Introduction of Modifications to the Nucleic Acid Scaffold for the Preparation of Circular Structures and Applications in Biochemistry

Gabrielle Juneau

A Thesis
In
The Department
of
Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Science (Biochemistry) at

Concordia University
Montreal, Quebec, Canada

August 2022

Gabrielle Juneau, 2022

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

by: Gabrielle Juneau

Entitled: Introduction of Modifications to the Nucleic Acid Scaffold for the Preparation of Circular Structures and Applications in Biochemistry

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biochemistry)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____Chair

Dr. Marek Majewski

_____Examiner

Dr. Dajana Vuckovic

_____Examiner

Dr. Xavier Ottenwaelder

_____Supervisor

Dr. Christopher Wilds

Approved by _____

Dr. Louis Cuccia

_____2022 _____

Dr. Pascale Sicotte, Dean of Arts and Science

Abstract

Introduction of Modifications to the Nucleic Acid Scaffold for the Preparation of Circular Structures and Applications in Biochemistry

Gabrielle Juneau

The use of oligonucleotides as therapeutic and diagnostic tools has grown significantly in recent decades. However, native oligonucleotides face numerous challenges such as poor cellular uptake and low resistance to nucleases hence necessitating the incorporation of modifications to their scaffold. Small circular oligonucleotides have shown promise in these fields; however, their formation is limited by their length and the proximity of the ligating ends. To address this challenge, we investigated a methodology to synthesize circular constructs containing a chemically stable butylene or base labile sulfonylethane interstrand cross-link to pre-organize the system to aid in the circularization via CNBr-assisted ligation or click chemistry. Moreover, nucleic acid crystallography has proven to be an asset in determining nucleic acid structures and complexes. Incorporation of a selenium atom to the scaffold of DNA has helped with phasing issues. Thus, we investigated a novel approach to introduce a 5'-methylphosphoroselenoate handle by the use of a methylphosphoramidite followed by treatment with potassium selenocyanate. Lastly, the O^6-position of guanine bases can be methylated by various exo- and endogenous agents. This adduct can be repaired by the protein O^6-Alkylguanine DNA Alkyltransferase (AGT). The activity of AGT can reduce the efficiency of some alkylating chemotherapeutic drugs. Thus, the discovery and evaluation of AGT inhibitors is desirable. Hence, we explored development of a method to evaluate the repair proficiency of AGT in the presence of inhibitors towards DNA duplexes which contain an O^6-methyl-2'-deoxyguanosine adduct.

Acknowledgements

First and foremost, I would like to say thank you to my supervisor Dr. Christopher Wilds for the opportunity to work in the lab, I enjoyed every moment of it. I would like to extend a big thank you to Dr. Anne Noronha for her help in the lab, between managing the lab and making sure I had all that was needed for my experiments to help with oligonucleotide synthesis, I really couldn't have done it without you.

To my committee members, Dr. Dajana Vuckovic and Dr. Xavier Ottenwaelder, thank you for the support as well as new ideas and perspective throughout the last few years.

Thank you as well to all past and present lab members who have helped make the darker days in the lab just a little brighter. A special shout-out goes out to Alex Pontarelli, my partner in crime in and out of the lab! Thank you for helping me with the synthesis of all small molecules and for the constant support and encouragements. I can't wait to see what the next chapter holds for us. Also thank you to Dr. Enrique Pedroso and Dr. Anna Grandas (Universitat de Barcelona) for their collaboration on the circular DNA project

Enfin, un grand merci à tous mes amis et à ma famille qui m'ont aidé, soutenu et encouragé durant ces dernières années. Vos mots d'encouragement m'ont fait chaud au cœur.

Thanks to all!

Merci à tous!

Table of Contents

List of Figures

List of Tables

List of Schemes

List of Abbreviations

2'-MOE	2'-O-Methoxyethyl
A	Adenine
ACN	Acetonitrile
AGT	O^6-Alkylguanine DNA Alkyltransferase
AMP	Adenosine Monophosphate
APS	Ammonium Persulfate
ASO	Antisense Oligonucleotide
ATP	Adenosine Triphosphate
BG	O^6-benzylguanine
C	Cytosine
CD	Circular Dichroism
CNBr	Cyanogen Bromide
CPR	Chemical Phosphorylating Reagent
CuAAC	Copper (I) Catalyzed Alkyl-Azide Cycloaddition
dA	2'-Deoxyadenosine
DBU	1,8-Diazabicyclo [5.4.0] undec-7-ene
dC	2'-Deoxycytidine
DCM	Dichloromethane
DDTT	3-[(Dimethylaminomethylene) amino]-3H-1,2,4-dithiazole-5-tione
dG	2'-Deoxyguanosine
DIPEA	N,N-Diisopropylethylamine
DMD	Duchene Muscular Dystrophy
DMF	Dimethylformamide
DMT	Dimethoxytrityl
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphate
dsDNA	Double Stranded DNA
EDTA	Ethylene diamine tetra-acetic acid
EtOAc	Ethyl acetate
EtOH	Ethanol
Eqs	Equivalents
FCC	Flash Column Chromatography
G	Guanine
hAGT	Human AGT
H-bond	Hydrogen Bond
HBr	Hydrobromic Acid
Hex	Hexanes
HPLC	High Performance Liquid Chromatography
HTH	Helix-Turn-Helix
IEX-HPLC	Ion Exchange HPLC
iPr-Pac-dG	2'-Deoxy-4-isopropyl-phenoxyacetyl-guanosine
KSeCN	Potassium Selenocyanate
LCAA-CPG	Long Chain Alkylamine Controlled Pore Glass
LNA	Locked Nucleic Acid
MAD	Multiple Anomalous Diffraction
MeOH	Methanol
MES	N-Morpholinoethanesulfonate

MGMT	O^6-Methylguanine DNA Methyltransferase
miRNA	microRNA
MMR	Mismatch Repair
mRNA	Messenger RNA
MS	Mass Spectrometry
NTP	Nucleotide Triphosphate
OD	Optical Density
Pac	Phenoxyacetyl
Pac-dA	2'-Deoxy-N6-phenoxyacetyl-adenosine
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PMDTA	Pentamethyldiethylenetriamine
PMO	Phosphorodiamidate Morpholio Oligomer
PNA	Peptide Nucleic Acid
PO	Phosphodiester
PPi	Pyrophosphate
PSe	Phosphoroselenoate
RCA	Rolling Circle Amplification
RCT	Rolling Circle Transcription
RNA	Ribonucleic Acid
RNAi	RNA Interference
ROS	Reactive Oxygen Species
SeNA	Selenium-derivatized Nucleic Acids
siRNA	Small Interfering RNA
ssDNA	Single Stranded DNA
SVDPE	Snake Venom Phosphodiesterase
T	Thymine
TBAF	Tetra-n-butylammonium fluoride
TBDMS	*tert*-Butyldimethylsilyl
TCA	Trichloroacetic acid
TEA	Triethylamine
TEMED	Tetramethylethylenediamine
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
T_m	Melting Temperature
U	Uracil
UPP	Ubiquitin-Proteasome Pathway
UV	Ultraviolet
V_f	Final Reaction Volume

Chapter 1 - General Introduction

1.1-Key Events in the History of Nucleic Acid Chemistry

Nucleic acids are biomolecules responsible for numerous roles within a cell. When they were initially discovered by Swiss scientist Frederich Meischer in the late 1860s, nucleic acids were called nuclein due to the fact that they were isolated from the nucleus of a cell. However, due to their acidic nature, these biomolecules became known as nucleic acids by the end of the 1880s.[1] By the 1930s, nucleic acids had further been divided into two categories: deoxyribonucleic acids (DNA) and ribonucleic acids (RNA). Then, two decades later, the biological and physiological role of nucleic acids were examined, followed by the hallmark discovery of the structure of the DNA double helix by Watson and Crick in 1953. To this day, experimental studies of nucleic acids still constitute a major part of modern biological and medical research as well as foundation for biotechnological and pharmaceutical industries.[2]

1.2-Composition and Structure of Nucleic Acids

1.2.1-Arrangement of Nucleic Acids

It is well known that DNA holds all the genetic information necessary for life while RNA is mainly used as a translator between DNA and proteins. DNA and RNA are both composed of nucleotides linked together by phosphodiester bonds between the 5' and the 3' positions of adjacent nucleotides where the difference between these two biomolecules lies in the absence or presence of a hydroxyl group at the 2'-position of the sugar moiety (**Figure 1.1-A)**. A nucleotide is composed of a pentose sugar, a heterocyclic nitrogenous base, and a phosphate group whereas a nucleoside lacks the phosphate group. The pentose sugar can adopt a variety of conformations, also known as sugar pucker, in order to maintain a stable helical form and minimize non-bonded

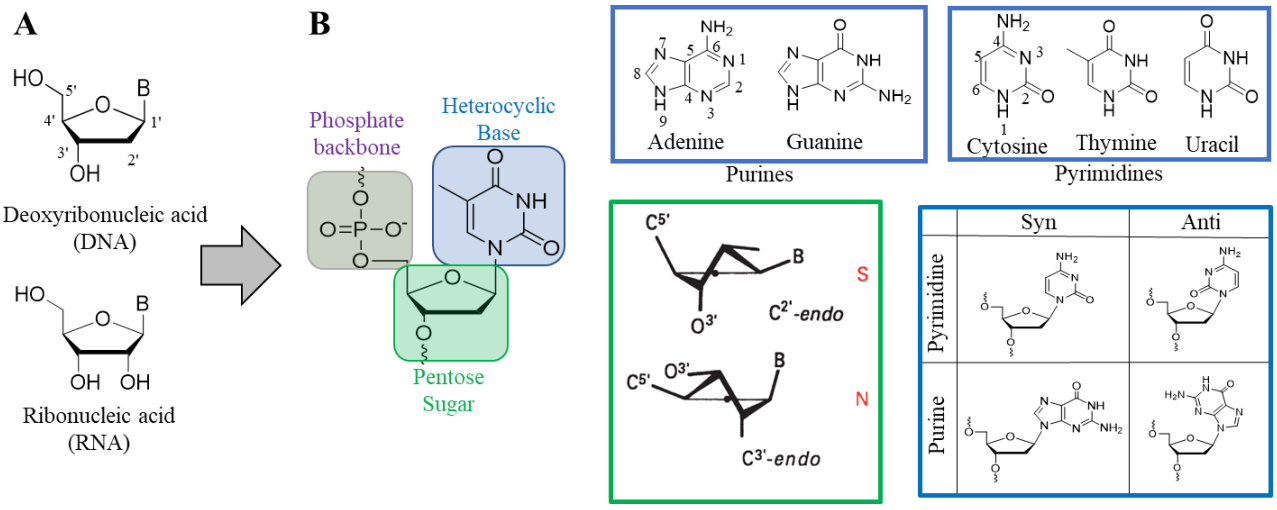


Figure 1.1-(**A**) Structure of DNA and RNA, (**B**) Components of nucleic acids (partial reproduction with permission from *reference 4*).

interactions of their substituents.[3] The most common sugar puckers are the C3'-endo and the C2'-endo, where the C3' or C2' atom, respectively resides on the same side of the five membered ring as the C5' or the nucleobase. The nitrogenous bases can be classified as purines, such as adenine (A) and guanine (G), or pyrimidines such as cytosine (C) and thymine (T) in the case of DNA or cytosine (C) and uracil (U) in the case of RNA. The nitrogen-containing base and pentose sugar are covalently attached through a β-glycosidic bond between C1' and N9 (purines) or N1 (pyrimidines).[4] The planes of the sugar and the base are almost perfectly perpendicular allowing the heterocycles to adopt one of two conformations, that is, the *syn*- and *anti*-conformations. In the *syn*-conformation the O2-atom of pyrimidines or the N3-atom of purines are directly above the sugar ring. However, in the *anti*-conformation, the H6-atom of pyrimidines or the H8-atom of purines sit above the sugar ring (**Figure 1.1-B**).[4]

1.2.2-DNA Polymorphism

DNA is polymorphic meaning that it can adopt a variety of different structures. Each base has a certain hydrogen bond (H-bond) donor/acceptor pattern which allows them to recognize their complementary partner on another strand resulting in base pairing. The most common type of H-

bonding pattern is Watson-Crick base pairing where A pairs with T forming 2 H-bonds and G pairs with C forming 3 H-bonds (**Figure 1.2**). In RNA, A base pairs with U also forming 2 H-bonds. The Watson-Crick base pairs result in the well-known anti-parallel double helix of DNA.[4,5]

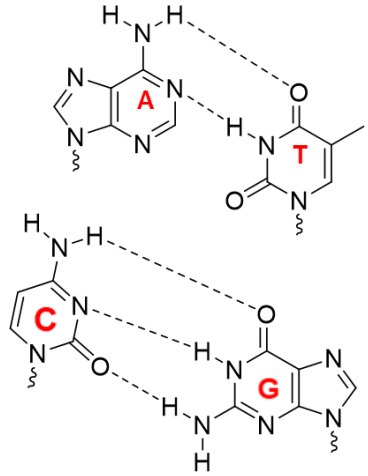


Figure 1.2-Watson Crick hydrogen bonds.

There exist three major secondary structures for DNA: the B-form, the A-form and the Z-form. Under physiological conditions, B-form DNA (**Figure 1.3-B**) is predominant. One model B-form DNA duplex that has been studied extensively using X-ray crystallography is the Drew-Dickerson dodecamer (dCGCGAATTCGCG).[6] In B-form DNA the base pairs sit directly on the helix allowing the major and minor groove to be of similar depths with both strands running anti-parellel to one another. Furthermore, the bases are stacked perpendicular to the axis and the sugars have a C2'-endo pucker resulting in a right-handed helix.[4,7] The A-form (**Figure 1.3-A**) is adopted under dehydrating conditions such as low humidity and high salt concentration which result in a more compact form compared to the B-form duplex. In the A-form duplex, DNA adopts a C3'-endo sugar pucker and like B-form duplexes, A-form duplexes are right-handed with anti-parallel complementary strands. However, the sugar rings are parallel to the helical axis and the bases are slightly tilted sideways allowing for the formation of a

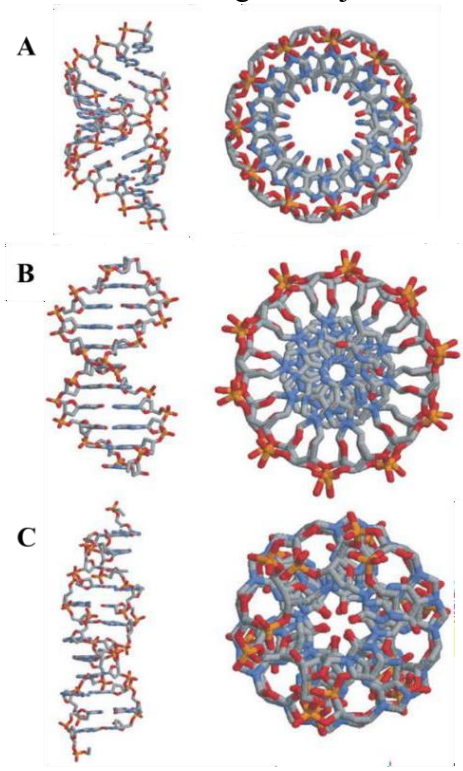


Figure 1.3-(**A**) A-form DNA, (**B**) B-form DNA, (**C**) Z-form DNA, (reproduced with permission from *reference 7*).

hollow core as well as a wide major groove and narrow minor groove.[4] Lastly, Z-DNA (**Figure 1.3-C**) is radically different than its sister conformers in the fact that it is a left-handed duplex with a zig-zag backbone. Initially, Z-DNA was discovered in oligonucleotides with a purine-pyrimidine alternating sequence with the most studied Z-DNA oligonucleotide being the 6-mer d(CGCGCG) however other sequences deviating from the purine-pyrimidine alternating sequence have been found to form Z-DNA.[4] The zig-zag backbone of Z-DNA is caused by the different conformations adopted by the sugars where cytosine adopts an *anti*-conformation with a C2'-endo sugar pucker and guanine adopts a *syn*-conformation with a C3'-endo sugar pucker.[4] Like B-DNA, this structure has very little differences in width in the minor and major groove of DNA however it is the least favored due to its low stability and transient nature.[4,8] Under the appropriate conditions, DNA can adopt all three of these forms.[4,9]

1.2.3-Other Structures

Other non-canonical base interactions exist like the Hoogsteen base pairs (**Figure 1.4**) which allows for the formation of other interesting DNA structures such as triplexes, G-quadruplexes, and i-motif. DNA triplexes can form relatively easily when a DNA duplex interacts with a third single-stranded DNA through Hoogsteen base pairing whereas G- quadruplexes and i-motifs are dependent on the base composition and environmental factors. G- quadruplexes require a neutral pH, a metal cation (Na^+ or K^+) and a G-rich sequence to form. The quartet is formed via both Watson-Crick and Hoogsteen base pairing between adjacent guanines resulting in a quasi-planar structure.[5,10] On the other hand, the i-motif, which stands for intercalating motif, is formed in the presence of cytosine-rich sequences and requires a low pH to form resulting in the hemi-protonation of cytosine residues.[5,11]

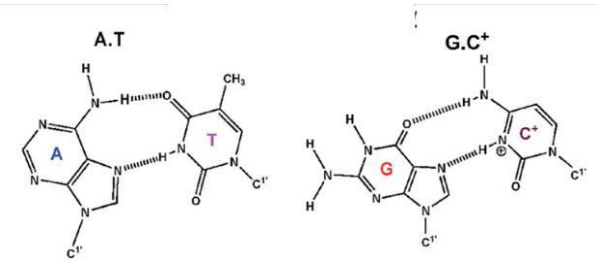


Figure 1.4-Hoogsteen base pairs (reproduced with permission from *reference 5*).

1.3-Chemical Synthesis of DNA and RNA

Nature has evolved a very efficient and sophisticated way to make and replicate DNA from a template where numerous enzymes and proteins work in conjunction to unwind, replicate and ligate the new DNA strands from the template strand. Unfortunately, we have yet to reach this level of sophistication when it comes to making DNA outside of the body. In a laboratory setting, DNA can be made in one of two ways: either via a series of enzymatic reactions using deoxynucleotide triphosphates (dNTPs) catalyzed by DNA polymerases which closely resembles natural DNA replication or DNA can also be made in an automated fashion on an instrument called a DNA synthesizer.

1.3.1-Enzymatic Synthesis of Oligonucleotides

Enzymatically, DNA and RNA are catalyzed by the enzymes DNA or RNA polymerases in the presence of dNTPs or nucleoside triphosphates (NTPs) respectively. The polymerase enzyme reads the template strand, which runs anti-parallel to the growing strand, to recruit the desired (d)NTP. The enzyme can then catalyze the reaction between the free 3'-hydroxyl and the α-phosphate of the (d)NTP yielding a phosphodiester bond.[12] This results in the release of pyrophosphate (PPi) which generates a larger amount of energy which is the driving force of this reaction.[12] The process proceeds with a 5' to 3' directionality.

1.3.2-Chemical Synthesis of Oligonucleotides

The conventional approach for chemical synthesis of an oligonucleotide chain uses phosphoramidite chemistry and follows a 4-step process involving solid-phase synthesis (**Figure 1.5**). This process is dubbed "solid-phase" due to the fact that the oligonucleotide is grown on a porous, insoluble support. Prior to synthesis, the first nucleoside of the sequence must be functionalized onto the solid support through a linkage such as a succinyl group. The two most

commonly employed types of solid support are glass beads known as long-chain amino-alkyl controlled pore glass (LCAA-CPG) and polystyrene. CPG is the most popular support that can enable the synthesis of oligomers of up to 100 nucleotides typically with higher loading (10-50 μmol/g) than polystyrene.[4] Polystyrene, on the other hand, has the advantage of being more efficient for small scale synthesis (typically around 40 nmol) due to its moisture exclusion properties.[4]

Contrarily to what is seen in nature, conventional solid-phase synthesis of oligonucleotides made in a lab occurs from the 3'-to the 5'-position. The cycle (**Figure 1.5**) begins with the detritylation of the 5'O-position of the first nucleotide which is bound to the solid support using 3% trichloroacetic acid/dichloromethane (TCA/DCM, v/v). This is followed by coupling of the second nucleoside to the free 5'-hydroxyl group via a phosphoramidite building block in the presence of a weak acid activator such as ethylthiotetrazole allowing for the formation of a phosphodiester linkage between the first and second nucleotide. The third step involves capping any unreacted 5'-hydroxyl groups with a solution of acetic anhydride in pyridine and THF to avoid further elongation of failure sequences. The final step involves oxidizing the P(III) species to a P(V) species using iodine in water, pyridine and THF. This cycle is repeated until an oligonucleotide of desired length and composition is achieved. The coupling efficiency of each step is determined by comparing the intensity of the bright orange dimethoxytrityl (DMT) cation released following detritylation with the intensity of previous detritylations. Furthermore, stepwise yields (coupling efficiency) as well as the length of the oligonucleotide being synthesized are both determining factors in the overall yield of the synthesis: length is inversely proportional to yield while coupling efficiency is directly proportional to yield.[4]

Using this method requires the protection of the reactive positions on each individual base as well as the reactive positions of the backbone. Typically, a heated solution of NH_4OH/EtOH (3:1, v/v) can readily remove the N^4-acetyl of 2'-deoxycytidine (dC), N^6-benzoyl of 2'-deoxyadenosine (dA), N^2-isobutyryl on 2'-deoxyguanosine (dG) as well as the cyanoethyl (backbone) protective groups. In the case of RNA, an extra protecting group (*tert*-butyldimethylsilyl) is needed to protect the O2'-position and can be removed via treatment with a

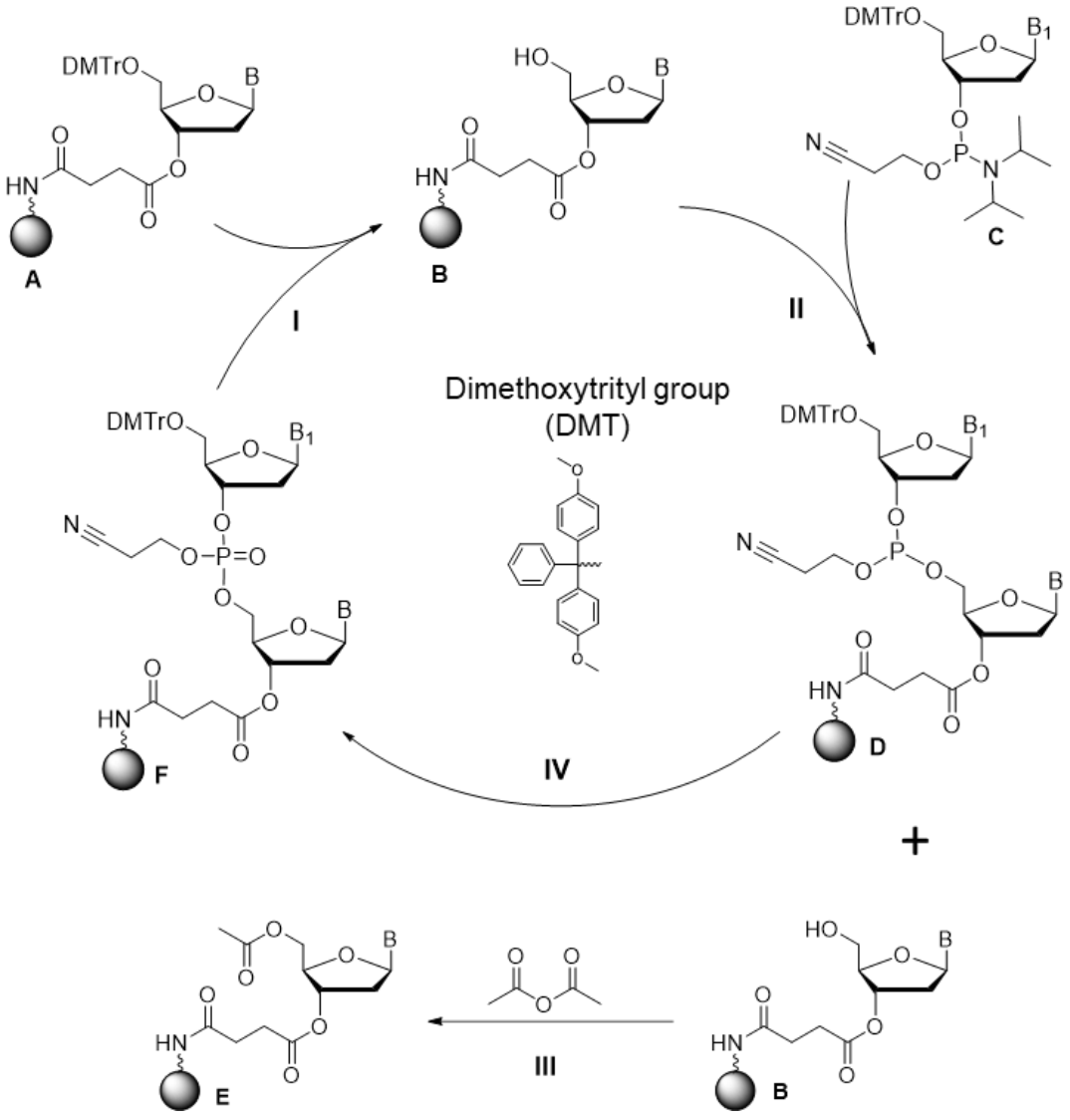


Figure 1.5-Oligonucleotide solid-phase synthesis cycle.

source of fluoride. Therefore, following synthesis, the oligonucleotide can then be cleaved off the solid support, deprotected and purified by denaturing polyacrylamide gel electrophoresis (PAGE) or high-performance liquid chromatography (HPLC).

1.4-Common Characterization Methods for Nucleic Acids

Oligonucleotides exhibit interesting properties which have been used to study them extensively. Oligonucleotides absorb strongly in the 240-280 nm range with a typical absorption maximum at 260 nm.[4] The molar extinction coefficients of individual nucleotides at 260 nm are well known and can therefore be used to approximate the molar extinction coefficient of the oligonucleotide using the nearest neighbor approximation.[13] Hence, by using the Beer-Lambert law it is possible to determine the molar concentration of a sample from its optical density (OD).

1.4.1-Thermal Denaturation Studies

Oligonucleotides exhibit hyperchromicity, which is an increase in absorbance, when a duplex is denatured to its constituent single strands. This property can be used to monitor the thermal stability of an oligonucleotide.[14,15] As temperature increases, for example, this leads to denaturation of the duplex due to the breaking of the H-bonds and disruption of stacking between the complementary base pairs.[14] This results in an increase in the absorbance of the observed species. However, if the temperature is reduced, the complementary strands reanneal resulting in a decrease in absorbance.[15]

From this, the melting temperature (T_m) can be determined, that is the temperature at which 50% of the species is in the duplex and 50% is in the single stranded form (**Figure 1.6**). T_m can primarily be affected by two factors: the GC content of the oligonucleotide and the salt concentration of the buffer.[4,16] The GC content of an oligonucleotide simply refers to what percentage of the oligonucleotide duplex is composed of GC base pairs. Since GC pairs, compared

to AT pairs, are held together via 3 H-bonds they require a higher energy input to be broken therefore raising the T_m.[16] The salt concentration of an oligonucleotide solution is important as well as it provides counter-ions to stabilize the negative charge of the backbone. Therefore, a lower salt concentration results in a lower T_m as there are less cations present and therefore a greater charge repulsion.[4] T_m values are important in order to access the stability of the secondary structure of the oligonucleotide and help to determine which sequences are viable for applications such as their use as gene silencing, antisense oligonucleotides or polymerase chain reaction (PCR) primers.[17]

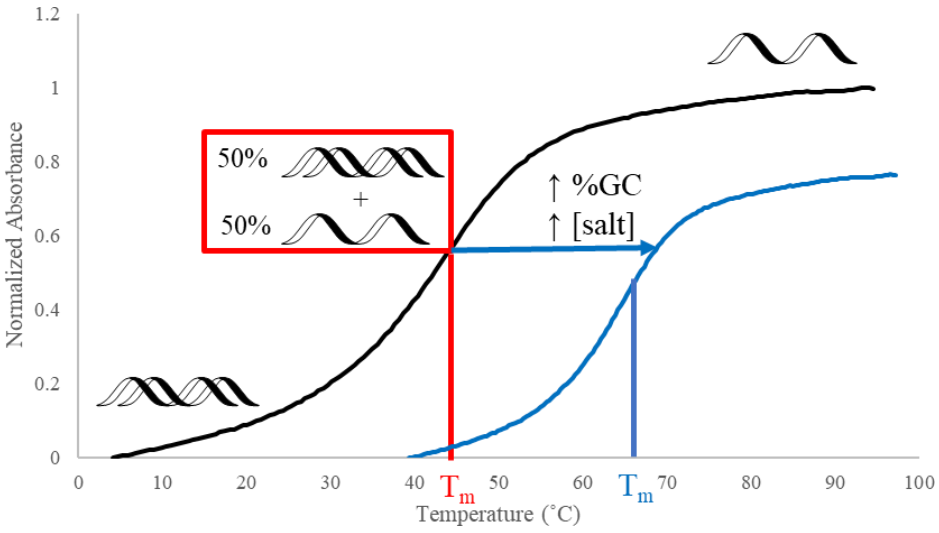


Figure 1.6-Visual representation of effect of salt and GC content on the T_m of an oligonucleotide.

1.4.2-Circular Dichroism Studies

Oligonucleotides are molecules with three main sources of chirality: the asymmetric sugar (C1'-position), the helicity of the secondary structures and tertiary ordering of the oligonucleotide (A- vs B- vs Z-form).[18] Due to this property, circular dichroism (CD) spectroscopy can be utilized to provide information on the tertiary structure of an oligonucleotide as well as monitor any

structural alterations due to changes in temperature, pH and ionic strength.[19] CD, in itself, refers to the absorption of either right-handed or left-handed circularly polarized light by chiral molecules.[18] The observed maximums and minimums allows for insight into the global structure of the oligonucleotide. For example, B-form DNA has a maximum around 240-280 nm and a minimum around 245 nm whereas A-form DNA has a maximum around 260 nm and a minimum around 210 nm. Furthermore, Z-form DNA has a maximum around 260 nm and a minimum around 290 nm.[20] Other structures, like the G-quadruplexes, can also be determined using this technique.

Chapter 2 - Optimization of Strategies to Prepare Circular DNA Using an Interstrand Cross-link

Abstract

Small circular oligonucleotides have shown great promise as both therapeutic and diagnostic tools due to their unique DNA and RNA recognition properties and their increased resistance to endonucleases when compared to their linear counterparts. Such circular oligonucleotides can be synthesized using chemical ligation with cyanogen bromide, enzymatic ligation with T4 DNA ligase or through non-canonical backbone ligation to form a triazole linkage. Major hurdles to the formation of these circular constructs are associated with their length, due to the reduced flexibility of shorter oligonucleotide strands, as well as the difficulty to bring the ligating ends together. Herein, we propose an approach to synthesize circular DNA containing an interstrand cross-link to assist in pre-organizing the system to facilitate ligation and circularization. Upon cleavage and deprotection, the cross-linked oligonucleotide self-assembles into a duplex due to complementary base pairing allowing the ligating ends to be in proximity. These cross-linked duplexes have been prepared with appropriate 5′ and 3′ moieties for ligation via chemical methods. Experiments to optimize the ligation efficiency have been performed. This strategy offers a practical approach to synthesize circular DNA constructs without the need for additional guide strands or purification.

2.1. Introduction

Between the 1990s and today, the accessibility of synthetic oligonucleotides has exploded allowing for the rapid advances in molecular biology and medical diagnostic technology for investigating and identifying specific genetic sequences.[21] Chemically modified oligonucleotides are a relatively new class of therapeutic agents with 15 drugs based on this family of molecules now on the market.[22] However, oligonucleotides face many challenges in their development for therapeutic applications including short half-lives, charged backbones and poor binding. Many research groups have focused on approaches to change the nature and/or the topology of the DNA backbone in an effort to reduce degradation, boost cellular absorption, or increase binding affinity in response to these challenges.[21] One of the highly interesting topological structures that has been examined is circular DNA.

2.1.1-Circular Oligonucleotides

Circular DNAs are a group of ssDNA defined by their closed topology.[23] On top of the well-known advantages of nucleic acid-based materials, that is, biocompatibility, low cost and ease of chemical modification, circular oligonucleotides exhibit a higher resistance to exonucleases compared with their linear counterparts which enhances their biostability.[24] Furthermore, they exhibit unique DNA binding properties allowing for the formation of highly stable triplexes. DNA triplexes are formed between a single stranded DNA (ssDNA) strand and double stranded DNA (dsDNA) where the third strand (ssDNA) can adopt a parallel or anti-parallel motif (**Figure 2.1-A**).[21,25] Triplexes formed with circular DNA can adopt two topologies: (1) the circular construct binds to itself on either the Watson-Crick or the Hoogsteen face and to the target ssDNA strand on the other face or (2) the circular dsDNA encapsulates the ssDNA allowing for binding on both the Watson-Crick and Hoogsteen faces of the ssDNA (**Figure 2.1-B & C**).[25]

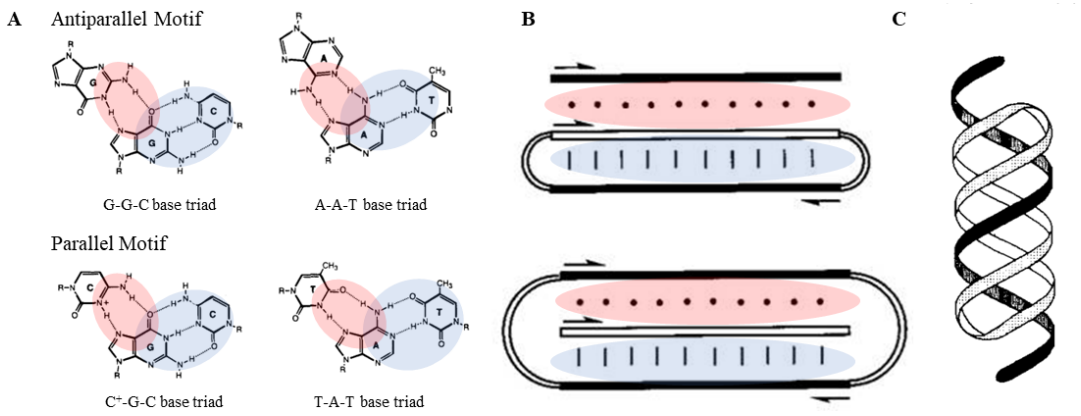


Figure 2.1-(**A**) Base triads highlighting the Watson-Crick (blue) and Hoogsteen (red) faces of antiparallel and parallel motifs, (**B**) Possible triplex formation of circular oligonucleotides with pairing on the WC and Hoogsteen faces, (**C**) Visual representation of circle formation a triplex with ssDNA (reproduced and adapted with permission from *references 21 and 25*).

Moreover, circular DNA and RNA constructs have been used for rolling circle amplification (RCA) and rolling circle transcription (RCT). These key biological techniques have been investigated as models for bioanalytical signal amplification and the manufacture of nano-, micro-, and macro-biomaterials.[23] RCA and RCT allow for the rapid and efficient amplification of DNA and RNA in the presence of appropriate (d)NTPs, primers and polymerases without the limitations of other amplification techniques like PCR.[26] Lastly, due to their unique binding abilities, circular DNA constructs have shown great promise as both a diagnostic and therapeutic tool. In the past few decades, groups have shown that circular constructs can be used as multitarget recognition probes, biological decoys, aptamers, a template for nanostructures, microRNA (miRNA) inhibitors, antisense oligonucleotides, scaffolds for DNA origami, CRISPR-Cas gene editing donors and to study noncanonical DNA structural motifs.[23,25,27,28]

2.1.2-Synthesis of Circular Constructs

The synthesis of single-stranded DNA circles, however, remains a challenge due to the unfavorable entropy of the macrocycle formation.[27] This is thought to be due to two major factors: the length of the oligonucleotide where shorter oligonucleotides exhibit reduced flexibility and the proximity of the ligating ends whereas if the ligating ends are far away (either due to the folding of the oligonucleotide or its length) then the reaction needed to ligate these ends is hindered.[23] To further assist the intramolecular cyclization, which competes severely with the intermolecular multimerization, low concentrations of oligonucleotides are used. Furthermore, splints, guide strands, cross-links and base complementarity are utilized to help pre-organize the system into the desired configuration as non-templated cyclization of oligonucleotides longer than a 20mer tend to be unsuccessful.[29]

Nevertheless, oligonucleotides larger than 28 nucleotides can be cyclized via enzymatic or chemical means. Enzymatically, circular constructs may be ligated through the use of a DNA or RNA ligase which results in a natural phosphodiester backbone at the ligation site. On the other hand, chemical agents may also be used to ligate the 5’ and 3’ ends of the oligonucleotides. Examples of these include cyanogen bromide and the copper (I) catalyzed alkyl-azide cycloaddition (CuAAC) reaction where the latter results in a natural backbone and the former in a triazole linkage.

2.1.3-Enzymatic Circularization Approaches Using Ligases

T4 DNA and RNA ligases are enzymes that catalyze the reaction between a 5’-phosphate and a 3’hydroxyl group in either DNA or RNA yielding a phosphodiester bond in the presence of adenosine triphosphate (ATP).[29] Specifically, T4 DNA ligase will join blunt or overhanging ends in dsDNA, however, it will not work with ssDNA. DNA ligases function via a 3-step mechanism:

(1) formation of an enzyme-adenosine monophosphate (AMP) intermediate via residue Lys159 and subsequent release of pyrophosphate, (2) transfer of AMP to the 5'-phosphate of the oligonucleotide resulting in a DNA-AMP intermediate and (3) formation of the phosphodiester bond with release of AMP.[30,31] Due to the simplicity and wide availability of this enzyme, its use has become widespread in biology and biotechnology however it does have some disadvantages. Large scale enzymatic ligation via T4 DNA ligase is limited in the fact that there is a certain ratio of enzyme-to-oligonucleotide that can be used and there is significant multimeric by-product formation.[28] Furthermore, due to the nature of the enzyme's active site, it can only accept oligonucleotides with regular topologies.[32]

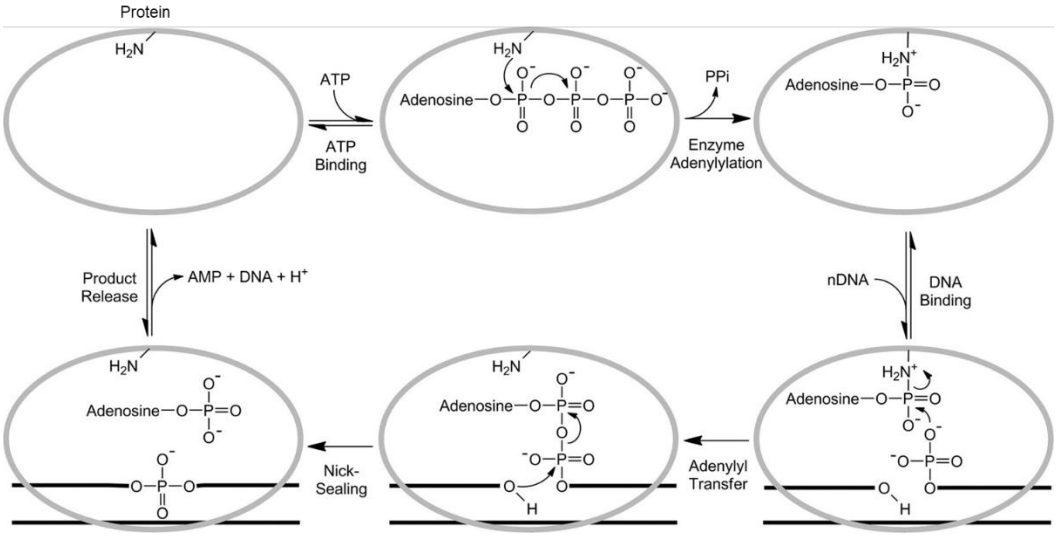


Figure 2.2-Mechanism of T4 DNA Ligase (reproduced and adapted from *reference 31*).

2.1.4-Chemical Circularization

Due to the unpredictable nature, low yield and poor scalability of enzymatic ligation, biochemists have turned their attention to chemical ligation as a means to achieve circularization. The process of chemical ligation was originally used as a means to assemble extended DNA fragments on a complementary strand and internucleotide bond formation would ensue. Today, this technique has been widely used in nucleic acid chemistry. This method allows for the

preparation of branched DNA, circular DNA and chimeric oligonucleotides.[33] Currently, there exist two options for phosphodiester bond formation via chemical ligation. The less common option involves the activation of a phosphomonoester group. Usually reactive imidazolide or N-hydroxybenzotriazolide-derived oligonucleotides are used as the terminal phosphomonoester resulting in an activated phosphomonoester.[33] The second, and more commonly used option is the use of a condensing reagent.

2.1.4.1-Cyanogen Bromide Ligation

Chemical ligation of oligonucleotides has been performed using condensing agents such as carbodiimides in order to activate a phosphate group at the ligation site.[33] However, this reaction is slow with some cases taking up to 6 days for efficient ligation.[34] More recently, cyanogen bromide (CNBr) has been used to ligate the 5' and 3' ends of oligonucleotides. Currently, there are two CNBr-induced ligation methods, the first employs the use of CNBr in the presence of imidazole and the other uses a N-morpholioethanesulfonate (MES) buffer. In the presence of imidazole, the ligation products are formed within 3 to 20 h, which is comparable to the carbodiimide method previously described.[33] This slow reaction rate allowed for the isolation of the condensing agent responsible for this reaction, namely the cyanoimidazole intermediate. On the other hand, in the presence of MES, the ligation products are obtained within minutes rather than hours and the ligation efficiency is 5x times higher (compared to similar experiments in the presence of imidazole) which proves to be a great advantage.[33,34]

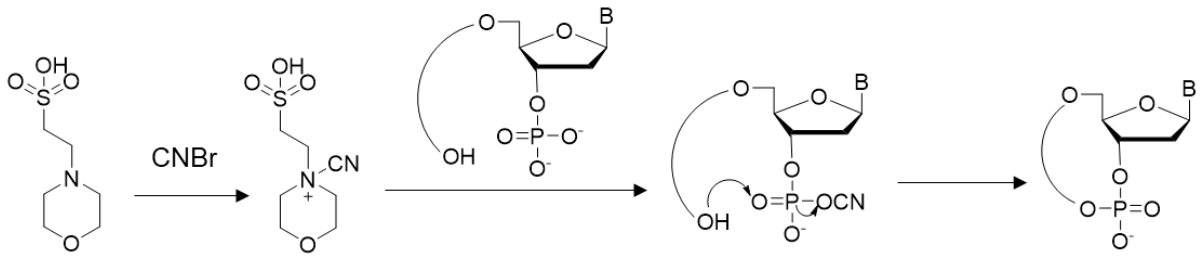


Figure 2.3-Circularization of an oligonucleotide using CNBr in MES buffer.

The ligation mechanism for CNBr in the presence of MES occurs in three steps (**Figure 2.3**). Initially, the cyanogen bromide with MES, a tertiary amine, forms a water-soluble quaternary ammonium base. This intermediate then reacts very quickly with the monosubstituted phosphate resulting in the formation of a phosphorocyanatidic acid-derived oligonucleotide. The free hydroxyl group can then attack the phosphate displacing the cyanate resulting in formation of a phosphodiester bond.[29,33]

This reaction evolves hydrobromic acid (HBr) which can be damaging to oligonucleotides however the Shabarova group found that if a concentration of MES of 0.25 M or higher is used then the pH of the reaction remains relatively stable hence reducing the risk of depurination and phosphodiester bond cleavage.[33,35]

2.1.4.2-Click Chemistry Ligation

The Sharpless group coined the phrase "click chemistry" in the late 1990s to describe a high yielding process for attaching a probe or substrate to a biomolecule via bioconjugation with no side products observed.[36] To date, click chemistry has been used in a wide array of fields ranging from polymer science to bioengineering. Specific guidelines have been set to classify a reaction or process as a click reaction. The process must be modular, wide in scope, very high yielding, generate only inoffensive by-products (which can be easily removed via non-chromatographic means) and be stereospecific (not necessarily enantioselective). Furthermore, simple reaction conditions insensitive to oxygen and water are needed, readily available starting materials and reagents, no or benign or easily removable solvents, simple product isolation and purification and the product should be stable under physiological conditions.[36] A handful of reactions fit this profile including the [3,2]-cycloaddition (Huisgen 1,3-dipolar cycloaddition), thiol-ene reaction and the Diels-Alder reaction.[36]

Over the past 30 years, the Huisgen 1,3-dipolar cycloaddition has been widely used and popularized across numerous fields including organic, biological and polymer chemistry. Initially, this reaction was not considered a click reaction as it yields a mix of two regio-isomers (the 1,4 and the 1,5-product) under thermodynamic conditions which is against the guideline about stereospecificity set forth by Sharpless.[37] The Meldal and Sharpless groups independently found that the addition of a copper (I) source to this reaction resulted in the production of only the 1,4-isomer rendering the reaction regiospecific.[37] The most studied metal ion catalyzed Huisgen reaction is the CuAAC reaction.[38]

Due to the unreactive nature of the azide and alkyne precursor as well as the easy incorporation of these groups into biological scaffolds, click chemistry has been used in a wide range of fields. In the realm of nucleic acids, researchers have been using this approach to synthesize base- or sugar-modified nucleotides, nucleosides, and oligonucleotides efficiently and with lower cost relative to other chemical and enzymatic techniques.[37] Applications include single strand DNA ligation, interstrand and intrastrand cross-links, self-templating DNA, cyclization and bioconjugation of dyes and small molecules.[36–39]

Even though the CuAAC reaction has been studied for two decades, the mechanism is still debated. Computational and kinetic studies have determined two possible mechanisms (**Figure 2.4**) either a slow, mononuclear Cu-catalyzed pathway or a fast, di-nuclear Cu-catalyzed pathway. In both cases, the first step is the replacement of the Cu(I) donor with the alkyne resulting in the formation of a π bond between the alkyne and the metal. Deprotonation of the alkyne to form a Cu-acetylide adduct ensues. Coordination of the azide to the Cu(I) core in the slow pathway or to the second Cu (I) core in the fast pathway allows for the endothermic addition of the third nitrogen to C2 of the acetylide. It has been said that copper plays a crucial role in promoting the C-N bond

formation by reducing the activation barrier. Lastly, the ring contracts to form the triazolyl-copper derivative which, after protolysis, releases the desired triazole product.[38]

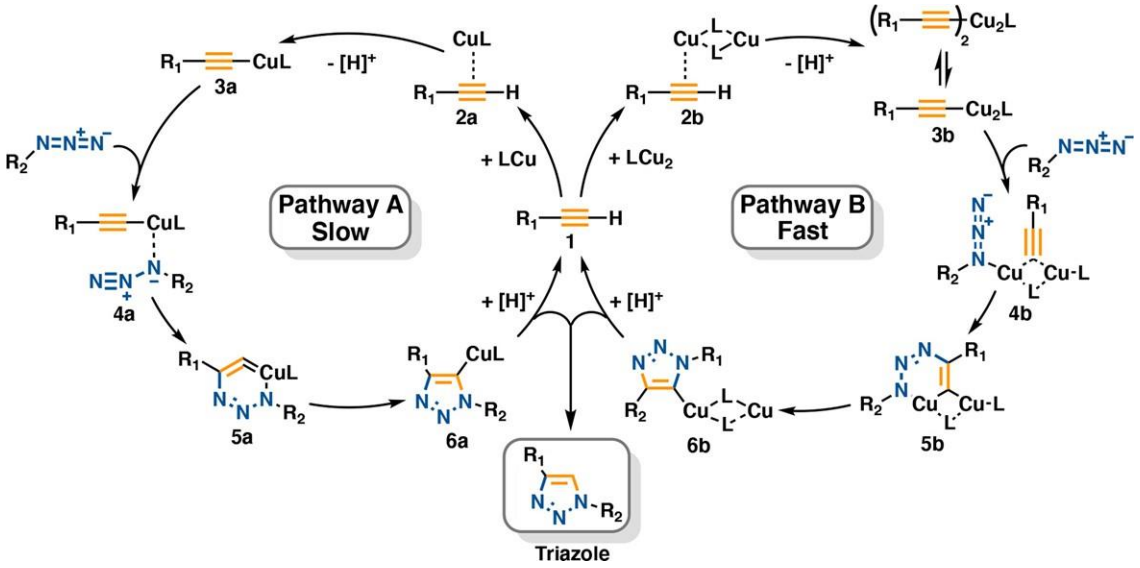


Figure 2.4-Proposed mechanisms of the click reaction (reproduced with permission from *reference 38*).

2.1.5-Project Objective

Despite many benefits presented by circular DNAs, major drawbacks regarding their preparation persist. Preparing pure and high yielding circular constructs remains the major challenge as most ligations require a high-fidelity template or splint which is difficult to remove due to base pairing.[28] This results in repeated gel electrophoresis and exonuclease digests to remove the extra guide strand which reduces the overall yield.

Hence the objective of this project is to develop a method to synthesize circular DNA constructs containing an interstrand cross-link (**Figure 2.5**). Initially, an uncleavable cross-link was studied for preliminary reaction optimization with hopes of introducing a cleavable sulfone cross-link. Starting from the solid support bound oligonucleotide, the cross-link is introduced during solid-phase synthesis, resulting in a covalent bond between the two oligonucleotide strands. Following solid-phase synthesis, the resulting oligonucleotide can then be deprotected and cleaved

from the solid support. Due to the base complementarity, the oligonucleotide is able to self-assemble which brings the 5' and 3' ligating ends much closer to each other. Upon ligation, the cross-link can be removed. This method would have great advantages including the pre-organization of the system to facilitate ligation in the absence of a splint or guide strand and the ability to build large circular constructs from shorter oligonucleotides.

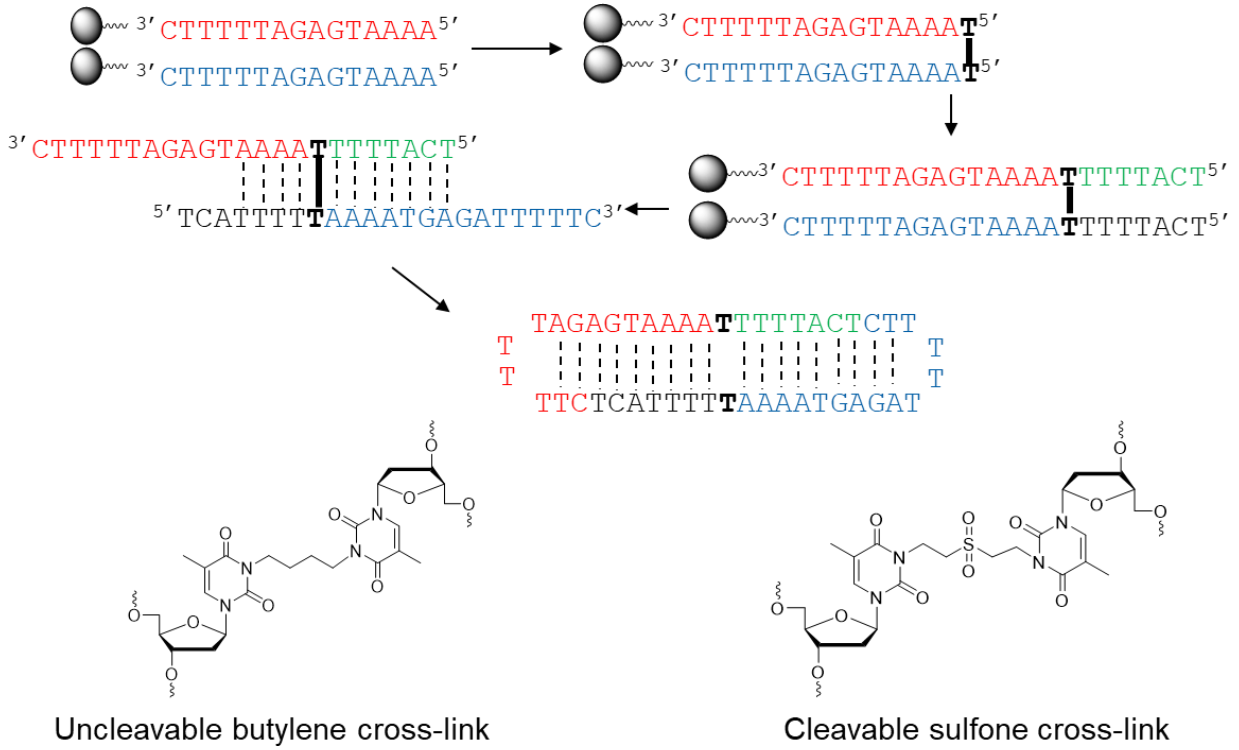


Figure 2.5-Visual representation of the Project Objectives for Chapter 2.

2.2-Experimental

Unless otherwise stated, all chemicals and solvents were purchased from MilliporeSigma (St. Louis, MO) or Thermo Fisher Scientific (Waltham, MA). All nucleoside derivatives were purchased from ChemGenes Corporation (Wilmington, MA). Dichloromethane (DCM), methanol (MeOH), ethyl acetate (EtOAc), hexanes (Hex), tetrahydrofuran (THF), dimethylformamide (DMF), and acetonitrile (ACN), were used as purchased unless otherwise stated. Flash column

chromatography (FCC) was performed using silica gel 60 (230–400 mesh) purchased from Silicycle (Quebec City, QC). Thin layer chromatography (TLC) was carried out with precoated TLC plates (Merck, Kieselgel 60 F_{254}, 0.25 mm) purchased from EMD Chemicals Inc. (Gibbstown, NJ). NMR spectra were recorded on a Varian 500 MHz NMR spectrometer at room temperature. ^{1}H NMR spectra were recorded at a frequency of 500.0 MHz and chemical shifts were reported in parts per million (ppm) downfield from tetramethylsilane. ^{13}C NMR spectra (^{1}H decoupled) were recorded at a frequency of 125.7 MHz and chemical shifts were reported in ppm with tetramethylsilane as a reference. ^{31}P NMR spectra (^{1}H decoupled) were recorded at a frequency of 202.3 MHz and chemical shifts were reported in ppm with H_3PO_4 used as an external standard. High resolution mass spectrometry of modified nucleosides was obtained using an 7T-LTQ FT ICR instrument (Thermo Scientific), at the Concordia University Centre for Structural and Functional Genomics. The mass spectrometer was operated in full scan, positive ion detection mode. The samples were dissolved in 50% ACN/MeOH and injected directly.

2.2.1-Synthesis of Modified Nucleosides

2.2.1.1-Preparation of N3-Butylene Linked Thymidine Bis-phosphoramidite

* Compounds 2-4 and were prepared and characterized as previously described.[40]

1-{N^3-[5'-O-(4,4'-dimethoxytrityl)-3'-O-*tert*-butyldimethylsilyl-2'-deoxythymidine]}-4-{N^3-[5'-O-(4,4'-dimethoxytrityl)-3'-O-*tert*-butyldimethylsilyl-2'-deoxythymidine]}-butane (**5**). To a solution of 5'-O-(4,4'-dimethoxytrityl)-N^3-(4-iodobutyl)-3'-O-*tert*-butyldimethylsilyl-2'-deoxythymidine[40] (**4**, 1.12 g, 1.33 mmol) in anhydrous ACN (5.0 mL), was added a solution of **2** (0.921 g, 1.40 mmol) and DBU (398 µl, 2.66 mmol) in anhydrous ACN (15.0 mL). The reaction was allowed to stir for 24 hours at room temperature. The resulting solution was concentrated *in vacuo* and taken up with DCM (70 mL) and subsequently washed with 3% sodium bicarbonate (2

x 70 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated *in vacuo*. The resulting oil was purified via flash column chromatography (FCC) using a gradient of DCM/MeOH (100/0 to 99/1) yielding a white powder (1.44 g, 1.05 mmol, 80%). R_f (SiO_2 TLC): 0.53 EtOAc/Hex (4/6). λ_{max} (ACN): 204 nm. ^{1}H NMR (500 MHz, $CDCl_3$, ppm): δ 7.64 (s, 1H, H6), 7.42-7.41 (m, 2H, Ar), 7.32-7.28 (m, 6H, Ar), 7.25-7.21 (m, 1H, Ar), 6.85-6.85 (m, 4H, Ar), 6.36 (t, 1H, H1′, $J_{1'-2'}$ = 6.4 Hz), 4.52-4.50 (m, 1H, H3′), 3.97-3.94 (m 3H, H4′, NCH_2), 3.78 (s, 6H, 2 x OCH_3), 3.47 (dd, 1H, H5′, $J_{5'-5''}$ = 10.7 Hz, $J_{5'-4'}$ = 2.7 Hz), 3.25 (dd, 1H, H5″, $J_{5''-5'}$ = 10.7 Hz, $J_{5''-4'}$ = 2.7 Hz), 2.35-2.18 (m, 2H, H2′ and H2″), 1.72-1.68 (m, 2H, $NCH_2C\mathbf{H_2}$), 1.49 (s, 3H, C5-Me), 0.82 (s, 9H, $Si(CH_3)_3$), 0.01 (s, 3H, $SiCH_3$), -0.04 (s, 3H, $SiCH_3$). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): δ 163.5, 158.7, 150.8, 144.3, 135.5, 135.5, 133.5, 130.1, 130.0, 128.1, 127.9, 127.1, 113.3, 113.2, 110.2, 86.8, 86.6, 85.4, 71.9, 62.8, 60.4, 55.2, 41.6, 41.1, 25.7, 25.4, 17.9, 14.2, 12.7, -4.7, -4.9. HRMS (ESI-MS) *m*/*z* calculated for $C_{78}H_{98}N_4O_{14}Si_2\,NH_4$ 1388.6962: found 1388.6964 [$M + NH_4^+$].

1-{N^3-[5'-O-(4,4'-dimethoxytrityl)-2'-deoxythymidine]}-4-{N^3-[5'-O-(4,4'-dimethoxytrityl)-2'-deoxythymidine]}-butane (6). To a solution of **(5)** (1.44 g, 1.05 mmol) in anhydrous THF (10 mL), was added 1.0 M TBAF in THF (2.62 mL, 2.62 mmol). The reaction was allowed to stir for 20 minutes at room temperature. The resulting solution was concentrated *in vacuo* and taken up with DCM (70 mL) and subsequently washed with 3% sodium bicarbonate (2 x 70 mL). The ensuing solution was dried over anhydrous sodium sulfate and concentrated *in vacuo*. The resulting oil was purified via FCC using EtOAc/Hex (7/3) yielding a white powder (0.843 g, 0.746 mmol, 71%). R_f (SiO_2 TLC): 0.17 EtOAc/Hex (7/3). λ_{max} (ACN): 204 nm. ^{1}H NMR (500 MHz, $CDCl_3$, ppm): δ 7.59 (s, 1H, H6), 7.40-7.18 (m, 9H, Ar), 6.82-6.80 (m, 4H, Ar), 6.42-6.39 (m, 1H, H1′), 4.55-4.54 (m, 1H, H3′), 4.05-4.03 (m, 1H, H4′), 3.98-3.88 (m, 2H, NCH_2), 3.75 (s, 6H, 2 x OCH_3),

3.47-3.44 (m, 1H, H5′), 3.36-3.33 (m, 1H, H5″), 2.42-2.27 (m, 2H, H2′ and H2″), 1.67-1.63 (m, 2H, $NCH_2C\mathbf{H_2}$), 1.44 (s, 3H, C5-Me). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): δ 171.3, 163.5, 158.6, 150.9, 144.4, 135.5, 135.4, 133.8, 130.1, 130.1, 128.1, 127.9, 127.1, 113.3, 113.3, 110.3, 86.8, 86.1, 85.3, 72.2, 63.6, 60.5, 55.2, 55.1, 41.1, 25.1, 21.0, 20.7, 14.2, 14.1, 12.6. HRMS (ESI-MS) *m*/*z* calculated for $C_{66}H_{70}N_4O_{14}\,NH_4$1160.5232: found 1160.5232 [M + NH_4^+].

1-{N^3-[3'-O-(2-cyanoethoxy(diisopropylamino)-phosphino)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxythymidine]}-4-{N^3-[3'-O-(2-cyanoethoxy(diisopropylamino)-phosphino)-5'-O-(4,4'-dimethoxytrityl)- 2'-deoxythymidine]}-butane (7). To a solution of **(6)** (0.30 g, 0.266 mmol) in anhydrous THF (5.0 mL), was added DIPEA (139 µl, 0.79 mmol) and diisopropylamino cyanoethyl phosphonamidic chloride (149 µl, 0.638 mmol) dropwise. The reaction was allowed to stir for 30 minutes at room temperature. The resulting solution was concentrated *in vacuo* and taken up in EtOAc (40 mL). The solution was washed with 3% sodium bicarbonate (2 x 30 mL) and brine (1 x 30 mL). The ensuing solution was dried over anhydrous sodium sulfate and concentrated *in vacuo*. The resulting oil was purified via FCC using EtOAc/Hex (1/1) yielding a white powder (0.26 g, 0.17 mmol, 64%). R_f (SiO_2 TLC): 0.186, 0.256, 0.326, ethyl acetate/hexanes (4.5/5.5). λ_{max} (ACN): 203 nm. ^1H NMR (500 MHz, acetone-d6, ppm): δ 7.67-7.64 (m, 1H, H6), 7.52-7.31 (m, 9H, Ar), 6.92-6.89 (m, 4H, Ar), 6.42-6.40 (m, 1H, H1′), 4.77-4.74 (m, 1H, H3′), 4.22-4.16 (m, 1H, H4′), 3.94-3.60 (m, 12H, OCH_2, NCH_2, 2 x NCH, 2 x OCH_3), 3.47-3.40 (m, 2H, H5′ and H5″), 2.77-2.61 (m, 2H, $CNCH_2$), 2.52-2.48 (m, 2H, H2′ and H2″), 1.64-1.60 (m, 2H, $NCH_2C\mathbf{H_2}$), 1.52-1.51 (m, 3H, C5-Me), 1.20-1.09 (m, 12H, 2 x $CH(C\mathbf{H_3})_2$). ^{31}P NMR (202.3 MHz, acetone-d6, ppm): δ 148.2, 148.1. ^{13}C NMR (125.7 MHz, acetone-d6, ppm): δ 162.75, 158.87, 158.85, 150.65, 144.94, 144.93, 135.65, 135.62, 135.58, 135.53, 133.99,

133.90, 130.17, 130.14, 128.18, 128.13, 127.86, 126.90, 126.88, 118.10, 117.93, 113.14, 109.45, 109.40, 86.64, 86.60, 85.33, 85.29, 85.10, 85.05, 73.77, 73.45, 63.40, 63.29, 58.69, 58.64, 58.54, 58.49, 54.70, 54.68, 43.10, 43.00, 40.48, 39.50, 39.44, 25.03, 24.05, 24.02, 24.00, 23.99, 23.96, 19.93, 19.87, 19.86, 19.80, 12.09. HRMS (ESI-MS) *m/z* calculated for $C_{85}H_{108}N_8O_{16}P_2$ $NH_4$1560.7389: found 1560.7318 [M + NH_4^+].

2.2.1.2-Preparation of N3-1-Sulfonylethane Linked Thymidine Bis-phosphoramidite

1-{N^3-[5'-O-(4,4'-dimethoxytrityl)-2'-deoxythymidine]}-2-(2-{N^3-[5'-O-(4,4'-dimethoxytrityl)-2'-deoxythymidine]}-1-sulfonyl)ethane (**8**)**.** To a solution of commercially available 5'-O-(4,4'-Dimethoxytrityl)-2'-deoxythymidine (**1**, 0.40 g, 0.73 mmol) in anhydrous ACN (2.0 mL), was added divinyl sulfone (331 µl, 0.33 mmol) and tributylphosphine (16.5 µl, 0.07 mmol), the reaction was allowed to stir overnight at 50 ˚C. The resulting solution was taken up in DCM (70 mL) and washed with 3% sodium bicarbonate (2 x 70 mL). The ensuing solution was dried over anhydrous sodium sulfate and concentrated *in vacuo*. The resulting oil was purified *via* FCC using a gradient of MeOH/DCM (0/100 to 3/100) yielding a white powder (0.36 g, 0.31 mmol, 92%). R_f (SiO_2 TLC): 0.38 MeOH/DCM (5/95). λ_{max} (ACN): 204 nm. ^{1}H NMR (500 MHz, $CDCl_3$, ppm): δ 7.60 (s, 1H, H6), 7.39-7.21 (m, 9H, Ar), 6.84-6.82 (m, 4H, Ar), 6.40-6.37 (m, 1H, H1′), 4.58-4.53 (m, 1H, H3′), 4.46-4.35 (m, 2H, NCH_2), 4.06-4.03 (m, 1H, H4′), 3.78 (s, 6H, 2 x OCH_3), 3.57-3.50 (m, 2H, SCH_2), 3.48-3.34 (m, 2H, H5′ and H5″), 2.44-2.27 (m, 3H, H2′, H2″, and OH), 1.47 (s, 3H, C5-Me). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): δ 163.15, 158.69, 150.61, 144.27, 135.37, 135.30, 134.22, 130.06, 128.10, 128.00, 127.14, 113.28, 113.27, 110.22, 86.93, 86.05, 85.51, 72.26, 63.43, 55.25, 49.04, 41.04, 35.00, 12.48. HRMS (ESI-MS) *m/z* calculated for $C_{66}H_{70}N_4O_{16}S$ $NH_4$1224.4851: found 1224.4879 [M + NH_4^+].

1- {N^3-[3'-O-(2-cyanoethoxy(diisopropylamino)-phosphino)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxythymidine]}-2-(2-{N^3-[3'-O-(2-cyanoethoxy(diisopropylamino)-phosphino)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxythymidine]}-1-sulfonyl)ethane (9). To a solution of **8** (0.36 g, 0.30 mmol) in anhydrous THF (4.0 mL), was added DIPEA (159 µL, 0.91 mmol) and diisopropylamino cyanoethyl phosphonamidic chloride (162 µL, 0.730 mmol) dropwise. The reaction was allowed to stir for 30 minutes at room temperature. The resulting solution was concentrated *in vacuo* and taken up in EtOAC (50 mL). The solution was washed with 3 % sodium bicarbonate (2 x 50 mL) and brine (1 x 50 mL). The ensuing solution was dried over anhydrous sodium sulfate and concentrated *in vacuo*. The resulting oil was purified via FCC using EtOAc/Hex (1/1) yielding a white powder (0.22 g, 0.14 mmol, 46%). R_f (SiO_2 TLC): 0.49, 0.56, 0.64, EtOAc/Hex (6/4). λ_{max} (ACN): 207 nm. ^{1}H NMR (500 MHz, acetone-d6, ppm): δ 7.71-7.68 (m, 1H, H6), 7.53-7.24 (m, 9H, Ar), 6.92-6.89 (m, 4H, Ar), 6.39-6.38 (m, 1H, H1′), 4.80-4.74 (m, 1H, H3′), 4.44-4.36 (m, 2H, NCH_2), 4.23-4.17 (m, 1H, H4′), 3.93-3.39 (m, 14H, OCH_2, SCH_2, 2 x NCH, 2 x OCH_3, H5′ and H5″), 2.76-2.61 (m, 2H, $CNCH_2$), 2.59-2.40 (m, 2H, H2′ and H2″), 1.54-1.53 (m, 3H, C5-Me), 1.21-1.09 (m, 12H, 2 x CH(C**H$_3$**)$_2$). ^{31}P NMR (202.3 MHz, acetone-d6, ppm): δ 148.2, 148.1. ^{13}C NMR (125.7 MHz, acetone-d6, ppm): δ 162.66, 158.87, 158.86, 150.51, 144.92, 135.63, 135.60, 135.55, 135.51, 134.42, 134.33, 130.19, 130.16, 128.19, 128.15, 127.88, 126.95, 126.92, 118.17, 117.99, 113.17, 109.46, 109.42, 86.67, 86.63, 85.54, 85.49, 85.43, 85.40, 85.14, 85.10, 73.38, 73.05, 63.27, 63.17, 58.69, 58.66, 58.53, 58.51, 54.76, 54.74, 49.22, 43.12, 43.02, 39.59, 39.46, 34.53, 28.61, 25.33, 24.10, 24.06, 24.04, 24.01, 19.98, 19.92, 19.90, 19.84, 12.05. HRMS (ESI-MS) *m*/*z* calculated for $C_{85}H_{108}N_8O_{18}P_2S$ $NH_4$1624.7008: found 1624.7013 [M + NH_4^+].

2.2.2-General Synthesis of Oligonucleotides

All oligonucleotides were synthesized using an Applied Biosystems Model 3400 automated DNA synthesizer via standard β-cyanoethylphosphoramidite chemistry. All reagents and standard protected nucleoside phosphoramidites were purchased from Glen Research (Sterling, VA). All oligonucleotide sequences were synthesized on a 1.5 μmole scale using long chain alkylamine- controlled pored glass (CPG, 500 Å). Oligonucleotides used for click ligation were synthesized on 500 Å LCAA-CPG derivatized with 5'-O-(4,4'-dimethoxytrityl)-3'-O-propargyl-2'-deoxycytidine at the N^4-atom. Standard nucleoside phosphoramidites were dissolved in anhydrous ACN to a final concentration of 0.1 M. Custom phosphoramidite dimers were dissolved at a concentration of 0.05 M, while all other custom phosphoramidites were dissolved at a concentration of 0.15 M. Oligonucleotides containing the sulfone cross-link were synthesized using fast-deprotecting amidites: phenoxyacetyl-dA (Pac-dA) and isopropyl-phenoxyacetyl-dG (iPr-Pac-dG) phosphoramidites from Glen Research at a concentration of 0.12 M. Phenoxyacetic anhydride was also used as a capping agent in this case. The assembly of the oligonucleotide sequences followed the standard synthetic cycle initiated by detritylation of the 5′-O-dimethoxytrityl protecting group of the support bound base using TCA/DCM (3 %, v/v) followed by coupling of the incoming phosphoramidite using 5-ethylthio-1H-tetrazole (0.45 M in ACN). A coupling time of 120 seconds was used for the standard phosphoramidites whereas coupling time was increased to 450 seconds for all custom phosphoramidites except for phosphoramidite dimers which were coupled for 30 minutes. Iodo-functionalized oligonucleotides were synthesized using 5′-iodo-thymidine-3′-[(2- cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (Glen Research) and used according to the manufacturers protocol. The capping step involved treatment with acetic anhydride/pyridine/THF (1:1:8, v/v) followed by oxidation using 0.02 M iodine in water/pyridine/THF (2:1:2.5, v/v).

2.2.3-Conversion of the 5'-Iodo to 5'-Azido Group

Post synthesis, oligonucleotides containing a terminal 5'iodo-thymidine moiety underwent treatment with a prepared solution of sodium azide (50 mg, 0.77 mmol) in 1 mL of DMF. The support bound oligonucleotide was transferred to a 1.5 mL screw cap vial, the above-mentioned solution was added, and the resulting suspension was vortexed, centrifuged and placed in a sand bath at 55°C for 5h. The resulting CPG was filtered, then washed with DMF (2 x 10 mL) followed by ACN (2 x 10 mL) and air dried.

2.2.4-Deprotection of Oligonucleotides

Following synthesis, the support bound oligonucleotides were transferred to a clean 1.5 mL screw cap vial. Deprotection and cleavage of the oligonucleotide was achieved through treatment with 1 mL of ammonia/ethanol (3:1, v/v). The resulting suspension was vortexed and centrifuged, then placed in a sand bath at 55°C for 16h. The supernatant was transferred to a new screw cap vial and the CPG was washed with 50% ACN/water (2 x 200 µL) transferring the supernatant after every wash. The combined supernatant was dried down in a Thermo Savant SC110A SpeedVac concentrator. The dried, crude oligonucleotide was dissolved in 1 mL of water and a crude quantification performed by measuring the absorbance at λ=260 nm using a Varian CARY model 3E spectrophotometer.

For oligonucleotides synthesized with fast-deprotecting amidites, a solution of 0.05 M K_2CO_3/MeOH was used to deprotect overnight at room temperature and neutralized with glacial acetic acid rather than the above-mentioned ammonia/ethanol (3:1, v/v).

2.2.5-Polyacrylamide Gel Electrophoresis

PAGE was used for the analysis and purification of all oligonucleotides used in this project. A 20% PAGE solution composed of 7 M urea (168 g, 2.8 mol), bisacrylamide (4 g, 26

mmol), acrylamide (76 g, 1.1 mol), 80 mL 5X TBE (0.45 M tris, 0.44 mM boric acid and 6.4 mM EDTA) and diluted to 400 mL with distilled water was prepared. For analytical purposes 5 mL of the above-mentioned solution was used, a 2:1 (20 μL : 10 μL) ratio of ammonium persulfate (APS, 20%, v/v) to tetramethylethylenediamine (TEMED) was used to initiate polymerization. The solution was poured between two glass plates and allowed to polymerize for 20 minutes. Following polymerization, the wells were washed with 1X TBE buffer and placed in the gel chamber filled with 1X TBE buffer. The wells are loaded with 0.1-0.5 ODs of material dissolved in 10 μL of formamide. The gel was run at 250V for the desired amount of time. Following electrophoresis, the oligonucleotides are visualized under UV light using a Spectroline model CL-150 ultraviolet fluorescence analysis cabinet. For purification purposes, 40 mL of the 20% PAGE solution is used with the same ratio of APS to TEMED (200 μL:100 μL). The resulting solution is poured between glass plates and the well loaded with a maximum of 60 ODs in 100 μL of formamide. The gel was run at 450V until separation is achieved. Following electrophoresis, the oligonucleotide is visualized under UV light using a Spectroline model CL-150 ultraviolet fluorescence analysis cabinet. The bands of interest are excised from the gel using a scalpel and transferred to a 15 mL Falcon™ tube. The tube was shaken using a nutator for 16h in a solution of 0.1 M sodium acetate (8 mL) to elute the oligonucleotide.

2.2.6-Desalting

The purified oligonucleotide was desalted using a C-18 SepPak (360 mg) cartridge fitted to a 10 mL syringe and a 3-way value (syringe-value-cartridge). The cartridge was equilibrated with 100% ACN (10 mL), 50% ACN/water (10 mL) and 0.1 M sodium acetate (10 mL). The oligonucleotide was loaded onto the C-18 column and the flow-through collected in the original 15 mL Falcon™ tube; this is repeated five times. The flow through was analyzed using a Varian

Cary Model 3E spectrophotometer at A_{260}. The column was washed with Milli-Q® ultrapure water (3 x 10 mL) with each flow through collected, combined, then analyzed using a Varian CARY model 3E spectrophotometer at 260 nm. The oligonucleotide was eluted from the cartridge using 5mL of an ACN/water/MeOH (2:1:1, v/v) mixture in a 5 mL Falcon™ tube. The ensuing solution was then dried down using a Thermo Savant SC110A SpeedVac concentrator. The dry oligonucleotide was dissolved in 18 Milli-Q® ultrapure water and quantified by monitoring the absorbance at 260 nm.

2.2.7-Ligation via CNBr

2.2.7.1- Ligation Using a 0.1 mM Concentration of Oligonucleotide

1000 pmol of cross-linked oligonucleotide was dried down using a Thermo Savant SC110A SpeedVac concentrator and re-suspended in 0.25 M MES to afford a final concentration of 0.1 mM. The solution was placed in a sand bath at 95˚C for 10 minutes and then allowed to slowly cool to room temperature and then was placed on ice for 15 minutes. Following, 1 µL of 5M CNBr was added on ice for 5 minutes. The reaction was quenched via ethanol precipitation.

*Optimization of this involved using different amounts of CNBr (1, 2, 5, 10 and 20 µL) as well as varying the reaction time (5, 10, 15, 30, 60 and 120 minutes) while keeping all other variables constant.

2.2.7.2-Ligation Using a 0.01 mM Concentration of Oligonucleotide

300 pmol of cross-linked oligonucleotide was dried down using a Thermo Savant SC110A SpeedVac concentrator and re-suspended in 0.25 M MES to afford a final concentration of 0.01 mM. The solution was placed in a sand bath at 95˚C for 10 minutes and then allowed to slowly cool to room temperature and then was placed on ice for 15 minutes. Following, 3 µL of 5M CNBr was added on ice for 5 minutes. The reaction was quenched via ethanol precipitation.

*Optimization of this involved varying the reaction time (5, 10, 15 and 20 minutes) while keeping all other variables constant.

2.2.7.3-Ethanol Precipitation

On ice, $1/10^{th}$ V_f (final volume of the reaction) of 3 M sodium ascorbate and 2.5 x V_f of 99% ethanol (EtOH) are added to the reaction vessel and allowed to rest on ice for 15 minutes. The sample is then subjected to 14 000 x G for 30 minutes at 4˚C using a cold centrifuge. The supernatant is carefully discarded without disturbing the resulting pellet. Said pellet is washed with 2.5x V_f 70% EtOH and subjected to 14 000 x G for 15 minutes at 4˚C. The supernatant is discarded, the pellet is dried in the Thermo Savant SC110A SpeedVac concentrator and re-suspended in the desired matrix.

2.2.8-Ligation via Click Chemistry

The oligonucleotide of interest was dried and re-suspended in 0.2 M NaCl solution to allow a final concentration of 6 μM. The solution was placed in a sand bath at 95˚C for 10 minutes and then allowed to slowly cool to room temperature. To this solution, sodium ascorbate, pentamethyldiethylenetriamine (PMDTA) and copper II acetate were added to allow a final concentration of 0.6 mM, 6 mM and 5 μM. The solution was degassed in a Thermo Savant SC110A SpeedVac concentrator for 2 minutes. The ensuing solution was let to react for 2h at room temperature. After 1h, the reaction was supplemented with sodium ascorbate and copper II acetate. After the full 2h, the solution was transferred to a 15 mL Falcon™ tube and diluted with 0.1 M sodium acetate to 8 mL and desalted as previously described.

2.2.9-Removal of the Sulfone Cross-link

0.2 ODs of the sulfone cross-linked oligonucleotide was dried and re-suspended in K_2CO_3/MeOH (0.05, 0.1, 0.5 and 1 M). The ensuing solutions were placed in the sand bath at 70˚C

overnight. They were neutralized with an appropriate amount of glacial acetic acid and desalted as previously described. 20 pmol of the resulting solutions were dried down in preparation for an analytical PAGE. The resulting gel was stained with SYBR gold. The gel was submerged in a solution of SYBR Gold (prepared according to the protocol supplied by ThermoFisher Scientific) for 20 minutes and rinsed twice with distilled water.[41] It was visualized using Amersham™ Typhoon™ 5 Biomolecular Imager (Cytiva).

2.2.10-Nuclease Digest Using Exonuclease 1

100 pmol of each oligonucleotide was dissolved in 52.5 μL of buffer (10mM Tris, 2mM $MgCl_2$, pH 8.0). 0.2 U of SVPDE enzyme was added. At various times, 7.5 μL of the reaction mixture was removed and added to 7.5 μL of stop buffer (20mM EDTA / 95% formamide) and heat killed at 80˚C for 5mins. Following electrophoresis, the gels were stained using SYBR Gold nucleic acid stain and visualized using an Amersham™ Typhoon™ 5 Biomolecular Imager (Cytiva).

2.2.11-Biophysical Studies

2.2.11.1-UV Thermal Denaturation Experiments

Molar extinction coefficients were calculated using the nearest-neighbour approximation using the molar extinction coefficient of the mono- and dinucleotides.[42] Duplexes were prepared by dissolving equimolar amounts of single stranded oligonucleotides in 90 mM NaCl, 10 mM sodium phosphate and 1 mM EDTA (pH 7) to give a final concentration of 3.5 μM. The samples were then heated to 95 °C for 10 minutes, cooled slowly to room temperature then placed in the fridge overnight. Prior to the experiment, samples were degassed for 2 minutes in a SpeedVac concentrator. A Varian CARY Model 3E spectrophotometer fitted with a 12-sample thermostated cell block was used to monitor the absorbance at 260 nm and 280 nm. Starting at 4 °C, the

temperature was increased at a rate of 0.5 °C/min until a temperature of 95 °C was reached. The obtained data was transferred and processed using Microsoft Excel™ as described by Puglisi and Tinoco.[43]

2.2.11.2-Circular Dichroism Experiments

Circular dichroism experiments were run on a Jasco J-815 spectropolarimeter equipped with a Julaba F25 circulating bath. Samples were prepared as previously described in the thermal melt section. Each spectrum was an average of 3 scans collected at a rate of 20 nm/min with a bandwidth of 1 nm. Samples were placed in a fused quartz cuvette (Starna 29-Q-10). Spectra were recorded from 320-210 nm at 18˚C. The molar ellipticity (φ) was calculated using a software supplied by the manufacturer (JASCO Inc) following the equation $\varphi = \varepsilon/(C*L)$ where ε is the relative ellipticity (mdeg), C is the concentration (mol/L) and L is the path length (cm). The data was then transferred to Microsoft Excel™ for viewing.

2.2.12-Mass Spectrometry

The purity and identity of all oligonucleotides were assessed by mass spectrometry (LC-MS) at the Concordia University Centre for Biological Application of Mass Spectrometry (CBAMS). LC-MS analyses were performed on an Agilent 1100 LC system coupled to a Thermo LTQ Orbitrap Velos mass spectrometer equipped with a heated electrospray ion source in negative mode. A Spurcil C18-EP column (50 mm x 2.1 mm with a 3 μm particle diameter, Dikma Technologies) was used and oligonucleotides were eluted using a 20-min gradient at an initial flow rate of 250 μL/min with mobile phase A (10 mM ammonium acetate and 1 mM ammonium fluoride water solution) and B (ACN). The gradient started at 2% B and was held for 3 minutes followed by a linear gradient to 50% B at 8 minutes and 90% B at 10 minutes, then isocratic for 2 minutes with 90% B. The column was reconditioned at 13 minutes with 2% B at a

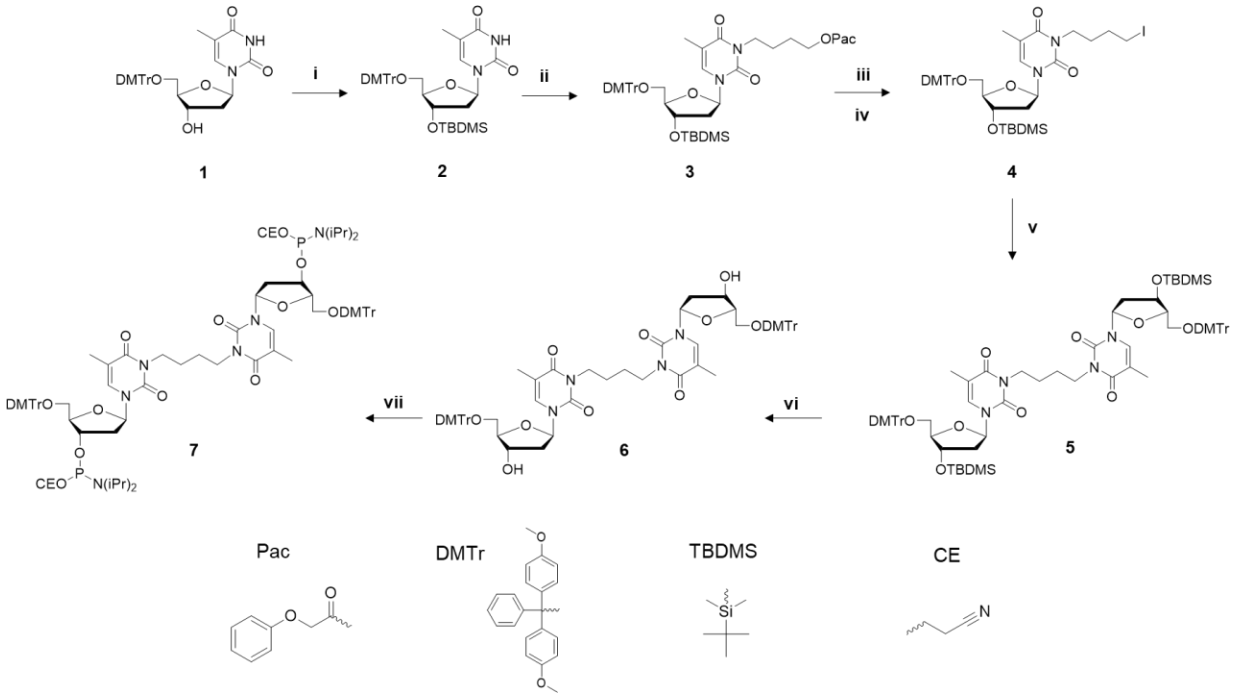
flow rate of 400 μL/min for 5 minutes and at 250 μL/min for an extra 2 minutes. Dried samples were reconstituted in 50 μL of mobile phase A and the injection volume was 10 μL. MS spectra (*m/z* 300-2000) were acquired in the Orbitrap at a resolution of 60 000. The uncharged monoisotopic mass of all oligonucleotides was calculated using Thermo Freestyle™ software (v1.7 SP1).

2.3-Results and Discussion

2.3.1-Synthesis of Small Molecules

2.3.1.1- N3-Butylene Linked Thymidine Bis-phosphoramidite

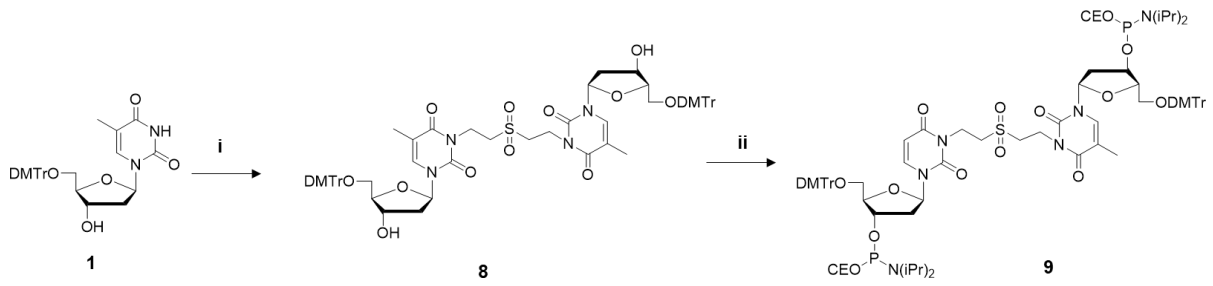
The synthesis of the N3-butylene linked thymidine bis-phosphoramidite for exploring the preparation and optimization of methods to circularize DNA (**Scheme 2.1**) began with the 2'O-silylation of compound **1** followed by the addition of a phenoxyacetyl (Pac)-protected butyl linker resulting in compound **3**. The Pac-protected hydroxyl group was then deprotected and subsequently iodinated giving rise to compound **4**. From here, dimerization was possible via the addition of 5'-O-(4,4'-dimethoxytrityl)-3'-O-*tert*-butyldimethylsilyl-2'-deoxythymidine (compound **2**) yielding compound **5**. Lastly, the *tert*-butyldimethylsilyl (TBDMS) protecting groups were removed and the 3'-hydroxyl groups were phosphitylated giving rise to the desired bis-amidite **7** (see Appendix I, **A1.1-7** for all NMRs pertaining to the synthesis of the N3-butylene linked thymine bis-phosphoramidite).



Scheme 2.1-Synthesis of N3-butylene linked thymidine bis-phosphoramidite. Compounds **2-4** and were prepared and characterized as previously described.[40] ***(i) tert***-butyldimethylsilyl chloride (1.8 eq), imidazole (3.6 eq), DCM, 16 h, rt, (*97* %). (**ii**) 4-iodobutyl-1-phenoxyacetate (1.1 eq), 1,8-Diazabicyclo[5.4.0]undec-7-ene (DBU, 2.0 eq), ACN, 24 h, rt, (*87* %). (**iii**) propylamine (16 eq), DCM, 48 h, rt. (**iv**) imidazole (5.1 eq), triphenylphosphine (2.5 eq), iodine (2.5 eq), THF, 4 h, rt, (*76 %*). (**v**) compound **2** (1.05 eq), DBU (2.0 eq), ACN, 24 h, rt, (*80 %*). (**vi**) tetrabutylammonium fluoride (1.0 M, THF, 2 eq), THF, 1 h, rt, (*71 %*). (**vii**) diisopropylamino cyanoethyl phosphonamidic chloride (2.4 eq), diisopropylethylamine (DIPEA, 2.8 eq), THF, 1h, rt, (*64 %*).

2.3.1.2- N3-1-Sulfonylethane Linked Thymidine bis-phosphoramidite

Following, the phosphoramidite derivative of bis 2-(N3-thymidine)-ethyl sulfone was prepared (**Scheme 2.2**) according to a procedure from Roger Tresánchez, PhD Thesis, University of Barcelona, 2013. Starting once again from compound **1**, dimerization was achieved via the addition of divinyl sulfone and tributylphosphine giving rise to compound **8**. This dimer was then phosphitylated resulting in compound **9** (see Appendix I, **A1.8-12** for all NMRs pertaining to the synthesis of the N3-1-sulfonylethane linked thymine bis-phosphoramidite).



Scheme 2.2-Synthesis of N3-1-sulfonylethane linked thymidine bis-phosphoramidite. (**i**) divinyl sulfone (0.5 eq), tributylphosphine (0.1 eq), ACN, 16 h, 50 °C, (*92* %). (**ii**) diisopropylamino cyanoethyl phosphonamidic chloride (2.4 eq), diisopropylethylamine (DIPEA, 2.8 eq), THF, 1h, rt, (*46 %*).

2.3.2-Oligonucleotide Synthesis

The oligonucleotides synthesized for this project were synthesized using standard β-cyanoethylphosphoramidite chemistry on a 1.5 μmol scale (**Table 2.1**). CJW 978, 1035 and 1238, which were used for click ligation, were prepared using 500 Å LCAA-CPG derivatized with 5'-O-(4,4'-dimethoxytrityl)-3'-O-propargyl-2'-deoxycytidine at the N^4-atom. CJW 997 and 1036 were prepared using standard 500 Å LCAA-CPG. Both CJW 1035 and 1036 were prepared with the N3-butylene linked thymidine bis-phosphoramidite whereas CJW 1238 was prepared with the phosphoramidite of bis 2-(N3-thymidine)-ethyl sulfone which required some modifications to the traditional SPS cycle. Furthermore, CJW 1238 was synthesized using fast deprotecting amidites since we were uncertain whether the sulfone cross-link would be labile under standard deprotection conditions. Under these conditions, phenoxyacetyl-dA (Pac-dA) and isopropyl-phenoxyacetyl-dG (iPr-Pac-dG) phosphoramidites from Glen Research were used rather than the standard protected nucleoside phosphoramidite with concentration of 0.12 M. Phenoxyacetic anhydride was used as the capping agent instead of acetic anhydride. For the cross-links, a concentration of 0.05 M in ACN and coupling time of 1800 seconds were required. Furthermore, all oligonucleotides were synthesized with terminal modifications. CJW 978, 1035 and 1238 have

a terminal 5'-I-dT which required a concentration of 0.15 M and a coupling time of 300 seconds. On the other hand, CJW 997 and 1036 were synthesized with a terminal phosphate group which was incorporated using a chemical phosphorylating reagent. This modification was also dissolved in ACN to afford a final concentration of 0.15 M and necessitated a coupling time of 450 seconds. The identity of all synthesized species were confirmed by MS (see Appendix I, **A1.13-15** and **A1.17** and **A1.21**).

Table 2.1-Oligonucleotides synthesized for Chapter 2 with their corresponding modifications.

Identifier	Sequence	Modifications
CJW 978	5′- TCA TTT TTA AAA TGA GAT TTT TC-3′	T: 5′-I-dT C: 3′-propargyl-dC
CJW 997	5′- YTCA TTT TTA AAA TGA GAT TTT TC-3′	Y: 5′-phosphate
CJW 1035 (b-series)	5′- TCA TTT TXA AAA TGA GAT TTT TC-3′	T: 5′-I-dT X: TN3-Butylene-N3T Cross-link C: 3′-propargyl-dC
CJW 1036 (a-series)	5′- YTCA TTT TXA AAA TGA GAT TTT TC-3′	Y: 5′-phosphate X: TN3-Butylene-N3T Cross-link
CJW 1238 (c-series)	5′- TCA TTT TXA AAA TGA GAT TTT TC -3′	T: 5′-I-dT X: TN3-sulfone-N3T Cross-link C: 3′-propargyl-dC

2.3.3-Post-synthetic Conversion of 5'-Iodo to 5'-Azide

As mentioned, oligonucleotides used for click chemistry where prepared with a terminal 5'-I moiety. Post-synthetically, this terminal moiety was substituted for a 5'-N_3 group. This additional step is necessary since, in the presence of a phosphite source, the azide will react resulting in the formation of an iminophosphorane.[44] This is known as the Staudinger reaction and is highly undesirable. Hence, following oligonucleotide synthesis, sodium azide was dissolved in

DMF and heated to 70°C for 20 minutes to aid in the dissolution. Once the sodium azide has dissolved, it was transferred to the vial containing the bound oligonucleotide. The oligonucleotide, still attached to the solid support, was treated with sodium azide immediately after synthesis in order to reduce the likelihood of de-iodination. Following this conversion, the oligonucleotide was deprotected under standard conditions described below.

2.3.4-Deprotection and Purification of Oligonucleotides

Following the synthesis of oligonucleotides and any necessary post-synthetic functionalization, CJW 1035 and 1036 still bound to CPG were transferred to a 2 mL screw cap vials and 1 mL of ammonia/ethanol (3:1, v/v) was added to each and heated to 55°C for 16h. These are standard conditions that follow solid-phase oligonucleotide synthesis that enable cleavage of the oligonucleotide from the CPG and removal of protecting groups on the nucleotides at the heterocyclic bases, sugars, and non-bridging O-atoms at the phosphodiester backbone. The supernatant, which contains the free and deprotected oligonucleotide, is then transferred to a clean 2 mL vial. The CPG is subsequently washed twice with 50% ACN/water and the supernatant is transferred each time to the new vial. The resulting solution is then dried down and resuspended in the desired matrix, usually water.

For CJW 1238, due to the use of fast-deprotecting amidites, 0.05 M K_2CO_3/MeOH was used for deprotection. This deprotection strategy is considered milder and allows for efficient deprotection of oligonucleotides containing more sensitive or base-labile modifications. Similarly, to CJW 1035 and 1036, the CPG is transferred to a 2 mL screw cap vial and 1 mL of 0.05 M K_2CO_3/MeOH is added and left at room temperature overnight. The base is neutralized with an equimolar amount of glacial acetic acid. Following, the supernatant is transferred to a clean vial and the CPG is washed. The merged supernatant is dried and re-suspended in the desired matrix.

All oligonucleotides used in this chapter were purified via PAGE. This technique is useful as it can be used for studying interactions or purification of oligonucleotides. Prior to purification, an analytical gel was often performed to analyze the crude oligonucleotide that was synthesized. These analytical gels are run for shorter periods of time (1.5h) and provide a general idea of the purity of the oligonucleotides. For purification purposes, a preparatory gel is performed which allows for better resolution of closely migrating species but typically require longer run times (3.5-4.5h) to achieve separation. After resolution of the oligonucleotide bands on the gel, the bands of interests are excised from the gel and eluted using 0.1 M sodium acetate to extract the oligonucleotides into solution in preparation for the desalting step.

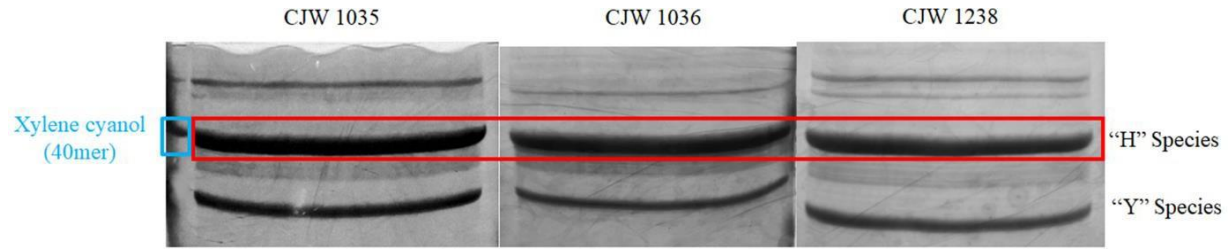


Figure 2.6-Purification of cross-linked species using 20% PAGE at 450V for 4.5 hours (xylene cyanol runs as 40mer).

As seen in **figure 2.6**, all cross-link species have two distinct bands with the one of interest being the top band (highlighted in the red box), referred to the "H" species. The bottom band corresponds to the oligonucleotide species that only coupled at one of the phosphoramidite groups of the dimer, which is also referred to as the "Y" species.

2.3.5-Circularization

2.3.5.1-CNBr-Assisted Ligation

Following purification, ligation attempts using cyanogen bromide (CNBr) ensued (**Figure 2.7**). As a starting point, 1000 pmol of CJW 1036 were dried and re-suspended in 0.25 M MES. The resulting solution was placed in a sand bath at 95˚C for 10 minutes to denature the oligonucleotide and allowed to cool to room temperature to assure proper annealing of the complementary strands. Following, since the CNBr reaction is exothermic, all reagents were placed in an ice bath for 15 minutes. CNBr was added to the reaction and allowed to react for 5 minutes before quenching the reaction via an ethanol precipitation. The ligation products were then visualized on a gel under UV light (**Figure 2.8**).

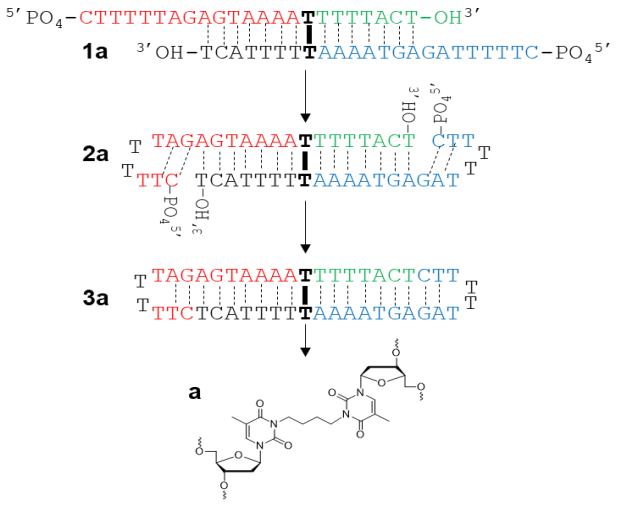


Figure 2.7-Visual representation of ligation using CNBr

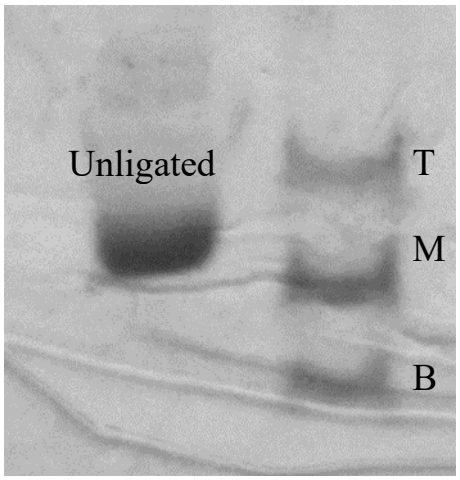


Figure 2.8-First circularization attempt using CNBr (20% PAGE, 250V, 1.5h). Left lane: unligated product, Right lane: ligation products.

Under these conditions, three products were formed. The left lane represents the linear unligated species. On the other hand, in the right lane, the three products were identified via mass spectrometry (MS) as the unligated product (top), the partially ligated product (middle) and the fully ligated product (bottom) (see Appendix I, **A1.18-20**).

In an effort to optimize the reaction, different conditions were investigated. First using the same concentration of oligonucleotide (0.1 mM), longer reaction

times were investigated (**Figure 2.9-A**). This was not fruitful as there was no improvement of the yield for the fully ligated product. The concentration of CNBr used was also investigated but did not produce any significant changes (**Figure 2.9-B**).

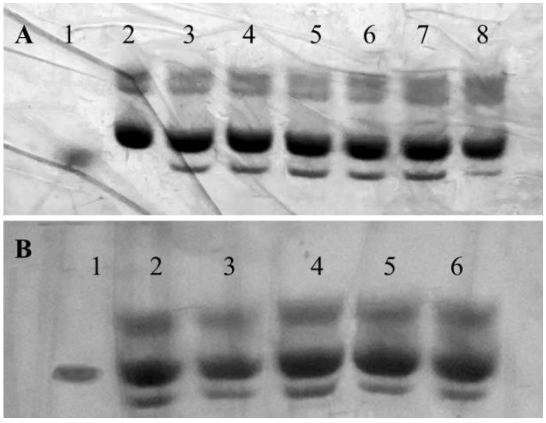


Figure 2.9-Optimization of the CNBr circularization using an oligonucleotide concentration of 0.1 mM (**A**) Investigation into the reaction time (1: Xylene cyanol dye (runs as a 40mer oligonucleotide), 2: Linear pre-cursor, 3-8: CNBr circularization with reaction times of 5, 10, 15, 30, 60 and 120 minutes), (**B**) Investigation into higher [CNBr] (1: Xylene cyanol dye (runs as a 40mer oligonucleotide), 2-6: 1, 2, 5, 10 and 20 µL of 5 M CNBr). 20% PAGE, 250V, 1.5h.

Next a lower concentration of oligonucleotide (0.01 mM) was used which produced a bottom band with more intensity than what had previously been seen. The reaction time was also investigated however no significant changes were observed when increasing the reaction time (**Figure 2.10-A**). From this, it was determined that the optimal ligation conditions for this system was 0.01 mM DNA in 0.25 M MES buffer with 3 µL of 5 M CNBr and a reaction time of 5 minutes. Following this, the reaction was scaled up in hopes of generating an improved yield of fully ligated product to conduct biophysical experiments. However, this was not the case. Instead of the bottom band being the major product, we saw that the top and middle bands dominated (**Figure 2.10-B**). All three bands were purified. The top and middle band were re-treated with CNBr in an attempt to drive the reaction forward, but this was unsuccessful. Nonetheless, enough of the desired material (bottom band) was obtained to push forward.

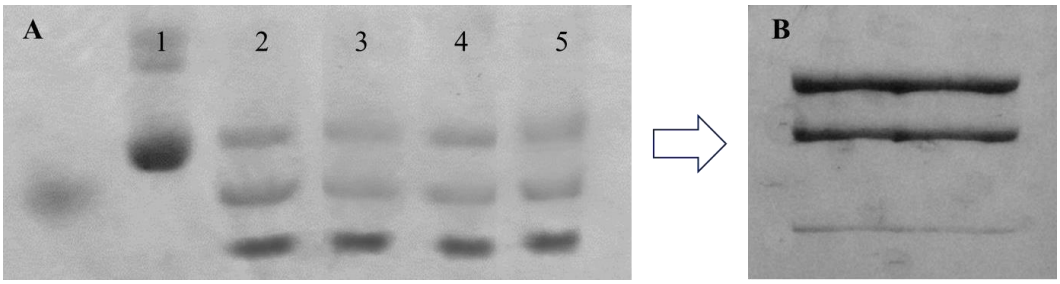


Figure 2.10- Optimization of CNBr circularization using a oligonucleotide concentration of 0.01 mM, (**A**) Investigation into the reaction time (1: Linear pre-cursor, 2-5: CNBr ligation with reaction times 5, 10, 15 and 20 minutes. 20% PAGE, 250V, 1.5h. (**B**) Purification of larger scale CNBr reaction using optimized conditions. 20% PAGE, 450V, 4.5h.

2.3.5.1-Circularization via Click Chemistry

Compared to CNBr ligation, ligation of the oligonucleotides by click chemistry (**Figure 2.11**) was more straightforward. A solution of 6 µM of cross-linked DNA in 0.2 M sodium chloride was prepared in a 5 mL Falcon™ tube. The solution was heated in the sand bath for 10 minutes at 95°C in order to remove any secondary structures the DNA may have in solution. To this, freshly prepared sodium ascorbate, PMDTA and copper acetate were added to allow a final concentration of 0.6 mM, 6 mM and 5 µM. The order of addition is important as copper (II) can cause DNA degradation hence it is important that the reducing agent and ligand be added first in order to reduce the Cu (II) to Cu (I) and to chelate any non-reduced Cu(II).[45] The solution was then degassed and allowed to proceed for 2h. From previous lab members' work, it is known that the solution contains too much salt to allow for proper resolution and visualization of the

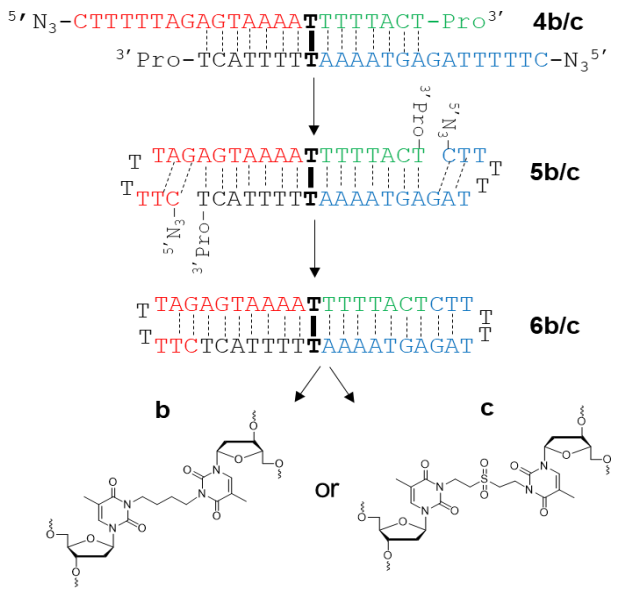


Figure 2.11-Visual representation of ligation via Click Chemistry using oligonucleotides containing a 5'-azide (N3) moiety and 3'-propargyl (Pro) moiety.

oligonucleotides by gel electrophoresis. Therefore, upon completion of the reaction, a desalting step was necessary in order to remove excess reagents.

To evaluate the effectiveness of the circularization of the linear oligonucleotides by the click reaction, an analytical gel was performed to resolve the circularized versus linear species. The circularized products run faster on a gel (has a higher mobility) compared to its linear counterpart because it can fold into a more compact structure which makes it easier for this species to pass through the pores of the gel matrix (**Figure 2.12-A)**. Compared to the CNBr reaction, the click reaction was near quantitative resulting in one major species and a minor species corresponding to the starting material. Furthermore, this was easily scaled up to allow for more fully ligated product to be generated. For this reason, we decided to include the sulfone linker into a click oligonucleotide only rather than in a CNBr oligonucleotide (**Figure 2.12-B**).

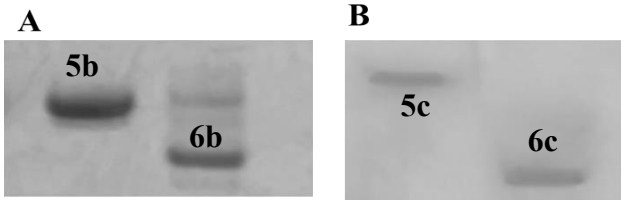


Figure 2.12-(**A**) Result of click reaction with butyl cross-link. Right: linear precursor, Left circularized product. 20% PAGE, 200V, 1.5h (**B**) Result of click reaction with sulfone cross-link. Right: linear precursor, Left: circularized product. 20% PAGE, 200V, 1.5h.

The identity of the circularized products was confirmed by MS (see Appendix I, **A1.16** and **A1.22).** From previous work in our lab, we know that circularization of oligonucleotides using click chemistry results in the same mass as the linear oligonucleotide. Therefore, the formation of the circle cannot be directly inferred from the MS results but rather requires a combination of techniques including gels, thermal melt experiments and nuclease digests.

2.3.6-Removal of the Sulfone Cross-link

Removal of the sulfone is thought to proceed following a β-elimination to restore natural DNA. The base abstracts the β-hydrogen which results in the release of divinyl sulfone and restoration of a natural thymine base (**Figure 2.13**).

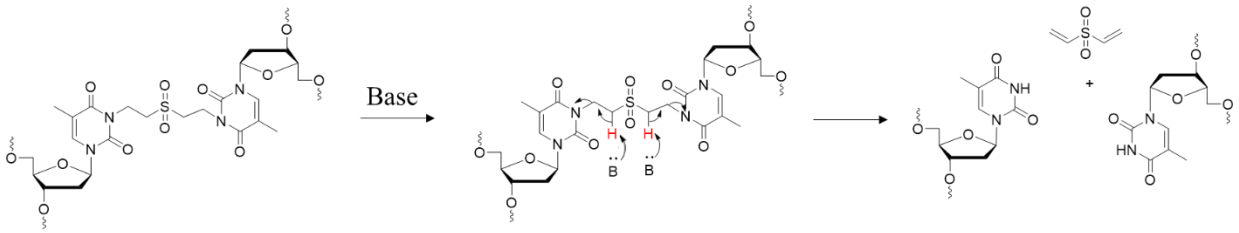


Figure 2.13-Removal of the sulfone cross-link via a β-elimination.

Previous work from Lauralicia Sacre (MSc 2017) showed that the sulfone cross-link was stable in 0.05 M K_2CO_3/MeOH for up to 72h at 65˚C. Knowing this, we were interested in determining if the sulfone cross-link could be fully removed at higher concentrations of potassium carbonate. To do this, 0.2 ODs of the linear sulfone cross-linked oligonucleotide was dried and re-suspended in K_2CO_3/MeOH (0.05, 0.1, 0.5 and 1 M). The resulting solutions were placed in the sand bath overnight at 70 ˚C. The solutions were then neutralized with acetic acid and desalted. 20 pmols of each sample was loaded on a gel and stained with SYBR gold. As shown in **Figure 2.14-A**, treatment with 1 M K_2CO_3/MeOH results in significant damage to the DNA resulting with the formation of numerous degradation products. On the other hand, with 0.05 M and 0.1 M K_2CO_3/MeOH little change in the oligonucleotide is observed. However, when using 0.5 M K_2CO_3/MeOH there is the appreciable formation of a new band. This new band has a higher

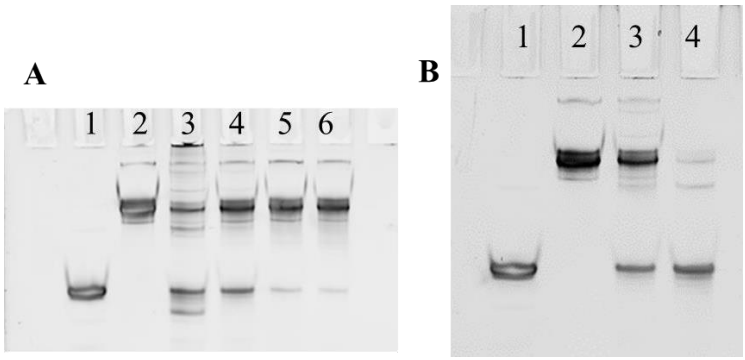


Figure 2.14-(**A**) Removal of the sulfone cross-link with K_2CO_3/MeOH (1: ssDNA linear 2: cross-linked DNA, 3-6: 1, 0.5, 0.1 and 0.05 M K_2CO_3/MeOH, 48h, 70˚C. 20% PAGE, 250V, 1.5h (**B**) Cleavage of the sulfone cross-link (1: ssDNA, 2: cross-linked DNA, 3: 0.5 M K_2CO_3/MeOH, 48h, 70˚C, 4: 0.5 M K_2CO_3/50% MeOH/50% H_2O, 48h, 70˚C. 20% PAGE, 250V, 1.5h.

mobility and runs similarly to the control ssDNA (which is believed to be the product of the linear oligonucleotide with the sulfone group removed).

With this information in hand, we decided to test a longer reaction time. Also, K_2CO_3 has limited solubility in MeOH. To improve the solubility of the K_2CO_3 for more optimal conditions for removal of the sulfone cross-link in the oligonucleotide, a solution of 0.5 M K_2CO_3 in 50% H_2O and 50% MeOH was prepared. Following the same experimental set-up as described above, the MeOH and 50% H_2O/50% MeOH potassium carbonate solutions were investigated for the removal of the sulfone cross-link. The time was also increased to 48h and the temperature to 70˚C to allow the reaction to go to completion (**Figure 2.14-B**).

Hence, it was determined that the optimal conditions for the removal of the sulfone are: 0.5 M K_2CO_3 in 50% H_2O and 50% MeOH, 48h at 70 ˚C. These conditions were then used to remove the sulfone cross-link from the circularized product. The circularized product with and without the cross-link exhibit the same mobility on a gel hence, the removal of the sulfone was confirmed by MS (see Appendix I, **A1.23**).

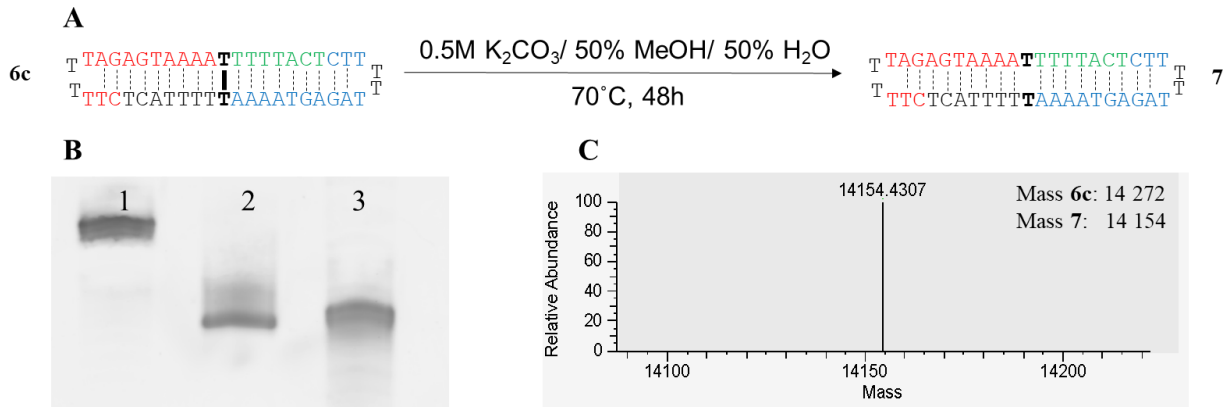


Figure 2.15-(A) Visual representation of the sulfone cleavage reaction, (**B**) Gel highlighting mobility of 6c vs 7 (lane 1: Linear precursor, Lane 2: 6c, Lane 3: 7). 20% PAGE, 250V, 1.5h. (**C**) MS data confirming sulfone cleavage.

2.3.7-Exonuclease 1 Digest

Snake venom phosphodiesterase (SVPDE) also known as exonuclease 1 is a 5'-nucleotidohyrolase.[46] In order words, it hydrolyses oligonucleotides from the 5'-termini. This type of assay is particularly useful for identifying circular DNA as they do not possess free 5'-terminus which greatly increases their resistance to exonucleases in general. To evaluate that circularization has occurred, 100 pmol of oligonucleotide is dried and then re-suspended in SVPDE buffer. Prior to the addition of enzyme, an aliquot is removed and diluted with stop buffer which serves as t=0 minutes. Following this, 0.2 U of SVDPE is added to the reaction and aliquots are removed, diluted with stop buffer and heat inactivated at times 5, 10, 15, 30, 60 and 120 minutes. The aliquots are then directly loaded onto a gel.

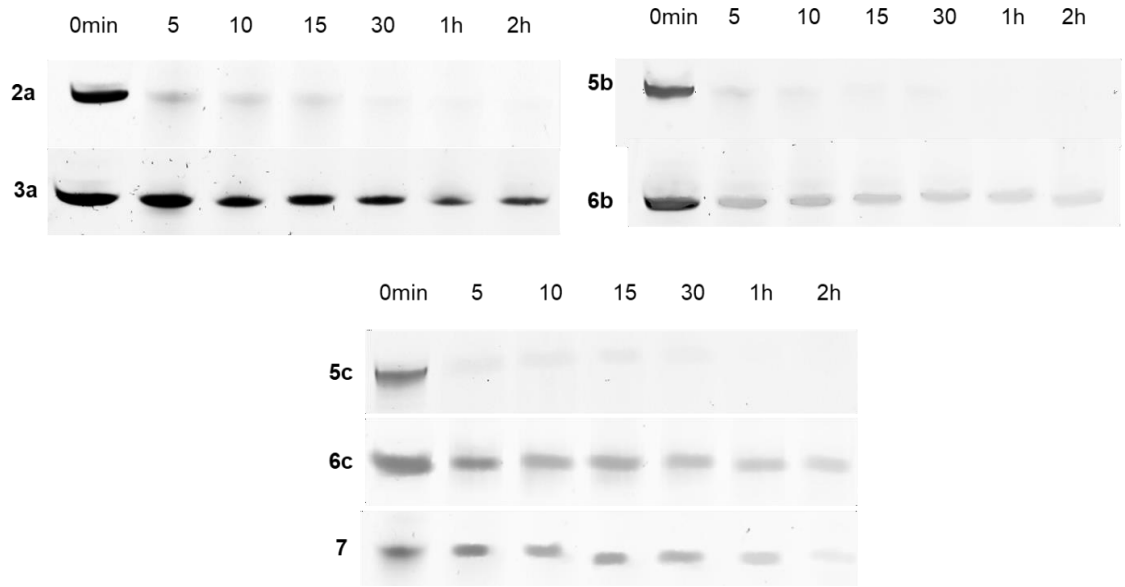


Figure 2.16-Nuclease digest of 2a vs 3a (top left), 5b vs 6b (top right) and 5c vs 6c vs 7 (bottom middle), 20% PAGE, 250V, 1.5h.

As seen in **Figure 2.16**, all linear cross-linked oligonucleotides (2a, 5b and 5c) are fully degraded within 30 minutes. However, the circularized products (3a, 6b and 6c) display an increased resistance to exonuclease 1 as they are not fully degraded even after 120 minutes.

Furthermore, oligonucleotide 3a, which has a natural phosphodiester backbone, displays an enhanced resistance compared to oligonucleotides 6b and 6c. This is consistent with what is seen in the literature where the triazole linkage is resistant to exonucleases but to a lesser extent than the natural phosphodiester backbone. Typically, the triazole backbone is used more to enhance membrane permeability and cellular uptake to its neutral nature.[47] Lastly, the presence or absence of the sulfone cross-link (6c vs 7) does not seem to have a significant impact on the nuclease stability of the oligonucleotide.

2.3.8-Biophysical Studies

2.3.8.1-UV Thermal Denaturation

The influence of the circular topology was investigated via ultraviolet (UV) thermal denaturation. A general trend consistent with what has been seen in the literature is observed where there is an increase in T_m when going from ssDNA to cross-linked DNA to circularized DNA. However, we do see a major drop in T_m when the sulfone cross-link is removed which can be attributed to a combination of factors. Upon removal of the sulfone, a T-T mismatch is introduced

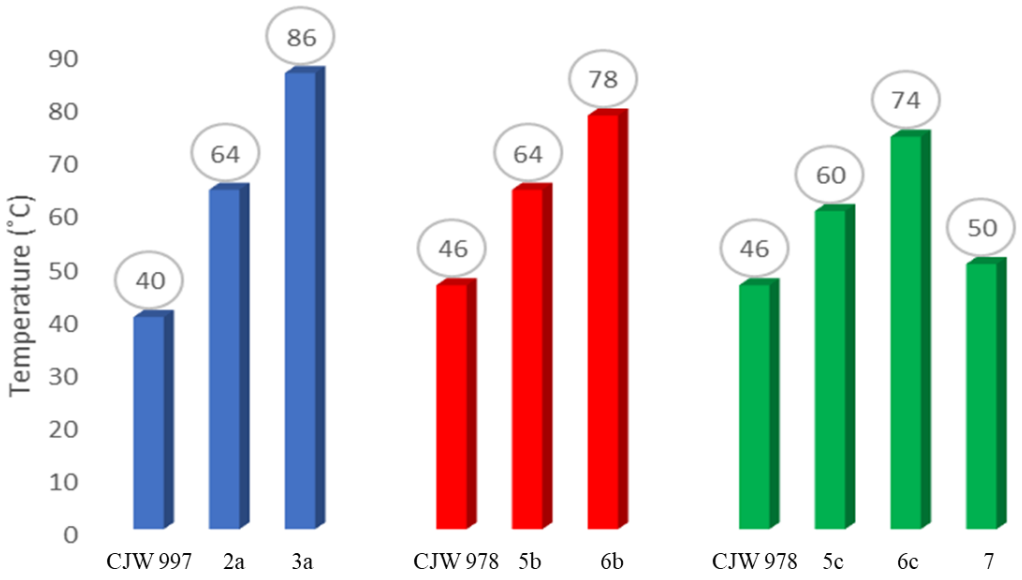


Figure 2.17-T_m values of ssDNA, cross-linked DNA and circularized DNA.

which has been known to lower duplex stability by 9-13˚C for oligonucleotides of similar length.[48] Furthermore, the rigid cross-link "locked" the oligonucleotide into a certain conformation however, by removing it, the oligonucleotide is free to adopt other structures some of which may not be favorable resulting in a lower T_m value.

2.3.8.2-Circular Dichroism

CD experiments were performed to determine the global structure of the circularized oligonucleotides. The CD profile of all species investigated show characteristic B-form maxima at 275-280nm and minima at 245-250 nm. Hence, the circular topology does not affect the global structure of the oligonucleotide.

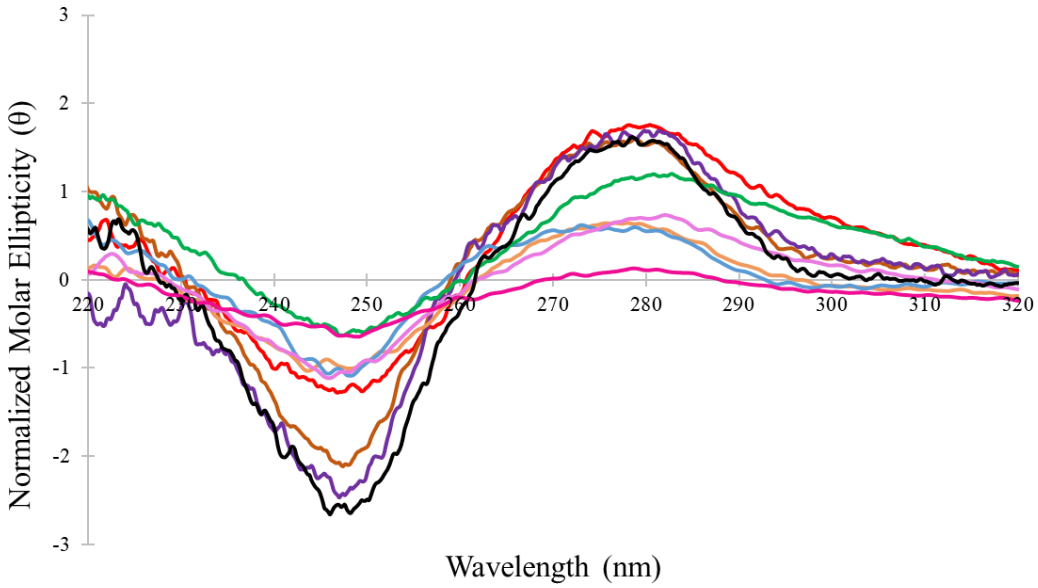


Figure 2.18-CD traces of ssDNA, cross-linked DNA and circularized DNA (CJW 997, 2a, 3a, CJW 978, 5b, 6b, 5c, 6c and 7).

2.4. Conclusion

Circular oligonucleotides containing both a cleavable and uncleavable cross-link were synthesized. The pre-organization of the system via the use of cross-links allowed for efficient ligation using click chemistry and CNBr-assisted ligation. Furthermore, a cleavable sulfone cross-link between the N3 atoms of thymine nucleobases was efficiently removed using 0.5 M K_2CO_3 in 50% water/MeOH for 48h at 70˚C. All circularized constructs showed an increased resistance to exonuclease 1 when compared to their linear counterparts. Circular dichroism spectroscopy revealed that the presence of these cross-links does not affect the global structure of the oligonucleotides with this circular topology.

Chapter 3 - Exploring Methodologies to Prepare Phosphoroselenoate-Modified DNA for Structure Determination

Abstract

Nucleic acid crystallography has proven itself to be an essential asset to determine the functional role of this biological molecule in various complexes and has provided mechanistic insights into nucleic acid related diseases.[49] However, two major hurdles exist in nucleic acid crystallography: the phase problem and crystallization. To address the former challenge, modified nucleobases containing halogens such as bromine have been employed however, this strategy presents numerous disadvantages such as tedious synthesis, radiation instability and structural perturbations.[50] Derivatization of nucleic acids with selenium has shown promise over the previously used halogen strategy. The introduction of a phosphoroselenoate internucleotide linkage, where one of the nonbridging oxygen atoms is replaced by selenium, offers a simple way to incorporate an anomalous scatter into oligonucleotides.

Herein, we explored a novel and facile way to introduce selenium at the terminal 5' position of an oligonucleotide by employing chemical phosphorylation of a support bound oligonucleotide followed by manual treatment with a saturated solution of KSeCN in acetonitrile. Initially, the incorporation of selenium was found to be unsuccessful as determined by HPLC and MS characterization of the products. Through rounds of optimization, introduction and retention of selenium was observed when a methyl group was incorporated at a non-bridging oxygen atom of the 5'-terminal selenophosphate.

3.1. Introduction

3.1.1-Importance of Modifications to the Nucleic Acid Scaffold

In the last decades, there has been significant growth in applications of oligonucleotides as a therapeutic and a diagnostic tool. The promise of oligonucleotides as a therapeutic platform involves gene therapy which targets disease-causing genes in a sequence- specific manner, allowing for more precise and personalized treatment of a wide range of life- threatening disorders. It is possible to downregulate, amplify or correct gene expression in a desired tissue.[51] Antisense oligonucleotides (ASO) as well as small interfering RNA (siRNA) are two examples of approaches used to inhibit gene expression.

A handful of these modified oligonucleotides have been approved to treat an array of illnesses. Recently, casimersen (brand name Amondys 45), a 22mer phosphorodiamidate morpholino oligomer (PMO), has been approved for the treatment of Duchene muscular dystrophy (DMD). This drug can bind to exon 45 on the DMD pre-messenger RNA (mRNA) gene which forces the mRNA processing machinery to skip exon 45 during mRNA translation resulting in a truncated yet functional protein.[52] Givosiran (brand name Givlaari) has recently been approved for the treatment of acute hepatic porphyria. This drug functions by entering the cell and interferes with the RNA interference (RNAi) pathway hence it hinders the expression of a specific gene with a complementary nucleotide sequence by degrading the mRNA after transcription therefore preventing translation.[51]

Challenges for *in vivo* and *in vitro* applications involving oligonucleotides is that they are highly susceptible to nuclease degradation, and they undergo rapid filtration from the bloodstream via renal clearance.[53] A way to address these issues is by the introduction of modifications to the nucleic acid scaffold without compromising their recognition properties.[53]

Due to the versatility of solid phase synthesis and interest in developing therapeutic agents, many modifications have been made to the scaffold of nucleic acids either at the phosphate backbone, the pentose sugar, or the heterocyclic base. Modifications to the backbone, also known as 1st generation modifications, consist of replacing one of the non-bridging oxygen atoms of the backbone with another atom or group. Examples include replacement with a sulfur atom (phosphorothioate), a methyl group (methyl phosphonates) or an amine moiety (phosphoroamidates).[54] These 1st generation modifications (**Figure 3.1**) showed great promise for therapeutic purposes as they increase nuclease resistance and exhibit longer half-lives both *in vitro* and *in vivo* when compared with native oligonucleotides. In fact, to date, most oligonucleotide drugs in clinical trial and on the market have a phosphorothioate backbone.[54] However, 1st generation oligonucleotides were found to have undesirable side effects such as decreased duplex stability and immune stimulation due to interactions with proteins.[53,54]

2nd generation modifications are focused on modifying the 2'-O position of the sugar with an alkyl group (**Figure 3.1)**. The most widely used 2nd generation modifications are the 2'-O-methyl and the 2'-O-methoxyethyl (2'-MOE) modifications. To circumvent this issue, chimeric oligonucleotides known as gapmers were created. The idea was to have a central window or "gap" composed of modification (s) that would induce RNase H mediated cleavage of target RNA, flanked by "arms" consisting of modifications that increase nuclease resistance.[54–56] These 2nd generation gapmers were shown to have a higher affinity for mRNA, increased resistance to nucleases, longer half-lives, better tissue uptake and lower toxicity compared to 1st generation oligonucleotides.[54]

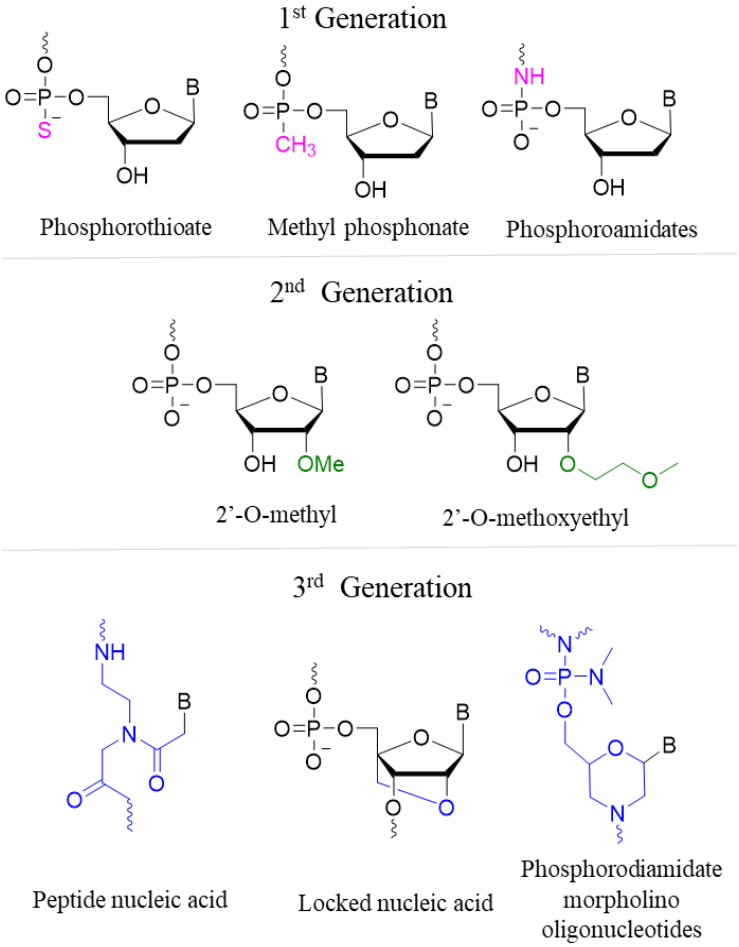


Figure 3.1-1st, 2nd, and 3rd generation modifications to the nucleic acid scaffold.

More recently, 3rd generation oligonucleotides have been developed (**Figure 3.1**). These oligonucleotides are highly heterogenous involving many different types of chemical modifications ranging from furanose ring modifications to phosphate linkage, and whole nucleotide modifications. Examples of such modifications include the locked nucleic acid (LNA), the peptide nucleic acid (PNA) and phosphorodiamidate morpholino oligonucleotides (PMO).[54,57] These 3rd generation oligonucleotides are highly resistant to nuclease degradation giving them greater stability in biological fluids, exhibit strong hybridization with their target and can even modulate gene expression to some extent, however, they do not activate RNAse H.[54] Therefore, 3rd generation gapmers were developed similarly to 2nd generation gapmers.

3.1.2-Adaptation of Conventional SPS to Produce Chemically Modified Oligonucleotides

The introduction of chemical modifications to the oligonucleotide scaffold can be accomplished by two strategies to incorporate the above-mentioned modifications into oligonucleotides when using solid phase synthesis. The first strategy is to synthesize a phosphoramidite with the desired modification (**Figure 3.2-A**). This method is typically used for base and sugar modifications. Taking for example the incorporation of a 2'-O-methyl moiety, once the phosphoramidite is synthesized it can be easily incorporated into an oligonucleotide by simply coupling it to the growing chain. The second strategy is commonly used for backbone modifications, which can involve the substitution of the oxidizing agent on the synthesizer with different reagents (**Figure 3.2-B**). Take for example the phosphorothioate backbone modification, which can be introduced by substituting the conventional oxidizing chemicals for a sulfurizing agent such as the Beaucage reagent or 3-[(dimethylaminomethylene)amino]-3H-1,2,4-dithiazole-5-tione (DDTT).

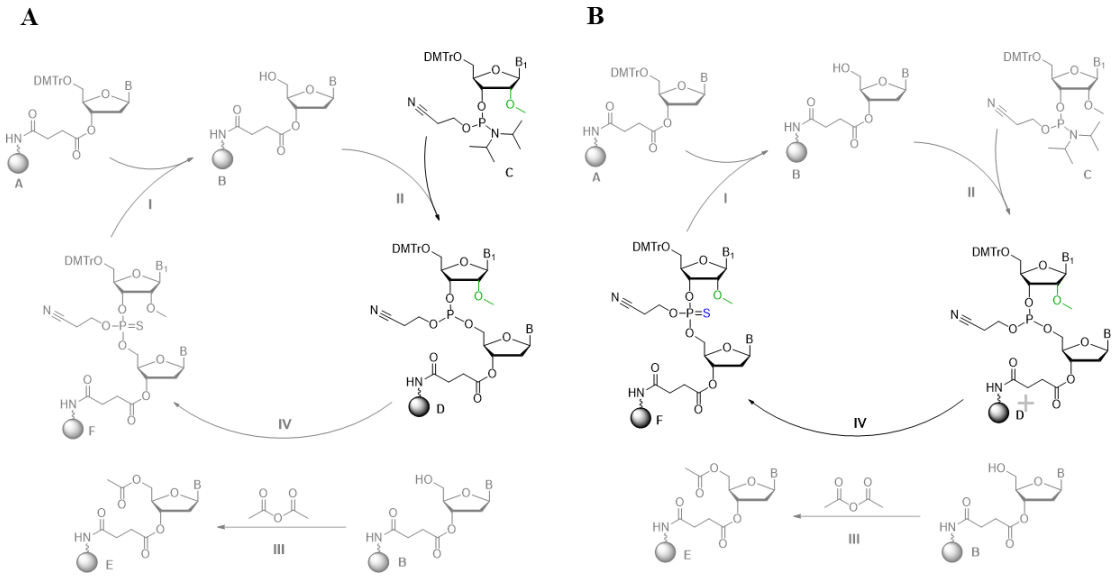


Figure 3.2-Methods to incorporate a modification into an oligonucleotide, (**A**) Through a modified phosphoramidite building block or (**B**) Changing the oxidizing solution.

3.1.3-X-ray Crystallography of Nucleic Acids

Extensive studies looking into how drugs bind and interact with their target are necessary to properly understand their mechanism of action. One important approach for these types of studies is X-ray crystallography. Nucleic acid crystallography has proven itself to be an essential asset to determine the functional role of nucleic acids in various complexes and has provided mechanistic insights into nucleic acid related diseases.[58,59] Structure determination at the atomic level is important to speed up nucleic acid drug discovery, as well as the boost our understanding of nucleic acid roles in biology and the enable future development of nucleic acid nanotechnology.[59]

Essentially, the three-dimensional structure of a molecule or a complex can be deduced from it's crystal by exposing said crystal to an X-ray beam. The result is a diffraction pattern which offers information about the size of the repeating unit as well as the packing symmetry.[60] The intensity of the spots can be extrapolated to generate an electron density map (also known as phasing) which can then be computationally fitted to produce an atomic model.[60] However, two major hurdles exist in nucleic acid crystallography. The first is known as the crystallization problem, specifically nucleic acids are difficult to crystallize and it is challenging to acquire a high-quality crystal.[59] This is due to the fact that nucleic acids are negatively charged, have a high solvent content and their structural dynamics hinder the formation of highly ordered crystals.[59] Alas, it is difficult to predict good nucleation and growth conditions therefore there is no rational approach other than trial and error screening which requires a large amount of pure material.[4] Hundreds, or even thousands, of crystallization solutions are prepared with varied temperature, concentration, buffer etc. in hopes of acquiring a high-quality crystal.[59] This is often the bottle neck of nucleic acid crystallography.

The other hurdle is known as the phase problem whereas there is a loss of information concerning the phase that occurs when making a physical measurement. In order to obtain an intelligible electron density map the amplitude and phase of a wave are necessary however, in a diffraction experiment, the phase cannot be directly measured.[61] Several methods have been employed to try and solve this problem. The direct method makes use of established phase correlations among specific groupings of reflections. This method requires a high-resolution data set and is most suitable for small molecules of less than 1000 non-hydrogen atoms.[58,61] The molecular replacement method is best suited when the structure being solved has a homologous structure as this homolog can be used to determine the orientation and position of the native molecules within the unit cell allowing the phase information to be extracted.[58,61]

There also exist experimental methods that can help solve the phase problem such as isomorphous replacement and the anomalous scattering method. The former entails preparing derivatives, using heavy atoms, of the native crystal with the same space group and unit cell without changing its crystalline order.[58] The idea is that the heavy atom derivative will introduce an intense difference in the diffraction pattern which is measurable via the direct method. Once its position is established, the corresponding phase can be calculated and used to deduce the phase of the native molecule.[58] It should be noted that two or more heavy atoms in separate crystals should be incorporated to avoid phase ambiguity. On the other hand, in anomalous scattering experiments (known as multiple wavelength anomalous diffraction, MAD), diffraction data are taken at several different wavelengths close to the absorption edge of an anomalous scatterer present in the crystal.[62] The anomalous scatterer exhibits different scattering factors at different wavelengths allowing for variations in the diffraction intensities. These differences can be used to calculate the phase.[62,63] This method offers several advantages when compared with isomorphous replacement:

only one crystal is needed thus reducing systemic error and improving the accuracy of the phase, the anomalous scatter remains strong at high resolution therefore the phase is more accurately determined at high resolution and since the electron density map is usually of higher quality, the initial model can be built faster.[63] Due to this, MAD is now the most widely used phasing techniques for X-ray crystallography of biomolecules.[62]

3.1.4-Selenium Modified Nucleic Acids

Typically, one of two methods - soaking the crystal in a solution containing a heavy atom compound or co-crystallization in the presence of a molecule that contains heavy atoms - is used to introduce the heavy atom into the crystal.[58] However, this has proved problematic for nucleic acids (when compared to proteins) as they lack specific binding sites for metal ions.[64] Conventional, halogen derivatization using bromine has been used (such as 5-bromouridine) for MAD phasing. These bromine derivatives are relatively stable; however, this strategy presents numerous disadvantages.[59,65] First, it is hard to incorporate bromine at other positions other than the 5-position of uracil (or cytosine) since the synthesis can be tedious and, due to the chemical

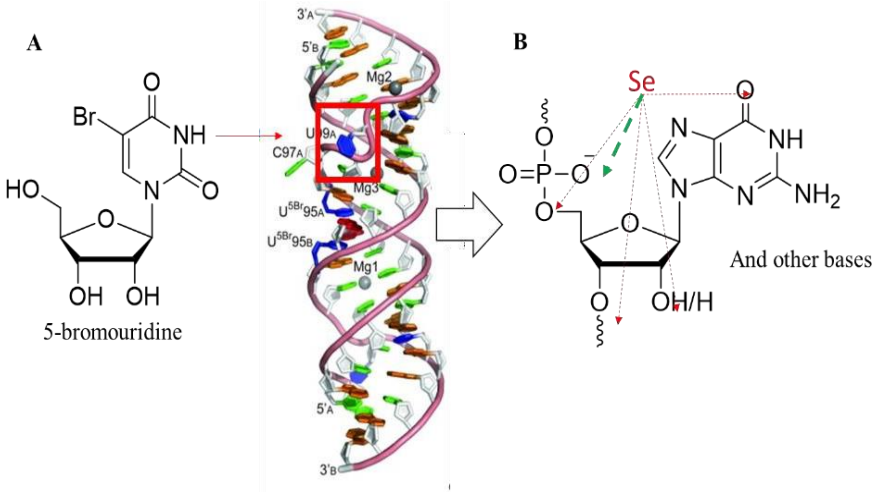


Figure 3.3-(**A**) Structure of 5-bromouridine and its effect in an oligonucleotide (reproduced with permission from *reference 64*). (**B**) Locations in the oligonucleotide scaffold for selenium introduction (adapted and reproduced with permission from *reference 57*).

nature of the halogen, it can act as a good leaving group (especially when placed at the 2'-position).[64] This constraint places the halogen in the major groove of DNA and RNA which, when in the A-form, is narrow deep, and causes structural perturbances (**Figure 3.3-A**).[64–66] Furthermore, the synthesis of 8-bromo-purine nucleosides have been explored however these tend to adopt a syn-conformation causing more structural perturbances as well as base-stacking disruptions and hydration pattern changes.[66] Lastly, bromine derivatives are light sensitive where prolonged exposure to X-ray and UV light causes decomposition.[64,67]

Recently, derivatization of nucleic acids with selenium has shown promise over the previously used halogen strategy. Essentially every oxygen atom of a nucleotide can be replaced with a selenium atom which offers a simple way to incorporate an anomalous scatter into oligonucleotides (**Figure 3.3-B**). Furthermore, this site-specific derivatization strategy has been shown to help the phase problem without causing structural perturbations which results in better crystal diffraction quality and is compatible with solid phase synthesis.[58] Moreover, Se-derivatized nucleic acids (SeNA) allow for better crystallization where, when compared to native nucleic acids, SeNA crystallized faster with a wider variety of conditions and yielded larger usable crystals.[59] Yet, most of these substitutions would still require some synthesis since, unlike the naturally occurring selenomethinonine for proteins, there are no natural selenium derived nucleic acids. A facile approach to by-pass this problem is via the introduction of a phosphoroselenoate (PSe) internucleotide linkage, where one of the nonbridging oxygen atoms is replaced by selenium.[68] Since one Se atom can provide phasing power for up to 30 nucleotides there is no need to have a completely modified backbone which has been shown to have reduced binding and antisense activity.[58,68] Furthermore, this approach allows for oxygen to selenium substitutions at the oligonucleotide level rather than constructing each individual building blocks of the oligomer.

Experimentally, the incorporation of a site-specific PSe linkage has been achieved by replacing the oxidizing solution used in the solid-phase cycle with a solution of saturated potassium selenocyanate (KSeCN).[68] This allows for the incorporation of a non-bridging selenium atom rather than a non-bridging oxygen atom. This methodology does have some disadvantages namely that the yield for these types of linkages is dramatically reduced when the PSe linkage is placed internally in an oligonucleotide. This was hypothesized to be due to the subsequent rounds of regular oxidization following the incorporation of this linkage reverting some PSe to the natural phosphodiester (PO) linkage.[68] However, the yield increases when the PSe linkage is placed between the last and second to last base.

3.1.5-Project Goals

In the interest of developing a simplified strategy to prepare selenium modified nucleic acids, we investigated a methodology to incorporate a terminal selenophosphate at the 5' end of an oligonucleotide (**Figure 3.4**). The proposed approach involves preparation of an oligonucleotide with a terminal chemical phosphorylating agent introduced by solid-phase synthesis followed with formation of the selenium - phosphorus bond using a solution of KSeCN to convert the P(III) to a P(V) species. Following this the oligonucleotide would be deprotected and analyzed. We hypothesized that this strategy would be a facile alternative to traditional incorporation of selenium since it requires no multi-step synthesis of modified nucleosides, should cause little to no structural perturbations in the oligonucleotide.

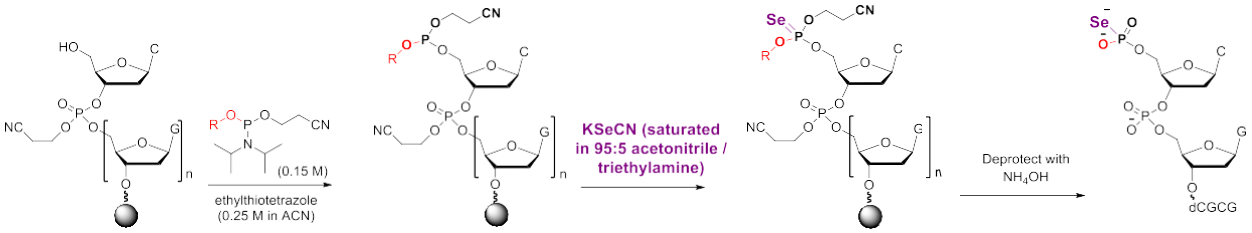


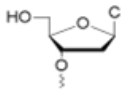
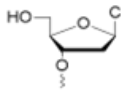
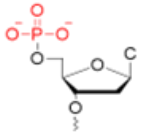
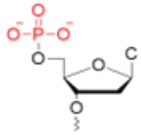
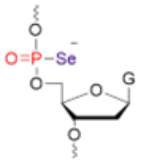
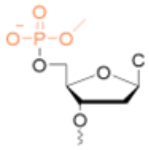
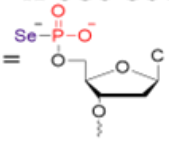
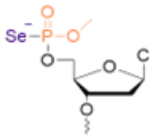
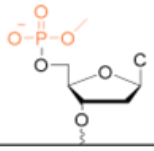
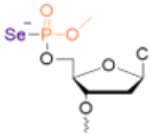
Figure 3.4-Project Objectives for Chapter 3.

3.2. Experimental

3.2.1-Synthesis and Deprotection of Oligonucleotides

The oligonucleotides were synthesized as previously described in section 2.2.2 with slight modifications. A concentration of 0.15 M was used for the chemical phosphorylation reagent (ChemGenes) as well as a coupling time 450 seconds. For KSeCN oxidation, the cycle was stopped after the coupling step. All oligonucleotides were deprotected using standard conditions as described in section 2.2.4 and purified by IEX-HPLC. Oligonucleotides synthesized for this project are shown in **Table 3.1**.

Table 3.1-Oligonucleotides synthesized for Chapter 3 with their corresponding terminal modifications.

Identifier	Sequence	Identifier	Sequence
6mer-OH	5'-CGC GCG-3' 	12mer-OH	5'-CGC GAA TTC GCG-3' 
6mer-PO	5'-**X**GC GCG-3' **X** = 	12mer-PO	5'-**X**GC GAA TTC GCG-3' **X** = 
6mer-PSeI	5'-C**X**C GCG-3' **X** = 	12mer-MePO	5'-**X**GC GAA TTC GCG-3' **X** = 
6mer-Pse	5'-**X**-CGC GCG-3' **X** = 	12mer-MePSe	5'-**X**GC GAA TTC GCG-3' **X** = 
6mer-MePO	5'-**X**GC GCG-3' **X** = 		
6mer-MePSe	5'-**X**GC GCG-3' **X** = 		

3.2.2-Oxidation Using Potassium Selenocyanate (KSeCN)

KSeCN (1.0 g, 6.94 mmol) was suspended in ACN/TEA (9.5/0.5, v/v, 1 mL) in an airtight vial flushed with argon and placed in a sand bath at 60 °C for 24 h in the dark. The following day, the chemical phosphorylating agent was coupled to freshly prepared support bound oligonucleotide as described above. Prior to oxidation, the automated DNA synthesizer was

paused, and the solid support was removed. The support bound oligonucleotide was then treated with the prepared solution of KSeCN (1 mL) for 24 hours in the dark. Following this, the selenized oligonucleotide was washed with 20 mL of anhydrous ACN, placed back on the synthesizer to finalize its cycle and deprotected as described in section 2.2.4.

3.2.3-HPLC Purification

Oligonucleotides were purified by ion exchange HPLC (IEX-HPLC) using a Dionex DNAPAC PA-100 column (0.4 cm x 25 cm) purchased from Dionex Corp using a gradient elution of 50% B over 20 minutes (Buffer A: 100 mM Tris-HCl, pH 7.5, 10% ACN and Buffer B: 100 mM Tris-HCl, pH 7.5, 10% ACN, 1 M NaCl) with a flow rate of 1 mL/min. The column was monitored at 260/280 nm. The purified oligonucleotides were diluted to 10 mL with 0.1 M sodium acetate and desalted as previously described. Purity was assessed to be over 95% for all oligonucleotides by analytical IEX-HPLC. All pure species were quantitated using a Varian Cary Model 3E spectrophotometer where concentrations were calculated using the Beer-Lambert law from absorbances measured at 260 nm. The molar extinction coefficients were calculated using the nearest-neighbour method.[13]

3.2.4-Biophysicial Studies

Thermal UV denaturation and CD samples were prepared and processed in a similar fashion as described in 2.2.11 however a buffer with 4 M sodium chloride, 10 mM sodium phosphate and 1 mM EDTA (pH 7.002) was used for the preparation of the Z-DNA samples.

3.3. Results and Discussion

3.3.1-Oligonucleotide Synthesis, Deprotection and Purification

The oligonucleotides used in this chapter were synthesized as described in section 2.2.2. During traditional solid phase synthesis, the terminal 5’-position is a hydroxyl group therefore

an extra step is necessary to achieve the desired 5'-phosphate moiety. The incorporation of a terminal 5'-phosphate group was achieved via the use of a chemical phosphorylating reagent (CPR). Currently there exist 4 different commercially available chemical phosphorylating agents each with different R-groups giving them then different properties. For example, CPR I, II and solid CPR II (**Figure 3.5**) all have a DMT handle which allows for quantitation of the coupling efficiency however bis(cyanoethyl)-N,N-diisopropylphosphoramidite has a the smallest R-group resulting in less crowding around the reactive site of the molecule which is helps improve coupling efficiency. For the purpose of this project, the latter was used as a phosphorylating reagent. All phosphoramidites were dissolved in ACN to give a final concentration of 0.15 M. For the phosphorylating reagent a coupling time of 450 seconds was required to achieve efficient coupling. Following treatment with KSeCN, the oligonucleotides were returned to the DNA synthesizer to complete their synthetic cycle. Following synthesis, all oligonucleotides were deprotected using standard ammonia deprotection and purified via IEX-HPLC. All species were characterized by MS.

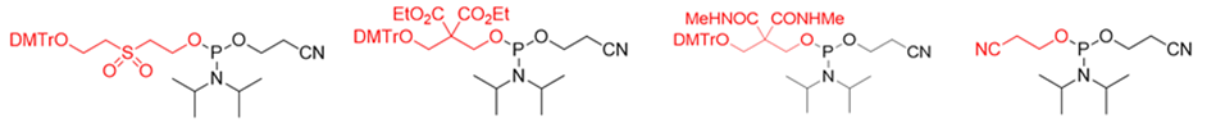


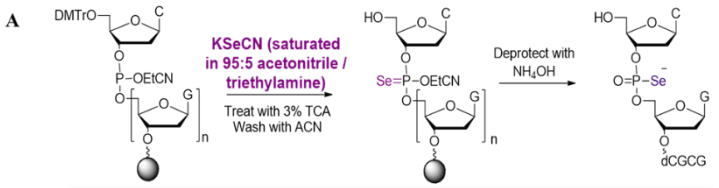
Figure 3.5-Different phosphorylating reagents and their R-groups.

3.3.2-Oxidation Using Potassium Selenocyanate

Following the successful synthesis and subsequent incorporation of the phosphate group to yield the 6mer OH and 6mer PO respectively (see Appendix II, **A2.1-4**), the 6mer PSeI oligonucleotide was made. This sequence was used as a test sequence to become accustomed to the technique of manual oxidation. The Egli group had previously shown good results using this

technique. [68] The selenizing solution was prepared 24h in advance of the synthesis of any oligonucleotide containing a phosphoroselenoate. This is done by dissolving 1 g of KSeCN in 95% ACN/5% TEA (v/v) and heating this solution overnight at 60˚C.

The synthesis of 6mer PSeI is like that of the 6mer OH except in this case, the cycle is stopped after the coupling of base C, the support bound oligonucleotide is manually selenized with KSeCN resulting in an internal phosphoroselenoate linkage. The resulting oligonucleotide is then treated with 3% TCA to remove the 5'-O-DMT group and subsequently washed and deprotected (**Figure 3.6-A**). Following, the 6mer PSeI was analyzed by IEX-HPLC. It had been previously reported in the literature that the introduction of a phosphorselenoate linkage introduces a chiral center in the oligonucleotide resulting in the R and S-isomers (**Figure 3.6-B**). The two diastereomers were separated and sent to MS for further analysis (see Appendix II, **A2.5-7**). As expected, peaks 1 and peak 2 from the internucleotide phosphorselenoate have the same mass proving that the two species are the same (**Figure 3.6- C**). Seeing that the methodology works as expected, the PSe oligonucleotide was made.



C

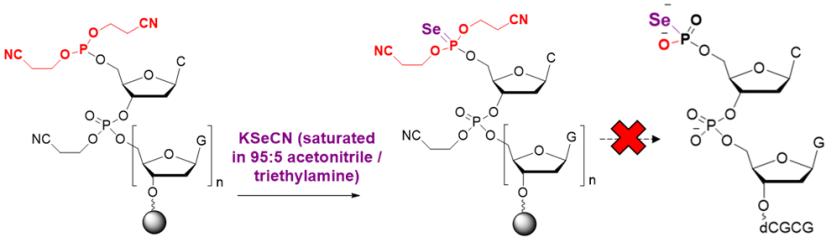
Sample	Expected MW	Obtained MW
5′HO–CGC GCG–3′	1793.2	1792.3
5′PO_4–CGC GCG–3′	1873.2	1872.3
5′–C(PO_3Se)GC GCG–3′ (1)	1856.3	**1856.3**
5′–C(PO_3Se)GC GCG–3′ (2)	1856.3	**1856.3**

Figure 3.6-(**A**) Synthesis of PSeI oligomer, (**B**) HPLC traces of 6mer-PSeI showing the R and S-isomers, (**C**) MS (negative ESI) results for 6mer PSeI.

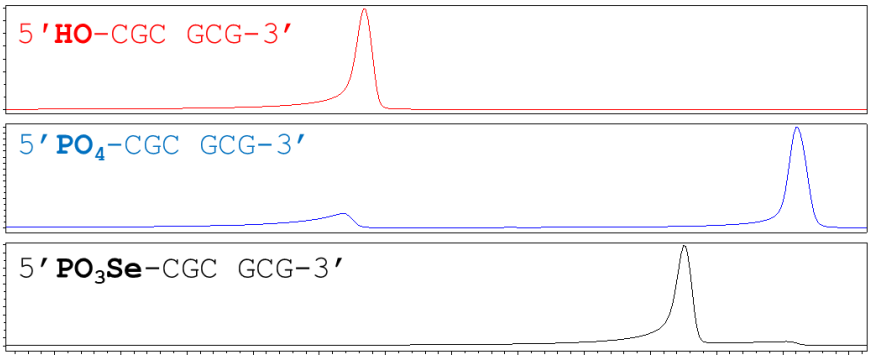
The synthesis of 6mer PSe required some slight modifications to the traditional solid phase synthesis cycle. For this oligonucleotide, the cycle proceeds as described until the coupling of the last base, which in this case is the phosphorylating agent. After coupling of this last

base, the cycle is stopped, and the column containing the support-bound oligonucleotide is removed from the synthesizer. The oligonucleotide is then washed with ACN and the oxidation/selenization step is performed manually using saturated KSeCN. The apparatus for the selenization step is composed of two 3 mL syringes each with 500 µL of the selenizing solution fixed to the top and bottom of the column (**Figure 3.7-A**). The apparatus is then left in the dark for 24h and periodically the KSeCN solution is passed through the column. The selenized oligonucleotide was then washed with 20 mL of ACN and deprotected as usual.

A



B



C

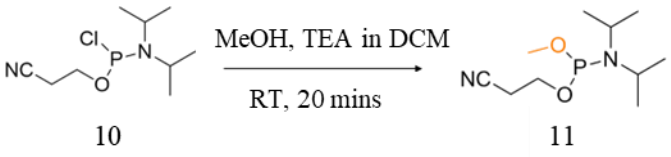
Sample	Expected MW	Obtained MW
5′**HO**-CGC GCG-3′	1793.2	1792.3
5′**PO_4**-CGC GCG-3′	1873.2	1872.3
5′**PO_3Se**-CGC GCG-3′	**1936.2**	**1871.3**

Figure 3.7-(**A**) Synthesis of 6mer-PSe, and experimental set-up for selenization, (**B**) HPLC trace of 6mer-PSe, (**C**) MS (negative ESI) results for 6mer-PSe.

There was a slight difference in retention time between the PO and PSe species (**Figure 3.7-B**) which was consistent with what we had seen with the internucleotide phosphoroselenaote. However, after MS analysis (see Appendix II, **A2.8-9**), it was revealed that these species have the same molecular weight which indicates to us that the selenization was not successful (**Figure 3.7-C**). This was repeated 2 more times, each time giving the same results. From this, it was then hypothesized that the addition of a small alkyl group on the other non-bridging oxygen atom would help retain the selenium by: (1) reducing the charge repulsion at the terminal position (-2 vs -1 charge) and (2) the R-group might mimic an internucleotide linkage (like the 6mer PSeI species) since the sugar moiety is essential a big R group. With that rationale, a methyl-phosphoramidite was synthesized.

3.3.3-Synthesis of a Methylphosphoramidite and its Incorporation Into an Oligonucleotide

The methylphosphoramidite **11** was synthesized in one step as previously described.[69] Starting from compound **10**, the addition of the methyl group was achieved using MeOH and TEA in DCM at room temperature for 20 minutes resulting in compound **11**. It is important to note that due to the volatility of compound **11**, it was isolated as a solution in ethyl acetate.



10 11

Scheme 3.1-Synthesis of methylphosphoramidite **11**.

Phosphoramidite **11** was incorporated into new oligonucleotides (6mer MePO and 6mer MePSe). The synthesis of the MePO species proceeded as described for the PO species where the phosphorylating reagent was substituted for the methylphosphoramidite. The synthesis for the

MePSe species proceeds as described for the PSe species with, once again, the substitution of the phosphorylating reagent for the methylphosphoramidite. The species were deprotected, analyzed and purified via IEX-HPLC. The purified samples were analyzed by MS which confirm the presence of a selenium atom (see Appendix II, **Supporting Figures A2.10-13**). This new modification was also incorporated into a 12mer oligonucleotide series (12mer OH, PO, MePO, MePSe) using the same strategy (see Appendix II, **Supporting Figures A2.14-21**).

3.3.4-Biophysical Studies

3.3.4.1-Thermal Denaturation Experiments

Thermal denaturation studies were run on these duplexes to assess their thermal stability. The 6mer ODN is known to form B-form DNA in low salt concentrations and Z-DNA in high salt concentration. At 90 mM NaCl, the addition of a phosphate, methyl phosphate and methyl phosphoroselenoate group causes slight decreases in the T_m value (**Table 3.2**) compared to the native oligonucleotide which is somewhat expected since there is a lot more charge repulsion from the negatively charged phosphate group. However, at higher salt concentration the opposite is observed where the high salt concentration is able to accommodate and stabilize charge repulsion resulting in a T_m increase. For the 12mer, there was a slight increase in the T_m values indicating that the farther away the negative charges are from one another, the more stabilizing an effect is observed for the duplex. In general, the thermal melt experiments indicate that the modifications do not cause significant change to the thermal stability of the DNA duplex.

Table 3.2-T_m values of 6mer and 12mer series at different salt concentrations.

	6mer Series				**12mer Series**	
Species	T_m in °C (90mM NaCl)	ΔT_m (°C)	T_m in °C (4M NaCl)	ΔT_m (°C)	T_m in °C (90mM NaCl)	ΔT_m (°C)
5′ OH	47.2	-	38.1	-	63.8	-
5′ PO	43.4	-3.8	44.6	+6.5	67	+3.2
5′ MePO	45.5	-1.7	45	+6.9	63.6	-0.2
5′ MePSe	42.6	-4.6	37.7	-0.4	67.8	+4

3.3.4.2-Cicrular Dichroism Spectroscopy

CD spectra were acquired on these oligonucleotides to observe if any deviations in global structure was observed as a result of the modifications. For the 6mer and the 12mer at low salt concentrations, the oligonucleotides exhibit characteristic B-form duplex maxima and minima at 280nm and 250nm respectively (**Figure 3.8-A and B**). On the other hand, the 6mer at high salt concentration exhibits a characteristic Z-form duplex signature with maxima and minima at 270 and 295nm respectively (**Figure 3.8-C**). Hence this data indicates that the presence of a terminal methyl-phosphoroselenoate does not change the global structure of the DNA duplexes whether it be in B-form or Z-form.

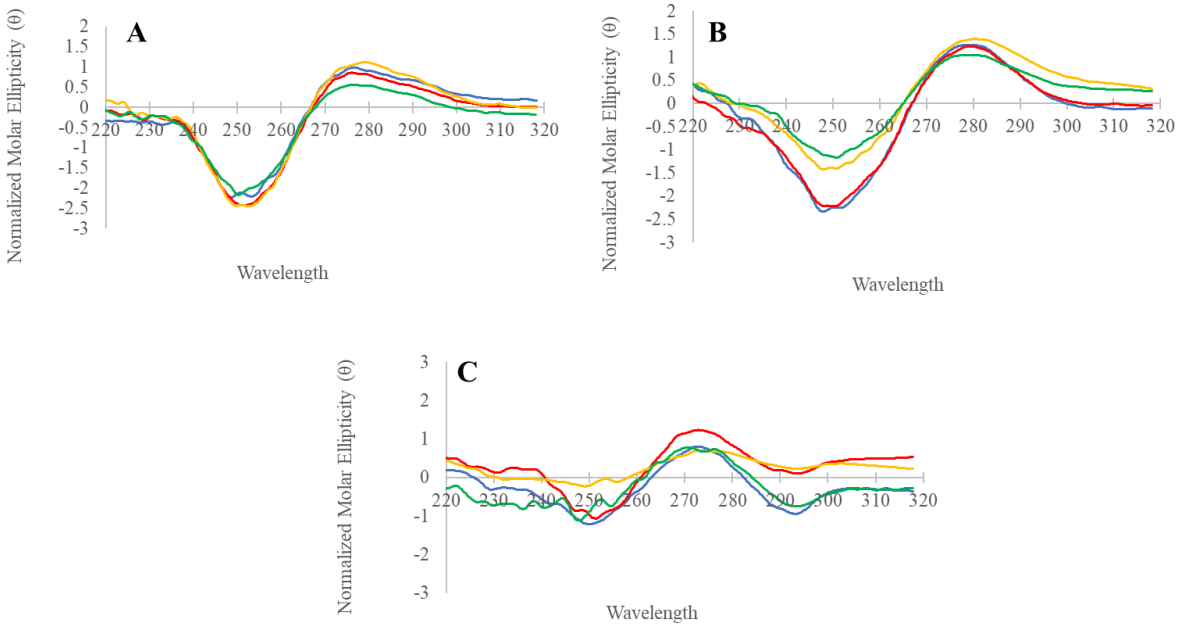


Figure 3.8-CD spectra (**A**) 6mer series with 90 mM NaCl (OH, PO, MePO and MePSe), (**B**) 12mer series with 90 mM NaCl (OH, PO, MePO and MePSe), (**C**) 6mer series with 4 M NaCl (OH, PO, MePO and MePSe).

3.4. Conclusion

Introduction of a selenium atom at a terminal phosphorus atom of an oligonucleotide was accomplished by coupling a methyl containing phosphoramidite to the 5'-end of the oligonucleotide followed by treatment with a saturated solution of KSeCN in ACN/TEA. Initially, the incorporation of the 5'-phosphoroselenoate was a challenge when the phosphorylating agent bis(cyanoethyl)-N,N-diisopropylphosphoramidite was used as the selenium atom was not retained possibly due to oxidation with air or water. However, this was by-passed via the synthesis of a methyl phosphoramidite which, following its incorporation into an oligonucleotide and subsequent selenization, the retention of the selenium atom was achieved. Furthermore, T_m and CD data suggest minimal perturbances compared to the native duplex and the presence of the MePSe handle does not affect formation of Z-DNA. This methodology may be useful to add to the approaches to prepare selenium modified oligonucleotides.

Chapter 4 - Development of an Assay to Evaluate O^6-Alkylguanine DNA Alkyltransferase Activity

Abstract

Exposure of DNA to various chemical agents that are endogenous and exogenous to the cell can result in the introduction of various adducts to the nucleobase. Such modifications may interfere with important cellular processes such as replication and translation. However, DNA repair can restore the DNA by excision of various lesions. One DNA repair protein, O^6-alkylguanine DNA alkyltransferase (AGT), removes alkyl groups from the O^6-atom of 2'-deoxyguanosine and the O^4-atom of thymidine. However, this protein may be inhibited by various O^6-alkylguanine derivatives. Assays to evaluate the repair activity of this protein assist in guiding the design of various substrates for applications related to biotechnology and health. Herein, a method to evaluate the repair proficiency of the AGT protein in the presence of an inhibitor will be described.

4.1. Introduction

4.1.1-DNA Damage

Every minute, DNA can undergo a variety of damage inducing events. DNA damage can be characterized as an alteration of the genetic material which can have an impact on the process of replication.[70] The result of these damaging events on DNA are called lesions. Depending on the damage event, examples of these lesions include single and double strand breaks, abasic sites, mismatches, bulky adducts as well as inter- and intrastrand cross-links (**Figure 4.1**).[71]

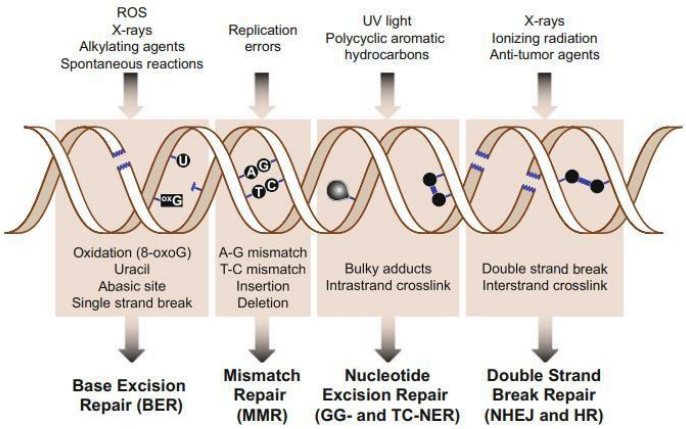


Figure 4.1-DNA damaging events and their resulting lesions (reproduced with permission from *reference 69*).

DNA damaging agents may result from exogenous or endogenous agents or processes. Endogenous DNA damage is caused by internal sources like replication stress, errors in cell metabolism or reactive oxygen species (ROS). In contrast, exogenous DNA damage is due to external factors such as radiation (UV and X-rays), viral infections and chemotherapy.[70] Each type of DNA damage gives rise to a distinct change in the chemical structure of DNA: double strand breaks originate from ionizing radiation and chemotherapeutic drugs, formation of a chemical bond between neighbouring nucleotides like a T-T dimer arises from UV light, a chemical modification of a single nucleotide and interstrand cross-links are due to ROS or

chemotherapeutic drugs.[72] If DNA damage is left untreated it can lead to cell death and tissue damage, neurodegenerative diseases, cancer formation and premature aging.[72,73] Depending on the extent of the damage, the cell cycle will be stopped to allow for repair or the induction of apoptosis in order to stop proliferation of damaged cells.

4.1.2-Alkylating Agents

Alkylating agents are an important class of chemotherapeutic drugs.[74] They act by inhibiting protein synthesis by physically blocking DNA transcription via the addition of alkyl groups at certain positions on the bases of DNA. Alkylating agents can be divided into 3 categories: mono-functional methylating agents (temozolomide and dacarbazine), bifunctional alkylating agents (nitrogen mustards) or chloroethylamine agents (carmustine and nimustine).[75] Monofunctional, or simple, methylating agents typically methylate the O and N atoms of DNA. The most common alkylating targets for these electrophilic carcinogenic molecules are the N^7, N^3 and O^6 positions of guanine as well as the N^1 and N^3 positions of adenine (**Figure 4.2**).[76]

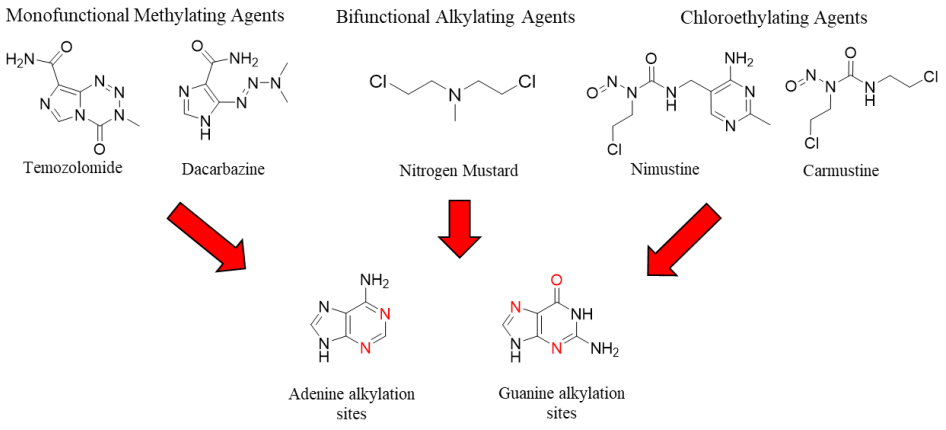


Figure 4.2-Common alkylating agents and their alkylation sites.

Over 80% of these alkylating events result in N-methylated adducts whereas O-methylated adducts account for less than 10%.[75] However, these lesions exhibit very different stabilities. Taking for example the most stable N-methylated adduct, N^7-methylguanine, which has a half-life

of 80h versus its O analogue (O^6-methylguanine) which will persist indefinitely until it is repaired.[75,77] This vast difference in stability points towards the highly mutagenic and genotoxic nature of O-methylated adducts and hence their need to be efficiently repaired. Such modifications can lead to a variety of alterations which are dependent on: (1) where the adduct was formed, (2) when in the cell cycle it was formed, and (3) which repair process can be used to repair the damage.[76]

As previously mentioned, temozolomide and dacarbazine are common monofunctional alkylating agents which result in the formation of the O^6-methylguanine lesion. In the body, both of these drugs undergo reactions which result in the formation of the reactive methanediazonium ion (**Figure 4.3**). It is a short-lived electrophilic reactant that has the ability to transfer a methyl group to a nucleophilic position in a molecule, i.e. the O^6-position of guanine.[78] If this lesion is not repaired via the direct reversal of damage pathway using the enzyme O^6-alkylguanine-DNA-alkyltransferase (AGT), also referred to as O^6-methylguanine-DNA-methyltransferase (MGMT), it can mispair with thymine rather than adenine during DNA replication which can then be recognized by enzymes involved in the mismatch repair (MMR) pathway. However, this pathway triggers ineffective cycles of thymine elimination and reinsertion resulting in strand breakage

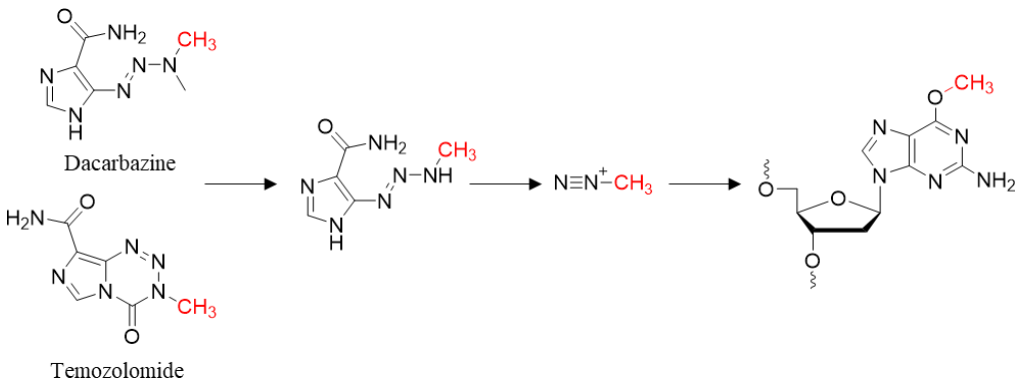


Figure 4.3-Alkylation of the O^6-position of 2'-deoxyguanosine by temozolomide or dacarbazine.

eventually leading to cell cycle arrest and apoptosis.[78] Temozolomide and dacarbazine can also cause N^7-methylguanine adducts however these are readily repaired via base excision repair (BER) pathway.[75,78]

4.1.3-O^6-Alkylguanine DNA Alkyltransferase

AGTs are small repair proteins present in bacteria, archaea and eukaryotes.[79] They are composed of two domains: a highly conserved C-terminal domain and a less conserved N-terminal domain (**Figure 4.4**). The human variant is known as human O^6-alkylguanine DNA alkyl transferase (hAGT). The N-terminal domain of this variant is composed of an α/β roll where a single α-helix is encapsulated within a four-stranded anti-parallel β-sheet and plays a regulatory role.[79–81] On the other hand, the functional roles of hAGT are found at the C-terminal domain. This moiety is composed of: a short two-stranded parallel β-sheet, five α-helices, the IPCHRV motif which contains the conserved active-site cysteine, the helix-turn-helix (HTH) DNA-binding motif and the O^6-alkylguanine-binding channel.[80] This protein also contain a zinc (II) ion bridge between the N- and C-terminal domains which has been hypothesized to aid in stabilizing the interface between the two domains.[80]

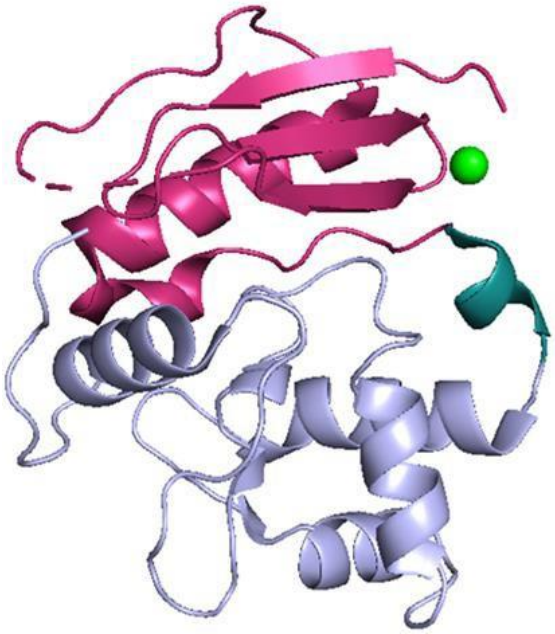


Figure 4.4-Structure of hAGT highlighting its N-terminal (pink), Mg^{2+} ion (green), helix 3 connecting the N and C-terminals (teal) and the C-terminal (light purple). The model was prepared using PyMol with PDB entry 1EH6.

4.1.3.1-DNA Repair with hAGT

It has been shown that hAGT can repair various O^6 adducts of guanine as well as certain O^4 adducts of thymine and certain cross-links.[82,83] This protein does so by physically flipping the alkylated base out of the duplex into the active site which contains proline, cysteine, and histidine residues. This is accomplished via binding of hAGT to the minor groove of DNA by the HTH motif via residues Ala127-Gly136 which allows for close-packing. Residues Tyr114-Ala121 as well as residues Phe96-Val105 can then interact with the phosphate backbone of DNA via helix dipoles further stabilizing the protein and DNA interaction. These interactions widen the minor groove and bend the DNA strand away from the protein. It has been hypothesized that these structural changes to DNA rather than to the protein help flip the alkylated nucleoside out from the major groove and allows for cooperative binding.[81,84] Following this, in a concerted S_N2-type

fashion, the methyl adduct is transferred to the sulfur atom of the active Cys145 residue (**Figure 4.5**). This step resembles closely the catalytic triad of serine proteases where His146 acts as a base to deprotonate Cys145 which can in turn, act as a nucleophile in order to dealkylate the O^6-position.[81] The repaired base is then flipped back into the major groove of DNA while the alkylated protein is marked for degradation via the ubiquitin-proteasome pathway (UPP).

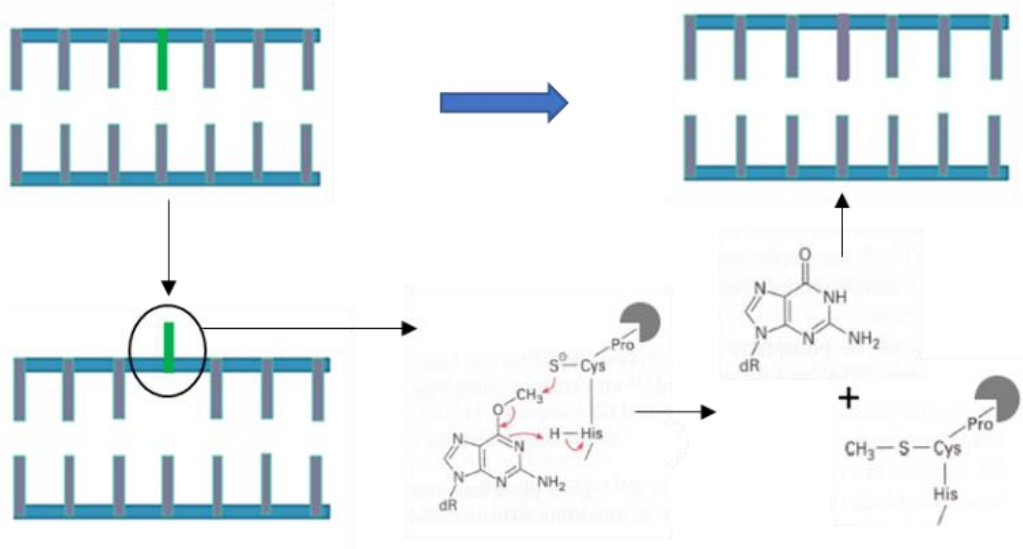


Figure 4.5-hAGT repair mechanism of O^6-methyl-dG adducts (partially adapted and reproduced from *reference 4*).

hAGT has also been shown to repair inter- and intrastrand cross-links.[83] Cross-links are considered toxic as they lead to distortion in the DNA structure which impairs cellular processes and cause helical bending which affects the affinity of DNA-binding proteins.[83] The proposed mechanism of action is similar to what has been previously described however, two repair events and two hAGT proteins are needed for complete repair to occur. The first repair event leads to the formation of a DNA-cross-link-protein intermediate. Following the second repair event, the native DNA is restored, and a protein-cross-link-protein species is formed.[83]

4.1.3.2-Inhibition of hAGT

It was previously mentioned that there is a need to efficiently repair O^6-methylguanine lesions that can arise, however, it can be a bit of a double-edged sword when it comes to cancer treatment involving alkylating agents where it is intended to induce DNA damage to elicit a therapeutic effect. For example, the alkylating therapeutic agent temozolomide which has the ability to alkylate the N^3 position of adenine as well as the N^7 and O^6 positions of guanine. This is desirable as a chemotherapeutic approach as these lesions prevent replication which eventually leads to cell death essentially killing the cancer cells. However, hAGT can repair these lesions quite efficiently which results in resistance to temozolomide. Therefore, during cancer treatments it is important to inhibit repair proteins such as hAGT to allow for the formation of these alkylated adducts.

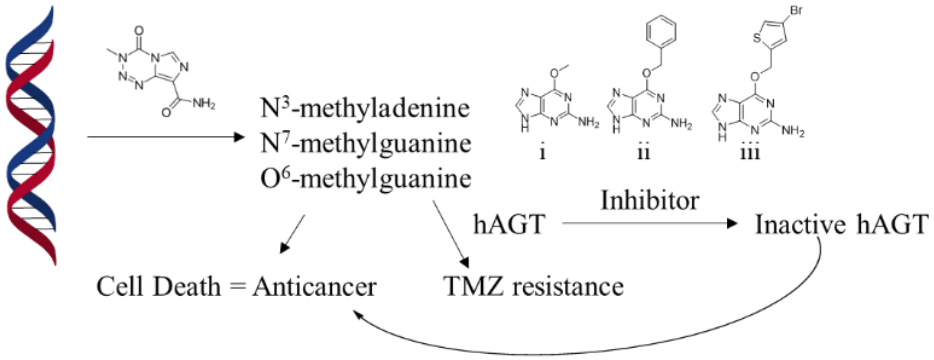


Figure 4.6- Inhibitors of hAGT for cancer treatment (i) O^6-methylguanine, (ii) O^6-benzylguanine and (iii) 6-[(4-bromo-2-thienyl)methoxy]-9H-purin-2-amine.

One of the first hAGT inhibitors was O^6-methylguanine. *In vitro*, the tumor cells treated with the alkylated nucleobase exhibited a loss of hAGT activity as well as an increased sensitivity to alkylating agents. However, when using a mouse model, no enhancement of therapeutic index was observed hence O^6-methylguanine never entered clinical trials.[85] The first inhibitor to enter clinical trials was O^6-benzylguanine (BG). This inhibitor, combined with temozolomide, resulted in the successful obliteration of brain tumor, melanoma, lymphoma, and colon cancer.[86] Yet,

there are therapeutic limitations to using BG including low affinity for hAGT compared to alkylated residues in a DNA duplex, its low bioavailability, poor water solubility and rapid plasma clearance.[80]

More recently, 6-[(4-bromo-2-thienyl)methoxy]-9H-purin-2-amine also known as PaTrin-2 (trade name lomeguatrib) has been explored as a potential inhibitor of hAGT. This drug exhibits a higher potency compared to BG (IC_{50} of 3.4 nM compared to 180 nM for BG), is less toxic, more bioavailable and has a slightly higher water solubility.[87,88] PaTrin-2 is currently in phase II clinical trials for the treatment of melanoma, brain, prostate, and colon cancers.[89]

These O^6-alkylguanine type derivatives operate in a similar fashion. Typically, cancer patients are given the inhibitor prior to their chemotherapy.[90] During this time, the inhibitor will scavenge to find and bind to hAGT. This results in the formation of the alkylated protein which is marked for degradation. Researchers have found that the efficacy of both temozolomide and dacarbazine increased slightly (3% increase) when administering BG prior to the chemotherapeutic drug, however the efficiency of these chemotherapeutic agents was increased to 16% when using PaTrin-2.[90] This small increase is attributed to the dose-limiting side effect of these two drugs where above a certain concentration, these drugs cause myelosuppression which is a condition that causes bone marrow activity to decrease resulting in fewer red blood cells, white blood cells and platelets.[87,90,91] This has been credited to the many off-target effects of chemotherapy. Since chemotherapeutic alkylating drugs are not specific, they will attack any and every cell especially any fast-replicating cells like those found in bone marrow.[92] Thus, the combination of targeted hAGT inhibition and chemotherapy has the potential to improve the effects of alkylating drugs.

4.1.4-Project Goals

The development of more potent and selective hAGT inhibitors would be desirable in order to increase the effectiveness of chemotherapeutic alkylating agents. To assist in this endeavor, an assay to screen for hAGT inhibitors was evaluated. In this approach, an assay our group has previously employed to examine AGT activity towards DNA duplexes containing various lesions at the O^6-position of guanine (and O^4-position of thymine) method was adapted with the addition of a known AGT inhibitor (O^6-benzylguanine). This methodology involves a DNA duplex containing a O^6-methylguanine residue located in in the middle of the sequence that is cleaved by the restriction enzyme PvuII. Upon incubation of the inhibitor with hAGT, one of two possible outcomes can occur: (1) No reaction would occur between the inhibitor and hAGT in which case the protein would repair the lesion which would restore the native DNA or (2) hAGT would react with the inhibitor rendering it inactive and the lesion would not be repaired. Then the restriction enzyme PvuII would be added. If repair occurs (i.e. the inhibitor is inactive), then PvuII would cleave the duplex and smaller oligonucleotide products would be visible by gel electrophoresis. However, if repair by AGT does not occur (i.e. the inhibitor was active) the alkylated DNA would remain intact as the restriction enzyme would not cleave the duplex and the full length oligonucleotide would be visible on the gel.

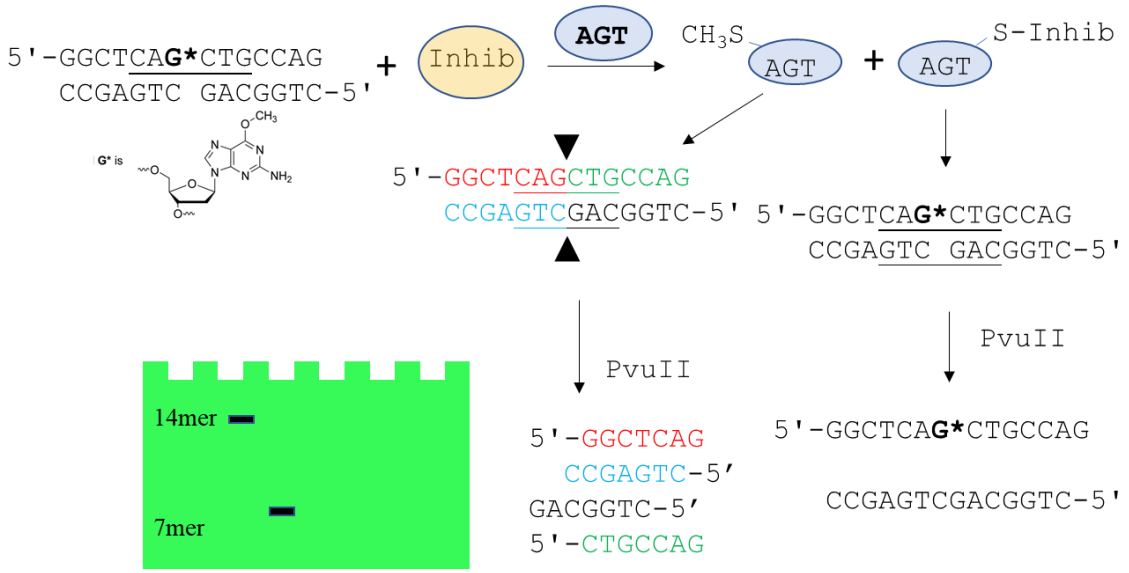


Figure 4.7-Project Objective for Chapter 4.

4.1. Experimental

4.2.1-Synthesis of Small Molecules

O^6-Benzylguanine (13). To a solution of benzyl alcohol (2.70 g, 25.0 mmol) in dioxane (30 mL) on ice, was added NaH (0.84 g, 21.0 mmol) slowly and the solution was left to stir vigorously for 20 minutes. The reaction was brought to room temperature and allowed to mix for an additional 40 minutes. 2-amino-6-chloropurine (1.70 g, 10.0 mmol) was added to the reaction mixture slowly and the solution was allowed to reflux for 20 hours. The reaction was then cooled to room temperature, filtered, and diluted to 100 mL with water. The solution was neutralized with acetic acid (pH = 7.0) and washed with hexanes (3 x 40 mL). The aqueous layer was concentrated and adhered to silica gel before purification by FCC using MeOH/DCM (0/100 to 2/98) as eluent to yield a white powder (1.39 g, 5.75 mmol, 57%). R_f (SiO_2 TLC): 0.13, MeOH/DCM (5/95). λ_{max} (ACN): 208 nm ^{1}H NMR (500 MHz, DMSO-d6, ppm): δ 12.48 (bs, 1H, NH), 7.84 (s, 1H, H8), 7.49-7.47 (m, 2H, Ar), 7.38-7.29 (m, 3H, Ar), 6.29 (2, 2H, NH_2), 5.47 (s, 2H, CH_2). ^{13}C NMR

(125.7 MHz, DMSO-d6, ppm): δ 160.28, 160.08, 155.64, 149.47, 141.75, 138.24, 137.21, 128.88, 128.82, 128.43, 113.94, 67.14. HRMS (ESI-MS) *m/z* calculated for $C_{12}H_{11}N_5O_2$ 242.1042: found 242.1040 [M + H^+].

4.2.2-Oligonucleotide Synthesis, Deprotection and Purification

CJW 1083 was synthesized as described in 2.2.2. CJW 1128 was synthesized with fast deprotecting amidites as described in section 2.2.2. An O^6-methyl-2'-deoxyguanosine modification was also incorporated into this oligonucleotide using O^6-methyl-2'-deoxyguanosine CED phosphoramidite (Glen Research) at a concentration of 0.15 M and a coupling time of 300 seconds. Both oligonucleotides were deprotected using the appropriate deprotection strategy (2.2.4). CJW 1083 was purified via PAGE whereas CJW 1128 was purified by IEX HPLC using a gradient of 50% B over 35 minutes as described in section 3.2.4.

4.2.3-Radioactivity Labelling and Duplex Formation

The target strand was labelled with ^{32}P, using T4 polynucleotide kinase from Fermentas (Burlington, ON), and [$\gamma^{32}P$] ATP from PerkinElmer (Woodbridge, ON). 100 pmol of the CJW 1128 was incubated with 1 µL of γ-[^{32}P]-ATP (10µCi/µL), 5 units of T4 PNK in 10 µL of 1x PNK buffer for 1 hour at 37 °C. After 1 h of incubation the samples were boiled for 10 minutes and let cool down to room temperature. 110 pmol of CJW 1083 was added and the solution was diluted to 100 µL to afford a stock of radio-labelled DNA duplex (dsDNA) of 1 pmol/ µL. The sample was boiled for 2 minutes and allowed to cool to room temperature.

4.2.4-Protein Purification and Overexpression

The hAGT protein was purified and overexpressed as described previously.[93] The mass of the hAGT protein was verified by ESI-MS run in positive ion mode on a Micromass Q-Tof Ultima API and is in accordance with the expected mass of 21 875 Da.[93]

4.2.5-O^6-Benzylguanine Assays

4.2.5.1-Determination of Working Concentration of hAGT

1 pmol of dsDNA was incubated at 37˚C for 1h with various amounts of hAGT and activity buffer (10 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, pH 7.6) to give final concentrations of 0.1, 0.2, 0.5, 1 and 3 pmol/μL of hAGT. The reaction was quenched by boiling at 95˚C for 10 minutes. $MgCl_2$ was added to give a final concentration of 10 mM. 5 units of PvuII (NEB) was added and the reaction was allowed to proceed for 2h at 37˚C. The reaction was quenched by the addition of stop buffer (20 mM EDTA in 95% formamide, pH 8.0) to give a final volume of 20 μL and heat killed at 95˚C for 5 minutes. 15 μL of the reaction mixture was directly loaded onto a 20% PAGE gel. The resulting gel was visualized by X-ray film. The film was exposed for 1h to the radioactive gel and then developed.

4.2.5.2-Inhibition of hAGT With O^6-Benzylguanine Using Radiolabelled Oligonucleotides

hAGT was incubated with the desired amount of the inhibitor at 37˚C for 30 minutes in activity buffer (10 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, pH 7.6) to give a final concentration of hAGT of 1 pmol/μL. dsDNA was then added to the reaction (final concentration of 0.1 pmol/μL) and incubated for 1h at 37˚C. The reaction was quenched by boiling at 95˚C for 10 minutes. $MgCl_2$ was added to give a final concentration of 10 mM. 5 units of PvuII (NEB) was added and the reaction was allowed to proceed for 2h at 37˚C. The reaction was quenched by the addition of stop buffer (20 mM EDTA in 95% formamide, pH 8.0) to give a final volume of 20 μL and heat killed at 95˚C for 5 minutes. 15 μL of the reaction mixture was directly loaded onto a 20% PAGE gel. The resulting gel was visualized by X-ray films. The films are exposed for 1h to the radioactive gel and then developed.

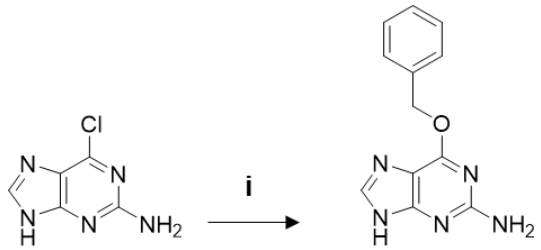
4.2.5.3- Inhibition of hAGT With O^6-Benzylguanine Using SYBR Gold Staining

A stock of dsDNA (5 pmol/μL) was prepared as previously described. hAGT was incubated with the desired amount of the inhibitor at 37˚C for 30 minutes in activity buffer (10 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, pH 7.6) to give a final concentration of hAGT of 6 pmol/μL. dsDNA was then added to the reaction (final concentration of 0.6 pmol/μL) and incubated for 1h at 37˚C. The reaction was quenched by boiling at 95˚C for 10 minutes. $MgCl_2$ was added to give a final concentration of 10 mM. 15 units of PvuII (NEB) was added and the reaction was allowed to proceed for 2h at 37˚C. The reaction was quenched by the addition of stop buffer (20 mM EDTA in 95% formamide, pH 8.0) to give a final volume of 35 μL and heat killed at 95˚C for 5 minutes. 20 μL of the reaction mixture was directly loaded onto a 20% PAGE gel. The resulting gel was visualized via SYBR Gold staining. The gel was submerged in a solution of SYBR Gold (prepared according to the protocol supplied by ThermoFisher Scientific) for 20 minutes and rinsed twice with distilled water.[41] Amersham™ Typhoon™ 5 Biomolecular Imager (Cytiva).

4.2. Results and Discussion

4.3.1-Synthesis of O^6-Methylguanine and O^6-Benzylguanine

O^6-Methylguanine was prepared as previously described.[94] The synthesis of O^6-benzylguanine (**Scheme 4.1**) was initiated by the deprotonation of benzyl alcohol with NaH in dioxane. As this is an exothermic reaction, it was allowed to proceed in an ice bath. Following, the solution was brought to room temperature in preparation for the addition of 2-amino-6-chloropurine. The resulting solution was refluxed for 20h affording compound **13**. This compound was purified by FCC and identified via ^1H and ^{13}C NMR (see Appendix III, **A3.1-2**).



12 **13**

Scheme 4.1-Synthesis of O^6-benzylguanine. (i) benzyl alcohol (2.5 eq), sodium hydride (2.1 eq), dioxane, 20 h, reflux, (57 %).

4.3.2-Oligonucleotide Synthesis

For this study, two oligonucleotides were synthesized: CJW 1083 and CJW 1128 where the former is the complement and the latter the target strand (**Table 4.1**). CJW 1083 was synthesized using standard SPS as previously described. However, some modifications were made for the SPS of CJW 1128. This oligonucleotide contains the O^6-methylguanine modification which, upon standard deprotection with ammonia can result in the conversion of the methyl to an amine moiety.[94] Hence, fast-deprotecting amidites were used to by-pass the need for ammonia during the deprotection step. Furthermore, since fast-deprotecting amidites were used, it was necessary to use phenoxyacetic anhydride as a capping agent rather than acetic anhydride. This is due to the fact that an exchange can occur between the Pac-protected dG and acetic anhydride during the capping step resulting in a stable impurity.[95] These fast-deprotecting amidites can be readily deprotected with 0.05M K_2CO_3/MeOH at room temperature overnight. Neutralization with glacial acetic ensues.

Purification of CJW 1083 was achieved via PAGE and its identity confirmed by MS (see Appendix III, **A3.3**). As for CJW 1126, purification was achieved via IEX-HPLC and its identity confirmed by MS (see Appendix III, **A3.4-5**). The choice of HPLC over PAGE in this case

resides in the fact that HPLC allows for higher recovery of purified product.[96] The synthesis of CJW 1128 did not proceed as smoothly as that of CJW 1083 which is expected when using modified phosphoramidites hence we wanted to ensure maximum recovery after purification.

Table 4.1-Oligonucleotides synthesized for Chapter 4.

Identifier	Sequence		Modification
CJW 1083	5'-CTG GCA GCT GAG CC-3'	Complement	–
CJW 1128	5'-GGC TCA **X**CT GCC AG-3'	Target	X: O^6-methyl-2'-deoxyguanosine

4.3.4-O^6-Benzylguanine AGT Inhibitor Assays

Following synthesis, a portion of the target oligonucleotide was radiolabelled using T4 polynucleotide kinase from and [γ^{32}P] ATP. This enzyme allows for the transfer of the γ-phosphate from ATP to the 5'-hydroxyl of single-stranded DNA or double-stranded DNA. A ratio of 1:1.1 target-to-complement is used to assemble the labelled duplex to ensure all the target is in duplex form (dsDNA).

4.3.4.1-Determination of Optimal Concentration of hAGT

The first assay that was performed was used to determine what concentration of hAGT to use for the subsequent assay. Concentrations of hAGT from 0.1-3 pmol/μL were used. In the first step in the assay, hAGT is incubated with the dsDNA (0.1 pmol/μL) in activity buffer for 1h at 37˚C which are the optimal conditions for repair of an O^6-methylguanine lesion.[93] Following, the solution is boiled at 95˚C for 10 minutes to ensure that hAGT is inactivated. The target oligonucleotide was constructed with the restriction site of the PvuII at its center which has been shown to be inhibited in the presence of a O^6-methylguanine lesion. Therefore, to determine if the lesion was by hAGT, the solution was treated with PvuII in the presence of 10mM $MgCl_2$ for 2h

at 37˚C. Once again, the enzyme was heat inactivated and diluted with stop buffer in preparation for visualization by PAGE. The gel was visualized via X-ray film. In total 8 samples were prepared (see Appendix III, **Supporting Figure A3.6**). Samples 1-3 served as control where sample 1 showed that in the absence of hAGT and restriction enzyme, the duplex is not repaired nor cleaved, sample 2 demonstrates that the restriction enzyme cannot cleave an unrepaired duplex and sample 3 shows that in the absence of the restriction enzyme, the duplex cannot be cleaved. As can be seen in **Figure 4.8**, all these have a lower mobility on the gel which indicates that they correspond to the intact 14-mer.

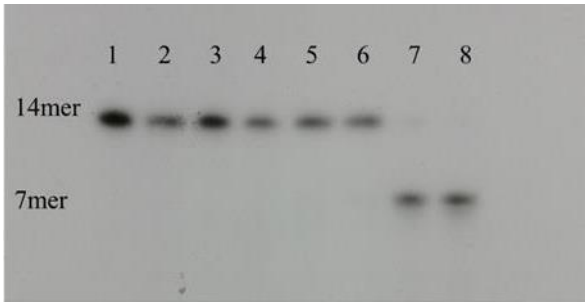


Figure 4.8 – Determination of optimal hAGT concentration for the assay. 20% PAGE, 250V, 1.5h, visualized by autoradiography.

Samples 4-8 contain varying amounts of hAGT and it can be seen that at a concentration of 1 pmol/μL, hAGT can fully repair the duplex which can be deduced by the higher mobility of the shorter oligonucleotide product produced by PvuII cleavage. This ratio of 1:10 DNA-to-protein was used for all subsequent assays.

4.3.4.2-Inhibitory Assays of hAGT Using O^6-Benzylguanine

Following these results, the first assay with BG was performed. Previous work from the Kool group showed that when hAGT was incubated with 2.5-80 equivalents (Eqs) with respect to hAGT of BG for 30 minutes at 37˚C, it produces a sigmoidal inhibition curve.[97] An assay similar to what was is described above (section 4.3.4.2) was set up with a pre-incubation step to inhibit hAGT (see Appendix III, **Supporting Figure A3.7**). This assay contains an additional control to ensure no reaction takes place between dsDNA and the inhibitor (**Figure 4.9-A, lane 4**). From the migration pattern on the gel, it was deduced that the amount of inhibitor is too high and thus a

lower concentration is necessary to see a transition from no inhibition to full inhibition. Another assay was set up (**Figure 4.9-B**), this time with much smaller Eqs (0.01-0.2). This assay demonstrated that these Eqs were too little as all the duplexes were repaired resulting in cleavage at the restriction site and faster migration in the gel. All these assays were performed in duplicates.

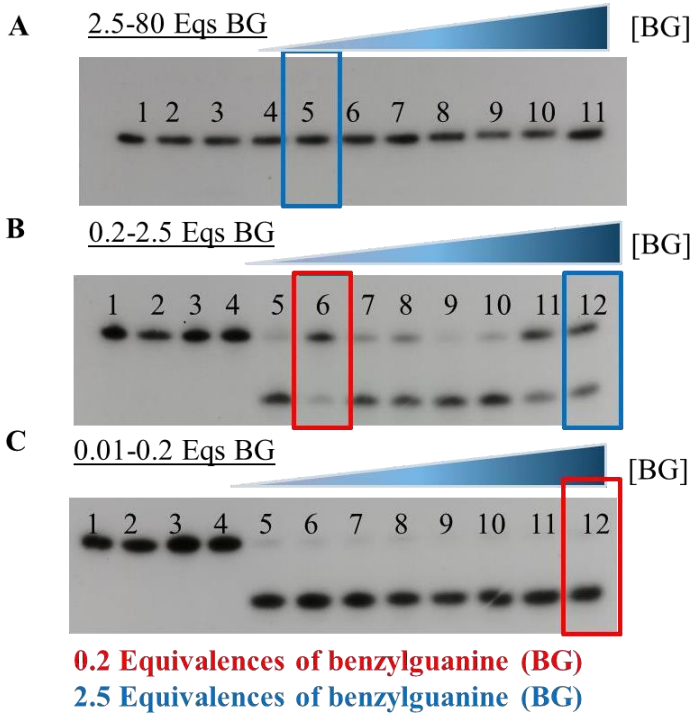


Figure 4.9-Inhibition of hAGT using (**A**) 2.5, 8, 10, 20, 25, 40 and 80 Eqs of BG (lanes 5-11) (**B**) 0.2, 0.5, 0.7, 1, 1.5, 2 and 2.5 Eqs of BG (lanes 6-12) (**C**) 0.01, 0.02, 0.05, 0.075, 0.1, 0.15 and 0.2 Eqs of BG (lanes 6-12). 20% PAGE, 200V, 1.5h.

Given the observation that 0.01 equivalents produce no inhibition of hAGT but 2.5 equivalents fully inhibit the protein, then preparing an assay with 0.01-2.5 Eqs (**Figure 4.9-**C) should produce a good inhibition profile. However, after analyzing the duplicate gels and comparing them with those from previous experiments, some inconsistencies were revealed. As can be seen in **Figure 4.9**, the lanes highlighted in a blue box (lane 5 in **A** and lane 12 in **B**) contain the same amount of inhibitor (2.5 Eqs) however the results are different in that in **A** the protein is

fully inhibited whereas in **B** the protein is only partially inhibited. A similar situation is observed when comparing the lanes in a red box (lane 6 in **B** and lane 12 in **C**). Both these samples contain 0.2 Eqs of BG however in **B** the protein is fully inhibited whereas in **C** the protein is not inhibited.

It was then decided to investigate a different visualization method that would allow for visualization of all the species present rather than only the target strand. Hence, staining of the gel with SYBR Gold which is a highly sensitive nucleic acid stain allowing for detection of pmol amount of DNA upon excitation with blue light was employed. This method is more convenient for benchtop work as it does not require radiolabeling of the target strand, has a longer shelf life and is less expensive.[98] Finally, a triplicate assay with 0.5-5 Eqs of BG (see Appendix III, **Supporting Figure A3.8** for reaction conditions) was performed and showed repeatedly that in order to achieve full inhibition of hAGT with BG, 2 Eqs of the inhibitor are necessary (**Figure 4.10-**lane 7).

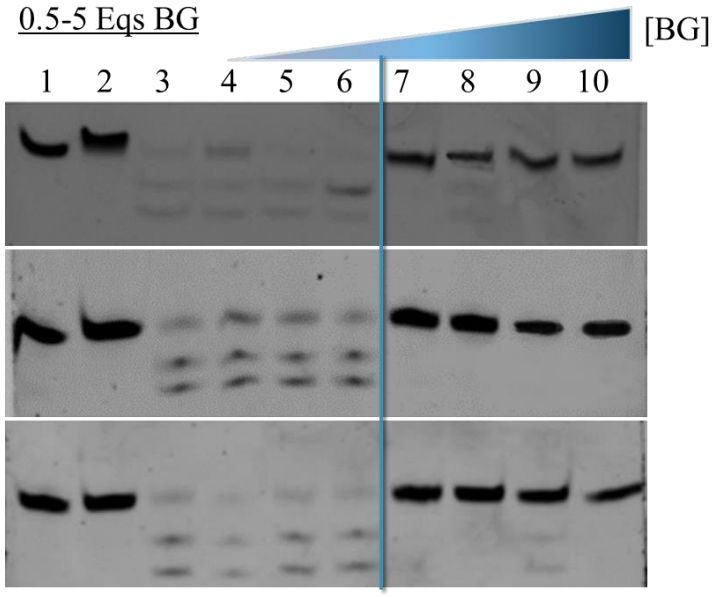


Figure 4.10-Triplicate gels for the Inhibition of hAGT using 0.5, 1, 1.5, 2, 2.5, 3 and 5 Eqs of BG (lanes 4-10) visualized by SYBR Gold staining. 20%, PAGE, 250V, 1.5h.

4.1. Conclusion

To conclude, a restriction enzyme-based assay was developed to screen for potential inhibitors of hAGT. The inhibitor used to develop this assay was O^6-benzylguanine and it was determined that two Eqs of this inhibitor are necessary to fully inhibit hAGT. Initially, some reproducibility issues were observed when using radioactivity which necessitated a switch to SYBR gold staining as a means of visualization. Upon switching to this method of visualization, the assay was very reproducible and reliable. In the future, this assay could be used with other known inhibitors of hAGT like PaTrin-2. Furthermore, it could also be used to screen for other alkylated-guanine derivatives that could be inhibitors of hAGT. This method offers a quick and cost-effective method to screen for various inhibitors of hAGT however optimization to determine the optimal concentration of each test inhibitor would need to be performed.

Chapter 5 - General Conclusions and Future Directions

5.1-General Conclusions

To summarize, a method to synthesize circular DNA constructs by pre-organization of the system with a cross-link without the need for splints, was developed. A circular oligonucleotide containing a labile cross-link was synthesized with ligation to seal the circular structure confirmed by experiments including increased resistance to exonuclease 1 relative to the unligated precursor. Successful removal of the labile cross-link was confirmed by MS and circular dichroism revealed that removal of the cross-link from the circular oligonucleotide did not result in a significant deviation from a B-form global structure. In chapter 3, the introduction of a selenium atom at the 5' position of an oligonucleotide was achieved followed by subsequent selenization using a saturated solution of KSeCN. Circular dichroism spectra revealed minimal structure perturbances for these selenium modified nucleic acid duplexes when compared to the native control. Finally, the new restriction enzyme base assay to screen for potential hAGT inhibitors was explored. Reproducibility of results were improved when visualization of the oligonucleotides was performed by staining the polyacrylamide gels using SYBR Gold.

5.2-Future Directions

Future work for chapter 2 includes synthesizing circular constructs of various sizes and evaluation of the influence of a sulfone cross-link on their properties such as stability and structure. Furthermore, the method explored for the preparation of circular oligonucleotides only allows for the synthesis of symmetrical/palindromic sequences at the cross-linked site. Incorporating different protecting groups at the 3' and 5' positions of the phosphoramidite dimer would enable synthesis of a non-symmetrical circular oligonucleotide (**Figure 5.1**). This oligonucleotide could be constructed with a sulfone cross-link with the methodology described in chapter 2, with

appropriate terminal moieties for click or CNBr-assisted circularization and, upon cleavage of the cross-link, give rise to a non-symmetrical circular oligonucleotide product.

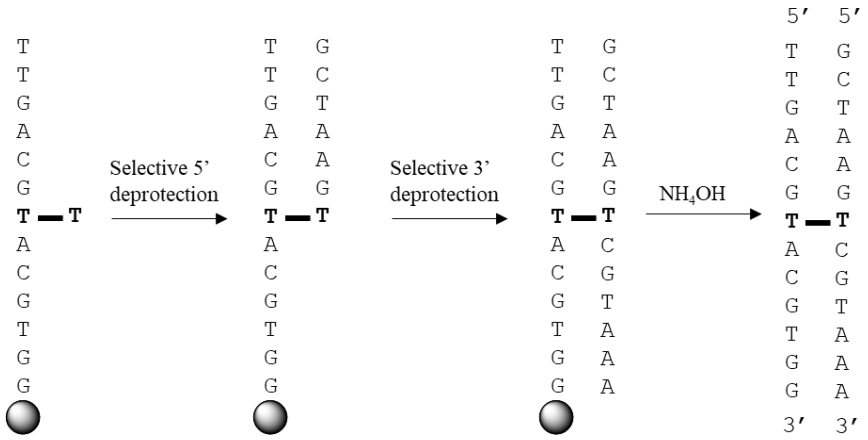


Figure 5.1-Synthesis of a non-symmetrical oligonucleotide using a sulfone cross-link.

Future directions for this synthesis of selenium modified nucleic acids containing the selenium group at the 5'-terminus would include investigating phosphorylating agents containing different alkyl groups to evaluate differences in selenium incorporation, structural effects on the oligonucleotide and ultimately attempting to crystallize and solve structures. With the assay developed to screen inhibitors of hAGT it would be of interest to evaluate other inhibitors to observe if similar trend of inhibition is observed.

It would be of interest to apply the methodology to prepare circular DNA construct containing the methyl phosphoroselenoate for structure determination. This could be accomplished by constructing a circular construct with a small overhang consisting of a protected hydroxyl group. With the appropriate protecting group, selective deprotection would liberate the hydroxyl group from which site-specific extension with a nucleoside phosphoramidite would result in a branched oligonucleotide.[99] Chemical phosphorylation followed by selenization would afford a circular construct with a selenium atom (**Figure 5.2**). This could in turn be used for X-ray

crystallization to determine the structure of the circular construct with and without the cross-links.

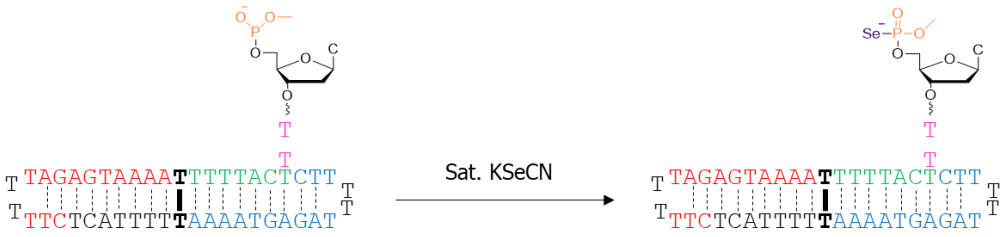


Figure 5.2-Proposed method to prepare a circular construct containing a 5'methylphosphoroselenoate.

Additionally, if the circular construct is recognized by hAGT, a crystal structure of this interaction could be obtained using the above-mentioned method. hAGT is known to remove alkyl cross-links from the O6-atom of 2'-deoxyguanosine. Hence it could be interesting to investigate whether it would be able to repair a sulfone cross-link engineered at the O6-position of a 2'-deoxyguanosine residue as another means to remove this cross-link. However, if hAGT does recognize the circular topology of this oligonucleotide but does not repair it, cross-linked circular DNA could be investigated as a potential inhibitor of hAGT.

References

(1) Dahm, R. *Hum Genet* **2008**, *122* (6), 565–581.

(2) Venter, J. C. et al. *Science* **2001**, *291* (5507), 1304–1351.

(3) Huang, M. et al. *J Chem Theory Comput* **2014**, *10* (4), 1538–1545.

(4) *Nucleic Acids in Chemistry and Biology*; 2006.

(5) Takahashi, S.; Sugimoto, N. *Acc. Chem. Res.* **2021**, *54* (9), 2110–2120.

(6) Drew, H. R. et al. *Proceedings of the National Academy of Sciences* **1981**, *78* (4), 2179–2183.

(7) Arnott, S. *Trends in Biochemical Sciences* **2006**, *31* (6), 349–354.

(8) Franklin, R. E.; Gosling, R. G. *Acta Crystallographica* **1953**, *6* (8–9), 673–677.

(9) D'Ascenzo, L. et al. *Nucleic Acids Res* **2016**, *44* (12), 5944–5956.

(10) Bochman, M. L. et al. *Nat Rev Genet* **2012**, *13* (11), 770–780.

(11) Day, H. A. et al. *Bioorganic & Medicinal Chemistry* **2014**, *22* (16), 4407–4418.

(12) Alberts, B. et al. *Molecular Biology of the Cell. 4th edition* **2002**.

(13) Cavaluzzi, M. J.; Borer, P. N. *Nucleic Acids Res* **2004**, *32* (1), e13.

(14) Lalan, M. et al. In *Challenges in Delivery of Therapeutic Genomics and Proteomics*; Misra, A., Ed.; Elsevier: London, 2011; pp 1–43.

(15) Pelley, J. W. In *Elsevier's Integrated Biochemistry*; Pelley, J. W., Ed.; Mosby: Philadelphia, 2007; pp 7–18.

(16) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22* (2), 256–263.

(17) Mergny, J.-L.; Lacroix, L. *Oligonucleotides* **2003**, *13* (6), 515–537.

(18) Kypr, J. et al. In *Comprehensive Chiroptical Spectroscopy*; Berova, N. et al., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2012; pp 575–586.

(19) Bishop, G. R.; Chaires, J. B. *Curr Protoc Nucleic Acid Chem* **2003**, *Chapter 7*, 7.11.1-7.11.8.

(20) Kypr, J. et al. *Nucleic Acids Res* **2009**, *37* (6), 1713–1725.

(21) Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25* (1), 1–28.

(22) Igarashi, J. et al. *Future Rare Diseases* **2022**, *2* (1), FRD19.

(23) Shen, T. et al. *Theranostics* **2022**, *12* (1), 35–47.

(24) Akhtar, S. In *Delivery Strategies for Antisense Oligonucleotide Therapeutics*; CRC Press, 1995.

(25) Kool, E. T. *Acc. Chem. Res.* **1998**, *31* (8), 502–510.

(26) Mohsen, M. G.; Kool, E. T. *Acc. Chem. Res.* **2016**, *49* (11), 2540–2550.

(27) Pedroso, E. et al. In *Oligonucleotide Synthesis*; Humana Press: New Jersey, 2004; Vol. 288, pp 101–126.

(28) Li, J. et al. *ChemBioChem* **2020**, *21* (11), 1547–1566.

(29) Petkovic, S.; Müller, S. *Nucleic Acids Res* **2015**, *43* (4), 2454–2465.

(30) Shuman, S. *Journal of Biological Chemistry* **2009**, *284* (26), 17365–17369.

(31) Taylor, M. R. et al. *Journal of Biological Chemistry* **2011**, *286* (26), 23054–23062.

(32) Dolinnaya, N. G. et al. *Nucl Acids Res* **1993**, *21* (23), 5403–5407.

(33) Fedorova, O. A. et al. *Nucleosides and Nucleotides* **1996**, *15* (6), 1137–1147.

(34) Sokolova, N. I. et al. *FEBS Letters* **1988**, *232* (1), 153–155.

(35) An, R. et al. *PLOS ONE* **2014**, *9* (12), e115950.

(36) Kolb, H. C. et al. *Angewandte Chemie International Edition* **2001**, *40* (11), 2004–2021.

(37) Amblard, F. et al. *Chem. Rev.* **2009**, *109* (9), 4207–4220.

(38) Fantoni, N. Z. et al. *Chem. Rev.* **2021**, *121* (12), 7122–7154.

(39) Kukwikila, M. et al. *Nat Chem* **2017**, *9* (11), 1089–1098.

(40) Sun, G. et al. *Current Protocols in Nucleic Acid Chemistry* **2012**, *51* (1), 5.11.1-5.11.17.

(41) *SYBR Safe - DNA Gel Stain - CA*. https://www.thermofisher.com/ca/en/home/life-science/dna-rna-purification-analysis/nucleic-acid-gel-electrophoresis/dna-stains/sybr-safe.html (accessed 2022-07-11).

(42) Borer, P. N. et al. *Journal of Molecular Biology* **1974**, *86* (4), 843–853.

(43) Puglisi, J. D.; Tinoco, I. In *Methods in Enzymology*; RNA Processing Part A: General Methods; Academic Press, 1989; Vol. 180, pp 304–325.

(44) Gololobov, Yu. G. et al. *Tetrahedron* **1981**, *37* (3), 437–472.

(45) Schweigert, N. et al. *Environ Mol Mutagen* **2000**, *36* (1), 5–12.

(46) Uzair, B. et al. *Protein Pept Lett* **2018**, *25* (7), 612–618.

(47) Baraniak, D.; Boryski, J. *Biomedicines* **2021**, *9* (6), 628.

(48) Tikhomirova, A. et al. *Biochemistry* **2006**, *45* (35), 10563–10571.

(49) Lin, L. et al. *Chem. Soc. Rev.* **2011**, *40* (9), 4591–4602.

(50) Zhang, W. et al. *Front. Chem. Sci. Eng.* **2016**, *10* (2), 196–202.

(51) Hu, B. et al. *Sig Transduct Target Ther* **2020**, *5* (1), 1–25.

(52) Shirley, M. *Drugs* **2021**, *81* (7), 875–879.

(53) Ochoa, S.; Milam, V. T. *Molecules* **2020**, *25* (20), 4659.

(54) Mansoor, M.; Melendez, A. J. *Gene Regul Syst Bio* **2008**, *2*, 275–295.

(55) Khvorova, A.; Watts, J. K. *Nat Biotechnol* **2017**, *35* (3), 238–248.

(56) Kupryushkin, M. S. et al. *Molecular Therapy - Nucleic Acids* **2022**, *27*, 211–226.

(57) Quemener, A. M. et al. *Molecules* **2022**, *27* (2), 536.

(58) Lin, L. et al. *Chem. Soc. Rev.* **2011**, *40* (9), 4591.

(59) Zhang, W. et al. *Front. Chem. Sci. Eng.* **2016**, *10* (2), 196–202.

(60) Smyth, M. S.; Martin, J. H. J. *Mol Pathol* **2000**, *53* (1), 8–14.

(61) Hauptman, H. *Current Opinion in Structural Biology* **1997**, *7* (5), 672–680.

(62) Liu, Y. et al. *Proceedings of the National Academy of Sciences* **2001**, *98* (19), 10648–10653.

(63) Ealick, S. E. *Current Opinion in Chemical Biology* **2000**, *4* (5), 495–499.

(64) Jiang, J. et al. *Nucleic Acids Res* **2007**, *35* (2), 477–485.

(65) Dibrov, S. et al. *Acta Crystallographica Section D* **2011**, *67* (2), 97–104.

(66) Tavale, S. S.; Sobell, H. M. *Journal of Molecular Biology* **1970**, *48* (1), 109–123.

(67) Ennifar, E. et al. *Acta Cryst D* **2002**, *58* (8), 1262–1268.

(68) Wilds, C. J. et al. *J. Am. Chem. Soc.* **2002**, *124* (50), 14910–14916.

(69) Qian, L. et al. *ChemMedChem* **2006**, *1* (3), 376–383.

(70) Moiseeva, T. N.; Bakkenist, C. J. *DNA Repair (Amst)* **2018**, *72*, 99–106.

(71) *DNA Repair of Cancer Stem Cells*; Mathews, L. A. et al., Eds.; Springer Netherlands: Dordrecht, 2013.

(72) Burkhalter, M. D. et al. *Ageing Res Rev* **2015**, *23*, 29–36.

(73) Matt, S.; Hofmann, T. G. *Cell Mol Life Sci* **2016**, *73* (15), 2829–2850.

(74) In *LiverTox: Clinical and Research Information on Drug-Induced Liver Injury*; National Institute of Diabetes and Digestive and Kidney Diseases: Bethesda (MD), 2012.

(75) Kondo, N. et al. *Journal of Nucleic Acids* **2010**, *2010*, e543531.

(76) Klaunig, J. E.; Kamendulis, L. M. In *Comprehensive Toxicology (Second Edition)*; McQueen, C. A., Ed.; Elsevier: Oxford, 2010; pp 117–138.

(77) Beranek, D. T. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **1990**, *231* (1), 11–30.

(78) Stevens, M. F. G. In *Cancer Drug Design and Discovery (Second Edition)*; Neidle, S., Ed.; Academic Press: San Diego, 2014; pp 145–164.

(79) Serpe, M. et al. *Sci Rep* **2019**, *9* (1), 6889.

(80) *The EMBO Journal* **2000**, *19* (7), 1719–1730.

(81) Daniels, D. S. et al. *Nat Struct Mol Biol* **2004**, *11* (8), 714–720.

(82) Kanugula, S.; Pegg, A. E. *Biochem J* **2003**, *375* (Pt 2), 449–455.

(83) O'Flaherty, D. K.; Wilds, C. J. *Chemistry–A European Journal* **2015**, *21* (29), 10522–10529.

(84) Pegg, A. E. *Mutation Research/Reviews in Mutation Research* **2000**, *462* (2), 83–100.

(85) Dolan, M. E.; Pegg, A. E. *Clinical Cancer Research* **1997**, *3* (6), 837–847.

(86) Yu, W. et al. *Frontiers in Oncology* **2020**, *9*.

(87) Turriziani, M. et al. *Pharmacological Research* **2006**, *53* (4), 317–323.

(88) Christmann, M.; Kaina, B. In *DNA Repair in Cancer Therapy (Second Edition)*; Kelley, M. R., Fishel, M. L., Eds.; Academic Press: Boston, 2016; pp 55–82.

(89) Fan, C.-H. et al. *Cell Death Dis* **2013**, *4* (10), e876–e876.

(90) Rabik, C. A. et al. *Cancer Treatment Reviews* **2006**, *32* (4), 261–276.

(91) *Definition of myelosuppression - NCI Dictionary of Cancer Terms - NCI*. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/myelosuppression (accessed 2022-06-02).

(92) *Low blood cell counts: Side effect of cancer treatment*. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/cancer/in-depth/cancer-treatment/art-20046192 (accessed 2022-06-02).

(93) McManus, F. P. et al. *Org. Biomol. Chem.* **2010**, *8* (19), 4414–4426.

(94) Spratt, T. E.; Campbell, C. R. *Biochemistry* **1994**, *33* (37), 11364–11371.

(95) Paredes, E. et al. In *Comprehensive Medicinal Chemistry III*; Chackalamannil, S. et al., Eds.; Elsevier: Oxford, 2017; pp 233–279.

(96) *PAGE & HPLC*. Integrated DNA Technologies. https://www.idtdna.com/pages/products/optional-services/page-hplc-purification (accessed 2022-07-11).

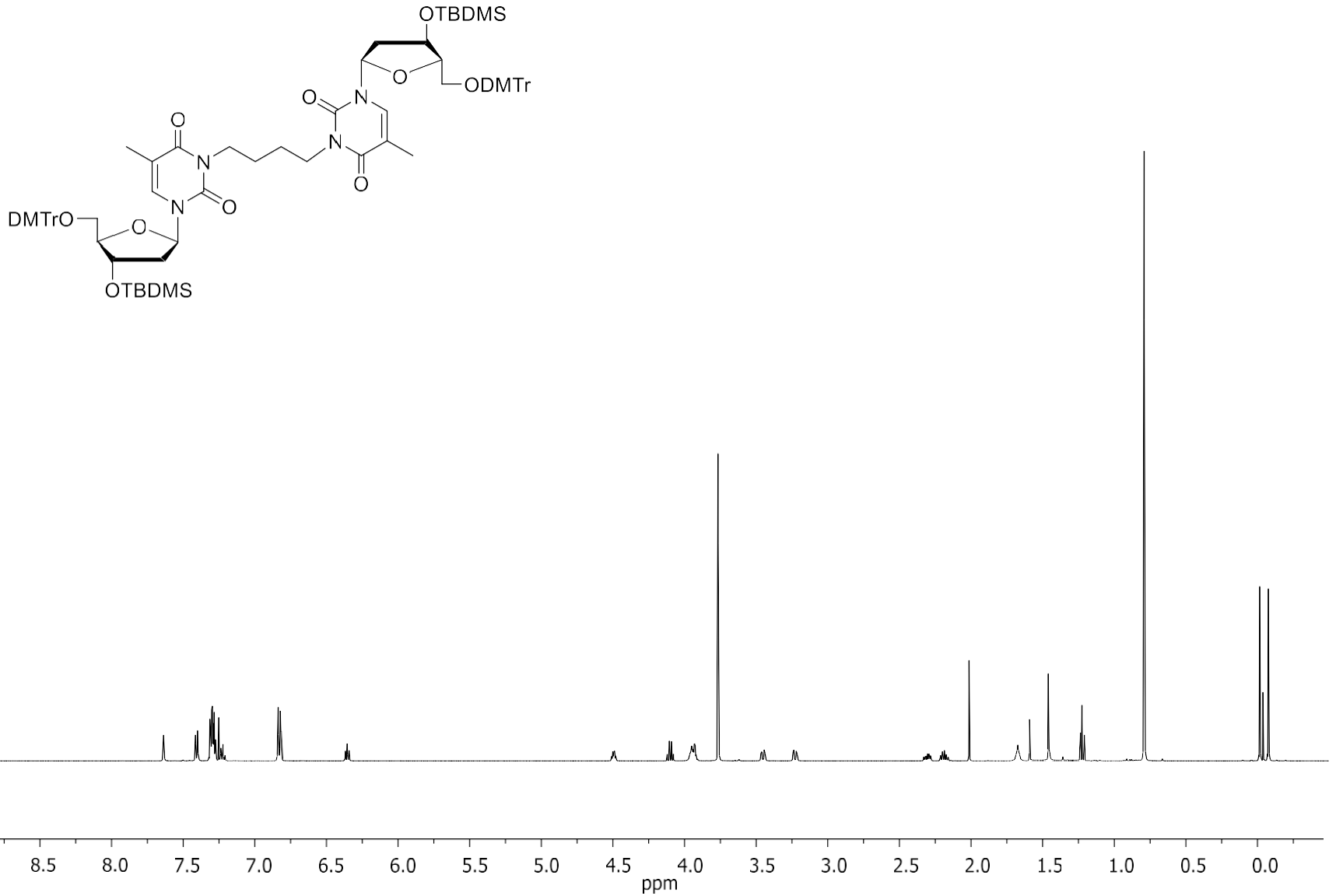
(97) Beharry, A. A. et al. *PLOS ONE* **2016**, *11* (4), e0152684.

(98) Guillen, D. et al. *PLoS One* **2020**, *15* (3), e0229527.

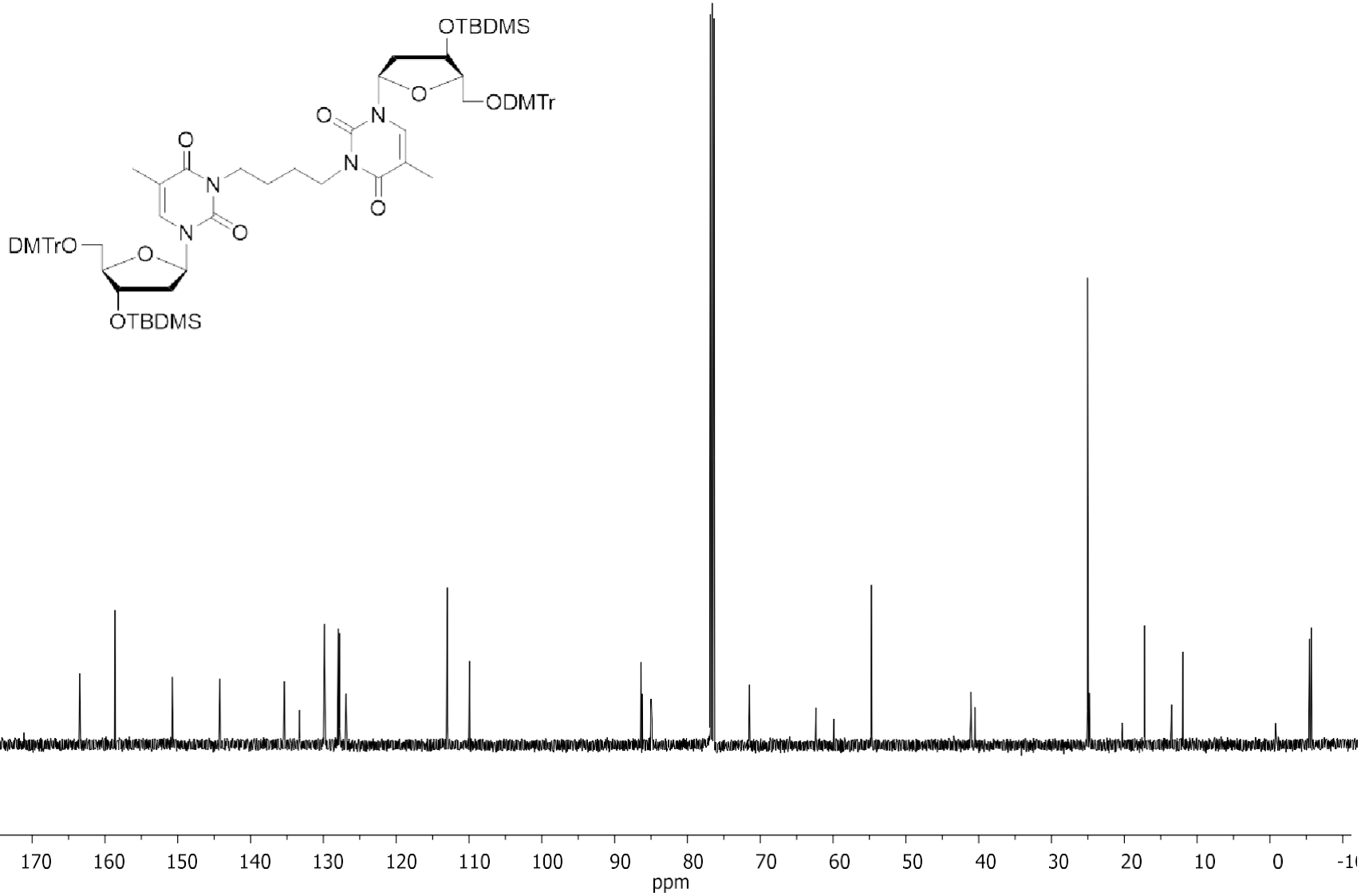
(99) Copp, W.; Wilds, C. J. *Chemistry* **2020**, *26* (65), 14802–14806.

Appendix I: Supporting Information for Chapter 2

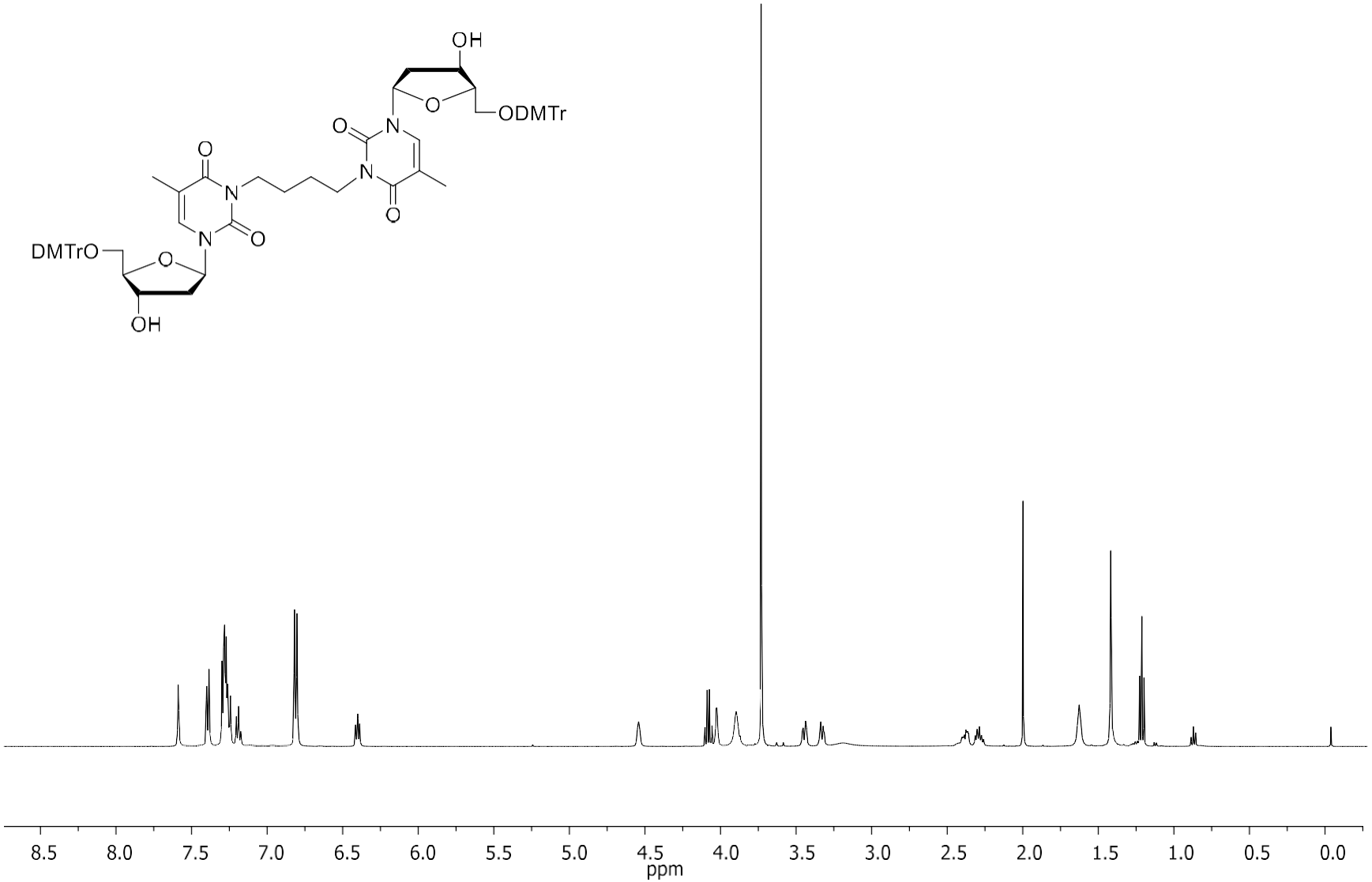
Supporting Figure A1.1: 500 MHz ^{1}H NMR spectrum of compound 5 in $CDCl_3$.



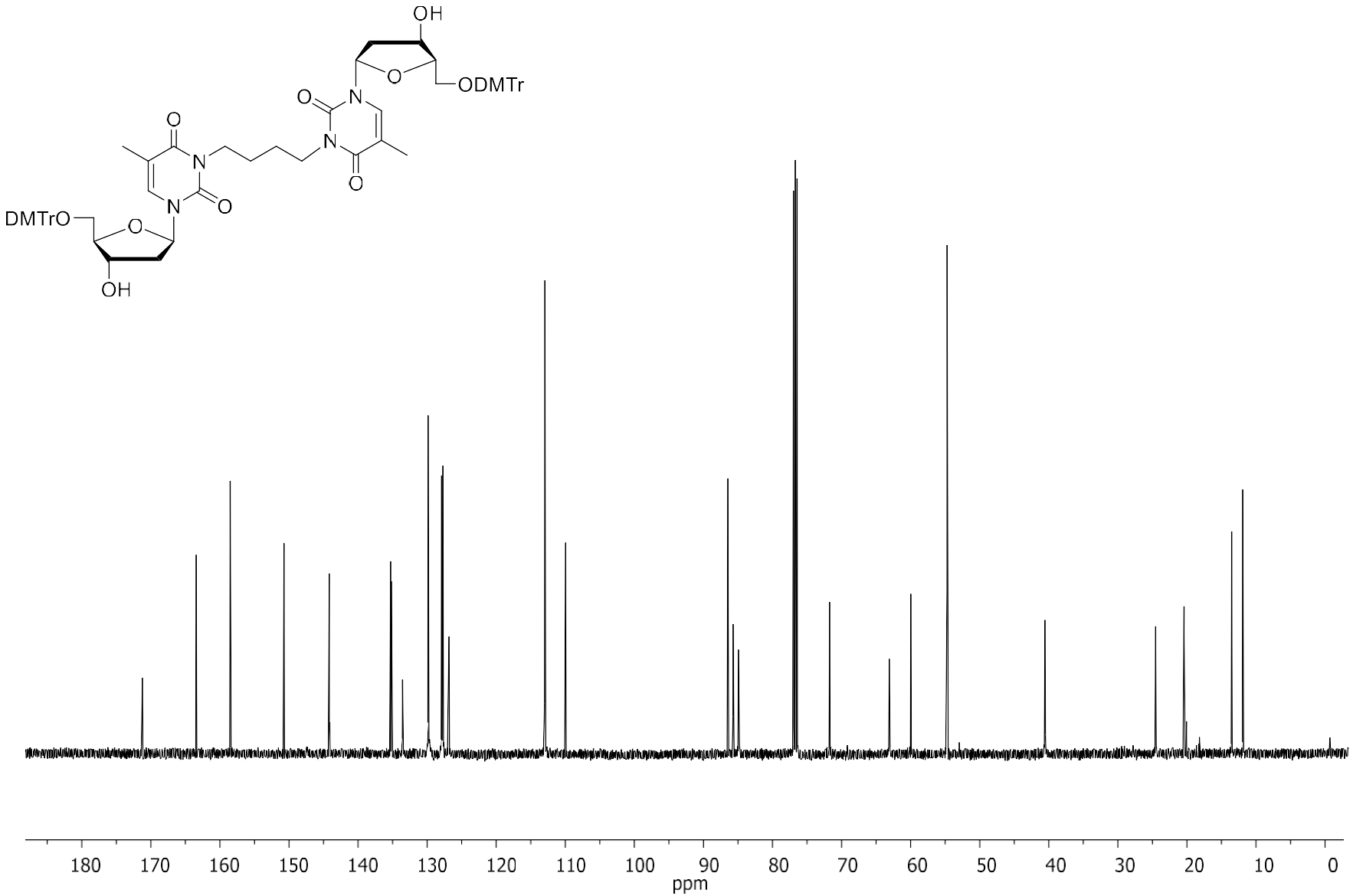
Supporting Figure A1.2: 125.7 MHz ^{13}C NMR spectrum of compound 5 in $CDCl_3$.



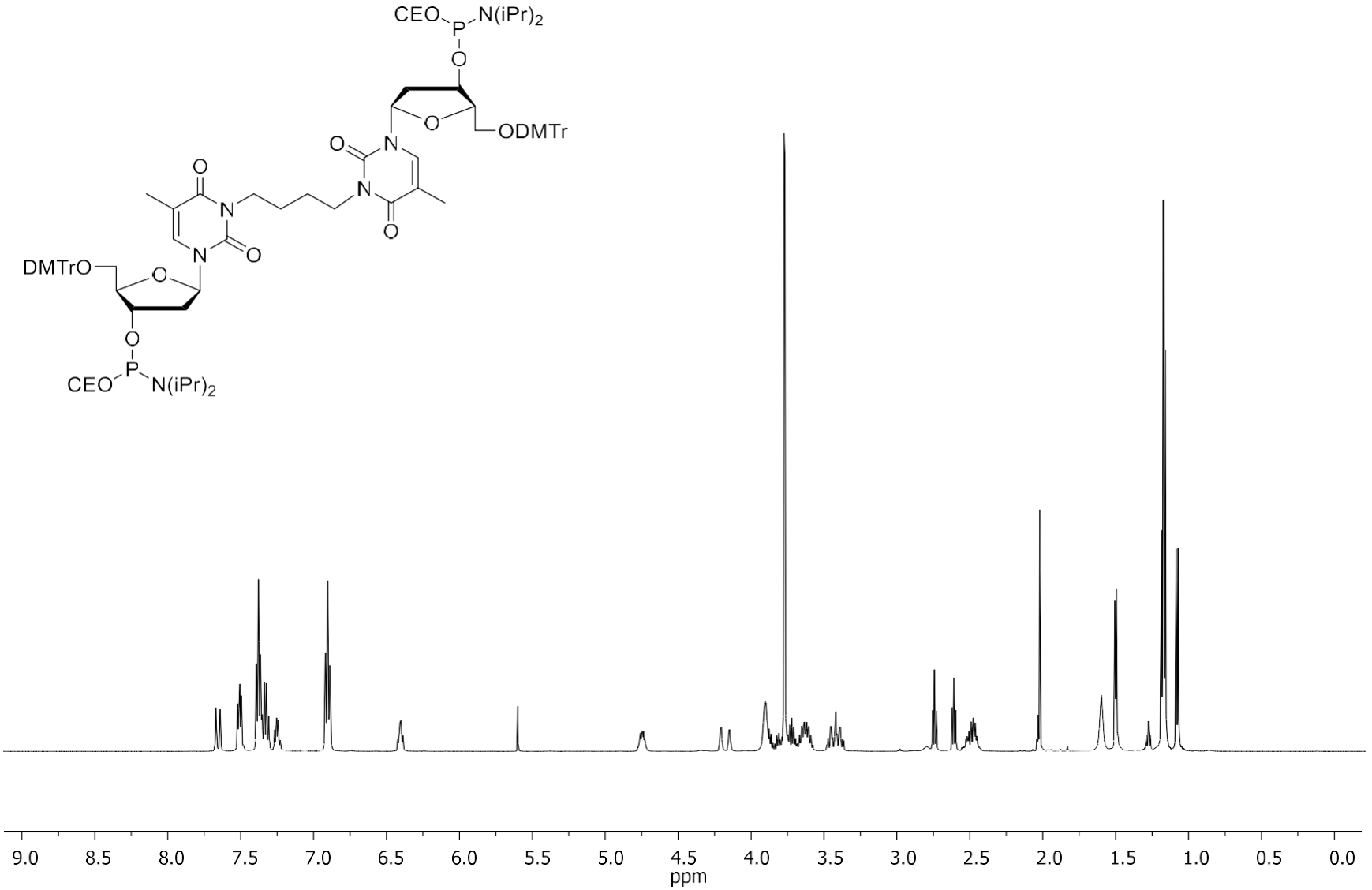
Supporting Figure A1.3: 500 MHz ^{1}H NMR spectrum of compound 6 in $CDCl_3$.



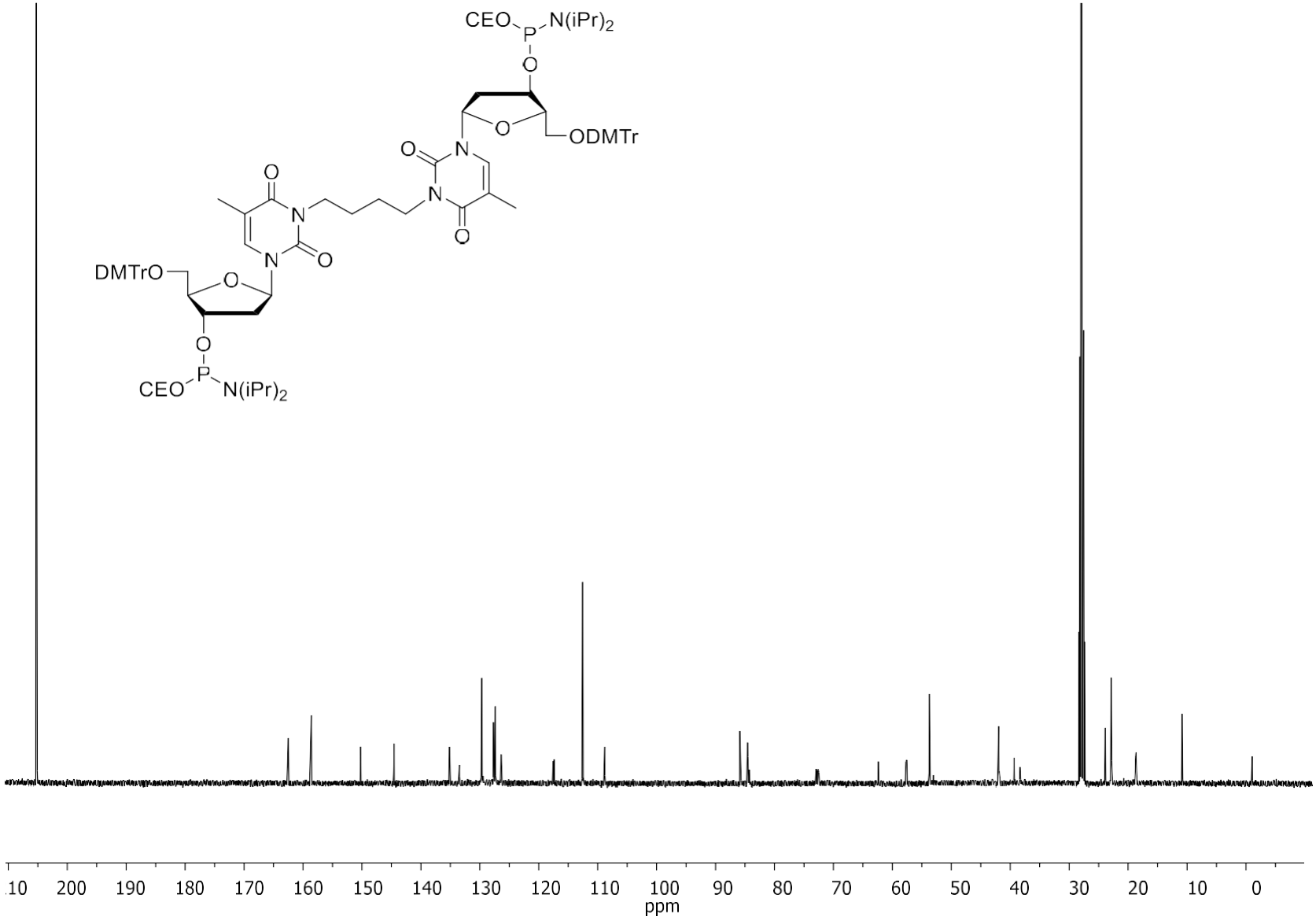
Supporting Figure A1.4: 125.7 MHz ^{13}C NMR spectrum of compound 6 in $CDCl_3$.



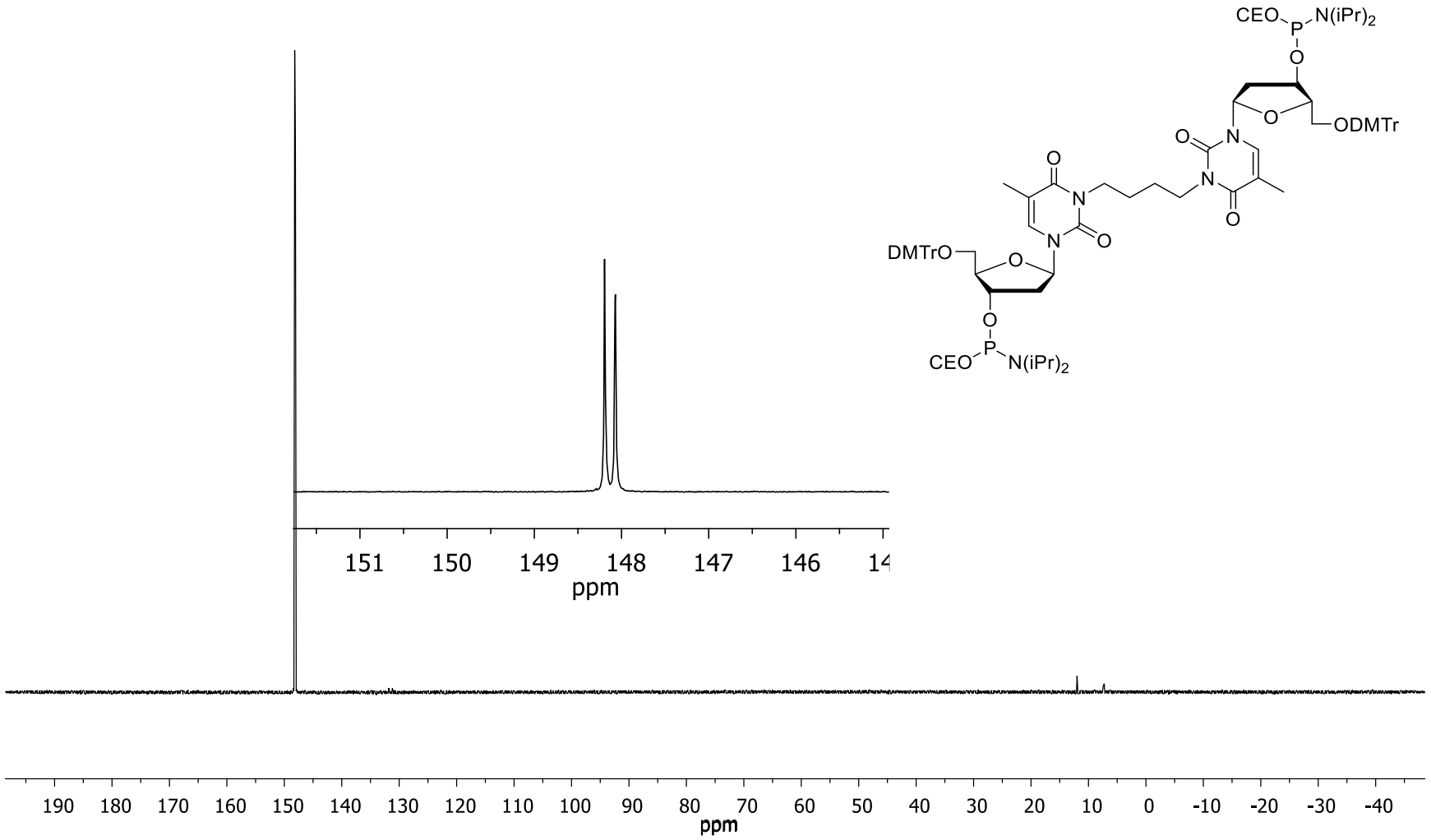
Supporting Figure A1.5: 500 MHz ^{1}H NMR spectrum of compound 7 in acetone-d6.



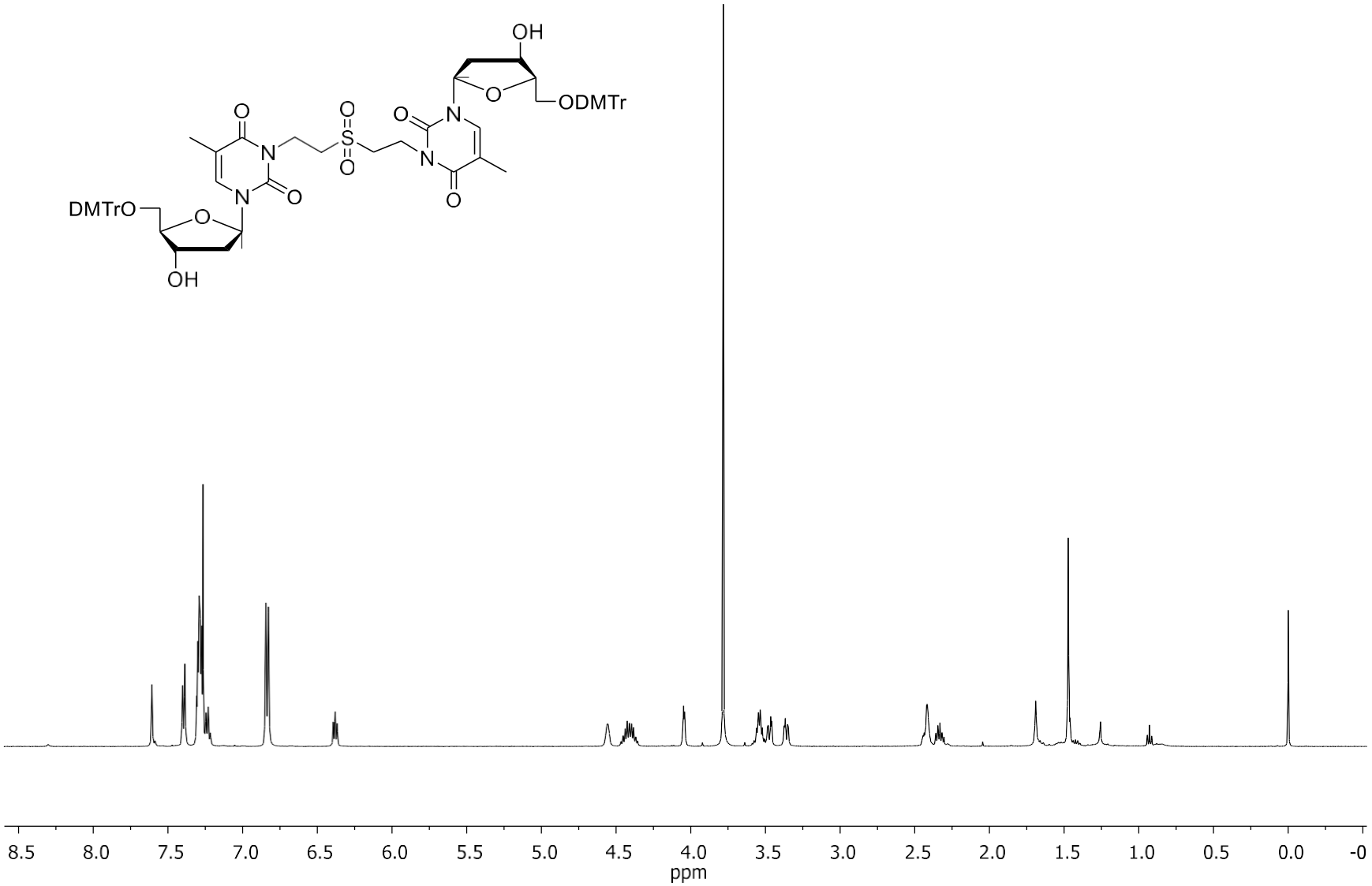
Supporting Figure A1.6: 125.7 MHz ^{13}C NMR spectrum of compound 7 in acetone-d6.



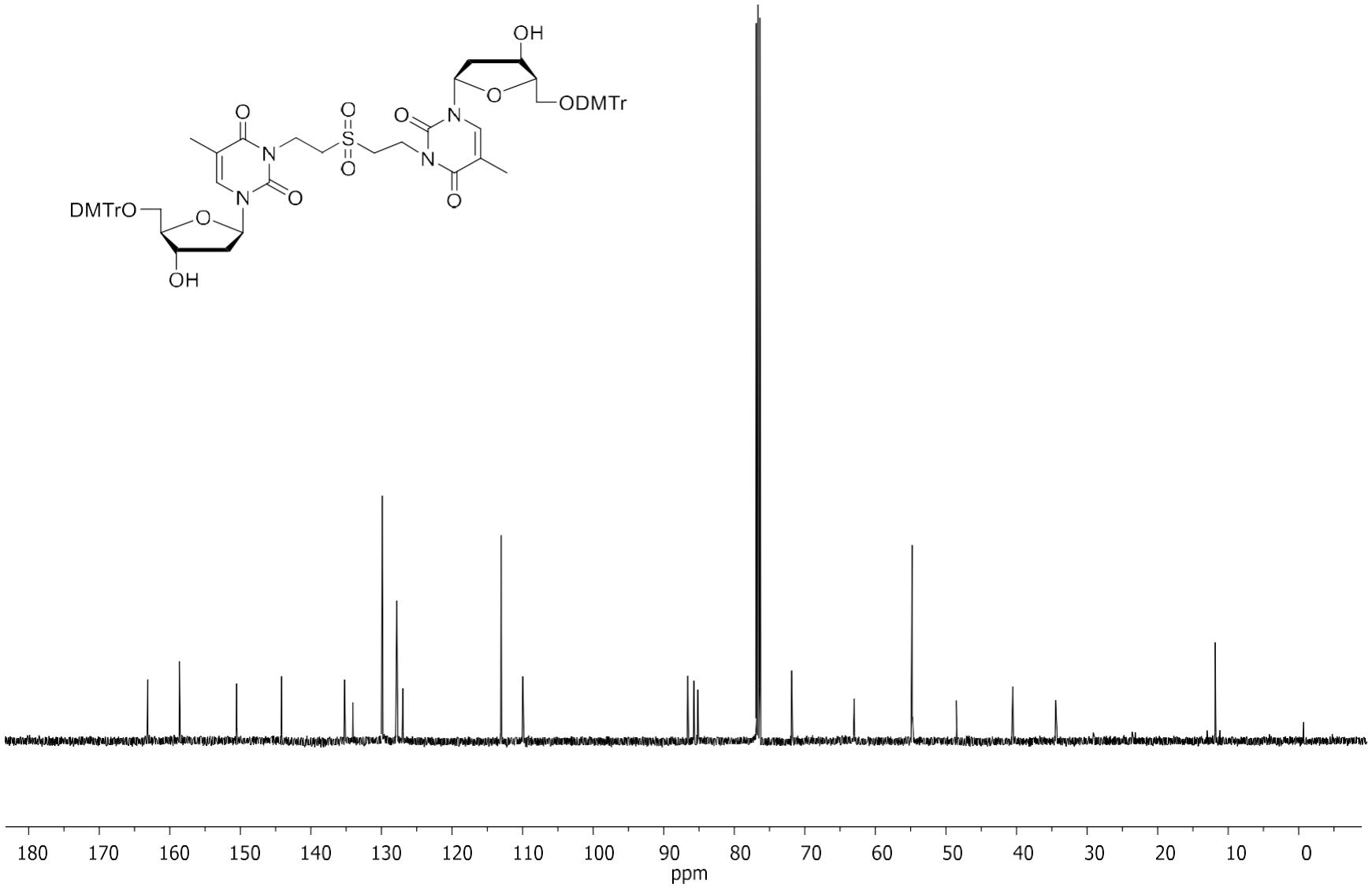
Supporting Figure A1.7: 202.3 MHz ^{31}P NMR spectrum of compound 7 in acetone-d6.



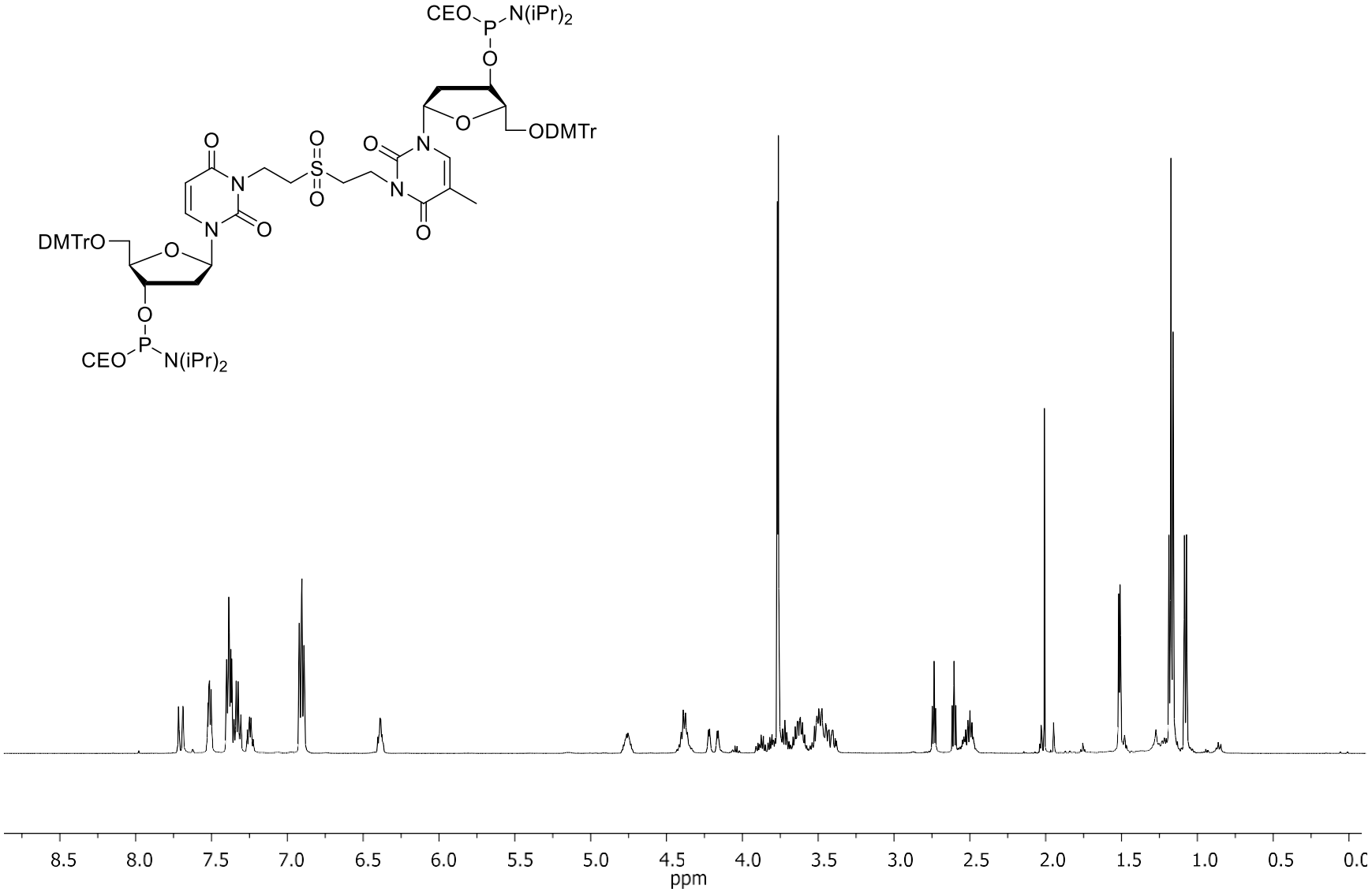
Supporting Figure A1.8: 500 MHz ^{1}H NMR spectrum of compound 8 in $CDCl_3$.



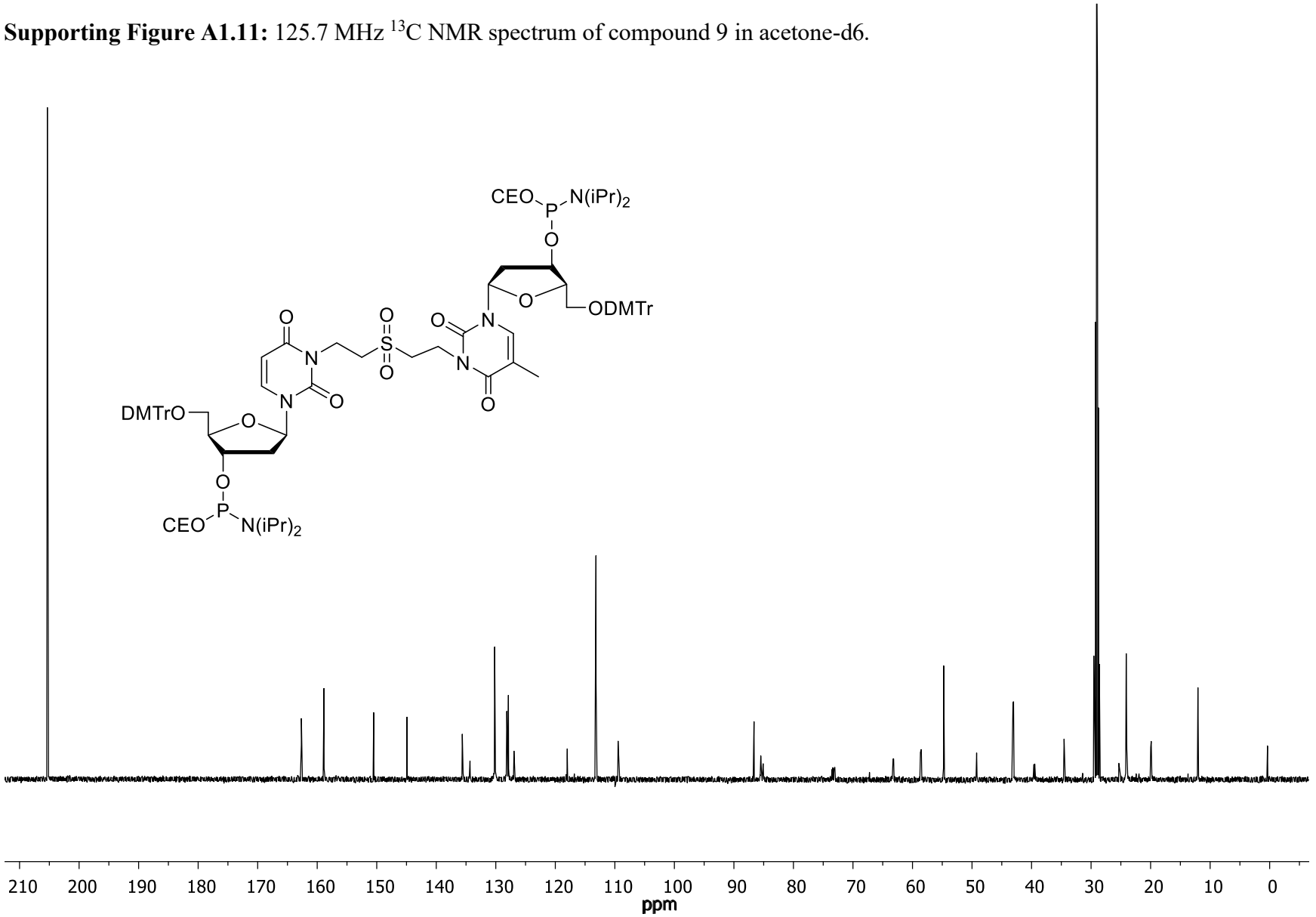
Supporting Figure A1.9: 125.7 MHz ^{13}C NMR spectrum of compound 8 in $CDCl_3$.



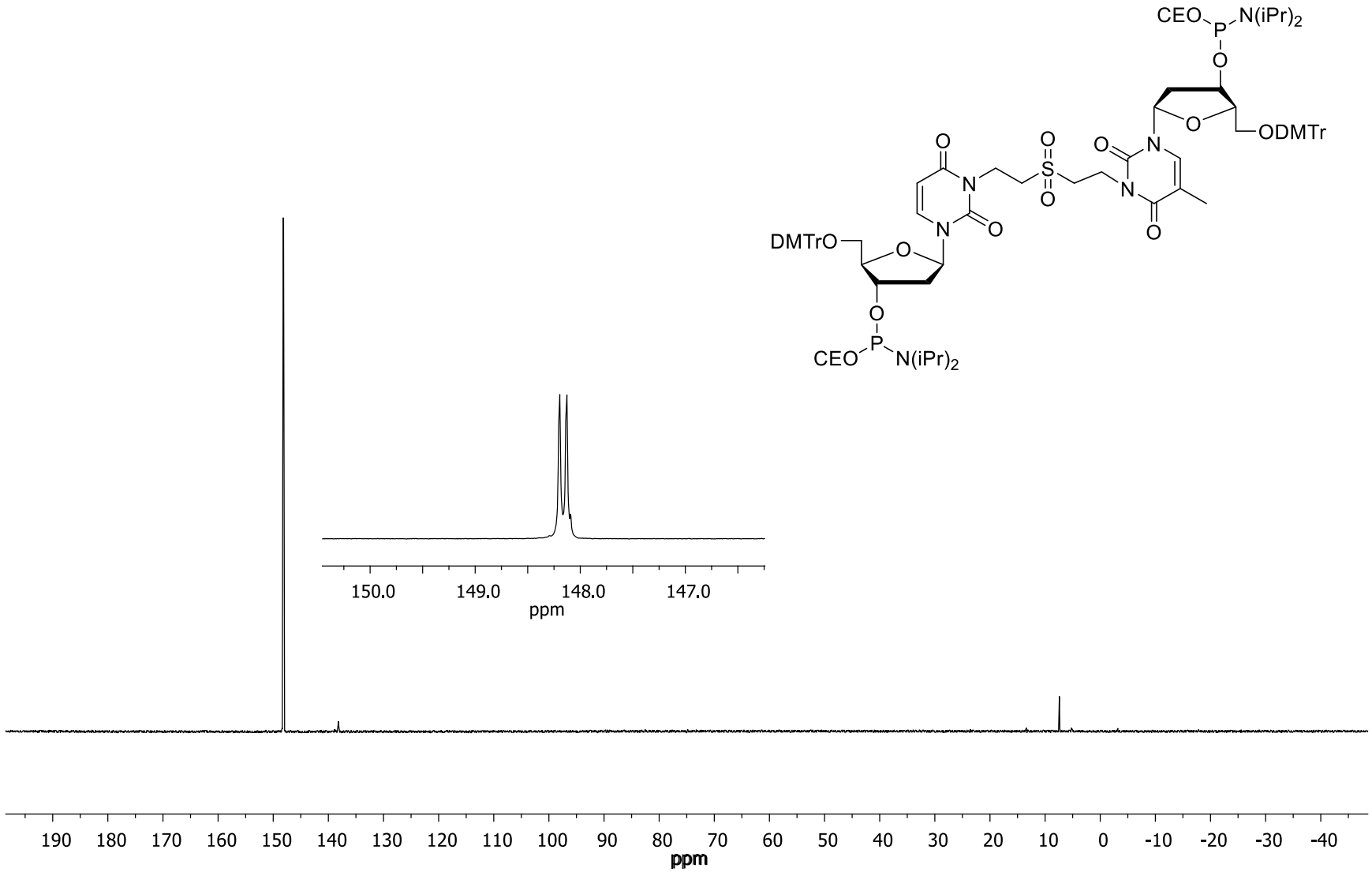
Supporting Figure A1.10: 500 MHz ^{1}H NMR spectrum of compound 9 in acetone-d6.



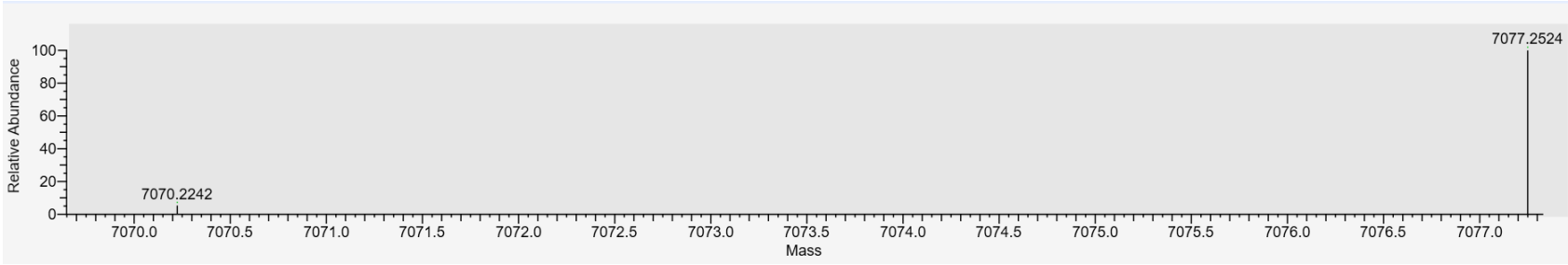
Supporting Figure A1.11: 125.7 MHz ^{13}C NMR spectrum of compound 9 in acetone-d6.



Supporting Figure A1.12: 202.3 MHz ^{31}P NMR spectrum of compound 9 in acetone-d6.

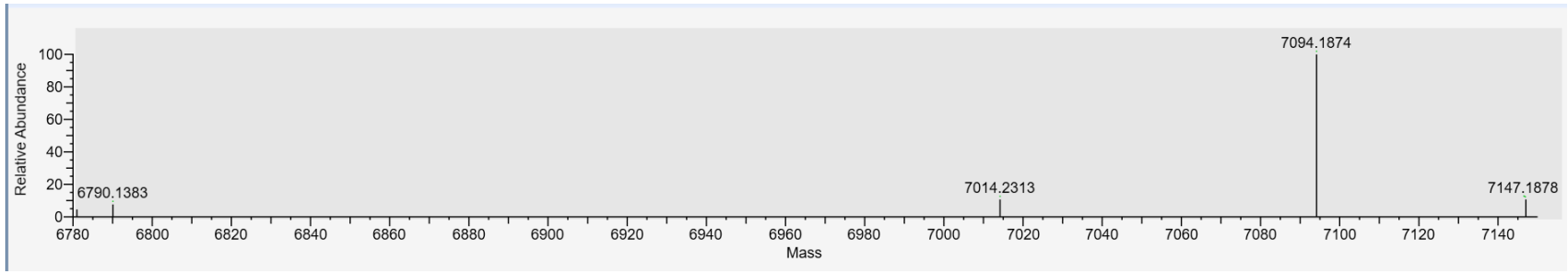


Supporting Figure A1.13: ESI-MS Spectrum of CJW 978.



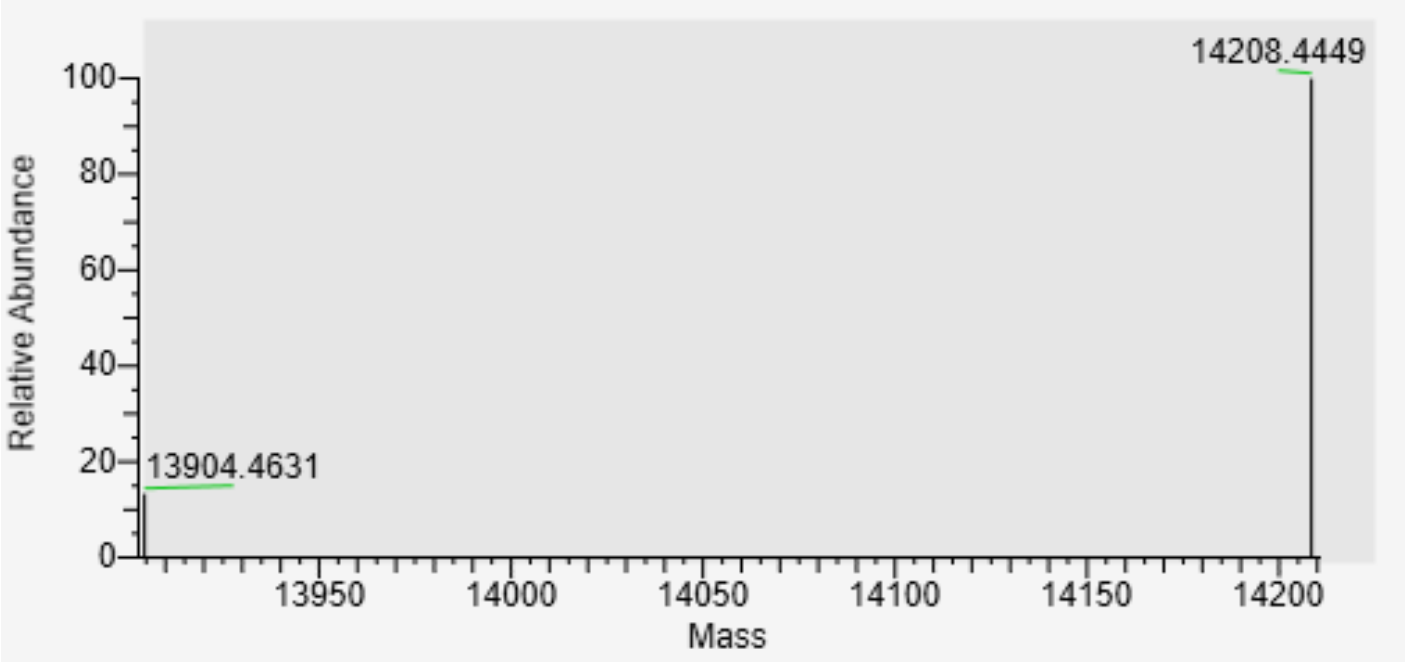
Expected: 7077.2 Obtained: 7077.3

Supporting Figure A1.14: ESI-MS Spectrum of CJW 997.



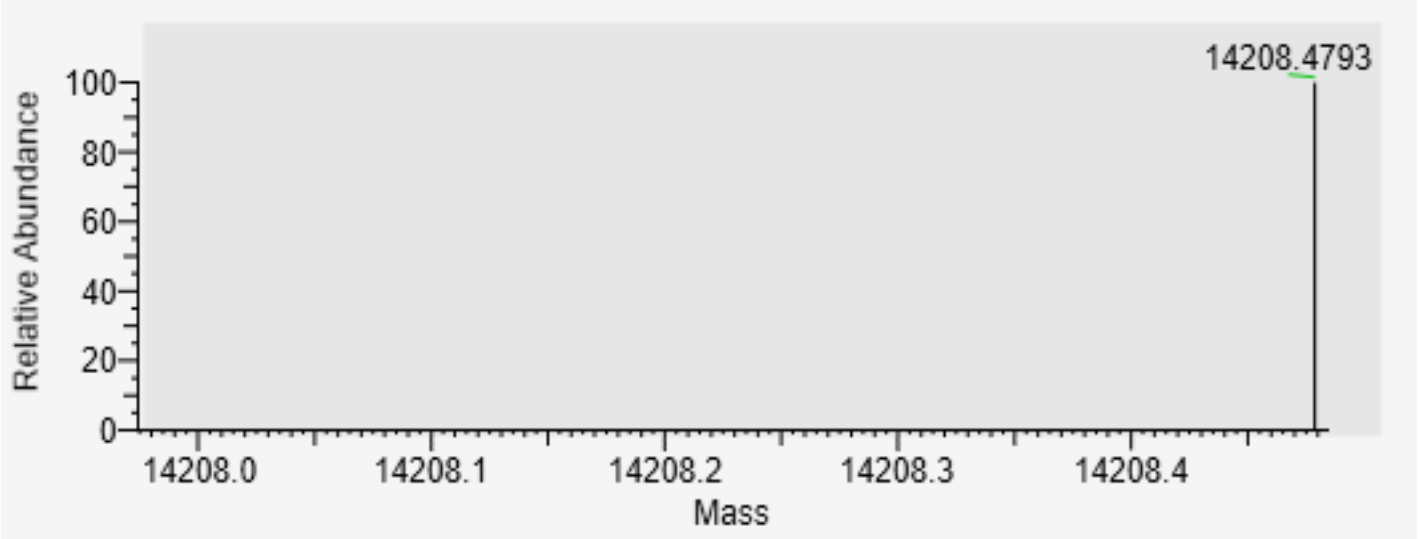
Expected: 7094.2 Obtained: 7094.2

Supporting Figure A1.15: ESI-MS Spectrum of CJW 1035.



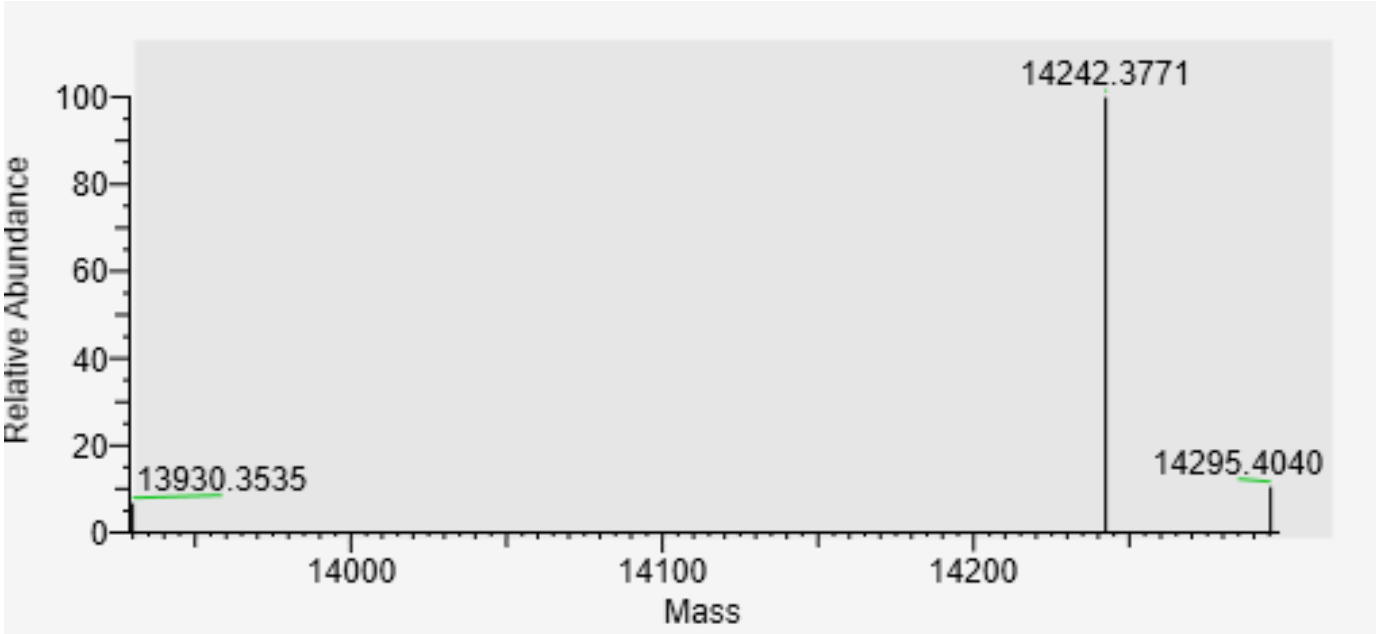
Expected: 14 208.4 Obtained: 14 208.4

Supporting Figure A1.16: ESI-MS Spectrum of CJW 1035 (Circularized).



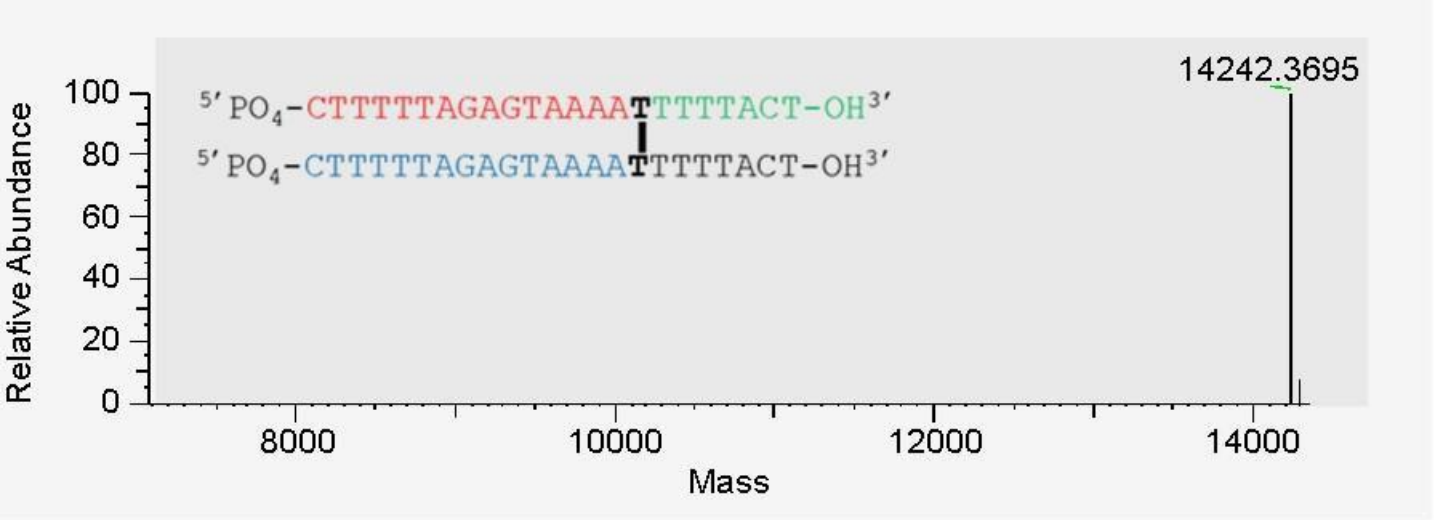
Expected: 14 208.5 Obtained: 14 208.5

Supporting Figure A1.17: ESI-MS Spectrum of CJW 1036.



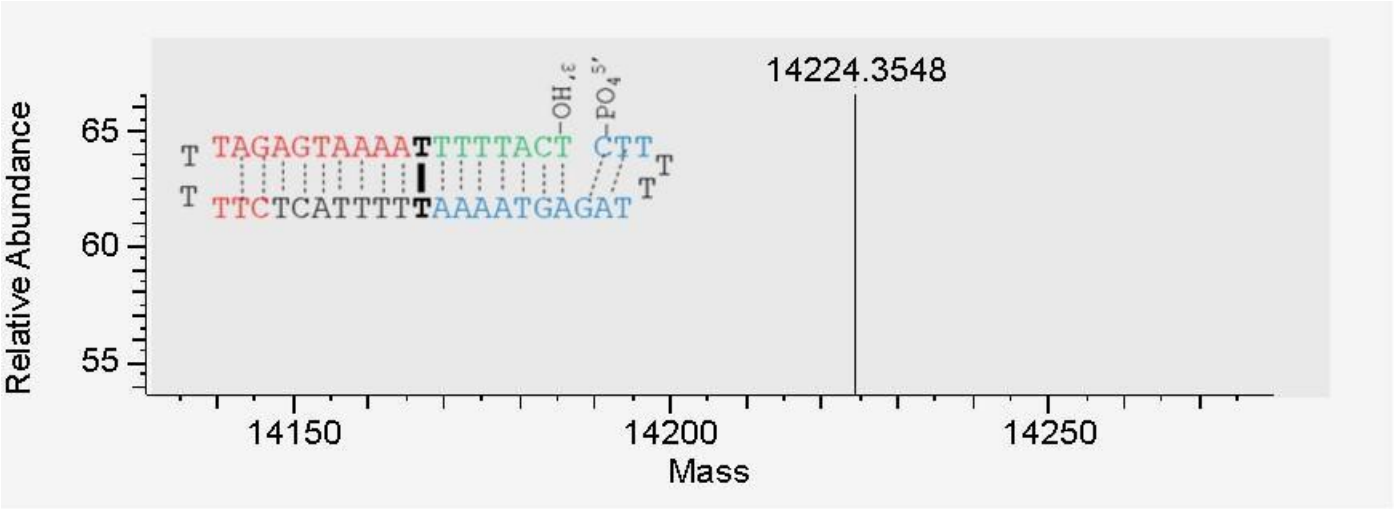
Expected: 14 242.4 Obtained: 14 242.4

Supporting Figure A1.18: ESI-MS Spectrum of CJW 1036 (CNBr ligation product, top band).



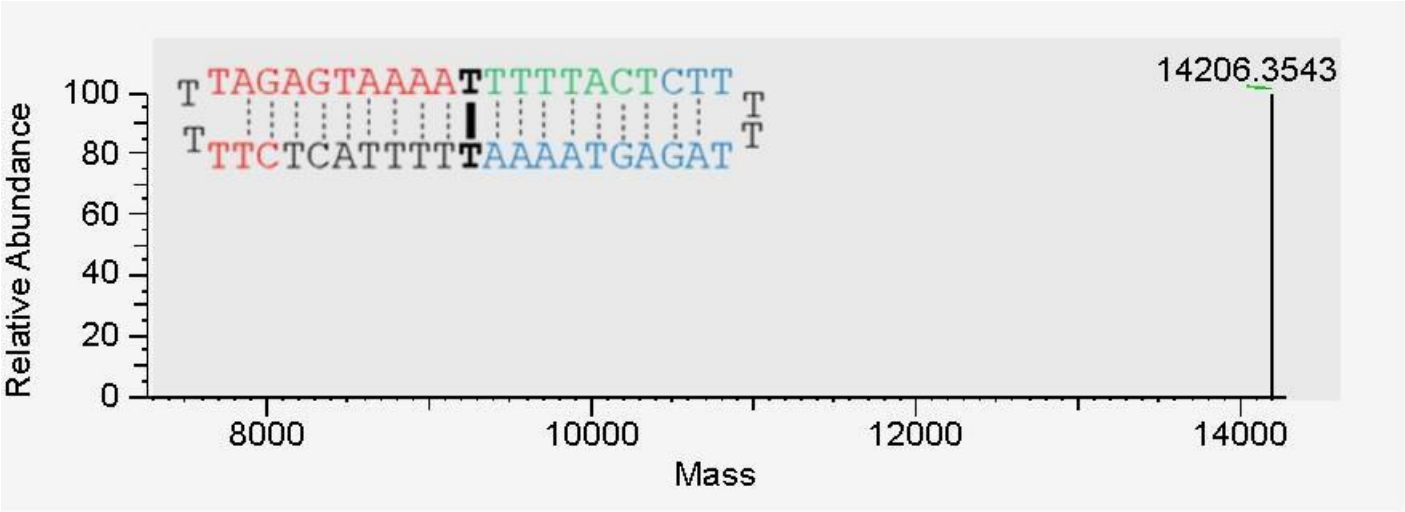
Expected: 14 206.4 Obtained: 14 242.4

Supporting Figure A1.19: ESI-MS Spectrum of CJW 1036 (CNBr ligation product, middle band).



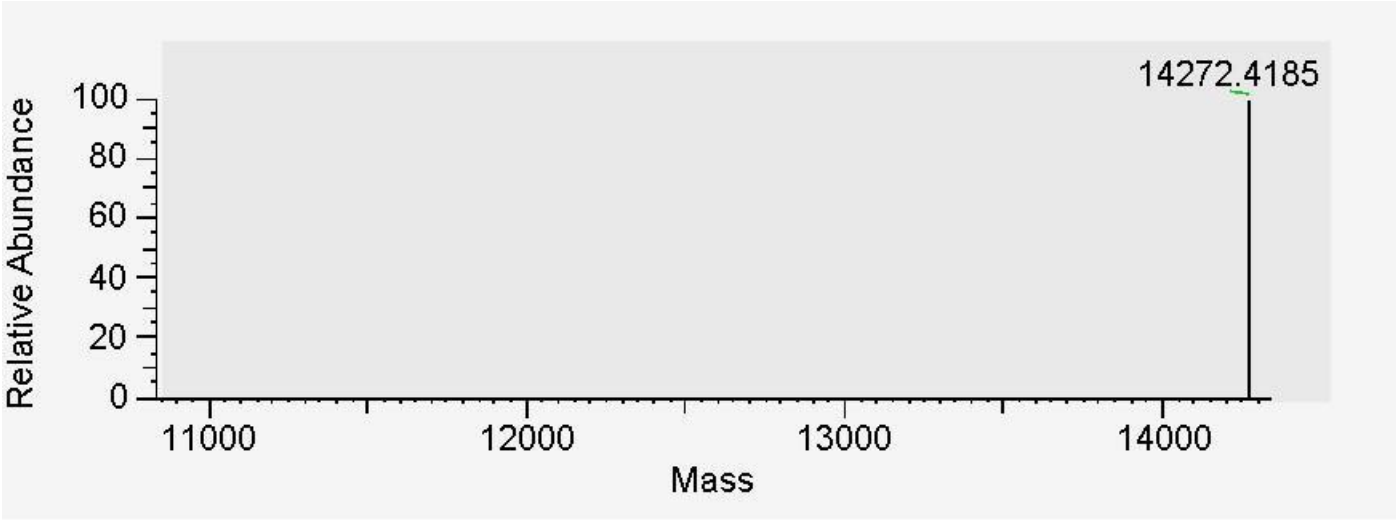
Expected: 14 206.4 Obtained: 14 224.4

Supporting Figure A1.20: ESI-MS Spectrum of CJW 1036 (Circularized, CNBr ligation product, bottom band).



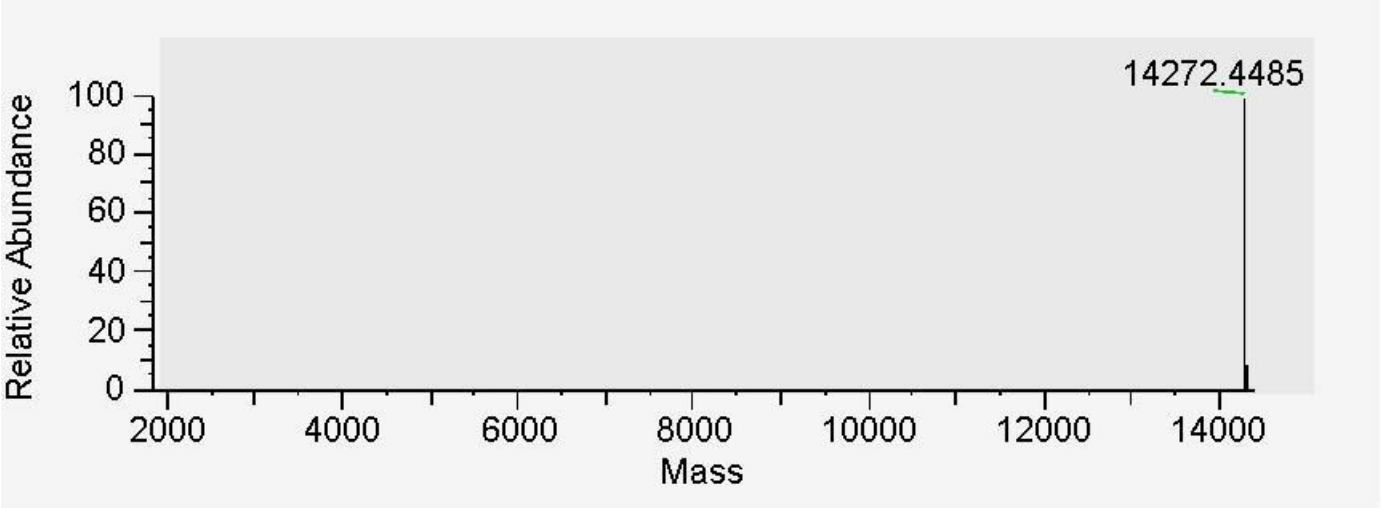
Expected: 14 206.4 Obtained: 14 206.4

Supporting Figure A1.21: ESI-MS Spectrum of CJW 1238.



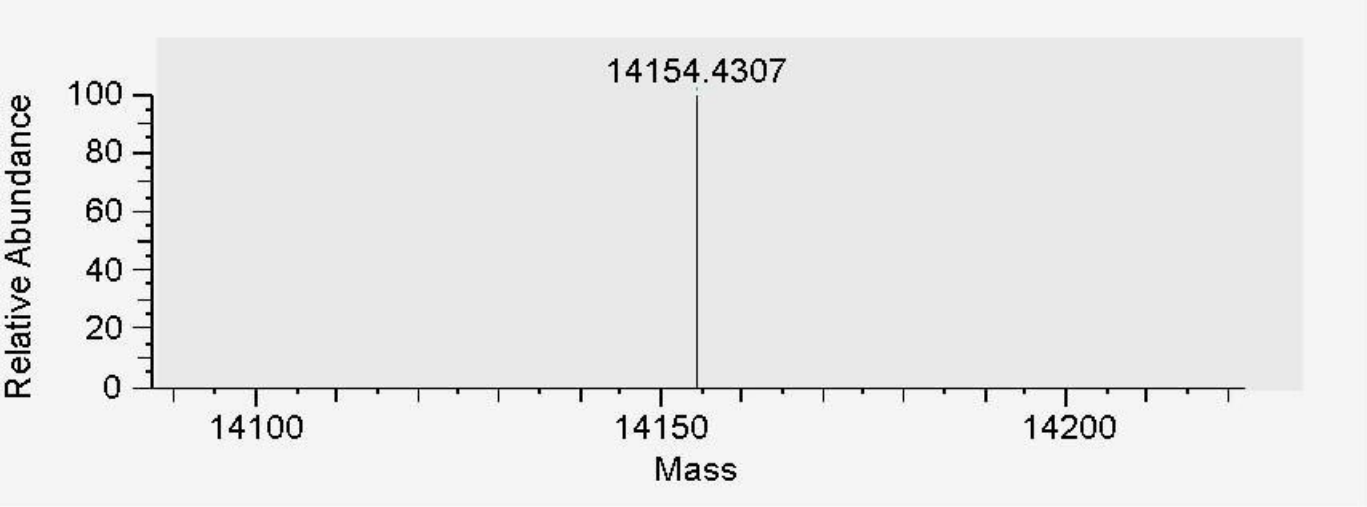
Expected: 14 272.4 Obtained: 14 172.4

Supporting Figure A1.22 ESI-MS Spectrum of CJW 1238 (Circularized).



Expected: 14 272.4 Obtained: 14 172.4

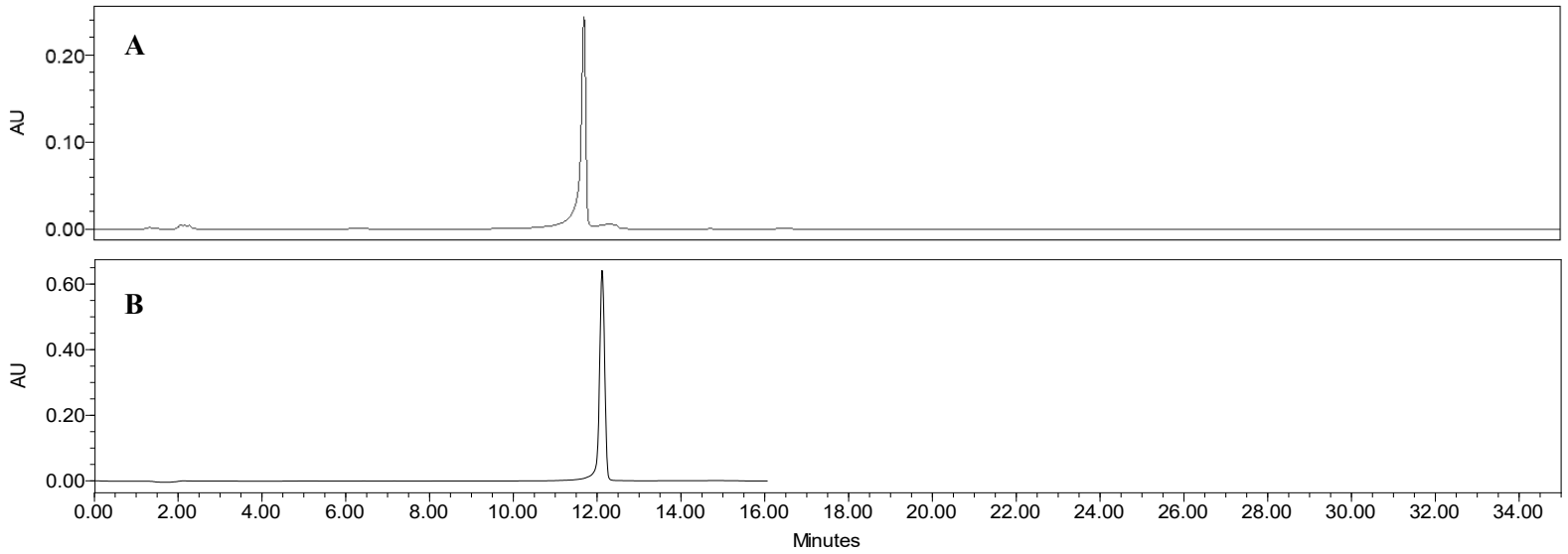
Supporting Figure A1.23: ESI-MS Spectrum of CJW 1238 (Circularized, cross-link removed).



Expected: 14 154.3 Obtained: 14 154.4

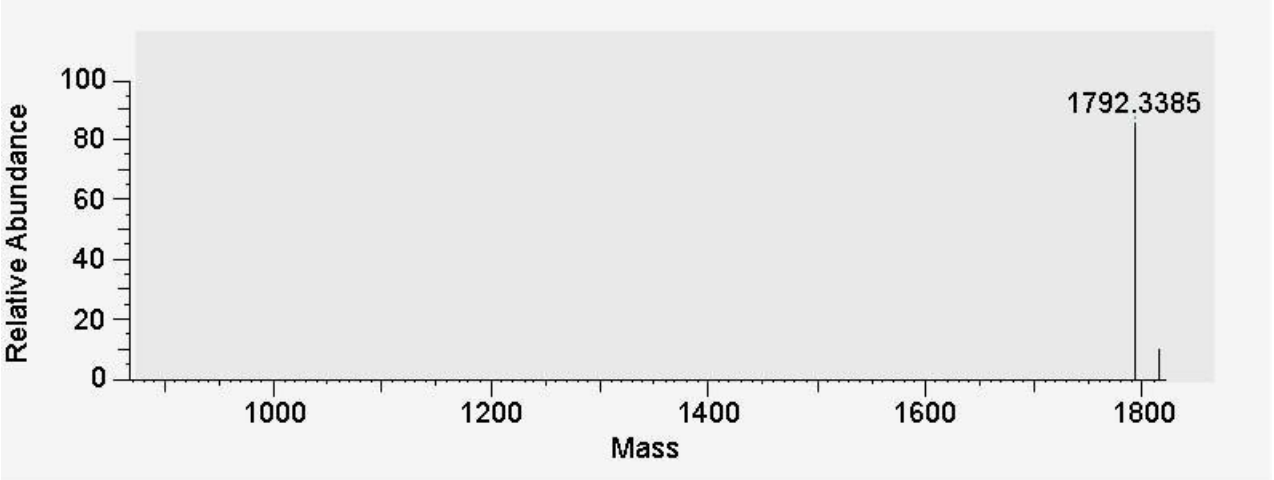
Appendix II: Supporting Information for Chapter 3

Supporting Figure A2.1: IEX-HPLC chromatograph of (A) crude and (B) pure 6mer OH.



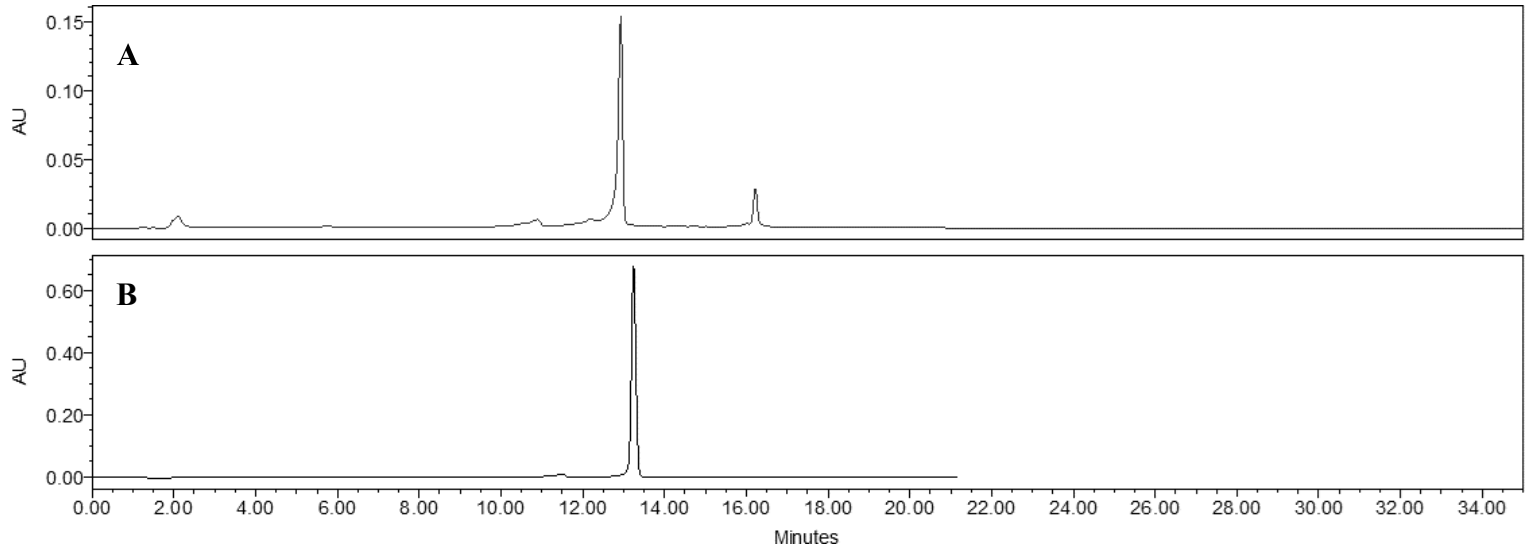
Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A2.2: ESI-MS Spectrum of 6mer OH.



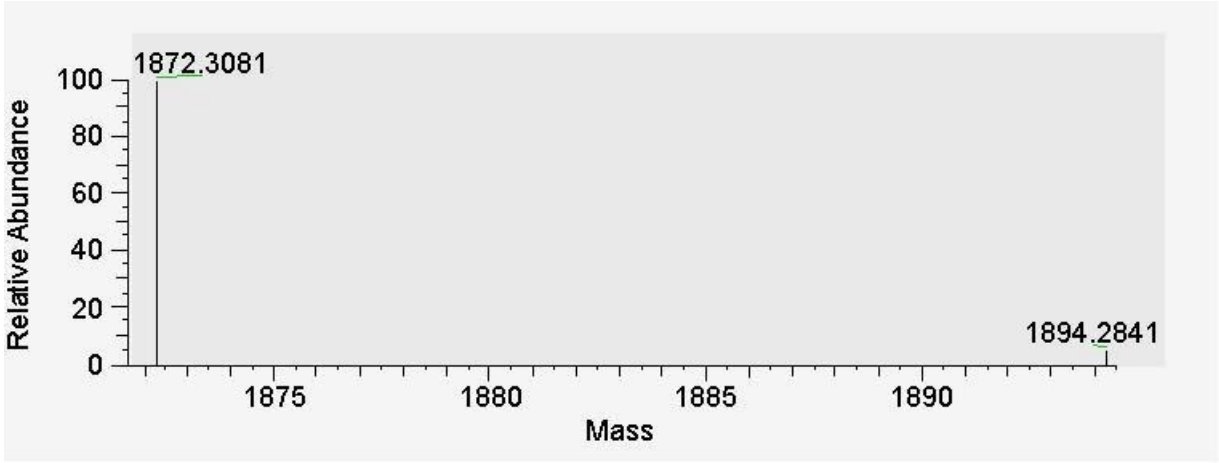
Expected: 1792.3 Obtained: 1792.3

Supporting Figure A2.3: IEX-HPLC chromatograph of (A) crude and (B) pure 6mer PO.



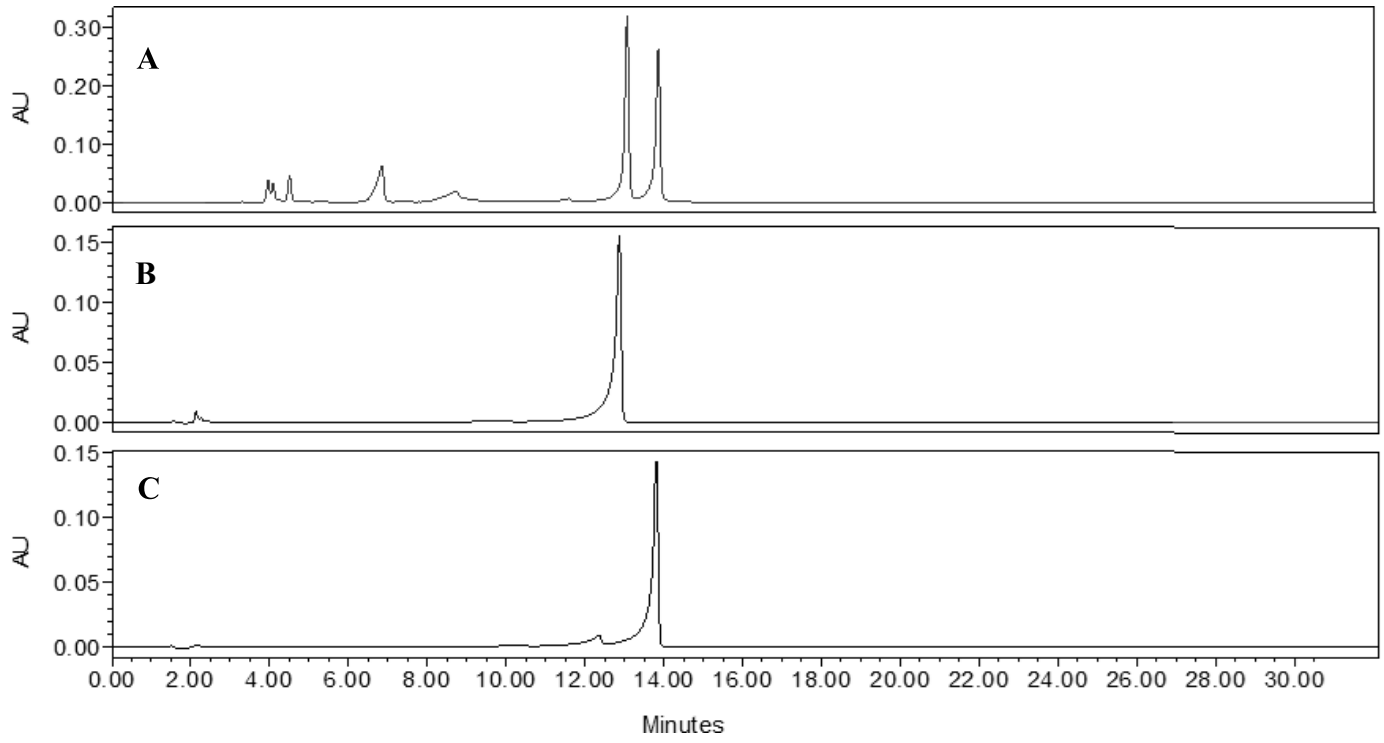
Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A2.4: ESI-MS Spectrum of 6mer PO.



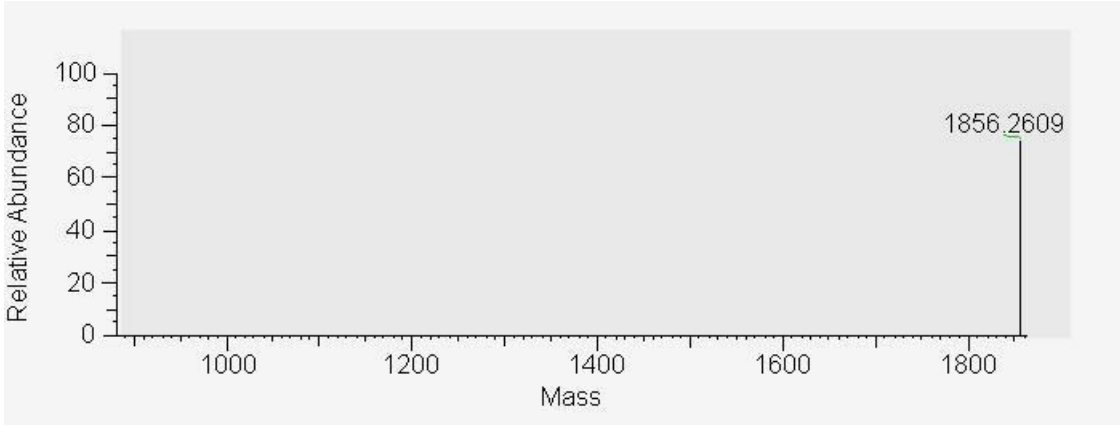
Expected: 1872.3 Obtained: 1872.3

Supporting Figure A2.5: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure 6mer PSeI Peak 1 (**C**) pure 6mer PSeI Peak 2.



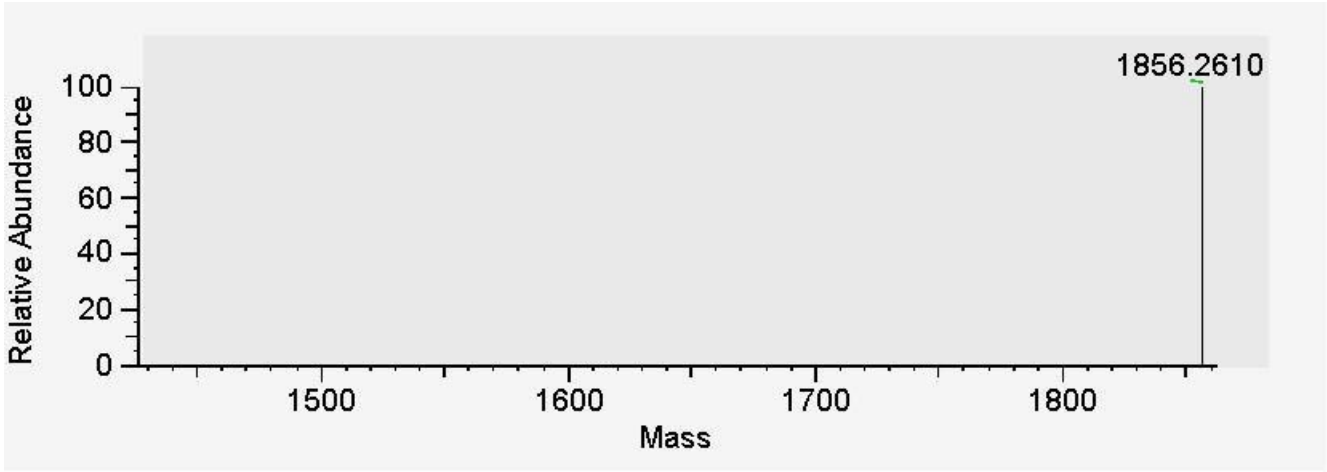
Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A2.6: ESI-MS Spectrum of 6mer PSeI (peak 1).



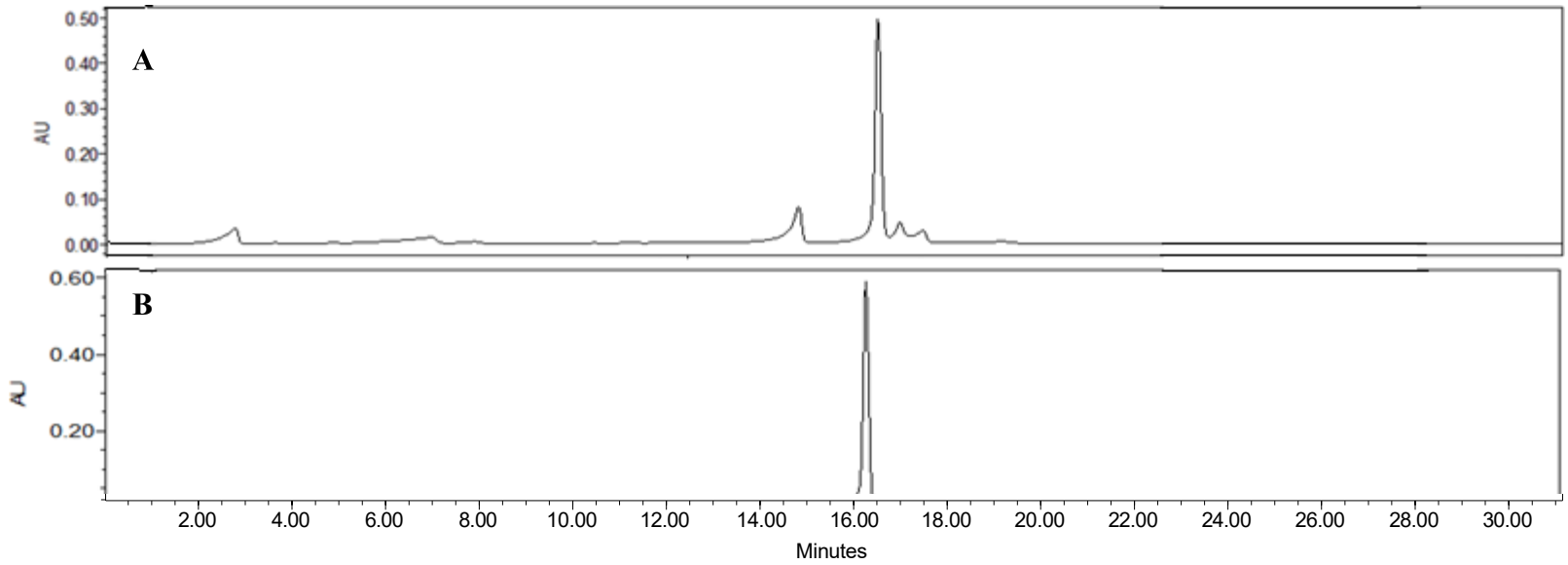
Expected: 1858.3 Obtained: 1856.3

Supporting Figure A2.7: ESI-MS Spectrum of 6mer PSeI (peak 2).



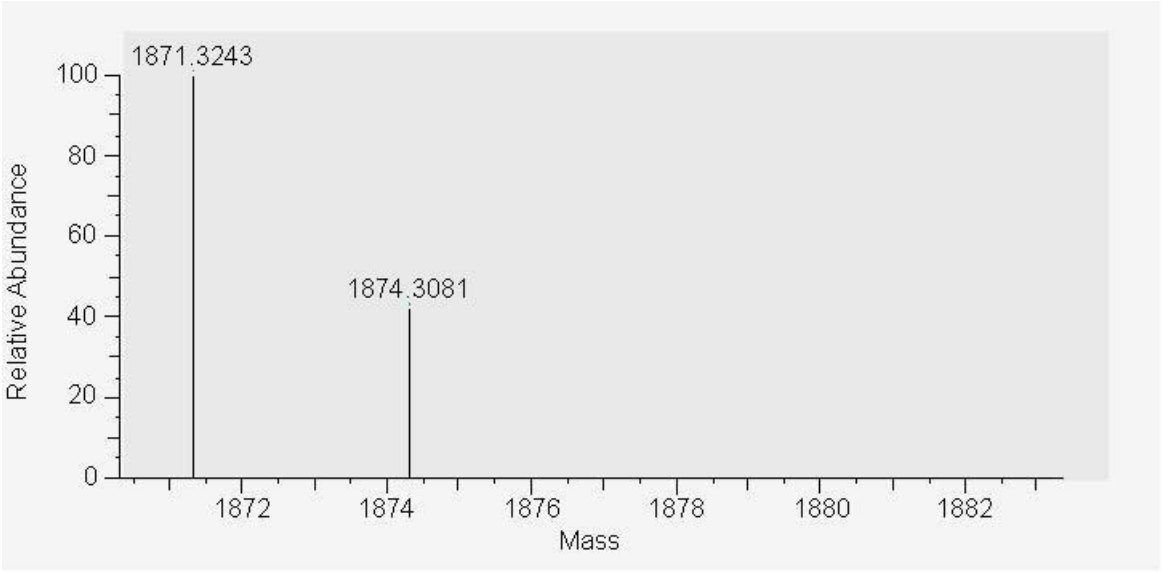
Expected: 1856.3 Obtained: 1856.3

Supporting Figure A2.8: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure 6mer PSe.



Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

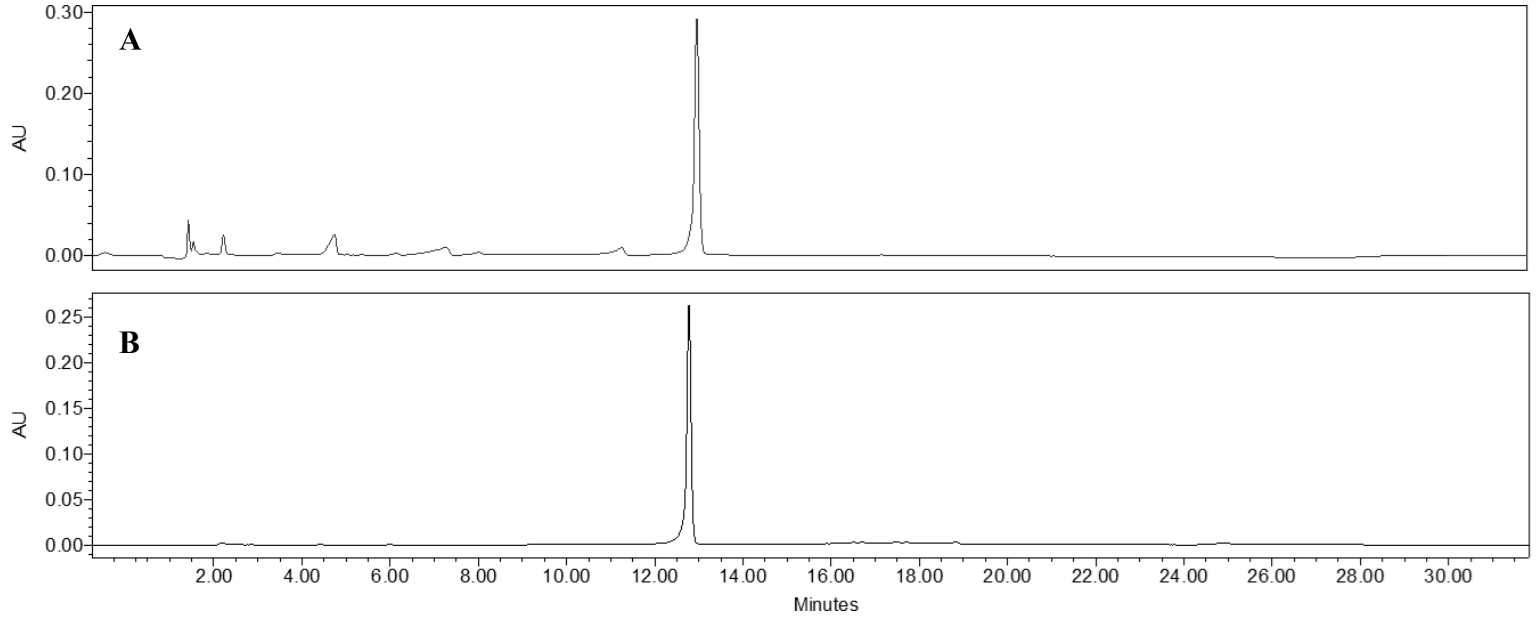
Supporting Figure A2.9: ESI-MS Spectrum of 6mer PSe.



Expected: 1936.1 Obtained: 1871.3

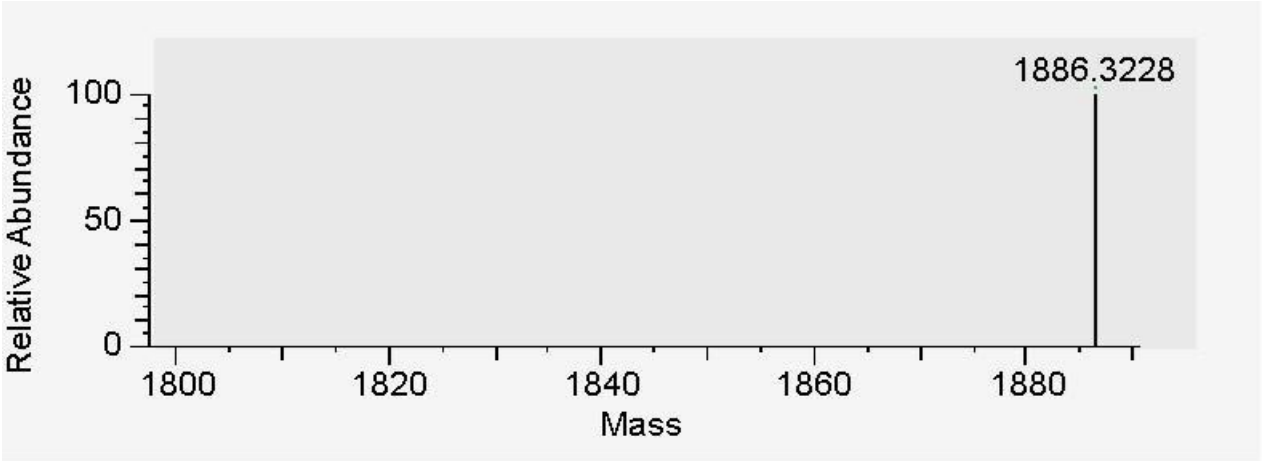
*The selenium atom was not retained as a weight of 1871.3 is the weight of the 6mer PO species.

Supporting Figure A2.10: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure 6mer MePO.



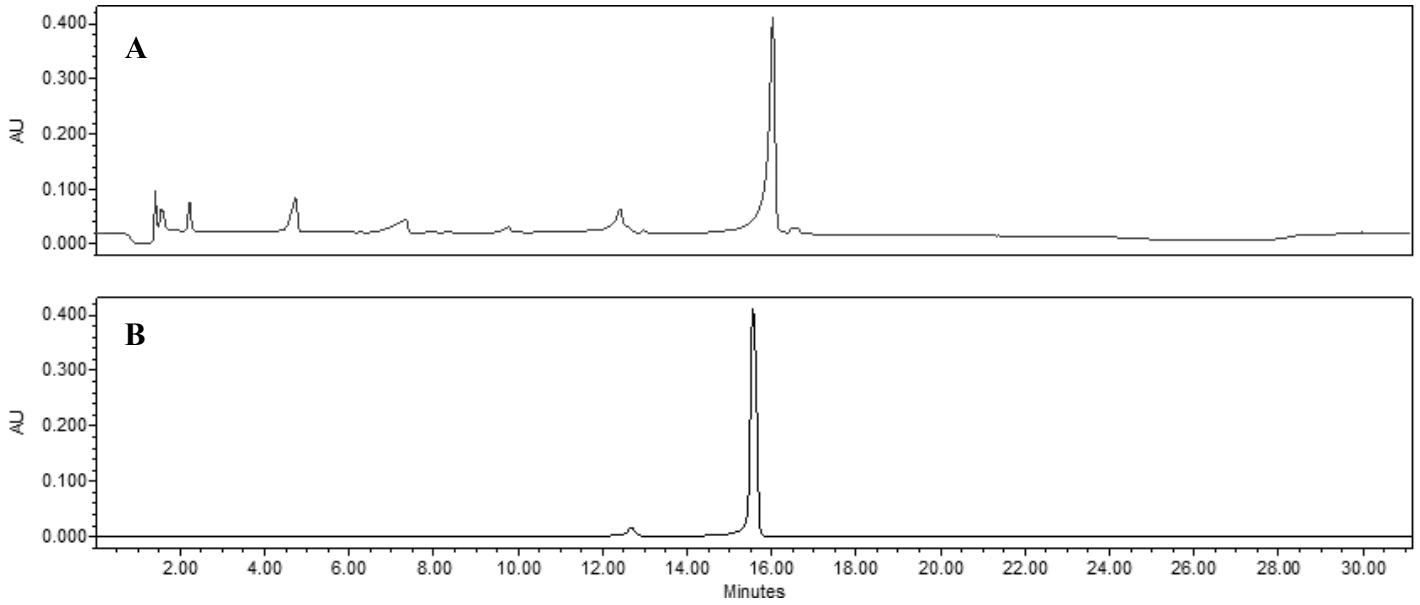
Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A2.11: ESI-MS Spectrum of 6mer MePO.



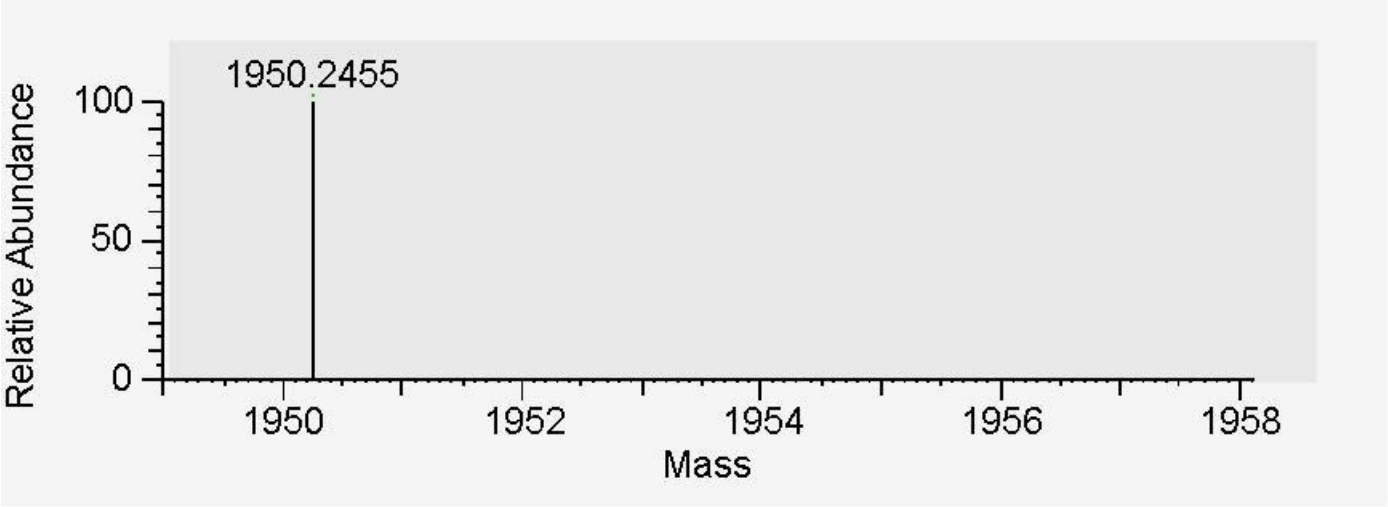
Expected: 1886.3 Obtained: 1886.3

Supporting Figure A2.12: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure 6mer MePSe.



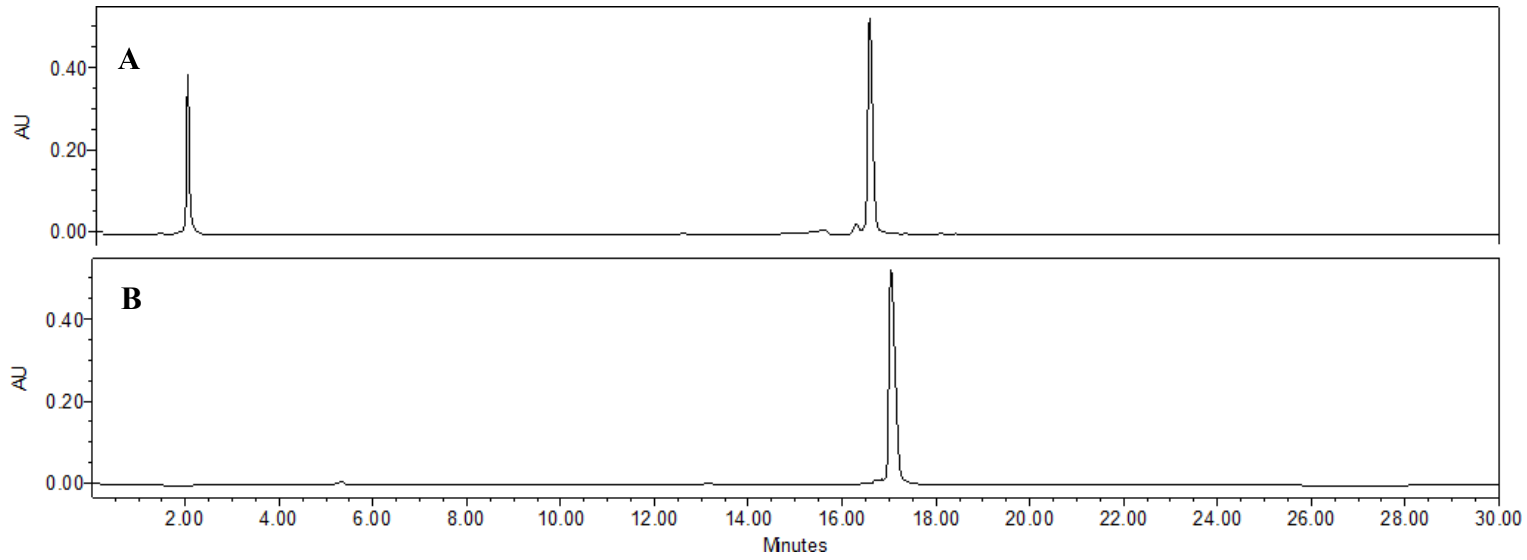
Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A2.13: ESI-MS Spectrum of 6mer MePSe.



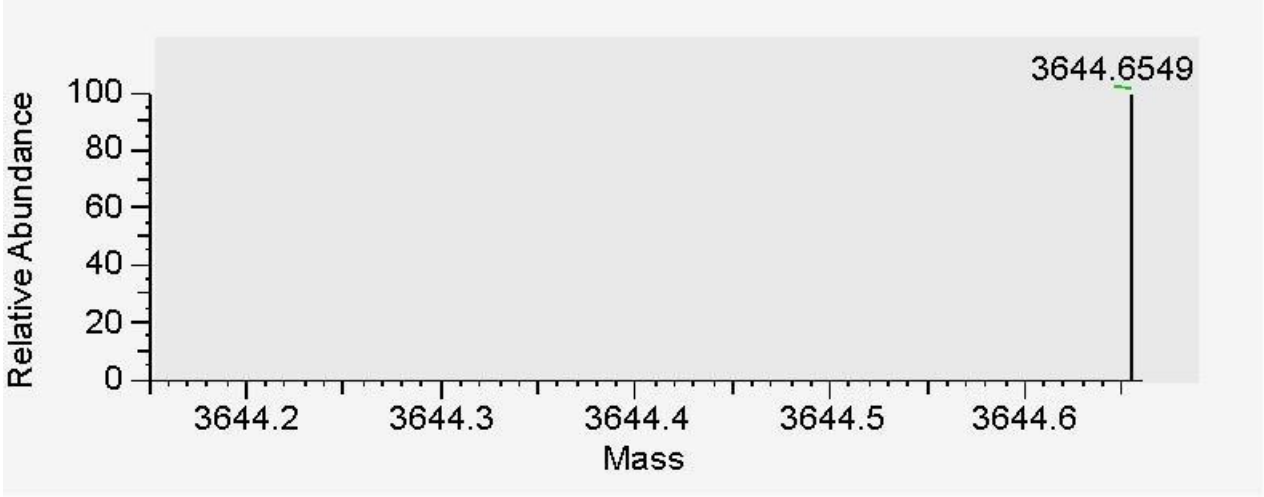
Expected: 1950.2 Obtained: 1950.2

Supporting Figure A2.14: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure 12mer OH.



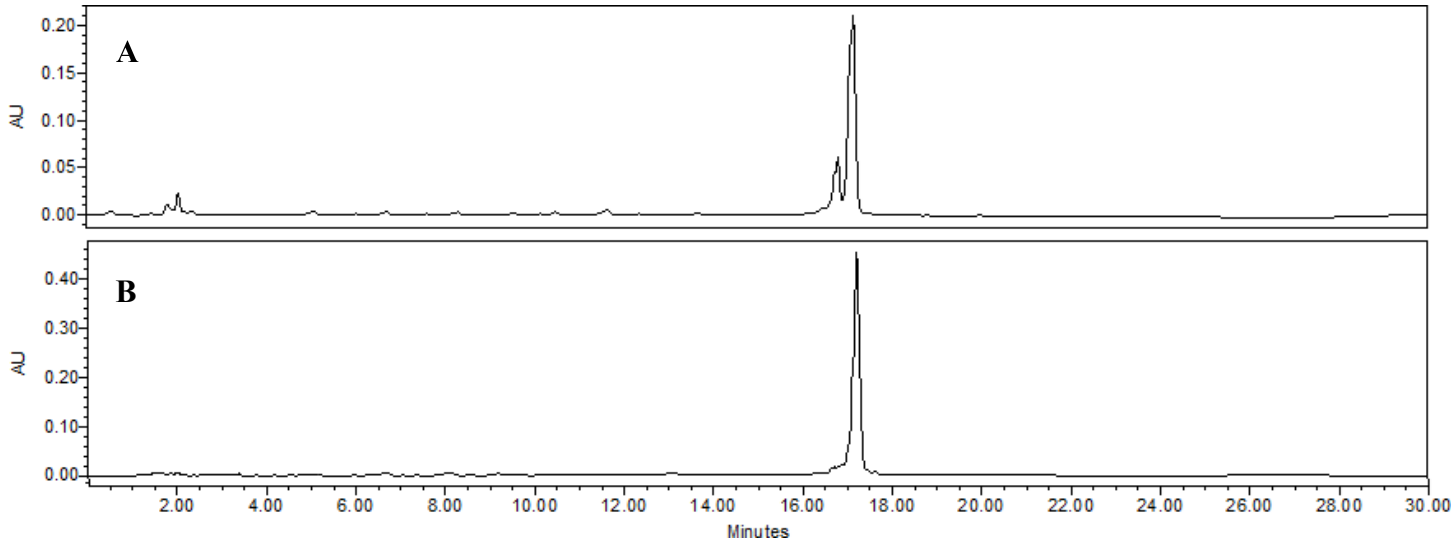
Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A2.15: ESI-MS Spectrum of 12mer OH.



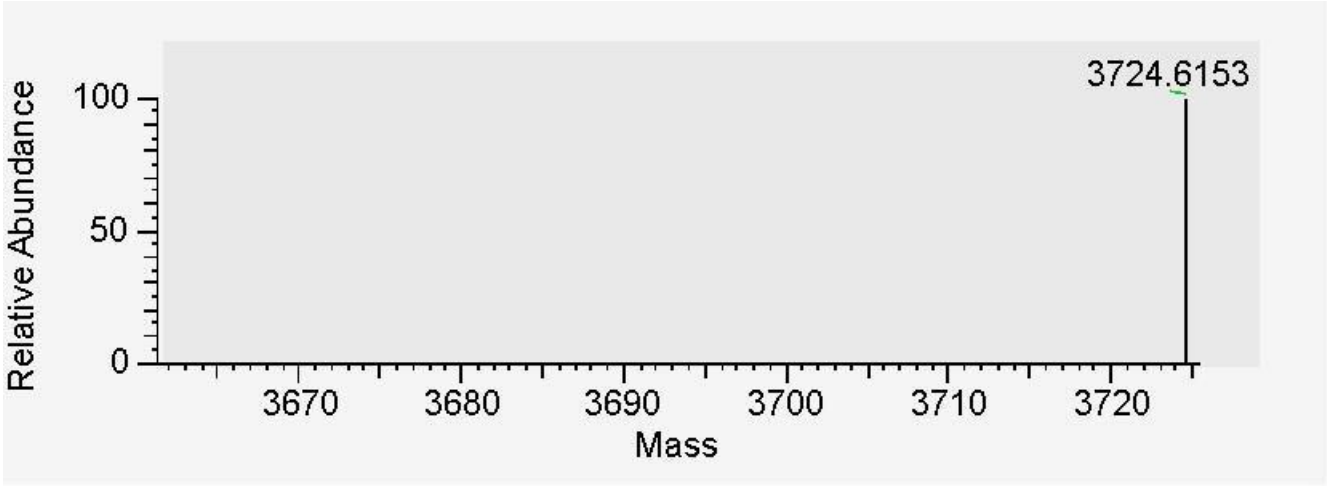
Expected: 3644.6 Obtained: 3644.6

Supporting Figure A2.16: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure 12mer PO.



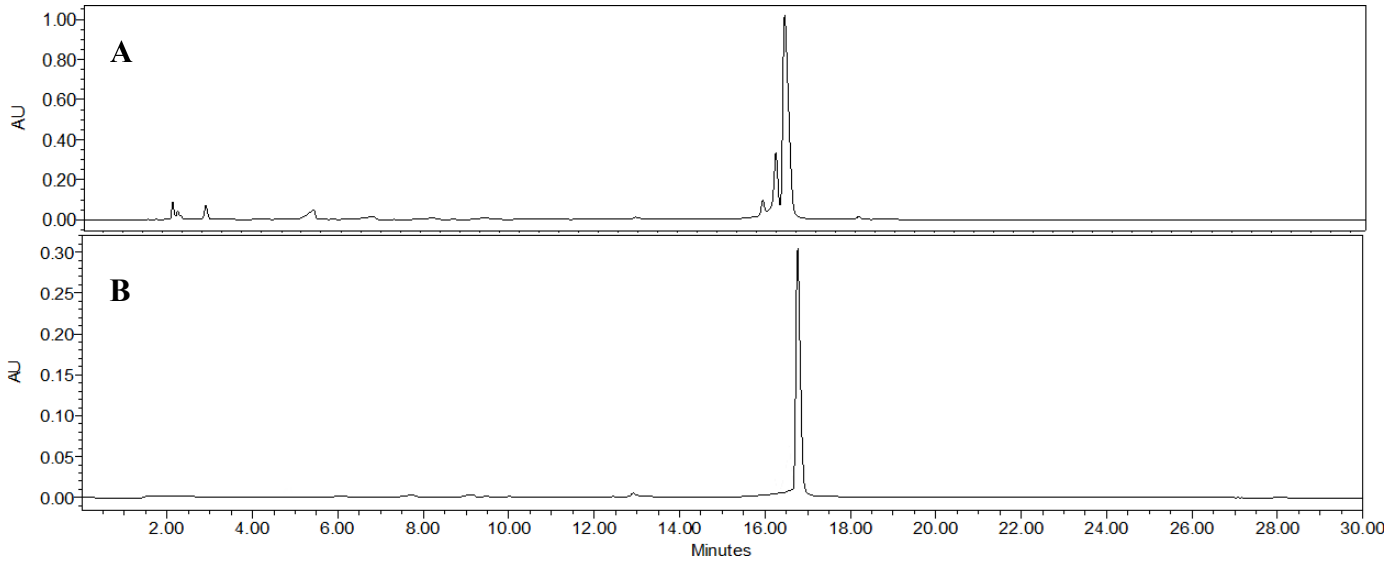
Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A2.17: ESI-MS Spectrum of 12mer PO.



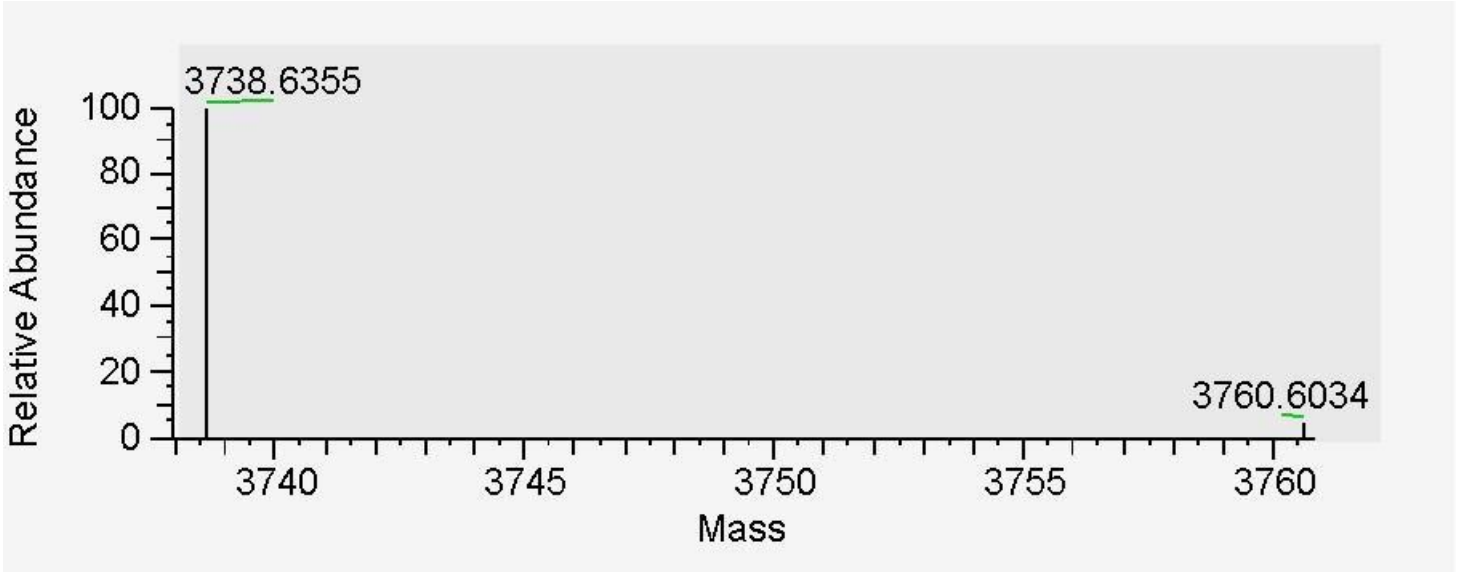
Expected: 3724.6 Obtained: 3724.6

Supporting Figure A2.18: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure 12mer MePO.



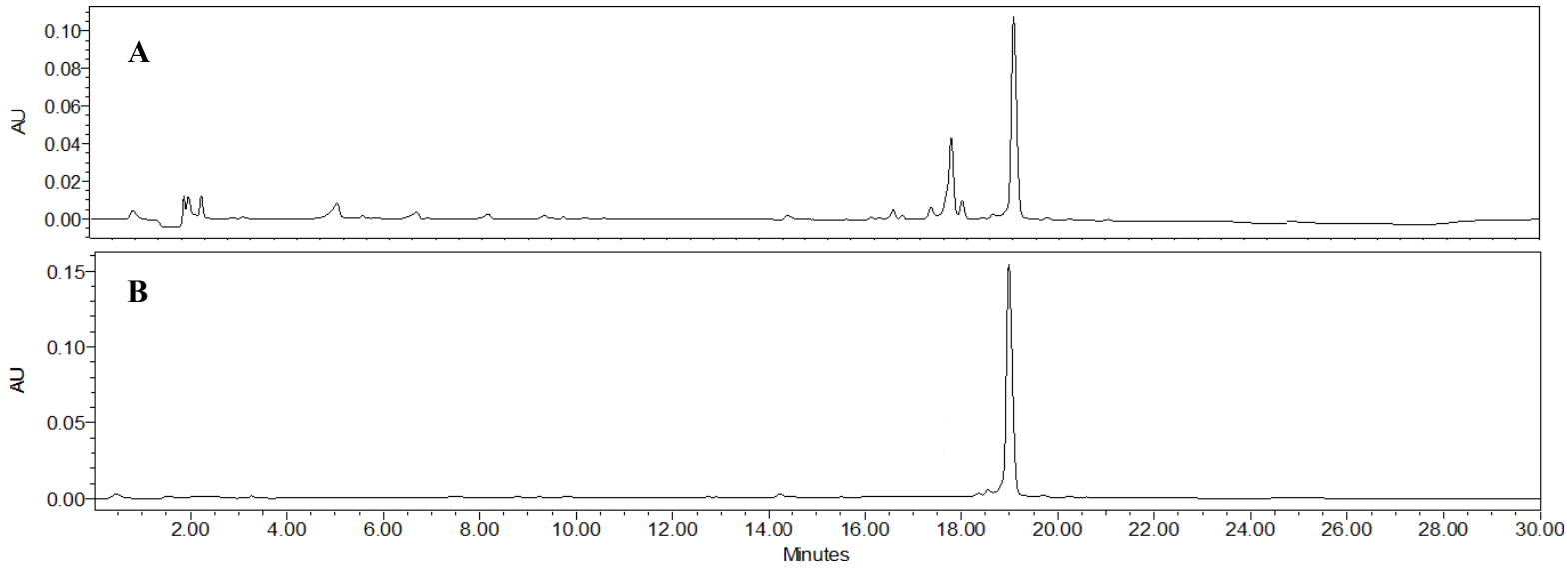
Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A2.19: ESI-MS Spectrum of 12mer MePO.



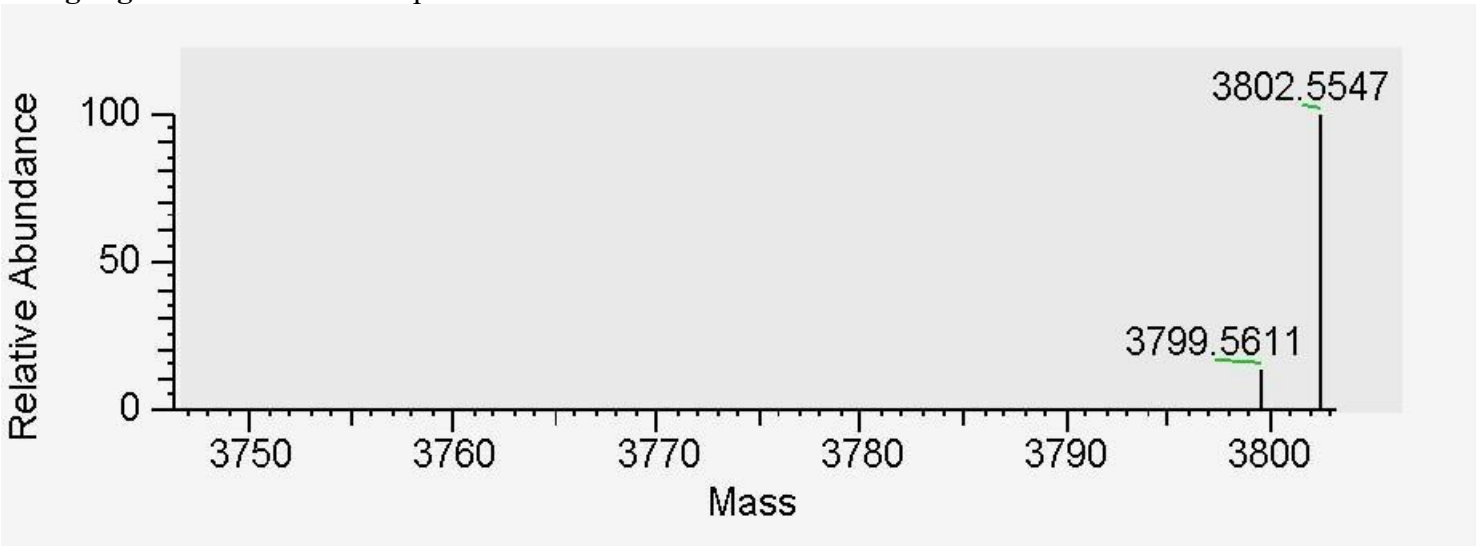
Expecting: 3738.6 Obtained: 3738.6

Supporting Figure A2.20: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure 12mer MePSe.



Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

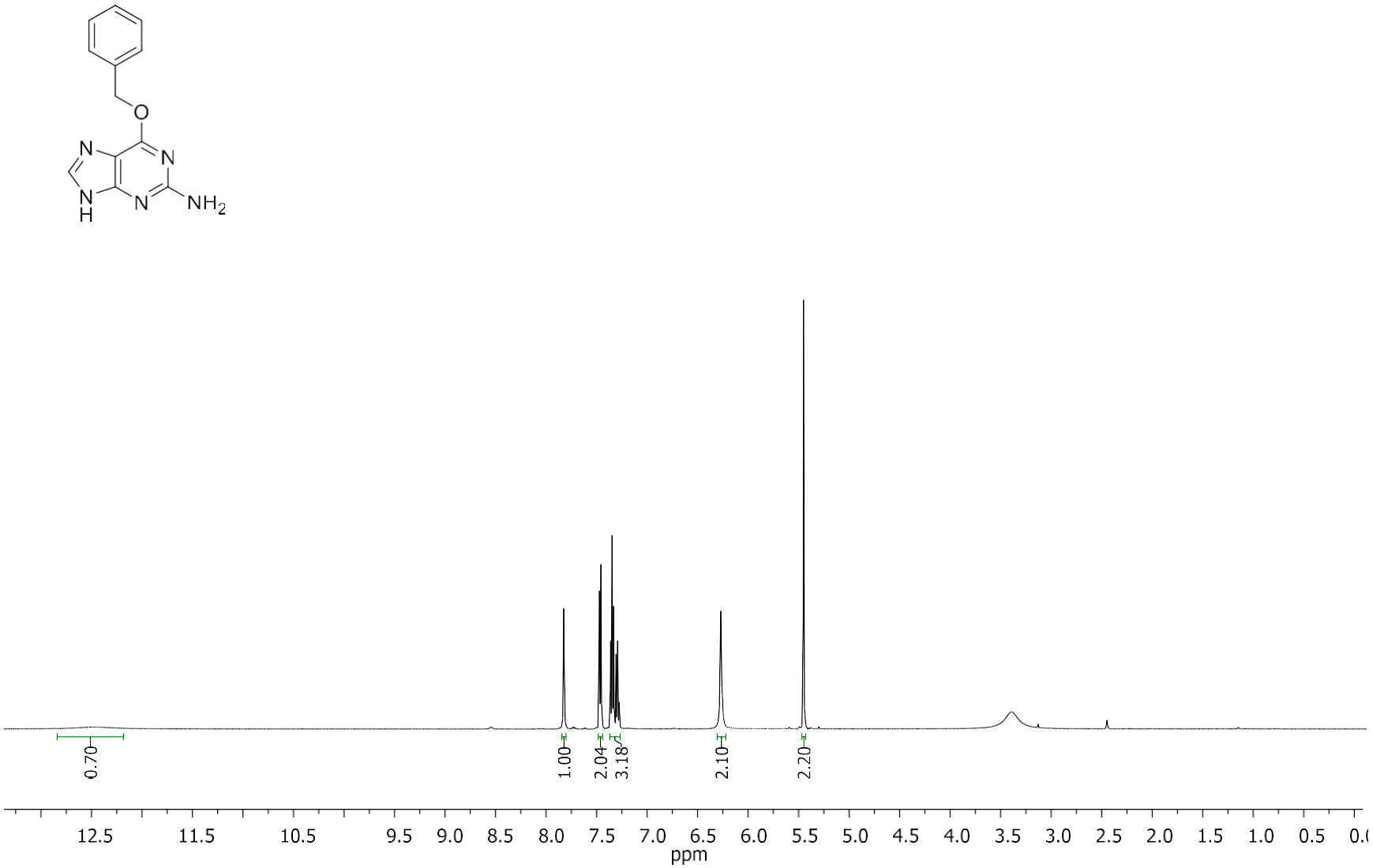
Supporting Figure A2.21: ESI-MS Spectrum of 12mer MePSe.



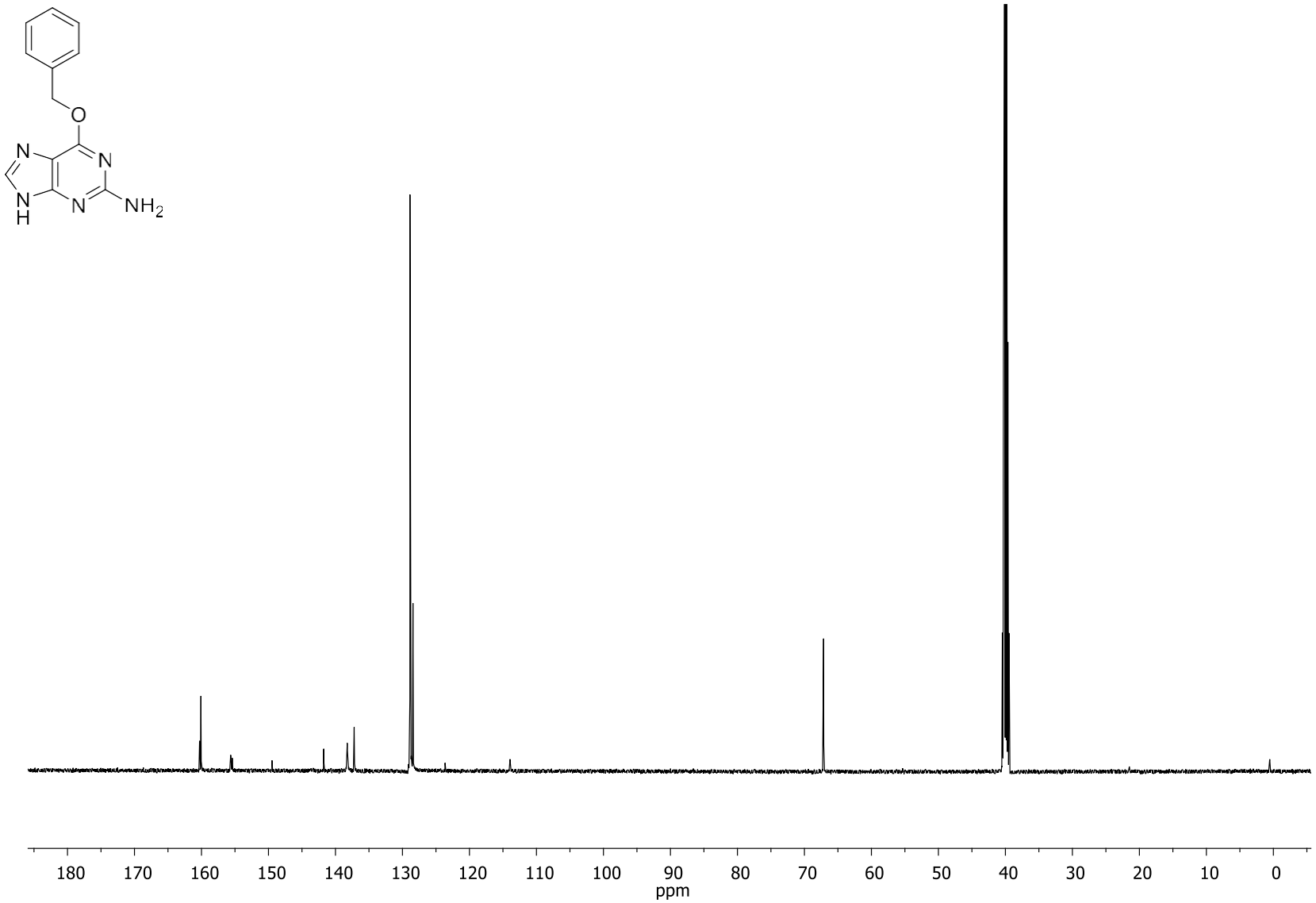
Expected: 3802.5 Obtained: 3802.5

Appendix III: Supporting Information for Chapter 4

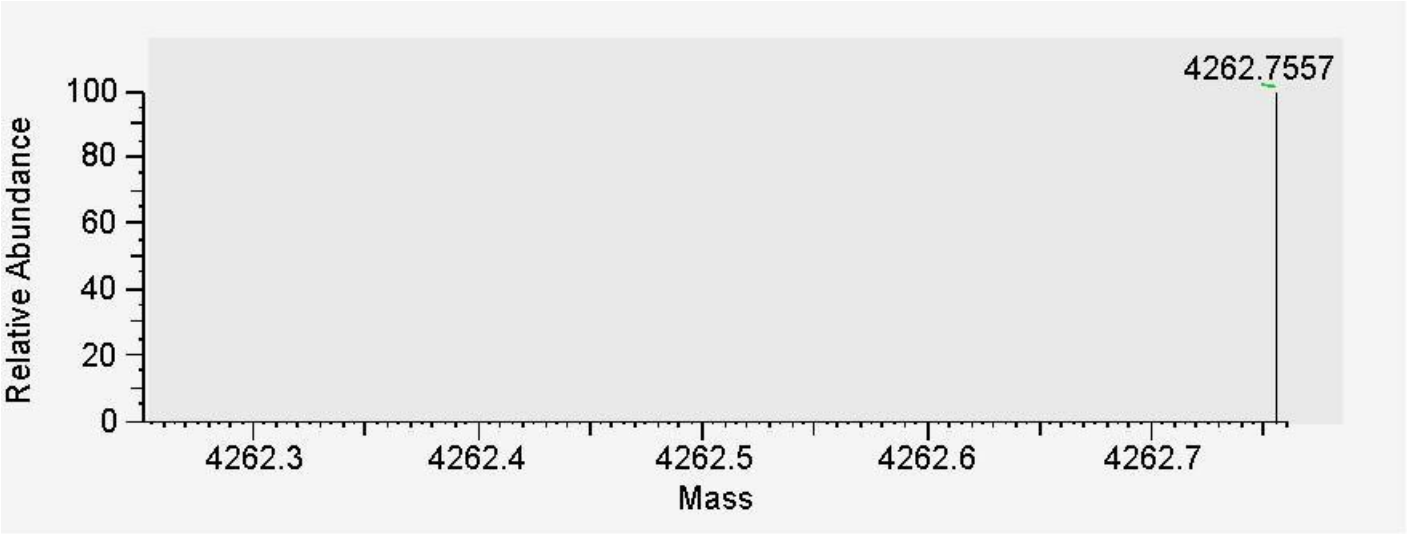
Supporting Figure A3.1: 500 MHz ^{1}H NMR spectrum of compound 13 in DMSO-d6.



Supporting Figure A3.2: 125.7 MHz ^{13}C NMR spectrum of compound 13 in DMSO-d6.

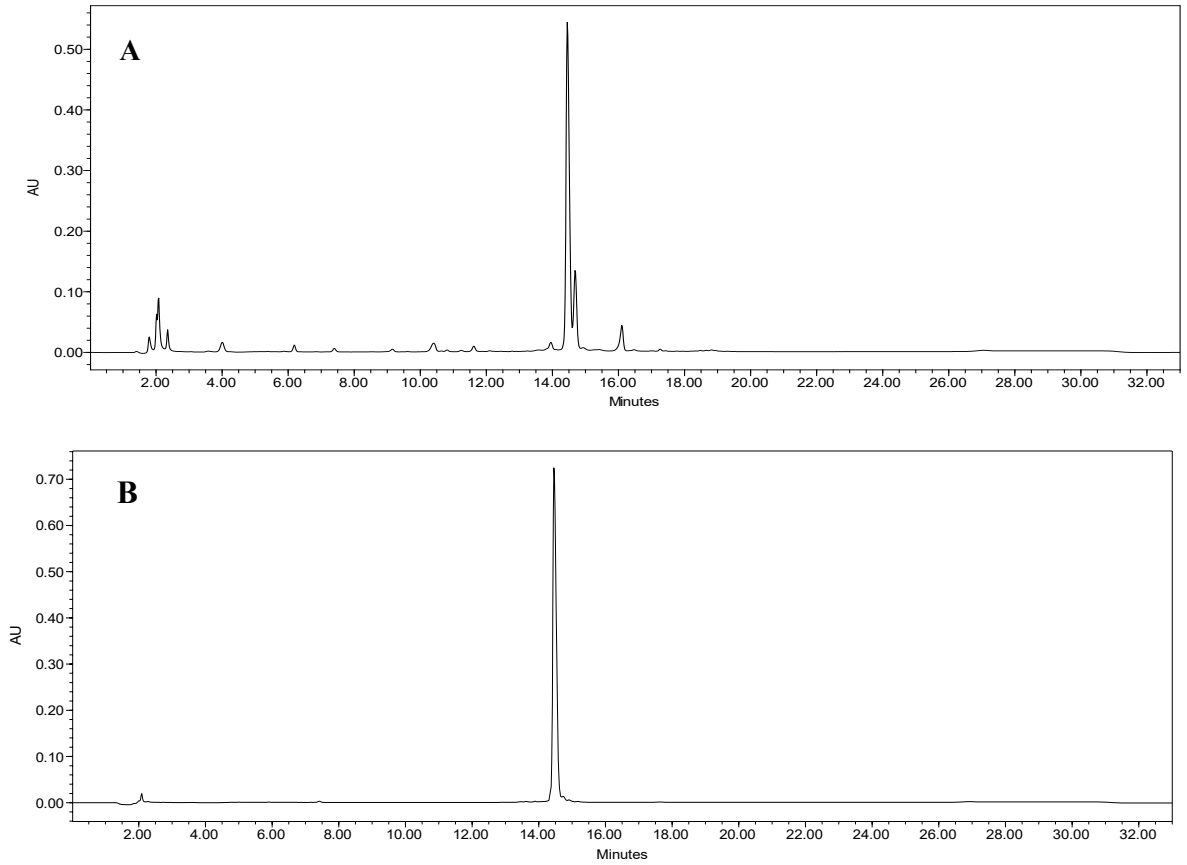


Supporting Figure A3.3: ESI-MS Spectrum of CJW 1083.



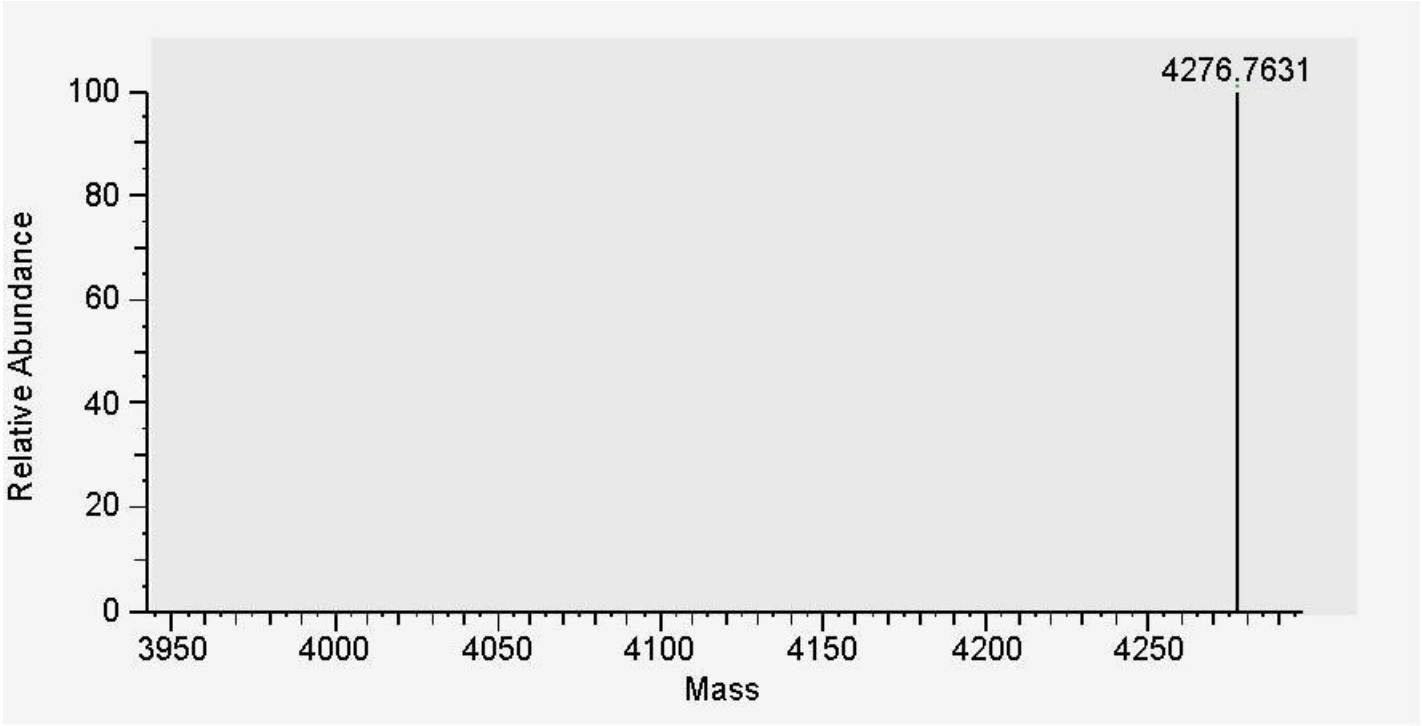
Expected: 4262.8 Obtained: 4262.8

Supporting Figure A3.4: IEX-HPLC chromatograph of (**A**) crude and (**B**) pure CJW 1128.



Elution at room temperature with a flow rate of 1 mL/min using a gradient of 0-50% buffer B over 20 minutes monitoring at 260 nm from an ion exchange HPLC column. Buffer A: 100 mM Tris-HCl pH 7.5 with 10% ACN and Buffer B: 100 mM Tris-HCl with 10% ACN and 1 M NaCl.

Supporting Figure A3.5: ESI-MS Spectrum of CJW 1128.



Expected: 4276.8 Obtained: 4276.8

Supporting Figure A3.6: Reaction conditions for gel in **Figure 4.8**.

Lane	DNA	Protein	Restriction Enzyme
1	✓		
2	✓		✓
3	✓	✓	
4-8	✓	✓	✓

Supporting Figure A3.7: Reaction conditions for gel in **Figure 4.9**.

Lane	DNA	Protein	Restriction Enzyme	Inhibitor
1	✓			
2	✓	✓		
3	✓	✓	✓	
4-10/11/12	✓	✓	✓	✓

Supporting Figure A3.8: Reaction conditions for gel in **Figure 4.10**.

Lane	DNA	Protein	Restriction Enzyme	Inhibitor
1	✓			
2	✓	✓		
3	✓	✓	✓	
4-10	✓	✓	✓	✓