C₁₈, 4.6 x 50 mm, 2.5 μM particle size) at 60 °C with a linear gradient (flow = 1.0 mL/min) from 4-14% acetonitrile in 30 minutes and 96-86% aqueous tritethylammonium acetate buffer (50 mM, pH=7.0).

With these two oligonucleotides in hand, the duplex stability was measured in the presence of two equivalents of both copper(II) and nickel(II). Since both the H-Cu⁺²-H and P-Ni⁺²-P base pairs increase the stability of the duplex when incorporated into the middle of a 15-mer GNA oligonucleotide, one would expect that combining these two base pairs would provide a duplex of even higher stability. As shown in Figure 4.7, the duplex containing one of each of these base pairs is actually destabilized compared to the duplex containing only one Cu(II) mediated hydroxypyridone homobase pair. This led us to believe that one of these base pairs might become destabilized in the presence of other

cations and tested accordingly as shown in Figure 4.8. For the Cu(II) mediated hydroxypyridone homobase pair, the presence of two equivalents of Cu(II) and two equivalents of Ni(II) had no effect on the thermal stability of the duplex. However, for the Ni(II) mediated pyridylpurine homobase pair, the presence of two equivalents of Cu(II) in the solution already containing two equivalents of Ni(II) has a destabilizing effect on the pre-formed duplex. Thereafter, due to the sensitivity of the pyridylpurine homobase pair to Cu(II), the incorporation of two metals selectively into one GNA duplex was determined to be more difficult than initially thought and our interest faded.

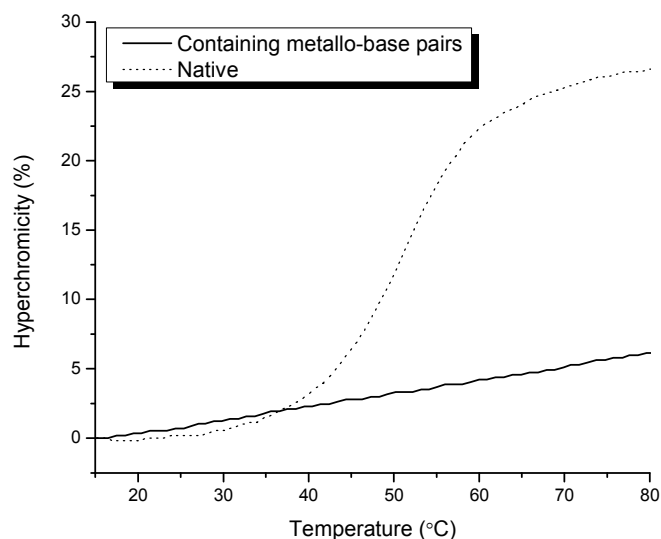


Figure 4.7. Thermal melting curves monitored by UV spectroscopy of 3'-AAT APT ATT AHT TTA-2' and its complement at 2 μ M duplex concentration (solid line). Also shown is the sequence 3'-AAT ATT ATT ATT TTA-2' and its complement at 2 μ M duplex concentration. Measurements were taken in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate.

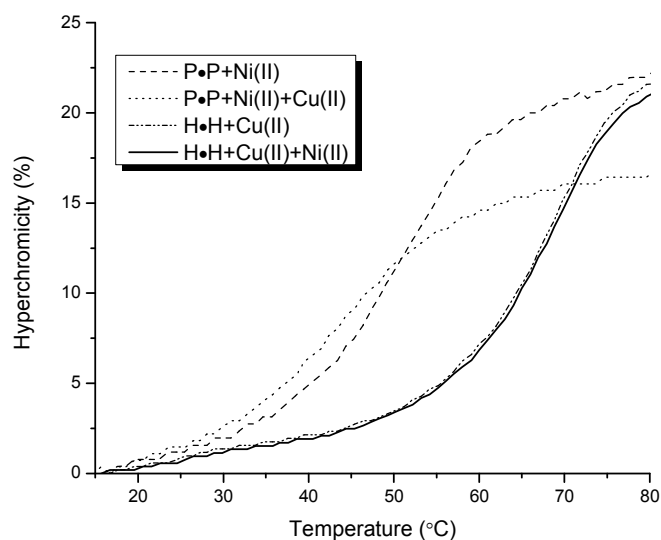


Figure 4.8. Thermal melting curves monitored by UV spectroscopy of 3'-AAT ATT APT ATT TTA-2' and its complement at 2 μ M duplex concentration with two equivalents of nickel. Also shown is the same duplex containing two equivalents of nickel followed by the addition of 2 equivalents of copper. The corresponding duplex composed of 3'-AAT ATT AHT ATT TTA-2' and its complement at 2 μ M duplex concentration with two equivalents of copper is also shown. Finally, the same duplex containing two equivalents of copper followed by two equivalents of nickel is also shown. Measurements were taken in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate.

Chapter 4.4. Metal-mediated base pairing as a tool for crystallography

Chapter 4.4.1. Duplex structure of a (S)-GNA duplex containing H-Cu-H

Based on the observed high stability of the H•H homo-base pair in GNA, we envisioned that the introduction of this base pair into GNA duplexes would be advantageous for solving the crystal structure of a GNA duplex. Not only could the highly stable duplexes of GNA be advantageous for growing high quality crystals, but the site-selective introduction of Cu^{+2} ions into the duplex has the possibility to be a convenient handle for phasing the crystallographic data as had been previously shown in the case of a DNA duplex.³⁴ Accordingly, we introduced the H•H homo-base pair into six self-complementary GNA duplexes to be used for crystallography (Table 4.4). Although the synthesis and HPLC purification of the “Trityl-ON” oligonucleotides proceeded smoothly, the subsequent purification after the removal of the trityl group was problematic. Again, because of the high stability, it was envisioned that the addition of copper(II) ions to the crude oligonucleotide solutions would aid in the purification by forming a highly stable duplex which could be purified at room temperature *as the duplex*. Unfortunately, upon the addition of excess Cu^{+2} ions, a precipitate formed which turned out to be some metal-GNA complex since no oligonucleotide could be observed in the remaining solution by HPLC. The four 8-mer GNA oligonucleotides were able to be recovered by the addition of strong acid, but the two 12-mer oligos were unrecoverable. After the 8-mer oligonucleotides were redissolved, HPLC purification proceeded at room temperature and the pure strands were isolated, presumably as the duplex.

Table 4.4. GNA oligonucleotides for crystallography utilizing metal-mediated base pairing

Name	Sequence	Molecular Weight	T _m (°C) ^[a]
MKS62	3'-CGCGHATHCGCG-2'	3129	-
MKS63	3'-CGCHAATTHGCG-2'	3128	-
MKS64	3'-CGHATHCG-2'	2060	78
MKS65	3'-CHAATTHG-2'	2059	51
MKS66	3'-AHGCGCHT-2'	2060	none
MKS67	3'-ATHCGHAT-2'	2059	none

[a] Melting temperatures determined at a duplex concentration of 4 μ M in 100 mM sodium phosphate buffer (pH = 7.0) with 25 mM sodium chloride.

In order to determine whether these purified 8-mer GNA oligonucleotide duplexes were interesting for crystallography, the thermal melting points (T_m) were determined (Figure 4.9). Thermal stability should play a large part in the crystallization of oligonucleotide duplexes considering a duplex that is not thermally stable enough probably does not possess the amount of order to produce high quality crystals. As shown in Figure 4.9, two of the four duplexes show thermal melting curves indicative of duplex formation. These two duplexes formed from MKS64 and MKS65 have thermal stabilities of 78 and 51 °C, which are higher than the stabilities of the related DNA and GNA sequences 5'(3')-CGA ATT CG-3'(2') of 26 and 40 °C, respectively.

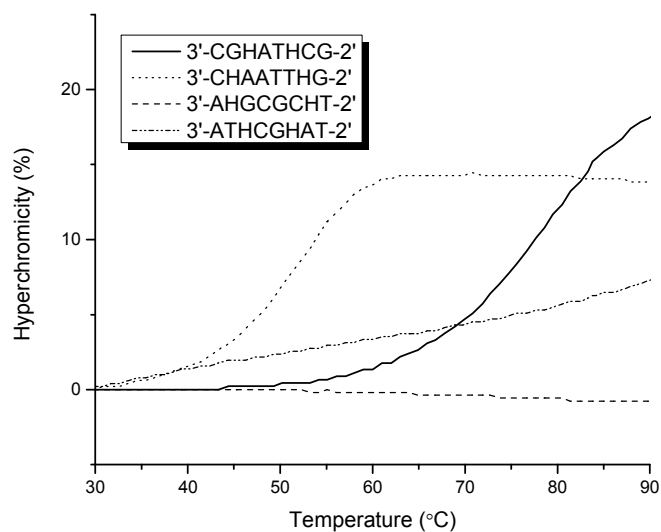


Figure 4.9. Thermal melting curves monitored by UV spectroscopy of (S)-GNA oligos containing H:H metallo-base pairs for crystallography at a duplex concentration of 2 μ M and 8 μ M added CuSO_4 . Measurements were taken in 10 mM sodium phosphate buffer (pH = 7.0) with 100 mM sodium nitrate.

With the duplexes in hand, the solution of purified oligonucleotide was concentrated and then redissolved at a concentration of 1 mM. The oligos were set up with the Nucleic Acid Mini Screen (Hampton) at 4 °C and monitored for crystal formation. The Nucleic Acid Mini Screen uses buffers containing a range of pH, polyamines, and cations. The buffer, also containing 10% 2-methyl-2,4-pentanediol, is mixed with the sample which is then equilibrated against a well containing 35% 2-methyl-2,4-pentanediol. This is in contrast to other screens which typically use the same buffer for the sample and well. Crystals developed for all four sequences in a time period of one week, however, the optically best looking crystals of the sequence 3'-CGHATHCG-2' (MKS64) were selected for data collection. After a fluorescence

scan to determine the precise copper absorption edge in the crystal, data was collected at three wavelengths for MKS64 with separate scans for high and low resolution reflections. Processing and phasing the data provided an initial map with density that was unambiguous for all the bases and phosphates of the duplex, along with one copper ion per GNA strand, 86 water molecules, two cobalt hexamine molecules, 2 magnesium ions, and 1 sodium ion (see appendix for crystallographic data).³⁵

Initial inspection of the refined crystal structure of the GNA oligonucleotide 3'-CGH ATH CG-2' shows that in the crystal individual duplexes are coaxially stacked, thereby forming a continuous helix within the crystal (Figure 4.10).³⁵ Minimal distortion of the terminal C:G base pairs allows the duplexes to stack on each other within the crystal. Also within the crystal, duplexes pack in a hexagonal fashion with contacts between neighboring duplexes mediated by cobalt hexamine binding sites between the central phosphate groups of neighboring strands and also between the 2'-terminal guanine nucleotides on neighboring strands (Figure 4.10c). Three of the six NH₃ groups of the cobalt hexamine are within hydrogen bonding distance (2.8-3.0 Å) to the O6 and N7 atoms on the guanine residues.

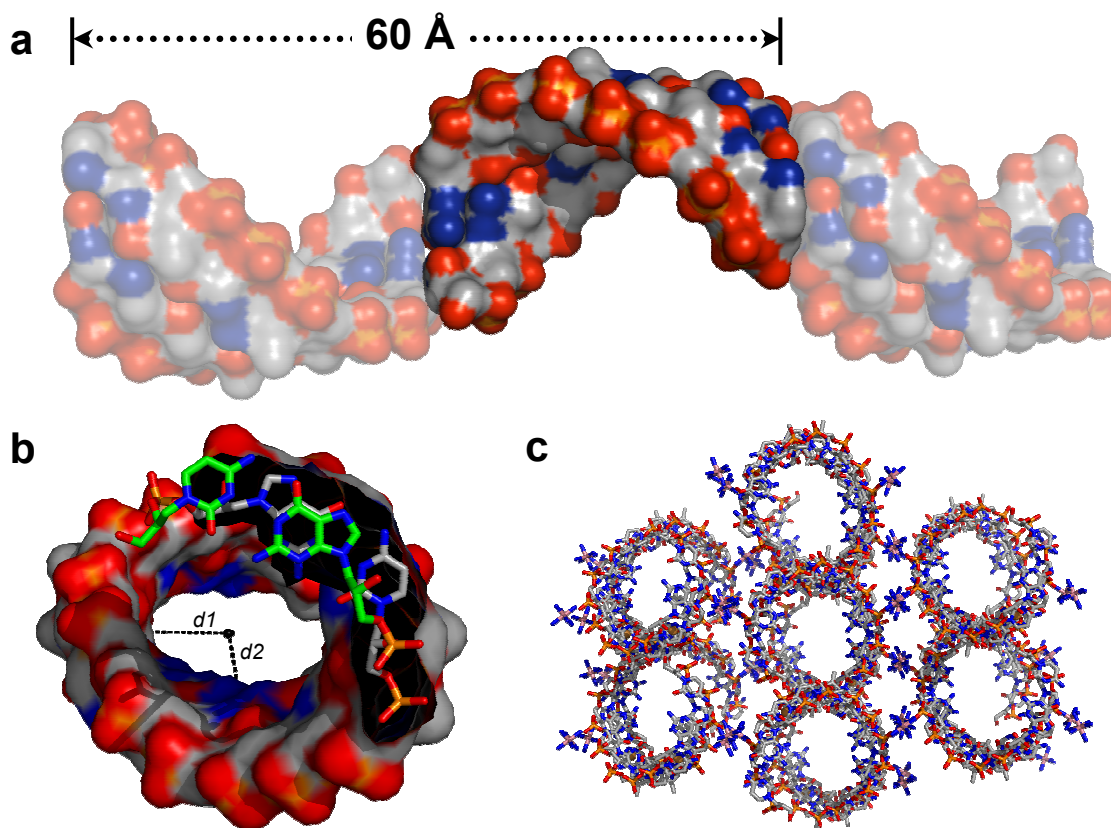


Figure 4.10. Overall GNA duplex structure. [a] Continuous packing of octamer duplexes along the crystallographic z-axis. [b] View along the z-axis with approximate distances $d1 = 7.0 \text{ \AA}$ and $d2 = 4.5 \text{ \AA}$ (defined from the helix axis to the center of the closest atom). [c] Packing of duplexes within the crystal. The structure has been deposited in the Protein Data Bank under PDB code 2JJA.

The overall structure of the right-handed (*S*)-GNA double helix differs significantly from the canonical A- and B-form nucleic acid helices possessing a very large helical pitch of 60 \AA with 16 residues per turn, resulting in a large helical rise (Table 4.5, Figure 4.10a). The base pairs are displaced from the helix axis (x-displacement) by 5.1 to 8.6 \AA , resulting in a very large elliptical hollow core (Figure 4.10b). The GNA helix possess only one large groove, corresponding to the canonical minor groove, whereas it lacks a major groove which is instead a convex surface.

Therefore, the GNA helix structure might be best described as a helical ribbon loosely wrapped around the helix axis; closely resembling the recently disclosed hexose containing nucleic acids HNA and homo-DNA rather than the canonical A- and B-form nucleic acids.^{36,37}

Table 4.5. Comparison of Average Helical Parameters for (S)-GNA, B-DNA, and A-DNA.

	(S)-GNA ^a	B-DNA ^b	A-DNA ^b
Helical sense	right	right	right
Residues per turn	16	10	12
Helical pitch (Å)	60	34	34
Helical rise (Å)	3.8	3.4	2.9
x-displacement (Å)	-7.0	0.1	-4.2
Tilt(°) ^b	0.0	0.1	-0.1
Roll (°) ^b	-2.7	0.6	8.0
Twist (°) ^b	23.5	36.0	31.0
Slide (Å) ^b	-3.5	0.2	-1.5
P-P distance (Å) ^c	5.4	7.0	5.9

[a] Data for GNA were calculated using the program CURVES.^{38,39} Data for B-DNA and A-DNA were taken from published values.^{40,41} [b] Local base pair step parameters. [c] Intrastrand P-P distances.

All natural base pairs are engaged in standard Watson-Crick hydrogen bonding, whereas the two hydroxypyridone bases coordinate to a central Cu(II) ion in a square planar fashion with a slight propeller twist of 15° . The hydroxypyridone metallo-base pair appears to fit well into the overall helix structure without any major distortions even though the $C_1'-C_1'$ distance of 12.7 \AA is expanded by 2.0 \AA compared to the standard Watson-Crick base pairs (Figure 4.11).

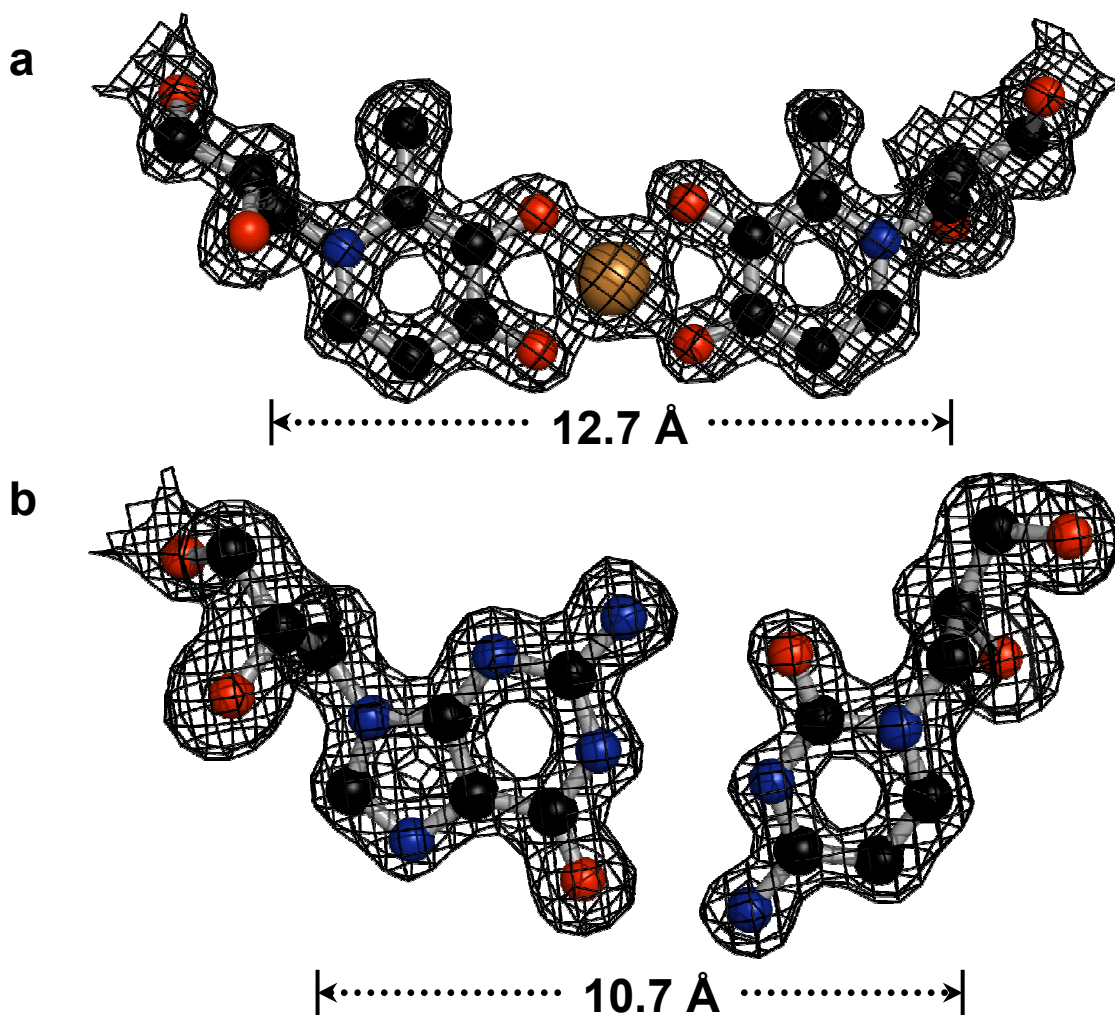


Figure 4.11. Electron density of the H-Cu-H base pair [a] and the terminal G-C base pair [b]. The $C_1'-C_1'$ distances are indicated.

Within the crystallized 8-mer duplex (Figure 4.12a), the propylene glycol nucleotides adopt two different conformations with respect to the torsional angles between C₂'-O and C₃'-O (Figure 4.12b). Whereas nucleotides of Watson-Crick base pairs maintain a *gauche* conformation with an average torsional angle γ of 70°, the glycol at the hydroxypyridone nucleotides assume an *anti* conformation ($\gamma = 165$ and 172°). As expected for such a simplified backbone, the distance between intrastrand phosphates are quite short, with an average distance of 5.4 Å, compared to around 7.0 Å for B-DNA and 5.9 Å for A-DNA.

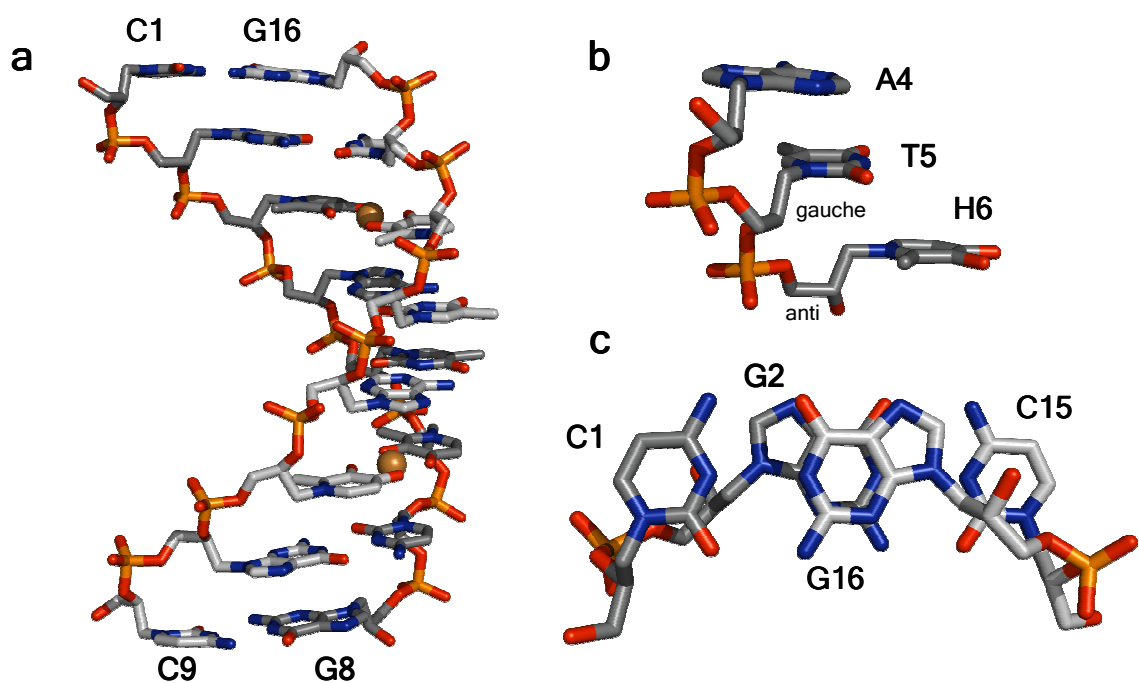


Figure 4.12. Details of the GNA duplex structure. [a] A single GNA octamer duplex. [b] The backbone conformation with *gauche* and *anti* referring to the torsional angle between C₂'-O and C₃'-O. [c] Interstrand stacking of two adjacent base pairs.

Maybe the most interesting feature of this GNA duplex structure is the large average slide between neighboring base pairs of 3.4 Å (Table 4.5, Figure 4.12c). This is a consequence of the large backbone-base inclination, ranging for this duplex from 42 to 50°. This is similar to the unnatural hexose nucleic acids but much different than the average slide of 0° for B-form DNA. The backbone-base inclination results in an almost complete absence of intrastrand base-base stacking, the predominant stacking interaction in A- and B-form nucleic acids, but extensive interstrand base-base stacking. In order to compensate for the solvent-exposed base resulting from the large base pair slide, the CH₂ group of the propylene glycol backbone is participating in packing against nucleobases of the same strand. Thus, in this simplified GNA double helix, the backbone is directly involved in hydrophobic interactions with the π -system, which might contribute to the high duplex stability of GNA.

Chapter 4.4.2. New insight into the high duplex stability of GNA

The newly solved crystal structure of an 8-mer (*S*)-GNA duplex allowed us to reevaluate old data in an attempt to understand the intrinsic stability of GNA duplexes. One of our prevailing theories is that the conformational preorganization of GNA oligonucleotides in solution is a major contributor to the higher duplex stability because of the lower entropic cost for duplex formation. More support for the helical preorganization of GNA single strands comes from the comparison of the crystallographic data of the GNA duplex with GNA nucleosides.^{18,35} In the duplex crystal structure, all natural nucleotides adopt the same staggered conformation with an *anti* orientation of the nucleobase and a *gauche* conformation with respect to the vicinal C-O bonds. A comparison of this backbone conformation with the conformation of crystallized single nucleosides reveals that this conformation is also preferred in the single nucleosides. This is shown in Figure 4.13 by the superimposition of the crystal structure of 1-(2,3-dihydroxypropyl)cytosine¹⁵ with the conformation of a cytosine nucleotide within the GNA duplex. The crystal structures of the adenine, guanine, and thymine nucleosides¹⁵ also show the preferred *gauche* conformation with respect to the vicinal C-O bonds even with different protection groups on the exocyclic amines or backbone hydroxyl groups. However, only the guanine and thymine nucleosides show an *anti* conformation with respect to the nucleobase, while the adenine nucleoside adopts an almost eclipsed conformation. It is quite possible that this difference could be the result

of the adenine nucleoside adopting a different conformation during crystal formation to maximize packing interactions between individual molecules.

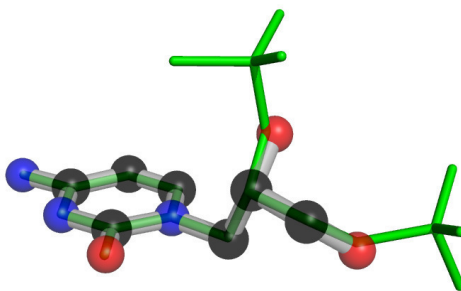


Figure 4.13. (S)-1-(2,3-dihydroxypropyl)cytosine (colored balls and sticks) superimposed with the conformation of a cytosine nucleotide within a GNA duplex (green sticks).

Furthermore, the crystal structure of the 8-mer GNA duplex sheds new light onto the stability studies with overhanging nucleotides in GNA duplexes. In GNA duplexes, overhanging nucleotides on the 3' end do not provide any increase in thermal or thermodynamic stability of the duplex, while overhanging nucleotides on the 2' end provide an exceptionally strong stabilization (Table 3.7).¹⁸ This was in stark contrast to the results for overhanging nucleotides in DNA duplexes, but similar to those results for RNA duplexes. This directionality can now be explained by the strong backbone-nucleobase inclination in GNA as show in Figure 4.12c. This prevents stacking on the 3' end (with the nucleobase almost completely exposed to solvent) but enables efficient stacking on the 2' end. Furthermore, the zipperlike interstrand stacking interactions probably contribute to the duplex stability of GNA as a result of attractive interstrand van der Waals forces.⁴²⁻⁴⁴ Finally, since GNA cannot maximize the intrastrand stacking of neighboring nucleobases because of the large backbone-base inclination and

preorganized conformation of the backbone, GNA oligonucleotides should benefit more from the conversion to the duplex where they can form stable interstrand stacking interactions.

Table 3.7. Thermal and thermodynamic stabilities of (*S*)-GNA and DNA duplexes with overhanging nucleotides^[a]

Entry	Overhanging nucleotide X ^[b]	XCGAATTCG GCTTAAGCX			CGAATTCGX XGCTTAAGC		
		T _m (°C)	ΔT _m (°C)	ΔΔG (298K, kcal/mol)	T _m (°C)	ΔT _m (°C)	ΔΔG (298K, kcal/mol)
<i>GNA duplexes</i>							
1	None	54.3	0	0	54.3	0	0
2	A	55.2	0.9	-0.2	68.9	14.6	4.0
3	G	55.5	1.2	0.1	64.9	10.6	2.4
4	C	54.8	0.5	-0.6	66.8	12.5	3.6
5	T	55.4	1.1	-0.6	62.6	8.3	2.3
<i>DNA Duplexes</i>							
6	None	35.5	0	0	35.5	0	0
7	A	45.0	9.5	1.8	39.1	3.6	0.6
8	G	41.1	5.6	0.9	36.8	1.3	0.0
9	C	42.6	7.1	1.3	39.4	3.9	0.7
10	T	40.1	4.6	0.6	36.9	1.4	-0.1

[a] Measured in 10 mM sodium phosphate buffer (pH = 7.0) with 500 mM NaCl and 2 μM of each duplex. [b] Sequence for the upper strand in the direction 3'→2' for GNA and 5'→3' for DNA.

Chapter 4.5. Conclusions

The synthetic ease associated with synthesizing phosphoramidites of GNA nucleotides makes this scaffold highly interesting for applications in the field of nanotechnology. Of the two phosphoramidites synthesized for metal-mediated base pairing in the context of GNA duplexes, only the (*S*)-**H** phosphoramidite could be synthesized in a higher yield than the corresponding 2'-deoxyribonucleotide phosphoramidite. Synthesis of the (*S*)-**P** phosphoramidite proceeds in a lower yield than the corresponding 2'-deoxyribonucleotide. However, it is believed this synthesis can be improved by modifying the conditions for the ring-opening step, or by using an alternate route. In any case, this synthesis is still attractive considering one has much easier access to the opposite enantiomer in GNA than the opposite anomer in DNA.

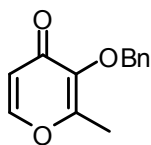
These new phosphoramidites (*S*)-**H** and (*S*)-**P** could be easily incorporated into GNA oligonucleotides by the standard procedures outlined in Chapter 2. At times, purification could be troublesome, especially with the incorporation of more than one metal-mediated base pair into a single oligonucleotide. However, greater than 98% purity was obtained in most cases for the GNA oligonucleotides synthesized for these studies. One should also consider that ion exchange chromatography may be a good means in which to separate these functionalized oligonucleotides based solely on charge and this should be considered for more complex oligos. Once incorporated, these metal-mediated base pairs in GNA displayed similar metal ion selectivities as the corresponding base pairs in DNA. In the case of the hydroxypyridone homo-base pair, the difference in

duplex stability between the base pair with and without copper is greater than in DNA. This property has already been utilized in GNA duplexes for the sensing of copper ions in solution.⁴⁵

Probably the greatest impact of these metal-mediated base pairs comes from the crystal structure of a GNA duplex determined using the hydroxypyridone homo-base pair as a handle for phasing the crystallographic data. This was important to gain a first insight into the structure of GNA duplexes in an attempt to understand their inherent greater stability compared to similar duplexes of DNA or RNA. The structure of this GNA duplex has little resemblance to either A- or B-form duplexes, but instead rather resembles the recently published structure of hexose containing nucleic acids. Finally, the structure shows that the metal-mediated base pair fits nicely within the overall duplex structure, rendering it an attractive means for gaining structural insight into other nucleic acid duplexes.

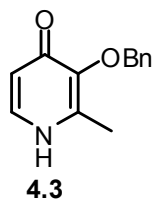
Chapter 4.6. Experimental procedures

General procedures and reagents. NMR spectra were recorded on a Bruker DRX-500 (500 MHz), DRX-400 (400 MHz), DMX-360 (360 MHz), or DMX-300 (300 MHz) spectrometer. High-resolution mass spectra were obtained with a Micromass AutoSpec or Thermo LTQ-FT instrument using ES ionization. Infrared spectra were recorded either on a Perkin Elmer 1600, Nicolet 510, or Bruker alpha series FTIR spectrometer. Solvents and reagents were used as supplied from Aldrich, Acros, Fluka, or TCI. Reactions were performed under an atmosphere of argon or nitrogen unless otherwise specified.

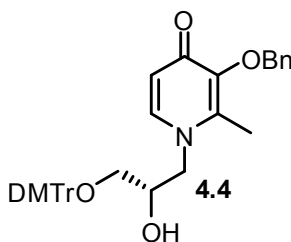


4.2

Compound 4.2. To an argon purged solution of 3-hydroxy-2-methyl-4-pyrone (10.2 g, 80.9 mmol) and 7.8 M aqueous NaOH (11 mL) in methanol (100 mL) at 0 °C was added benzyl bromide (11.1 mL, 93.0 mmol) slowly over 15 minutes. The solution was allowed to warm up gradually to room temperature and stir for two hours after which the solution was concentrated, redissolved in methylene chloride, washed with water, then with 1 M NaOH, dried over MgSO₄, and finally concentrated. The crude product was purified via flash chromatography over silica gel eluting with 1:1 Hexanes:EtOAc to afford compound **4.2** as a light yellow oil (15.6 g, 89%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 7.60 (d, *J* = 5.7 Hz, 1H), 7.41 (m, 2H), 7.38-7.30 (m, 3H), 6.37 (d, *J* = 5.6 Hz, 1H), 5.18 (s, 2H), 2.10 (s, 3H). Data matches that of previously published data.²³

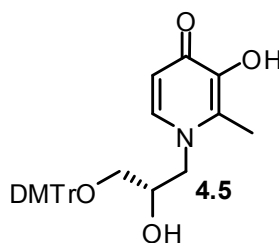


Compound 4.3. To a solution of compound **4.2** (6.01 g, 27.8 mmol) in ethanol (30 mL) was added 25% aqueous ammonium hydroxide (60 mL) and the solution allowed to stir in a closed flask at room temperature for three days. The solution was then concentrated, dry loaded onto silica gel, and purified via flash chromatography over silica gel starting with 10:1 EtOAc:MeOH, then eluting with 20:3 EtOAc:MeOH to afford compound **4.3** as a fluffy white solid (5.20 g, 87%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 7.39 (d, *J* = 7.0 Hz, 1H), 7.34-7.26 (m, 5H), 6.35 (d, *J* = 7.0 Hz, 1H), 5.07 (s, 2H), 2.16 (s, 3H). Data matches that of previously published data.²³



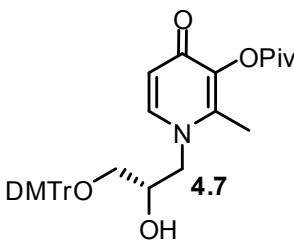
Compound 4.4. Compound **4.3** (1.3 g, 6.0 mmol) and NaH (60% in mineral oil, 50 mg, 1.3 mmol) were combined in DMF (12 mL) and stirred at room temperature under argon for one hour. A solution of compound **2.11** (2.1 g, 5.7 mmol) in DMF (12 mL) was added and the resulting mixture was heated to 110 °C overnight. The DMF was then evaporated, the residue taken up in ethyl acetate and concentrated to dryness. The crude product was purified by column chromatography starting with 100:1 EtOAc:Et₃N, then eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **4.4** as a light yellow foam

(2.2 g, 65%). $^1\text{H-NMR}$ (500 MHz, CDCl_3) δ (ppm) 7.42 (d, $J = 8.1$ Hz, 2H), 7.37-7.20 (m, 13H), 6.83 (d, $J = 8.2$ Hz, 4H), 6.11 (d, $J = 7.3$ Hz, 1H), 5.08 (d, $J = 11.3$ Hz, 1H), 4.90 (d, $J = 11.3$ Hz, 1H), 4.14 (m, 2H), 3.79 (s, 6H), 3.52 (dd, $J = 14.6, 9.7$ Hz, 1H), 3.35 (dd, $J = 9.2, 4.3$ Hz, 1H), 3.10 (t, $J = 8.4$ Hz, 1H), 2.16 (s, 3H). $^{13}\text{C-NMR}$ (125 MHz, CDCl_3) δ (ppm) 173.0, 158.8, 145.9, 144.9, 142.2, 140.4, 137.8, 136.2, 136.0, 130.3, 129.1, 128.5, 128.4, 128.2, 128.1, 127.1, 116.5, 113.4, 86.5, 73.4, 68.9, 65.5, 58.6, 55.4, 12.9. IR (film) ν (cm^{-1}) = 3197, 3063, 2933, 2836, 1626, 1608, 1561, 1508, 1462, 1398, 1301, 1251, 1220, 1176, 1073, 1033, 980, 910, 827, 790, 753, 728, 701. HRMS calcd for $\text{C}_{37}\text{H}_{38}\text{NO}_6$ ($\text{M}+\text{H}$) $^+$ 592.2699, found ($\text{M}+\text{H}$) $^+$ 592.2727.

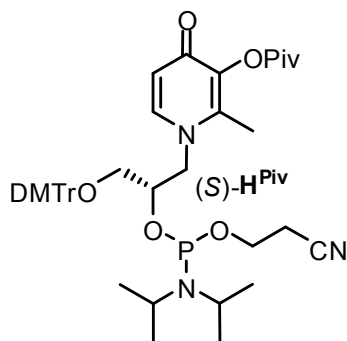


Compound 4.5. Compound **4.4** (1.5 g, 2.5 mmol) and Pd/C (10% on carbon, 750 mg) were combined in ethyl acetate (70 mL) and the resulting suspension was purged with argon, then hydrogen, and allowed to stir under a hydrogen atmosphere for two hours. The suspension was then filtered through Celite using 100:1 EtOAc:Et₃N to afford compound **4.5** as a tan foam (1.25 g, 98%) which was used without further purification. $^1\text{H-NMR}$ (500 MHz, CDCl_3) δ (ppm) 7.45 (d, $J = 7.8$ Hz, 2H), 7.37-7.15 (m, 8H), 6.86 (d, $J = 8.5$ Hz, 4H), 5.98 (br m, 1H), 5.42 (br s, 2H), 4.25 (d, $J = 13.7$ Hz, 1H), 4.03 (br s, 1H), 3.81 (s, 6H), 3.65 (m, 1H), 3.41 (m, 1H), 3.18 (m, 1H), 2.39 (s, 3H). $^{13}\text{C-NMR}$ (90 MHz, CDCl_3) δ (ppm) 168.7, 158.9, 145.7, 144.8, 138.5, 136.0, 135.8, 130.2, 129.4,

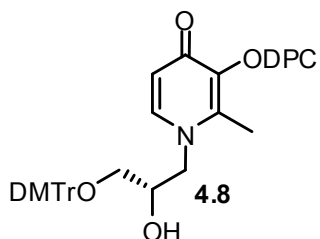
128.24, 128.18, 127.2, 113.5, 113.4, 110.8, 86.7, 69.2, 65.2, 58.0, 55.5, 12.4. IR (film) ν (cm^{-1}) = 3199, 3071, 2933, 2825, 1626, 1607, 1577, 1508, 1464, 1444, 1366, 1302, 1248, 1174, 1076, 1032, 909, 825, 791, 727, 702, 643, 584. HRMS calcd for $\text{C}_{30}\text{H}_{32}\text{NO}_6$ ($\text{M}+\text{H}$)⁺ 502.2229, found ($\text{M}+\text{H}$)⁺ 502.2248.



Compound 4.7. To an argon purged solution of compound **4.5** (1.55 g, 3.1 mmol) and *N,N*-diisopropylethylamine (0.65 mL, 3.7 mmol) in THF (12.5 mL) was added pivalic anhydride (0.69 mL, 3.4 mmol) dropwise and allowed to stir at room temperature overnight. The next morning, the solution was diluted with methylene chloride, washed with brine, dried over Na_2SO_4 , and finally concentrated. The crude product was purified via flash chromatography over silica gel starting with 100:1 EtOAc:Et₃N, then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound **4.7** as a white foam (1.05 g, 58%). ¹H-NMR (500 MHz, CDCl_3) δ (ppm) 7.44 (m, 2H), 7.35-7.21 (m, 8H), 6.85 (m, 4H), 6.01 (br, 1H), 4.18 (br, 1H), 4.07 (m, 1H), 3.81 (s, 6H), 3.64 (m, 1H), 3.41 (m, 1H), 3.12 (m, 1H), 2.27 (s, 3H), 1.39 (s, 9H).

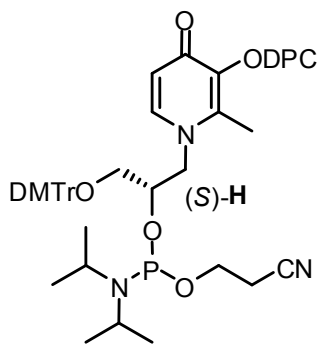


Compound (S)-H^{Piv}. To an argon purged solution of compound **4.7** (1.02 g, 1.7 mmol) and *N,N*-diisopropylethylamine (1.75 mL, 10.0 mmol) in methylene chloride (29 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.80 mL, 3.6 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The solution was diluted with methylene chloride and washed one time with saturated aqueous NaHCO₃ dried over Na₂SO₄, and finally concentrated. The crude product was purified by column chromatography eluting with 1:1:0.01 Hexanes:Acetone:Et₃N to afford compound (S)-H^{Piv} as a white foam (1.10 g, 80%). Only one product peak was observed by phosphorus NMR. ³¹P NMR (121 MHz, CDCl₃) δ (ppm) 150.3.



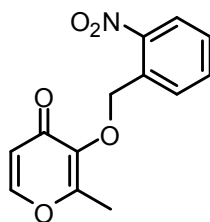
Compound 4.8. To an argon purged solution of compound **4.5** (415 mg, 0.83 mmol) and diphenylcarbamoyl chloride (230 mg, 0.99 mmol) in anhydrous pyridine (6.6 mL) was added *N,N*-diisopropylethylamine (175 μL, 0.99 mmol) and the solution stirred for one hour at room temperature under argon. The solution was diluted with methylene chloride,

washed with saturated aqueous NaHCO₃, dried over Na₂SO₄, and finally concentrated. After coevaporation with toluene, the crude product was purified by column chromatography starting with 100:1 EtOAc:Et₃N, then 50:1:0.01 EtOAc:MeOH:Et₃N, and finally eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **4.8** as a light yellow foam (370 mg, 64%). ¹H-NMR (500 MHz, 373K, DMSO-*d*₆) δ (ppm) 7.50-7.20 (m, 20H), 6.89 (d, *J* = 8.8 Hz, 4H), 6.08 (d, *J* = 7.6 Hz, 1H), 4.13 (d, *J* = 11.4 Hz, 1H), 3.84 (m, 2H), 3.76 (s, 6H), 3.14 (dd, *J* = 9.3, 3.2 Hz, 1H), 3.00 (dd, *J* = 9.3, 6.1 Hz, 1H), 2.25 (s, 3H). ¹³C-NMR (125 MHz, 373K, DMSO-*d*₆) δ (ppm) 169.2, 157.9, 150.9, 144.2, 142.3, 140.6, 140.4, 139.5, 135.3, 129.2, 128.3, 127.4, 127.2, 126.4, 126.1, 125.6, 114.5, 112.9, 85.5, 68.6, 64.9, 62.5, 55.2, 54.7, 12.0. IR (film) ν (cm⁻¹) = 3400 (br), 3059, 2928, 2833, 1742, 1644, 1608, 1598, 1512, 1494, 1352, 1306, 1250, 1218, 1200, 1176, 1151, 1053, 1025, 1002, 819, 757, 693. HRMS calcd for C₄₃H₄₁N₂O₇ (M+H)⁺ 697.2908, found (M+H)⁺ 697.2897.



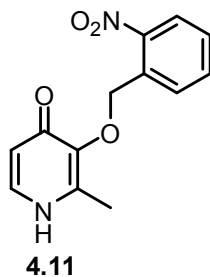
Compound (S)-H. To an argon purged solution of compound **4.8** (355 mg, 0.51 mmol) and *N,N*-diisopropylethylamine (0.51 mL, 2.9 mmol) in dichloromethane (8.5 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.24 mL, 1.1 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The

solution was diluted with methylene chloride, washed one time with saturated aqueous NaHCO₃, dried over Na₂SO₄, and finally concentrated by rotary evaporation. The crude product was purified by column chromatography eluting with 1:1:0.01 Hexanes:Acetone:Et₃N to afford compound (*S*)-**4.10** as a white foam (395 mg, 86%). ³¹P NMR (121 MHz, CDCl₃) δ (ppm) 150.6, 150.0. HRMS calcd for C₅₂H₅₈N₄O₈P (M+H)⁺ 897.3987, found (M+H)⁺ 897.3973.

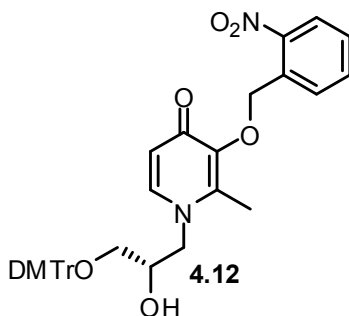


4.10

Compound 4.10. To an argon purged suspension of 3-hydroxy-2-methyl-4-pyrone (1.54 g, 12.2 mmol) and 9.8 M aqueous NaOH (1.5 mL) in methanol (5 mL) at 0 °C was added a solution of *o*-nitrobenzyl bromide (3.17 g, 14.7 mmol) in methanol (10 mL) slowly over 15 minutes. The solution was allowed to warm up gradually to room temperature and stir for two hours after which the solution was concentrated, redissolved in methylene chloride, washed with water, then with 1 M NaOH, dried over Na₂SO₄, and finally concentrated. The crude product was purified via flash chromatography over silica gel starting with 2:1 Hexanes:EtOAc, then eluting with 1:1 Hexanes:EtOAc to afford compound **4.10** as a light yellow solid (2.10 g, 66%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 8.10 (d, *J* = 8.2 Hz, 1H), 7.98 (d, *J* = 7.7 Hz, 1H), 7.69 (t, *J* = 7.5 Hz, 1H), 7.65 (d, *J* = 5.7 Hz, 1H), 7.50 (t, *J* = 7.8 Hz, 1H), 6.39 (d, *J* = 5.7 Hz, 1H), 5.49 (s, 2H), 2.31 (s, 3H).

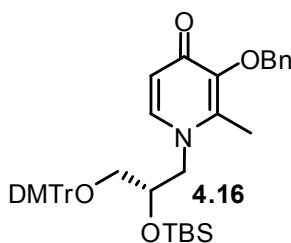


Compound 4.11. To a solution of compound **4.10** (1.71 g, 6.6 mmol) in ethanol (7 mL) was added 25% aqueous ammonium hydroxide (14 mL) and the solution allowed to stir in a closed flask at room temperature for five days. TLC indicated only ~75% conversion, so another portion of 25% aqueous ammonium hydroxide (7 mL) was added to the suspension and heated to 50 °C overnight. The suspension was then concentrated, dry loaded onto silica gel, and purified via flash chromatography over silica gel eluting with 20:3 EtOAc:MeOH to afford compound **4.11** as a tan solid (1.17 g, 69%). ¹H-NMR (360 MHz, MeOD) δ (ppm) 8.06 (dd, *J* = 8.2, 1.2 Hz, 1H), 7.95 (m, 1H), 7.73 (dt, *J* = 7.6, 1.2 Hz, 1H), 7.59-7.51 (m, 2H), 6.46 (d, *J* = 7.1 Hz, 1H), 5.44 (s, 2H), 2.24 (s, 3H).

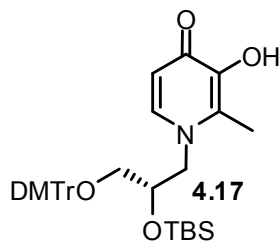


Attempted synthesis of compound 4.12. Compound **4.12** (1.03 g, 4.0 mmol) and NaH (60% in mineral oil, 33 mg, 0.83 mmol) were combined in DMF (8 mL) and stirred at room temperature under argon for one hour. A solution of **2.11** (1.42 g, 3.8 mmol) in DMF (8 mL) was added and the resulting mixture was heated to 110 °C overnight. The

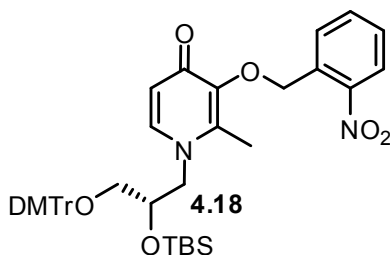
DMF was then evaporated, the residue taken up in ethyl acetate and concentrated to dryness. The crude product was purified by column chromatography starting with 100:1 EtOAc:Et₃N, then eluting with 40:3:0.01 EtOAc:MeOH:Et₃N. ¹H NMR indicated a mixture of products.



Compound 4.16. Compound **4.4** (1.61 g, 2.7 mmol) was dissolved in dichloromethane (12 mL) under argon and *tert*-butyldimethylsilyl chloride (1.03 g, 6.8 mmol), imidazole (1.67 g, 24.5 mmol), and catalytic DMAP added to the solution. This was allowed to stir overnight at room temperature and then concentrated by rotary evaporation the next morning. The crude product was purified by column chromatography starting with 1:2:0.01 Hexanes:EtOAc:Et₃N, then eluting with 100:1 EtOAc:Et₃N to afford compound **4.16** as a white foam (920 mg, 48%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 7.44 (m, 2H), 7.40 (m, 2H), 7.36-7.22 (m, 10H), 7.18 (d, *J* = 7.6 Hz, 1H), 6.84 (m, 4H), 6.35 (d, *J* = 7.5 Hz, 1H), 5.22 (m, 2H), 4.24 (dd, *J* = 14.5, 2.6 Hz, 1H), 3.83-3.77 (m, 7H), 3.63 (dd, *J* = 14.5, 9.0 Hz, 1H), 3.20 (dd, *J* = 9.7, 4.0 Hz, 1H), 3.03 (dd, *J* = 9.6, 8.1 Hz, 1H), 2.21 (s, 3H), 0.76 (s, 9H), -0.21 (s, 3H), -0.29 (s, 3H).

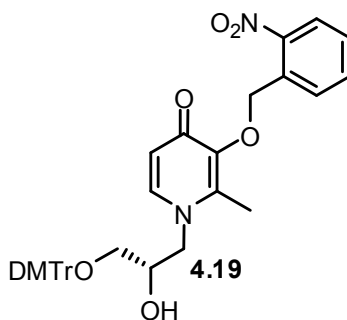


Compound 4.17. Compound **4.16** (920 mg, 1.3 mmol) and Pd/C (10% on carbon, 460 mg) were combined in ethyl acetate (36 mL) and the resulting suspension was purged with argon, then hydrogen, and allowed to stir under a hydrogen atmosphere for two hours. The product was then filtered through Celite using 100:1 EtOAc:Et₃N to afford compound **4.17** as a tan foam (810 mg, 100%) which was used without further purification. ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 7.42 (m, 2H), 7.34-7.23 (m, 7H), 7.21 (d, *J* = 7.3 Hz, 1H), 6.85 (d, *J* = 8.7 Hz, 4H), 6.32 (d, *J* = 7.3 Hz, 1H), 4.34 (dd, *J* = 14.4, 2.6 Hz, 1H), 3.89 (m, 1H), 3.80 (s, 6H), 3.72 (dd, *J* = 14.4, 8.9 Hz, 1H), 3.25 (dd, *J* = 9.7, 4.0 Hz, 1H), 3.07 (dd, *J* = 9.5, 8.2 Hz, 1H), 2.43 (s, 3H), 0.76 (s, 9H), -0.20 (s, 3H), -0.30 (s, 3H).



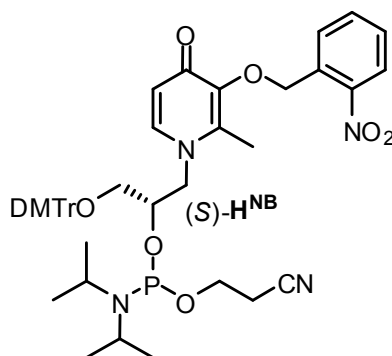
Compound 4.18. A solution of compound **4.17** (502 mg, 0.82 mmol) and *o*-nitrobenzyl bromide (338 mg, 1.6 mmol) in 3.5 M aqueous sodium hydroxide (0.28 mL) and methanol (2.9 mL) was allowed to stir at room temperature under argon for two hours. The solution was dissolved in methylene chloride, washed with 1 M NaOH, dried over

Na₂SO₄, and finally concentrated. The crude product was purified via flash chromatography over silica gel starting with 100:1 EtOAc:Et₃N, then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound **4.18** as a light yellow foam (525 mg, 86%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 8.11 (d, *J* = 7.7 Hz, 1H), 8.07 (d, *J* = 7.7 Hz, 1H), 7.68 (t, *J* = 7.5 Hz, 1H), 7.46 (t, *J* = 7.8 Hz, 1H), 7.41 (m, 2H), 7.34-7.20 (m, 8H), 6.85 (d, *J* = 8.7 Hz, 4H), 6.36 (d, *J* = 7.6 Hz, 1H), 5.48 (s, 2H), 4.30 (dd, *J* = 14.5, 2.6 Hz, 1H), 3.89 (m, 1H), 3.80 (s, 6H), 3.71 (dd, *J* = 14.6, 9.0 Hz, 1H), 3.25 (dd, *J* = 9.7, 4.0 Hz, 1H), 3.06 (dd, *J* = 9.6, 8.3 Hz, 1H), 2.39 (s, 3H), 0.77 (s, 9H), -0.18 (s, 3H), -0.25 (s, 3H).

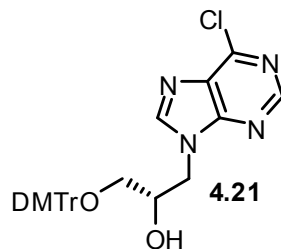


Compound 4.19. To an argon purged solution of compound **4.18** (525 mg, 0.70 mmol) in anhydrous THF (14 mL) was added TBAF (1 M in THF, 1.54 mL, 1.54 mmol) and the solution stirred for 30 minutes at room temperature. The solution was washed with water, extracted into ethyl acetate, dried over Na₂SO₄, and concentrated by rotary evaporation. The crude product was purified by flash chromatography over silica gel starting with 100:1 EtOAc:Et₃N, then eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **4.19** as a light yellow foam (435 mg, 98%). ¹H-NMR (500 MHz, CDCl₃) δ (ppm) 8.05 (dd, *J* = 8.2, 1.0 Hz, 1H), 7.98 (d, *J* = 7.5 Hz, 1H), 7.61 (dt, *J* = 7.6, 1.1 Hz, 1H), 7.45-7.39 (m, 3H), 7.34-7.26 (m, 8H), 6.83 (m, 4H), 6.11 (d, *J* = 7.5 Hz, 1H), 5.37 (m, 2H),

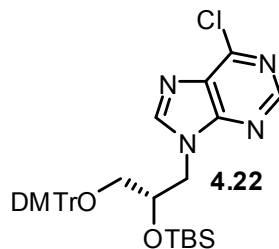
4.17 (dd, $J = 14.3, 2.2$ Hz, 1H), 4.07 (m, 1H), 3.79 (s, 6H), 3.63 (dd, $J = 14.4, 9.4$ Hz, 1H), 3.34 (dd, $J = 9.4, 4.6$ Hz, 1H), 3.11 (dd, $J = 9.3, 7.7$ Hz, 1H), 2.34 (s, 3H).



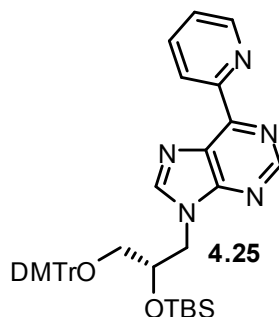
Compound (S)-H^{NB}. To an argon purged solution of compound **4.19** (435 mg, 0.68 mmol) and *N,N*-diisopropylethylamine (0.72 mL, 4.1 mmol) in methylene chloride (11 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.31 mL, 1.4 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The solution was diluted with methylene chloride and washed one time with saturated aqueous NaHCO₃, dried over Na₂SO₄, and finally concentrated. The crude product was purified by flash chromatography over silica gel loading with 3:2:0.01 Hexanes:Acetone:Et₃N, then with 1:1:0.01 Hexanes:Acetone:Et₃N, and finally eluting with 2:3:0.01 Hexanes:Acetone:Et₃N to afford compound (S)-H^{NB} as a white foam (365 mg, 64%) ³¹P NMR (121 MHz, CDCl₃) δ (ppm) 150.5, 150.3.



Compound 4.21. 6-chloropurine (2.0 g, 12.9 mmol) and K_2CO_3 (230 mg, 1.7 mmol) were combined in DMF (20 mL). A solution of compound **2.11** (4.7 g, 12.5 mmol) in DMF (20 mL) was added and the resulting mixture was heated to 90 °C overnight. The DMF was then evaporated, the residue taken up in ethyl acetate and concentrated to dryness. The crude product was purified by flash chromatography over silica gel starting with 1:1:0.01 Hexanes:EtOAc:Et₃N, then eluting with 1:2:0.01 Hexanes:EtOAc:Et₃N to afford compound **4.21** as a white foam (2.5 g, 38%). ¹H-NMR (360 MHz, CDCl₃) δ (ppm) 8.69 (s, 1H), 8.16 (s, 1H), 7.40 (m, 2H), 7.34-7.20 (m, 7H), 6.83 (m, 4H), 4.51 (dd, $J=14.3, 3.0$ Hz, 1H), 4.35 (dd, $J=14.3, 7.1$ Hz, 1H), 4.22 (m, 1H), 3.81 (s, 6H), 3.48 (b, 1H), 3.20 (d, $J=5.6$ Hz, 2H). ¹³C-NMR (90 MHz, CDCl₃) δ (ppm) 159.0, 152.2, 152.0, 151.3, 146.8, 144.6, 135.73, 135.71, 131.7, 130.2, 128.31, 128.25, 127.4, 87.0, 69.5, 64.8, 55.6, 48.0. IR (film) ν (cm⁻¹) = 3317, 2933, 2835, 1666, 1602, 1592, 1562, 1508, 1464, 1445, 1405, 1336, 1302, 1253, 1179, 1076, 1032, 948, 909, 830, 727, 702, 643. HRMS calcd for C₂₉H₂₇N₄O₄Cl (M+H)⁺ 531.1799, found (M+H)⁺ 531.1805.

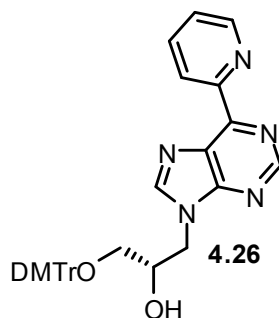


Compound 4.22. Compound **4.21** (2.6 g, 4.9 mmol) was dissolved in dichloromethane (22 mL) under argon and *tert*-butyldimethylsilyl chloride (1.7 g, 11.3 mmol), imidazole (3.0 g, 44.1 mmol), and catalytic DMAP added to the solution. This was allowed to stir overnight at room temperature and then concentrated by rotary evaporation the next morning. The crude product was purified by flash chromatography over silica gel starting with 6:1:0.01 Hexanes:EtOAc:Et₃N, then eluting with 4:1:0.01 Hexanes:EtOAc:Et₃N to afford compound **4.22** as a white foam (2.7 g, 86%). ¹H-NMR (360 MHz, CDCl₃) δ (ppm) 8.73 (s, 1H), 8.12 (s, 1H), 7.43 (d, *J*=7.3 Hz, 2H), 7.35-7.19 (m, 7H), 6.82 (m, 4H), 4.56 (dd, *J*=14.1, 3.9 Hz, 1H), 4.44 (dd, *J*=14.0, 6.6 Hz, 1H), 4.15 (m, 1H), 3.80 (s, 6H), 3.14 (dd, *J*=9.8, 4.2 Hz, 1H), 3.00 (dd, *J*=9.6, 7.1 Hz, 1H), 0.79 (s, 9H), -0.16 (s, 3H), -0.33 (s, 3H). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm) 158.76, 158.74, 152.3, 151.9, 150.9, 146.5, 144.6, 135.8, 135.7, 131.4, 130.0, 128.2, 128.0, 127.1, 113.3, 86.6, 69.8, 64.7, 55.3, 48.0, 25.8, 17.9, -4.8, -5.4. IR (film) ν (cm⁻¹) = 2954, 2927, 2858, 1608, 1590, 1560, 1508, 1464, 1442, 1403, 1333, 1302, 1246, 1176, 1080, 1036, 997, 940, 831, 779, 726, 700. HRMS calcd for C₃₅H₄₁N₄O₄SiCl (M+H)⁺ 645.2664, found (M+H)⁺ 645.2683.



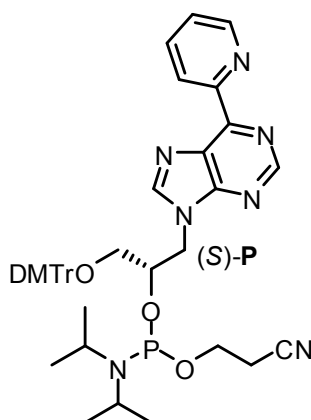
Compound 4.25. To an argon purged solution of compound **4.22** (2.7 g, 4.2 mmol) and $\text{PdCl}_2(\text{PPh}_3)_2$ (265 mg, 0.4 mmol) in anhydrous THF (85 mL) was added 2-pyridylzinc bromide (0.5 M in THF, 13.5 mL) and the solution heated to 65 °C for two hours. After cooling to room temperature, the solution was washed with saturated aqueous NaHCO_3 , extracted into dichloromethane, dried over Na_2SO_4 , and concentrated by rotary evaporation. The crude product was purified by flash chromatography over basic alumina starting with 3:1:0.01 Hexanes:EtOAc:Et₃N, then with 1:1:0.01 Hexanes:EtOAc:Et₃N, and finally eluting with 100:1 EtOAc:Et₃N to afford compound **4.25** as a tan foam (2.3 g, 81%). ¹H-NMR (400 MHz, CDCl₃) δ (ppm) 9.12 (s, 1H), 8.97 (m, 1H), 8.86 (d, *J* = 8.0 Hz, 1H), 8.23 (s, 1H), 7.93 (td, *J* = 7.7, 1.6 Hz, 1H), 7.45 (m, 3H), 7.31 (m, 6H), 7.21 (t, *J* = 7.2 Hz, 1H), 6.82 (dd, *J* = 8.9, 2.5 Hz, 4H), 4.60 (dd, *J* = 14.1, 3.7 Hz, 1H), 4.50 (dd, *J* = 14.1, 6.7 Hz, 1H), 4.21 (m, 1H), 3.77 (s, 6H), 3.16 (dd, *J* = 9.8, 4.2 Hz, 1H), 3.07 (dd, *J* = 9.8, 7.0 Hz, 1H), 0.79 (s, 9H), -0.16 (s, 3H), -0.34 (s, 3H). ¹³C-NMR (125 MHz, CDCl₃) δ (ppm) 158.73, 158.71, 154.0, 153.7, 153.3, 152.6, 150.6, 146.8, 144.7, 136.8, 136.0, 135.9, 130.1, 128.2, 128.0, 127.0, 126.0, 124.9, 113.3, 86.6, 69.9, 64.9, 55.3, 47.7, 25.8, 17.9, -4.8, -5.3. IR (film) ν (cm⁻¹) = 3499, 2956, 2929, 2853, 2362, 1612, 1581, 1510, 1463, 1441, 1384, 1325, 1301, 1251, 1174, 1116, 1068,

1030, 988, 929, 825, 773, 720, 692, 635, 570, 531. HRMS calcd for C₄₀H₄₅N₅O₄Si (M+H)⁺ 688.3319, found (M+H)⁺ 688.3314.



Compound 4.26. To an argon purged solution of compound **4.25** (1.25 g, 1.8 mmol) in anhydrous THF (38 mL) was added TBAF (1 M in THF, 4.0 mL, 4.0 mmol) and the solution stirred for 30 minutes at room temperature. The solution was washed with water, extracted into ethyl acetate, dried over Na₂SO₄, and concentrated by rotary evaporation. The crude product was purified by column chromatography over basic alumina starting with 100:1 EtOAc:Et₃N, then eluting with 40:3:0.01 EtOAc:MeOH:Et₃N to afford compound **4.26** as a tan foam (840 mg, 81%). ¹H-NMR (400 MHz, CDCl₃) δ (ppm) 9.04 (s, 1H), 8.95 (m, 1H), 8.82 (m, 1H), 8.22 (s, 1H), 7.93 (td, *J* = 7.8, 1.8 Hz, 1H), 7.46-7.38 (m, 3H), 7.31-7.25 (m, 6H), 7.21 (m, 1H), 6.81 (m, 4H), 4.54 (dd, *J* = 14.4, 2.8 Hz, 1H), 4.39 (dd, *J* = 14.3, 7.0 Hz, 1H), 4.22 (m, 1H), 3.76 (s, 6H), 3.24 (dd, *J* = 9.8, 5.9 Hz, 1H), 3.17 (dd, *J* = 9.7, 5.6 Hz, 1H). ¹³C-NMR (100 MHz, CDCl₃) δ (ppm) 158.8, 153.8, 153.5, 153.4, 152.3, 150.5, 146.9, 144.6, 136.9, 135.7, 135.6, 130.0, 128.08, 128.07, 127.1, 125.9, 125.0, 113.4, 86.7, 69.6, 64.7, 55.3, 47.9, 25.8. IR (film) ν (cm⁻¹) = 3369, 3060, 2953, 2931, 2833, 1612, 1582, 1509, 1462, 1444, 1328, 1302, 1252, 1209, 1175,

1152, 1069, 1031, 905, 828, 726, 697, 637, 577. HRMS calcd for $C_{34}H_{31}N_5O_4$ (M+H)⁺ 574.2454, found (M+H)⁺ 574.2449.



Compound (S)-P. To an argon purged solution of compound **4.26** (705 mg, 1.2 mmol) and *N,N*-diisopropylethylamine (1.0 mL, 5.9 mmol) in dichloromethane (20 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.56 mL, 2.5 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The solution was diluted with methylene chloride, washed once with saturated aqueous NaHCO₃, dried over Na₂SO₄, and finally concentrated. The crude product was purified by flash chromatography over basic alumina starting with hexanes: 1:1:0.01 EtOAc:Et₃N, then with 1:2:0.01 Hexanes:EtOAc:Et₃N, and finally with 100:1 EtOAc:Et₃N to afford compound (S)-P as a light yellow foam (620 mg, 65%). ³¹P NMR (162 MHz, CDCl₃) δ (ppm) 150.5, 150.0. HRMS calcd for $C_{43}H_{48}N_7O_5P$ (M+H)⁺ 774.3533, found (M+H)⁺ 774.3527.

Crystallization and data collection: Crystals of self-complementary duplex GNA were grown using the sitting drop vapor diffusion method with buffers from the Nucleic Acid

Mini Screen (Hampton Research). Crystallization conditions consisted of 1 mM duplex GNA (2 μ L) and buffer (4 μ L) against a reservoir of 35% MPD in water (1 mL). Flat, square-shaped crystals generally appeared after 4-7 days at 4 °C in Buffer #2 consisting of 10% 2-methyl-2,4-pentanediol, 40 mM sodium cacodylate (pH = 5.5), 20 mM cobalt hexamine, 80 mM sodium chloride, and 20 mM magnesium chloride. These crystals were cryoprotected by raising the concentration of MPD to 30% and subsequently picked from the drop with nylon loops and frozen in liquid N₂.

MAD data from a single duplex GNA crystal were recorded at beamline ID23-1, ESRF Grenoble. After performing a fluorescence scan to determine the precise copper absorption edge in the crystal, data was collected at three wavelengths with separate scans for high and low resolution reflections. All data was integrated and merged using XDS and phased by SHELXE and SHARP. The initial map provided density that was unambiguous for all the bases and phosphates of the duplex, along with one copper ion per GNA strand. Automated and manual refinements were performed using REFMAC⁵⁴⁶ and COOT.⁴⁷

Chapter 4.7. References

- (1) Seeman, N. C. *Chem. Biol.* **2003**, *10*, 1151.
- (2) Seeman, N. *Mol. Biotechnol.* **2007**, *37*, 246.
- (3) Freier, S. M.; Altmann, K. H. *Nucl. Acids Res.* **1997**, *25*, 4429-4443.
- (4) Eschenmoser, A. *Science* **1999**, *284*, 2118.
- (5) Kool, E. T. *Acc. Chem. Res.* **2002**, *35*, 936-943.
- (6) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727.
- (7) Benner, S. A. *Acc. Chem. Res.* **2004**, *37*, 784-797.
- (8) Hans-Achim, W. *Angew. Chem. Int. Ed.* **2003**, *42*, 3204-3206.
- (9) Shionoya, M.; Tanaka, K. *Curr. Opin. Chem. Biol.* **2004**, *8*, 592.
- (10) Clever, G. H.; Kaul, C.; Carell, T. *Angew. Chem. Int. Ed.* **2007**, *46*, 6226-6236.
- (11) Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shionoya, M. *Science* **2003**, *299*, 1212-1213.
- (12) Tanaka, K.; Clever, G. H.; Takezawa, Y.; Yamada, Y.; Kaul, C.; Shionoya, M.; Carell, T. *Nat Nano* **2006**, *1*, 190.
- (13) Clever, G. H.; Carell, T. *Angew. Chem. Int. Ed.* **2007**, *46*, 250-253.
- (14) Zhang, L.; Peritz, A.; Meggers, E. *J. Am. Chem. Soc.* **2005**, *127*, 4174-4175.
- (15) Zhang, L.; Peritz, A.; Carroll, P. J.; Meggers, E. *Synthesis* **2006**, 645-653.
- (16) Schlegel, M. K.; Peritz, A. E.; Kittigowittana, K.; Zhang, L.; Meggers, E. *ChemBioChem* **2007**, *8*, 927-932.
- (17) Schlegel, M. K.; Meggers, E. *J. Org. Chem.* **2009**, *74*, 4615-4618.

- (18) Schlegel, M. K.; Xie, X.; Zhang, L.; Meggers, E. *Angew. Chem. Int. Ed.* **2009**, *48*, 960-963.
- (19) Schlegel, M. K.; Zhang, L.; Pagano, N.; Meggers, E. *Org. Biomol. Chem.* **2009**, *7*, 476-482.
- (20) Zhang, L.; Meggers, E. *J. Am. Chem. Soc.* **2005**, *127*, 74-75.
- (21) Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shiro, M.; Shionoya, M. *J. Am. Chem. Soc.* **2002**, *124*, 12494-12498.
- (22) Switzer, C.; Sinha, S.; Kim, P. H.; Heuberger, B. D. *Angew. Chem. Int. Ed.* **2005**, *44*, 1529-1532.
- (23) Piyamongkol, S.; Liu, Z. D.; Hider, R. C. *Tetrahedron* **2001**, *57*, 3479.
- (24) Kamimura, T.; Tsuchiya, M.; Koura, K.; Sekine, M.; Hata, T. *Tetrahedron Lett.* **1983**, *24*, 2775.
- (25) Seela, F.; Kroschel, R. *Bioconjugate Chem.* **2001**, *12*, 1043-1050.
- (26) Gold, A.; Sangaiah, R. *Nucleosides, Nucleotides and Nucleic Acids* **1990**, *9*, 907 - 912.
- (27) Greenberg, M. M.; Gilmore, J. L. *The Journal of Organic Chemistry* **1994**, *59*, 746-753.
- (28) Dussy, A.; Meyer, C.; Quennet, E.; Bickle, T. A.; Giese, B.; Marx, A. *ChemBioChem* **2002**, *3*, 54-60.
- (29) Bai, X.; Li, Z.; Jockusch, S.; Turro, N. J.; Ju, J. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 409-413.

- (30) Aujard, I.; Benbrahim, C.; Gouget, M.; Ruel, O.; Baudin, J.-B.; Neveu, P.; Jullien, L. *Chemistry - A European Journal* **2006**, *12*, 6865-6879.
- (31) Max, H. *Chem. Ber.* **1960**, *93*, 2777-2781.
- (32) Kazimierczuk, Z.; Cottam, H. B.; Revankar, G. R.; Robins, R. K. *J. Am. Chem. Soc.* **1984**, *106*, 6379-6382.
- (33) Miyake, Y.; Togashi, H.; Tashiro, M.; Yamaguchi, H.; Oda, S.; Kudo, M.; Tanaka, Y.; Kondo, Y.; Sawa, R.; Fujimoto, T.; Machinami, T.; Ono, A. *J. Am. Chem. Soc.* **2006**, *128*, 2172-2173.
- (34) Atwell, S.; Meggers, E.; Spraggon, G.; Schultz, P. G. *J. Am. Chem. Soc.* **2001**, *123*, 12364-12367.
- (35) Schlegel, M. K.; Essen, L.-O.; Meggers, E. *J. Am. Chem. Soc.* **2008**, *130*, 8158-8159.
- (36) Declercq, R.; Aerschot, A. V.; Read, R. J.; Herdewijn, P.; Meervelt, L. V. *J. Am. Chem. Soc.* **2002**, *124*, 928-933.
- (37) Egli, M.; Pallan, P. S.; Pattanayek, R.; Wilds, C. J.; Lubini, P.; Minasov, G.; Dobler, M.; Leumann, C. J.; Eschenmoser, A. *J. Am. Chem. Soc.* **2006**, *128*, 10847-10856.
- (38) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63-91.
- (39) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 655-667.
- (40) Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67-111.
- (41) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X.-J.; Neidle, S.; Shakked, Z.; Sklenar, H.; Suzuki, M.;

- Tung, C.-S.; Westhof, E.; Wolberger, C.; Berman, H. M. *J. Mol. Biol.* **2001**, *313*, 229.
- (42) Brotschi, C.; Häberli, A.; Leumann, C. J. *Angew. Chem. Int. Ed.* **2001**, *40*, 3012-3014.
- (43) Brotschi, C.; Leumann, C. J. *Angew. Chem. Int. Ed.* **2003**, *42*, 1655-1658.
- (44) Zahn, A.; Brotschi, C.; Leumann, C. J. *Chemistry - A European Journal* **2005**, *11*, 2125-2129.
- (45) Zhou, H.; Ma, X.; Wang, J.; Zhang, L. *Org. Biomol. Chem.* **2009**, *7*, 2297-2302.
- (46) *Acta Crystallog. Sect. D* **1994**, *50*, 760-763.
- (47) Emsley, P.; Cowtan, K. *Acta Crystallog. Sect. D* **2004**, *60*, 2126-2132.

Appendix to Chapter 4

^1H , ^{13}C , and ^{31}P NMR spectra

IR spectra

Crystallographic Tables

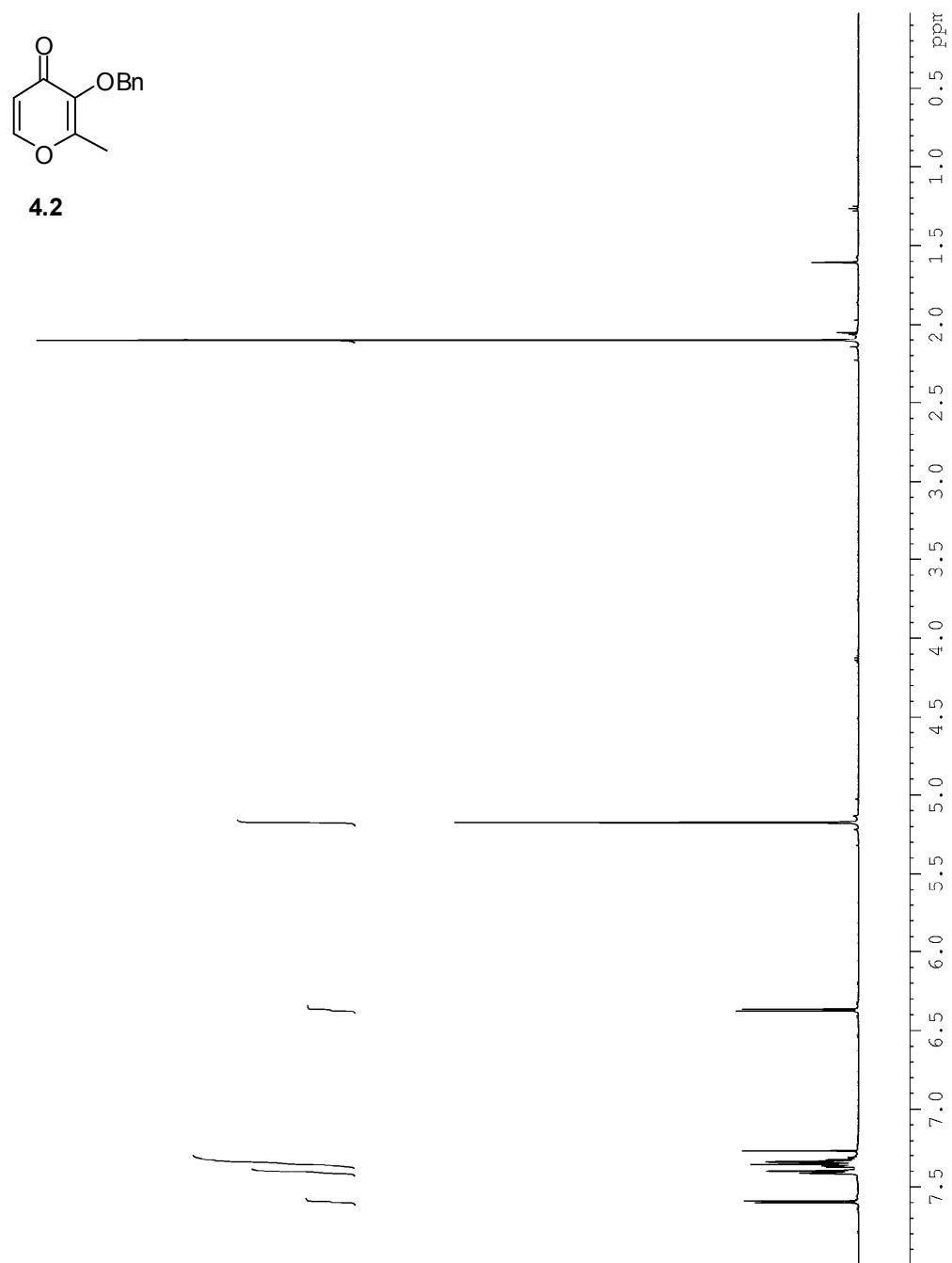


Figure A4.1.1. ^1H NMR spectrum of compound **4.2** (500 MHz, CDCl_3).

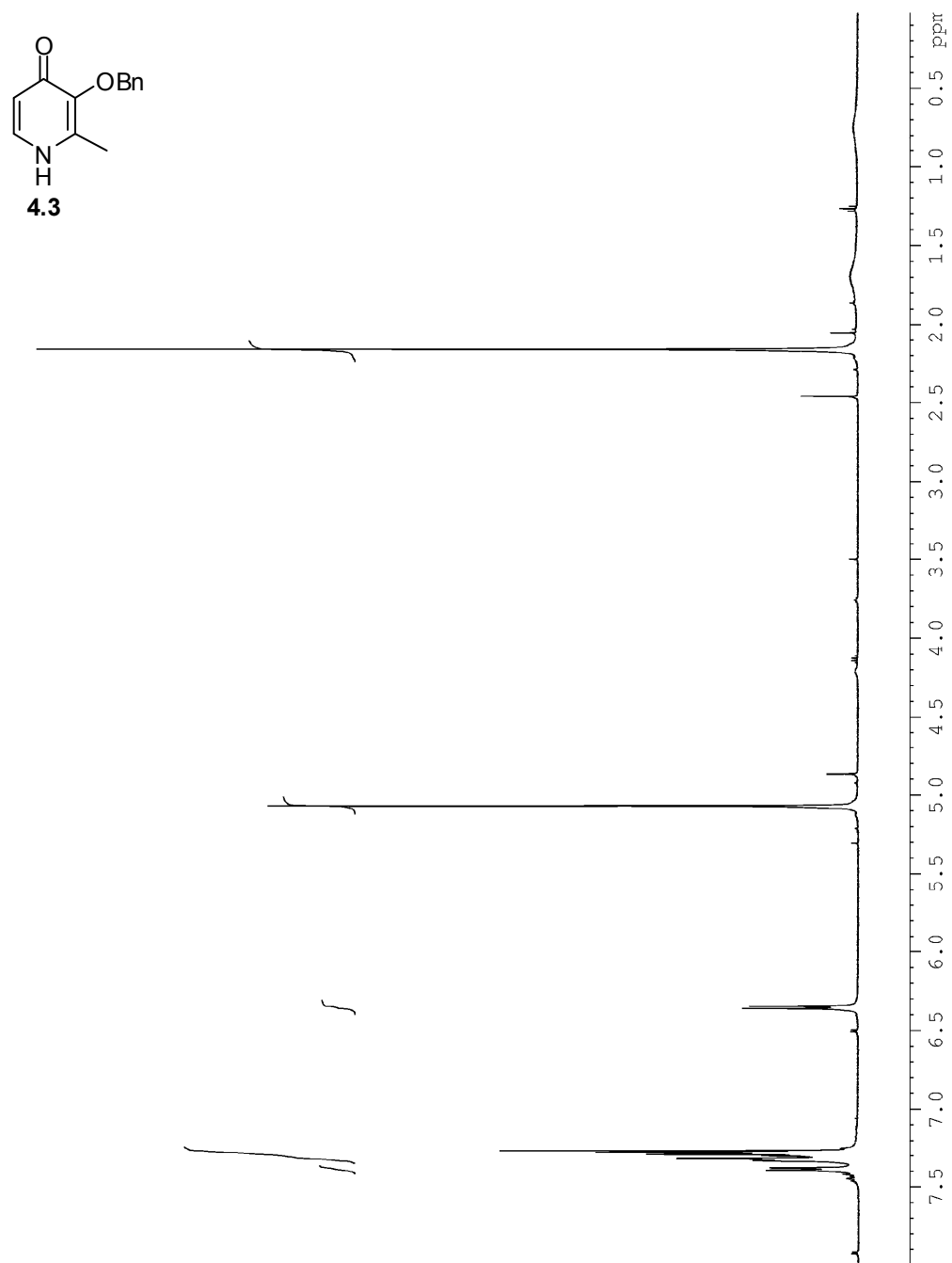


Figure A4.2.1. ¹H NMR spectrum of compound **4.3** (500 MHz, CDCl₃).

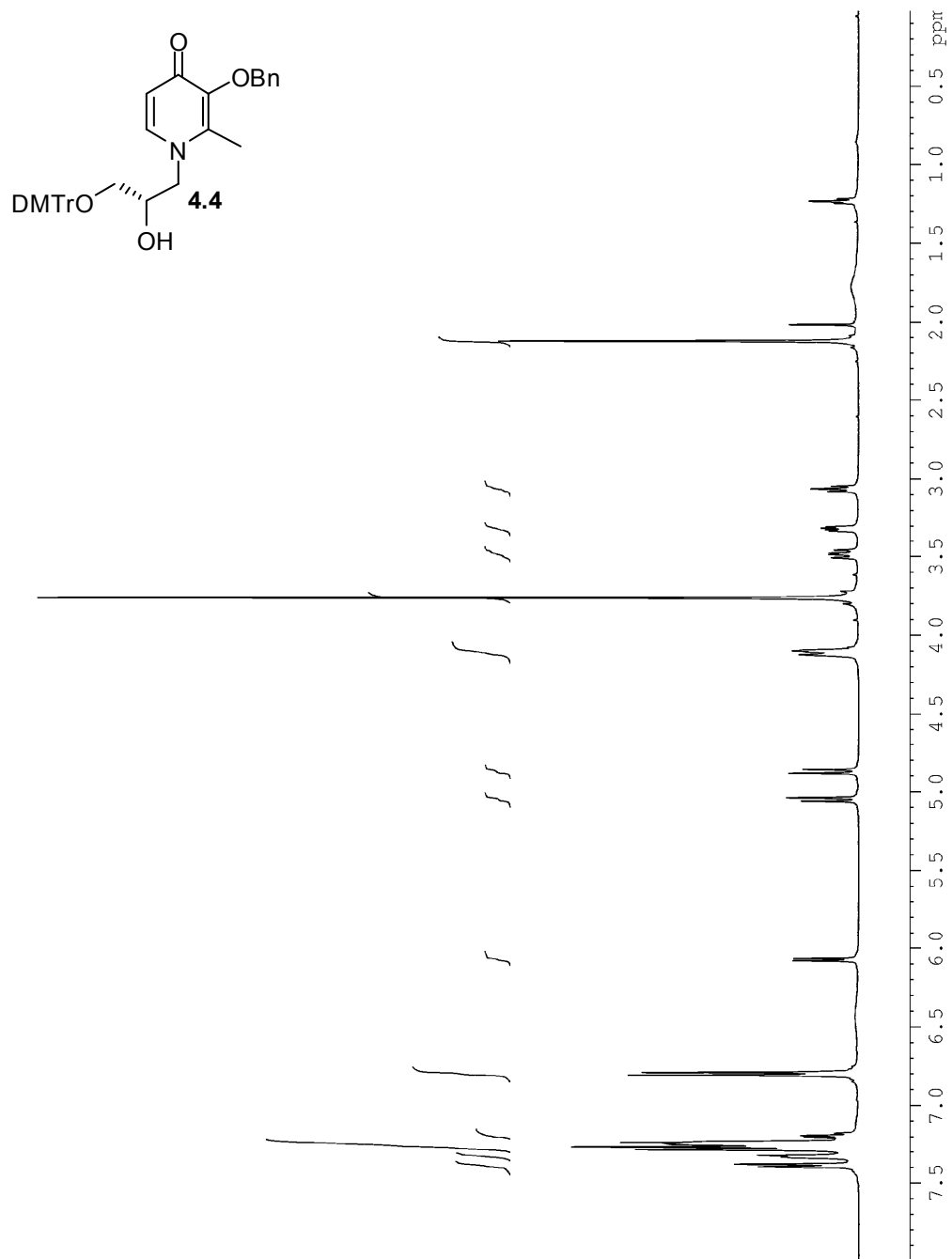


Figure A4.3.1. ^1H NMR spectrum of compound **4.4** (500 MHz, CDCl_3).

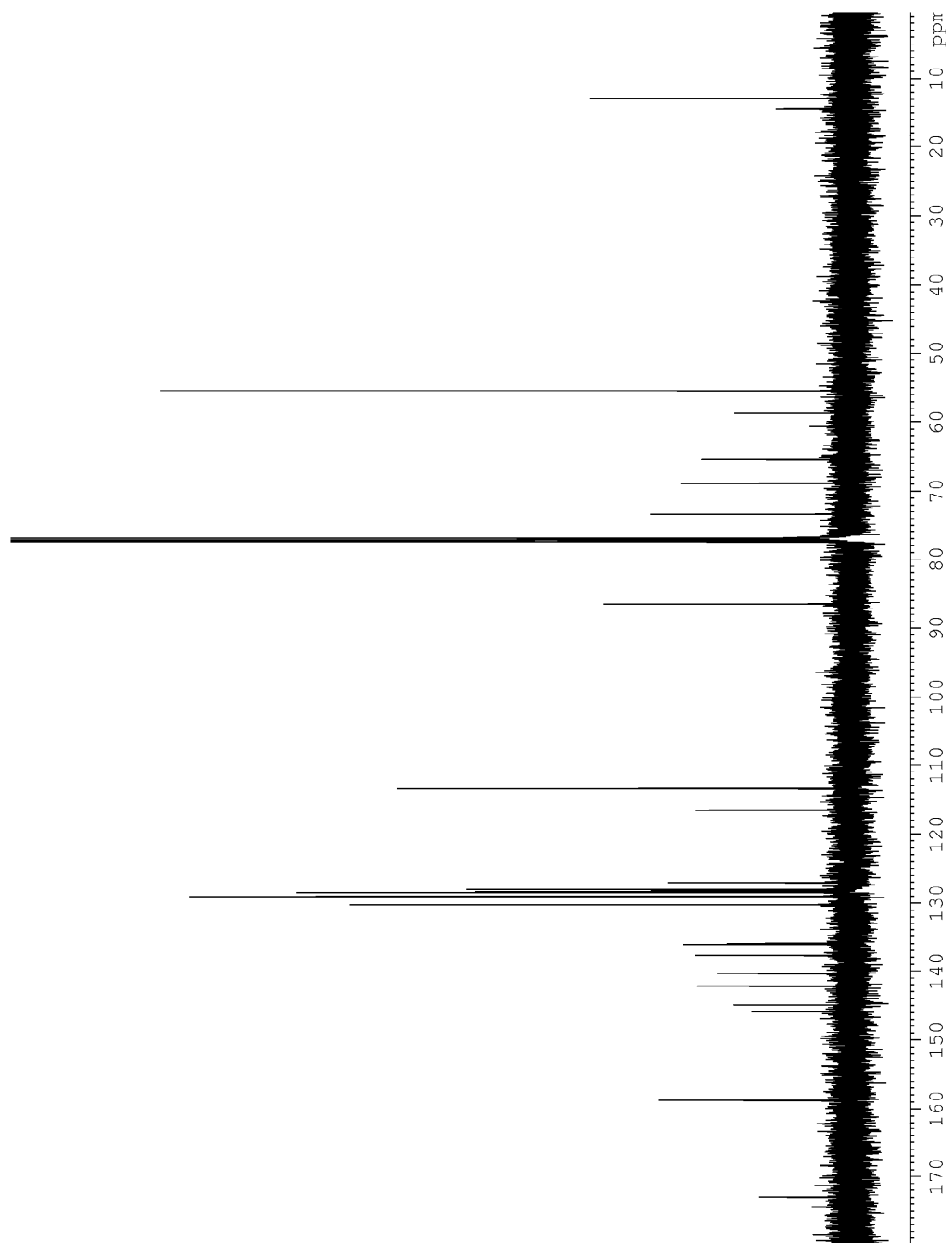


Figure A4.3.2. ^{13}C NMR spectrum of compound **4.4** (125 MHz, CDCl_3).

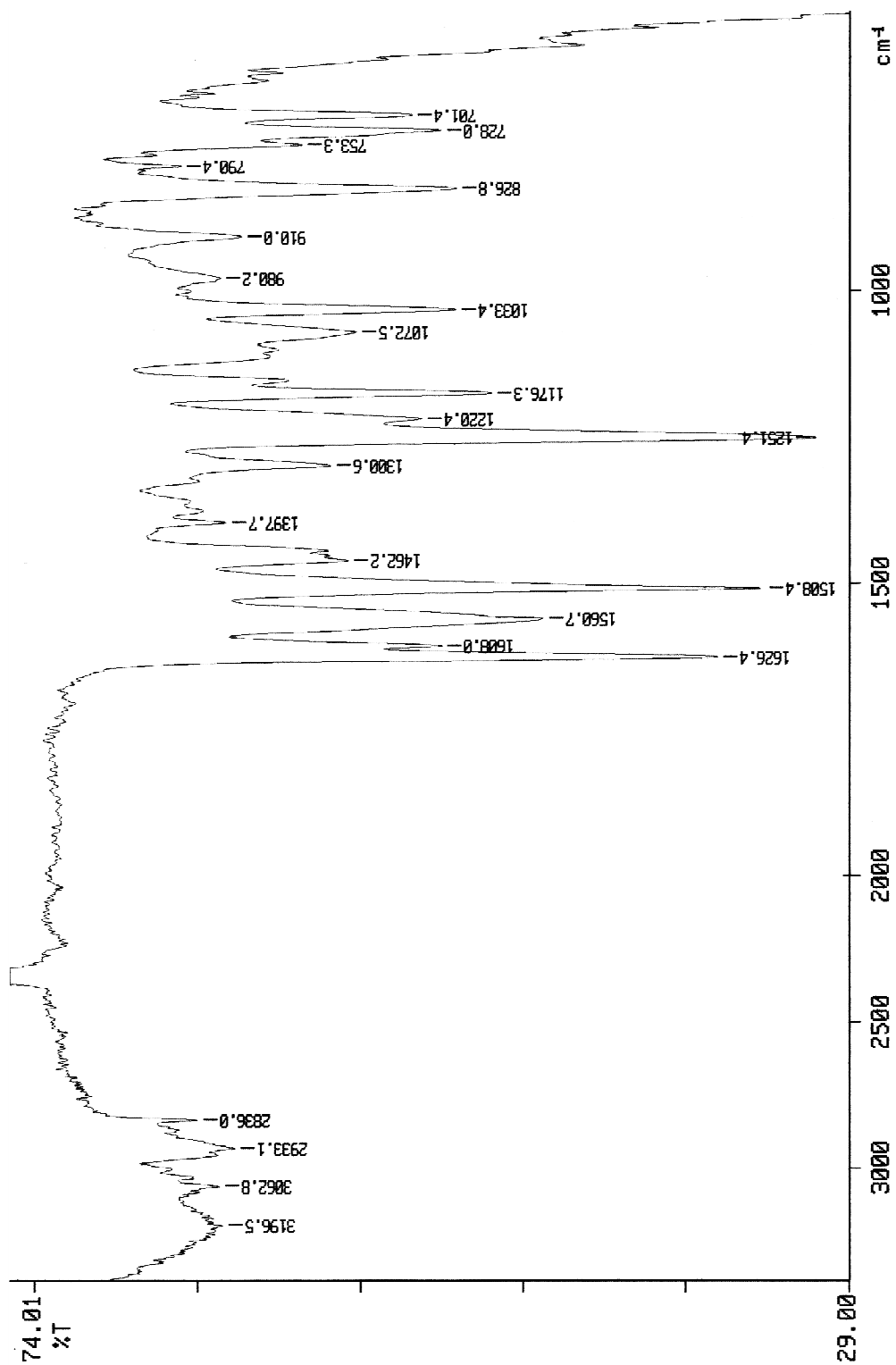


Figure A4.3.3. IR spectrum of compound 4.4 (film).

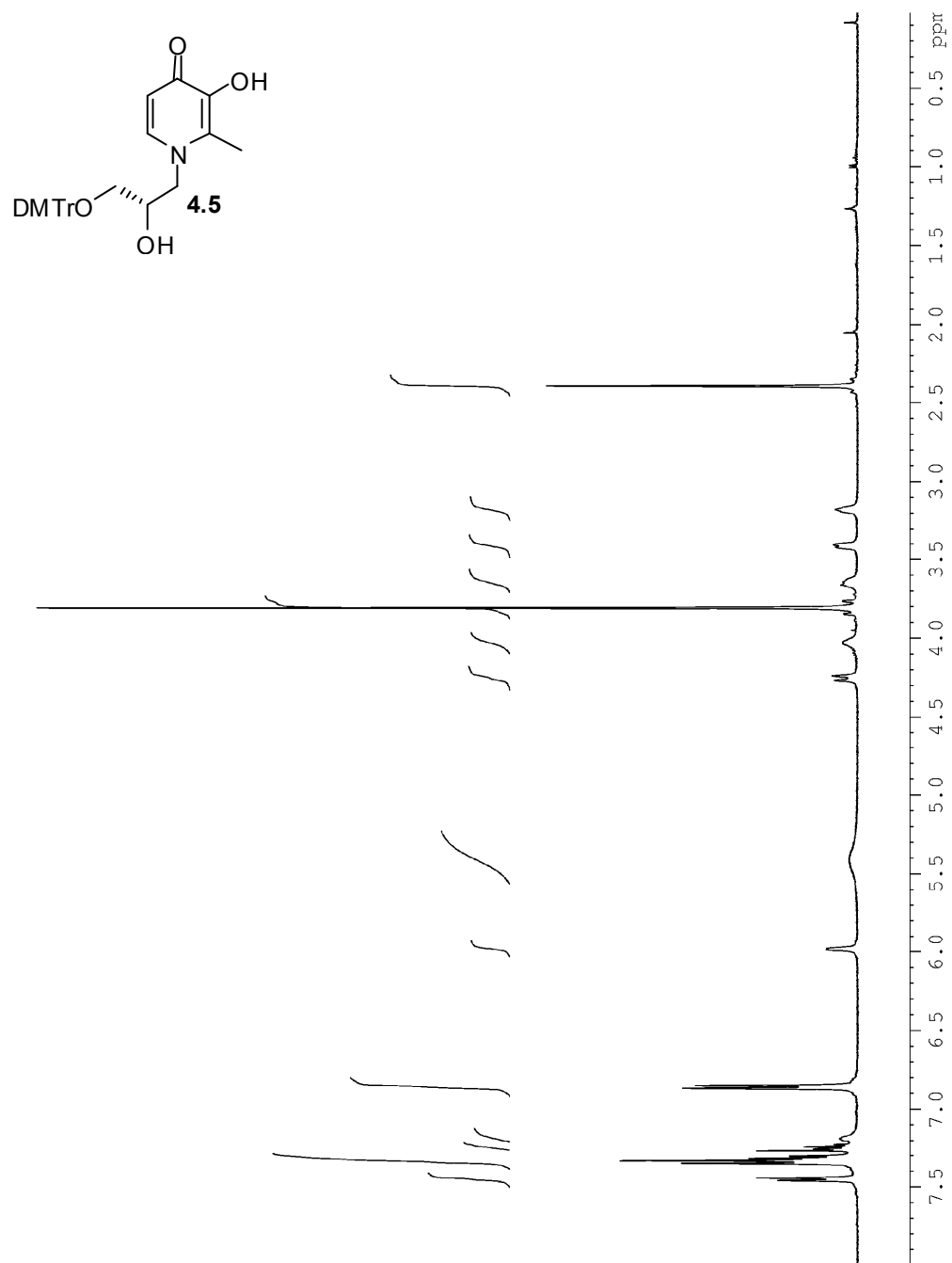


Figure A4.4.1. ¹H NMR spectrum of compound 4.5 (500 MHz, CDCl₃).

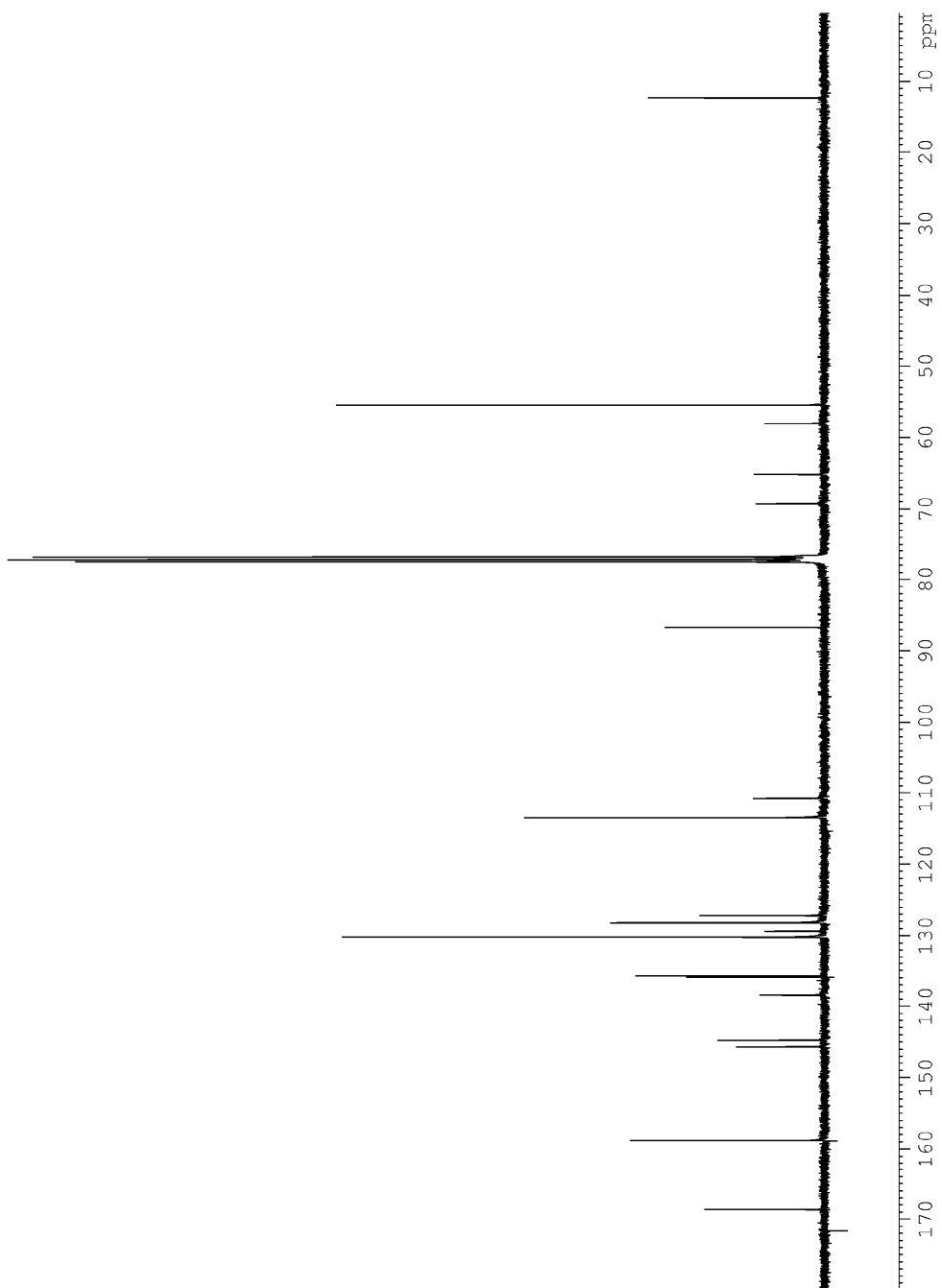


Figure A4.4.2. ^{13}C NMR spectrum of compound 4.5 (90 MHz, CDCl_3).

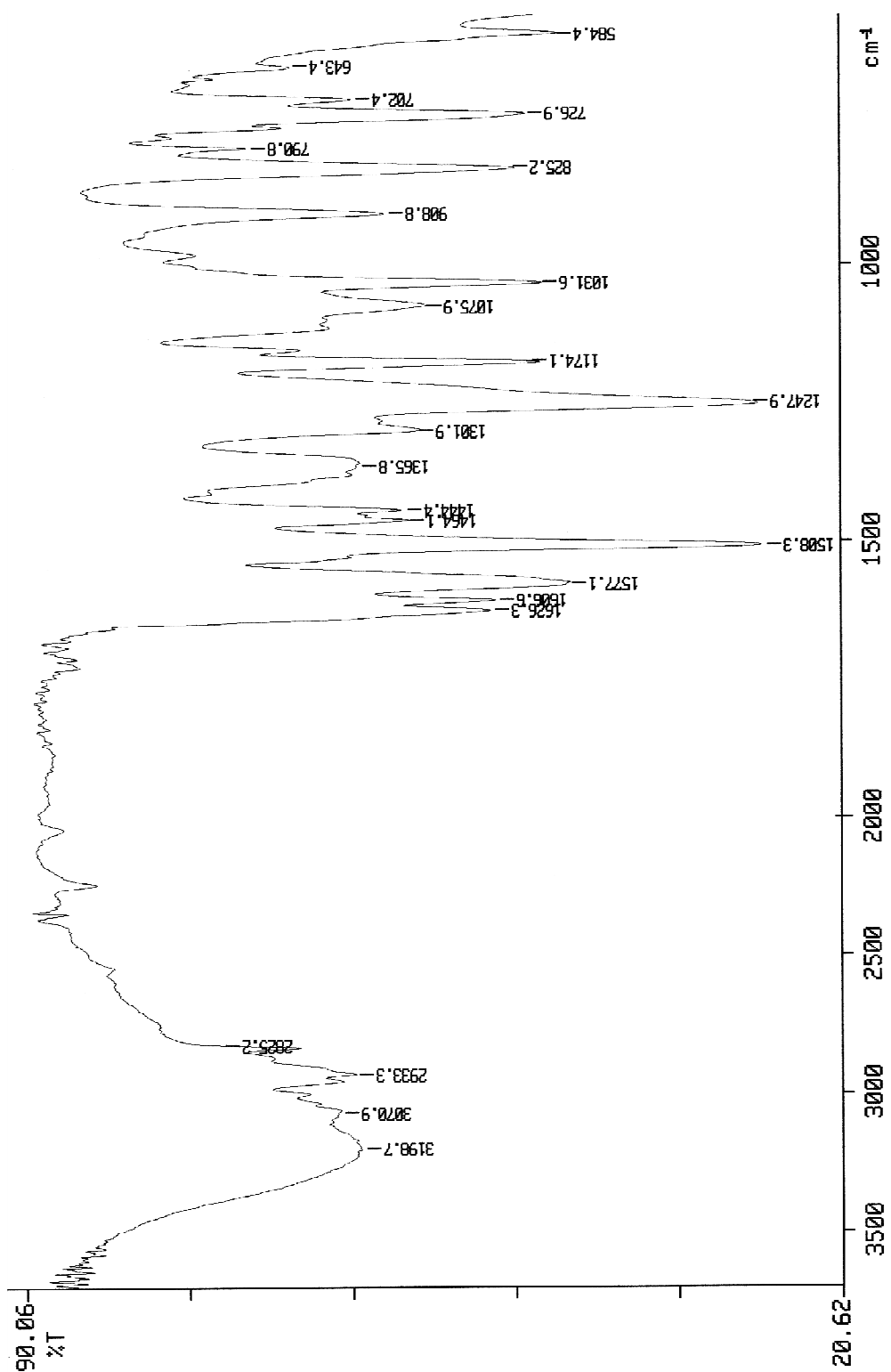


Figure A4.4.3. IR spectrum of compound 4.5 (film).

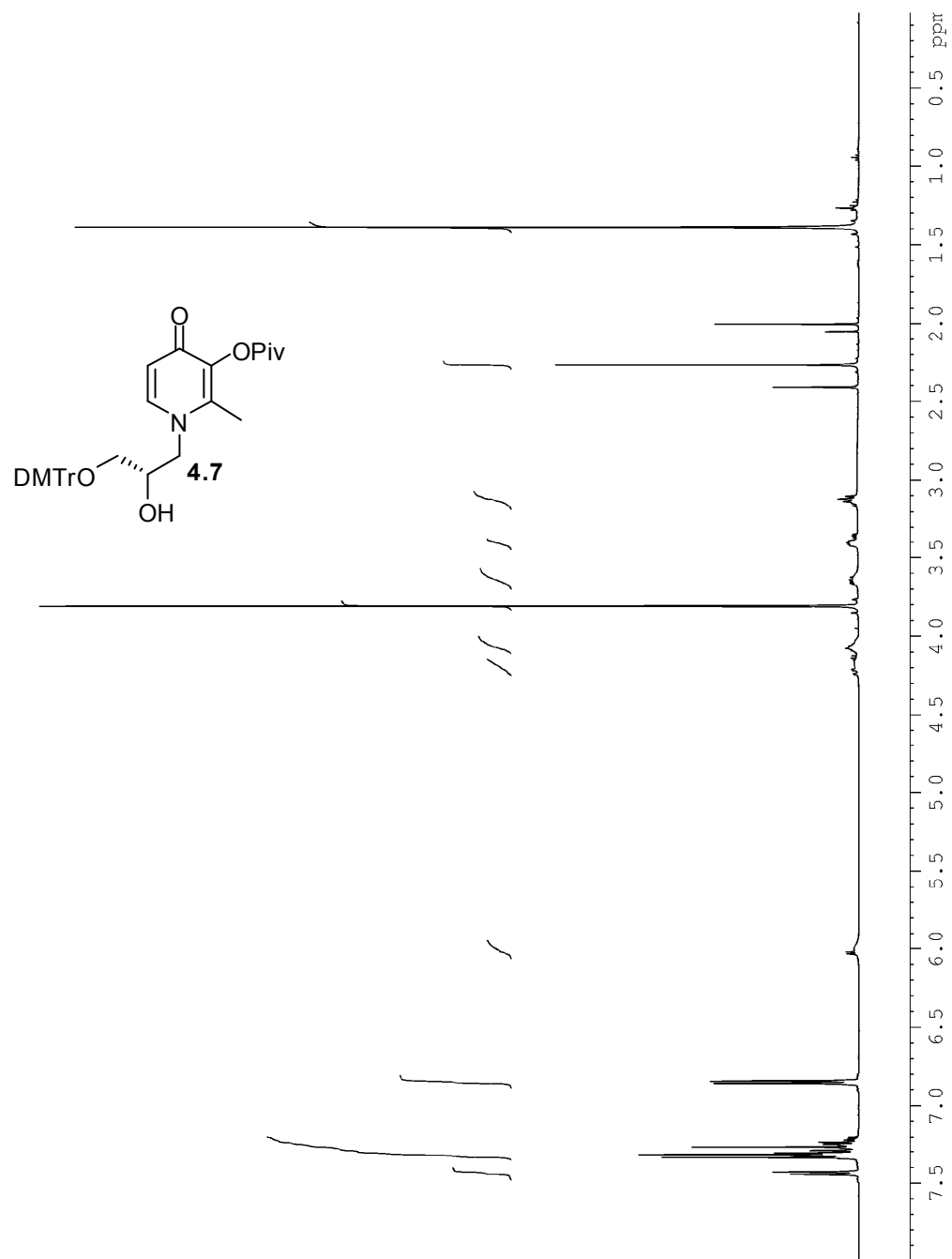


Figure A4.5.1. ^1H NMR spectrum of compound **4.7** (500 MHz, CDCl_3).

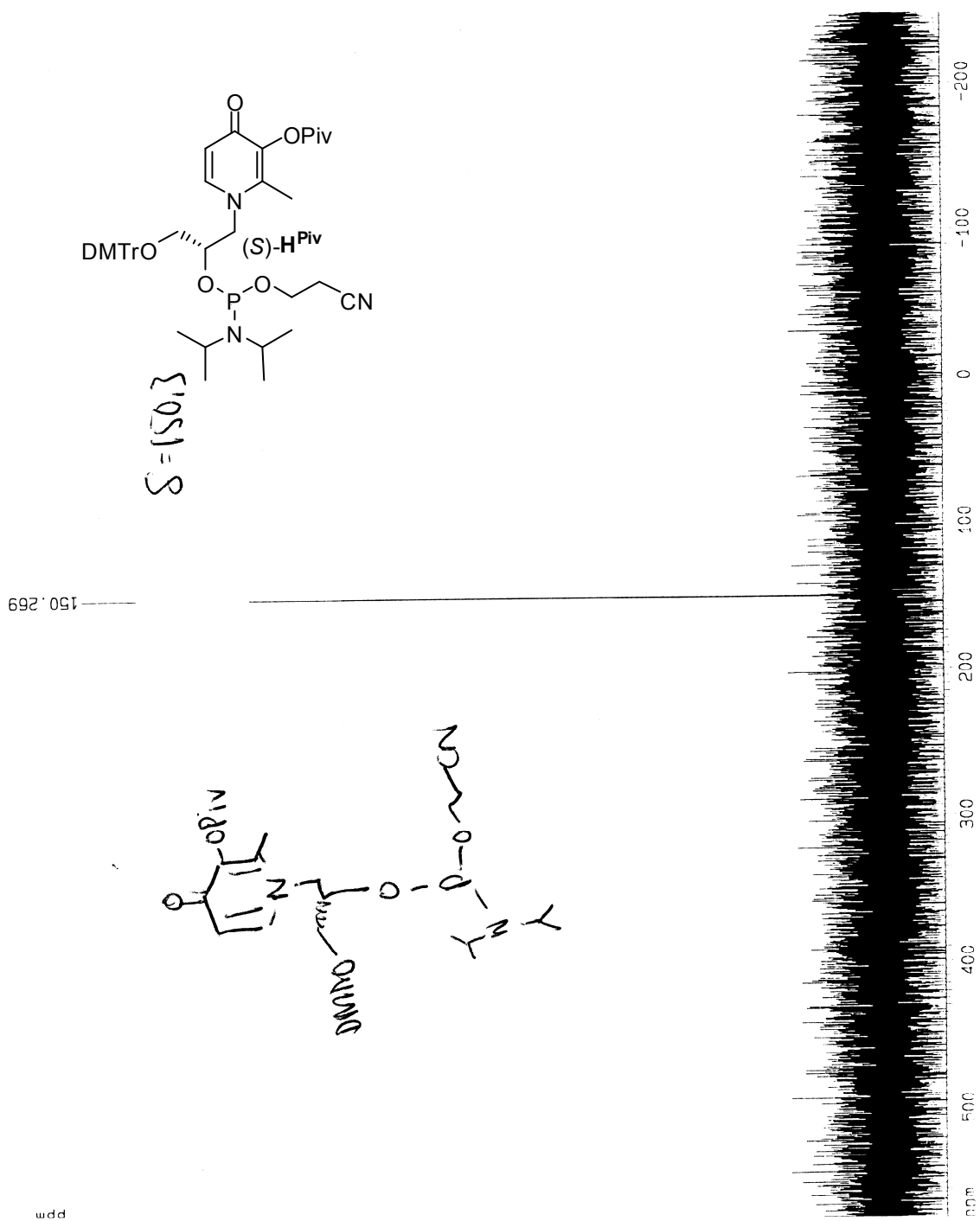


Figure A4.6.1. ^{31}P NMR spectrum of phosphoramidite (*S*)-H^{Piv} (121 MHz, CDCl₃).

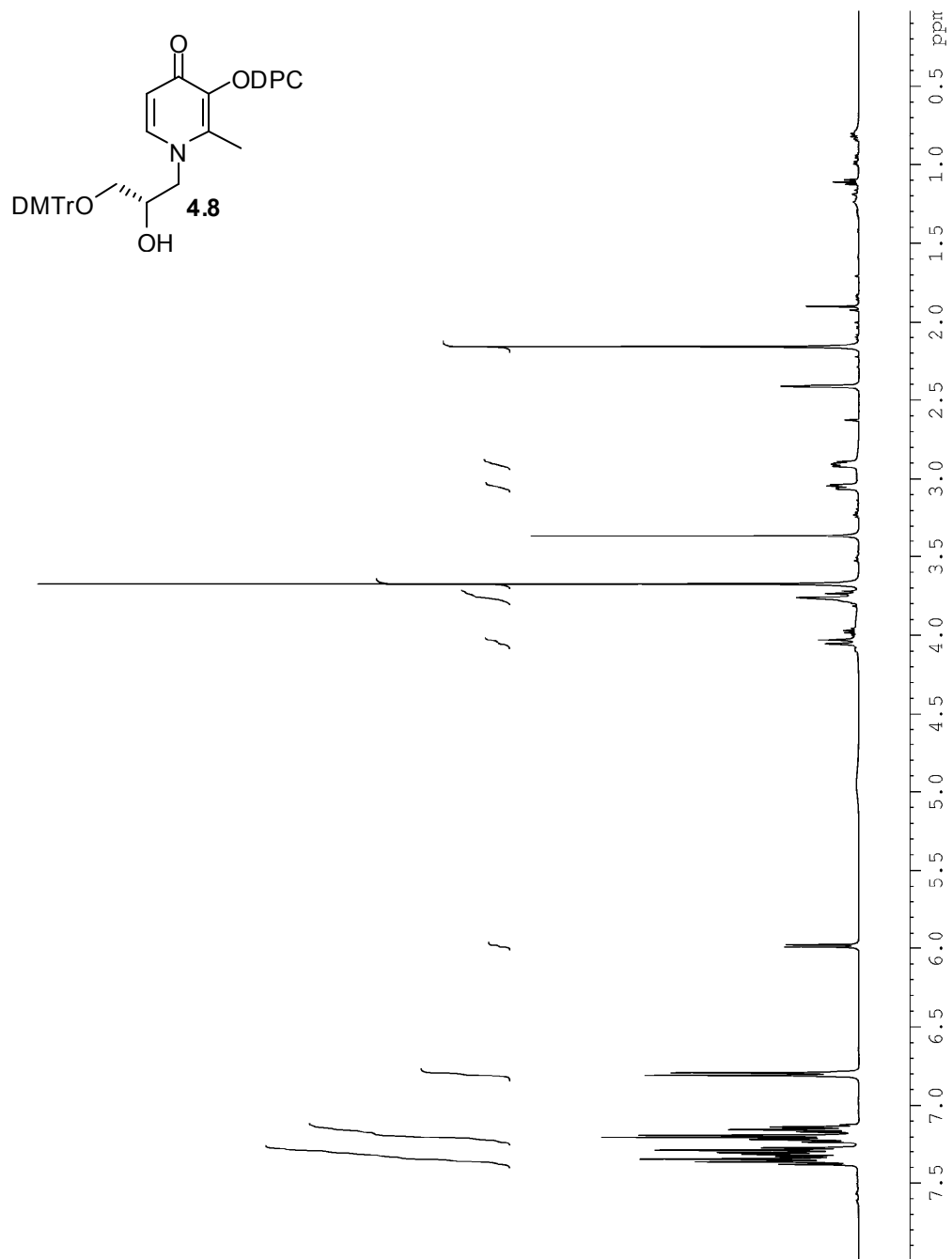


Figure A4.7.1. ^1H NMR spectrum of compound **4.8** (500 MHz, 373K, $\text{DMSO-}d_6$).

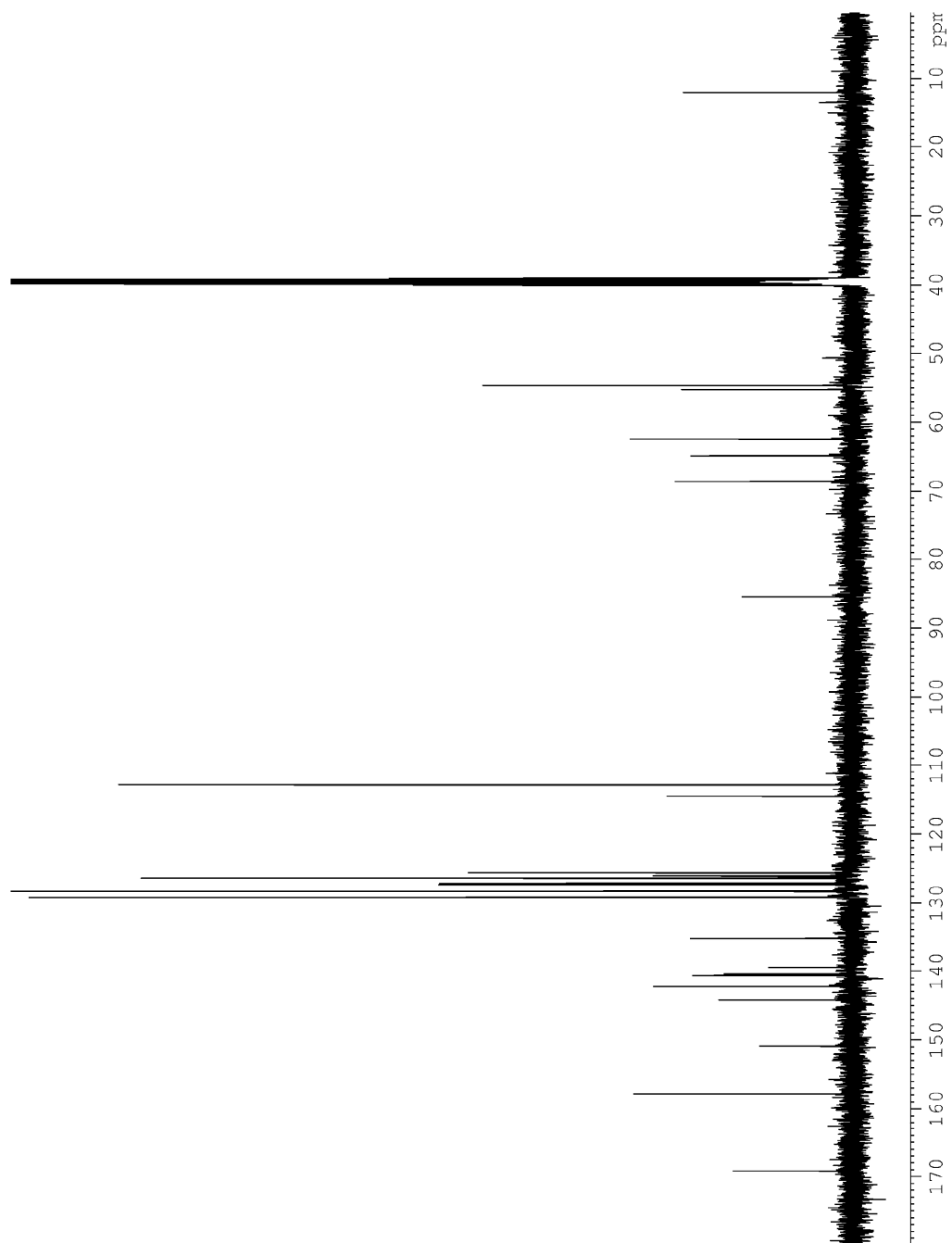


Figure A4.7.2. ^{13}C NMR spectrum of compound **4.8** (125 MHz, 373K, $\text{DMSO-}d_6$).

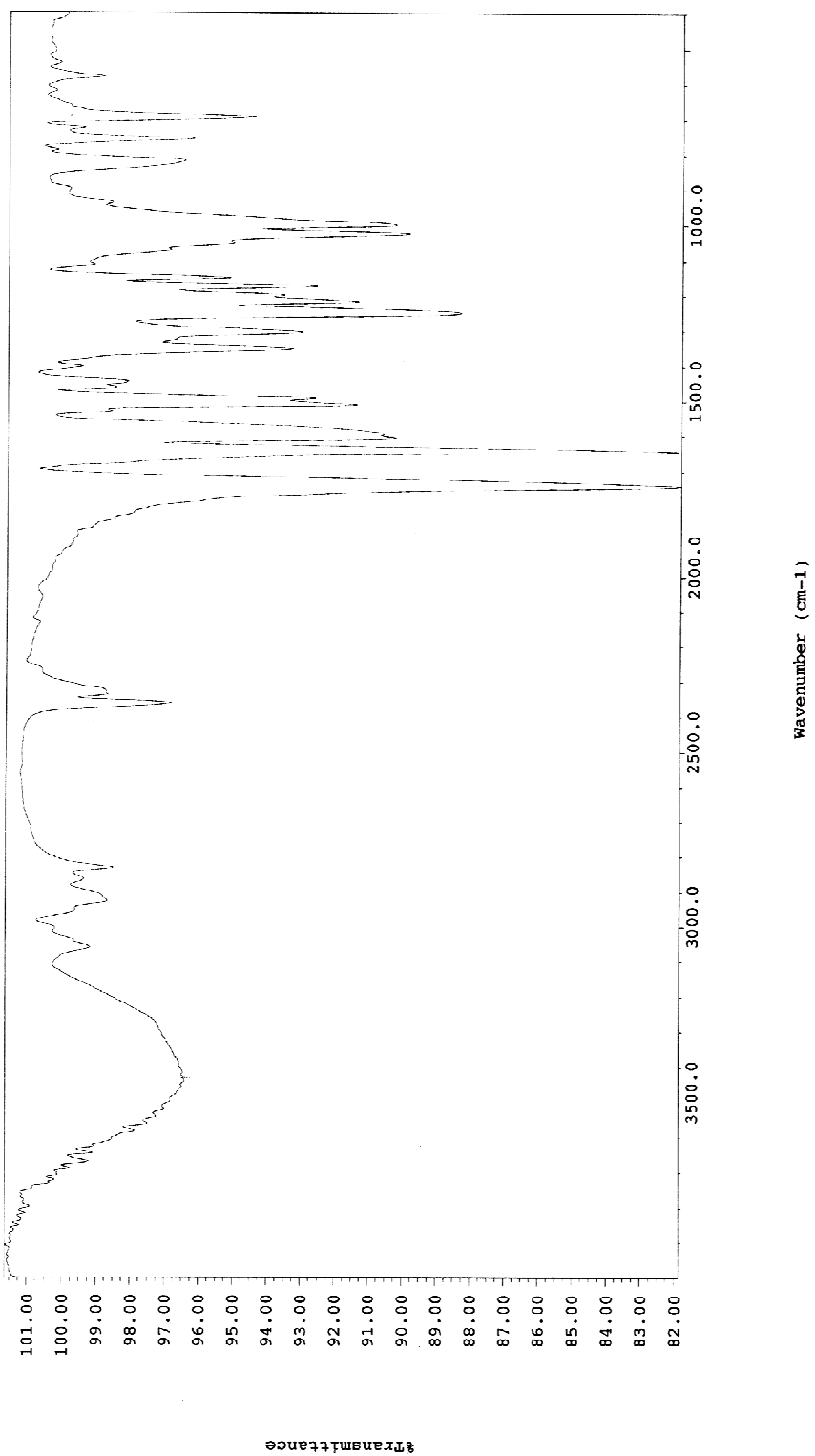


Figure A4.7.3. IR spectrum of compound 4.8 (film).

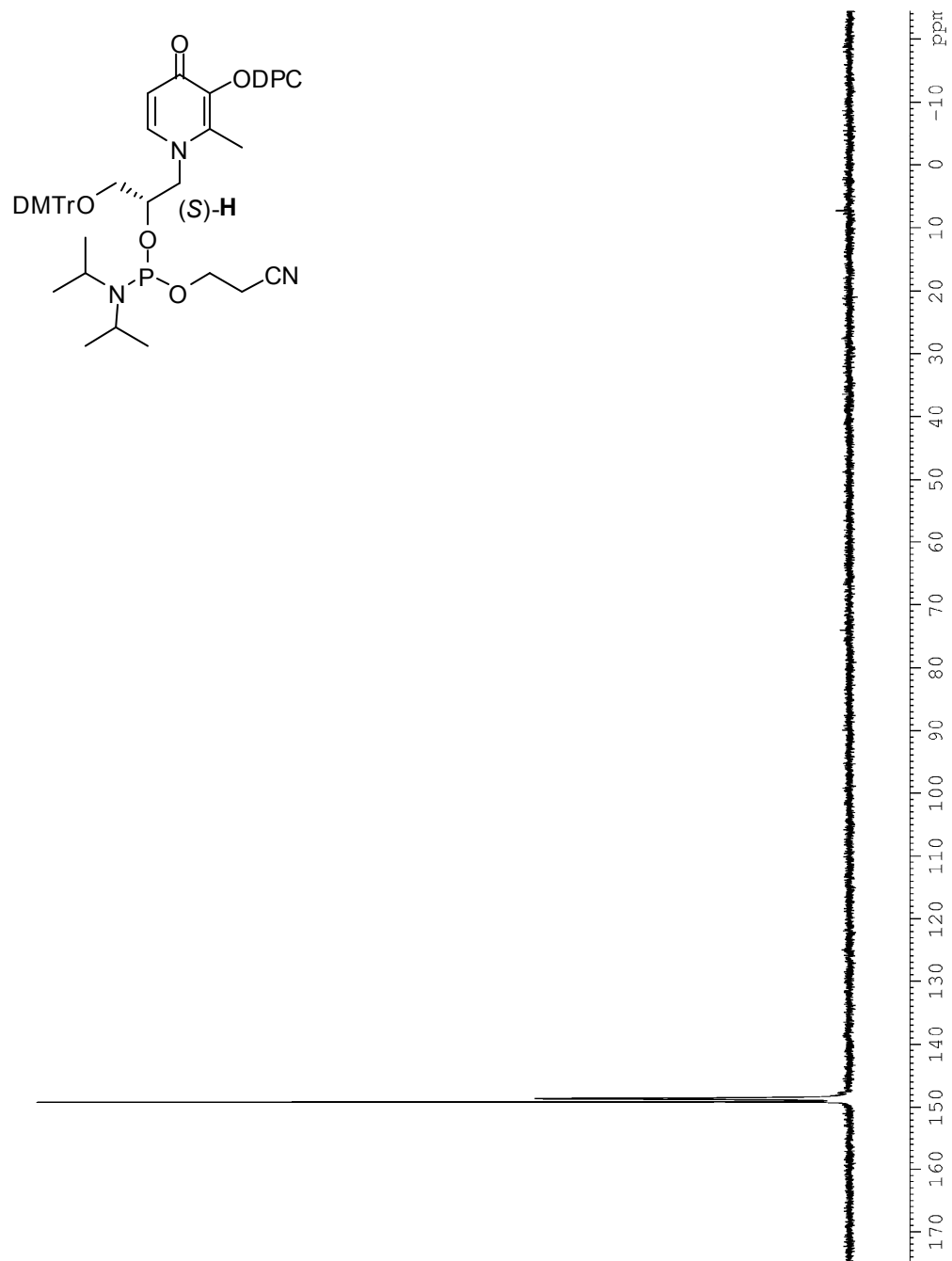


Figure A4.8.1. ³¹P NMR spectrum of phosphoramidite (S)-H^{DPC} (121 MHz, CDCl₃).

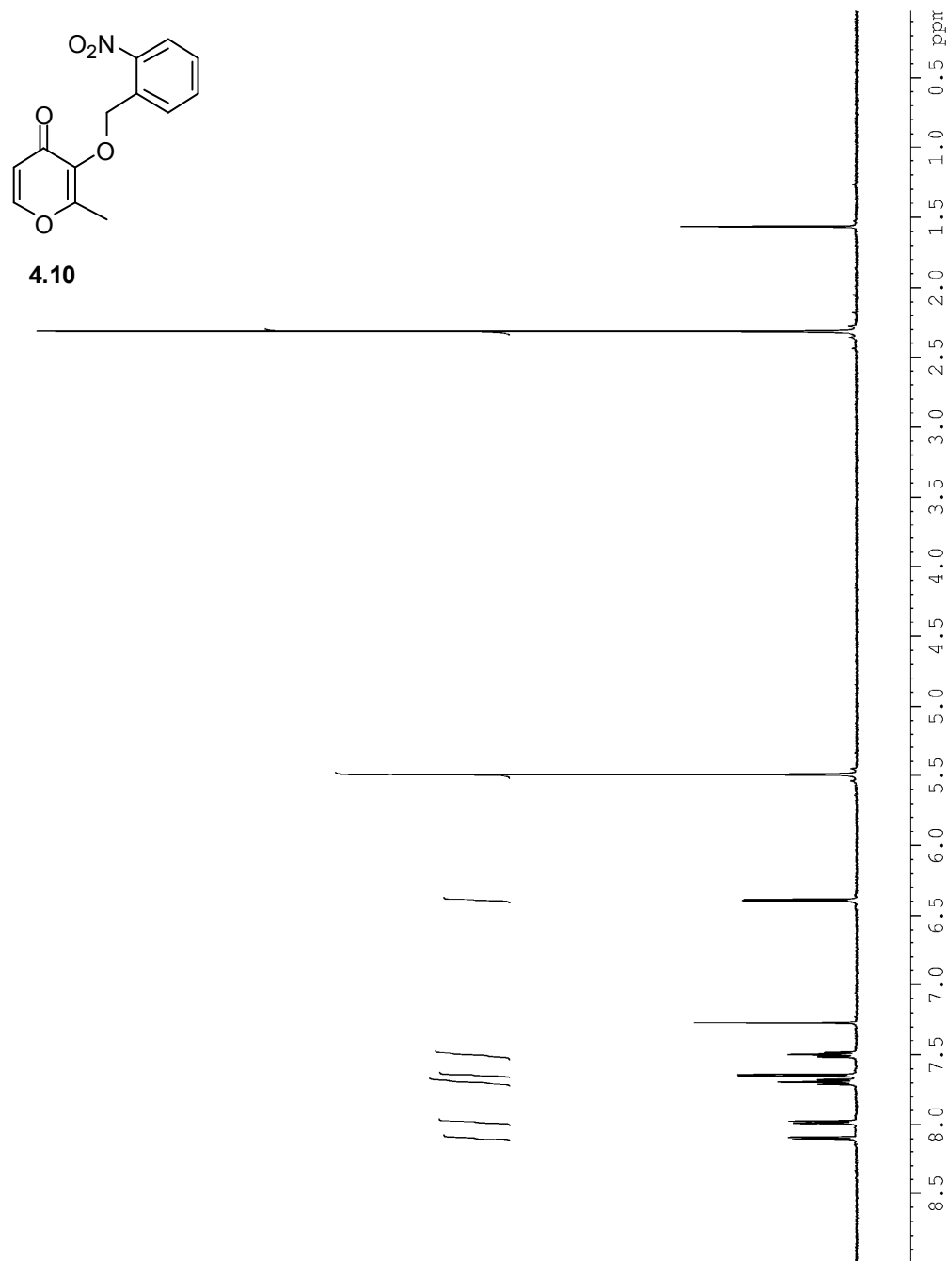


Figure A4.9.1. ^1H NMR spectrum of compound **4.10** (500 MHz, CDCl_3).



Figure A4.10.1. ¹H NMR spectrum of compound **4.11** (360 MHz, MeOD).

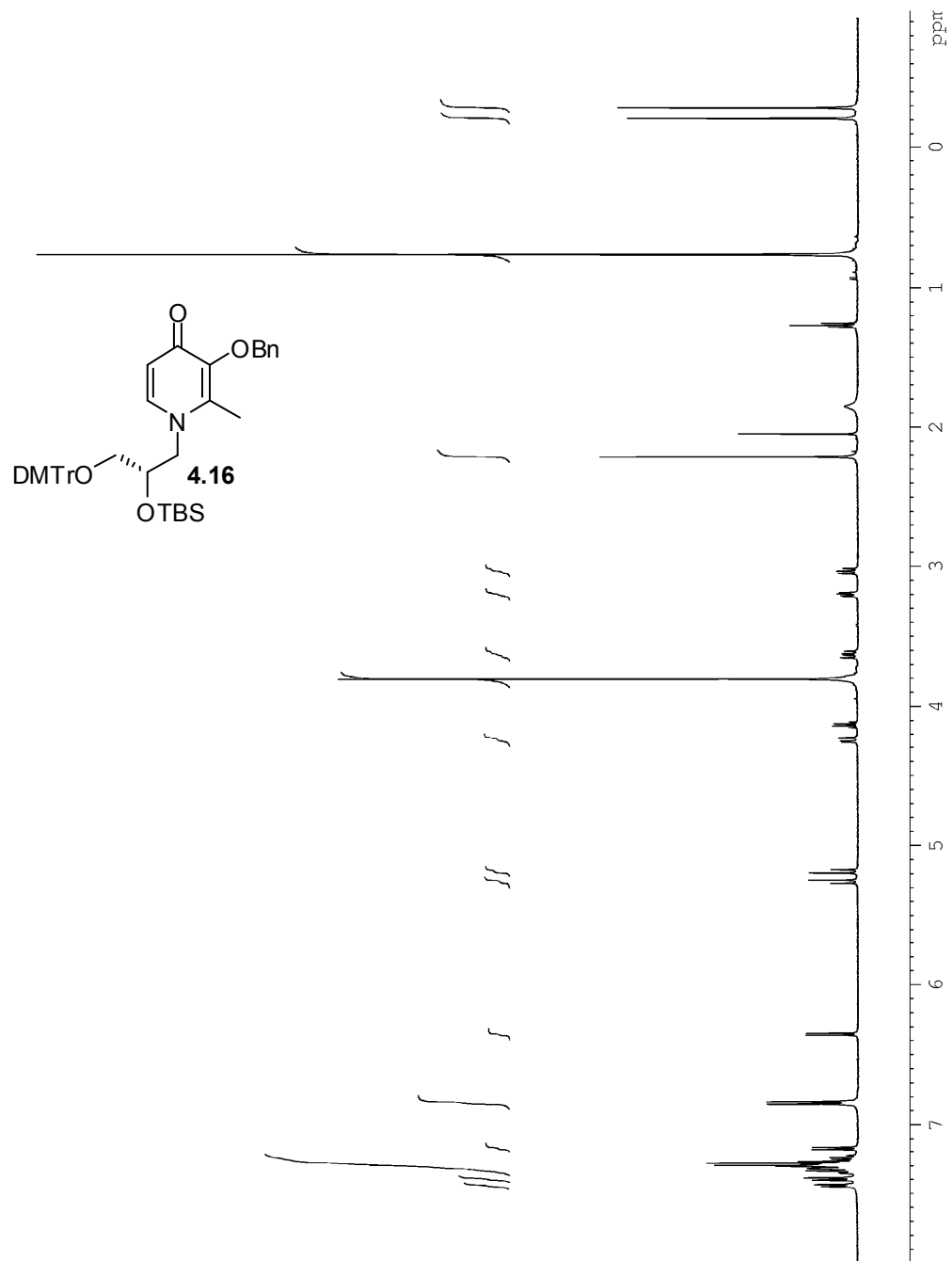


Figure A4.11.1. ^1H NMR spectrum of compound **4.16** (500 MHz, CDCl_3).

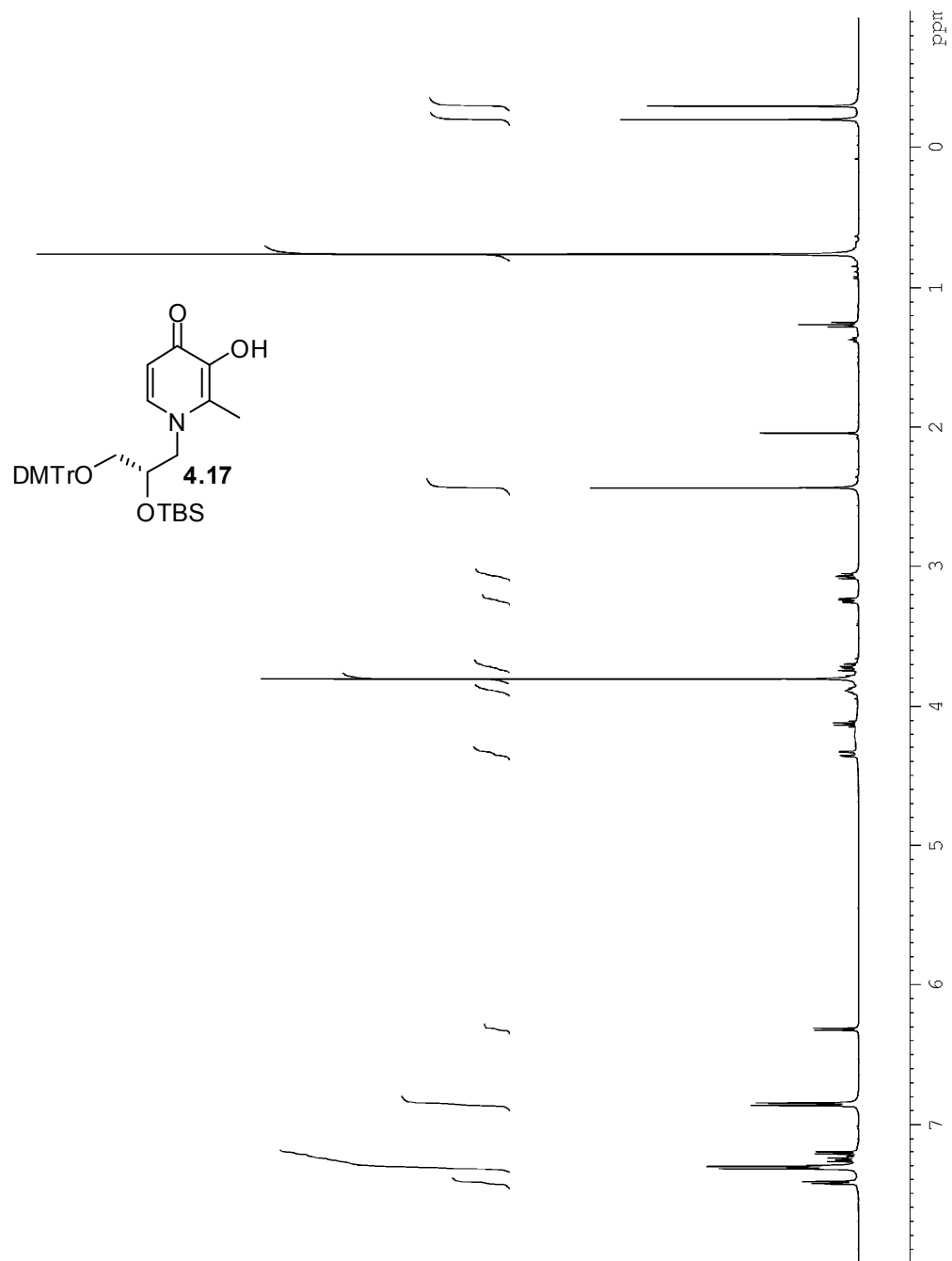


Figure A4.12.1. ^1H NMR spectrum of compound **4.17** (500 MHz, CDCl_3).

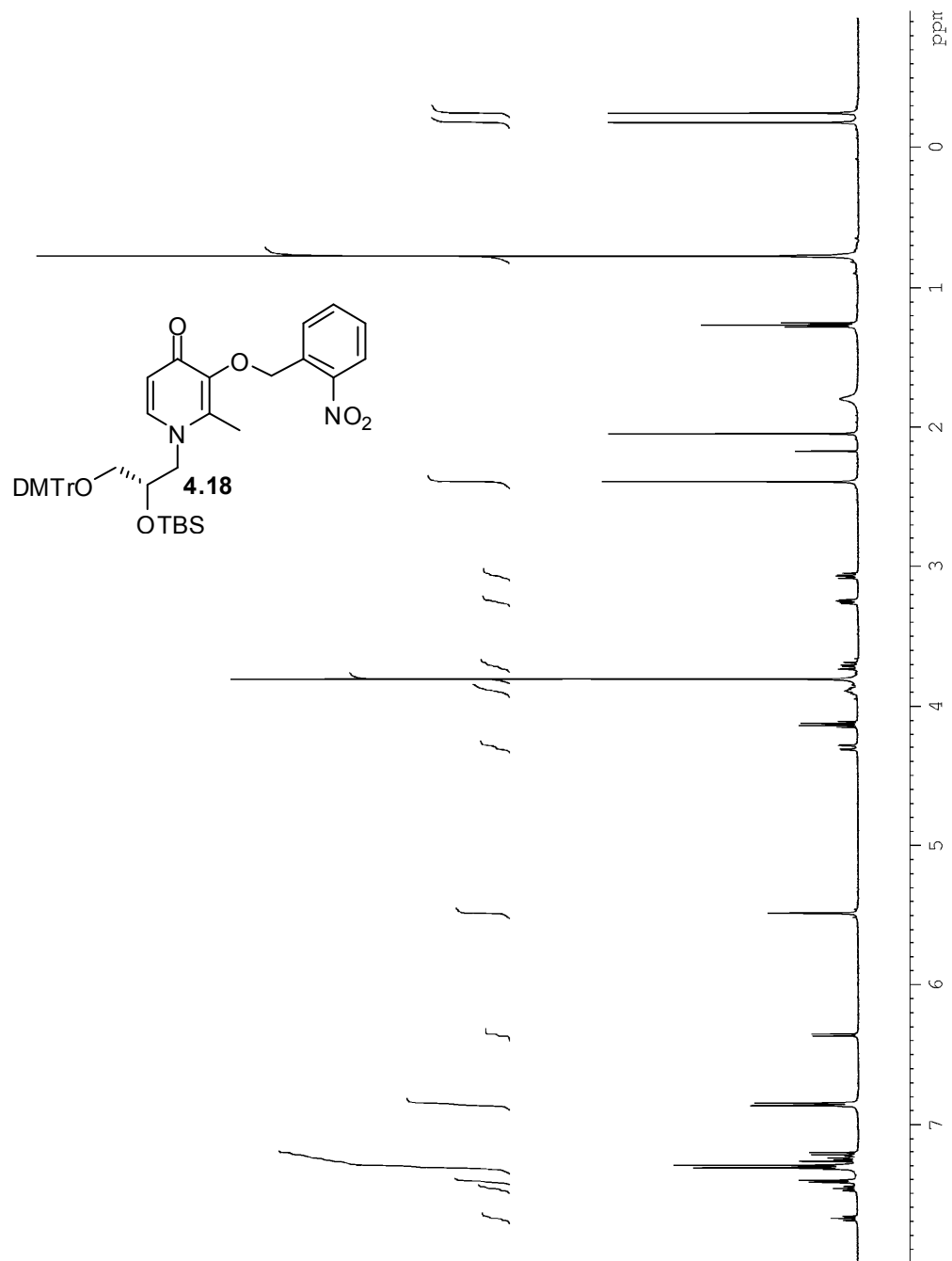


Figure A4.13.1. ^1H NMR spectrum of compound **4.18** (500 MHz, CDCl_3).

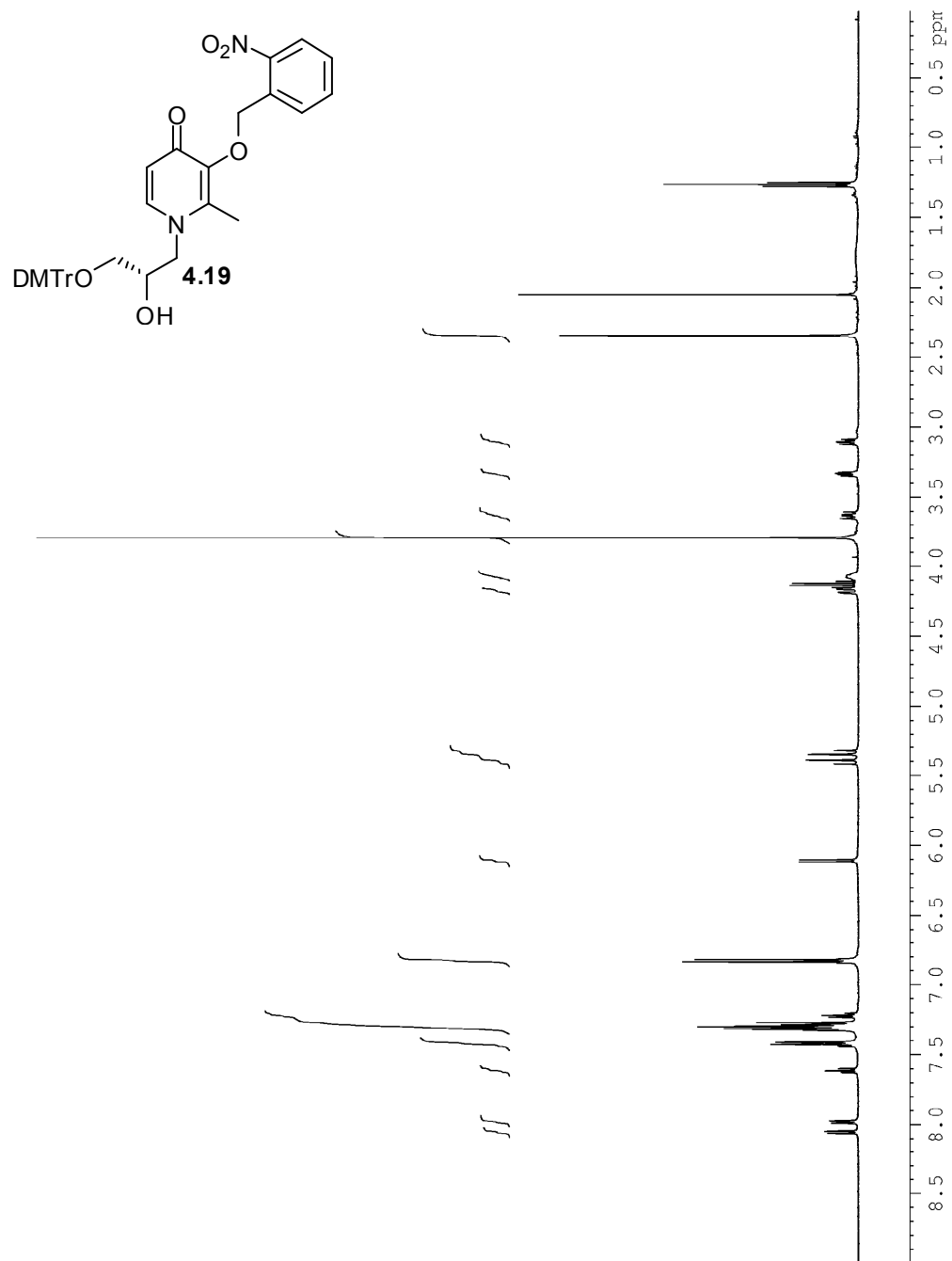


Figure A4.14.1. ¹H NMR spectrum of compound **4.19** (500 MHz, CDCl₃).

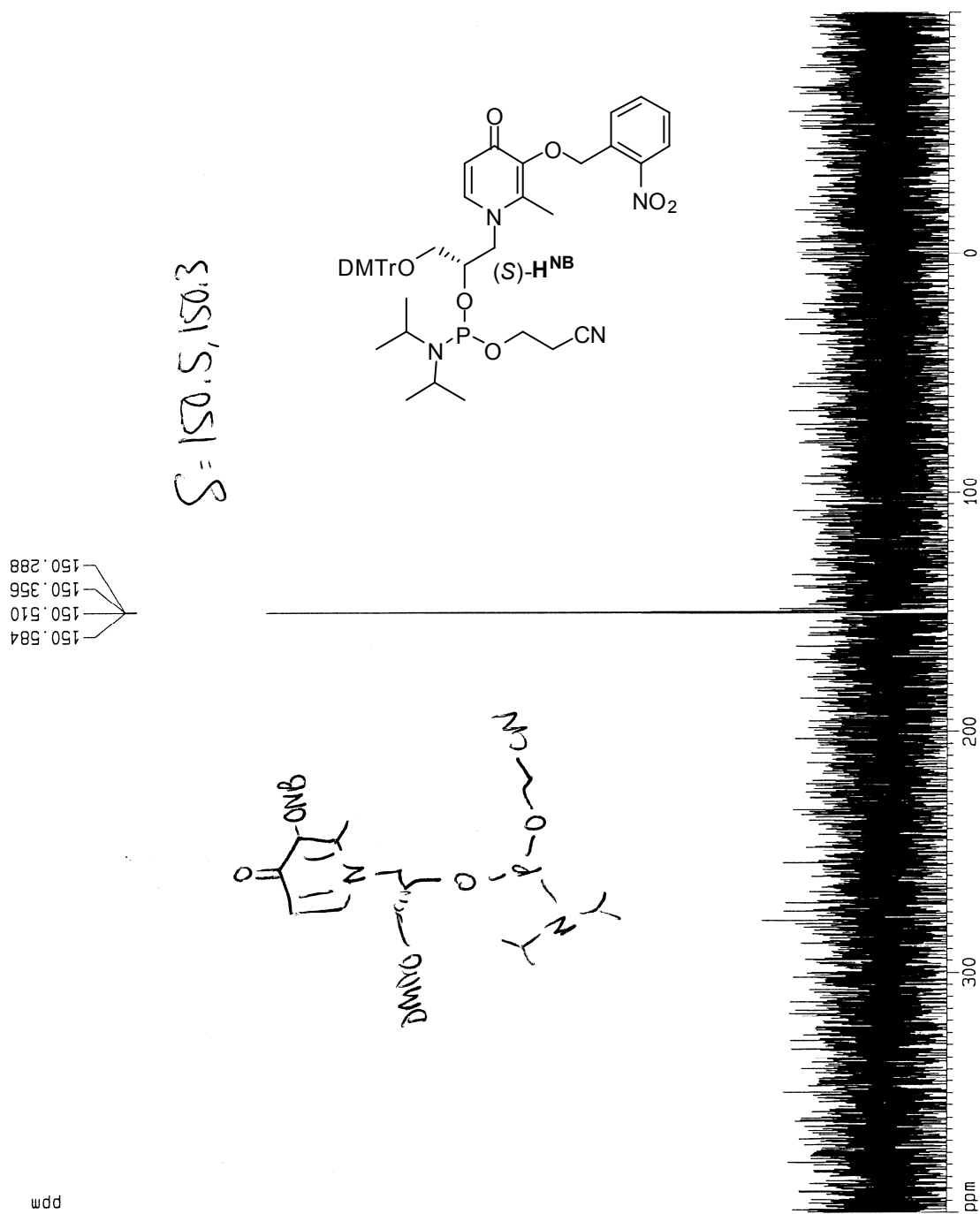


Figure A4.15.1. ^{31}P NMR spectrum of phosphoramidite (*S*)- H^{NB} (121 MHz, CDCl_3).

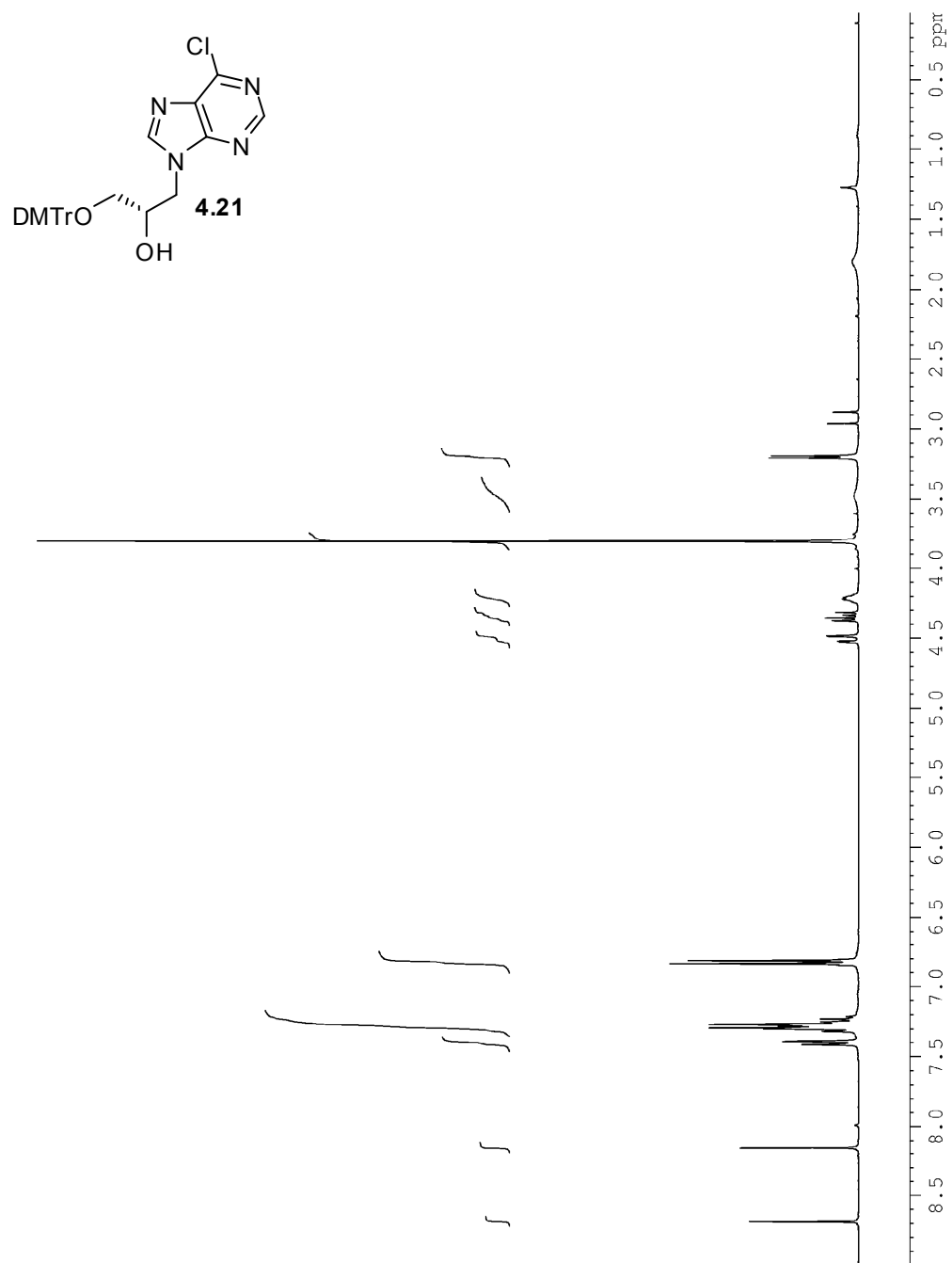


Figure A4.16.1. ^1H NMR spectrum of compound **4.21** (360 MHz, CDCl_3).

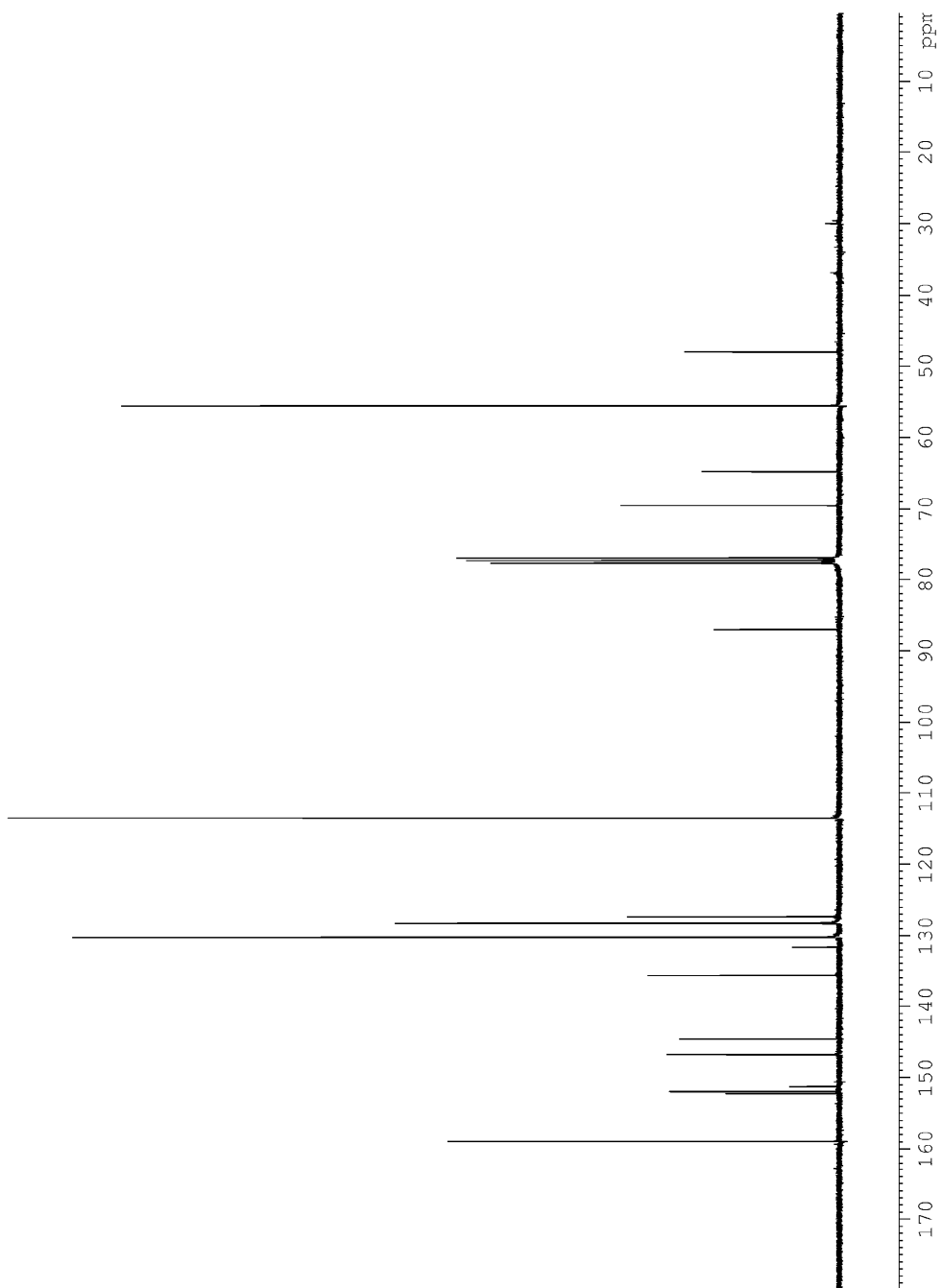


Figure A4.16.2. ^{13}C NMR spectrum of compound **4.21** (90 MHz, CDCl_3).

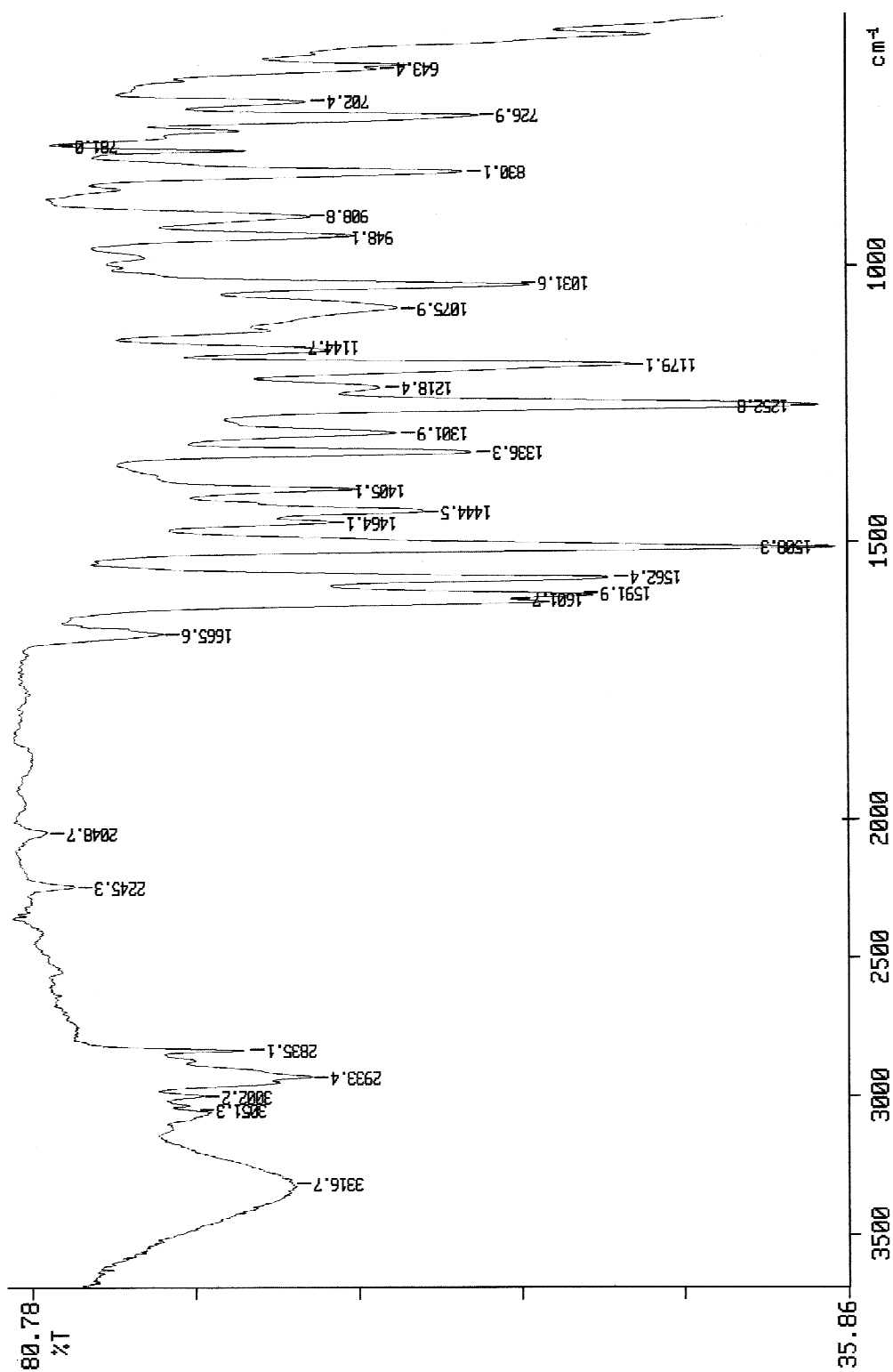


Figure A4.16.3. IR spectrum of compound 4.21 (film).

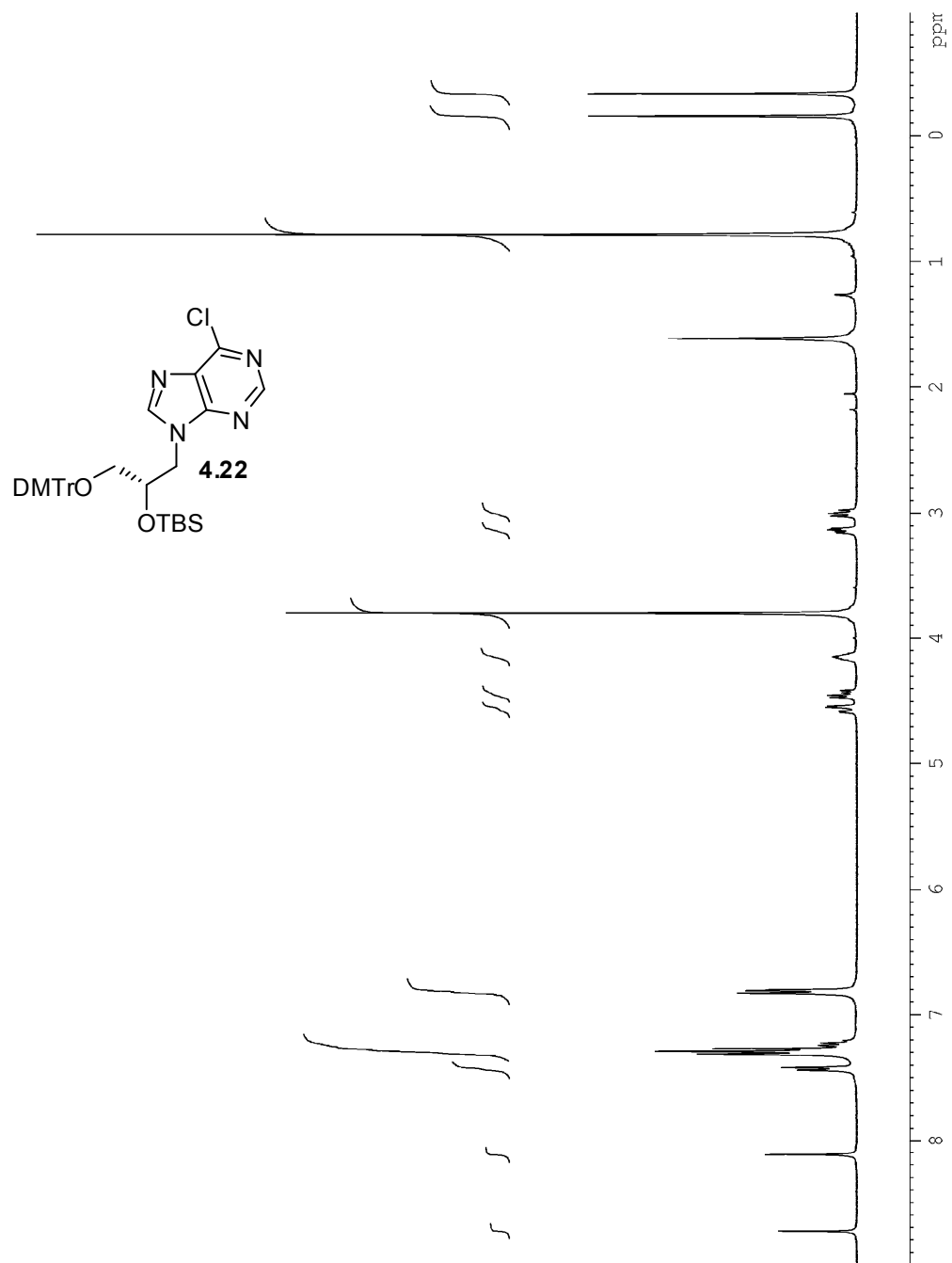


Figure A4.17.1. ^1H NMR spectrum of compound **4.22** (360 MHz, CDCl_3).

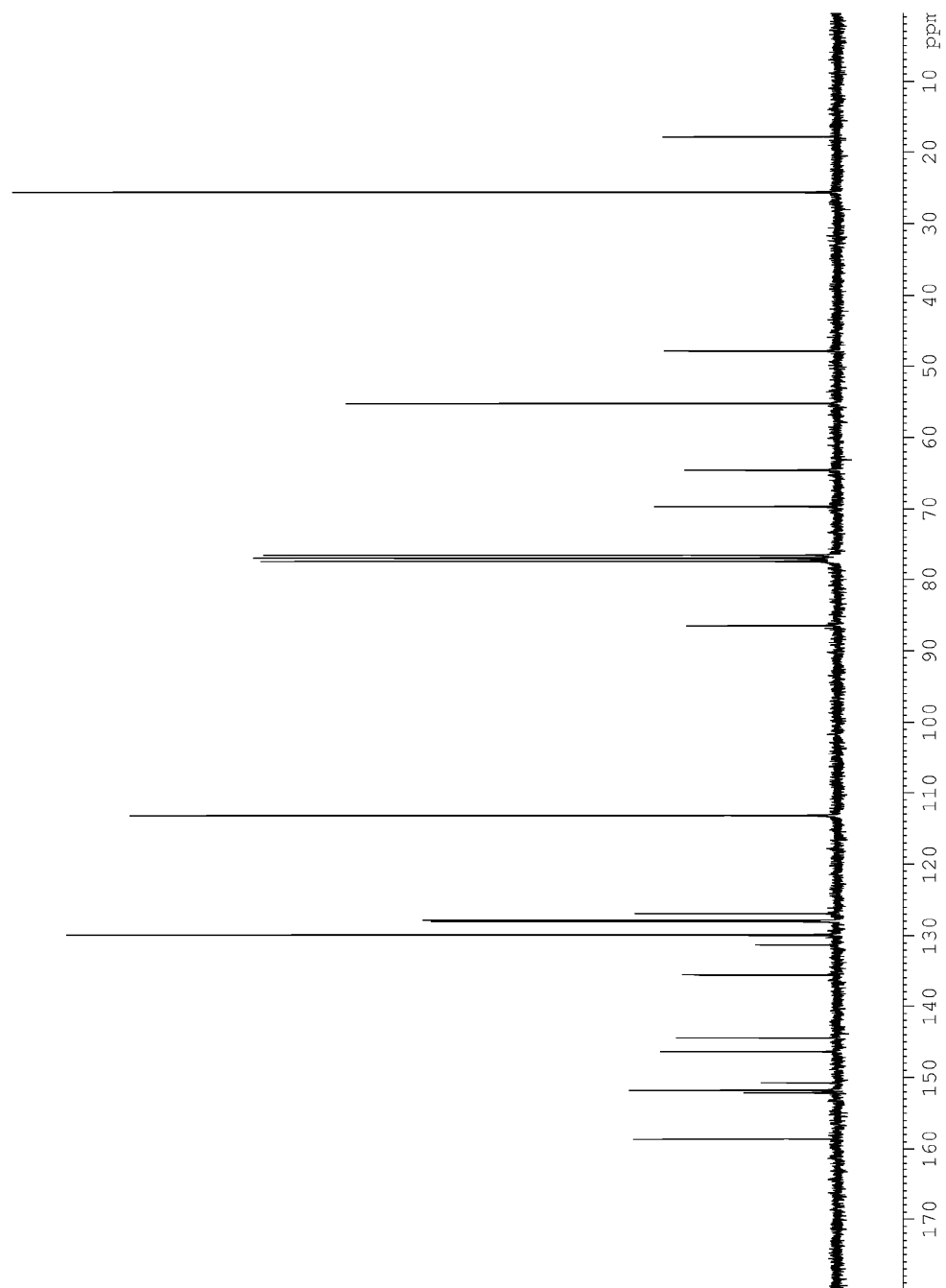


Figure A4.17.2. ^{13}C NMR spectrum of compound **4.22** (75 MHz, CDCl_3).

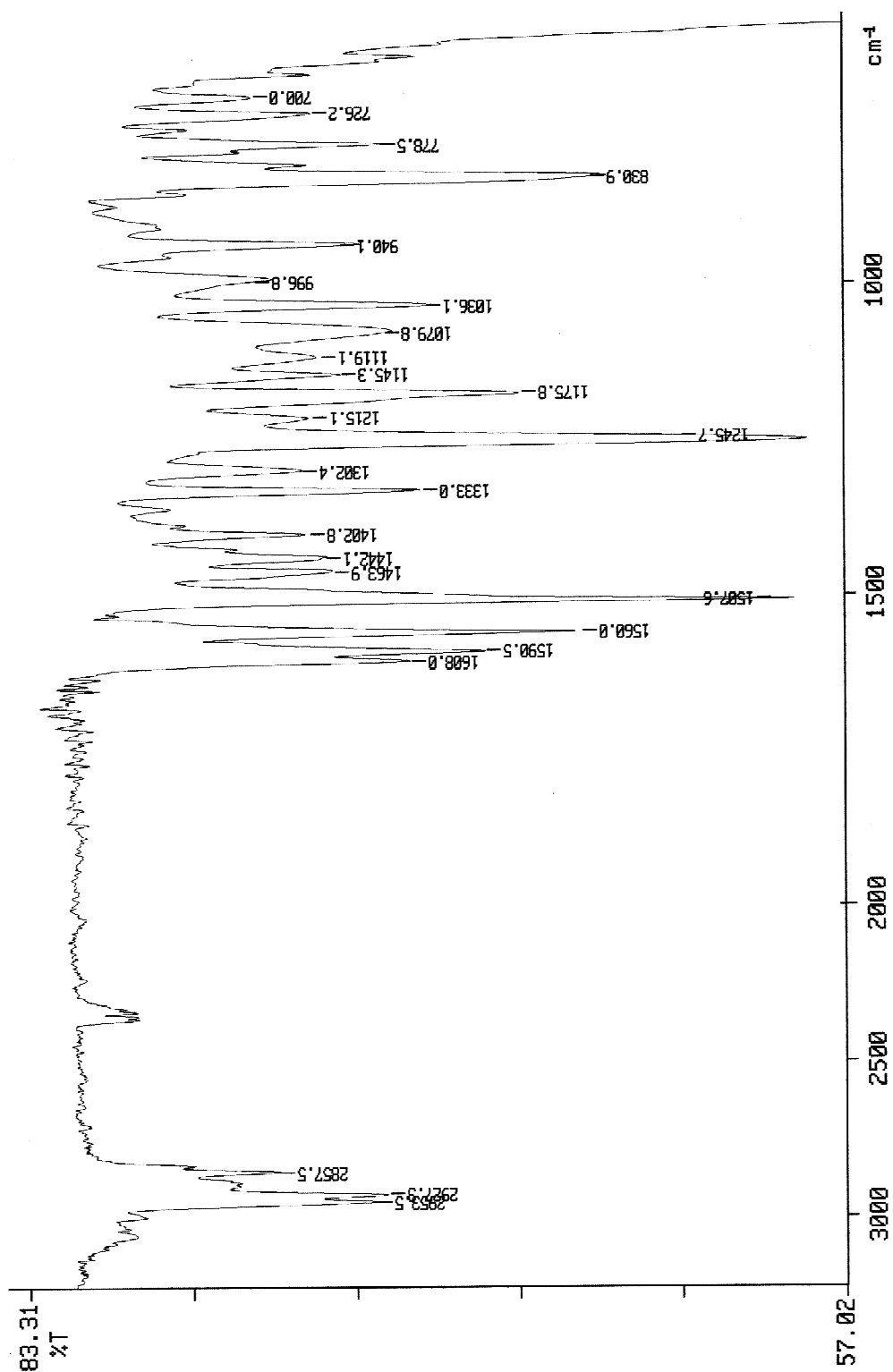


Figure A4.17.3. IR spectrum of compound 4.22 (film).

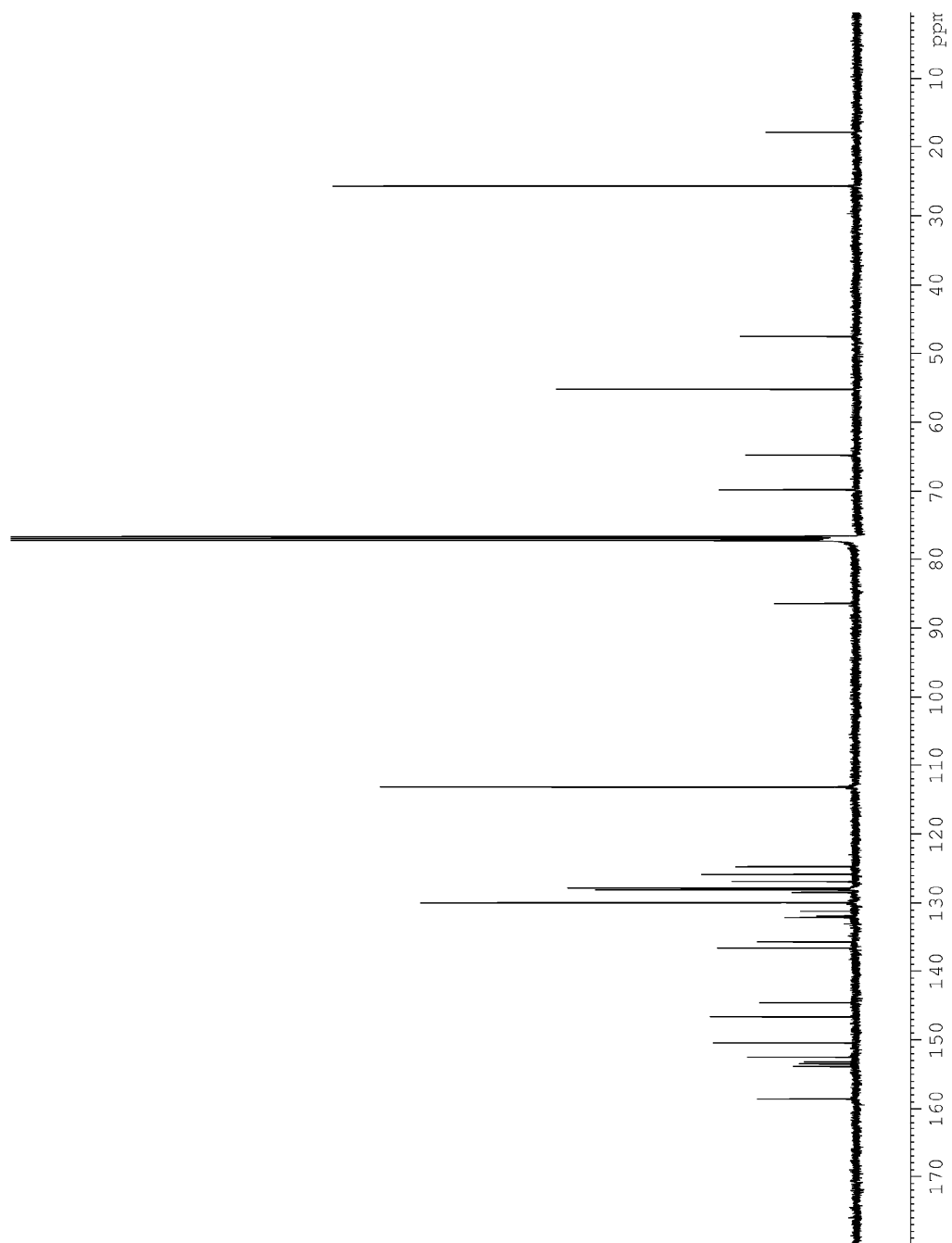
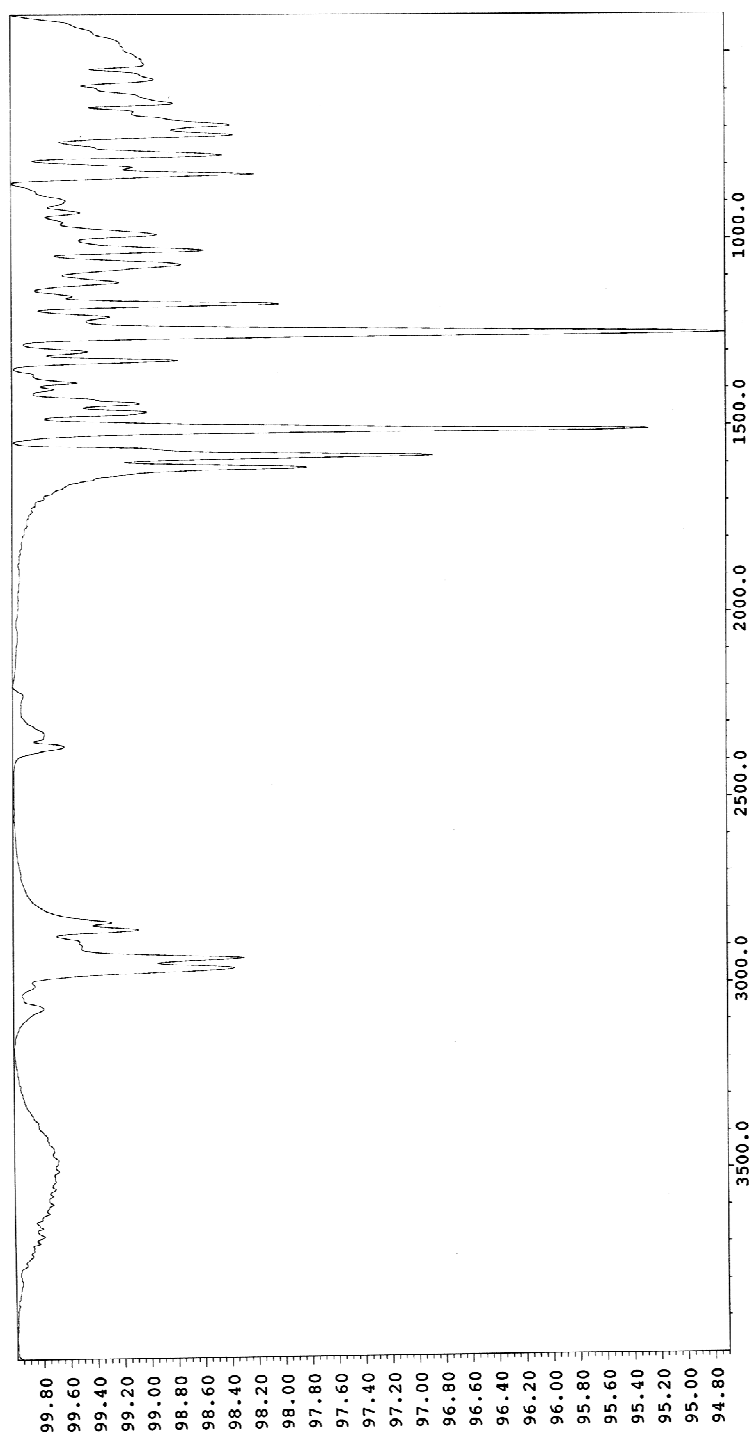


Figure A4.18.2. ^{13}C NMR spectrum of compound **4.25** (125 MHz, CDCl_3).



Wavenumber (cm-1) Nic/IR

Figure A4.18.3. IR spectrum of compound 4.25 (film).

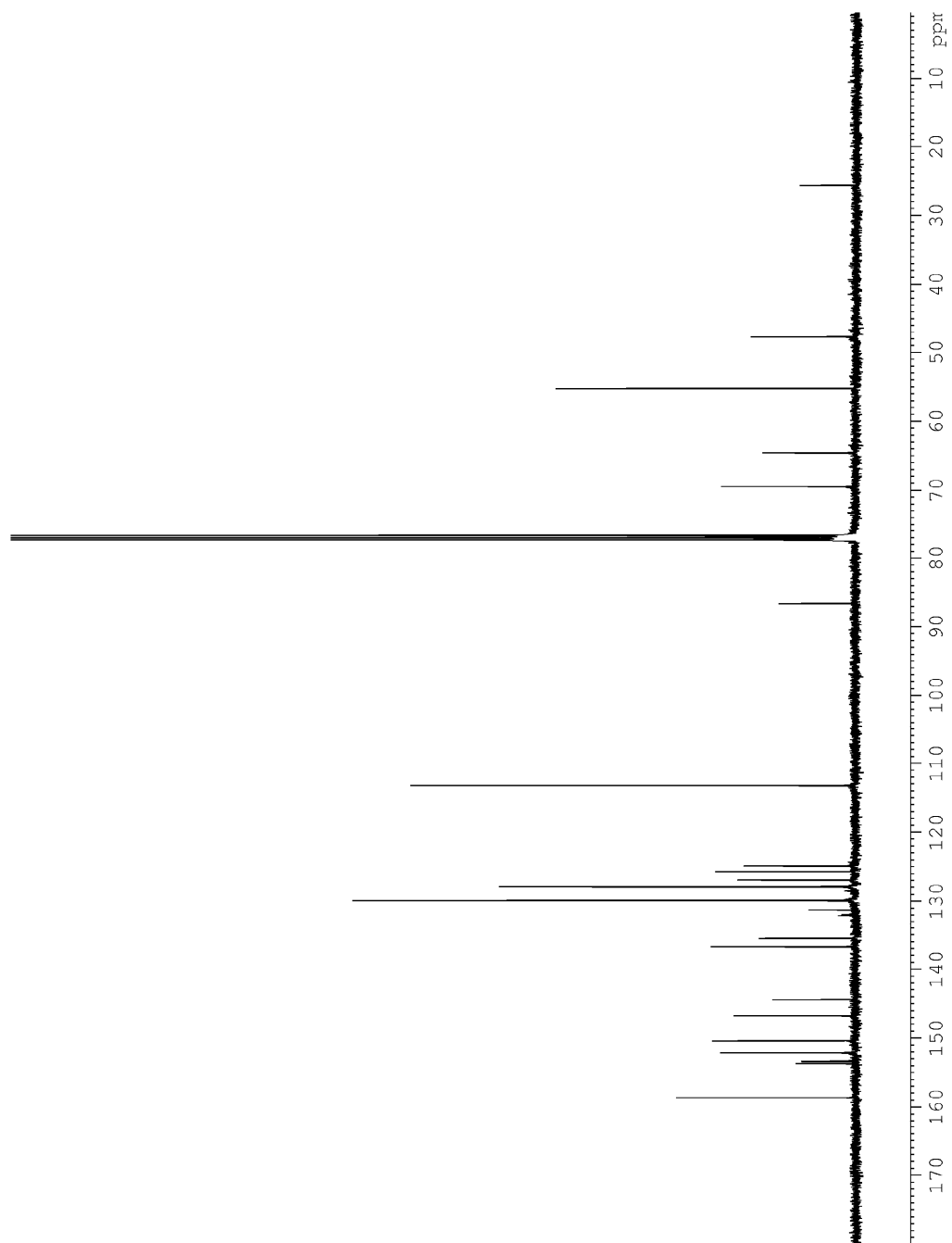


Figure A4.19.2. ^{13}C NMR spectrum of compound **4.26** (100 MHz, CDCl_3).

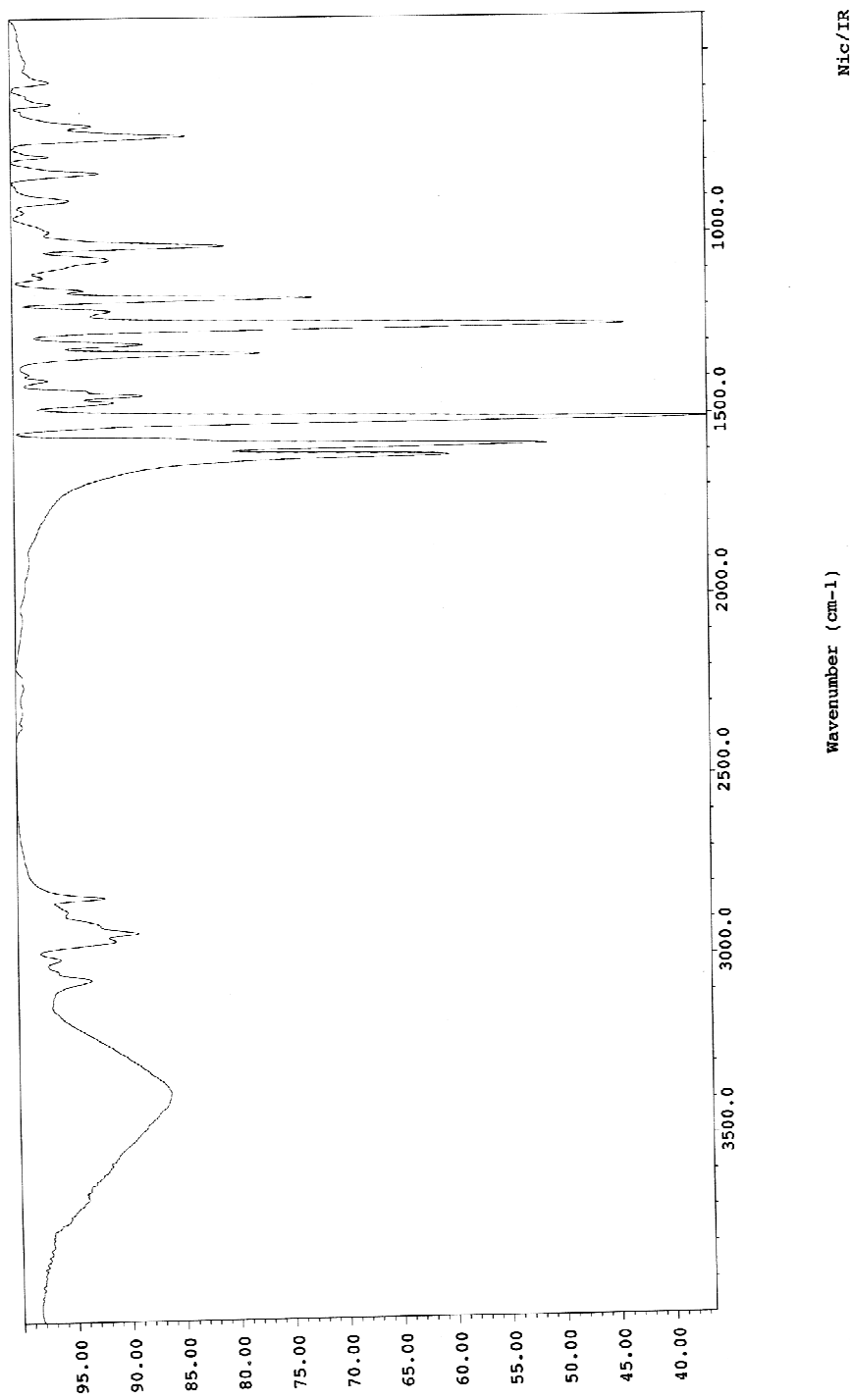


Figure A4.19.3. IR spectrum of compound 4.26 (film).

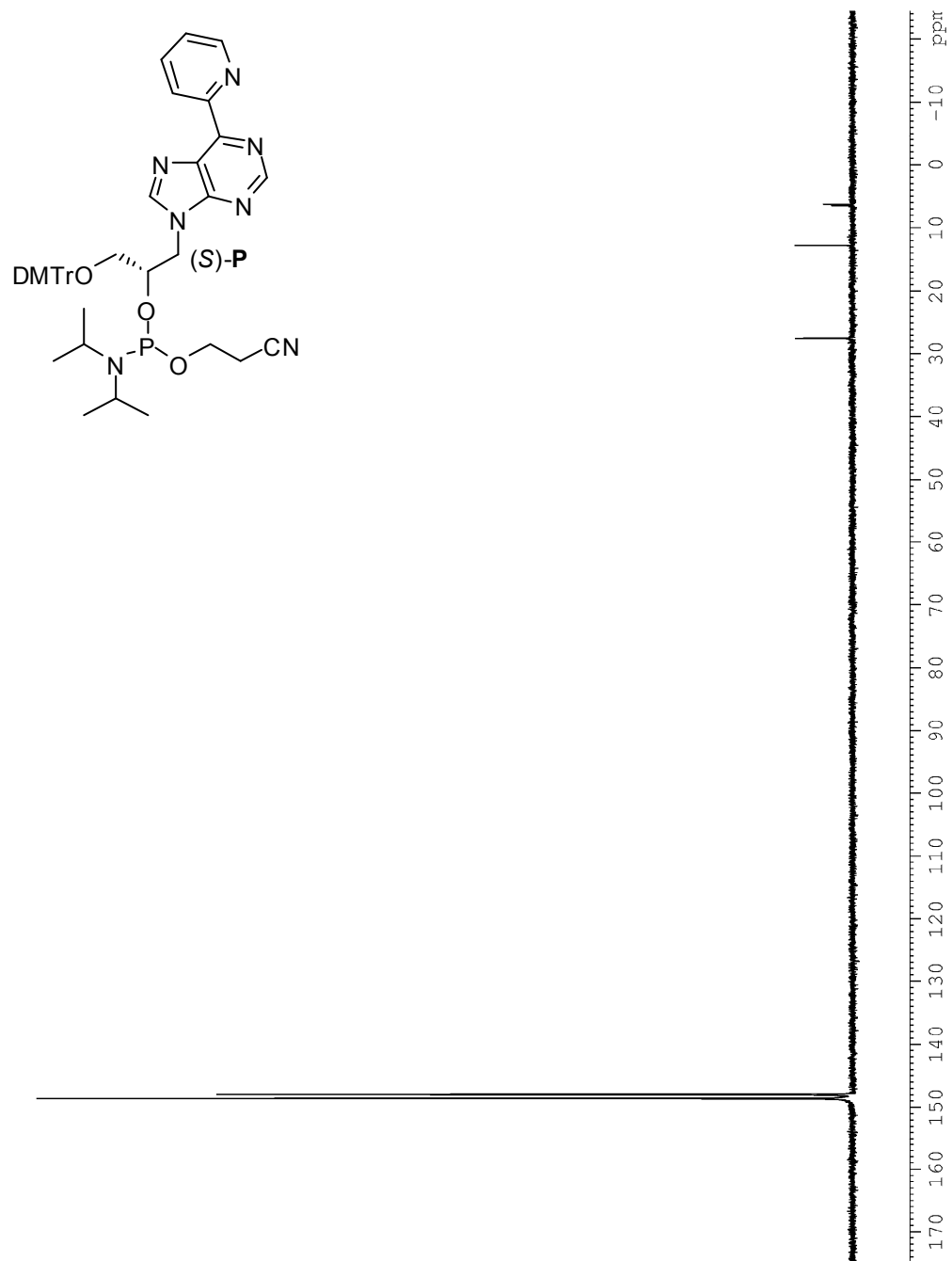


Figure A4.20.1. ^{31}P NMR spectrum of phosphoramidite (S)-P (162 MHz, CDCl_3)

Table A4.1.1. Data collection, phasing, and refinement statistics of MKS64

Data collection									
Dataset (wavelength Å)	Cell (Å)	Resol. (Å)	Measured, unique reflections	R _{merge} ^[a]	I/σ(I) ^[b]	B _{Wilson} (Å ²)	Completeness		
Native (0.80000)	a=18.16, b=48.59, c=59.87	12.7-1.28	79804, 7095	0.079 (0.296)	20.4 (4.0)	13.1	0.991 (1.00)		
Peak (1.37950)	a=18.15, b=48.59, c=59.84	24.3-1.80	34577, 2651	0.043 (0.124)	39.6 (18.6)	18.8	0.997 (1.00)		
Infl. (1.37975)	a=18.15, b=48.59, c=59.84	24.3-1.80	34586, 2649	0.042 (0.130)	40.4 (18.2)	19.3	0.997 (1.00)		
Remote (1.37010)	a=18.15, b=48.59, c=59.84	24.3-1.80	34659, 2656	0.055 (0.282)	31.3 (10.1)	20.5	0.999 (1.00)		
Phasing in Sharp									
Resolution (Å)	24	5.6	4.0	3.3	2.8	2.5	2.3	2.1	2.0
Figure of Merit (FOM)	0.796	0.818	0.807	0.808	0.780	0.771	0.781	0.761	0.706
Total FOM	0.754								
Refinement									
Resolution range (Å)	9.4-1.3								
Reflections (work, test)	6289, 479								
R-factor / R _{free} ^[c]	0.125 (0.184)/0.147 (0.197)								
Bases, phosphates	8, 7								
Water molecules	86								
R.m.s. deviation bonds (Å)	0.012								
R.m.s. deviation angles (°)	1.764								

[a] $R_{\text{merge}} = (\sum \sum |I(h) - \langle I(h) \rangle|) / (\sum \sum I_j(h)) \times 100$; values in parentheses correspond to highest resolution shell. [b] As calculated with the program SCALA. [c] $R = \sum \|F_o\| - k \sum \|F_c\| / \sum \|F_o\|$ with k as scaling factor; R_{free} calculated with test set.

Chapter 5

Crystallography of GNA duplexes

Chapter 5.1. Derivatizing glycol nucleic acids for phasing crystallographic data

The biggest challenge of solving de novo crystal structures is being able to figure out the problem of phasing the crystallographic data. Like any other form of electromagnetic radiation, a focused beam of x-rays is scattered by its interaction with the electron clouds of individual atoms. Unlike other forms of microscopy, which use lenses to refocus scattered visible light rays and produce an image, it is not possible to focus the scattered x-rays and produce an image. Therefore, one needs different methods to recreate the molecular image from the diffraction pattern produced by shooting x-rays at a particular sample crystal. For de novo crystal structure, one typically incorporates a heavy-atom into the structure of interest to produce phase estimates using the methods of isomorphous replacement and anomalous dispersion. In this manner, computers can then be used to recreate an image of the crystallized molecule based on its diffraction pattern.

In this respect, nucleic acid crystallography is no different from protein crystallography in that one must develop ways in which to solve the problem of phases. Although molecular replacement using duplex models may be seen as a general method in which to solve new structures, it does not always work with nucleic acid structures. Part of the reason for this may come from the fact that most of the structures of solved duplexes maintain similar sequences and it is possible that this set of structures is not representative of all geometries of double helices.¹ Therefore, researchers must develop means in which to incorporate heavy atoms within the oligonucleotide in order to phase the data. This is especially important when one has the task of solving a de novo nucleic

acid duplex structure which contains significant structural modifications, for instance of a nucleic acid with a sugar analog. One of the simplest and most classical ways to incorporate heavy atoms is to use halogenated derivatives of the Watson-Crick nucleobases. In this way, one merely needs to incorporate these nucleotides across from their Watson-Crick base pairing partner to form a derivatized duplex. Another means to phase the crystallographic data is to incorporate heavy atoms in the crystallization buffer, or by soaking the pre-formed crystals with solutions containing metal cations capable of anomalous dispersion. Unfortunately, this does not often produce crystals containing heavy metal cations at defined spots within the unit cell. Recently, there has been a significant amount of research directed towards methods in which selenium is incorporated in nucleic acid duplexes by replacing a non-bridging atom with selenium in the phosphate linker (phosphoroselenoate),^{2,3} in the thymine and guanine nucleobases,^{4,5} or the 2'-oxygen atom of the ribose sugar.⁶ Finally, there have been two reports in which researcher use artificial metal-mediated base pairs to site-specifically incorporate copper ions into a nucleic acid duplex to phase the crystallographic data.^{7,8}

Chapter 5.1.1. Synthesis of brominated nucleoside phosphoramidite derivatives

Along these lines, it was interesting to develop halogenated derivatives of nucleosides for phasing the crystallographic data of GNA duplexes. For nucleic acid crystallography, the most commonly used derivatives are the 5-bromo derivatives of uracil and cytosine. In this way, one only needs to synthesize the phosphoramidites of these modified nucleosides and incorporate them beside their Watson-Crick pairing partner since this modification should not affect the hydrogen bonding between the bases (Figure 5.1). For GNA, the syntheses of (*S*)-^{Br}U and (*S*)-^{Br}C proceed in a similar manner as the other glycol nucleoside phosphoramidites.

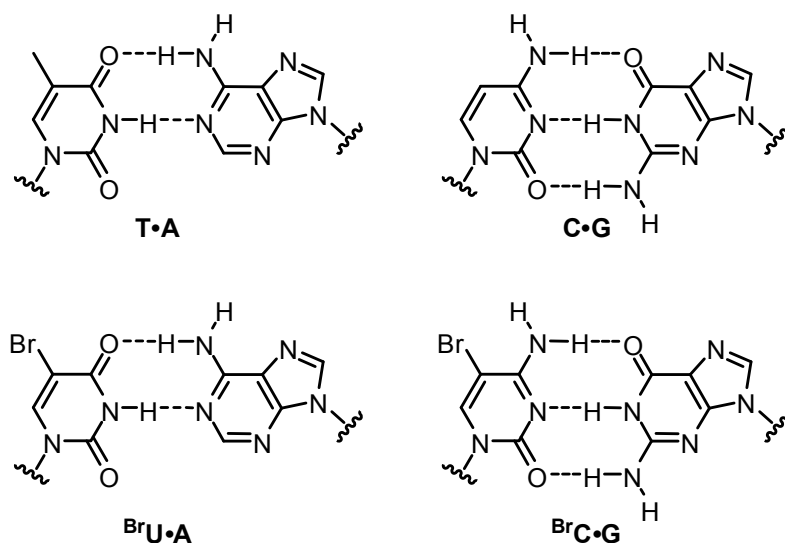
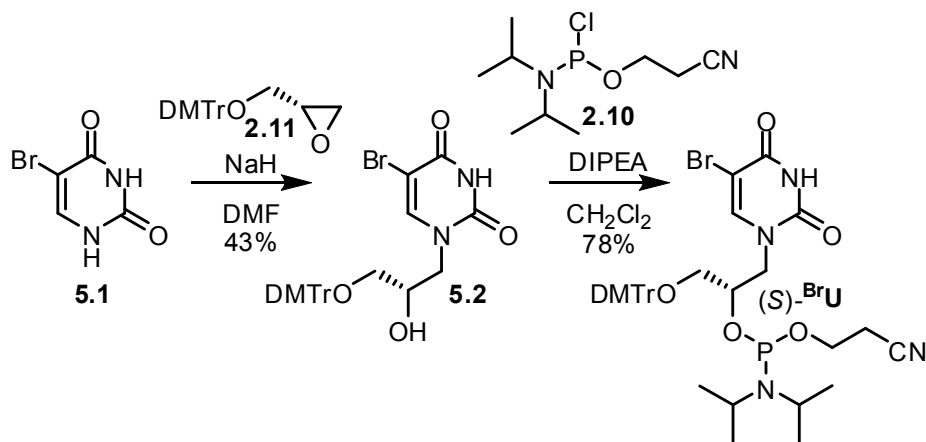


Figure 5.1. Comparison of the T•A and C•G Watson-Crick base pairs with the ^{Br}U•A and ^{Br}C•G base pair used in this study.

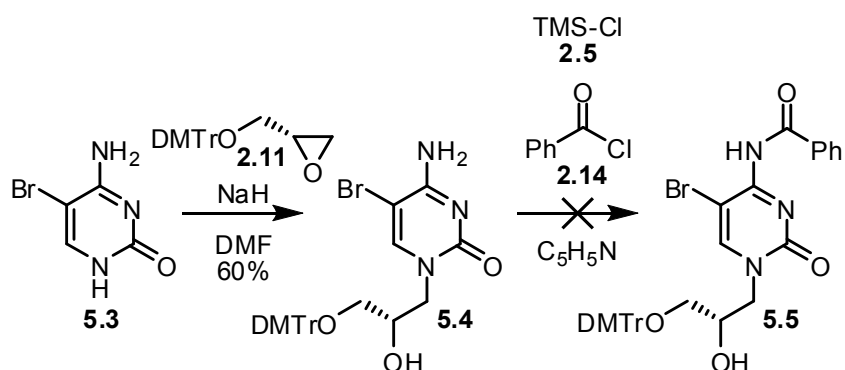
For the synthesis of (*S*)-^{Br}U, 5-bromouracil (**5.1**) and 0.2 equivalents of sodium hydride are used in the ring opening of compound **2.11** to produce compound **5.2** in 43% yield (Scheme 5.1). Compound **5.2** is then converted to the phosphoramidite (*S*)-^{Br}U using 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess *N,N*-diisopropylethylamine in 78% yield.



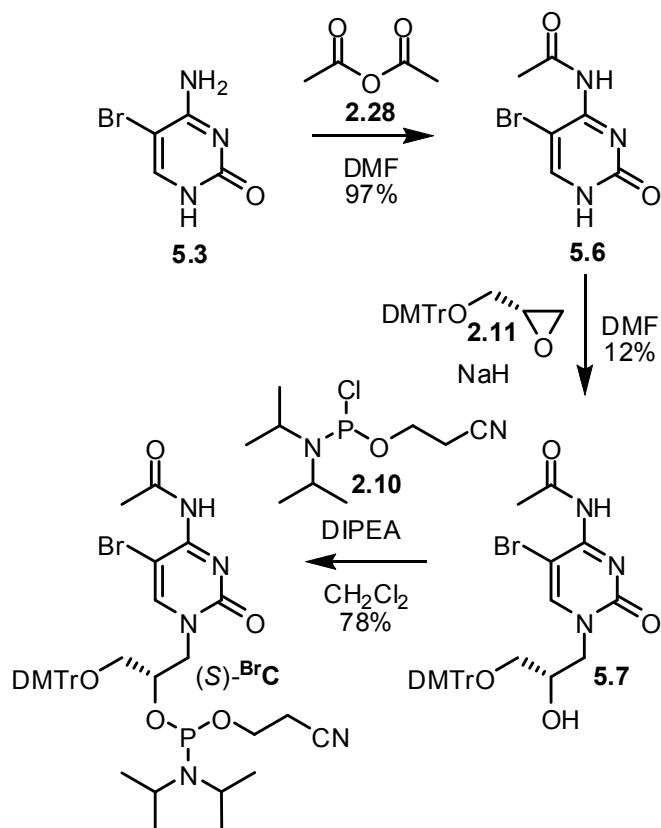
Scheme 5.1. Synthesis of (*S*)-^{Br}U phosphoramidite.

An initial attempt at the synthesis of (*S*)-^{Br}C started with the ring opening of compound **2.11** using 5-bromocytosine (**5.3**) and 0.2 equivalents of sodium hydride to produce compound **5.4** in 60% yield (Scheme 5.2). Similar to the reactions used to protect the exocyclic amines of adenine and guanine, compound **5.4** was first reacted with trimethylsilyl chloride (**2.5**) in pyridine and then with benzoyl chloride (**2.14**) in an attempt to make compound **5.5**. Unfortunately, ¹H NMR indicated several inseparable products. Therefore, an alternate route was envisioned starting with an acetyl protected version of 5-bromocytosine, compound **5.6**. Accordingly, 5-bromocytosine was reacted with acetic anhydride (**2.28**) in DMF to produce compound **5.6** in 97% yield (Scheme

5.3). Ring opening of compound **2.11** using compound **5.6** and 0.2 equivalents of sodium hydride proceeds, albeit poorly, to the formation of compound **5.7** in 12% yield. A major side product of this reaction is the ring-opened product that has lost the acetamide protection group, compound **5.4**. Further reaction of compound **5.7** with 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (**2.10**) and excess *N,N*-diisopropylethylamine affords compound (*S*)-^{Br}C in 78% yield. Unfortunately, this phosphoramidite was not completely pure after flash chromatography, indicating sensitivity towards silica gel, however, it was used as such for oligonucleotide synthesis.



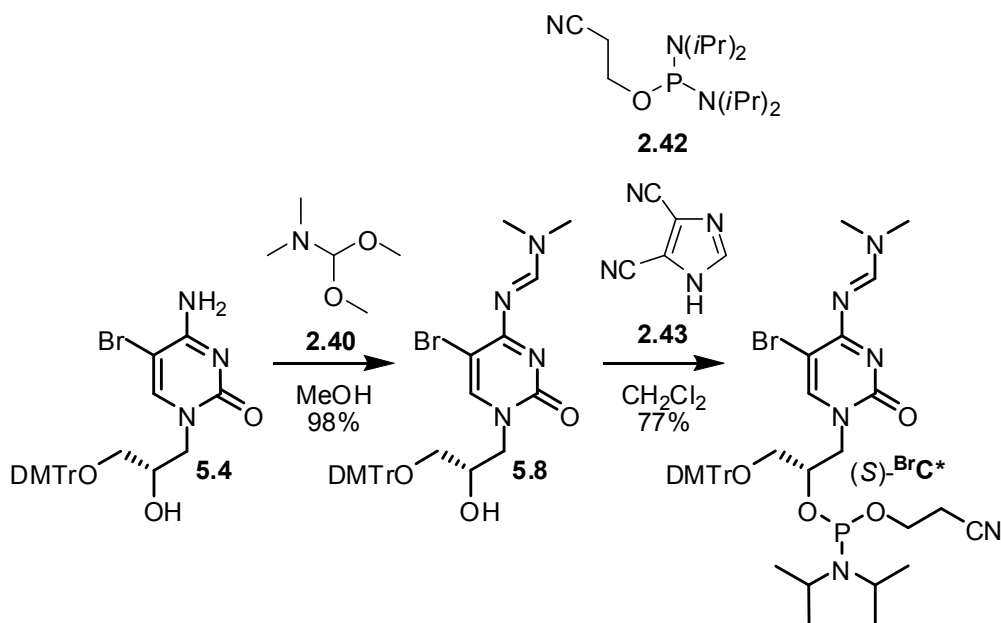
Scheme 5.2. Attempted synthesis of (*S*)-^{Br}C phosphoramidite.



Scheme 5.3. Synthesis of (*S*)-^{Br}C phosphoramidite.

Based on the disappointing yields for the synthesis of (*S*)-^{Br}C, an improved route was developed based on the success of using amidine protection of exocyclic amines of the phosphoramidites (*S*)-**A*** and (*S*)-**G***.⁹ Compound **5.3** was first reacted with dimethylformamide dimethylacetal (**2.40**) in methanol to afford compound **5.8** in 98% yield (Scheme 5.4). This reaction proceeds in only one hour and without side products allowing it to be purified quickly via a simple filter column. Compound **5.8** could then be converted to the phosphoramidite (*S*)-^{Br}C* using 2-cyanoethyl *N,N,N',N'*-tetraisopropylphosphordiamidite (**2.42**) and 0.7 equivalents of 4,5-dicyanoimidazole

(**2.43**) in 77% yield. Unlike its counterpart using acetyl protection (S)- BrC , (S)- BrC^* is completely stable towards flash chromatography and the product could be isolated without any side products. Furthermore, the route towards (S)- BrC^* proceeds with an overall yield of 45% in three steps versus an overall yield of 9% in three steps for the corresponding (S)- BrC , rendering the former a superior synthetic scheme.



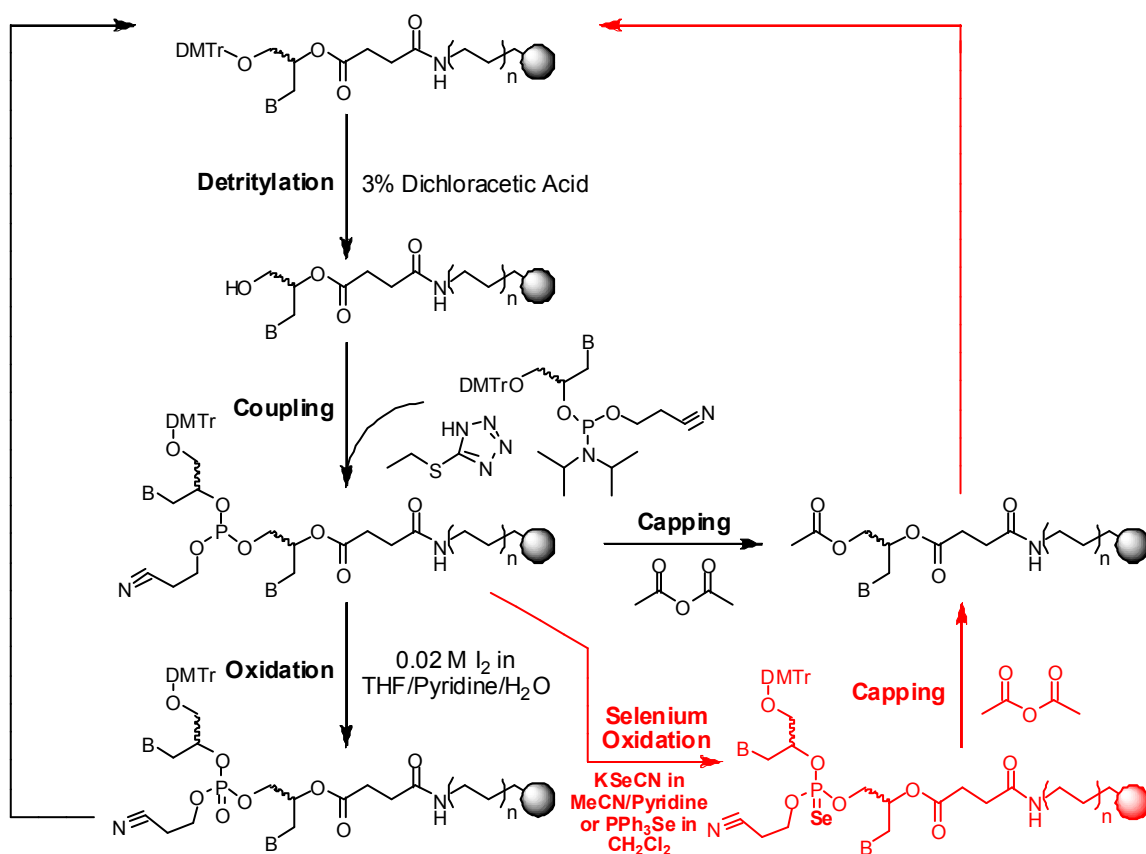
Scheme 5.4. Synthesis of (S)- BrC^* phosphoramidite.

The stabilities of phosphoramidites (S)- BrU and (S)- BrC^* were very similar to other phosphoramidites of GNA nucleosides and could be handled in the same manner. During their synthesis, no sensitivity towards ambient ultraviolet light was observed, a common concern of working with halogenated derivatives.

Chapter 5.1.2. Phosphoroselenoate derivatives of GNA oligonucleotides

Another means of derivatizing nucleic acids for crystallographic phasing is by the formation of a phosphoroselenoate linkage in the backbone.¹⁰⁻¹² Selenium has been used in this manner, in addition to selenated derivatives of nucleosides, to phase crystallographic data of DNA duplexes.²⁻⁶ The phosphoroselenoate differs from the phosphate linkage in that one of the non-bridging oxygen atoms is replaced by a selenium atom, therefore providing a handle to phase the crystallographic data. In this manner, the modification is introduced during oligonucleotide synthesis, avoiding the tedious work of synthesizing individual phosphoramidite derivatives. Accordingly, phosphoroselenoate GNA could be synthesized similar to published procedures which uses potassium selenocyanate as a selenium source.^{2,3,10,12} During oligonucleotide synthesis, the process is interrupted between the coupling and capping steps of the desired linkage leaving a phosphorus(III) intermediate (Scheme 5.5). The synthesis column was then oxidized with selenium by incubating with a saturated solution of KSeCN in 95:5 acetonitrile:triethylamine instead of the normal iodine and water, thereby producing the phosphoroselenoate linkage. After washing with acetonitrile and manual capping, the synthesis column is returned to the oligonucleotide synthesizer to continue extension of the sequence. Oligonucleotides were then deprotected in a normal fashion using 25% aqueous ammonium hydroxide at 55°C, purified using reverse phase HPLC in “Trityl ON” mode, the trityl group cleaved, and the subsequent crude oligo further purified using reverse phase HPLC. An example crude HPLC trace is shown in Figure 5.2 in which a

phosphoroselenoate linkage was incorporated between the first and second nucleotides of the sequence 3'-A_{PSe}TGCGCAT-2'. It should be noted that oxidizing with selenium produces an oligonucleotide mixture composed of two diastereomers, resulting from the newly established chiral phosphorus atom and the presence of the chiral center in the dihydroxypropyl backbone. Surprisingly, these two diastereomers were separated quite easily using reverse phase HPLC.



Scheme 5.5. Overview of the procedure for synthesizing GNA oligonucleotides containing phosphoroselenoate linkages.

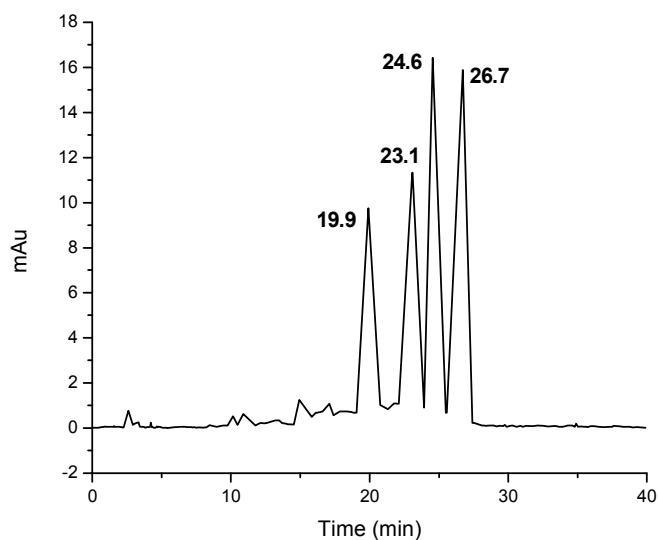


Figure 5.2. Crude HPLC trace of the GNA oligonucleotide 3'-A_{PSe}TGCGCAT-2'. The peak at 19.9 minutes represents the n-1 product. The peak at 23.1 minutes represents the GNA oligo containing all phosphate linkages. The peaks at 24.6 and 26.7 minutes represent the two diastereomers resulting from the phosphoroselenoate linkage. The crude oligo was eluted using a Waters Xterra column (MS C₁₈, 4.6 x 50 mm, 2.5 μM particle size) at room temperature with a linear gradient (flow = 1.0 mL/min) from 5-15% acetonitrile in 40 minutes and 95-85% aqueous tritethylammonium acetate buffer (50 mM, pH=7.0).

Although this method was successful in introducing the phosphoroselenoate linkage, an improved procedure was desired based on the toxicity and low solubility of KSeCN, making it difficult to work with. Furthermore, KSeCN has a low reactivity, the most likely cause for the formation of a significant peak for the n-1 product; in the case of Figure 5.2, the oligonucleotide 3'-TGCGCAT-2'. Therefore, another procedure was followed which describes the use of triphenylphosphine selenide as the selenium source for oxidation of the phosphorus(III) intermediate.^{11,13} Similar to the procedure using KSeCN, oligonucleotide synthesis is again halted after coupling of the phosphoramidite

to the growing chain. The solid support is transferred from the synthesis column to a screw cap vial and then incubated with a 25 mM solution of PPh_3Se in methylene chloride. Afterwards, the solid support was washed once with methylene chloride, transferred back to the synthesis column, manually capped on the synthesizer, and then the extension of the oligonucleotide was continued in a normal fashion. Deprotection and “Trityl ON” purification were performed normally and the crude GNA oligos purified via reverse phase HPLC. In this manner, a phosphoroselenoate linkage was incorporated in six out of the seven possible position of the sequence 3'-CTCTAGAG-2' and four of the five positions of 3'-GCGCGC-2'. An example HPLC trace of the crude oligonucleotide mixture using this method of selenium oxidation is shown in Figure 5.3 demonstrating the high yield of selenium incorporation in most cases. Incorporation between the seventh and eighth nucleotides proceeded poorly (3'-CTCTAGAP_{Se}G-2'), resulting in very little phosphoroselenoate product. It was never attempted to oxidize between the fifth and sixth nucleotides of the 6-mer sequence (3'-GCGCGP_{Se}C-2') based on the poor results with the 8-mer sequence. Again, reverse phase HPLC afforded separation of the two diastereomers, albeit to different extents as demonstrated in Figure 5.4.

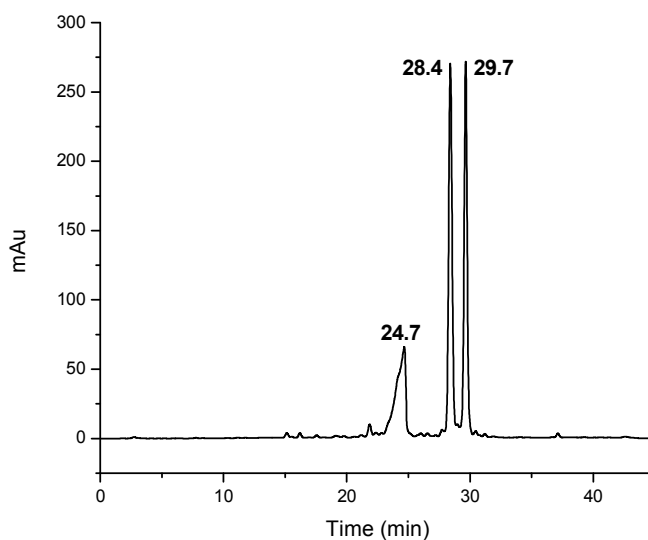


Figure 5.3. Crude HPLC trace of the GNA oligonucleotide 3'-CTC_{PSe}TAGAG-2'. The peak at 24.7 minutes represents the GNA oligo containing all phosphate linkages. The peaks at 28.4 and 29.7 minutes represent the two diastereomers resulting from the phosphoroselenoate linkage. The crude oligo was eluted using a Waters Xterra column (MS C₁₈, 4.6 x 50 mm, 2.5 μ M particle size) at 50 °C with a linear gradient (flow = 1.0 mL/min) from 2-6% acetonitrile in 40 minutes and 98-94% aqueous triethylammonium acetate buffer (50 mM, pH=7.0).

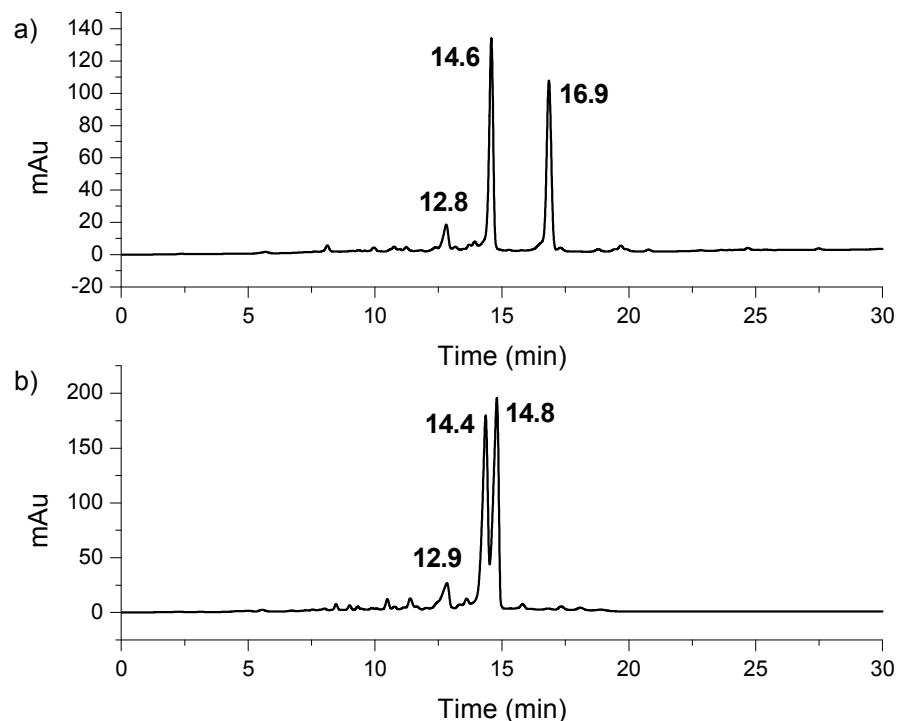


Figure 5.4. Crude HPLC trace of GNA oligonucleotides functionalized using a phosphoroselenoate linkage demonstrating the differences in separation ability of the two diastereomers. a) GNA oligonucleotide 3'-GC_{PSe}GCGC-2'. The peak at 12.8 minutes represents the GNA oligo containing all phosphate linkages. The peaks at 14.6 and 16.9 minutes represent the two diastereomers resulting from the phosphoroselenoate linkage. b) GNA oligonucleotide 3'-GCGC_{PSe}GC-2'. The peak at 12.9 minutes represents the GNA oligo containing all phosphate linkages. The peaks at 14.4 and 14.8 minutes represent the two diastereomers resulting from the phosphoroselenoate linkage. The crude oligos were eluted using a Waters Xterra column (MS C₁₈, 4.6 x 50 mm, 2.5 μ M particle size) at 60 °C with a linear gradient (flow = 1.0 mL/min) from 1-7% acetonitrile in 30 minutes and 99-93% aqueous tritethylammonium acetate buffer (50 mM, pH=7.0).

Although this new method of selenium incorporation proceeds with high yield and results in two phosphoroselenoate diastereomers to be tested for crystallography, the stability of these oligos is poor and more prone to oxidation compared to a normal phosphate linkage. This is consistent with previous results from Egli in which

phosphoroselenoate derivatives of DNA had low stability under normal conditions and acceptable phasing data could only be collected in the time period of one week between crystallization and data collection.³ From our experience, phosphoroselenate derivatives of GNA oligonucleotides are stable to all conditions during oligonucleotide synthesis, deprotection, and HPLC purification. However, it seems as though decomposition occurs during concentration of the freshly purified samples. It is uncertain whether this decomposition occurs because of effects during the concentration of the triethylammonium acetate buffer (i.e. – differential evaporation of the triethylamine and acetic acid components) or from the manner in which the oligos are concentrated. Lyophilization under high vacuum of the oligos was also used in comparison to evaporation under low vacuum with centrifugation, but provided mixed results as far as the stability of the phosphoroselenoate linkage. Our inability to obtain pure, unoxidized samples of phosphoroselenoate derivatives resulted in our disinterest with this method of phasing the crystallographic data. It may be possible that selenium derivatives of nucleosides are more stable towards oxidation and are more interesting for the phasing of nucleic acid duplexes.^{4,5}

Chapter 5.2. Crystallographic conditions

The purification of GNA oligonucleotides to be used for crystallography was generally followed by desalting over Sep-Pak C₁₈ columns. This allowed for the removal of any residual buffer and/or any excess metal cations. For some sequences, this process was found to be more important than in others. For example, crystals of 3'-ATGCGCAT-2' (MKS42) could only be grown in conditions using cobalt hexamine after desalting over Sep-Pak columns. Even though crystals usually developed in other conditions not containing cobalt hexamine, the highest quality crystals (for all GNA oligos) were observed in the buffers that contained the hexamine salt. After desalting, the oligonucleotide was concentrated and then redissolved in enough water to make a 1-2 mM stock solution. The concentration was also found to be an important parameter. For example, in some cases increasing the concentration to 2 mM produced crystals whereas no crystals appeared with a concentration of 1 mM. The stock solution was finally filtered to remove any small, insoluble particles.

Since crystallization set-ups were all performed at 4 °C, the stock solution of the oligo was placed in the 4 °C cold room for several hours prior to use to make sure the GNA oligonucleotide solution was equilibrated. Furthermore, the crystallization buffers were also equilibrated (generally stored in the cold room) at 4 °C overnight to ensure little temperature fluctuation during the crystallization set-up. The hanging drop vapor diffusion method was initially used for crystallization, but was later switched to sitting drop vapor diffusion based on the greater ease during set-up. It was found that almost

identical results were obtained with both methods. Crystallization buffers were obtained by using the Nucleic Acid Mini Screen (Hampton Research, HR2-118). This screen consists of 24 buffers with differing pH, salt, and polyamine conditions (Table 5.1). The GNA oligonucleotide sample (2 μ L) was then mixed with these buffers (4 μ L) and equilibrated against a well of 35% 2-methyl-2,4-pentanediol (MPD, 1 mL) after the plate was sealed with clear, transparent tape.

Table 5.1. Formulation of the Nucleic Acid Mini Screen (Hampton Res.)

Tube Number	Precipitant	Tube Number	Buffer	Tube Number	Polyamine	Tube Number	Monovalent Ion	Tube Number	Divalent Ion
1.	10% v/v MPD	1.	40 mM Na Cacodylate pH 5.5	1.	20 mM Cobalt Hexamine	1.	None	1.	20 mM Magnesium Chloride
2.	10% v/v MPD	2.	40 mM Na Cacodylate pH 5.5	2.	20 mM Cobalt Hexamine	2.	80 mM Sodium Chloride	2.	20 mM Magnesium Chloride
3.	10% v/v MPD	3.	40 mM Na Cacodylate pH 5.5	3.	20 mM Cobalt Hexamine	3.	12 mM Sodium Chloride / 80 mM Potassium Chloride	3.	None
4.	10% v/v MPD	4.	40 mM Na Cacodylate pH 5.5	4.	20 mM Cobalt Hexamine	4.	40 mM Lithium Chloride	4.	20 mM Magnesium Chloride
5.	10% v/v MPD	5.	40 mM Na Cacodylate pH 6.0	5.	12 mM Spermine tetra-HCl	5.	80 mM Potassium Chloride	5.	20 mM Magnesium Chloride
6.	10% v/v MPD	6.	40 mM Na Cacodylate pH 6.0	6.	12 mM Spermine tetra-HCl	6.	80 mM Potassium Chloride	6.	None
7.	10% v/v MPD	7.	40 mM Na Cacodylate pH 6.0	7.	12 mM Spermine tetra-HCl	7.	80 mM Sodium Chloride	7.	20 mM Magnesium Chloride
8.	10% v/v MPD	8.	40 mM Na Cacodylate pH 6.0	8.	12 mM Spermine tetra-HCl	8.	80 mM Sodium Chloride	8.	None
9.	10% v/v MPD	9.	40 mM Na Cacodylate pH 6.0	9.	12 mM Spermine tetra-HCl	9.	80 mM Sodium Chloride / 12 mM Potassium Chloride	9.	20 mM Magnesium Chloride
10.	10% v/v MPD	10.	40 mM Na Cacodylate pH 6.0	10.	12 mM Spermine tetra-HCl	10.	12 mM Sodium Chloride / 80 mM Potassium Chloride	10.	None
11.	10% v/v MPD	11.	40 mM Na Cacodylate pH 6.0	11.	12 mM Spermine tetra-HCl	11.	80 mM Sodium Chloride	11.	20 mM Barium Chloride
12.	10% v/v MPD	12.	40 mM Na Cacodylate pH 6.0	12.	12 mM Spermine tetra-HCl	12.	80 mM Potassium Chloride	12.	20 mM Barium Chloride
13.	10% v/v MPD	13.	40 mM Na Cacodylate pH 6.0	13.	12 mM Spermine tetra-HCl	13.	None	13.	80 mM Strontium Chloride
14.	10% v/v MPD	14.	40 mM Na Cacodylate pH 7.0	14.	12 mM Spermine tetra-HCl	14.	80 mM Potassium Chloride	14.	20 mM Magnesium Chloride
15.	10% v/v MPD	15.	40 mM Na Cacodylate pH 7.0	15.	12 mM Spermine tetra-HCl	15.	80 mM Potassium Chloride	15.	None
16.	10% v/v MPD	16.	40 mM Na Cacodylate pH 7.0	16.	12 mM Spermine tetra-HCl	16.	80 mM Sodium Chloride	16.	20 mM Magnesium Chloride
17.	10% v/v MPD	17.	40 mM Na Cacodylate pH 7.0	17.	12 mM Spermine tetra-HCl	17.	80 mM Sodium Chloride	17.	None
18.	10% v/v MPD	18.	40 mM Na Cacodylate pH 7.0	18.	12 mM Spermine tetra-HCl	18.	80 mM Sodium Chloride / 12 mM Potassium Chloride	18.	20 mM Magnesium Chloride
19.	10% v/v MPD	19.	40 mM Na Cacodylate pH 7.0	19.	12 mM Spermine tetra-HCl	19.	12 mM Sodium Chloride / 80 mM Potassium Chloride	19.	None
20.	10% v/v MPD	20.	40 mM Na Cacodylate pH 7.0	20.	12 mM Spermine tetra-HCl	20.	80 mM Sodium Chloride	20.	20 mM Barium Chloride
21.	10% v/v MPD	21.	40 mM Na Cacodylate pH 7.0	21.	12 mM Spermine tetra-HCl	21.	80 mM Potassium Chloride	21.	20 mM Barium Chloride
22.	10% v/v MPD	22.	40 mM Na Cacodylate pH 7.0	22.	12 mM Spermine tetra-HCl	22.	40 mM Lithium Chloride	22.	80 mM Strontium Chloride / 20 mM Magnesium Chloride
23.	10% v/v MPD	23.	40 mM Na Cacodylate pH 7.0	23.	12 mM Spermine tetra-HCl	23.	40 mM Lithium Chloride	23.	80 mM Strontium Chloride
24.	10% v/v MPD	24.	40 mM Na Cacodylate pH 7.0	24.	12 mM Spermine tetra-HCl	24.	None	24.	80 mM Strontium Chloride / 20 mM Magnesium Chloride

Crystals generally appeared within 1-2 weeks and they were subsequently picked from the drop using nylon loops after raising the concentration of MPD (for cryoprotection) in the buffer to 30%. The optimal cryoprotectant concentration of 30% was obtained by freezing samples of buffer containing different concentrations of MPD and picking the optically best looking condition. Initial crystal quality was judged by the size and shape of the growing crystals. In general, crystals of GNA oligonucleotides were square shaped or long rectangular rods with a plate like appearance. The crystals were then frozen and stored under liquid nitrogen until data collection.

Chapter 5.3. Structure of the GNA duplex 3'-GCGCGC-2'

The most recent crystal structure of a GNA oligonucleotide was solved using the self-complementary brominated derivative 3'-G^{Br}CGCGC-2'. Initially, diffraction quality crystals of the self-complementary sequence 3'-GCGCGC-2' (MKS5) were obtained from condition #20 of the Nucleic Acid Mini Screen and provided diffraction data that could be processed up to 1.20 Å. After this initial data was collected, the native sequence was derivatized using both bromine and selenium in an attempt to phase the data (see Table 5.2). Incorporating selenium via a phosphoroselenoate linkage into the structure proved to be trivial. However, as stated above, this linkage was unstable and prone to oxidation. Therefore, we focused on incorporating the bromo-cytosine glycol nucleotide into the sequence to phase the crystallographic diffraction data. Crystals were obtained for both sequences 3'-GCG^{Br}CGC-2' (MKS118) and 3'-G^{Br}CGCGC-2' (MKS132), albeit in different crystallization conditions.

Table 5.2. GNA oligos synthesized to solve the structure of 3'-GCGCGC-2'

Name	Sequence	Molecular Weight	Outcome
MKS5/MKS153	3'-GCGCGC-2'	1541	Data set collected up to 1.2 Å
MKS141	3'-G _{PSe} CGCGC-2'	1604	Only synthesized
MKS140	3'-GC _{PSe} GCGC-2'	1604	Only synthesized
MKS139	3'-GCG _{PSe} CGC-2'	1604	Only synthesized
MKS138	3'-GCGC _{PSe} GC-2'	1604	Only synthesized
MKS118/MKS131	3'-GCG ^{Br} CGC-2'	1620	Crystals obtained, poor diffraction
MKS119/MKS132	3'-G ^{Br} CGCGC-2'	1620	Structure solved

Initially, MKS118 provided the nicest looking crystals in similar crystallization conditions (containing spermine as the polyamine) as the native MKS5. Unfortunately, the data did not provide very high resolution diffraction (up to 2.6 Å) and it was not adequate to generate phases to solve the structure. Thereafter, crystals of the sequence 3'-G^{Br}CGCGC-2' (MKS132) were obtained in a crystallization condition containing cobalt hexamine as the polyamine providing diffraction data up to 0.97 Å.¹⁴ Several data sets were collected at the bromine absorption edge using separate scans for high and low resolution which allowed us to solve the structure using SAD phasing. The electron density generated provided unambiguous density for all six nucleotides, 48 water molecules, and three cations (see appendix for complete crystallographic data). Hexagonal packing is mediated by both cobalt hexamine cations and the interaction of bromines between adjacent duplexes (Figure 5.5).

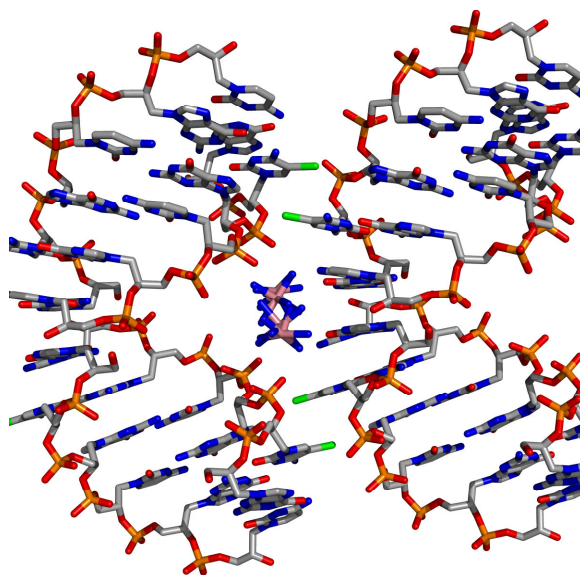


Figure 5.5. Packing contacts between duplexes in the crystal structure of 3'-G^{Br}CGCGC-2'. Bromine atoms are highlighted in green and are separated by 3.64 angstroms. A single, cobalt (pink) hexamine molecule taking on two conformations in the crystal is shown in the middle.

The overall structure (Figure 5.6) of this right-handed (*S*)-GNA double helix (from now on referred to as Type N) resembles that of a previously reported (*S*)-GNA double helix containing an artificial hydroxypyridone base pair (from now on referred to as Type M),⁸ but differs significantly from the canonical A- and B-form nucleic acid helices. A comparison of the two (*S*)-GNA helices shows that the Type N helix is compressed along the z-axis relative to the Type M helix (Figure 5.6). Although this compression is accompanied by very little change in the helix diameter, it causes the Type N helix to adopt a shallower pitch of 26 Å with 10 residues per turn versus 60 Å with 16 residues per turn for the Type M helix (Table 5.3). This also forces the

phosphates of opposing strands to be closer in the crystal as it wraps around the helix axis. Furthermore, the base pairs of the Type N helix are displaced from the helix axis (x-displacement) by 5.4 to 6.8 Å, resulting in a large circular hollow core similar to the Type M helix. Since the x-displacements are more regular for the Type N helix, the hollow core is circular rather than oblong as in the Type M helix which has greater variance of x-displacement values (from 5.1 to 8.6 Å). Also consistent with the Type M structure, the Type N (*S*)-GNA helix possesses one groove, corresponding to the canonical minor groove, and lacks a major groove, which is instead a convex surface.

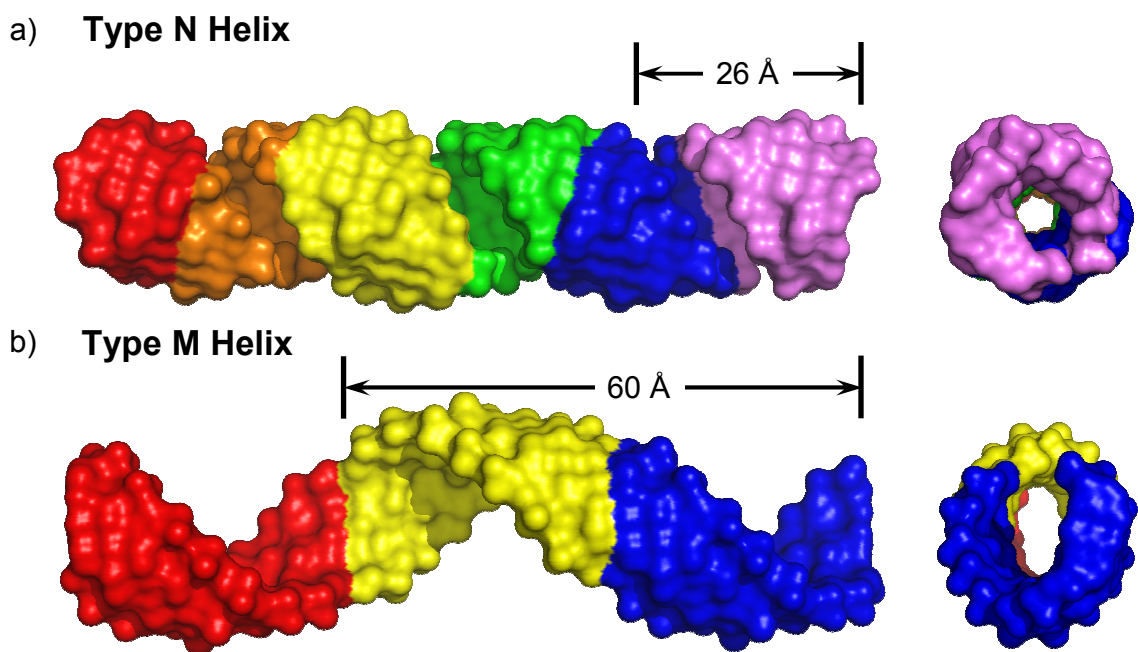


Figure 5.6. Overall structures of (a) Type N and (b) Type M⁸ GNA oligonucleotide helices with the length of a single duplex turn highlighted. Figure generated using PyMOL.

Table 5.3. Comparison of Average Helical Parameters for (S)-GNA, B-DNA, and A-DNA

	(S)-GNA Type M ^[a]	(S)-GNA Type N ^[a]	B-DNA ^[b]	A-DNA ^[b]
Helical sense	right	right	right	right
Residues per turn	16	10	10	12
Helical pitch (Å)	60	26	34	34
Helical rise (Å)	3.8	2.6	3.4	2.9
x-displacement (Å)	-7.0	-6.0	0.1	-4.2
Tilt(°) ^b	0.0	0.5	0.1	-0.1
Roll (°) ^b	-2.7	6.4	0.6	8.0
Twist (°) ^b	23.5	35.7	36.0	31.0
Slide (Å) ^b	-3.5	-3.4	0.2	-1.5
P-P distance (Å) ^c	5.4	5.4	7.0	5.9

[a] Data for GNA were calculated using the program CURVES^{15,16}. Data for B-DNA and A-DNA were taken from published values.^{17,18} [b] Local base pair step parameters. [c] Intrastrand P-P distances.

All base pairs are engaged in standard Watson-Crick hydrogen bonding patterns with the 5-bromocytosine nucleotide appearing to have little, if any, distorting effect. The distances between C_{1'}-C_{1'} carbons range from 10.71 to 10.85 Å which is in agreement for average values found in DNA duplexes (10.85 Å) and with the values previously reported for Watson-Crick base pairs in the Type M (S)-GNA 8-mer duplex.

Similar to the Type M GNA structure, the propylene glycol nucleotides adopt two different conformations with respect to the torsional angles between C_{3'}-O and C_{2'}-O (Figure 5.7). In contrast to the Type M structure in which all of the nucleotides involved in Watson-Crick base pairs maintain a *gauche* conformation, the nucleotides in the Type N structure adopt alternating *gauche* and *anti* conformations with average torsional angles γ of -66° and -174° , respectively. This also results in the opposing nucleotide of the Watson-Crick base pair to adopt the opposite conformation in the self-complementary duplex. Consistent with the Type M structure, and also what one would expect for such a simplified backbone, the average intrastrand phosphate distance of 5.4 Å is quite short compared to A- and B-form helices. Another interesting feature of the Type N GNA duplex is the large average slide between neighboring base pairs of 3.4 Å, resulting from the large backbone inclination ranging from -46° to -53° (see Figure 5.7).³ Similar to the Type M structure, this results in extensive interstrand base-stacking interactions and almost a complete absence of intrastrand base stacking, the major form in A- and B-form nucleic acids. Furthermore, this causes the C_{1'}-H₂ group of the propylene glycol backbone to participate in packing interactions with nucleotides of the same strand.

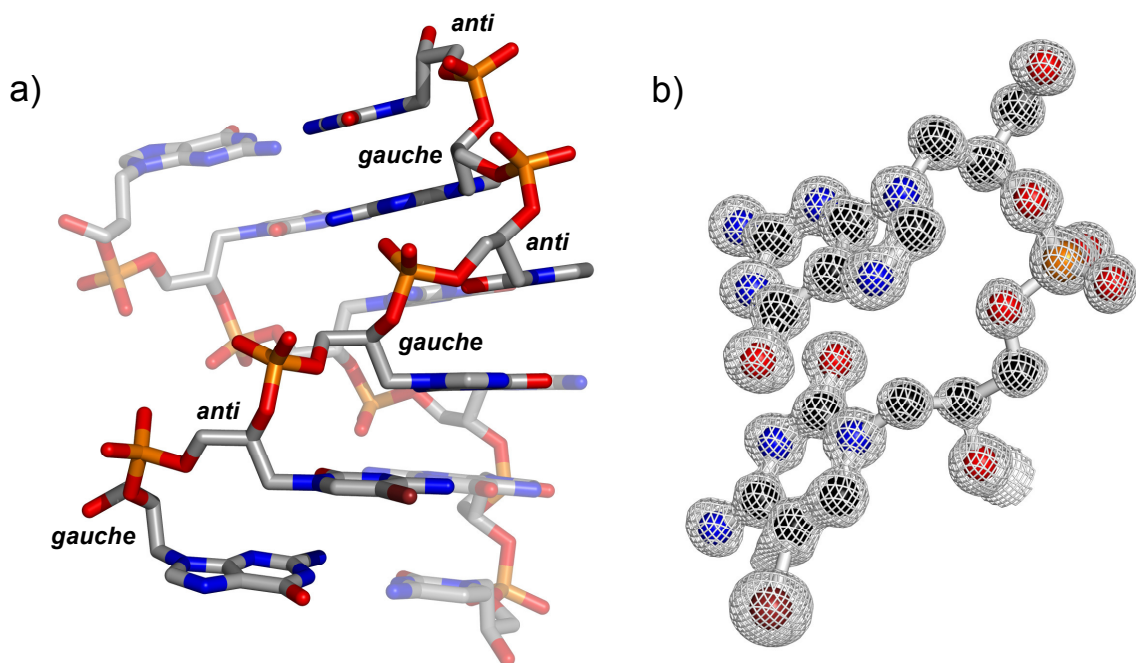


Figure 5.7. Structure of a single duplex of 3'-G^{Br}CGCGC-2' (a) and electron density of the 3'-terminal glycol nucleotides (b) at 1.5 sigma. Figure generated using PyMOL.

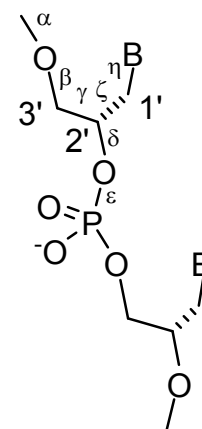
Finally, the high resolution of this structure allows us to determine for the first time accurate bond lengths, torsional angles, and phosphate bond lengths and angles. This is important for generating a more accurate stereochemical library to describe GNA nucleotides in crystal structures.¹⁴ For example, Table 5.4 shows the torsional angles for the GNA nucleotides in this new structure. The bond angles of the phosphate-oxygen bonds of each phosphate group determined solely from the electron density are shown in Table 5.5. These were obtained by refining the structure without any restraints on the bond angles, therefore allowing us to determine the values very accurately from the crystallographic data. The phosphate bond angles after refinement using the newly generated stereochemical library are also shown in parentheses in Table 5.5,

demonstrating the similarity between the unrefined and refined values for these bond angles. This will serve as the basis for solving and refining additional GNA duplex structures of different sequences and length in order to gain further insight into the exceptional duplex formation abilities of this minimal nucleic acid backbone.

Table 5.4. Backbone torsional angles.^[a]

Nucleotide	α	β	γ	δ	ϵ	ζ	η
G-1			-63	-97	-72	-65	-89
^{Br} C-2	-171	-151	-178	-92	-68	-56	-76
G-3	149	152	-68	-103	-95	-67	-91
C-4	-170	-119	-173	-109	-44	-62	-81
G-5	127	146	-68	-89	-76	-63	-87
C-6	-112	178	171			-59	-86

[a] Measured in the 3' to 2' direction



(S)-GNA

Table 5.5. Phosphate bond angles.^[a]

Linkage	O3G-P- O2G	O3G-P- O1P	O3G-P- O2P	O1P-P- O2P	O2G-P- O1P	O2G-P- O2P
G-1 – ^{Br} C-2	99.1 (99.6)	109.5 (107.0)	111.7 (112.1)	114.4 (116.3)	107.5 (108.8)	113.5 (111.7)
^{Br} C-2 – G-3	100.3 (100.4)	111.6 (105.8)	110.3 (110.5)	116.8 (117.1)	105.1 (111.5)	111.4 (110.2)
G-3 – C-4	99.9 (100.0)	110.8 (107.5)	110.9 (111.4)	114.1 (116.7)	107.8 (109.9)	112.4 (110.1)
C-4 – G-5	103.0 (103.6)	105.3 (100.6)	108.7 (108.0)	129.3 (126.1)	99.3 (107.0)	108.3 (109.2)
G-5 – C-6	104.7 (104.0)	111.6 (107.8)	107.4 (110.1)	113.7 (114.0)	106.5 (111.1)	112.6 (109.4)

[a] Determined by refining the structure with no restraints on bond angles. Bond angles after refinement using the new stereochemical library shown in parentheses.

With this new crystal structure in hand, it was interesting to attempt to solve the structure of the native sequence 3'-GCGCGC-2' (MKS5) via molecular replacement. Unfortunately, all attempts at molecular replacement failed, providing electron density that was poorly defined with respect to the model. This may be due to differences in the crystal forms, or the high crystallographic B factor of 64.5 Å² observed in the data (see appendix for complete crystallographic data). However, the overall structure is likely to be very similar owing to the fact that the 5-bromocytosine derivative is only a minimal perturbation of the native structure. Figure 5.8 shows the similar shapes of the CD spectra of the bromo-derivative (MKS132) and the native (MKS5) GNA duplex, confirming that the secondary structure of this bromo-derivative is similar to that of the native.

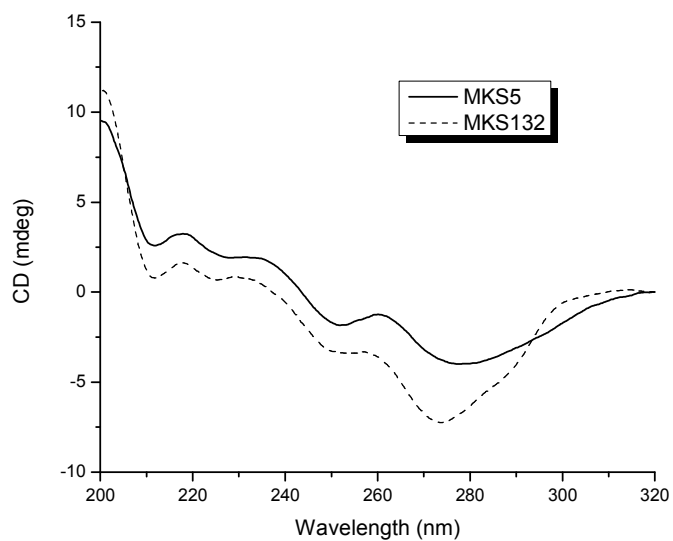


Figure 5.8. CD spectra of the self-complementary GNA duplexes 3'-GCGCGC-2' (MKS5) and 3'-G^{Br}CGCGC-2' (MKS132) at a duplex concentration of 20 μ M. Measurements were performed in 10 mM sodium phosphate buffer (pH=7.0) with 100 mM sodium chloride at 15°C. Each curve is the average of five measurements.

Chapter 5.4. Structure of the GNA duplex 3'-CTCTAGAG-2'

The second solved crystal structure of a GNA duplex was accomplished using a bromo-uracil derivatized glycol nucleotide (^{Br}U) for phasing. Initially, diffraction quality crystals (long rods) were obtained with the native sequence 3'-CTCTAGAG-2' and a data set was subsequently collected with diffraction up to 2.0 Å. At this point, several different GNA oligonucleotides were synthesized in an attempt to obtain phases for solving this structure. Both bromo-uracil derivatives, 3'-C^{Br}UCTAGAG-2' (MKS79) and 3'-CTC^{Br}UAGAG-2' (MKS80), were synthesized and screened for their crystallization properties (see Table 5.6). Fortunately, similar looking crystals were obtained under the same crystallization conditions for the sequence 3'-CTC^{Br}UAGAG-2' (MKS80) allowing for the collection of several data sets for MAD phasing. The data was processed and scaled in two different space groups, with the ultimate choice of I4₁22.

The initial map for the structure 3'-CTC^{Br}UAGAG-2' provided clear electron density for the backbone and all eight base pairs (see appendix for complete crystallographic data). During refinement of the structure, there was a significant portion of spherical electron density which resembled a duplex type structure. Unfortunately, the density could not be accommodated by either the oligonucleotide, or any components of the crystallization buffer. It was suspected that this density was the result of data collected on twinned crystals, however, all attempts with data manipulation to solve this problem failed. Therefore, we decided that although the data could not be refined to

provide good statistics, the electron density provided a model for the GNA duplex structure which should still be representative of the crystallized structure.

Table 5.6. GNA oligos synthesized to solve the structure of 3'-CTCTAGAG-2'

Name	Sequence	Molecular Weight	Outcome
MKS22	3'-CTCTAGAG-2'	2073	Solved structure
MKS79	3'-C ^{Br} UCTAGAG-2'	2138	Crystals obtained
MKS80	3'-CTC ^{Br} UAGAG-2'	2138	Solved Structure
MKS76	3'-C _{PSe} TCTAGAG-2'	2136	Only synthesized
MKS75	3'-CT _{PSe} CTAGAG-2'	2136	Only synthesized
MKS74/MKS95	3'-CTC _{PSe} TAGAG-2'	2136	Best separation of diastereomers, structure solved w/first diastereomer
MKS73	3'-CTCT _{PSe} AGAG-2'	2136	Only synthesized
MKS72	3'-CTCTA _{PSe} GAG-2'	2136	Only synthesized
MKS71	3'-CTCTAG _{PSe} AG-2'	2136	Only synthesized
MKS70	3'-CTCTAGA _{PSe} G-2'	2136	Poor synthesis

The structure of this GNA duplex formed from the self-complementary strand 3'-CTC^{Br}UAGAG-2' is shown in Figure 5.9. The helix is similar to a Type N helix, having an approximate pitch of 28 Å and approximately 10 residues per turn (obtained from modeling). Duplexes are stacked end-to-end, but have a staggered conformation rather than forming a continuous helix within the crystal. The helices are packed in a cubic fashion with the bromine atoms creating close contacts between neighboring duplexes (Figure 5.10). All base pairs are engaged in standard Watson-Crick base pairing with the

5-bromouracil appearing to have little distorting effect to the overall structure. The $C_{1'}$ - $C_{1'}$ distances range from 10.48-10.78 Å with an average of 10.63 Å for this duplex. Consistent with the previous Type N duplex, the nucleotides adopt alternating *gauche* and *anti* conformations, with average torsional angles γ of -56° and -167° , respectively. Finally, as with the 6-mer duplex structure, this duplex is also characterized by the large slide between neighboring base pairs resulting in mostly interstrand stacking, the packing of the propylene glycol backbone $C_{1'}$ - H_2 against the neighboring nucleobase, and the short distance between intrastrand phosphates.

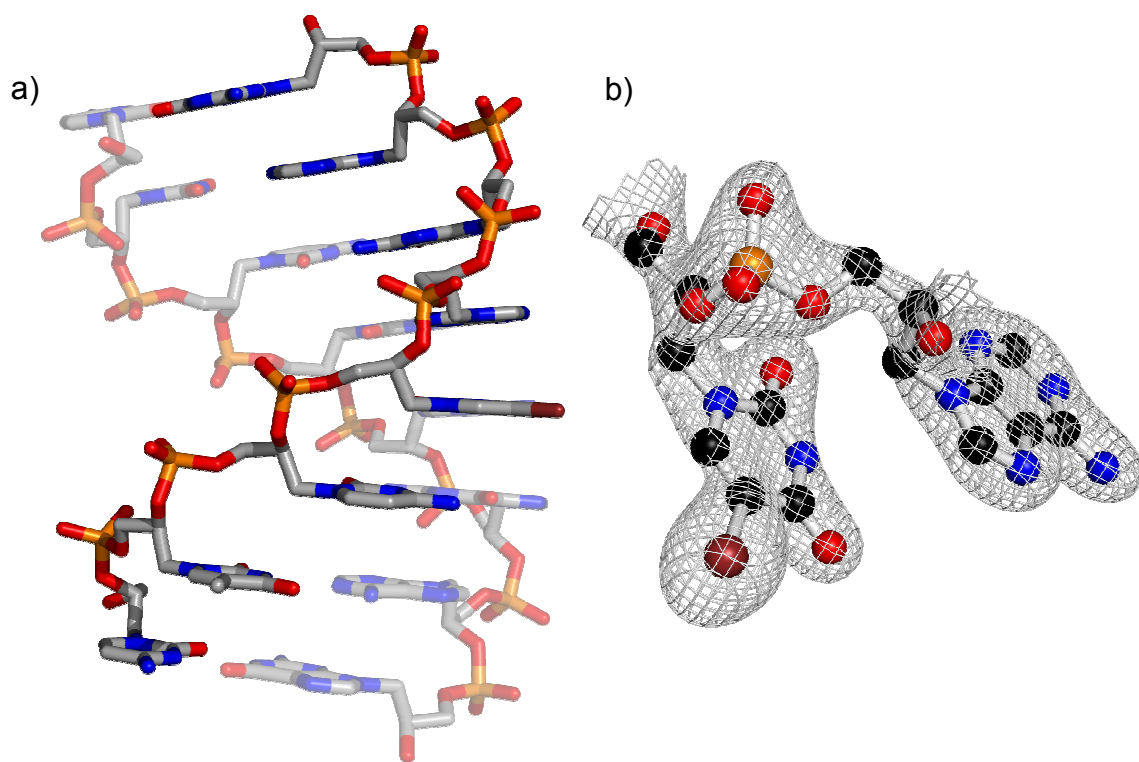


Figure 5.9. Structure of a single duplex of 3'-CTC^{Br}UAGAG-2' (a) and electron density of the center two glycol nucleotides (b) at 1.5 sigma. Figure generated using PyMOL.

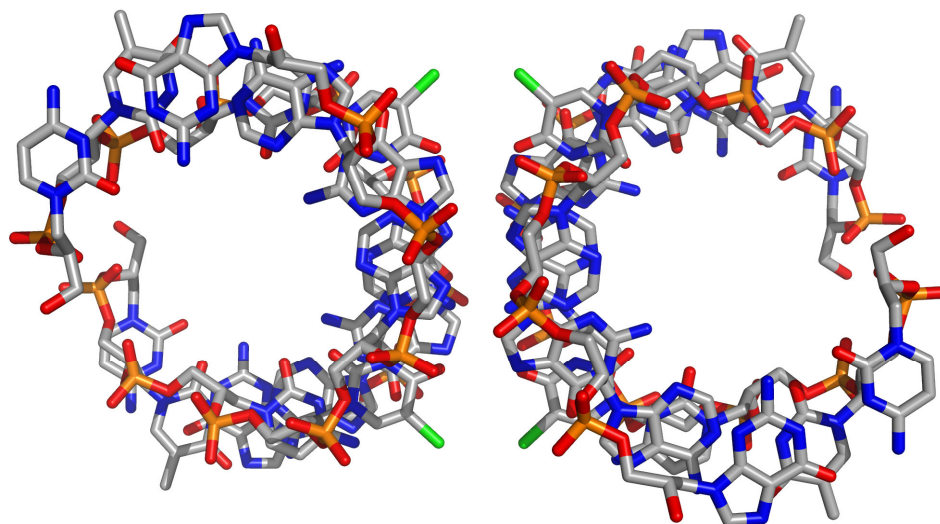


Figure 5.10. Packing contacts between two duplexes in the crystal structure of 3'-CTC^{Br}UAGAG-2'. Bromine atoms are highlighted in green and are separated by 3.66 angstroms.

During the same time the structure of the brominated derivative (MKS80) was being solved, phosphoroselenoate derivatives were also being used in an attempt to obtain phase information. Table 5.6 shows that six phosphoroselenoate derivatives of this sequence were synthesized, however, the sequence 3'-CTC_{PSe}TAGAG-2' (MKS74) was ultimately pursued for crystallography based on the ease in which the two diastereomers could be separated. As mentioned previously, the phosphoroselenoate linkage is slightly unstable in GNA, making it difficult to obtain pure material for crystallography. Nonetheless, a sufficiently pure sample of each diastereomer of MKS74 was obtained and crystals developed for the first diastereomer, MKS74-1, in similar conditions to the native (also containing cobalt hexamine as the polyamine). Data was collected on a single crystal at three different wavelengths after a fluorescence scan to

determine the precise selenium absorption edge (see appendix for complete crystallographic data). Unfortunately, the data did not provide a strong enough anomalous signal for phasing the data. However, since these two duplexes crystallized in the same space group with very similar cell constants, we were able to solve the structure via molecular replacement. There are very few differences between the structures of MKS80 and MKS74-1, but an anomalous density difference map calculated for these two data sets clearly shows the strong density for the selenium atom in the phosphoselenoate backbone of MKS74-1 (Figure 5.11). Furthermore, it demonstrates that the crystallized diastereomer was the one with an (*R*)- configuration around the phosphate.

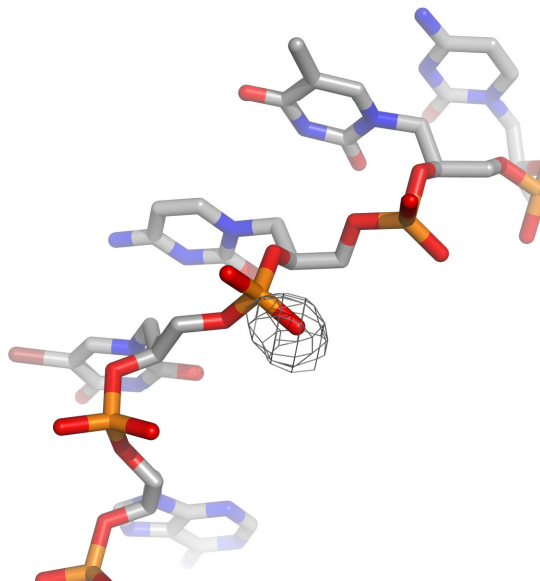


Figure 5.11. Anomalous difference density map at 5.0 sigma for peak data of 3'-CTC_{PS_e}TAGAG-2' using phases from the brominated derivative (MKS80). Figure generated using PyMOL.

As with the solved 6-mer (MKS132) structure, having a good model for the structure of 3'-CTC^{Br}UAGAG-2' was also advantageous for attempts at molecular replacement of the native structure, 3'-CTCTAGAG-2' (MKS22). Fortunately, these two GNA oligonucleotides crystallized in the same condition with the same space group and similar cell constants. This allowed us to simply perform molecular replacement by doing a refinement using the structure of 3'-CTC^{Br}UAGAG-2' and the processed data for MKS22. A clear solution emerged, providing the structure of 3'-CTCTAGAG-2' and acceptable electron density (see appendix for complete crystallographic data). However, as with the brominated derivative MKS80, the solution suffers from the presence of unknown helical electron density, making the refinement to provide good statistics difficult. In spite of this, the data provides a clear picture of the native duplex which is almost identical to the MKS80 duplex (Figure 5.12a). Furthermore, as also shown in Figure 5.12b, these two duplexes adopt a similar conformation to the high resolution structure of 3'-G^{Br}CGCGC-2' (MKS132), demonstrating the validity of these models.

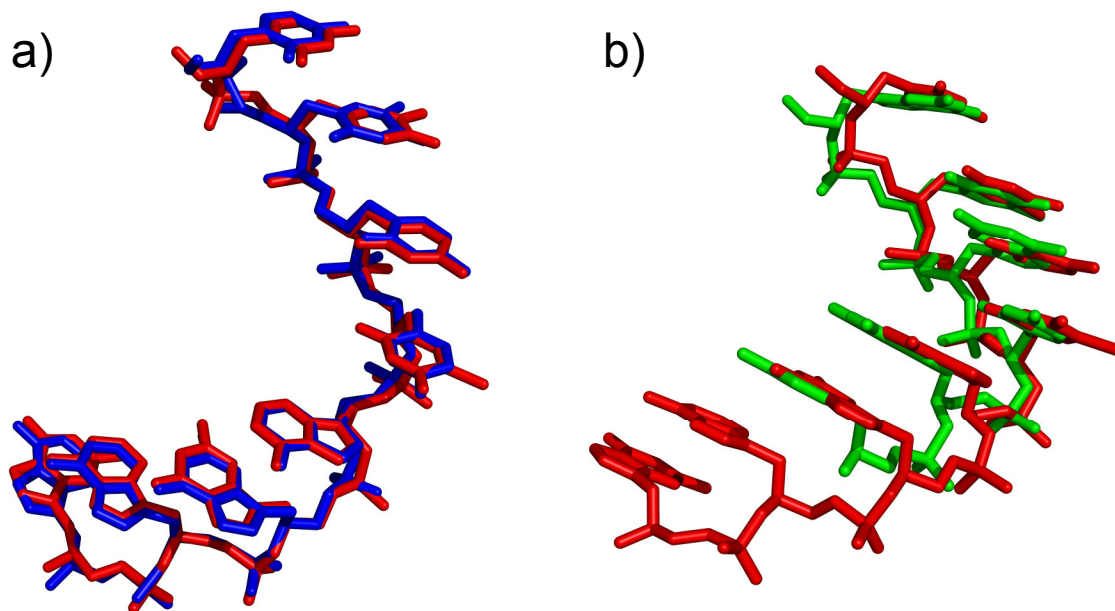


Figure 5.12. Superimposed models of: a) The structure of MKS80 (red) and MKS22 (blue). b) The structure of MKS80 (red) and MKS132 (green). Figures generated using PyMOL.

Chapter 5.5. Structure of the GNA duplex 3'-ATGCGCAT-2'

Another self-complementary GNA duplex, 3'-ATGCGCAT-2' (MKS42), provided diffraction quality crystals in many different conditions. The best diffracting crystals (square plates) were only obtained when cobalt hexamine was included in the crystallization buffer as has been previously observed with other sequences. Data was collected up to 2.00 Å for these crystals and processed in the $P2_12_12_1$ spacegroup. Thereafter, many different GNA oligonucleotide derivatives were synthesized in an attempt to collect phase information. Both oligonucleotides containing bromo-uracil, both containing bromo-cytosine, and one phosphorselenoate derivative were synthesized as shown in Table 5.7. Although crystals were obtained for all five of these derivatives, none of them provided diffraction with sufficient quality for solving the structure.

Table 5.7. GNA oligos synthesized to solve the structure of 3'-ATGCGCAT-2'

Name	Sequence	Molecular Weight	Outcome
MKS39/MKS42	3'-ATGCGCAT-2'	2073	Solution generated using MR
MKS52/MKS121	3'-A ^{Br} UGCGCAT-2'	2138	Crystals obtained
MKS116/MKS134	3'-ATG ^{Br} CGCAT-2'	2152	Crystals obtained
MKS115/MKS133	3'-ATGCG ^{Br} CAT-2'	2152	Crystals obtained
MKS54/MKS122	3'-ATGCGCA ^{Br} U-2'	2138	Crystals obtained
MKS60	3'-A _{PSe} TGCGCAT-2'	2136	Crystals obtained

At the time, based on the recently solved crystal structure of the 8-mer duplexes with the general sequence 3'-CTCTAGAG-2', molecular replacement was used in an attempt to solve the crystal structure of 3'-ATGCGCAT-2'. An initial search model based on the crystal structure of MKS80 was generated by simply reconstructing the appropriate base pairs. Molecular replacement with two duplexes per asymmetric symmetry unit provided an initial map with electron density for all phosphates and all but one terminal adenine of this duplex (see appendix for complete crystallographic data). The density of the propylene glycol backbone is absent in some cases, rendering this only an approximate structural solution. However, each strongly diffracting phosphate group is represented by prominent electron density in the map. This results in data that does not provide brilliant statistics, but nonetheless gives us more structural insight and a model for this sequence 3'-ATGCGCAT-2'.

The structural model for the self-complementary GNA duplex 3'-ATGCGCAT-2' is shown in Figure 5.12. One can observe, again, that the main features associated with GNA duplexes are present in this structure. These include a large x-displacement of the base pairs from the helix axis, strong interstrand stacking, and an alternation between an *anti* and *gauche* conformation of the propylene glycol backbone. As noted above, the electron density of the backbone is absent in some cases, so the backbone conformations cannot be assigned definitively for each case and the alternating conformation may just be representative of bias towards the model used for molecular replacement. Unfortunately, since the electron density for all non-hydrogen atoms is not present for this structure, one can not gain a whole lot of insight from it. The poor electron density

may be due to inherent disorder in the crystal based on the high crystallographic temperature factor (B_{Wilson}) of 38.3 \AA^2 . This also manifests itself during refinement, with an average B-factor of 58.7 \AA^2 .

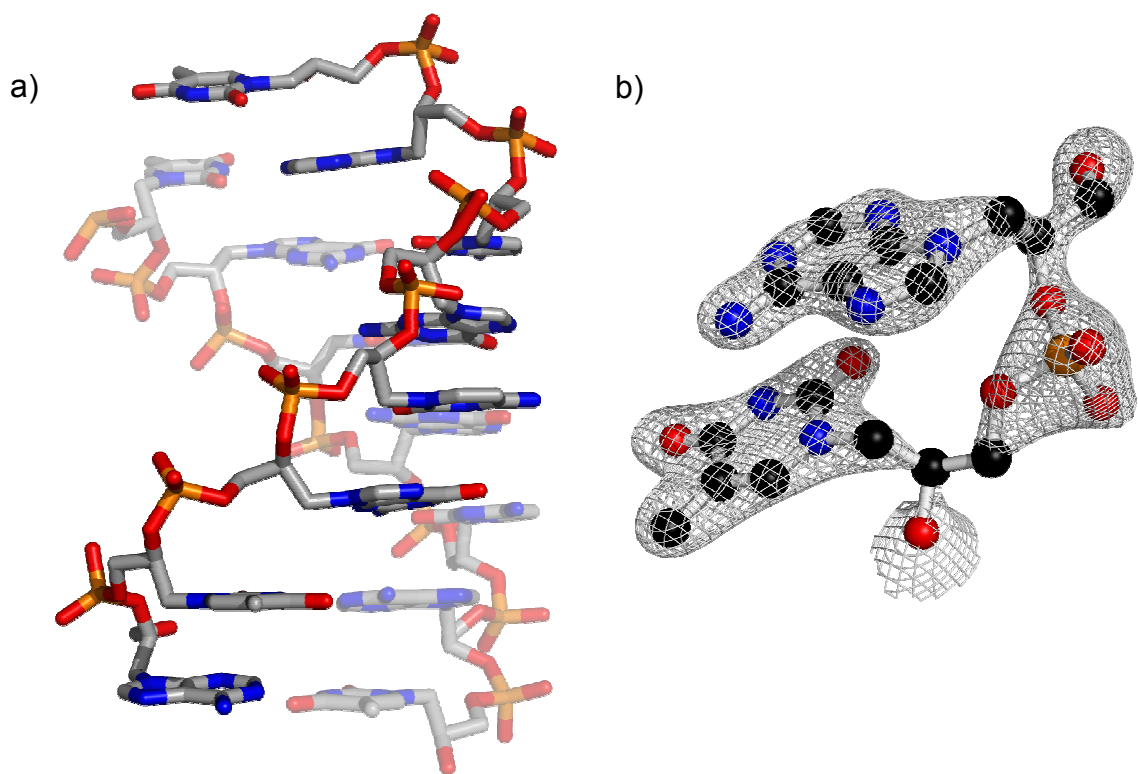


Figure 5.13. Structure of a single duplex of 3'-ATGCGCAT-2' missing one terminal adenine nucleotide (a) and electron density of the terminal two nucleotides (b) at 1.5 sigma. Figure generated using PyMOL.

Chapter 5.6. Conclusions

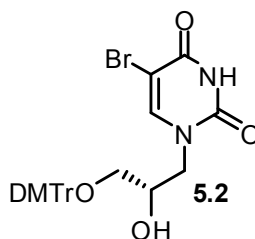
Significant progress has been made in the crystallography of glycol nucleic acid duplexes. Unfortunately, progress has been slow because of the inherent phase problem that comes with solving de novo crystal structures. In most cases, it was not an issue of being able to obtain crystals, as can be shown by the number of data sets that have been collected over the past couple of years. Some of the crystals were of poor quality, rendering solution of the structure difficult, but optimization of the crystallization conditions would easily solve this issue. However, the development of the 5-bromocytosine (^{Br}C*) and 5-bromouracil (^{Br}U) phosphoramidites have allowed us to gain increased structural insight into the nucleic acid duplexes formed by GNA. There are mixed reports as to whether bromine derivatives of nucleotides have a significant impact on the ability of nucleic acid duplexes to crystallize.^{1,19-21} From our experience the crystallization behavior of the bromine derivatives varied depending on the sequence. However, close inspection of the two crystal structures solved using bromo-derivatives (MKS80 and MKS132) suggest that, in the case of these two GNA structures, the bromine derivatives actually aid in crystallization by forming contacts between neighboring duplexes within the crystal.

The five crystal structures of GNA duplexes containing exclusively Watson-Crick base pairs presented here (MKS22, MKS46, MKS74-1, MKS80, and MKS132) all maintain very similar structural features. The large x-displacement values, strong slide of neighboring base pairs, large backbone-base inclinations, and the stacking of the C_{1'}-CH₂

group on the neighboring base are all consistent with the previously reported structure of a GNA duplex containing metal-mediated base pairs (MKS64).⁸ However, there are two significant differences between the Type N and Type M structures. Figure 5.6 demonstrates that the Type N helix is compressed along the z-axis, resulting in a much shallower pitch with fewer base pairs per complete turn. Furthermore, the Type N structures adopt alternating *gauche* and *anti* conformations with respect to the backbone, in contrast to the Type M structure in which all Watson-Crick nucleotides adopt a *gauche* conformation and the hydroxypyridone nucleotides adopt an *anti* conformation. Finally, the high resolution of the structure 3'-G^{Br}CGCGC-2' allows us to tabulate accurate values for bond lengths, torsional angles, and phosphate bond angles for the refinement of future glycol nucleic acid structures.

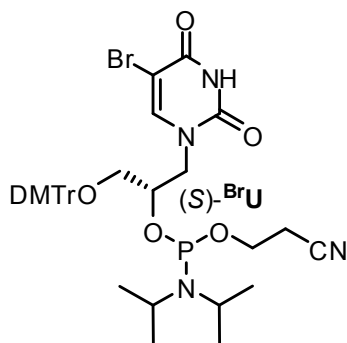
Chapter 5.7. Experimental procedures

General procedures and reagents. NMR spectra were recorded on a Bruker DMX-300 (300 MHz) spectrometer. High-resolution mass spectra were obtained with a Micromass AutoSpec or Thermo LTQ-FT instrument using ES ionization. Infrared spectra were recorded either on a Bruker alpha series FTIR spectrometer. Solvents and reagents were used as supplied from Aldrich, Acros, Fluka, or TCI. Reactions were performed under an atmosphere of argon or nitrogen unless otherwise specified.

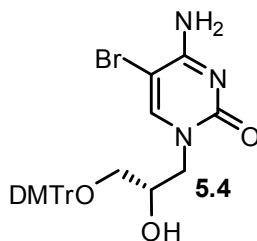


Compound 5.2. To a suspension of 5-bromouracil (1.04 g, 5.5 mmol) in anhydrous DMF (11 mL) under argon was added NaH (44 mg, 1.1 mmol, 60% in mineral oil) and the solution was allowed to stir under argon for one hour. A solution of compound **2.11** (1.95 g, 5.2 mmol) in DMF (11 mL) was added to the first solution and then heated to 85 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via flash chromatography over silica gel starting with 2:1:0.01 Hexanes:Acetone:Et₃N, then 3:2:0.01 Hexanes:Acetone:Et₃N, and finally eluting with 1:1:0.01 Hexanes:Acetone:Et₃N to afford compound **5.2** as a white foam (1.32 g, 43%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 7.65 (s, 1H), 7.46 (d, *J* = 7.4 Hz, 2H), 7.39-

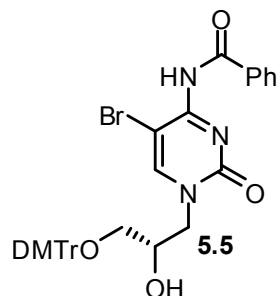
7.17 (m, 7H), 6.85 (d, $J = 8.8$ Hz, 4H), 4.14 (m, 2H), 3.79 (s, 6H), 3.64 (dd, $J = 14.5, 8.6$ Hz, 1H), 3.21 (d, $J = 4.5$ Hz, 2H). ^{13}C -NMR (75 MHz, CDCl_3) δ (ppm) 160.0, 158.6, 151.0, 145.8, 144.6, 135.7, 130.0, 128.07, 128.01, 127.0, 113.3, 95.8, 86.5, 68.7, 64.6, 55.3, 52.3. IR (solid) ν (cm^{-1}) = 3439 (br), 3168 (br), 3059, 2931, 2835, 1675, 1606, 1506, 1443, 1347, 1300, 1245, 1174, 1070, 1029, 906, 826, 727, 701, 621, 582, 526, 423. HRMS calcd for $\text{C}_{28}\text{H}_{27}\text{N}_2\text{O}_6\text{BrNa}$ ($\text{M}+\text{Na}$) $^+$ 589.0945, found ($\text{M}+\text{Na}$) $^+$ 589.0956.



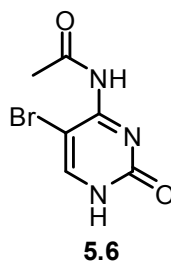
Compound (S)-BrU. To an argon purged solution of compound **5.2** (1.00 g, 1.8 mmol) and *N,N*-diisopropylethylamine (1.85 mL, 10.6 mmol) in methylene chloride (30 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.59 mL, 2.6 mmol) dropwise and the solution stirred for two hours at room temperature under argon. The solution was diluted with methylene chloride and washed once with saturated aqueous NaHCO_3 , dried over Na_2SO_4 , and finally concentrated. The crude product was purified by flash chromatography over silica gel using 3:2:0.01 Hexanes:Acetone: Et_3N to afford compound (S)-BrU as a white foam (1.05 g, 78%). ^{31}P -NMR (162 MHz, CDCl_3) δ (ppm) 150.6, 150.2. HRMS calcd for $\text{C}_{37}\text{H}_{45}\text{N}_4\text{O}_7\text{BrP}$ ($\text{M}+\text{H}$) $^+$ 767.2204, found ($\text{M}+\text{H}$) $^+$ 767.2206.



Compound 5.4. To a suspension of 5-bromocytosine (1.00 g, 5.3 mmol) in anhydrous DMF (10.5 mL) under argon was added NaH (42 mg, 1.1 mmol, 60% in mineral oil) and the solution was allowed to stir under argon for one hour. A solution of compound **2.11** (1.88 g, 5.0 mmol) in DMF (10.5 mL) was added to the first solution and then heated to 100 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via flash chromatography over silica gel starting with 100:1 EtOAc:Et₃N, then eluting with 50:1:0.01 EtOAc:MeOH:Et₃N to afford compound **5.4** as a light yellow solid (1.7 g, 60%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 7.87 (br, 1H), 7.52 (s, 1H), 7.41 (m, 2H), 7.29 (m, 6H), 7.21 (m, 1H), 6.83 (m, 4H), 5.64 (br, 1H), 4.24 (dd, *J* = 13.8, 2.5 Hz, 1H), 3.78 (s, 6H), 3.73 (dd, *J* = 14.0, 6.6 Hz, 1H), 3.23 (dd, *J* = 9.7, 5.1 Hz, 1H), 3.02 (dd, *J* = 9.7, 6.4 Hz, 1H). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm) 162.8, 158.7, 156.8, 147.4, 144.8, 135.9, 135.8, 130.0, 128.10, 128.03, 127.1, 113.4, 86.8, 86.6, 69.5, 64.5, 55.4, 54.3. IR (solid) ν (cm⁻¹) = 3442 (br), 2989, 2836, 1679, 1656, 1597, 1444, 1348, 1299, 1246, 1175, 1071, 1031, 948, 827, 727, 701, 622, 582, 417. HRMS calcd for C₂₈H₂₈N₃O₅BrNa (M+Na)⁺ 588.1105, found (M+Na)⁺ 588.1115.

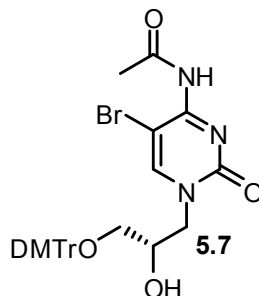


Attempted synthesis of compound 5.5. To an argon purged solution of compound **5.4** (260 mg, 0.46 mmol) in anhydrous pyridine (3.7 mL) was added trimethylsilyl chloride (0.23 mL, 1.8 mmol) and the solution allowed to stir under argon for two hours. The solution was then cooled to 0 °C, benzoyl chloride (80 μL, 0.69 mmol) added dropwise, and the solution allowed to warm up gradually to room temperature and stir for another two hours. The solution was cooled again to 0 °C and quenched by the addition of water (1 mL) and stirred for 15 minutes. Afterwards 25% aqueous ammonium hydroxide (2 mL) was added and stirring continued for another 30 minutes. All solvent was then removed, the oil coevaporated with toluene, and the crude product purified via flash chromatography over silica gel starting with 3:2:0.01 Hexanes:Acetone:Et₃N, then eluting with 1:1:0.01 Hexanes:Acetone:Et₃N. ¹H NMR indicated multiple products.

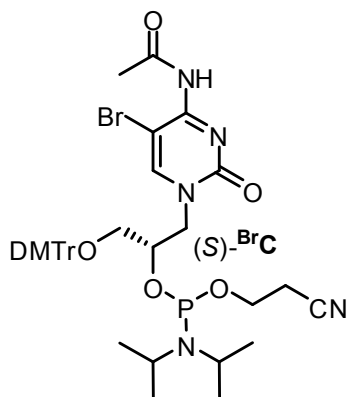


Compound 5.6. To a suspension of 5-bromocytosine (750 mg, 3.95 mmol) in anhydrous DMF (4.9 mL) was added acetic anhydride (1.10 mL, 11.8 mmol) and the suspension heated to 100 °C for 1 hour. After cooling to room temperature, all the solvent was

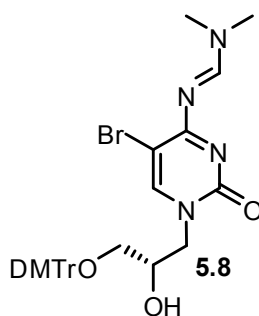
removed and the solid resuspended in 1:1 Hexanes:EtOAc and put in the -20 °C freezer for 30 minutes. The solid was filtered, washed three times with ice cold 1:1 Hexanes:EtOAc, and then three times with diethyl ether. The light yellow solid was used as crude for the next reaction (890 mg, 97%). ¹H-NMR (300 MHz, DMSO-*d*₆) δ (ppm) 12.01 (br, 1H), 9.67 (br, 1H), 8.20 (s, 1H), 2.22 (s, 3H).



Compound 5.7. To a suspension of compound **5.6** (890 mg, 3.8 mmol) in anhydrous DMF (7.5 mL) under nitrogen was added NaH (31 mg, 0.80 mmol, 60% in mineral oil) and the solution was allowed to stir under argon for one hour. A solution of compound **2.11** (1.38 g, 3.6 mmol) in DMF (7.5 mL) was added to the first solution and then heated to 110 °C overnight. The next morning, the solution was cooled, all solvent removed, the resulting oil coevaporated with toluene, redissolved in ethyl acetate and concentrated again. The product was purified via flash chromatography over silica gel starting with 2:1:0.01 Hexanes:Acetone:Et₃N, then eluting with 3:2:0.01 Hexanes:Acetone:Et₃N to afford compound **5.7** as a light yellow foam (290 mg, 12%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 7.78 (s, 1H), 7.41 (m, 2H), 7.34-7.19 (m, 7H), 6.83 (m, 4H), 4.29 (dd, *J* = 13.7, 2.6 Hz, 1H), 4.16 (m, 1H), 3.82-3.70 (m, 7H), 3.22 (dd, *J* = 9.8, 5.5 Hz, 1H), 3.14 (dd, *J* = 9.7, 5.6 Hz, 1H), 2.64 (s, 3H).

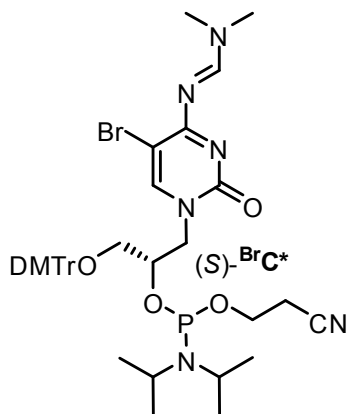


Compound (S)-^{Br}C. To a nitrogen purged solution of compound **5.7** (290 mg, 0.48 mmol) and *N,N*-diisopropylethylamine (0.50 mL, 2.9 mmol) in methylene chloride (8.0 mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.16 mL, 0.71 mmol) dropwise and the solution stirred for two hours at room temperature under nitrogen. The solution was diluted with methylene chloride and washed once with saturated aqueous NaHCO₃, dried over Na₂SO₄, and finally concentrated. The crude product was purified via flash chromatography over silica gel using 2:1:0.01 Hexanes:Acetone:Et₃N to afford compound (S)-^{Br}C as a white foam (300 mg, 78%). ³¹P-NMR (121 MHz, CDCl₃) δ (ppm) 149.1, 148.6.



Compound 5.8. To a solution of compound **5.4** (435 mg, 0.77 mmol) in 2.5 mL of MeOH was added dimethylformamide-dimethylacetal (360 μL, 2.69 mmol) and heated to

50 °C for one hour. After cooling and removal of the MeOH, the residue was redissolved in methylene chloride, washed once with water, dried over Na₂SO₄, and finally concentrated. The product was purified via flash chromatography over silica gel starting with 100:1 EtOAc:Et₃N, then eluting with 20:1:0.01 EtOAc:MeOH:Et₃N to afford compound **5.8** as a white foam (470 mg, 98%). ¹H-NMR (300 MHz, CDCl₃) δ (ppm) 8.72 (s, 1H), 7.64 (s, 1H), 7.42 (m, 2H), 7.29 (m, 6H), 7.21 (m, 1H), 6.83 (m, 4H), 4.29 (dd, *J* = 13.9, 2.4 Hz, 1H), 4.23 (d, *J* = 4.8 Hz, 1H), 4.13 (br, 1H), 3.89-3.76 (m, 7H), 3.28 (dd, *J* = 9.7, 5.3 Hz, 1H), 3.21 (s, 3H), 3.18 (s, 3H), 2.98 (dd, *J* = 9.6, 6.9 Hz, 1H). ¹³C-NMR (75 MHz, CDCl₃) δ (ppm) 168.1, 159.0, 158.7, 157.7, 147.6, 144.8, 136.0, 135.8, 130.0, 128.09, 128.01, 127.0, 113.4, 96.8, 86.6, 70.0, 64.4, 55.3, 54.6, 41.6, 35.6. IR (solid) ν (cm⁻¹) = 3269 (br), 2930, 2835, 1651, 1607, 1586, 1484, 1447, 1383, 1347, 1299, 1246, 1173, 1114, 1070, 1030, 982, 905, 827, 777, 725, 701, 643, 582. HRMS calcd for C₃₁H₃₄N₄O₅Br (M+H)⁺ 621.1707, found (M+H)⁺ 621.1701.



Compound (S)-BrC*. To a solution of compound **5.8** (450 mg, 0.72 mmol) in 3.6 mL of anhydrous methylene chloride under nitrogen was added a 1 M solution of 4,5-dicyanoimidazole (0.51 mL in acetonitrile). 2-cyanoethyl *N,N,N',N'*-

tetraisopropylphosphordiamidite (0.24 mL, 0.76 mmol) was then added dropwise and the solution stirred at room temperature. After two hours, the reaction mixture was diluted with methylene chloride, washed twice with saturated aqueous NaHCO₃, dried over Na₂SO₄, and then concentrated. The product was purified via flash chromatography over silica gel starting with 3:2:0.01 Hexanes:Acetone:Et₃N, then eluting with 1:1:0.01 Hexanes:Acetone:Et₃N to afford compound (S)-^{Br}C* as a white foam (460 mg, 77%). ³¹P-NMR (162 MHz, CDCl₃) δ (ppm) 150.1, 149.9. HRMS calcd for C₄₀H₅₁N₆O₆BrP (M+H)⁺ 821.2786, found (M+H)⁺ 821.2814.

Crystallization and data collection of MKS5 and MKS132: Crystals of self-complementary duplex GNA were grown using the sitting drop vapor diffusion method with buffers from the Nucleic Acid Mini Screen (Hampton Research). Crystallization conditions consisted of 1 (MKS5) or 2 mM (MKS132) duplex GNA (2 μL) and buffer (4 μL) against a reservoir of 35% MPD in water (1 mL). The brominated derivative (3'-G^{Br}CGCGC-2') crystallized as rectangular rods which appeared after 1-2 weeks at 4 °C in Buffer #2 consisting of 10% 2-methyl-2,4-pentanediol, 40 mM sodium cacodylate (pH = 5.5), 20 mM cobalt hexamine, 80 mM sodium chloride, and 20 mM magnesium chloride. The native sequence (3'-GCGCGC-2') also crystallized as rectangular rods and appeared within 1 week at 4 °C in Buffer #20 consisting of 10% 2-methyl-2,4-pentanediol, 40 mM sodium cacodylate (pH = 7.0), 12 mM spermine tetra-HCl, 80 mM sodium chloride, and 20 mM barium chloride. All crystals were cryoprotected by raising the concentration of MPD in the buffer to 30% and subsequently picked from the drop

with nylon loops and frozen in liquid N₂. Data was recorded on a single crystal of each GNA oligonucleotide at beamline ID29, ESRF in Grenoble. After performing a fluorescence scan to determine the precise bromine absorption edge in the crystal, SAD data was collected with separate scans for high and low resolution reflections for the brominated derivative. All data was integrated and merged using XDS and phased by SHELXE and SHARP. The initial map of the brominated derivative provided density that was unambiguous for all the bases and phosphates of the duplex. Automated and manual refinements were performed using REFMAC⁵²² and COOT.²³

Crystallization and data collection of MKS22, MKS74-1, and MKS80: Crystals of self-complementary duplex GNA were grown using the sitting drop vapor diffusion method with buffers from the Nucleic Acid Mini Screen (Hampton Research). Crystallization conditions consisted of 1 mM duplex GNA (2 μ L) and buffer (4 μ L) against a reservoir of 35% MPD in water (1 mL). The native (MKS22, 3'-CTCTAGAG-2') and brominated (MKS80, 3'-CTC^{Br}UAGAG-2') duplexes crystallized as long rectangular rods which appeared within one week at 4 °C in Buffer #4 consisting of 10% 2-methyl-2,4-pentanediol, 40 mM sodium cacodylate (pH = 5.5), 20 mM cobalt hexamine, 40 mM lithium chloride, and 20 mM magnesium chloride. The selenium derivative (MKS74-1, 3'-CTC_{PSe}TAGAG-2') also crystallized as long rectangular rods and appeared within 1 week at 4 °C in Buffer #2 consisting of 10% 2-methyl-2,4-pentanediol, 40 mM sodium cacodylate (pH = 5.5), 20 mM cobalt hexamine, 80 mM sodium chloride, and 20 mM magnesium chloride. All crystals were cryoprotected by

raising the concentration of MPD in the buffer to 30% and subsequently picked from the drop with nylon loops and frozen in liquid N₂. Data was recorded on a single crystal of each GNA oligonucleotide at beamline A1, CHESS (MKS22, MKS74-1) or ID23-1, ESRF in Grenoble (MKS80). A fluorescence scan was performed to determine the precise bromine or selenium absorption edges in the crystal of MKS80 or MKS74-1, respectively. Data was collected at three different wavelengths for both derivatives, with separate scans for high and low resolution reflections for the brominated derivative. All data was integrated and merged using XDS and phased by SHELXE and SHARP. The initial map of the brominated derivative provided density that was unambiguous for all the bases and phosphates of the duplex. However, there was also a significant amount of helical electron density that could not be filled with any component of the oligonucleotide or crystallization buffer. Automated and manual refinements were performed using REFMAC5 and COOT.

Crystallization and data collection of MKS42: Crystals of the self-complementary duplex GNA were grown using the sitting drop vapor diffusion method with buffers from the Nucleic Acid Mini Screen (Hampton Research). Crystallization conditions consisted of 1 mM duplex GNA (2 μ L) and buffer (4 μ L) against a reservoir of 35% MPD in water (1 mL). The native sequence (3'-ATGCGCAT-2') crystallized as rectangular plates which appeared within one week at 4 °C in Buffer #4 consisting of 10% 2-methyl-2,4-pentanediol, 40 mM sodium cacodylate (pH = 5.5), 20 mM cobalt hexamine, 40 mM lithium chloride, and 20 mM magnesium chloride. Crystals were cryoprotected by

raising the concentration of MPD in the buffer to 30% and subsequently picked from the drop with a nylon loop and frozen in liquid N₂. Data was recorded on a single crystal of the GNA oligonucleotide at beamline ID29, ESRF in Grenoble. All data was integrated and merged using XDS. An initial search model for molecular replacement was obtained by reconfiguring the base pairs from the crystal structure of MKS80. The PHASER search provided a solution containing two duplexes per asymmetric symmetry unit. After an initial rigid body refinement, subsequent automated and manual refinements were performed using REFMAC5 and COOT. The electron density could not be completely represented by the duplex model with the terminal adenine and several backbone carbon atoms missing.

Crystallization and data collection of MKS111: Crystals of the self-complementary duplex GNA were grown using the sitting drop vapor diffusion method with buffers from the Nucleic Acid Mini Screen (Hampton Research). Crystallization conditions consisted of 1 mM duplex GNA (2 μL) and buffer (4 μL) against a reservoir of 35% MPD in water (1 mL). The native sequence (3'-CGCAAATTTGCG-2') crystallized as rectangular plates which appeared within one week at 4 °C in Buffer #2 consisting of 10% 2-methyl-2,4-pentanediol, 40 mM sodium cacodylate (pH = 5.5), 20 mM cobalt hexamine, 80 mM sodium chloride, and 20 mM magnesium chloride. Crystals were cryoprotected by raising the concentration of MPD in the buffer to 30% and subsequently picked from the drop with a nylon loop and frozen in liquid N₂. Data was recorded on a single crystal of

the GNA oligonucleotide at beamline ID23-1, ESRF in Grenoble. All data was integrated and merged using XDS.

Chapter 5.8. References

- (1) Egli, M. *Current Opinion in Chemical Biology* **2004**, *8*, 580.
- (2) Wilds, C. J.; Pattanayek, R.; Pan, C.; Wawrzak, Z.; Egli, M. *Journal of the American Chemical Society* **2002**, *124*, 14910-14916.
- (3) Egli, M.; Pallan, P. S.; Pattanayek, R.; Wilds, C. J.; Lubini, P.; Minasov, G.; Dobler, M.; Leumann, C. J.; Eschenmoser, A. *Journal of the American Chemical Society* **2006**, *128*, 10847-10856.
- (4) Salon, J.; Sheng, J.; Jiang, J.; Chen, G.; Caton-Williams, J.; Huang, Z. *Journal of the American Chemical Society* **2007**, *129*, 4862-4863.
- (5) Salon, J.; Jiang, J.; Sheng, J.; Gerlits, O. O.; Huang, Z. *Nucl. Acids Res.* **2008**, *36*, 7009-7018.
- (6) Sheng, J.; Jiang, J.; Salon, J.; Huang, Z. *Organic Letters* **2007**, *9*, 749-752.
- (7) Atwell, S.; Meggers, E.; Spraggon, G.; Schultz, P. G. *J. Am. Chem. Soc.* **2001**, *123*, 12364-12367.
- (8) Schlegel, M. K.; Essen, L.-O.; Meggers, E. *J. Am. Chem. Soc.* **2008**, *130*, 8158-8159.
- (9) Schlegel, M. K.; Meggers, E. *Journal of Organic Chemistry* **2009**, *74*, 4615-4618.
- (10) Mori, K.; Boiziau, C.; Cazenave, C.; Matsukura, M.; Subasinghe, C.; Cohen, J. S.; Broder, S.; Toulme, J. J.; Stein, C. A. *Nucl. Acids Res.* **1989**, *17*, 8207-8219.
- (11) Holloway, G. A.; Pavot, C.; Scaringe, S. A.; Lu, Y.; Rauchfuss, T. B. *ChemBioChem* **2002**, *3*, 1061-1065.

- (12) Pallan, P. S.; Egli, M. *Nat. Protocols* **2007**, 2, 640.
- (13) Bollmark, M.; Stawinski, J. *Chemical Communications* **2001**, 771-772.
- (14) Schlegel, M. K.; Essen, L.-O.; Meggers, E. **2009**, In Preparation.
- (15) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, 6, 63-91.
- (16) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, 6, 655-667.
- (17) Dickerson, R. E. *Methods Enzymol.* **1992**, 211, 67-111.
- (18) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X.-J.; Neidle, S.; Shakked, Z.; Sklenar, H.; Suzuki, M.; Tung, C.-S.; Westhof, E.; Wolberger, C.; Berman, H. M. *Journal of Molecular Biology* **2001**, 313, 229.
- (19) Tippin, D. B.; Sundaralingam, M. *Journal of Molecular Biology* **1997**, 267, 1171.
- (20) Ortiz-Lombardia, M.; Gonzalez, A.; Eritja, R.; Aymami, J.; Azorin, F.; Coll, M. *Nat Struct Mol Biol* **1999**, 6, 913.
- (21) Rhee, S.; Han, Z.-j.; Liu, K.; Miles, H. T.; Davies, D. R. *Biochemistry* **1999**, 38, 16810-16815.
- (22) *Acta Crystallog. Sect. D* **1994**, 50, 760-763.
- (23) Emsley, P.; Cowtan, K. *Acta Crystallog. Sect. D* **2004**, 60, 2126-2132.

Appendix to Chapter 5

^1H , ^{13}C , and ^{31}P NMR spectra

IR spectra

Crystallographic Tables

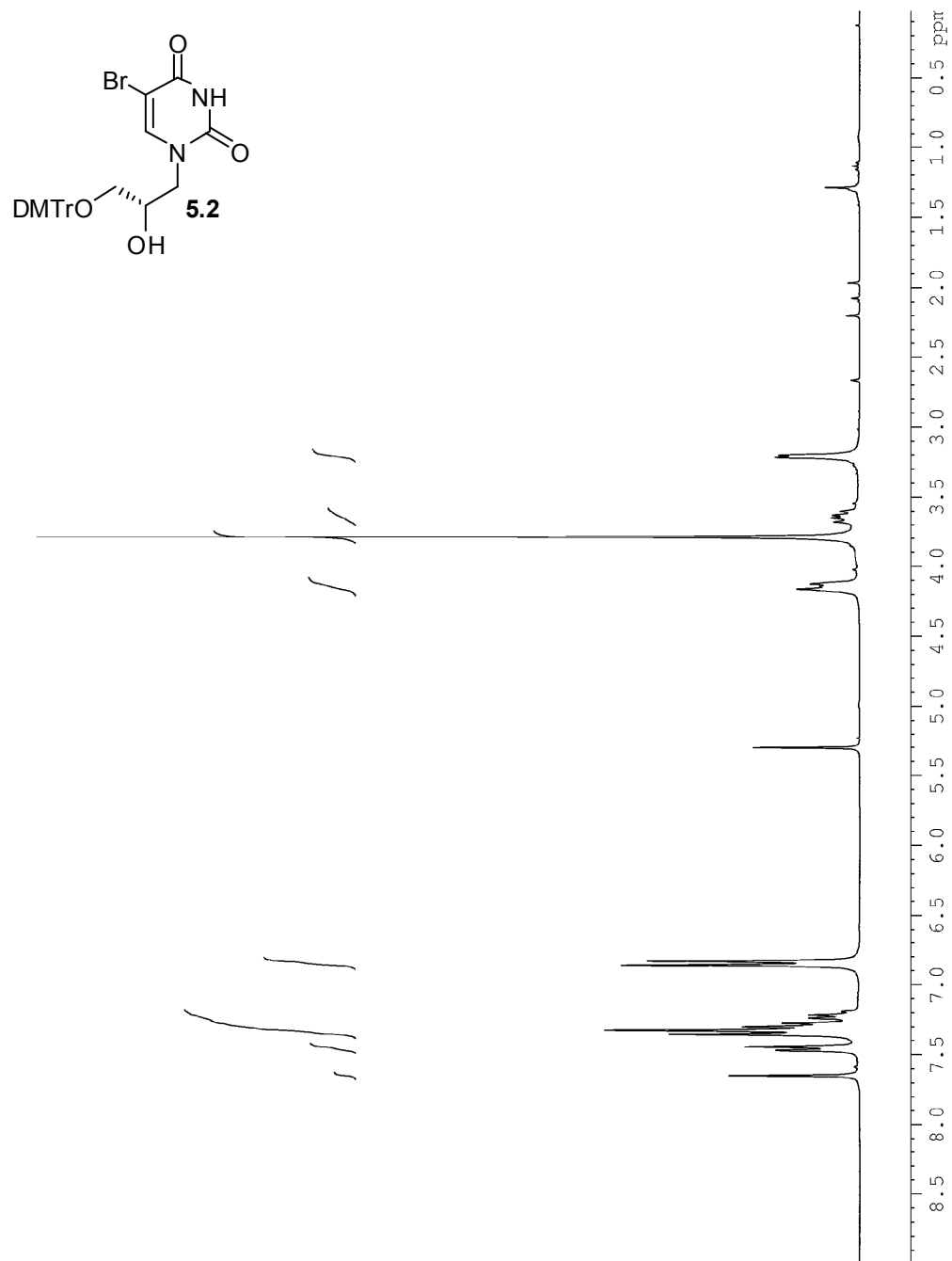


Figure A5.1.1. ^1H NMR spectrum of compound **5.2** (300 MHz, CDCl_3).

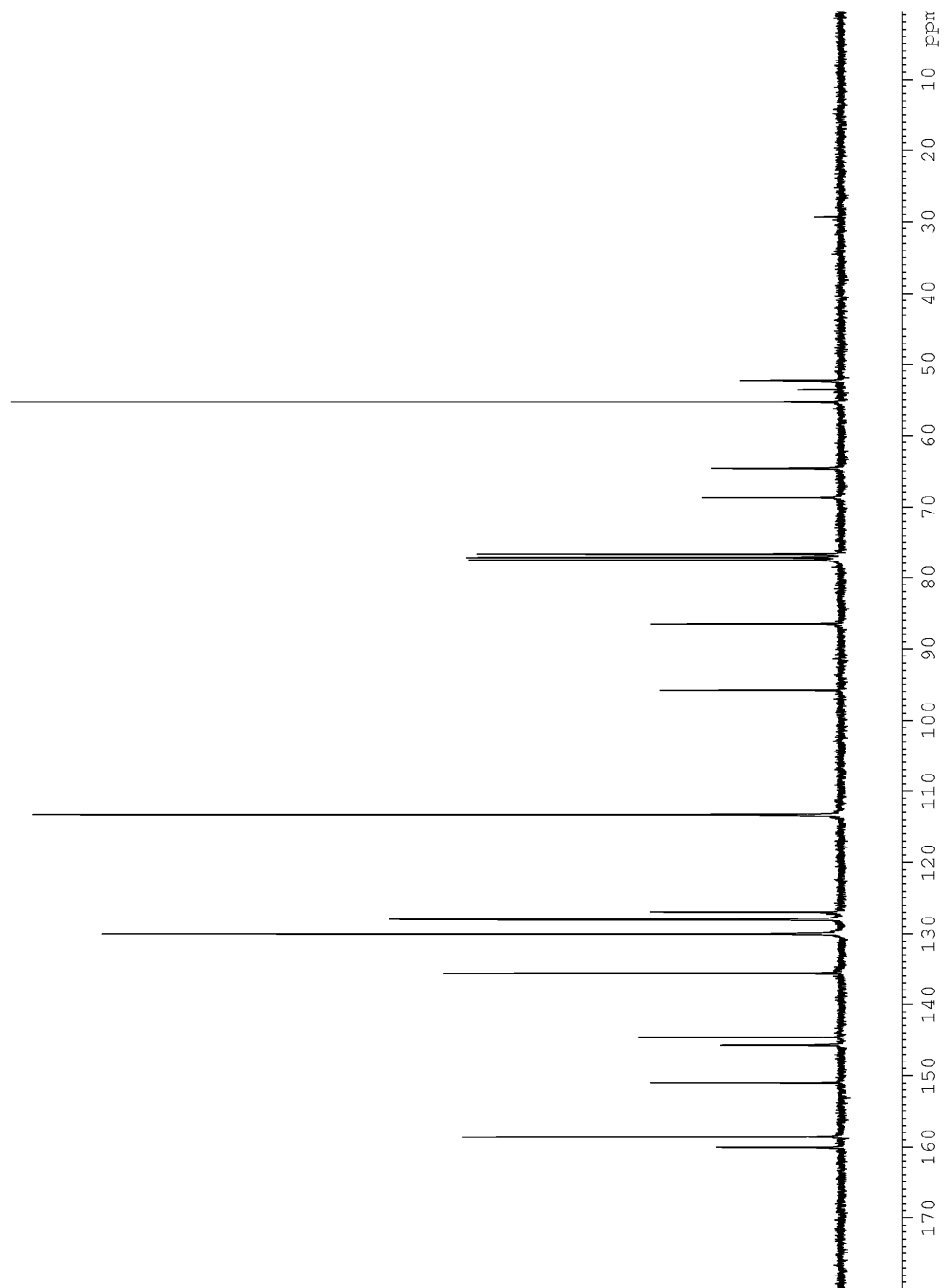


Figure A5.1.2. ^{13}C NMR spectrum of compound **5.2** (75 MHz, CDCl_3).

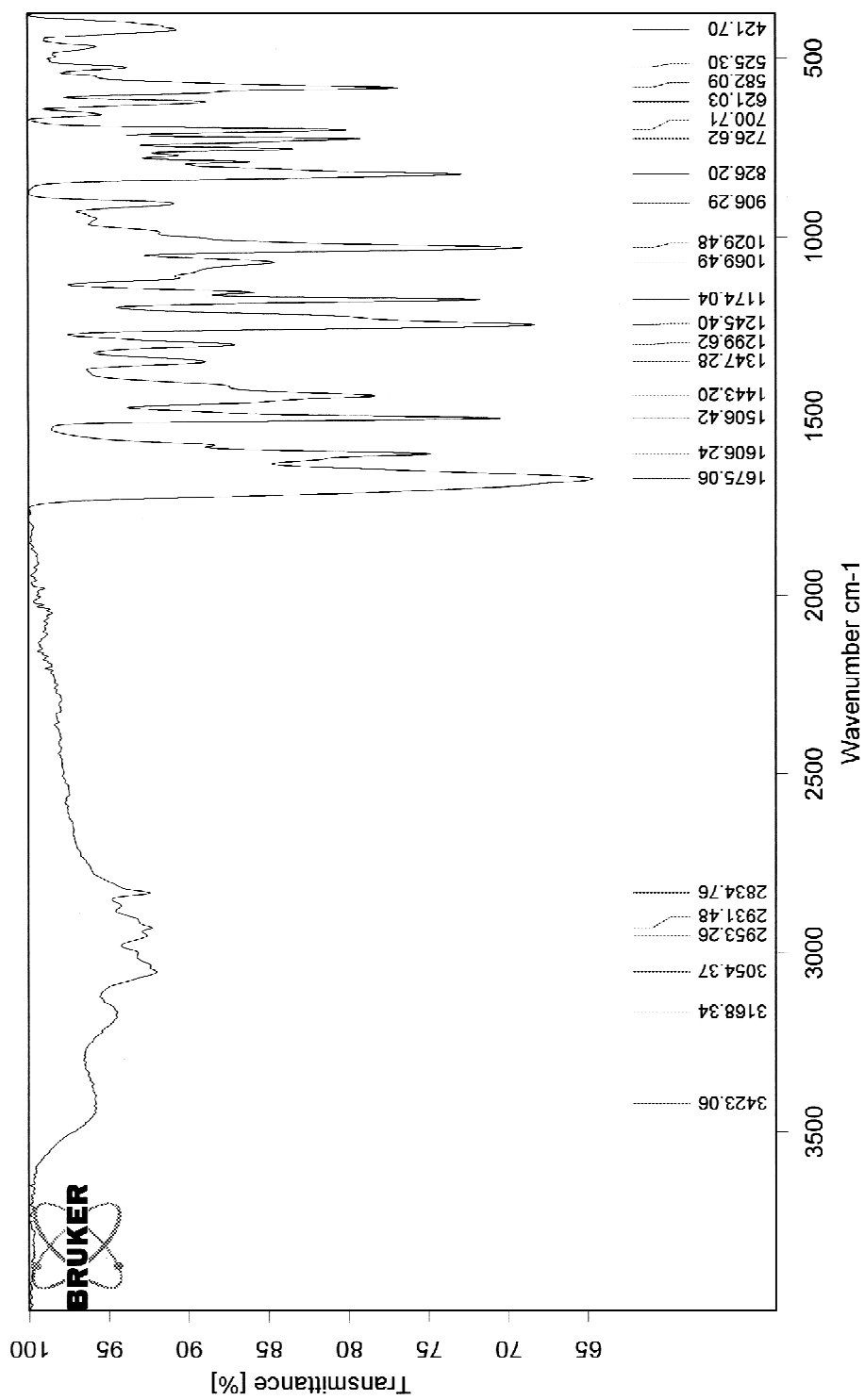


Figure A5.1.3. IR spectrum of compound 5.2 (solid).

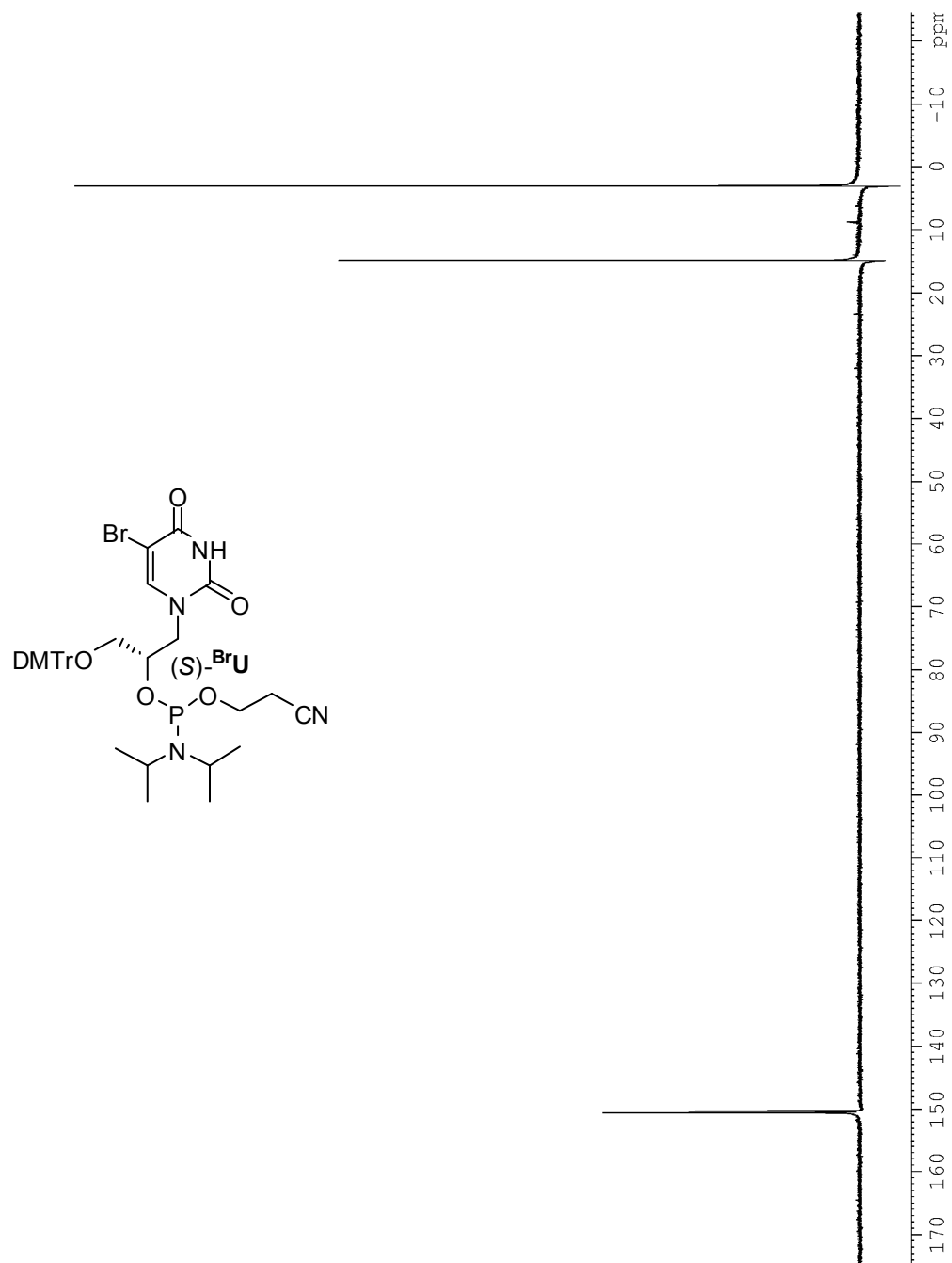


Figure A5.2.1. ³¹P NMR spectrum of phosphoramidite (S)-**BrU** (162 MHz, CDCl₃) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).

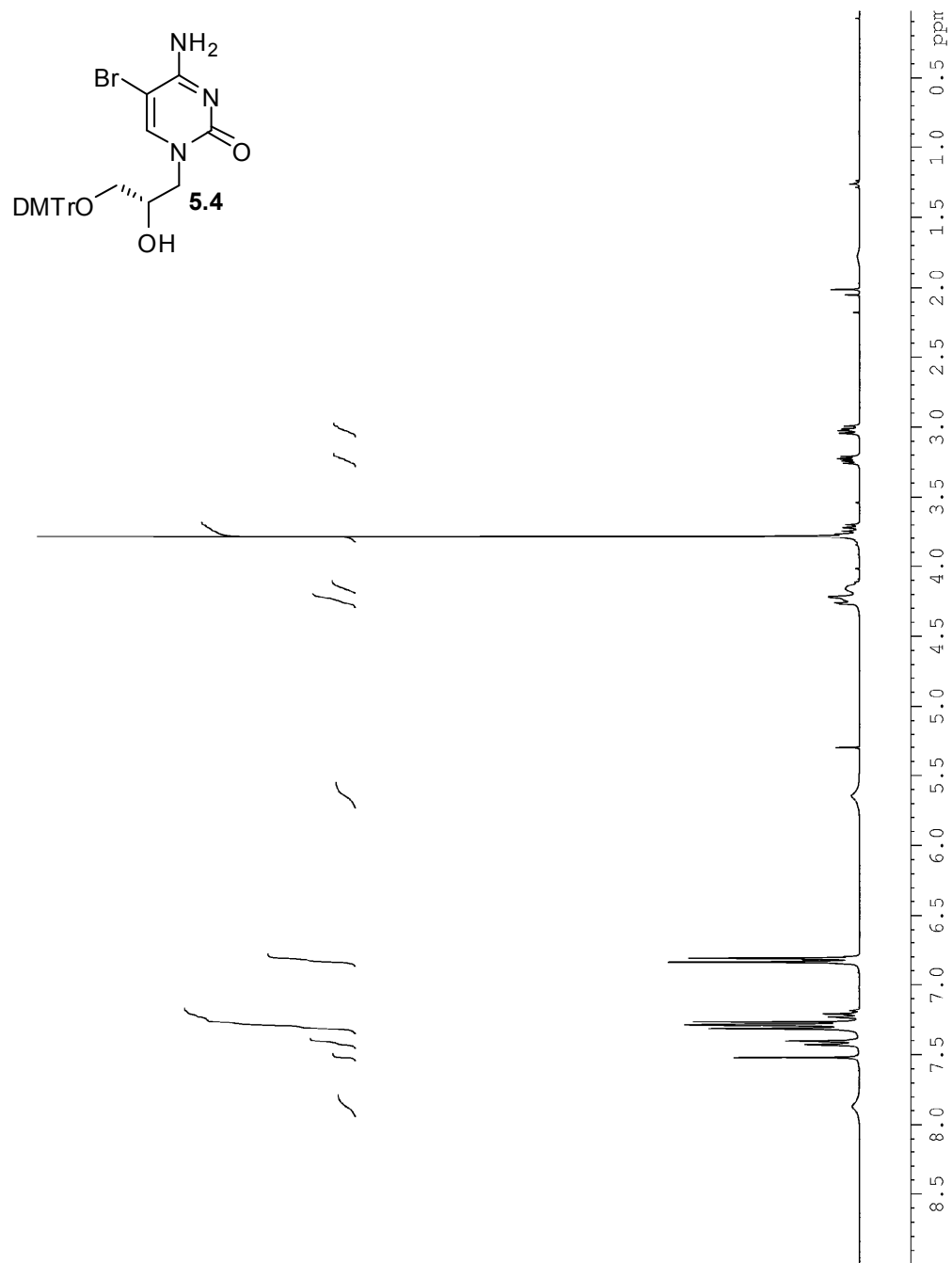


Figure A5.3.1. ¹H NMR spectrum of compound **5.4** (300 MHz, CDCl₃).

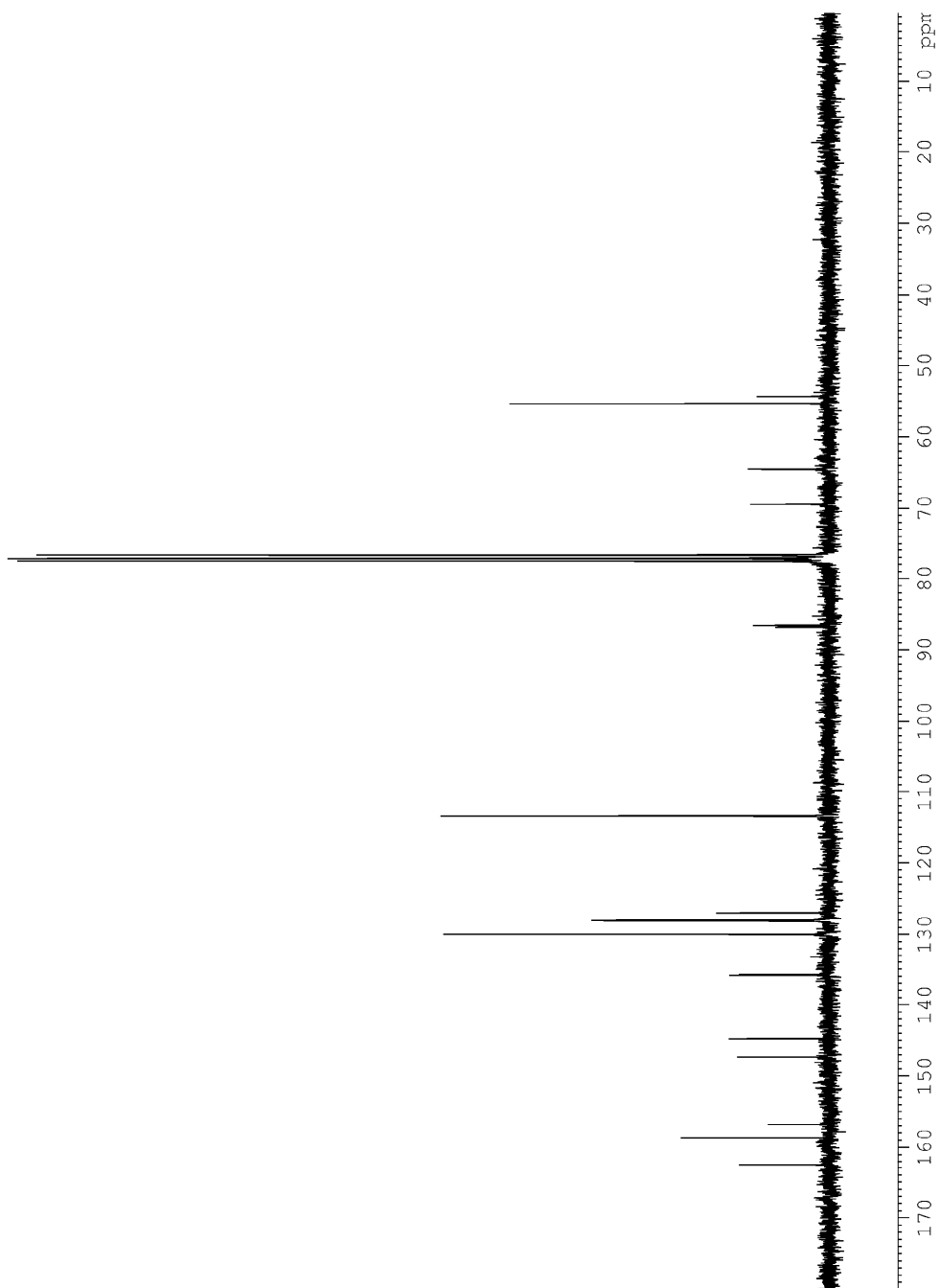


Figure A5.3.2. ^{13}C NMR spectrum of compound **5.4** (75 MHz, CDCl_3).

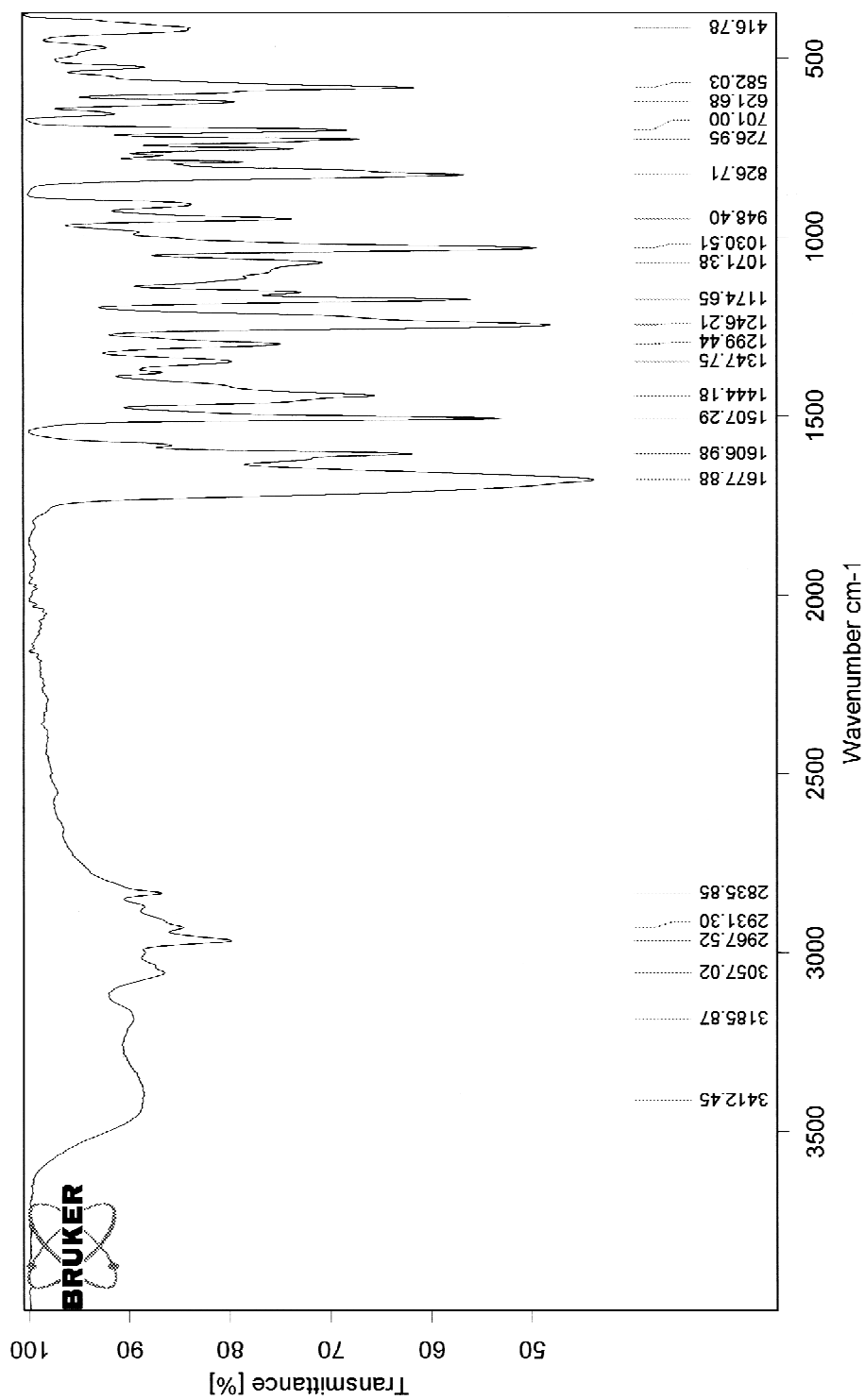


Figure A5.3.3. IR spectrum of compound 5.4 (solid).



Figure A5.4.1. ^1H NMR spectrum of compound **5.6** (300 MHz, $\text{DMSO}-d_6$).

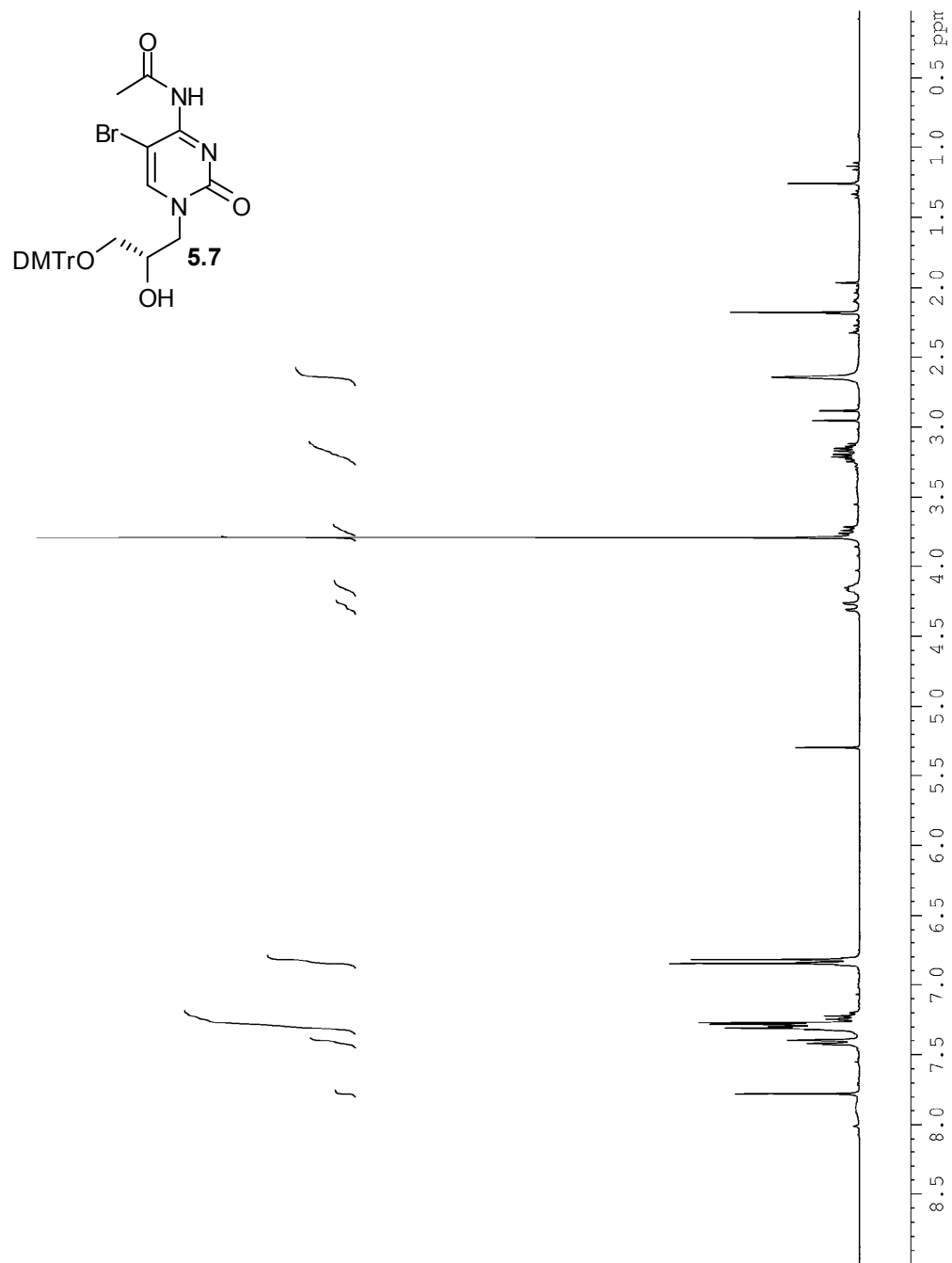


Figure A5.5.1. ^1H NMR spectrum of compound 5.7 (300 MHz, CDCl_3).

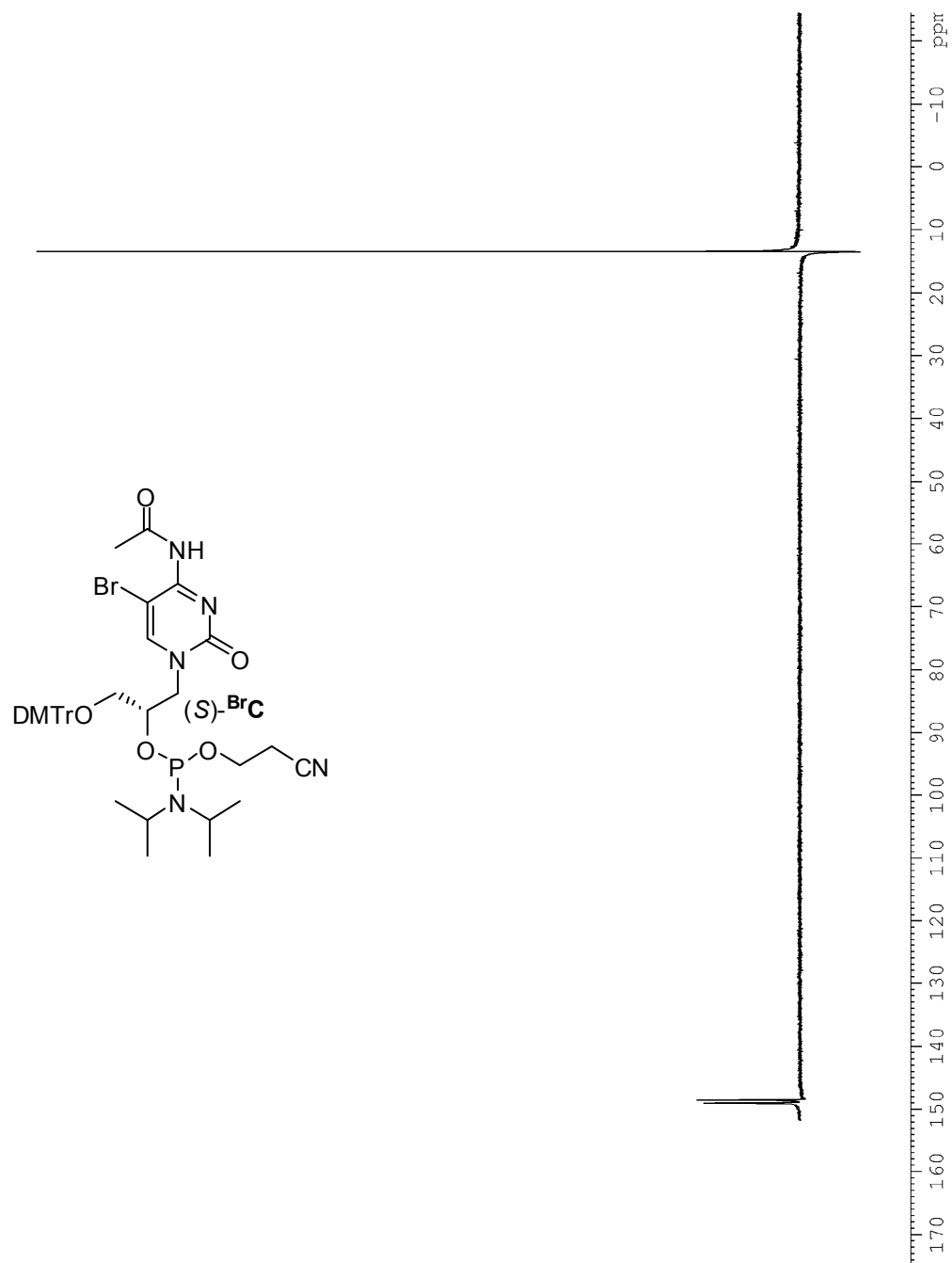


Figure A5.6.1. ³¹P NMR spectrum of phosphoramidite (*S*)-BrC (121 MHz, CDCl₃).

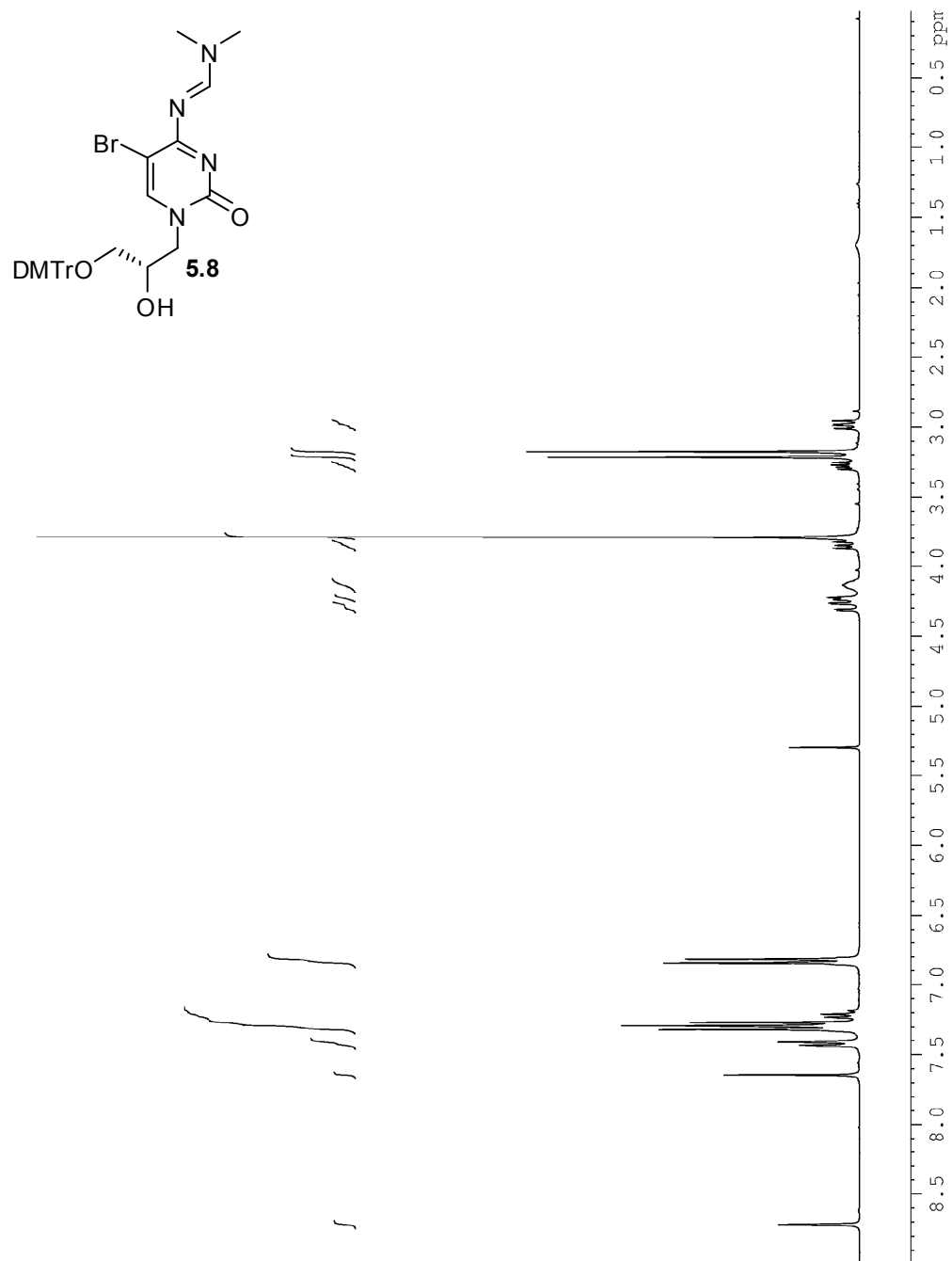


Figure A5.7.1. ¹H NMR spectrum of compound **5.8** (300 MHz, CDCl₃).

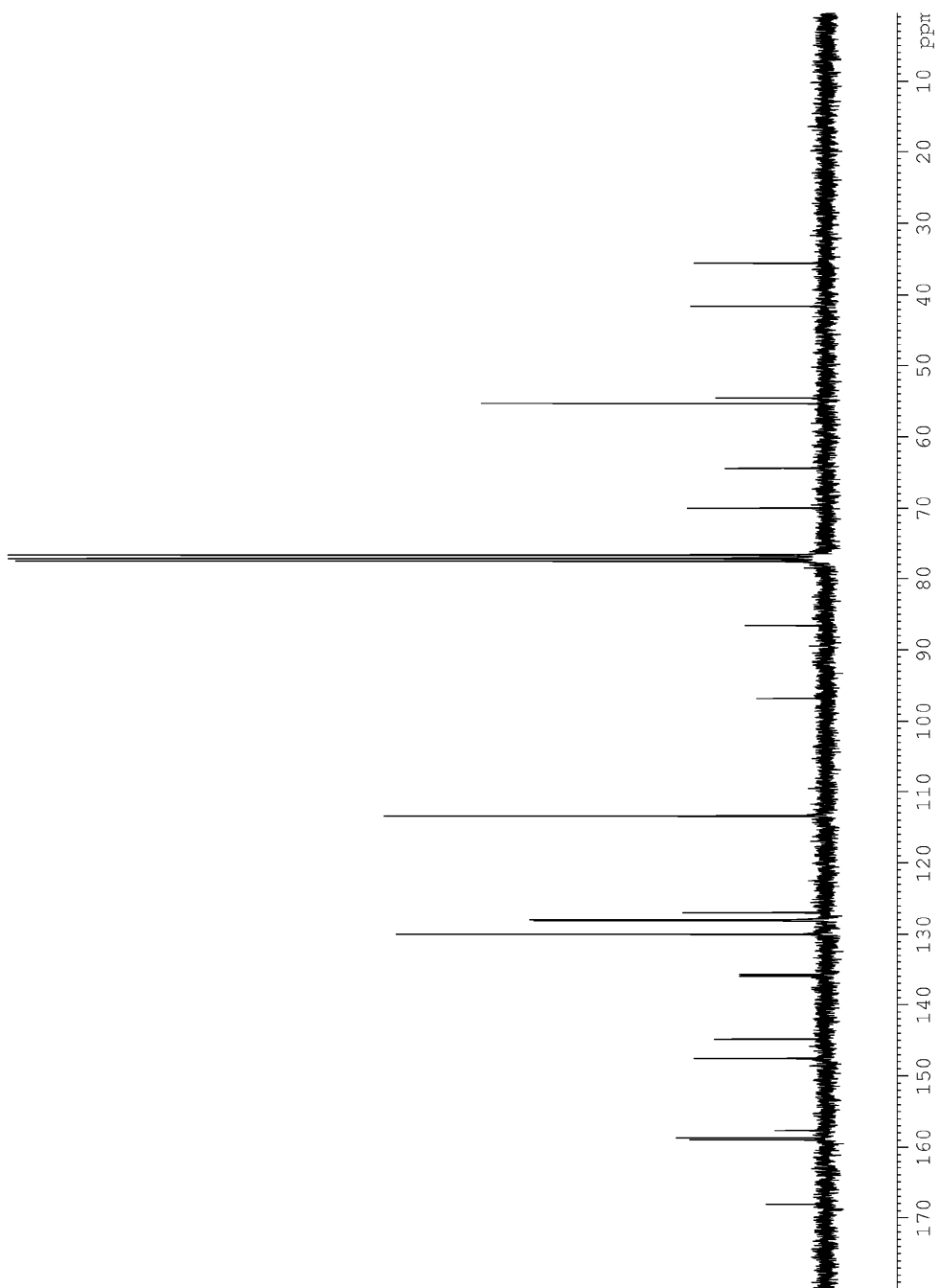


Figure A5.7.2. ^{13}C NMR spectrum of compound **5.8** (75 MHz, CDCl_3).

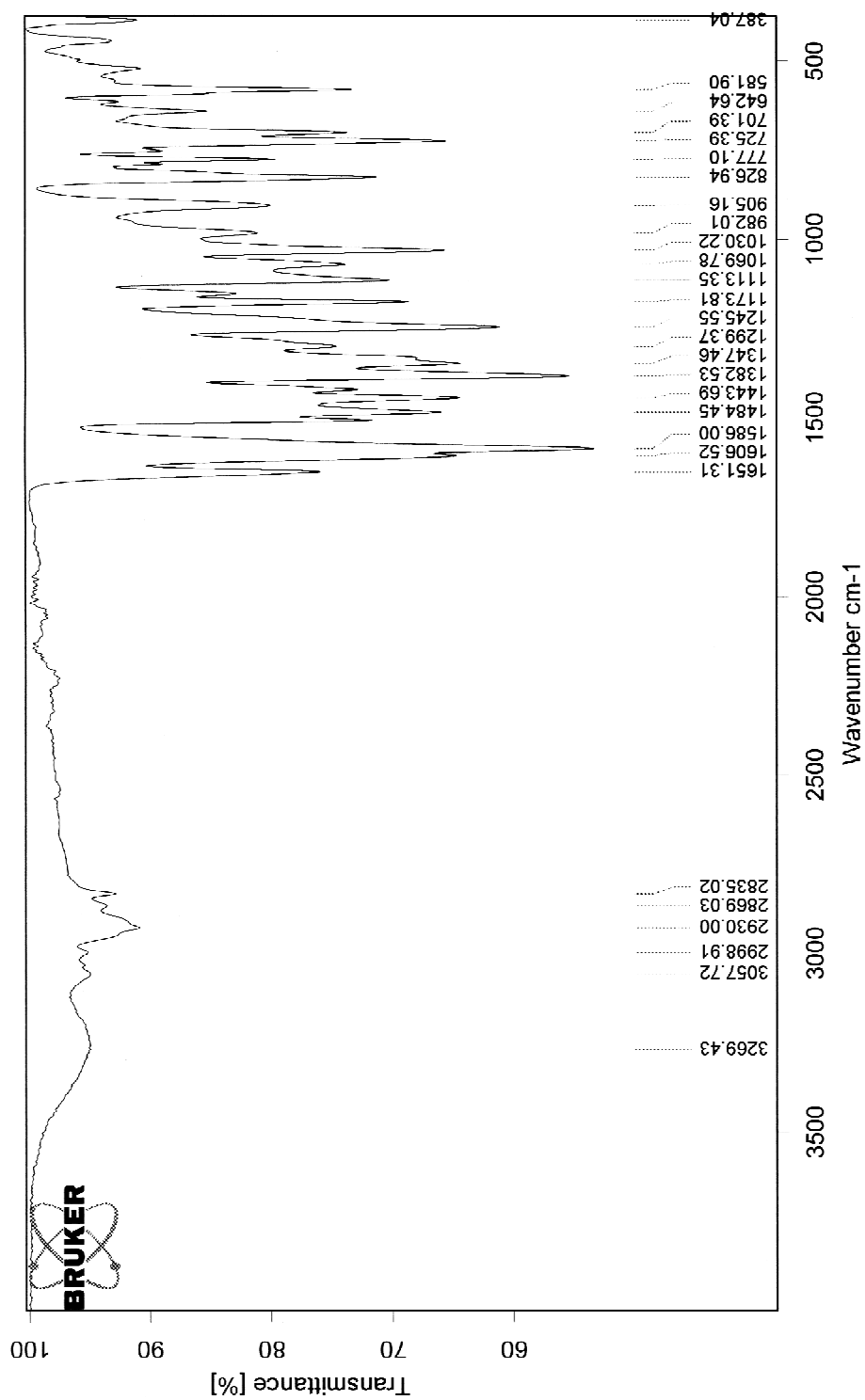


Figure A5.7.3. IR spectrum of compound 5.8 (solid).

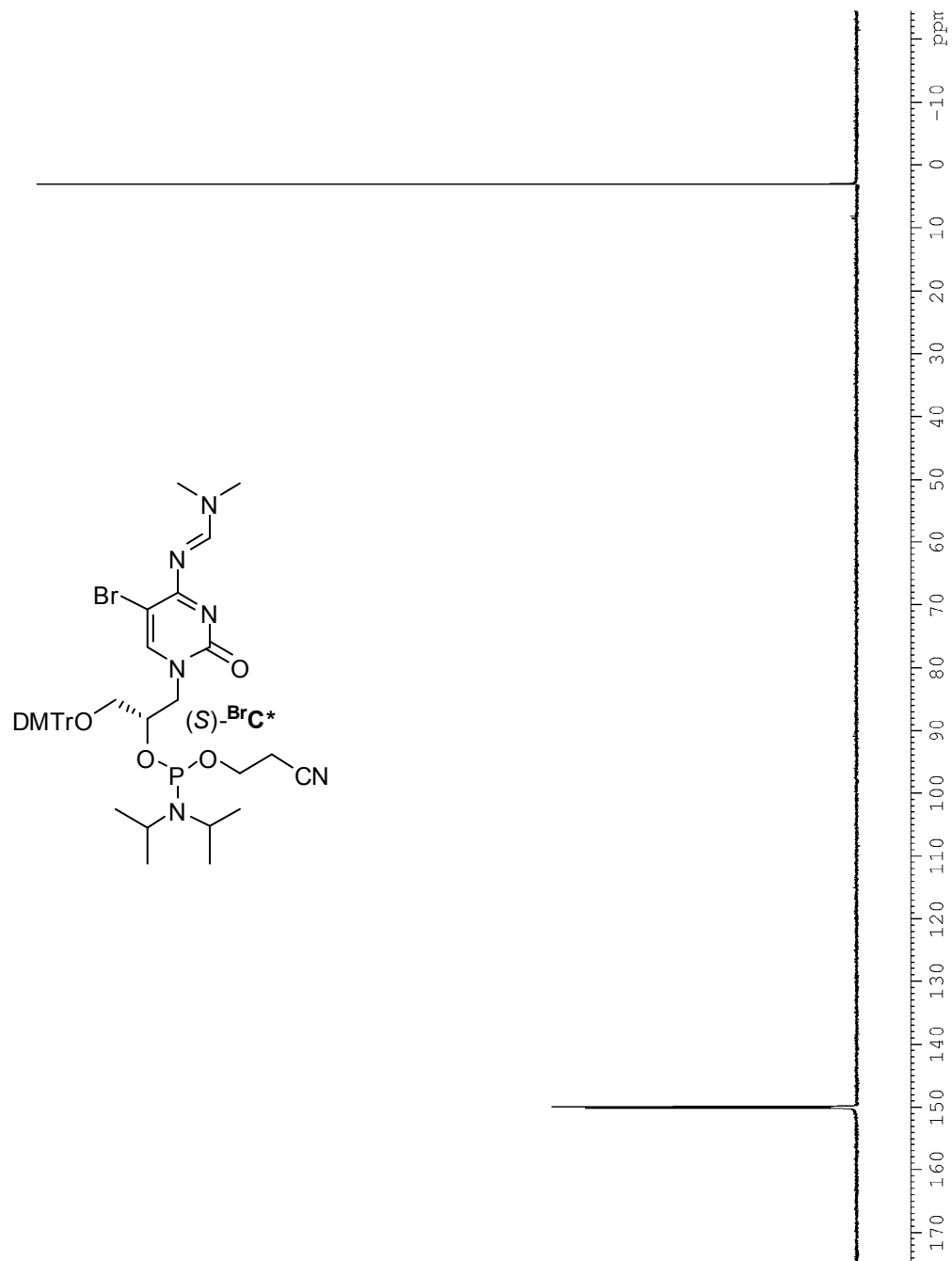


Figure A5.8.1. ^{31}P NMR spectrum of phosphoramidite (S) - BrC^* (162 MHz, CDCl_3) with trimethyl phosphate as an internal standard ($\delta = 3.06$ ppm).

Table A5.1.1. Data collection statistics of unsolved structures

Dataset (wavelength Å)	Cell (Å)	Resol. (Å)	Measured, unique reflections	$R_{\text{merge}}^{\text{[a]}}$	$I/\sigma(I)$ [b]	B_{wilson} (Å ²)	Completeness
MKS5-Native (0.8377)	a=18.55, b=37.91, c=40.19	9.27-1.20	37584, 9033	0.090 (0.350)	12.4 (4.0)	64.5	0.973 (0.989) ^[c]
MKS39-Native (0.9777)	a=42.23, b=42.23, c=96.97	19.4-2.35	28268, 4040	0.140 (0.689)	13.5 (1.8)	64.5	0.998 (1.000)
MKS111-Native (0.9162)	a=39.34, b=37.33, c=84.43	14.7-1.8	110698, 22498	0.147 (0.497) ^[d]	7.8 (3.2)	17.2	0.981 (0.986)
MKS118-Native (0.9196)	a=22.63, b=68.72, c=127.28	35-2.6	25642, 6514	0.037 (0.338)	19.6 (4.1)	86.8	0.982 (0.997)

[a] $R_{\text{merge}} = ((\sum \sum I_i(h) - \langle I(h) \rangle) / (\sum \sum I_i(h))) \times 100$; values in parentheses correspond to highest resolution shell. [b] As calculated with the program SCALA. [c] Only 77.5 % completeness in resolution shell 9.3-3.8 Å. [d] Large variation in mosaicity during data collection.

Table A5.2.1. Data collection, phasing, and refinement statistics of MKS132

Data collection							
Dataset (wavelength Å)	Cell (Å)	Resol. (Å)	Measured, unique reflections	$R_{\text{merge}}^{\text{[a]}}$	$I/\sigma(I)^{\text{[b]}}$	$B_{\text{Wilson}} (\text{Å}^2)$	Completeness
Peak (0.91826 Å)	a=20.42, b=42.00, c=28.67	21.0-0.965	30224, 7576	0.059 (0.132)	17.7 (7.2)	4.62	0.982 (0.914)
SAD Phasing in SHELX							
Resolution (Å)	21-2.14	-1.68	-1.33	-1.24	-1.11	-1.06	-0.97
CC (map)	0.903	0.958	0.934	0.945	0.951	0.941	0.942
Total FOM	0.850						
Refinement							
Resolution range (Å)	8.00-0.965						
Reflections (work, test)	7025, 533						
R-factor / $R_{\text{free}}^{\text{[c]}}$	0.104 / 0.128						
Average <i>B</i> -factor	4.87						
Bases, phosphates, ions	6, 5, 3						
Water molecules	48						
R.m.s. deviation bonds (Å)	0.006						
R.m.s. deviation angles (°)	1.622						

[a] $R_{\text{merge}} = ((\sum \sum I_i(h) - \langle I(h) \rangle) / (\sum \sum I_i(h))) \times 100$; values in parentheses correspond to highest resolution shell. [b] As calculated with the program SCALA. [c] $R = \sum ||F_o| - k|F_c|| / \sum |F_o|$ with *k* as scaling factor; R_{free} calculated with test set.

Table A5.3.1. Data collection, phasing, and refinement statistics of MKS80

Data collection									
Dataset (wavelength Å)	Cell (Å)	Resol. (Å)	Measured, unique reflections	R _{merge} ^[a]	I/σ(I) ^[b]	B _{Wilson} (Å ²)	Completeness		
Native (0.9777)	a=57.30, b=57.30, c=29.19	19.1-2.0	9210, 1607	0.055 (0.291)	22.2 (3.5)	34.0	0.926 (0.914)		
Peak (0.9199)	a=28.87, b=80.48, c=80.81	19.1-1.9	46136, 3881	0.076 (0.410)	18.1 (4.9)	36.2	0.998 (1.000)		
Peak (0.9199)	a=56.93, b=56.93, c=28.97	15.9-1.8	38960, 2242	0.054 (0.571)	32.2 (8.9)	34.2	0.990 (0.983)		
Inflection (0.9203)	a=57.27, b=57.27, c=29.31	14.7-2.1	41040, 1556	0.072 (0.435)	29.2 (9.6)	42.4	0.997 (1.000)		
Inflection (0.9203)	a=28.87, b=80.48, c=80.81	19.1-2.1	59840, 2903	0.083 (0.469)	23.1 (7.7)	45.7	0.998 (1.000)		
Remote (0.9165)	a=28.87, b=80.48, c=80.81	18.1-1.9	54809, 3877	0.085 (0.663)	18.4 (4.7)	36.2	0.997 (1.000)		
Phasing in SHELX									
Resolution (Å)	15-4.2	-3.3	-2.6	-2.4	-2.2	-2.1	-2.0	-1.9	
CC (map)	0.883	0.961	0.852	0.854	0.903	0.877	0.847	0.770	
Total FOM	0.648 (DM: 0.800)								
Refinement									
Resolution range (Å)	10.0-2.10								
Reflections (work, test)	1431, 112								
R-factor / R _{free} ^[c]	0.236 (0.196), 0.325 (0.421)								
Bases, phosphates	8, 7								
Water molecules									
R.m.s. deviation bonds (Å)	0.013								
R.m.s. deviation angles (°)	2.024								

[a] $R_{\text{merge}} = ((\sum \sum |I(h) - \langle I(h) \rangle|) / (\sum \sum I(h))) \times 100$; values in parentheses correspond to highest resolution shell. [b] As calculated with the program SCALA. [c] $R = \sum \|F_o - k|F_c| / \sum \|F_o$ with k as scaling factor; R_{free} calculated with test set.

Table A5.4.1. Data collection, phasing, and refinement statistics of MKS74-1

Data collection							
Dataset (wavelength Å)	Cell (Å)	Resol. (Å)	Measured, unique reflections	R _{merge} ^[a]	I/σ(I) ^[b]	B _{Wilson} (Å ²)	Completeness
Native (0.9777)	a=57.30, b=57.30, c=29.19	19.1-2.0	9210, 1607	0.055 (0.291)	22.2 (3.5)	34.0	0.926 (0.914)
Remote (0.9788)	a=57.03, b=57.03, c=29.22	14.6-2.0	22140, 1763	0.077 (0.519)	25.3 (5.3)	41.7	0.995 (0.999)
Inflection (0.9797)	a=57.30, b=57.30, c=29.19	14.6-2.0	22250, 1781	0.083 (0.429)	24.6 (5.9)	39.3	0.994 (0.995)
Peak (0.9791)	a=57.07, b=57.07, c=29.24	18.0-2.0	23155, 1778	0.089 (0.366)	24.2 (8.0)	39.7	0.999 (1.000)
Refinement							
Resolution range (Å) ^[d]	18.12-2.00						
Reflections (work, test)	1652, 118						
R-factor / R _{free} ^[c]	0.266, 0.318						
Bases, phosphates	8, 7						
Water molecules							
R.m.s. deviation bonds (Å)	0.016						
R.m.s. deviation angles (°)	2.366						

[a] $R_{\text{merge}} = ((\sum \sum I_i(h) - \langle I(h) \rangle) / (\sum \sum I_i(h))) \times 100$; values in parentheses correspond to highest resolution shell. [b] As calculated with the program SCALA. [c] $R = \sum ||F_o| - k|F_c|| / \sum |F_o|$ with k as scaling factor; R_{free} calculated with test set. [d] Refinement done with the inflection point data.

Table A5.5.1.1. Data collection and refinement statistics of MKS22

Data collection							
Dataset (wavelength Å)	Cell (Å)	Resol. (Å)	Measured, unique reflections	$R_{\text{merge}}^{\text{[a]}}$	$I/\sigma(I)$ [b]	BWilson (Å ²)	Completeness
Native (0.9777)	a=57.30, b=57.30, c=29.19	19.1-2.0	9210, 1607	0.055 (0.291)	22.2 (3.5)	34.0	0.926 (0.914)
Refinement							
Resolution range (Å)	40.0-2.01						
Reflections (work, test)	2910, 134						
R-factor / $R_{\text{free}}^{\text{[c]}}$	0.313 (0.398)/0.336 (0.203)						
Bases, phosphates	8, 7						
Water molecules							
R.m.s. deviation bonds (Å)	0.028						
R.m.s. deviation angles (°)	4.836						

[a] $R_{\text{merge}} = ((\sum \sum |I_i(h) - \langle I(h) \rangle|) / (\sum \sum I_i(h))) \times 100$; values in parentheses correspond to highest resolution shell. [b] As calculated with the program SCALA. [c] $R = \sum \|F_o - k|F_c|\| / \sum \|F_o\|$ with k as scaling factor; R_{free} calculated with test set.

Table A5.6.1.1. Data collection and refinement statistics of MKS42

Data collection							
Dataset (wavelength Å)	Cell (Å)	Resol. (Å)	Measured, unique reflections	R _{merge} ^[a]	I/σ(I)[b]	BWilson (Å ²)	Completeness
Native (0.9762)	a=22.50, b=50.61, c=65.62	19.6-2.0	24213, 5346	0.057 (0.447)	15.0 (4.0)	38.3	0.985 (0.999)
Refinement							
Resolution range (Å)	19.6-2.00						
Reflections (work, test)	4935, 410						
R-factor / R _{free} ^[c]	0.339 (0.398), 0.369 (0.420)						
Bases, phosphates ^[d]	?						
Water molecules	?						
R.m.s. deviation bonds (Å)	0.011						
R.m.s. deviation angles (°)	2.156						

[a] $R_{\text{merge}} = ((\sum \sum |I_i(h) - \langle I(h) \rangle|) / (\sum \sum I_i(h))) \times 100$; values in parentheses correspond to highest resolution shell. [b] As calculated with the program SCALA. [c] $R = \sum \|F_o\| - k \|\sum F_c\| / \sum \|F_o\|$ with k as scaling factor; R_{free} calculated with test set. [d]