Exploring Methodologies to Prepare Selenium Modified Oligonucleotides

Christopher Liczner

A Thesis in

The Department of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements

for the Degree of Master Science (Chemistry) at

Concordia University

Montréal, Québec, Canada

March 2018

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Christopher Liczner

Entitled: Exploring Methodologies to Prepare Selenium Modified Oligonucleotides

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

Master of Science (Chemistry)

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

Signed by the final examining committee:

Dr. Marek Majewski_____ Chair

Dr. Sébastien Robidoux_____ Examiner

Dr. Brandon Findlay_____ Examiner

Dr. Christopher Wilds_____ Supervisor

Approved by _____
Chair of Department or Graduate Program Director

_____ 20____ _____
Dean of Faculty

ABSTRACT

Exploring Methodologies to Prepare Selenium Modified Oligonucleotides

Christopher Liczner

X-ray crystallography has been central in determining structure-function relationships for biomacromolecules, such as nucleic acids and proteins, leading to a greater understanding of biochemical processes. The major barriers to nucleic acid structure determination by X-ray crystallography are the processes of crystallization and phasing. One strategy to overcome the latter challenge involves replacement of oxygen with selenium in a nucleotide. In the past two decades, selenium has been incorporated into oligonucleotides at many different sites, each with their specific advantages and limitations. Preparation of selenium modified nucleic acids often requires the multi-step synthesis of specialized phosphoramidites incorporated into oligonucleotides by solid-phase synthesis. It would be desirable to simplify the synthetic process to prepare these selenium containing oligonucleotides, which may result in their greater accessibility for crystallographers engaged in the determination of high resolution structural studies of nucleic acid structures.

In one approach, a simplified, greener and more versatile methodology to incorporate selenium into a solid support-bound oligonucleotide at the O-4 position of thymidine was explored. A range of reaction and deprotection conditions were evaluated, with promise of this approach demonstrated when selenium incorporation was attempted at a 5'-end residue of an oligonucleotide, followed by deprotection using a sterically hindered base. Another approach investigated the preparation of an oligonucleotide with 2'-deoxy-6-selenoinosine, a derivative of interest which was inspired

by the base pairing properties of 2'-deoxyinsoine towards other nucleobases. In the isolation and purification of this oligonucleotide, it was observed that the major product obtained was a non-complementary homodimer oligonucleotide connected by an interstrand diselenide bridge. Conditions to remove this cross-link were evaluated and expand the potential of selenium modified nucleic acids beyond X-ray crystallography for applications as a responsive nanomaterial.

Acknowledgements

First and foremost, I'd like to thank my wife, Cynthia, for her unwavering love, support and patience while I wrote this thesis. Also, thank you for showing me where to put my commas.

I'd like to express my appreciation to Dr. Christopher Wilds for allowing me to work in his lab and Dr. Anne Noronha for always making sure the lab ran smoothly. Your belief in me has really made a profound difference in my life.

To my committee members, Dr. Sébastien Robidoux and Dr. Brandon Findlay, thank you for pushing me to reach my potential and inspiring me with new ideas to investigate.

I'd like to also thank my past and present lab members for all their help and useful discussions. I promise to pay it forward.

Dr. Alexey Denisov, thank you for the help with selenium NMR.

A special thanks to Alain Tessier for his help with the mass spectrometry analysis of my oligonucleotides. Without him, the discovery of the diselenide bridged oligonucleotide would not have been possible.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

ACN - Acetonitrile
AcOH - Acetic acid
BOP - (Benzotriazol-1-yloxy)-tris(dimethylamino)phosphonium hexafluorophosphate
CD - Circular dichroism
CPG - Controlled-pore glass
dA - 2'-Deoxyadenosine
DBU - 1,8-Diazabicycloundec-7-ene
dC - 2'-Deoxycytidine
DCM - Dichloromethane
dG - 2'-Deoxyguanosine
dI - 2'-Deoxyinosine
DIPEA - *N*,*N*-Diisopropylethylamine
DMAP - 4-Dimethylaminopyridine
DMF - *N*,*N*-Dimethylformamide
DMT - 4,4'-Dimethoxytrityl
DNA - 2'-Deoxyribonucleic acid
dNTP - Deoxyribose nucleoside triphosphate
dT - Thymidine
DTT - Dithiothreitol
dU - 2'-Deoxyuridine
EDTA - Ethylenediaminetetraacetic acid
ESI-MS - Electrospray ionization mass spectrometry
EtOAc - Ethyl acetate
EtOH - Ethanol
GSH - Glutathione
HRMS - High resolution mass spectrometry
HPLC - High performance liquid chromatography
*i*PrOH - Isopropanol
IR - Infrared
MAD - Multiple wavelength anomalous diffraction
MeOH - Methanol
MIR - Multiple isomorphous replacement
mRNA - Messenger ribonucleic acid
MS - Molecular sieves
NMR - Nuclear magnetic resonance
NTP - Nucleoside triphosphate
P-Se - Phosphoroselenoate
PAGE - Polyacrylamide gel electrophoresis
PyBOP - (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
R_f - Retention factor
RNA - Ribonucleic acid

RP	- Reverse phase
rRNA	- Ribosomal ribonucleic acid
SAX	- Strong anion exchange
Se	- Selenium
S_N2	- Nucleophilic substitution, second order
TBDMS/TBS	- *tert*-Butyldimethylsilyl
TEA	- Triethylamine
THF	- Tetrahydrofuran
TLC	- Thin-layer chromatography
T_m	- Melting temperature
TREAT-HF	- Triethylamine trihydrofluoride
tRNA	- Transfer ribonucleic acid
UV	- Ultraviolet
Vis	- Visible

Contribution of Authors

Shown below are the contributions of each of the authors in this dissertation.

Chapter 3:

"Reversible Diselenide Cross-links are Formed Between Oligonucleotides Containing 2'-Deoxy-6-Selenoinosine".

Vincent Grenier: Initial synthesis of phosphoramidite 3.5.

Christopher Liczner: Wrote the first draft of the manuscript and corrected subsequent versions. Performed the remainder of the experiments: resynthesized phosphoramidite 3.5, characterized pre-amidite 3.4 and phosphoramidite 3.5, performed solid-phase synthesis of oligonucleotides (with the assistance of Dr. Anne Noronha), discovered and characterized the diselenide cross-linked oligonucleotide, performed the circular dichroism and UV thermal denaturation studies.

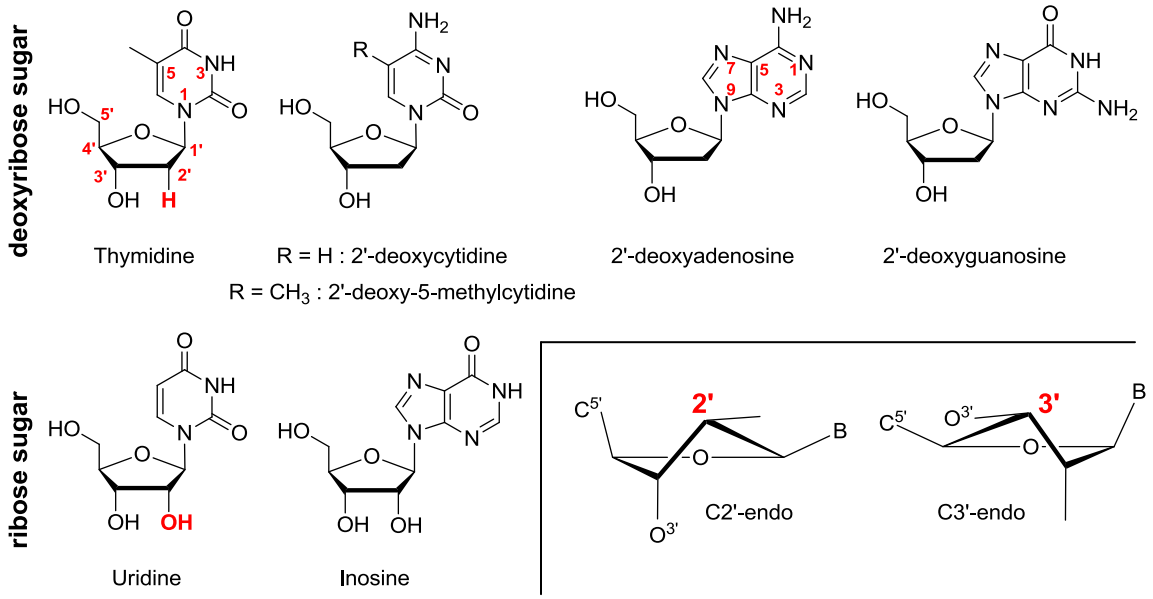
Dr. Christopher Wilds: Project supervisor responsible for correcting the manuscript.

Chapter 1. General Introduction

1.1. The Structure and Properties of Nucleic Acids

2'-Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are polymeric biological molecules common in all living organisms. These molecules are composed of nucleotide monomers, each consisting of three components; a nitrogenous heterocyclic base, a pentose sugar and a phosphate group. The pentose sugar is 2'-deoxyribose in DNA and ribose in RNA, with a differentiating hydroxyl group at the 2'-position, as shown in **Figure 1.1**. The pentose sugar adopts different conformations (also referred to as sugar puckers) in order to alleviate eclipsing interactions between their substituents. Two common twisted conformation sugar puckers are C2'-endo (Southern - S) and C3'-endo (Northern -N), shown in **Figure 1.1**. The heterocyclic bases are derivatives of purines and

Figure 1.1: Nucleoside structures with common sugar puckers, indicating the numbering system of the sugar, purine and pyrimidine bases.



pyrimidines. In DNA, the purine heterocyclic bases are adenine and guanine and the pyrimidine heterocyclic bases are cytosine and thymine. These bases are common to RNA, but with uracil replacing thymine. Other heterocyclic bases exist, such as hypoxanthine (commonly found in tRNA, used during the translation process) and 5-methyl-cytosine (used for regulating gene transcription), amongst others. The structures of these heterocycles are shown in **Figure 1.1**, depicted as nucleosides (i.e. without phosphate groups). The heterocyclic bases are connected to the C-1' position of the pentose sugar by a β-glycosidic bond so that the heterocycle and 5' carbon are on the same side of the pentose sugar. The heterocyclic bases can exist in any conformation, but they are normally found in the *syn*- or *anti*-conformations depending on steric influences. The *syn*-conformation has the C-4 (purines) or C-2 (pyrimidines) atoms of the heterocyclic base oriented cis with respect to the O-4' of the pentose sugar, whereas the *anti*-conformation has the C-4 or C-2 atoms of the heterocyclic base oriented trans with respect to the O-4' of the pentose sugar.[1]

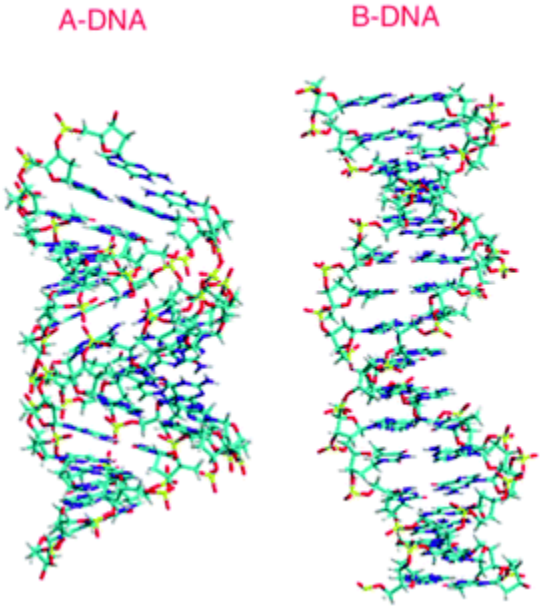
The primary structure of nucleic acids consists of nucleoside monomers connected through 3'-5' phosphodiester linkages. Natural nucleic acid sequences consist of a phosphate group at the 5'-end which is not involved in a phosphodiester linkage and a hydroxyl group at the 3'-end. Under physiological conditions, the phosphodiester backbone is negatively charged, allowing for optimal solubility in aqueous media. The primary sequence gains its uniqueness from the heterocyclic bases present. To designate such a sequence, single letter codes are used for each nucleotide monomer, namely A, G, C, T, U and I for

adenosine, guanosine, cytidine, thymidine, uridine and inosine, respectively. To differentiate DNA from RNA in strand naming, a “d” is used for deoxyribose and an “r” is used for ribose.[1]

Nucleic acids are polymorphic, capable of forming many different secondary structures. Recognition between the strands of nucleic acids is primarily attributed to hydrogen bonding between the heterocyclic bases. The hydrogen bonding pattern of the heterocycle’s dominant tautomeric form lead to cytosine base pairing with guanine and adenine base pairing with thymine/uracil. These interactions are known as Watson-Crick base pairing, which is the classical interaction between antiparallel DNA strands forming the double helix. The Hoogsteen face of the heterocyclic bases is also capable of hydrogen bonding, with an example depicted in **Figure 1.3A**. Inosine, a universal base, is capable of hydrogen bonding with any of the canonical bases.[2]

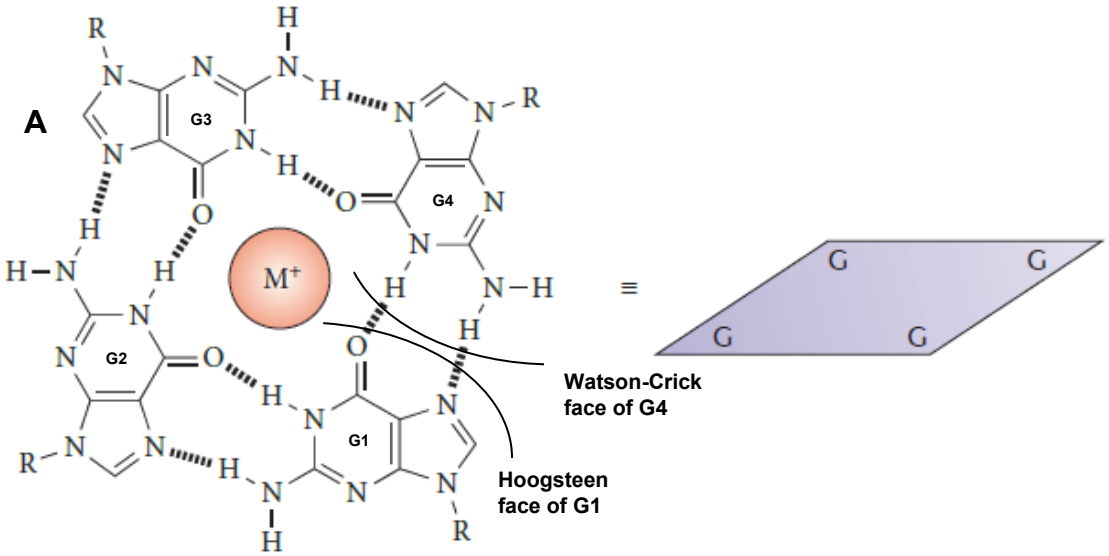
Strand interaction is not solely dependent on hydrogen bonding; sequence identity and environmental conditions also play a key role. DNA can adopt different anti-parallel double helical forms, the most common being the A-form and B-form,[3] as shown in **Figure 1.2**.[4] At low humidity and high salt

Figure 1.2: Molecular models of A-form and B-form DNA (taken with permission from reference 4).



concentration the A-form dominates. This double helix is right-handed, the sugars are C3'-endo and adjacent phosphates are separated by 6 Å.[1] At high humidity and low salt concentrations the B-form dominates which is the most common form of DNA found in nature. This double helix is right-handed, the sugars are C2'-endo and adjacent phosphates are separated by 6.77 Å.[1] The relatively closer negatively charged phosphate groups in A-form DNA are tolerated due to the high salt concentration, which masks the electrostatic repulsions. Although RNA is more commonly known to exist in its single-stranded form (mRNA) or in various structures such as hairpin loops in tRNA or rRNA, it can also form a Watson-Crick anti-parallel double helix.[5] This duplex is right-handed and the sugars adopt a C3'-endo sugar pucker, with an overall structure resembling that of A-form DNA.[1] Nucleic acids can also adopt more complex secondary structures than the double helix. Generally, the requirements for the formation of such structures depend on nucleobase composition and/or environmental factors. For example, the G-quadruplex[6,7] forms with G-rich sequences at neutral pH and requires specific metal cations, such as Na^+ and K^+, to be present for stabilization. In this, four guanines associate so that each base is engaged in four hydrogen bonds through Hoogsteen and Watson-Crick base pairing to form a nearly planar structure known as the G-quartet, shown in **Figure 1.3A**.[8] When two or more quartets are stacked, they form a G-quadruplex, shown in **Figure 1.3B**.[8] Each guanine directs its O-6 carbonyl oxygen towards the center of the quartet which interacts with a positively charged cation at the center. This secondary structure can form by both intramolecular and

Figure 1.3: (A) The interactions of a G-quartet illustrating the Watson-Crick and Hoogsteen faces of guanine. (B) Schematic diagrams of an intramolecular (left) and intermolecular (right) G-quadruplex (adapted with permission from reference 8).



intermolecular means, in a parallel or anti-parallel fashion. X-ray crystallography (described in **Chapter 1.3**) and NMR have been paramount in determining the nucleic acid structures discussed herein.

Specific properties of nucleic acids can be utilized in order to study these molecules. Firstly, the heterocyclic bases absorb strongly, with a molar extinction coefficient of 10^4 M^{-1} cm^{-1} at 260 nm. The (deoxy)ribose-phosphate backbone is essentially transparent to light in the ultraviolet region. The relatively strong absorption at this characteristic wavelength can be used to quantify the amount

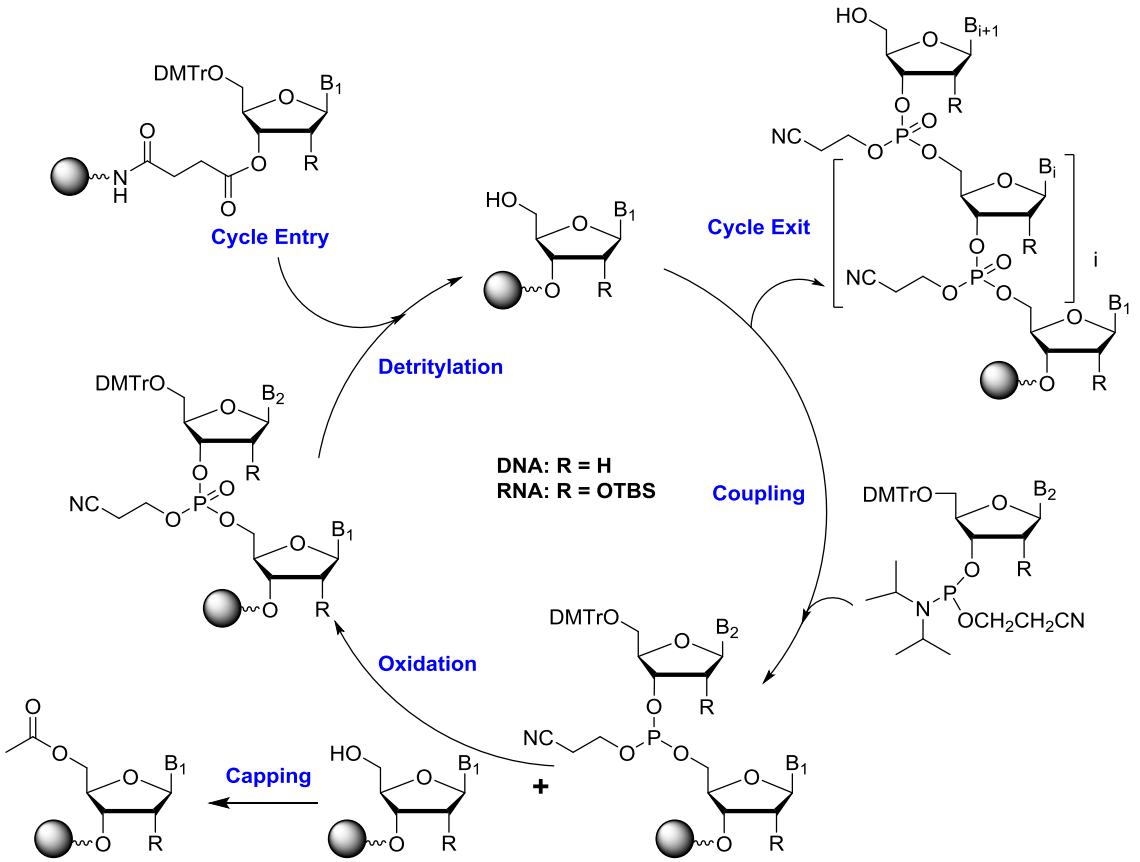
of nucleic acid present in solution using the Beer-Lambert law. Another crucial property is the difference in absorbance between nucleic acids strands which are base-paired and dissociated. In secondary structures, the bases are stacked with one another, causing a decrease in the molar extinction coefficient due to dipole-dipole interactions between the bases. When the strands dissociate due to environmental changes, such as an increase in temperature, the base stacking interactions are reduced which results in an increase in the molar extinction coefficient and therefore an increase in absorbance. This phenomenon is known as hyperchromicity and is the general basis for UV thermal denaturation experiments where absorbance at a specific wavelength is monitored as the temperature is increased. A sigmoidal transition denotes a two-state cooperative process: one population of nucleic acid secondary structures dissociates to another population. A common example of this phenomenon is DNA duplex dissociation into its single strands. The melting temperature (T_m) of such a transition is the temperature at which the ratio of the different populations in solution is 1:1. This occurs at the temperature where there is the greatest change in absorbance (inflection point), determined by taking the first derivative of the curve. This parameter can be used to assess the relative stability of secondary structures, however, for a reliable comparison, the experimental conditions must be identical since the T_m is sensitive to many factors, such as the buffer used and pH. Finally, nucleic acids are chiral molecules. The heterocyclic bases of nucleic acids are not chiral, however, the attachment between the N-9 of purines or N-1 of pyrimidines to the C-1' atom of the sugar by the glycosylic bond results in a

chiral disruption to their absorption of UV light. Nucleic acids are therefore able to be studied by circular dichroism (CD) spectroscopy. A CD signal is produced from the differential absorption of left and right circularly polarized light. The many secondary structures of nucleic acids (B-form, A-form, G-quadruplex DNA, etc.) exhibit characteristic CD profiles, allowing for rapid qualitative determination of structural information.[1]

1.2. Solid-Phase Synthesis of Oligonucleotides

One of the most widely used approaches for the assembly of oligonucleotides (sequences of less than 100 nucleotides) is solid-phase

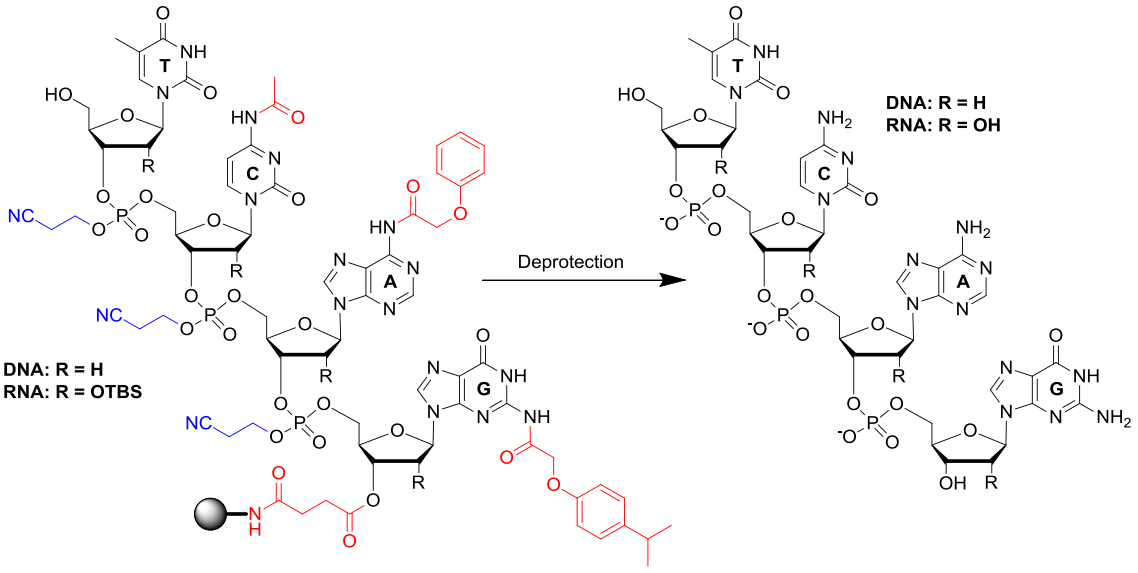
Scheme 1.1: The oligonucleotide solid-phase synthesis cycle.



synthesis using phosphoramidite chemistry.[9,10,11] Solid-phase synthesis is especially useful in that it can be adapted to automated systems, impurities and excess reagents can be washed away without the need for purification and large excess of reagents can be used to drive reactions to completion in short time periods.[12] Through the solution synthesis of modified phosphoramidites, novel nucleotides can be incorporated into an oligonucleotide strand by solid-phase synthesis. These modifications can be used in order to alter the properties of nucleic acids, thereby expanding their range of applications.

In this approach,[13] the DNA or RNA sequence is grown from a solid support, such as controlled-pore glass (CPG), typically in the 3' to 5' direction. The nucleoside at the 3'-end of the oligonucleotide is synthesized (in solution) with a 4,4'-dimethoxytrityl (DMT) protecting group at the 5'-hydroxyl position along with protecting groups on the heterocyclic base. The 3'-hydroxyl is then converted into the active succinate ester which can be coupled to the amino group attached by a long alkyl chain of the functionalized CPG (this step is commonly not required as there exists commercially available CPG bound to each of the protected 5'-O-DMT nucleosides). A long chain is crucial in order to extend the reactive groups away from the surface of the solid-support thereby making them accessible for subsequent reactions. Acid treatment removes the 5'-O-DMT group and exposes a free hydroxyl group. This hydroxyl group then couples to an incoming phosphoramidite via an S_N2 mechanism (activated by a tetrazole catalyst), creating the phosphite linkage. After this step, any unreacted 5'-hydroxyl groups on the support-bound nucleotide are capped by acetylation to

prevent the continued synthesis of undesired sequences. The phosphite triester linkage is oxidized by iodine in the presence of water and pyridine. The sequence then re-enters the cycle to continue the chain assembly (the cycle is shown in **Scheme 1.1**). Once the full length sequence is complete, it must be cleaved off the solid-support and deprotected (**Scheme 1.2**). Basic conditions are required in



Scheme 1.2: A protected support-bound oligonucleotide (left) and a deprotected free oligonucleotide (right).

order to remove the cyanoethyl group from the phosphate backbone via a β-elimination mechanism (shown in blue in **Scheme 1.2**). Nucleophilic conditions are used to remove the protecting groups on the heterocyclic bases and cleave the oligonucleotide from the solid-support (shown in red in **Scheme 1.2**). Commonly, a solution of NH_4OH/EtOH (3:1, v/v) is used at elevated temperatures in order to achieve deprotection and cleavage in one step. In the case of RNA, a subsequent treatment with a fluoride source is required to remove the *tert*-butyldimethylsilyl (TBDMS/TBS) group at the 2'-position. When compared to

naturally occurring sequences, solid-phase synthesized oligonucleotides lack the 5' terminal end phosphate group. They may, however, have a 5' terminal DMT which can aid in the purification process. After deprotection and cleavage, purification of the desired sequence is required. Common methods include denaturing polyacrylamide gel electrophoresis (PAGE) and strong anion exchange (SAX) HPLC.

The efficiency of the phosphoramidite coupling step can be determined from the intensity of the absorbance of the cleaved DMT cation at 495 nm. A high average stepwise coupling efficiency is required in order to acquire a high yield of desired oligonucleotide. The coupling efficiency is never 100 %, thereby instilling a length limitation to the solid-phase synthesis methodology. The pore size of the solid-support also influences the length of oligonucleotide which can be synthesized. CPG with pore sizes of 500 Å are common for the synthesis of shorter oligonucleotides, however, for longer sequences, steric influences prevent proper diffusion of reagents causing poor coupling reactions. The loading of the solid-support, or the moles of nucleoside bound per gram of solid-support, is also an important parameter. With high loading supports, steric hindrance between adjacent oligonucleotides strands can become an issue as the oligonucleotide is elongated.[12]

1.3. Nucleic Acid X-ray Crystallography

The technique of X-ray crystallography is used in order to determine high-resolution structures of biomacromolecules, such as proteins

and nucleic acids. A diffraction pattern is generated by irradiating homogenous crystals of the desired molecule with X-rays. In order to generate a three-dimensional electron density map of the molecule, three parameters are required, specifically the amplitude, frequency and phase of the diffracted X-rays.[14] From the experimental diffraction pattern, the amplitude and frequency can easily be measured. The amplitude is determined from the intensity of the X-ray reflections and the frequency from the position of the reflections in reciprocal space. The phase information, however, is absent and must be determined by independent means. This is known as the phase problem in X-ray crystallography and constitutes one of the major bottlenecks of the technique. Several methods have been developed in order to solve this problem and acquire the phase information. The direct method iterates through all possible phase angles for each reflection of the diffraction pattern in order to determine the phase information by brute-force. This method is not always viable due to computational power limitations and currently only works for smaller molecules (< 1000 non-hydrogen atoms) with atomic resolution diffraction data.[15] The molecular replacement method relies on the use of a search model, a known structure that is homologous to the target molecule. The phase information is determined by rotating and translating the search model until it produces a simulated diffraction pattern that best matches the observed diffraction pattern. Should a homology model not be available, there exist experimental approaches to determine phase information, such as multiple isomorphous replacement (MIR) and multiple wavelength anomalous diffraction (MAD). In the MIR method,

heavy atoms are introduced into the crystal which do not cause any structural perturbations, retaining the same crystalline order as the native, unmodified crystals. Intensity differences between the X-ray diffraction patterns generated from native and heavy atom derivatized crystals allow for the calculation of the phases of their associated reflections. It is important to introduce at least two different heavy atoms into separate crystals in order to prevent phase ambiguity. The MAD method only requires a single heavy atom derivatized crystal which is irradiated with X-rays at multiple wavelengths in order to deduce the phase information, making it the more attractive and widely used method. In both MIR and MAD, techniques to introduce heavy atoms into the crystal include covalent modification before crystallization, soaking the native crystal in a solution of heavy-atom compounds or growing the crystals in the presence of heavy-atom containing molecules (co-crystallization). Commonly, the heavy atoms introduced are electron-dense transition metals (third period or lower) capable of significantly scattering X-rays, but even atoms such as selenium and bromine have successfully been used.[14]

With the experimental diffraction data and phase information collected, a three-dimensional electron density map of the structure is generated computationally. It is then necessary to solve the X-ray crystal structure by model building and model refinement. A model is built to best fit the experimental electron density map, however inherent errors with the initial phase values must be corrected. These corrections are made by model refinement, the process of computationally moving atoms of the model in real space, generating a new

diffraction pattern in reciprocal space and then determining how well this diffraction pattern matches the experimentally determined one. After many rounds of model building and refinement, the model converges on the best solution.[14]

With regard to phasing in nucleic acid X-ray crystallography, the direct method may not be ideal for longer oligonucleotides where only low resolution diffraction data is available. Molecular replacement is challenging with nucleic acids due their highly repetitive units, making the simulated diffraction pattern less reliable.[15] Furthermore, molecular replacement is not possible for determining the phase information of novel structures. The techniques used in MIR to introduce heavy atoms into the crystal are frequently problematic for nucleic acids. Co-crystallization and soaking with heavy metal cation solutions can cause the random cleavage of the phosphate backbone, are prone to non-specific binding and can cause structural perturbations. Additionally, nucleic acid crystals frequently lack large solvent channels for efficient diffusion of heavy atoms when the crystal is soaked.[16] A classic technique to determine the phase problem of nucleic acids by MAD has been bromine derivatization; however this strategy also suffers from limitations. Due to the chemical nature of bromine, the incorporation sites are fairly limited and have mostly been at the 5-position of uracil and cytosine. The carbon-halogen bond is also light sensitive, susceptible to radiolysis during X-ray diffraction experiments.[15] Furthermore, incorporation of bromine commonly causes structural perturbations and even prevents crystallization altogether in some instances.[16]

Due to all these limitations, there has been great interest in using selenium containing nucleic acids for phasing in X-ray crystallography. Not only has the use of selenium found great success in protein X-ray crystallography with selenomethionine, it is also ideal for nucleic acids due to the many sites available for modification in the scaffold (selenium is a heavier homologue of oxygen). It is also much more stable to X-ray irradiation and there are many incorporation sites which do not cause significant structural perturbations. Crystallization, the other major bottleneck of X-ray crystallography, has also been shown to be facilitated with selenium-containing nucleic acids.[15]

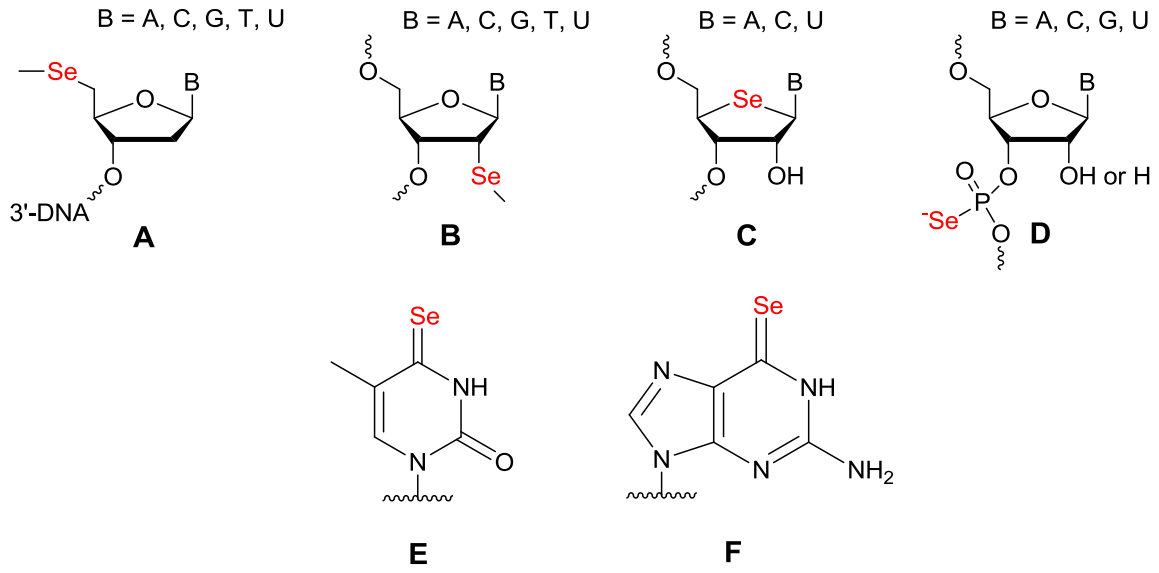
1.4. Selenium Incorporation into Oligonucleotides

Selenium can be incorporated into nucleic acids by two approaches: enzymatic and chemical synthesis. Enzymatic synthesis utilizes DNA and RNA polymerases with selenium analogues of deoxyribose nucleoside triphosphates (dNTPs) and nucleoside triphosphates (NTPs) as their substrates, respectively.[15] In contrast to solid-phase oligonucleotide synthesis, enzymatic synthesis grows the strand in the 5' to 3' direction. This method also requires that the polymerases actually recognize the selenium analogues as substrates in order to catalyze their incorporation. Furthermore, replacement of any one of the four dNTP or NTP building blocks with selenium-modified substrates results in approximately 25 % nucleotide substitution after the polymerization reaction. In contrast to proteins, which only have a 2.2 % natural abundance of methionine,[17] this is quite excessive and may lead to potential problems such as undesirable

structural perturbations. The other approach involves the solid-phase synthesis of oligonucleotides from phosphoramidite building blocks (described in **Chapter 1.2**). This method is advantageous as it provides control of the number of selenium atoms incorporated as well as their position in the sequence. Since solid-phase synthesis has a length limitation for both DNA and RNA, combined strategies have been developed to ligate longer enzymatically synthesized sequences to selenium-modified solid-phase synthesized oligonucleotides.[18,19]

Through chemical synthesis, selenium has successfully been introduced into many sites of nucleic acids. The first instance of selective incorporation of selenium into an oligonucleotide by solid-phase synthesis was at the 5' position (**Figure 1.4A**).[20] This was accomplished by the multi-step solution synthesis of the 5'-methylseleno phosphoramidite and subsequent solid-phase incorporation into an oligonucleotide. This work demonstrated that selenium-containing phosphoramidites were compatible with solid-phase synthesis as well as the

Figure 1.4: Nucleic acid selenium modification sites.



stability of selenium-containing oligonucleotides during the deprotection and purification steps. Following this work, the first nucleic acid duplex structure was solved by MAD phasing using 2'-methylseleno deoxyuridine (**Figure 1.4B**).[21] A further example is the preparation of a phosphoramidite with selenium replacing the sugar oxygen (**Figure 1.4C**).[22] In a unique approach, during solid-phase synthesis certain phosphite backbone positions were converted to the corresponding phosphoroselenoate (**Figure 1.4D**).[23] This was achieved by replacing the oxidation step in the solid-phase synthesis cycle (involving iodine in water / THF / pyridine) with manual treatment of the support-bound oligonucleotide with a saturated solution of KSeCN. Each of these incorporation sites have been useful in their own right, however they all possess certain limitations. Modifications at the 5'-position are limited to the 5' terminal end of the oligonucleotide. Substitutions of the sugar ring (4'-) oxygen are known to be synthetically challenging and modifications at the 2'-position may cause undesirable structural changes, such as promoting an A-form C3'-endo sugar pucker. Finally, replacing the non-bridging phosphate oxygen with selenium introduces a new chiral center in the molecule, which necessitates the separation of diastereomers by HPLC. This selenium modification was also best tolerated near the 5'-end of the oligonucleotide due to deselenization observed after multiple cycles of oxidation. More recently, selenium has been introduced at different nucleobase sites[24,25] through the multi-step synthesis of phosphoramidite derivatives, which essentially eliminate the limitations discussed above (**Figure 1.4E and 1.4F**). The X-ray structure of a DNA duplex containing

4-selenothymidine ($d^{4Se}T$) demonstrated that the larger selenium atom is well accommodated as the structure was virtually identical to the native duplex. The duplex stability and $d^{4Se}T$/dA base pair were also essentially the same as in the unmodified duplex. In a duplex containing 2'-deoxy-6-selenoguanosine ($d^{6Se}G$), the $d^{6Se}G$/dC base pair was shifted by approximately 0.3 Å, causing a decrease in duplex stability; however the overall structure was still very similar to the native duplex. Neither one of these cases showed any evidence of deviation from normal hydrogen bonding patterns when selenium was present.

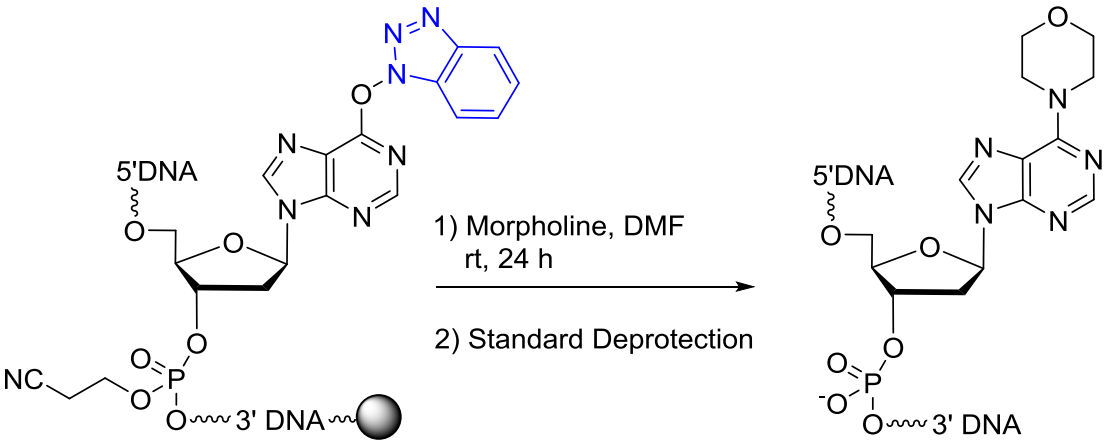
1.5. Project Objectives

To date, X-ray crystallography of biomacromolecules is still a long and arduous process and there is great interest in making the technique less time consuming. To this end, the present work has gone towards simplifying selenium incorporation into oligonucleotides. The previous methods have almost exclusively relied on the multi-step synthesis of modified phosphoramidites, followed by solid-phase synthesis. This strategy has been quite successful, but there remains the possibility for improvement.

The Lakshman group[26] have described the synthesis of O^6-(benzotriazol-1-yl) inosine and 2'-deoxyinosine derivatives, a class of easily convertible nucleoside analogues. A variety of commercially available sugar-protected and -unprotected inosine and 2'-deoxyinosine nucleosides were treated with benzotriazol-1-yloxy-tris(dimethylamino)phosphonium hexafluorophosphate (BOP) in order to introduce the benzotriazolyl group at the 6-position of

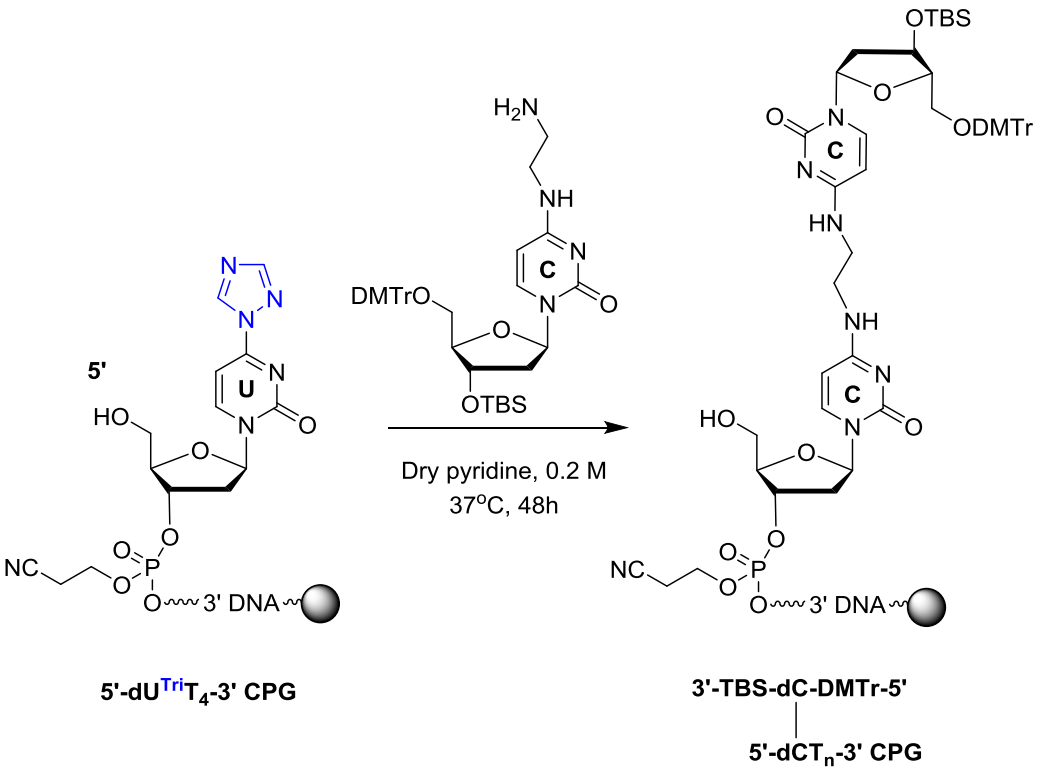
hypoxanthine. This modification was shown to tolerate different substitutions of the sugar oxygen atoms and the benzotriazolyl group could be displaced by a variety of nucleophiles. More importantly, these nucleosides could be elaborated

Scheme 1.3: Incorporation of morpholine into a support-bound oligonucleotide



to the corresponding phosphoramidite and incorporated into an oligonucleotide through solid-phase synthesis while retaining the benzotriazolyl group. Furthermore, post-synthetic modification was demonstrated as a proof-of-concept. This was accomplished by the treatment of a benzotriazolyl-containing solid-support bound oligonucleotide with nucleophilic morpholine (**Scheme 1.3**).

The Miller group[27] also demonstrated the feasibility of post-synthetic modifications (**Scheme 1.4**). A CPG-bound oligonucleotide containing O^4-triazolyl-2'-deoxyuridine was first prepared. The triazole group was displaced by the aminoethyl of 5'-*O*-DMT-3'-*O*-TBDMS-N^4-(2-aminoethyl)deoxycytidine, generating a cross-link. The DMT group could then be removed and the oligonucleotide further extended by solid-phase synthesis.



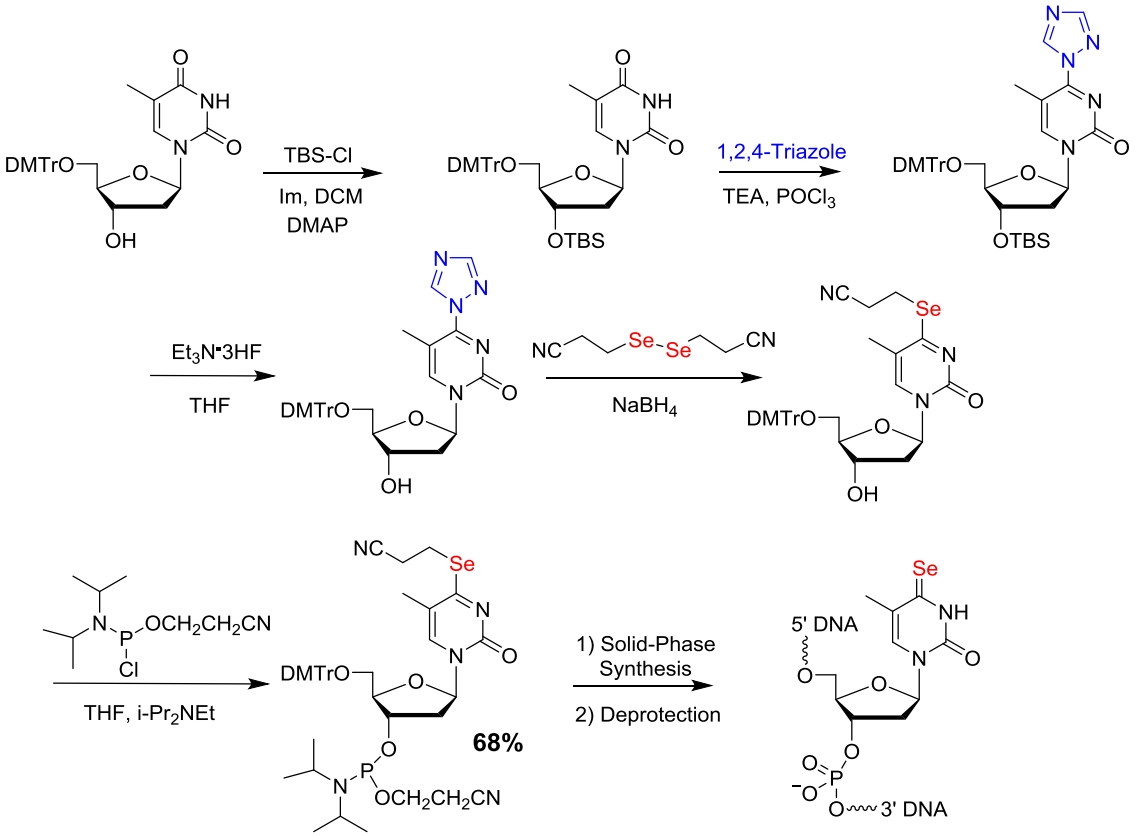
Scheme 1.4: Incorporation of an interstrand cross-link into a support-bound oligonucleotide.

Independently, the Huang group[28] developed an effective nucleophilic selenium species for incorporation of selenium at the nucleobase. Selenium powder is reduced with sodium borohydride in the presence of 3-bromopropionitrile to generate di(2-cyanoethyl) diselenide. Further reduction of the diselenide with sodium borohydride produces nucleophilic 2-cyanoethyl selenide. This nucleophile was successfully employed in the synthesis of oligonucleotides containing d^{4Se}T and d^{6Se}G nucleotides. The synthetic route for d^{4Se}T oligonucleotides is shown in **Scheme 1.5** (off-column selenium incorporation). This selenium nucleophile is especially useful in that the cyanoethyl functionality can be removed in the same step as phosphate

backbone deprotection, although special care needs to be taken during this step. Deselenization of nucleoside analogs bearing the 2-cyanothyl-seleno functionality is possible when treated with strong base. This makes it necessary to use phosphoramidites which can be deprotected under ultra-mild conditions, such as 50 mM K_2CO_3 in methanol, during solid-phase synthesis.

Based on these previous findings, it should be feasible to treat a support-bound oligonucleotide containing a convertible nucleoside (with either a triazolyl or benzotriazolyl group) with 2-cyanoethyl selenide in order to achieve selenium

Scheme 1.5: Huang approach for the synthesis of oligonucleotides containing $d^{4Se}T$.



incorporation (on-column approach). This would open up a greener and more efficient synthetic route to generating selenium modified oligonucleotides for

X-ray crystallography. This strategy would also have the added advantage of producing stock oligonucleotide which could be reused for other purposes due its convertible nature. Beyond selenium incorporation, post-synthetic fluorophore probes and cross-links could be introduced, to name a few possibilities.

Inosine is ideal for this approach, as its nucleobase can be readily accommodated across either C, A, U/T, and G, thereby reducing the sequence identity requirement of nucleic acid structures. This implies that in a nucleic acid sequence, any position can be modified with minimal structural perturbation. As a proof-of-concept of this approach, it is however preferable to work with DNA. Compared to RNA, DNA is less susceptible to chemical degradation, less likely to be enzymatically degraded, less expensive and easier to synthesize, both in solution and solid-phase. Furthermore, due to previous issues faced in our lab with the post-synthetic (on-column) selenium incorporation using 2'-deoxyinosine, the methodology was first optimized using thymidine. This was practical since the synthesis and properties of $d^{4Se}T$ oligonucleotides are well established,[24,29] allowing for a straightforward evaluation of the methodology.

Chapter 2 (unpublished work) explores this topic by describing the conditions found to incorporate selenium at the C-4 position of thymidine on a support-bound oligonucleotide and retain the selenium functionality after deprotection.

In our efforts to adapt the on-column Se incorporation approach to 2'-deoxyinosine, it was first necessary to synthesize an authentic oligonucleotide containing 2'-deoxy-6-selenoinosine. Unexpectedly, a diselenide bridge was

discovered to spontaneously form between two non-complementary strands after deprotection and cleavage from the solid-support. The formation and reversibility of this cross-link, along with biophysical studies of the cross-linked oligonucleotide, will be described in **Chapter 3**.

Chapter 2.
Selenium Incorporation into a Solid Support-Bound Oligonucleotide with Convertible 4-triazolylthymidine

Abstract

We have explored a simplified methodology for the incorporation of selenium into oligonucleotides for use in X-ray crystallography. This strategy involves treating a CPG-bound oligonucleotide containing a 5'-terminal 4-triazolylthymidine with a nucleophilic selenolate species. This approach has the advantage of generating a CPG-bound oligonucleotide which can be recycled for other purposes, beyond X-ray crystallography. Conditions were also developed to retain the selenium atom during deprotection and cleavage of the oligonucleotide from the solid-support.

2.1. Introduction

Covalent modification of oligonucleotides with selenium (Se) has become the gold standard method for solving the phase problem in nucleic acid X-ray crystallography. Since Se is a heavier homologue of oxygen, almost all the oxygen sites in nucleic acids have synthetically been replaced with Se with relative ease. Classically, Se has been incorporated through the multi-step synthesis of a phosphoramidite building block bearing a Se functionality in the sugar or the nucleobase, followed by solid-phase synthesis of the oligonucleotide.[20–22,24,25] In one unique case, the non-bridging phosphate oxygen was replaced with Se to yield a phosphoroselenoate (P-Se), prepared by modification of the conventional solid-phase oxidation step by treatment of the support-bound oligonucleotide with the Se reagent KSeCN.[30] Although this methodology suffered from some limitations, including the necessity for the HPLC separation of P-Se diastereomers (effective only for short sequences) and deselenization, it was the first successful instance of on-column Se incorporation.

The Lakshman[26] and Miller[27] groups demonstrated the feasibility for on-column incorporations at the nucleobase, utilizing convertible nucleotides, 2'-deoxy-6-benzotriazolylinosine and 2'-deoxy-4-triazolyluridine, respectively. In their approaches, treatment of control-pore glass (CPG) bound oligonucleotides with a nucleophilic species caused displacement of the convertible group and efficient on-column incorporation. Furthermore, the Huang group[28] developed a method to generate the nucleophilic 2-cyanoethyl selenide species (from the

reduction of di-(2-cyanoethyl) diselenide), which has successfully been implemented in the synthesis of Se nucleobase modified phosphoramidites and oligonucleotides.[24,25]

Inspired by this previous work and the recent push towards greener chemistry, we sought out to simplify the Se incorporation process. Herein, we describe the development of an on-column nucleobase Se incorporation approach utilizing convertible 4-triazolylthymidine. Conditions such as reagent concentration, temperature, reaction time and solvent were screened. Using an authentically synthesized Se oligonucleotide standard (off-column synthesis), we have also determined optimal deprotection conditions for retaining Se in near quantitative amounts.

By incorporating Se at the nucleobase, no new chiral centers are produced and the protected Se functionality, SeEtCN, has been shown to be resistant to deselenization during solid-phase synthesis,[24] thus avoiding previous issues faced with phosphoroselenoates. Furthermore, the convertible CPG-bound oligonucleotide generated by this process can be recycled for other purposes, depending on the nucleophile used for incorporation.

2.2. Experimental

2.2.1. General Methods for the Preparation and Characterization of Nucleosides

N,N'-Diisopropylamino cyanoethyl phosphonamidic chloride (Cl-

POCEN*i*Pr$_2$) and 5'-*O*-(4,4'-dimethoxytrityl)thymidine (compound **2.1**) were purchased from ChemGenes Corporation (Wilmington, MA). 5'-*O*-Dimethoxytrityl-2'-deoxyribonucleoside-3'-*O*-(β-cyanoethyl-*N,N'*-diisopropyl) phosphoramidites (including **2.8**) and protected 2'-deoxyribonucleoside-CPG supports were purchased from Glen Research (Sterling, VA). Compounds **2.2**, **2.3**, **2.4**, **2.5** and **2.6** were synthesized following previously published procedures.[24,28] All other chemicals and solvents were purchased from the Aldrich Chemical Company (Milwaukee, WI) or EMD Chemicals Inc. (Gibbstown, NJ). Flash column chromatography was performed using silica gel 60 (230–400 mesh) purchased from Canadian Life Science (Pointe-Claire, 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. ^{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 were obtained using an 7T-LTQ FT ICR mass spectrometer (Thermo Scientific) at the Concordia University Centre for Structural and Functional Genomics. The mass spectrometer was operated in full scan, positive ion detection mode.

2.2.2. Chemical Synthesis of Nucleosides

3'-*O*-(2-cyanoethyl-*N,N'*-diisopropyl)phosphoramidityl-4-(β-2-cyanoethyl)seleno-5'-*O*-(4,4'-dimethoxytrityl)-thymidine (**2.7**)

To a stirred solution of compound **2.6** (0.264 g, 0.400 mmol) in anhydrous THF (2 ml) was added dropwise DIPEA (0.209 ml, 1.20 mmol) followed by the dropwise addition of Cl-POCEN*i*Pr_2 (0.177 ml, 0.799 mmol). After 30 minutes the solvent was evaporated in vacuo. The content was diluted with EtOAc (50 ml) and washed with 3 % (aq., w/v) $NaHCO_3$ (2 x 50 ml) and brine (50 ml). The organic layer was dried over anhydrous Na_2SO_4, decanted and the solvent evaporated in vacuo. Crude material was purified by short flash column chromatography using EtOAc/hexanes (4:1, v/v) (with 0.1 % NEt_3, v/v) to afford 0.295 g (86 %) of compound **2.7** as a colorless foam. **^{31}P NMR** (202.3 MHz, d_6-acetone, ppm): 148.40 and 148.16.

3'-*O*-(2-cyanoethyl-*N,N'*-diisopropyl)phosphoramidityl-4-(1,2,4-triazolyl)-5'-*O*-(4,4'-dimethoxytrityl)-thymidine (**2.9**)

To a stirred solution of compound **2.8** (0.516 g, 0.693 mmol) in anhydrous DCM/ACN (1:1, v/v) (10 ml) was added 1,2,4-triazole (2.79 g, 40.42 mmol) followed by dropwise TEA (6.09 ml, 43.63 mmol). The reaction was cooled to 0^{o}C on ice, then dropwise $POCl_3$ (0.813 ml, 8.73 mmol) was added. After 1 hour the reaction flask was removed from the ice bath and allowed to stir at room temperature for 1 additional hour. The solvent was evaporated in vacuo. The content was diluted with DCM (100 ml) and washed with 3 % (aq., w/v) $NaHCO_3$ (3 x 100 ml). The organic layer was dried over anhydrous Na_2SO_4, decanted and

the solvent evaporated in vacuo. Crude material was purified by short flash column chromatography using EtOAc/hexanes (4:1, v/v) (with 0.1 % NEt_3, v/v) to afford 0.452 g (82 %) of compound **2.9** as a colorless foam. **^{31}P NMR** (202.3 MHz, d_6-acetone, ppm): 148.48, 148.23.

2.2.3. Oligonucleotide Solid-phase Synthesis

All oligonucleotide sequences were assembled with an Applied Biosystems Model 3400 synthesizer on a 2 μmol scale using standard β-cyanoethyl phosphoramidite cycles supplied by the manufacturer with slight modifications to coupling times as described below. The Se containing mixed-base oligonucleotide (**ODN$_2$**) was prepared using commercially available 3'-*O*-2'-deoxynucleoside phosphoramidites, containing "fast-deprotecting" groups (N6-phenoxyacetyl-2'-deoxyadenosine, N4-acetyl-2'-deoxycytidine and N2-phenoxyacetyl-2'-deoxyguanosine), were dissolved in anhydrous ACN to a concentration of 0.1 M. The modified 3'-*O*-2'-deoxynucleoside phosphoramidites (compounds **2.7 and 2.9**) were dissolved to a concentration of 0.15 M. All other oligonucleotides were prepared using a commercially available 3'-*O*-thymidine phosphoramidite, which was dissolved in anhydrous ACN to a concentration of 0.1 M. Oligonucleotide sequence assembly was carried out as previously described.[31,32] The capping step for the assembly of **ODN$_2$** was carried out using phenoxyacetic anhydride/pyridine/THF 1:1:8 (v/v/v; solution A) and 1-methyl-imidazole/THF 16:84 (w/v; solution B). All other oligonucleotides used acetic anhydride/pyridine/THF 1:1:8 (v/v/v, solution A) and 1-methyl-imidazole/THF

16:84 (w/v; solution B). The coupling time for phosphoramidites **2.7 and 2.9** were extended to 600 seconds (compared to 120 seconds for the commercially available phosphoramidites). All oligonucleotides had their 5'-terminal trityl groups removed on the synthesizer by acid treatment.

2.2.4. Deprotection Optimizations of Selenium Containing Oligonucleotides

The solid-phase column was removed from the synthesizer and the CPG-bound oligonucleotide was transferred to a 2 ml Teflon lined screw cap vial. The various conditions and reagents screened for the optimization of the deprotection and cleavage of the Se containing oligonucleotides from the solid-support are summarized in **Table 2.1**. After the incubation period, the supernatant was transferred to a separate, clean screw cap vial and the CPG washed twice with 200 µl of aqueous ACN (50 %, v/v). The crude oligonucleotides were lyophilized in a speed-vac concentrator.

2.2.5. Purification of Selenium Containing Oligonucleotides by Denaturing PAGE

The gel consisted of 20 % acrylamide (19:1, w/w, acrylamide to bis-acrylamide) and 7 M urea in a pH 8.0 TBE buffer (89 mM Tris base, 89 mM boric acid, 0.2 mM ethylenediaminetetraacetic acid (EDTA)) with dimensions 20 cm x 20 cm x 0.75 cm (L x H x W). The gels were run for 2 hours at a constant 10 W using the TBE buffer described above. Markers for the gel were a mixture of

bromophenol blue (8 nucleotides) and xylene cyanol FF (28 nucleotides). Oligonucleotides were prepared by lyophilizing up to 25 OD_{260} of sample and dissolving them in 100 µl formamide. The gel was visualized by placing it on a TLC plate (Merck, Kieselgel 60 F_{254}, 0.25 mm) which was irradiated with 254 nm light (UV shadowing). The band of interest was then excised from the gel, suspended in 0.1 M sodium acetate and shaken gently overnight. The purified oligonucleotides were desalted using C-18 SEP PAK cartridges (Waters).

2.2.6. Quantification of Selenium Containing Oligonucleotides

After the deprotection trials, each oligonucleotide was purified as stated previously and then analyzed by reverse phase (RP) HPLC. The Symmetry® C-18 5 µm column (0.46 x 15 cm) was eluted with a linear gradient of 0-70 % buffer B over 30 min (buffer A: 50 mM sodium phosphate, pH 5.8, 2 % ACN and buffer B: 50 mM sodium phosphate, pH 5.8, 50 % ACN). The column was monitored at 260 nm and 369 nm. The Se containing oligonucleotide was identified by its characteristic absorbance at 369 nm and the yield estimated from the area under the corresponding peak at 260 nm.

After the on-column Se incorporation trials, the crude oligonucleotide was analyzed by SAX HPLC. The Dionex DNAPAC PA-100 column (0.4 cm x 25 cm), purchased from Dionex Corp (Sunnyvale, CA), was eluted at room temperature using a linear gradient of 0-52 % buffer B over 30 minutes (buffer A: 100 mM Tris HCl, pH 7.5, 10 % ACN; buffer B: 100 mM Tris HCl, pH 7.5, 10 % ACN, 1 M NaCl). The column was monitored at 260 nm and 369 nm. The Se containing

oligonucleotide was identified by its characteristic absorbance at 369 nm and the yield estimated from the area under the corresponding peak at 260 nm.

2.2.7. Oligonucleotide Characterization by ESI-MS

ESI mass spectra for oligonucleotides were acquired at the Concordia University Centre for Biological Applications of Mass Spectrometry (CBAMS) using a Micromass Qtof2 mass spectrometer (Waters) equipped with a nanospray ion source. The mass spectrometer was operated in full scan, negative ion detection mode and the raw data were deconvoluted.

2.2.8. On-Column Selenium Incorporation Trials

Method A: The Se nucleophile was first generated in a round bottom flask. The solution was transferred to a syringe which was attached to the solid-phase column containing the CPG-bound oligonucleotide. The solution was pushed through the column over the specified time.

Method B: The Se nucleophile was first generated in a round bottom flask. The solution was transferred to a flame-dried glass vial containing the CPG-bound oligonucleotide and sealed.

Method C: Before the Se nucleophile treatment, neat triethylamine (TEA) was washed through the solid-phase column containing the CPG-bound oligonucleotide over 2 hours. The column was washed twice with ACN (2x 10 ml). Next, the column was placed under vacuum to dry for 1 hour.

Method D: The generation of the Se nucleophile was done in a glass vial already containing the CPG-bound oligonucleotide.

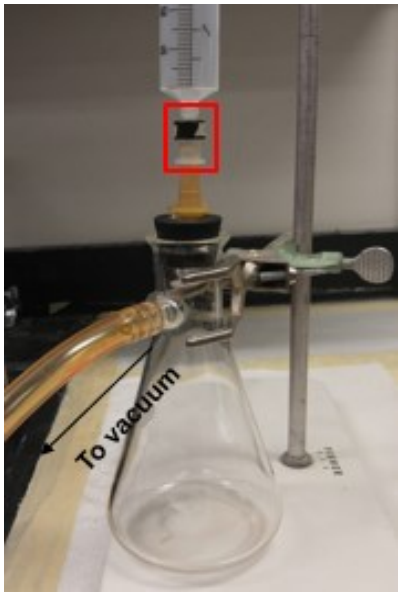
Method E: The Se nucleophile was first generated in a round bottom flask. 1 ml of the solution was added to a syringe and washed through the solid-phase column containing the CPG-bound oligonucleotide over 20 minutes. Another 1 ml of the Se nucleophile was then washed through the column over 20 minutes. This was repeated for a total of 2 hours.

Method F: The Se nucleophile was first generated in a round bottom flask. The solution was transferred to a screw-cap vial containing the CPG-bound oligonucleotide and sealed.

When 4 Å molecular sieves (MS) are specified, these were placed in the round bottom flask during the Se nucleophile synthesis as well as the vessel containing the CPG-bound oligonucleotide.

After each Se nucleophile treatment method, the CPG was transferred to a clean solid-phase column on vacuum (see apparatus in **Figure 2.1**). ACN (2x 10 ml) was washed through the column (shown in the red box in **Figure 2.1**). Next, the column containing the washed CPG was placed under vacuum to dry. The CPG was then transferred to a screw cap vial and deprotected with *t*-butylamine/H_2O (1:3, v/v) at 60 ^{o}C for 6 hours, followed by equimolar acetic acid treatment. After the incubation period, the supernatant was transferred to a separate, clean screw cap vial and the CPG washed twice with

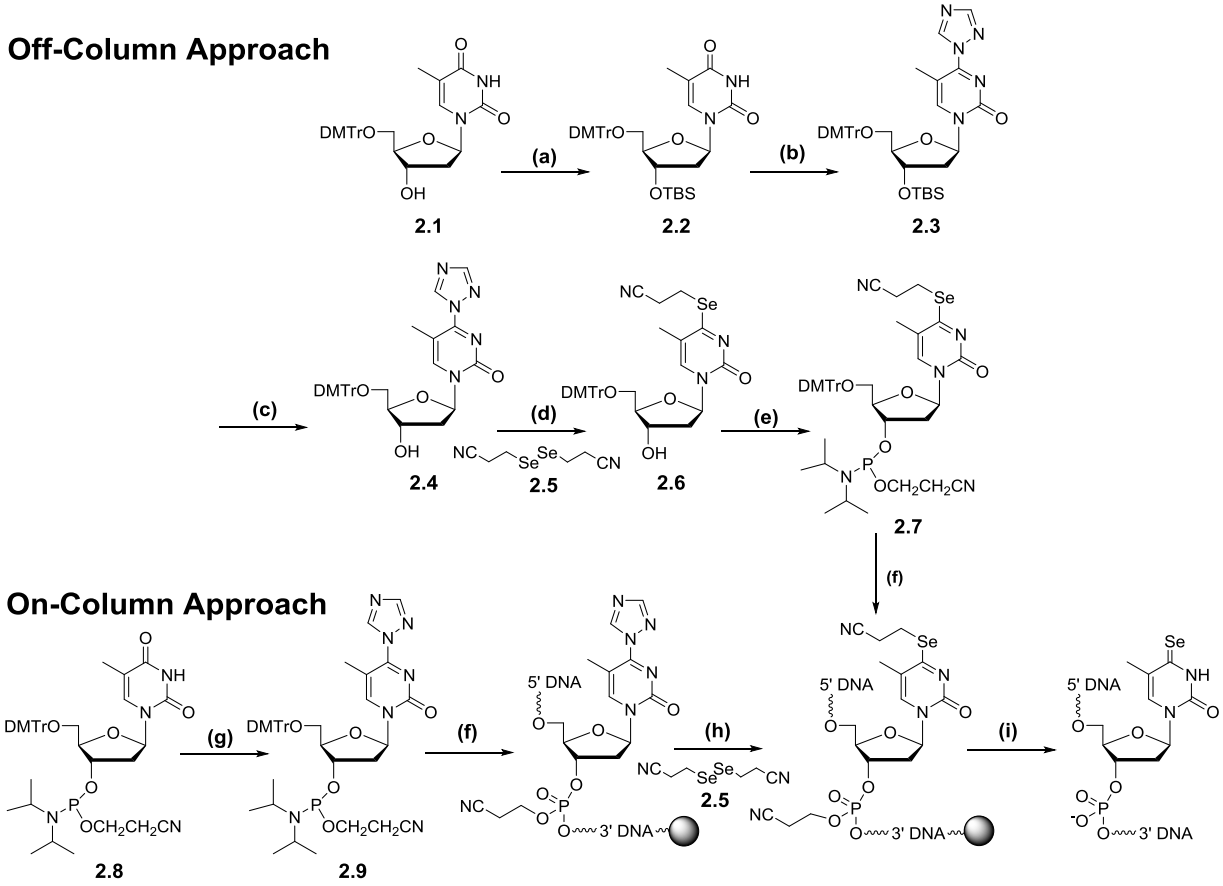
Figure 2.1: Setup for collecting and washing CPG-bound oligonucleotides.



200 µl aqueous ACN (50 %, v/v). The crude oligonucleotides were lyophilized in a speed-vac concentrator.

2.3. Results and Discussion

To ensure the deprotection condition used caused minimal deselenization



Scheme 2.1: On- and off-column approaches for the synthesis of $d^{4Se}T$ containing oligonucleotides.

Reagents and conditions: (a) TBS-Cl, DMAP, Imidazole, DCM, overnight, rt; (b) 1,2,4-triazole, TEA, $POCl_3$, DCM/ACN (1:1, v/v), 4 h, 0 °C to rt; (c) TREAT-HF, THF, overnight, 40 °C; (d) $NaBH_4$, EtOH, 1 h, 0 °C to rt; (e) Cl-POCEN*i*Pr_2, DIPEA, THF, 30 mins, rt; (f) Solid-phase synthesis (see experimental section); (g) 1,2,4-triazole, TEA, $POCl_3$, DCM/ACN (1:1, v/v), 3 h, 0 °C to rt; (h) On-column Se incorporation; (i) Deprotection.

or undesirable adduct formation, the $d^{4SeEtCN}T$ phosphoramidite **2.7** was synthesized following the Huang protocol[29] and incorporated into an oligonucleotide sequence by solid-phase synthesis (off-column approach in **Scheme 2.1**). First, the 3'-position of commercially available 5'-*O*-dimethoxytrityl-thymidine **2.1** was protected with a TBS group, generating the protected nucleoside **2.2**, with a yield of 94 %. Convertible nucleoside **2.3** was synthesized by adding a triazolyl group to the C-4 position of protected nucleoside **2.2**. The TBS group of unpurified convertible nucleoside **2.3** was removed, producing the convertible nucleoside **2.4**, with a yield of 96 %. Reduction of diselenide[28] **2.5** by sodium borohydride, generating nucleophilic selenolate, allowed for the displacement of the triazolyl group to produce pre-amidite **2.6**, with a yield of 71 %. Phosphitylation of pre-amidite **2.6** afforded phosphoramidite **2.7**, with a yield of 86 %. Phosphoramidite **2.7** was incorporated into the support-bound, protected oligonucleotide sequence, 5'-$dT_4(d^{4SeEtCN}T)dT_5$-3' (**ODN_1**), by solid-phase synthesis (all trityl values above 90 %). This sequence was chosen in order to use the simplest model sequence possible during the deprotection trials, as thymidine nucleotides generally do not require nucleobase protection, unlike the other canonical nucleotides.

The results of the deprotection trials of sequence **ODN_1** (optimization of step (i) in **Scheme 2.1**) are summarized in **Table 2.1**. The commonly used treatment of NH_4OH/EtOH (3:1, v/v) was found to be too harsh, causing complete deselenization and conversion of the modified nucleobase to 5-methyl-cytosine. Treatment with 50 mM K_2CO_3/MeOH, did retain the Se atom in low yield, but

caused methoxy adduct formation. In an attempt to prevent this adduct formation, the bulkier alcohol, *i*PrOH, was chosen to generate the nucleophile for deprotection with K_2CO_3. There were, however, solubility issues with the K_2CO_3 and no cleavage of the oligonucleotide from the solid-support was observed. Neutralization of the base after the incubation period with an equimolar amount of acetic acid resulted in a higher yield of Se containing oligonucleotide, however adduct formation was still observed. Finally, treatment with *t*-butylamine/H_2O (1:3, v/v) followed by acetic acid neutralization completely abolished adduct formation and caused minimal deselenization (**Figures A2.7** and **A2.8**). The utility of this deprotection condition was then evaluated for the mixed-base oligonucleotide sequence, 5'-dCGCGAA(d^{4SeEtCN}T)dTCGCG-3' (**ODN$_2$**). The same outcome was observed, as the Se containing oligonucleotide was the major species present (**Figures A2.9** and **A2.10**), validating the use of this deprotection condition for mixed-base sequences.

Table 2.1: Deprotection trials of ODN$_1$ in order to retain Se

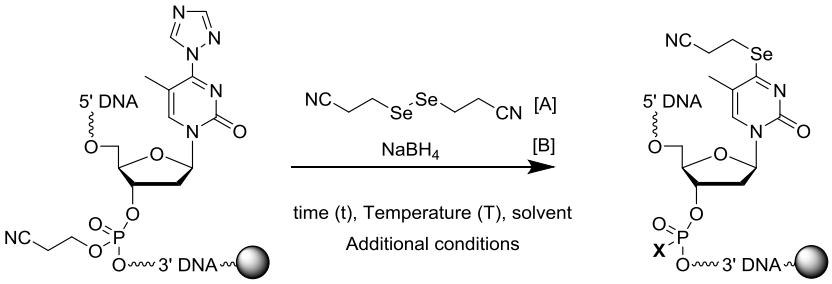
Entry	Sequence	Deprotection Conditions			Additional Step	Yield of Se Oligo (%)*
		Reagents	t (h)	T (^oC)		
1	**ODN$_1$**	NH_4OH/EtOH (3:1)	4	55	-	0
2		50 mM K_2CO_3/MeOH	4	23 (rt)	-	12
3		50 mM K_2CO_3/*i*PrOH	4	23 (rt)	-	0
4		50 mM K_2CO_3/MeOH	4	23 (rt)	100 % Glacial Acetic Acid	43
5		*t*-butylamine/H_2O (1:3)	6	60	-	60
6		*t*-butylamine/H_2O (1:3)	6	60	100 % Glacial Acetic Acid	100

* Relative yield estimated by RP HPLC

With the deprotection step optimized for optimal retention of the Se atom,

on-column Se incorporation trials were performed (optimization of step (h) in **Scheme 2.1**). First, it was necessary to synthesize the convertible thymidine phosphoramidite. Inspired by previous work,[33] phosphoramidite **2.9** was synthesized in one step, with a yield of 82 %. Phosphoramidite **2.9** was first introduced into the sequence 5'-$dT_4(d^{4Tri}T)dT_5$-3' (**ODN_3**), which upon successful Se incorporation would generate the same Se oligonucleotide sequence as deprotected **ODN_1**. Various methods and conditions for Se incorporation into support-bound **ODN_3** were performed. In each trial, it was necessary to reduce diselenide **2.5** to the reactive nucleophilic selenolate species. This was typically achieved by following the Huang protocol,[28] using sodium borohydride in ethanol.

Table 2.2: On-column Se incorporation trials



Entry	Sequence	Solvent	t (h)	T (^oC)	[A] (M)	[B] (M)	Additional Conditions			Yield of Se Oligo (%)*
							Under Argon	4 Å MS	Stirring	
1^A	**ODN_3**	99 % Ethanol	0.5	rt	0.15	0.30	N/A**	N/A**	N/A**	0
2^A			1	rt	0.15	0.30				0
3^A			2	rt	0.15	0.30				0
4^A			18.5	rt	0.15	0.30				0
5^{CE}			2	rt	0.15	0.30				0
6^B			1	rt	0.15	0.30	Yes	No	No	0
7^B			1	rt	0.15	0.30	Yes	No	Yes	0
8^B			1	40	0.15	0.30	Yes	No	Yes	0
9^B			1	60	0.15	0.30	Yes	No	No	0
10^B			4	60	0.15	0.30	Yes	No	No	0
11^{BC}			1	60	0.15	0.30	Yes	No	No	0

12[B]			0.5	60	0.15	0.90	Yes	No	No	0
13[B]			26	60	0.15	0.90	Yes	No	No	0
14[BC]			26	60	0.15	0.90	Yes	No	No	0
15[BD]			1	rt	0.10	0.20	Yes	Yes	Yes	0
16[F]			22	-20	0.20	0.40	No	No	No	0
17[BD]		Degassed THF	4	40	0.50	0.20	Yes	Yes	Yes	0
18[BC]		Anhydrous pyridine	19	60	0.15	0.30	Yes	Yes	No	0
19[F]	**ODN_4**	99 % Ethanol	22	-20	0.20	0.40	No	No	No	29
20[F]			48	-20	0.20	0.40	No	No	No	0
21[F]			48	rt	0.20	0.40	No	No	No	0
22[F]			48	40	0.20	0.40	No	No	No	0
23[B]			4	-20	0.20	0.40	No	Yes	No	0
24[B]			22	-20	0.20	0.40	No	Yes	No	0
25[F]			22	-20	0.20	0.40	No	Yes	No	trace
26[F]			1	-20	0.20	0.40	No	No	No	0
27[F]			4	-20	0.20	0.40	No	No	No	5
28[F]	**ODN_5**	99 % Ethanol	22	-20	0.20	0.40	No	No	No	20
29[F]			4	-20	0.20	0.40	No	No	No	7.5

***** Relative yield estimated by SAX HPLC
** N/A: Not applicable
A: CPG in solid-phase column treated with Se nucleophile solution
B: CPG in dried glass reaction vial treated with Se nucleophile solution
C: Prior to the Se nucleophile reaction, treated CPG by washing TEA through the solid-phase column over 2 hours
D: Se nucleophile generated *in situ* in glass reaction vial containing CPG
E: Washed Se nucleophile solution through solid-phase column containing CPG, adding more equivalencies every 20 minutes
F: CPG in screw cap vial treated with Se nucleophile solution

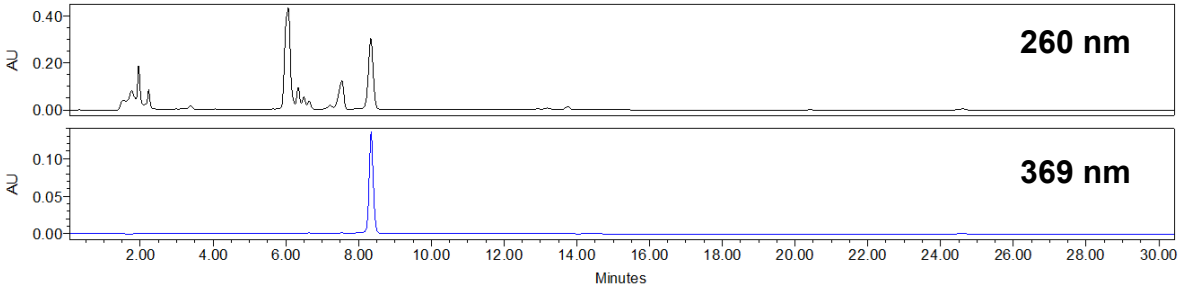
In some trials (entries 5, 11, 14 and 18), CPG-bound **ODN_3** was first treated with TEA before treatment with the Se nucleophile. This was done in order to remove the cyanoethyl groups from the phosphate backbone, generating a negatively charged backbone which would be more resistant to cleavage from the Se nucleophile. The Se containing oligonucleotide species was identified by its characteristic absorbance at 369 nm (**Figure A2.11**), providing a rapid means to identify these oligonucleotides. None of the trials with **ODN_3** proved fruitful,

although from the mass spectrometry analysis of the crude oligonucleotide after deprotection, identification of the side-products allowed for more focused optimizations going forward. Competing nucleophiles, specifically H_2O, hydride and *t*-butylamine were believed to be responsible for these side-products (**Figure A2.12**). The *t*-butylamine adduct was interesting as it was never observed during the deprotection optimizations, however it was believed to form due to the presence of unreacted triazole after treatment with the Se nucleophile. Furthermore, oxidation of the Se nucleophile produces water as a side-product. Using degassed THF as the solvent in order to lower the O_2 content was attempted, although this did not help, potentially due to solubility issues.

After failed attempts with **ODN_3**, phosphoramidite **2.9** was incorporated at the 5'-terminal end of a shorter sequence, 5'-($d^{4Tri}T$)dT_4-3' (**ODN_4**). It was suspected that with the reactive triazolyl group at an internal position of an oligonucleotide, as in **ODN_3**, steric hindrance may have been preventing proper Se incorporation. After Se nucleophile treatment of CPG-bound **ODN_4** in a screw

Figure 2.2: SAX HPLC profiles of crude ODN_4 after on-column Se nucleophile treatment and deprotection (entry 19).

The column was eluted at room temperature using a linear gradient of 0-52 % buffer B over 30 minutes (buffer A: 100 mM Tris HCl, pH 7.5, 10 % ACN; buffer B: 100 mM Tris HCl, pH 7.5, 10 % ACN, 1 M NaCl).



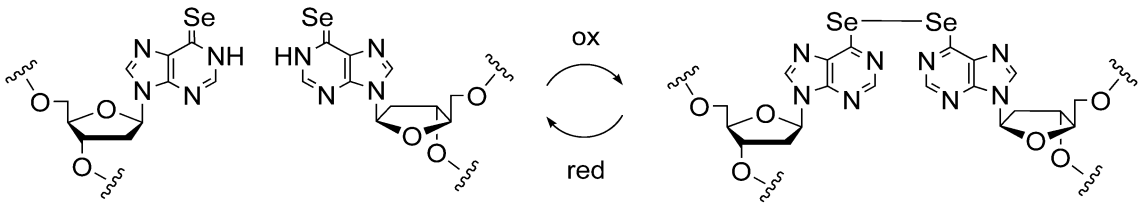
cap vial at -20 ^{o}C and deprotection (entry 19), successful on-column Se incorporation was achieved and verified by SAX HPLC (**Figure 2.2**). From the mass spectrometry analysis, the species formed by hydride attack was now absent (**Figure A2.13**). These colder conditions were repeated with **ODN$_3$** as well as another 10-mer, 5'-(d^{4Tri}T)dT$_9$-3' (**ODN$_5$**), in order to confirm if it was the position of the convertible nucleotide or the length of the oligonucleotide which was the main factor preventing on-column Se incorporation. Following treatment with the Se nucleophile, no Se containing oligonucleotide was detected when using **ODN$_3$**, however there was a 20 % yield when using **ODN$_5$**. This suggests the position of the convertible nucleotide is vital for proper on-column incorporation under these conditions. Finally, lowering the reaction time to 4 hours at -20 ^{o}C significantly lowered the yield of Se oligonucleotide.

2.4. Conclusion

In summary, modified oligonucleotides bearing d^{4Se}T nucleotides were isolated by two convergent approaches. Deprotection of an authentically synthesized Se oligonucleotide with *t*-butylamine/H_2O (1:3) at 60 ^{o}C for 6 hours, followed by the addition of equimolar acetic acid, proved to be the optimal conditions for retention of Se in homothymidine and mixed-base sequences. The on-column Se incorporation approach, although resulting in a lower yield of Se oligonucleotide, was shown to be viable in this study for sequences with d^{4Tri}T and the 5'-terminal end. The position of the convertible group in the oligonucleotide sequence was found to be more important than the length of the

oligonucleotide, however, a lower yield of Se oligonucleotide was found when exposing the longer sequence, **ODN_5**, with the same conditions as **ODN_4**, suggesting length does still play a role in efficient incorporation. Further optimizations for on-column Se incorporation should be explored in order to increase the yield of Se oligonucleotide and make this a viable approach. Potentially, by working in a glove box under an inert atmosphere, oxidation of the Se nucleophile would be limited, resulting in a more efficient reaction with the convertible nucleotide residue. In addition, reaction with water as a competing nucleophile would be diminished.

Chapter 3.
Reversible Diselenide Cross-links are Formed Between Oligonucleotides Containing 2'-Deoxy-6-Selenoinosine



Graphical Abstract:

Reversibility of the diselenide bridge under oxidizing and reducing conditions

Published as:

Liczner, C.; Grenier, V.; Wilds, C. *Tetrahedron Letters*. **2018**, *59*, 38-41.

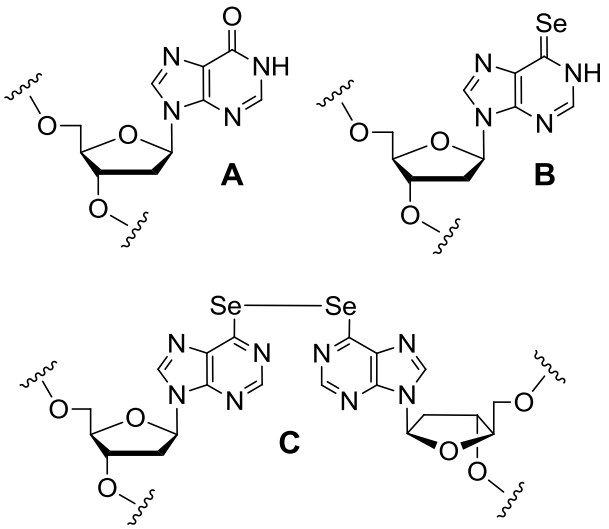
Abstract

We have synthesized and characterized a phosphoramidite derivative of 2'-deoxy-6-selenoinosine ($d^{6Se}I$) and incorporated this modification into an oligonucleotide by solid-phase synthesis. During cleavage from the solid-support and deprotection, spontaneous dimerization of this oligonucleotide occurs via formation of a diselenide cross-link between the modified nucleobases. This cross-link can be readily reduced to restore the single-stranded oligonucleotide. UV thermal denaturation and circular dichroism spectroscopy of duplexes with $d^{6Se}I$ paired against all four native nucleobases revealed minor differences in stability and structure relative to 2'-deoxyinosine. This selenium containing nucleobase modification may be useful for applications in DNA nanomaterials and X-ray crystallography.

3.1. Introduction

Selenium (Se) derivatized nucleic acids, prepared by either enzymatic or solid-phase synthesis, are a powerful tool in solving the phase problem of X-ray crystallography due to the anomalous scattering properties of this element.[15] The substitution of oxygen (O) with Se for this purpose within the nucleic acid scaffold has been reported at the phosphate backbone,[23] the 2', 4' and 5' positions of the sugar[21,22,20] and at the nucleobases.[24,25] Incorporation of Se at the nucleobase has advantages over other sites, including minimal structural perturbation, avoiding the introduction of an additional chiral center which would necessitate the separation of diastereomers and that incorporation is not limited at the termini of the oligonucleotide. Substitution of the O-6 atom of guanine[25] and the O4-atom of thymine[24] with Se are examples of nucleoside modifications that have been prepared and incorporated into oligonucleotides. The crystal structures of Se-nucleobase modified oligonucleotides have been shown to be virtually identical to that of their native counterparts.[15] In addition, a nucleic acid-protein complex consisting of RNase H/Se-DNA/RNA, has been solved utilizing

Figure 3.1: Structures of (A) 2'-deoxyinosine, (B) 2'-deoxy-6-selenoinosine, and (C) diselenide bridged 2'-deoxy-6-selenoinosine.



nucleobase Se incorporations.[34]

2'-Deoxyinosine (**Figure 3.1A**), containing the base hypoxanthine, is capable of pairing with the other natural nucleobases found in DNA.[35] Analogs of 2'-deoxyinosine have been used to introduce site-specific post-synthetic modifications within an oligonucleotide[36] and for applications as fluorescent probes.[37]

Herein, the preparation of a 2'-deoxy-6-selenoinosine ($d^{6Se}I$, **Figure 3.1B**) phosphoramidite is described. This was achieved by using a convertible 2'-deoxyinosine nucleoside[26] and the Huang reagent (di(2-cyanoethyl) diselenide).[28] The Se containing phosphoramidite was incorporated into an oligonucleotide by solid-phase synthesis, deprotected and purified by strong anion exchange (SAX) HPLC to prepare the sequence 5'-dGGCT^{6Se}IGATCACCAG-3' (**dI-Se-ODN**). Interestingly, during the deprotection of the oligonucleotide from the controlled pore glass (CPG) solid-support, the serendipitous discovery of a cross-linked oligonucleotide homodimer (**dI-Se-ODN$_2$**) connected by a diselenide bridge (**Figure 3.1C**) between the modified nucleobases was made. This is the first instance of a Se derivatized nucleic acid where two oligonucleotides are cross-linked between the nucleobases with a diselenide linkage. **dI-Se-ODN$_2$** was studied by UV thermal denaturation, denaturing polyacrylamide gel electrophoresis (PAGE) and circular dichroism (CD) spectroscopy. Time trials of oligonucleotide dimer formation after deprotection were monitored by SAX HPLC.

3.2. Experimental

3.2.1. General Methods for the Preparation and Characterization of Nucleosides

N,N'-Diisopropylamino cyanoethyl phosphonamidic chloride (Cl-POCEN*i*Pr$_2$) and 5'-*O*-dimethoxytrityl-2'-deoxyinosine (compound **3.1**) were purchased from ChemGenes Corporation (Wilmington, MA). 5'-*O*-Dimethoxytrityl-2'-deoxyribonucleoside-3'-*O*-(β-cyanoethyl-*N,N'*-diisopropyl) phosphoramidites and protected 2'-deoxyribonucleoside-CPG supports were purchased from Glen Research (Sterling, VA). Compounds **3.2** and **3.3** were synthesized following previously published procedures.[26,28] All other chemicals and solvents were purchased from the Aldrich Chemical Company (Milwaukee, WI) or EMD Chemicals Inc. (Gibbstown, NJ). Flash column chromatography was performed using silica gel 60 (230–400 mesh) purchased from Canadian Life Science (Pointe-Claire, 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. ^{77}Se NMR spectra (^{1}H decoupled) were recorded at a frequency of 95.3 MHz and chemical shifts were reported in ppm

with dimethyl selenide used as an external standard. ^{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 were obtained using an 7T-LTQ FT ICR mass spectrometer (Thermo Scientific) at the Concordia University Centre for Structural and Functional Genomics. The mass spectrometer was operated in full scan, positive ion detection mode.

3.2.2. Chemical Synthesis of Nucleosides

6-(2-cyanoethyl)seleno-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxyinosine (**3.4**)

In separate round bottom flasks, diselenide **3.3** (0.475 g, 1.79 mmol) was dissolved in 99 % EtOH (10 ml), $NaBH_4$ (0.169 g, 4.45 mmol) was suspended in 99 % EtOH (2.5 ml) and compound **3.2** (0.294 g, 0.438 mmol) was dissolved in anhydrous THF (2.5 ml). The atmosphere in all three flasks was exchanged with argon. The flask containing diselenide **3.3** (yellow solution) was stirred and immersed in an ice bath to cool for 10 minutes. The $NaBH_4$ suspension was injected dropwise over 5 minutes. After the cessation of gas evolution, the flask was removed from the ice bath (solution now clear). The solution of compound **3.2** was then injected dropwise. After 2 hours the reaction mixture was diluted with EtOAc (100 ml) and washed with brine (2 x 100 ml). The organic layer was dried over anhydrous Na_2SO_4, decanted and the solvent evaporated in vacuo. Crude material was purified by flash column chromatography using a gradient of MeOH/DCM (0 % to 3 % increased by increments of 0.5 %, v/v) to afford 0.192 g

(66 %) of compound **3.4** as a colorless foam. $\mathbf{R_f}$ (SiO_2 TLC): 0.64 (9:1 EtOAc/hexanes, v/v). $\boldsymbol{\lambda_{max\ (ACN)}}$: 286 nm. **^{1}H NMR** (500 MHz, $CDCl_3$, ppm): δ 8.63 (s, 1H, H8), 8.16 (s, 1H, H2), 7.40-7.38 (m, 2H, Ar), 7.29-7.19 (m, 7H, Ar), 6.79 (m, 4H, Ar), 6.47 (dd, 1H, H1' *J* = 6.5 Hz), 4.70 (m, 1H, H3'), 4.17 (m, 1H, H4'), 3.78 (s, 6H, 2x OCH_3), 3.55 (t, 2H, CH_2-Se, *J* = 7.4 Hz), 3.44-3.37 (m, 2H, H5'', H5'), 3.02 (t, 2H, CH_2-CN, *J* = 7.3 Hz), 2.86 (m, 1H, H2'), 2.56 (m, 1H, H2''), 2.49 (s, 1H, OH). **^{13}C NMR** (125.7 MHz, $CDCl_3$, ppm): 158.58, 157.27, 151.92, 147.72, 144.43, 141.95, 135.59, 135.56, 134.57, 129.98, 128.04, 127.90, 126.97, 118.85, 113.20, 86.64, 86.25, 84.64, 72.48, 63.66, 55.24, 40.20, 19.34, 18.49. **^{77}Se NMR** (95.3 MHz, $CDCl_3$, ppm): 323.60. **IR** (thin film); v_{max} (cm^{-1}) 3354, 2932, 2251, 1608, 1566, 1508, 1333, 1251, 1178, 1033. **HRMS** (ESI-MS) *m/z* calculated for $C_{34}H_{33}N_5O_5NaSe^+$: 694.1539; found 694.1548 $[M + Na]^+$.

3'-*O*-(2-cyanoethyl-*N,N'*-diisopropyl)phosphoramidityl-6-(β-2-cyanoethyl)seleno-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxyinosine (**3.5**)

To a stirred solution of compound **3.4** (0.139 g, 0.207 mmol) in anhydrous THF (2 ml) was added dropwise DIPEA (0.108 ml, 0.620 mmol) followed by the dropwise addition of Cl-POCEN*i*Pr_2 (0.092 ml, 0.416 mmol). After 30 minutes the solvent was evaporated in vacuo. The content was diluted with EtOAc (50 ml) and washed with 3 % (aq., w/v) $NaHCO_3$ (2 x 50 ml) and brine (50 ml). The organic layer was dried over anhydrous Na_2SO_4, decanted and the solvent evaporated in vacuo. Crude material was purified by short flash column chromatography using EtOAc/hexanes (4:1, v/v) (with 0.1 % NEt_3, v/v) to afford

0.158 g (88 %) of compound **3.5** as a colorless foam. $\mathbf{R_f}$ (SiO_2 TLC): 0.89, 0.84 (4:1 EtOAc/hexanes, v/v). $\boldsymbol{\lambda_{max\ (ACN)}}$: 286 nm. **^{1}H NMR** (500 MHz, d_6-acetone, ppm): δ 8.64 (s, 0.5H, H8), 8.63 (s, 0.5H, H8), 8.49 (s, 0.5H, H2), 8.48 (s, 0.5H, H2), 7.46-7.43 (m, 2H, Ar), 7.34-7.18 (m, 7H, Ar), 6.85-6.79 (m, 4H, Ar), 6.60-6.56 (m, 1H, H1'), 5.03-4.94 (m, 1H, H3'), 4.37-4.30 (m, 1H, H4'), 3.97- 3.63 (m, 12H, CH_2OP, 2x NCH, 2x OCH_3, CH_2-Se), 3.48-3.39 (m, 2H, H5', H5''), 3.29-3.22 (m, 1H, H2'), 3.16-3.13 (m, 2H, $OCH_2C\underline{H_2}CN$), 2.79-2.66 (m, 3H, $SeCH_2C\underline{H_2}CN$, H2''), 1.24-1.14 (m, 12H, 4 x CH_3). **^{13}C NMR** (125.7 MHz, d_6-acetone, ppm): 158.69, 158.67, 157.02, 151.57, 147.89, 147.87, 145.14, 143.59, 143.58, 135.87, 135.85, 135.80, 134.76, 134.74, 130.08, 130.05, 130.04, 130.00, 128.09, 128.04, 127.63, 127.62, 126.64, 126.62, 118.98, 118.09, 117.96, 112.95, 112.93, 86.14, 85.80, 85.76, 85.64, 85.59, 84.94, 84.93, 74.00, 73.87, 73.41, 73.28, 63.72, 63.54, 58.87, 58.77, 58.72, 58.62, 54.66, 54.64, 43.15, 43.12, 43.05, 43.02, 38.23, 38.20, 38.18, 38.14, 24.08, 24.06, 24.04, 24.02, 24.00, 23.98, 19.95, 19.89, 19.84, 18.70, 18.40. **^{31}P NMR** (202.3 MHz, d_6-acetone, ppm): 148.31, 148.20. **^{77}Se NMR** (95.3 MHz, d_6-acetone, ppm): 334.39. **IR** (thin film); v_{max} (cm^{-1}) 2965, 2927, 2365, 2338, 2251, 1608, 1560, 1508, 1333, 1250, 1179, 1077, 1034. **HRMS** (ESI-MS) *m/z* calculated for $C_{43}H_{51}N_7O_6PSe^+$: 872.2798; found 872.2795 $[M + H]^+$.

3.2.3. Oligonucleotide Solid-phase Synthesis

All oligonucleotide sequences were assembled with an Applied Biosystems Model 3400 synthesizer on a 2 μmol scale using standard β-

cyanoethyl phosphoramidite cycles supplied by the manufacturer with slight modifications to coupling times as described below. Se containing oligonucleotides were prepared using commercially available 3'-*O*-2'-deoxynucleoside phosphoramidites, containing "fast-deprotecting" groups (N6-phenoxyacetyl-2'-deoxyadenosine, N4-acetyl-2'-deoxycytidine and N2-phenoxyacetyl-2'-deoxyguanosine), which were dissolved in anhydrous ACN to a concentration of 0.1 M. The modified 3'-*O*-2'-deoxynucleoside phosphoramidite (compound **3.5**) used in the preparation of these oligonucleotides was dissolved to a concentration of 0.15 M. All other oligonucleotides were prepared using commercially available 3'-*O*-2'-deoxynucleoside phosphoramidites, containing standard protecting groups (N6-benzoyl-2'-deoxyadenosine, N4-benzoyl-2'-deoxycytidine and N2-isobutyryl-2'-deoxyguanosine), which were dissolved in anhydrous ACN to a concentration of 0.1 M. Oligonucleotide sequence assembly was carried out as previously described.[31,32] The capping step of the assembly was carried out using phenoxyacetic anhydride/pyridine/THF 1:1:8 (v/v/v; solution A) and 1-methyl-imidazole/THF 16:84 (w/v; solution B). The coupling time for phosphoramidite **3.5** was extended to 600 seconds (compared to 120 seconds for the commercially available phosphoramidites). All oligonucleotides had their 5'-terminal trityl groups removed on the synthesizer by acid treatment.

3.2.4. Oligonucleotide Deprotection and Purification

The solid-phase column was removed from the synthesizer and the CPG-bound oligonucleotide was transferred to a 2 ml Teflon lined screw cap vial. The

Se containing oligonucleotide was deprotected and cleaved from the solid support by treatment with 1 ml of freshly prepared *t*-butylamine/H_2O (1:3, v/v) containing 50 mM dithiothreitol (DTT) for 6 hours at 60 ^oC, followed by the addition of an equimolar amount of glacial acetic acid. All other oligonucleotides were deprotected and cleaved from the solid support with a treatment of 1 ml NH_4OH/EtOH (3:1) for 4 hours at 55 ^oC. After the incubation periods, the supernatant was transferred to a separate, clean screw cap vial and the CPG washed twice with 200 µl aqueous ACN (50 %, v/v). The crude oligonucleotides were lyophilized in a speed-vac concentrator.

All oligonucleotides were purified by strong anion exchange (SAX) HPLC using a Dionex DNAPAC PA-100 column (0.4 cm x 25 cm) purchased from Dionex Corp (Sunnyvale, CA). The column was eluted at room temperature using a linear gradient of 0-52 % buffer B over 30 minutes (buffer A: 100 mM Tris HCl, pH 7.5, 10 % ACN; buffer B: 100 mM Tris HCl, pH 7.5, 10 % ACN, 1 M NaCl). The column was monitored at 260 nm and 352 nm for analytical runs and 260 nm and 280 nm for preparative runs. The purified oligonucleotides were desalted using C-18 SEP PAK cartridges (Waters). The purified **dI-Se-ODN** sample was stored in an aqueous (18 MΩ) solution of 50 mM DTT. The purified **dI-Se-ODN$_2$** sample was stored in 18 MΩ H_2O.

3.2.5. Oligonucleotide Characterization by ESI-MS

ESI mass spectra for oligonucleotides were acquired at the Concordia University Centre for Biological Applications of Mass Spectrometry (CBAMS)

using a Micromass Qtof2 mass spectrometer (Waters) equipped with a nanospray ion source. The mass spectrometer was operated in full scan, negative ion detection mode and the raw data were deconvoluted.

3.2.6. Time Trials of Oligonucleotide Dimer Formation after Deprotection

The CPG-bound oligonucleotide was transferred to a 2 ml Teflon lined screw cap vial. The CPG was treated with 1 ml of freshly prepared *t*-butylamine/H_2O (1:3, v/v) at 60^{o}C. At different time points, a 20 µl aliquot was diluted to 500 µl using 18 MΩ H_2O with the addition of an equimolar amount of glacial acetic acid. After 18 hours, a sample was prepared with the extra addition of 50 µl of 0.5 M DTT which was allowed to incubate for 2 hours at room temperature before analysis. Each sample was vortexed, centrifuged and injected into the SAX HPLC. The column was eluted as described above.

3.2.7. Denaturing Polyacrylamide Gel Electrophoresis (PAGE) of Oligonucleotides

The gel consisted of 20 % acrylamide (19:1, w/w, acrylamide to bis-acrylamide) and 7 M urea in a pH 8.0 TBE buffer (89 mM Tris base, 89 mM boric acid, 0.2 mM ethylenediaminetetraacetic acid (EDTA)) with dimensions 10 cm x 7.5 cm x 0.75 mm (L x H x W). The gel was run for 1.5 hours at 200 V using the TBE buffer described above. Markers for the gel were a mixture of bromophenol blue (8 nucleotides) and xylene cyanol FF (28 nucleotides). Oligonucleotides

were prepared by lyophilizing 0.1 OD_{260} of sample and dissolving in 10 µl formamide. Prior to loading on the gel, **dI-Se-ODN$_2$** samples were treated with 10 mM DTT, 10 mM glutathione (GSH) or 10 mM $NaBH_4$ for 2 hours at room temperature, lyophilized in a speed-vac concentrator and dissolved in 10 µl formamide. The gel was visualized by placing it on a TLC plate (Merck, Kieselgel 60 F_{254}, 0.25 mm) which was irradiated with 254 nm light (UV shadowing).

3.2.8. UV Thermal Denaturation Studies of DNA Duplexes

Molar extinction coefficients for the unmodified and Se modified oligonucleotides were calculated from those of the mononucleotides using nearest-neighbor approximations (M^{-1} cm^{-1}). The molar extinction coefficient of 2'-deoxyinosine was used for the Se modified nucleotide. Oligonucleotide **dI-Se-ODN$_2$** (1 equiv.) was mixed with the complement strands (2 equiv.). All other duplexes were prepared by mixing equimolar amounts of the interacting strands. The mixtures were then lyophilized to dryness in a speed-vac concentrator. The resulting pellet was dissolved in 1 ml buffer (pH 7.0) containing 90 mM NaCl, 10 mM sodium phosphate and 1 mM EDTA. Samples containing **dI-Se-ODN** and **dI-ODN** had an additional 2.5 mM DTT in solution with a final duplex concentration of 3.7 µM. Samples of **dI-Se-ODN$_2$** did not contain DTT and had a final oligonucleotide dimer concentration of 1.0 µM. Samples were heated at 95 ^oC for 10 minutes then allowed to slowly come down to room temperature for 1 hour and placed in the fridge overnight. Prior to the thermal run, samples were degassed by placing them in a speed-vac concentrator for 2 minutes. Thermal

melting (T_m) curves were acquired at 260 nm starting at 15 ^{o}C and increasing the temperature at a rate of 0.5 ^{o}C/min until 95 ^{o}C, using a Varian CARY Model 3E spectrophotometer fitted with a 6-sample thermostated cell block and a temperature controller. Denaturing data processing was carried out as described by Puglisi and Tinoco[38] and transferred to Microsoft Excel™ software.

3.2.9. Circular Dichroism (CD) Spectroscopy of DNA Duplexes

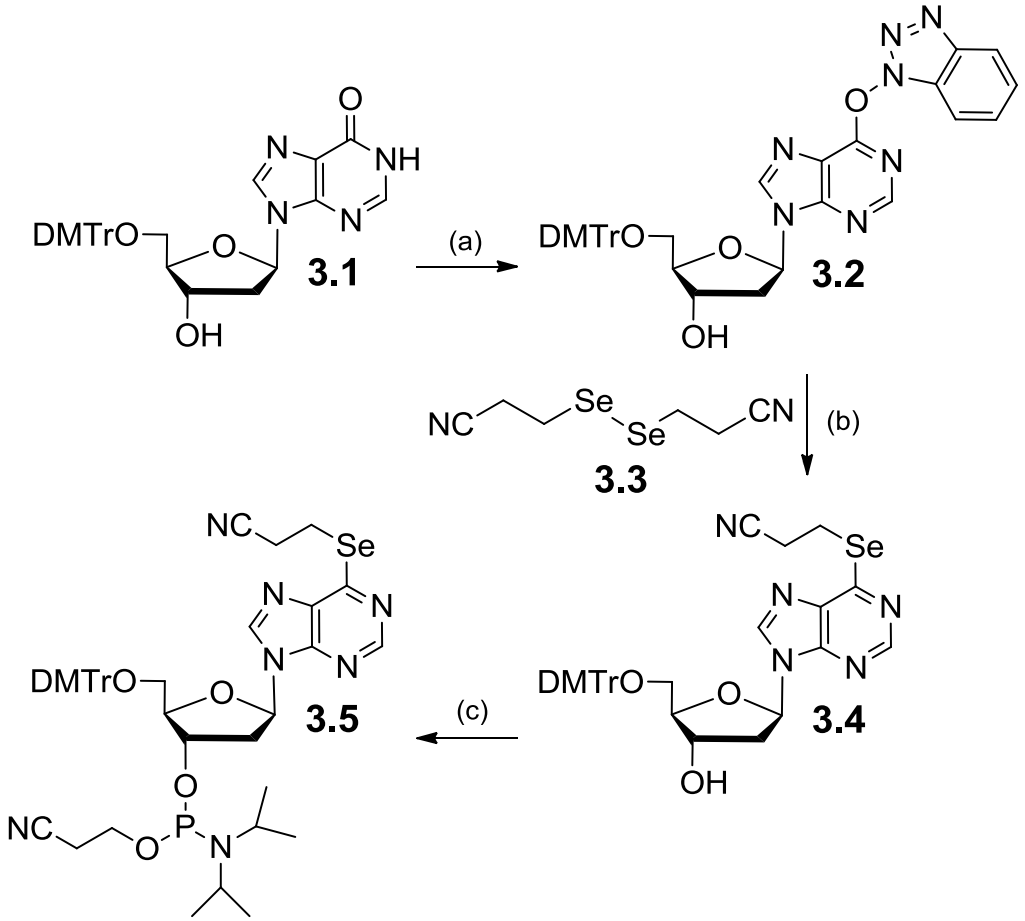
Circular dichroism spectra were obtained on a Jasco J-815 spectropolarimeter equipped with a Julaba F25 circulating bath. Samples were allowed to equilibrate for 5 minutes at 15 ^{o}C in 90 mM NaCl, 10 mM sodium phosphate, 1 mM EDTA (pH 7.0). Samples of **dI-Se-ODN** and **dI-ODN** had an additional 2.5 mM DTT in solution with a final duplex concentration of 3.7 μM. Samples of **dI-Se-ODN$_2$** did not contain DTT and had a final oligonucleotide dimer concentration of 1.0 μM. Each spectrum was an average of 5 scans, collected at a rate of 20 nm/min, with a bandwidth of 1 nm and sampling wavelength of 0.2 nm using fused quartz cells (Starna 29-Q-10). The spectra were recorded from 320 to 220 nm at 15 ^{o}C. The molar ellipticity [θ] was calculated from the equation [θ] = ε/Cl, where ε is the relative ellipticity (mdeg), C is the molar concentration of DNA duplex (M), and l is the path length of the cell (cm). The data were processed using software supplied by the manufacturer (JASCO, Inc.) and transferred to Microsoft Excel™ software.

3.3. Results and Discussion

Electrophilic O^6-(benzotriazol-1-yl)-2'-deoxyinosine[26] **3.2**, first described by Lakshman, was synthesized from commercially available 5'-*O*-dimethoxytrityl-2'-deoxyinosine **3.1**, with a yield of 74 % (**Scheme 3.1**). Reduction of diselenide[28]

Scheme 3.1: Synthesis of the $d^{6Se}I$ phosphoramidite (3.5).

Reagents and conditions: (a) PyBOP, DBU, DMF, 4 h, rt; (b) $NaBH_4$, EtOH, 2 h, 0 °C to rt; (c) Cl-POCEN*i*Pr_2, DIPEA, THF, 30 mins, rt.

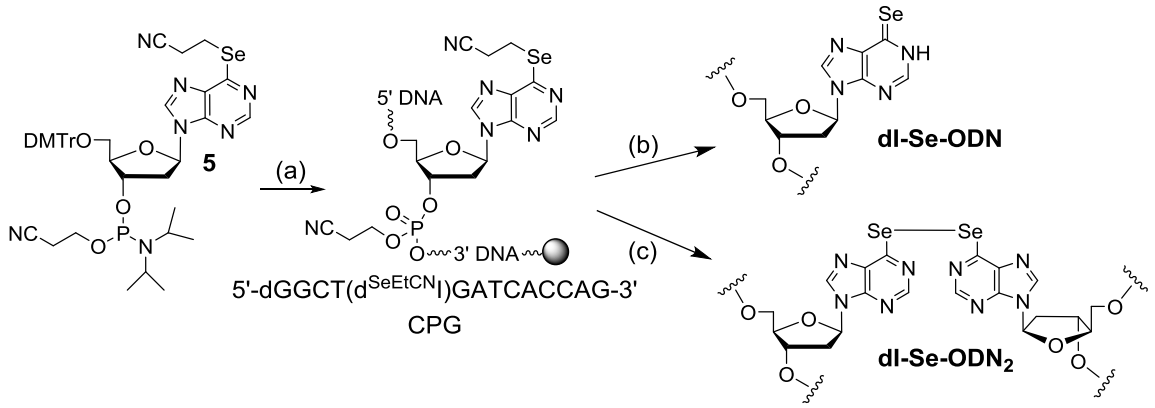


3.3, generating nucleophilic selenolate, allowed for the displacement of the benzotriazolyl group to produce **3.4**, with a yield of 66 %. Phosphitylation of **3.4** afforded phosphoramidite **3.5**, with a yield of 88 %. The phosphoramidite was

introduced into the CPG-bound DNA sequence 5'-dGGCT^{6Se}IGATCACCAG-3' by solid-phase synthesis (**Scheme 3.2**). All trityl values after phosphoramidite coupling were above 90 % indicating successful incorporation of the modified phosphoramidite.

Scheme 3.2: Synthesis of the dI-Se-ODN and formation of dI-Se-ODN$_2$.

Reagents and conditions: (a) Solid-phase synthesis (see experimental section); (b) (i) *t*-butylamine/H_2O (1:3, v/v) in 50 mM DTT, 6 h, 60°C; (ii) glacial acetic acid; (c) (i) *t*-butylamine/H_2O (1:3, v/v), 6 h, 60 °C (ii) glacial acetic acid.

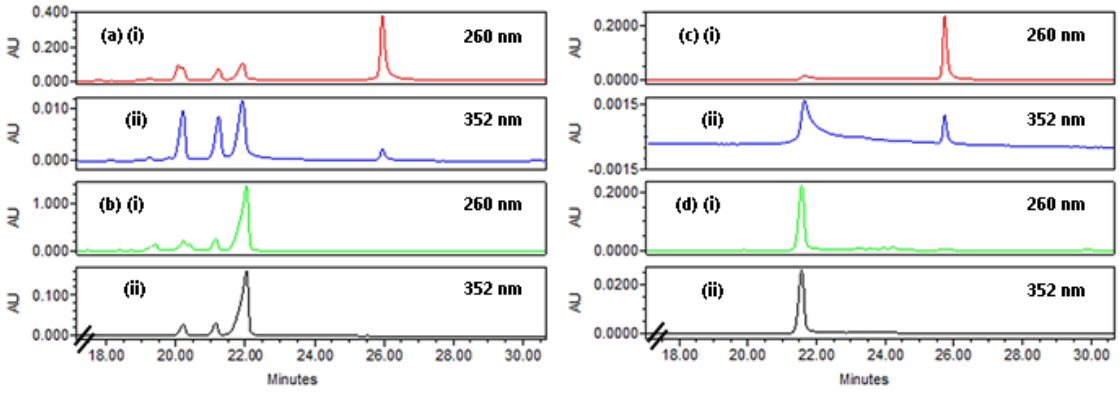


Conventional deprotection with NH_4OH/EtOH (3:1, v/v) was found to be too harsh to accomplish removal of the protective groups while retaining Se in the oligonucleotide, necessitating the use of a milder *t*-butylamine/H_2O (1:3, v/v) treatment followed by the addition of an equimolar amount of glacial acetic acid to neutralize the amine base (**Scheme 3.2**). Unexpectedly, rather than the desired oligonucleotide **dI-Se-ODN**, the dimerized **dI-Se-ODN$_2$** was the major product after deprotection and liberation from the solid-support. The formation of this dimer was first suspected during the SAX HPLC purification of the crude oligonucleotide. Elution times for sequences 14 nucleotides in length are expected to be in the range of 18-22 minutes under the buffer conditions

employed by SAX HPLC (specified in **Figure 3.2**); however, the major species generated had an elution time closer to 26 minutes (**Figure 3.2a**). Furthermore, 4-selenothymidine and 2'-deoxy-6-selenoguanosine containing oligonucleotides have characteristic UV absorptions at 369 nm and 360 nm, respectively.[25,29] Absorption at 352 nm was observed for the d^{6Se}I containing oligonucleotides (**Figure A3.10**); however, the major species generated had minimal absorbance at this wavelength (**Figure 3.2a**), especially when compared to the minor species. In order to account for such a drastic decrease in absorbance intensity

Figure 3.2: SAX HPLC traces of (a) crude dI-Se-ODN$_2$ monitored at (i) 260 nm and (ii) 352 nm (b) crude dI-Se-ODN monitored at (i) 260 nm and (ii) 352 nm (c) pure dI-Se-ODN$_2$ monitored at (i) 260 nm and (ii) 352 nm (d) pure dI-Se-ODN monitored at (i) 260 nm and (ii) 352 nm

The column was eluted at room temperature using a linear gradient of 0-52 % buffer B over 30 minutes (buffer A: 100 mM Tris HCl, pH 7.5, 10 % ACN; buffer B: 100 mM Tris HCl, pH 7.5, 10 % ACN, 1 M NaCl).



at 352 nm, it was hypothesized that a modification occurred at or near the Se position of the oligonucleotide. The major species was collected, desalted, and analysed by ESI-MS, which confirmed a higher than expected molecular mass. The mass observed corresponded to a diselenide bridged oligonucleotide homodimer (**Figure A3.12**). Unfortunately, attempts to digest **dI-Se-ODN$_2$** by

methods previously used in our group were not successful.[39] This may be due to the nuclease not being able to digest the oligonucleotide at the cross-linked site or that buffer used for digestion compromises the Se-Se cross-link.

In order to isolate **dI-Se-ODN** and confirm the presence of a diselenide bond, deprotection of the CPG-bound oligonucleotide was repeated, but with the addition of the reducing agent dithiothreitol (DTT). The major species in the SAX HPLC profile now had a reduced elution time of approximately 22 minutes and an intense absorbance at 352 nm (**Figure 3.2b**), which were characteristic of the $d^{6Se}I$ containing 14-mer. No peak at 26 minutes was observed. After collecting and desalting the major species, SAX HPLC indicated the dimer had reformed (**Figure 3.2c**). Re-addition of DTT to the solution allowed for the isolation of pure **dI-Se-ODN** (**Figure 3.2d**), which was confirmed by ESI-MS (**Figure A3.14**). Storage of **dI-Se-ODN** at 4 ^{o}C in the absence of DTT was shown to prevent dimerization, even after 1 week (data not shown).

To test if the deprotection condition used contributed to dimerization, deprotection was repeated with another mild treatment: 0.05 M K_2CO_3/MeOH for 4 h at room temperature, with and without the addition of an equimolar amount of glacial acetic acid. In both instances, the major species was once again **dI-Se-ODN$_2$** (data not shown). Furthermore, removal of DTT during desalting of crude **dI-Se-ODN** initiated re-dimerization. These results suggest spontaneous (without the addition of an oxidizing agent) diselenide bridge formation is occurring. Interestingly, 2'-deoxy-6-selenoguanosine containing oligonucleotides have not been reported to readily self-dimerize.[25] Potentially, the C-2 exocyclic amino

group of guanine sterically prevents diselenide bond formation between the 6-seleno atoms. Thio-nucleobase-modified oligonucleotides have been shown to form disulfide bridges in complementary sequences where the bridging atoms are in proximity to one another after hybridization.[40,41]

In order to observe the timescale of dimerization after cleavage from the solid-support, time trials were performed by analysing the crude **dI-Se-ODN** oligonucleotide by SAX HPLC after deprotection in the absence of DTT. Dimerization was however, too fast under these experimental conditions. After only 2 hours of deprotection, **dI-Se-ODN$_2$** was already the predominant species, with no evidence of **dI-Se-ODN** accumulating over time (**Figure A3.17**). This suggests that at elevated temperatures (60 ^{o}C), diselenide bridge formation is rapid and occurs soon after cleavage from the solid-support and removal of the cyanoethyl group protecting the 6-seleno atom. After 18 hours of deprotection, incubation of the crude oligonucleotide with 50 mM DTT for 2 hours quantitatively reverted the dimer (**dI-Se-ODN$_2$**) to the single-stranded species (**dI-Se-ODN**) (**Figure A3.17**).

UV thermal denaturation of duplexes between **dI-Se-ODN** with complementary strands where d^{6Se}I is base paired with dA, dC, dG and dT nucleotides were performed. The buffers for these experiments contained 2.5 mM DTT in order to avoid the self-dimerization of the d^{6Se}I containing oligonucleotide. The denaturation profiles revealed T_m values which were comparable to duplexes containing unmodified 2'-deoxyinosine (**Figure A3.18**). This suggests there is no significant destabilization of the duplex upon Se

modification at the 6-position of 2'-deoxyinosine. Furthermore, the CD spectra of the same **dI-Se-ODN** containing duplexes showed typical B-form DNA signatures, with a wavelength maximum at 280 nm, an intercept near 260 nm and a minimum around 250 nm, similar to the control 2'-deoxyinosine containing duplexes (**Figure A3.19**). This suggests the Se modification has minimal influence on the global DNA structure.

UV thermal denaturation experiments of **dI-Se-ODN$_2$** in the presence of 2 equivalents of the complementary strands produced T_m curves which were biphasic in nature and the CD spectra also showed typical B-form DNA signatures (**Figures A3.20** and **A3.21**). The biphasic T_m curves may be due to the dissociation of the first, then second equivalent, of the complementary strand hybridized to **dI-Se-ODN$_2$**. Denaturing PAGE was used in order to evaluate the efficiency of other reducing agents on diselenide bridge removal of **dI-Se-ODN$_2$**, specifically $NaBH_4$ and glutathione (GSH) (**Figure A3.22**). Both of these reagents caused the loss of the dimer band after only 2 hours of incubation at room temperature, with migrations similar to **dI-Se-ODN**. A significant smearing in the GSH lane was observed, perhaps due to incomplete removal of the GSH in loading of the oligonucleotide onto the gel. In the untreated **dI-Se-ODN$_2$** lane, two bands were observed. Electrolytic reduction of the diselenide bond is believed to occur, producing the lower band which migrates with the **dI-Se-ODN** control in lane 4; a significant amount of dimer remained. Electrolytic reduction of insulin disulfide bonds has previously been reported.[42]

3.4. Conclusion

In summary, the synthesis and characterization of a novel Se modified 2'-deoxyinosine phosphoramidite has been described. In the absence of reducing agents such as DTT, the crude oligonucleotide forms a homodimer linked by a diselenide bridge. This dimer can readily be reduced to restore the single stranded oligonucleotide upon treatment with agents such as DTT, $NaBH_4$ and GSH. Duplexes containing $d^{6Se}I$ paired with any of the four natural nucleobases of DNA revealed minimal perturbation to duplex thermal stability and structure relative to 2'-deoxyinosine.

Future work will involve crystallizing and solving the X-ray crystal structure of duplexes containing this novel Se containing 2'-deoxyinosine derivative. Formation and crystallization of a diselenide cross-linked DNA duplex and structural determination by X-ray crystallography may be advantageous in solving the structures of larger nucleic acid assemblies given that the Se phasing power is proportional to the number of Se atoms incorporated (one Se atom per 30 nucleotides).[43]

Beyond X-ray crystallography, these diselenide cross-linked oligonucleotides may find applications as stimuli responsive materials given their reactivity relative to species containing disulfide bonds.[44,45,46] These materials may be attractive as a molecular switch able to respond to the intracellular reducing environment.[47]

Chapter 4. Conclusions and Future Work

4.1. General Conclusions

Using an authentically synthesized $d^{4SeEtCN}T$ containing oligonucleotide, the optimal deprotection conditions were determined for retaining Se in near quantitative amounts. This was found to be a treatment of *t*-butylamine/H_2O (1:3; v/v) for 6 hours at 60 ^oC, followed by equimolar 100 % acetic acid to neutralize the amine. By minimizing deselenization and adduct formation, pure samples of Se oligonucleotides can be obtained which should facilitate crystallization. Conditions to incorporate Se into solid-support bound oligonucleotides containing a convertible 4-triazolyl-thymidine were also determined, however, the best yield was only 29 % (estimated by SAX HPLC), necessitating further optimization. The on-column Se incorporation was finally achieved by performing the reaction at -20 ^oC.

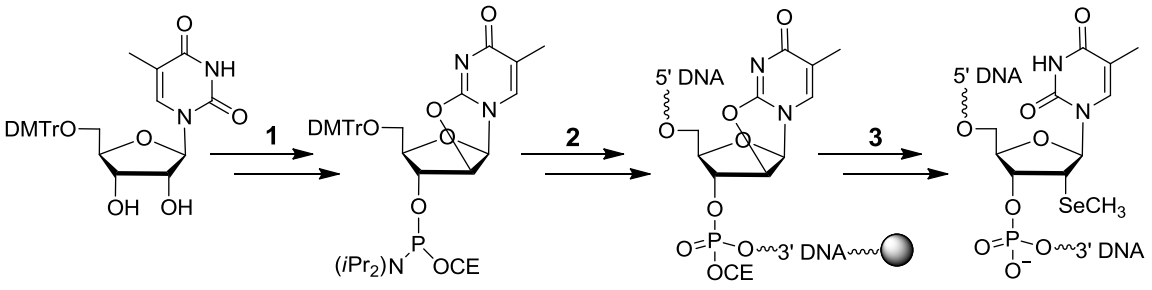
Through the synthesis of an authentic $d^{6Se}I$ containing oligonucleotide, unexpected cross-linking through the formation of a diselenide bond was observed between two non-complementary strands. This phenomenon was interesting as it was not reported in oligonucleotides containing $d^{6Se}G$ nucleotides, suggesting that the C-2 exocyclic amino group of guanine sterically prevents dimerization. This spontaneous diselenide bridge forms soon after deprotection at 60 ^oC and can be reversed with treatment of DTT, $NaBH_4$ or GSH. Under reducing conditions, it was found that $d^{6Se}I$ containing

oligonucleotides possess a characteristic absorbance at 352 nm. This property can be exploited in the HPLC purification of these oligonucleotides. Thermal denaturation and CD studies showed that the Se modification was well tolerated in the duplex, with and without the diselenide cross-link.

4.2. Future Work

Further optimization of the on-column Se incorporation conditions is required in order to maximize the yield of the desired oligonucleotide product. Working under an inert atmosphere in a glove box may prevent undesirable oxidation of Se or deselenization due to ambient moisture. Longer reaction times should be investigated to ensure complete conversion of the triazolyl functionality to prevent undesired adduct formation. Furthermore, using a CPG support with a lower nucleoside loading could decrease the steric hindrance between growing strands and potentially increase the efficiency of Se incorporation. Once on-column Se incorporation has been optimized at the 5'-end, it would be ideal to explore the introduction of Se at internal positions of the oligonucleotide. Additionally, it may be of interest to test the viability of further extension of the

Scheme 4.1: An alternative approach for on-column Se incorporation.

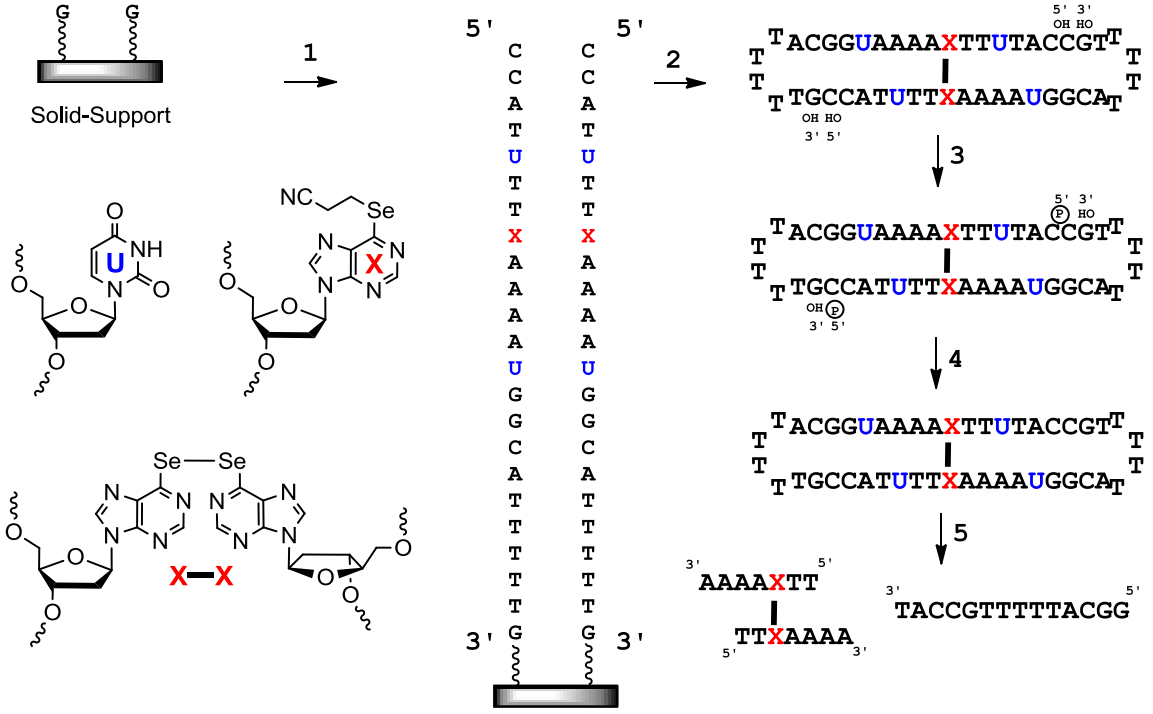


oligonucleotide after on-column Se incorporation at the terminal end. Both of these approaches could enable Se incorporation at any position in the sequence.

Should optimizations of the on-column methodology using a convertible nucleoside ultimately afford low yielding results, a new approach, illustrated in **Scheme 4.1**, could be attempted. This approach would involve the synthesis of an anhydrothymidine phosphoramidite, as shown in step 1, followed by solid-phase synthesis of the oligonucleotide in step 2 and finally, in step 3, the on-column introduction of the Se functionality at the 2'-position, deprotection and purification.[43] It has previously been shown that anhydro nucleosides are compatible with solid-phase synthesis conditions and they are even suitable for post-synthetic modifications.[48] This approach may have the advantage of being less sensitive to moisture resulting in deselenization.

With regard to the diselenide cross-linked oligonucleotides, it would be desirable to harness the dimerization reaction in order to expand to scope of this fundamental finding to prepare other well-defined cross-linked species. The determination of conditions that prevent self-dimerization and instead promote cross-linking to a different, complementary oligonucleotide would broaden the range of nucleic acid assemblies that can be constructed. It was shown that lower temperatures can prevent spontaneous diselenide bridge formation when the strands are non-complementary.[49] When incubating two complementary oligonucleotides with $d^{6Se}I$ nucleotides opposing one another at 4 ^{o}C in an aqueous DTT solution, duplex formation without cross-linking occurs. Quickly desalting the oligonucleotide solution (to remove DTT) and resuspending in cold

water would reinitiate duplex formation, placing the Se atoms in close proximity. Over time, diselenide bridge formation could occur. If the temperature was found



Scheme 4.2: Synthesis of small circular DNA utilizing the diselenide cross-link.

to be too low for the diselenide cross-link to form, it could be increased slowly, but never above the duplex T_m value to ensure the major species in solution was the duplex. This strategy could be explored to produce diselenide cross-linked heterodimeric duplexes. With crystals of the duplex, it would be possible to determine its high-resolution structure by X-ray crystallography, utilizing the 2 Se atoms for phasing. This would shed light on how the local environment of the duplex shifts in order to accommodate the cross-link.

With a better understanding of the spatial arrangement of the diselenide cross-link, it could potentially be utilized to design small circular DNA. One such design strategy is shown in **Scheme 4.2**. The d^{6Se}I monomer could be introduced

into an oligonucleotide sequence by solid-phase synthesis (step 1). The sequence illustrated is only one example; sequences of different nucleobase composition and length could also be evaluated. After deprotection of the oligonucleotide in the presence of reducing dithiothreitol (DTT), the crude oligonucleotide would be purified by SAX HPLC. By monitoring the eluent at 352 nm, the Se containing oligonucleotide can be selectively identified and collected. During the desalting process, DTT would be removed, which initiates spontaneous oxidation to form the diselenide cross-link between the strands (step 2). The strands should then hybridize as illustrated. The free 5'-OH groups would be phosphorylated and ligated to give the circular species (steps 3 and 4). It is anticipated that the structure of the circular DNA will be different in the presence and absence of the cross-link. Structural differences can readily be evaluated by denaturing PAGE and CD spectroscopy. It has previously been difficult to confirm circularization, but this issue could be addressed by engineering 2'-deoxyuridine monomers into the sequence. Treatment with the enzyme uracil deglycosylase followed by piperidine should allow for controlled cleavage into products that will confirm successful ligation and circularization (step 5). This can be accomplished by denaturing PAGE, providing fast visualization of the oligonucleotide cleavage products present. Furthermore, electrospray ionization mass spectrometry of the cleavage products would be possible.

Diselenide cross-linked circular DNA between non-complementary strands could be used to bind a target sequence with higher affinity and selectivity than

standard Watson Crick pairing.[50] These small circular DNA have been shown to be more resistant to degradation in human serum when compared to single stranded or duplex DNA.[51] A recent study described the utilization of small circular oligonucleotides in the formation of DNA nanotubes.[52]

Going forward, it may be possible to introduce the diselenide cross-link into other, more complex topologies, such as triplexes and tetrameric species, in order to stabilize them. Multiple cross-link incorporations may be needed in these instances, but this shouldn't pose a challenge owing to the synthetic ease by which the diselenide can be incorporated. This would also be where the dual functionality of Se is best displayed: by promoting the formation of novel nucleic acid structures and facilitating the determination of their high resolution X-ray crystal structure, providing information unattainable by PAGE and CD spectroscopy.

References

(1) Blackburn, M.; Gait, M.; David, L.; Williams, D. *Nucleic Acids in Chemistry and Biology*, 3rd ed.; Royal Society of Chemistry: Cambridge, UK, 2006.

(2) Crick, F. H. C. *J. Mol. Biol.* **1966**, *19*, 548–555.

(3) Franklin, R. E.; Gosling, R. G. *Acta Crystallogr.* **1953**, *6*, 673–677.

(4) Wood, B. R. *Chem. Soc. Rev.* **2016**, *45*, 1980–1998.

(5) Rich, A.; Watson, J. D. *Nature* **1954**, *173*, 995.

(6) Gellert, M.; Lipsett, M. N.; Davies, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013–2018.

(7) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871–880.

(8) Bochman, M. L.; Paeschke, K.; Zakian, V. A. *Nat. Rev. Genet.* **2012**, *13*, 770–780.

(9) Letsinger, R. L.; Mahadevan, V. *J. Am. Chem. Soc.* **1965**, *87*, 3526–3527.

(10) Caruthers, M. H. *Science.* **1985**, *230*, 281–285.

(11) Caruthers, M. H. *Acc. Chem. Res.* **1991**, *24*, 278–284.

(12) Brown, T.; Brown, T. J. Solid-Phase Oligonucleotide Synthesis https://www.atdbio.com/content/17/Solid-phase-oligonucleotide-synthesis (accessed Jan 31, 2018).

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

(14) Rhodes, G. *Crystallography Made Crystal Clear*, 3rd ed.; Academic Press: San Diego, 2006.

(15) Lin, L.; Sheng, J.; Huang, Z. *Chem. Soc. Rev.* **2011**, *40*, 4591–4602.

(16) Harp, J.; Pallan, P.; Egli, M. *Crystals* **2016**, *6*, 1–26.

(17) Jagannadham, M. V. *Proteomics Insights* **2009**, *2*, 27–31.

(18) Höbartner, C.; Micura, R. *J. Am. Chem. Soc.* **2004**, *126*, 1141–1149.

(19) Höbartner, C.; Rieder, R.; Kreutz, C.; Puffer, B.; Lang, K.; Polonskaia, A.; Serganov, A.; Micura, R. *J. Am. Chem. Soc.* **2005**, *127*, 12035–12045.

(20) Carrasco, N.; Ginsburg, D.; Du, Q.; Huang, Z. *Nucleosides. Nucleotides Nucleic Acids* **2001**, *20*, 1723–1734.

(21) Du, Q.; Carrasco, N.; Teplova, M.; Wilds, C. J.; Egli, M.; Huang, Z. *J. Am. Chem. Soc.* **2002**, *124*, 24–25.

(22) Watts, J. K.; Johnston, B. D.; Jayakanthan, K.; Wahba, A. S.; Pinto, B. M.; Damha, M. J. *J. Am. Chem. Soc.* **2008**, *130*, 8578–8579.

(23) Wilds, C. J.; Pattanayek, R.; Pan, C.; Wawrzak, Z.; Egli, M. *J. Am. Chem. Soc.* **2002**, *124*, 14910–14916.

(24) Salon, J.; Sheng, J.; Jiang, J.; Chen, G.; Caton-Williams, J.; Huang, Z. *J. Am.*

Chem. Soc. **2007**, *129*, 4862–4863.

(25) Salon, J.; Jiang, J.; Sheng, J.; Gerlits, O. O.; Huang, Z. *Nucleic Acids Res.* **2008**, *36*, 7009–7018.

(26) Bae, S.; Lakshman, M. K. *J. Am. Chem. Soc.* **2007**, *129*, 782–789.

(27) Noll, D. M.; Noronha, A. M.; Miller, P. S. *J. Am. Chem. Soc.* **2001**, *123*, 3405–3411.

(28) Logan, G.; Igunbor, C.; Chen, G.-X.; Davis, H.; Simon, A.; Salon, J.; Huang, Z. *Synlett* **2006**, *2006*, 1554–1558.

(29) Sheng, J.; Huang, Z. *Curr. Protoc. Nucleic Acid Chem.* **2008**, *32*, 1.19.1-1.19.13.

(30) Wilds, C. J.; Pattanayek, R.; Pan, C.; Wawrzak, Z.; Egli, M. *J. Am. Chem. Soc.* **2002**, *124*, 14910–14916.

(31) McManus, F. P.; Wilds, C. J. *ChemBioChem* **2014**, *15*, 1966–1977.

(32) Xu, Y. Z.; Swann, P. F. *Nucleic Acids Res.* **1990**, *18*, 4061–4066.

(33) Swann, P. F.; Yao-zhong, X.; Zheng, Q. *J. Org. Chem.* **1992**, *57*, 3839–3845.

(34) Abdur, R.; Gerlits, O. O.; Gan, J.; Jiang, J.; Salon, J.; Kovalevsky, A. Y.; Chumanevich, A. A.; Weber, I. T.; Huang, Z. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014**, *70*, 354–361.

(35) Martin, F. H.; Castro, M. M.; Aboul-ela, F.; Tinoco, I. *Nucleic Acids Res.* **1985**,

13, 8927–8938.

(36) Coleman, R. S.; Arthur, J. C.; McCary, J. L. *Tetrahedron* **1997**, *53*, 11191–11202.

(37) Seela, F.; Chen, Y. *Nucleic Acids Res.* **1995**, *23*, 2499–2505.

(38) Puglisi, J. D.; Tinoco Jr., I. *Methods Enzymol.* **1989**, *180*, 304–325.

(39) Sun, G.; Noronha, A.; Wilds, C. *Tetrahedron* **2012**, *68*, 7787–7793.

(40) Milton, J.; Connolly, B. a.; Nikiforov, T. T.; Cosstick, R. *J. Chem. Soc. Chem. Commun.* **1993**, *0*, 779–780.

(41) Coleman, R. S.; McCary, J. L.; Perez, R. J. *Tetrahedron* **1999**, *55*, 12009–12022.

(42) Markus, G. *J. Biol. Chem.* **1964**, *239*, 4163–4170.

(43) Teplova, M.; Wilds, C. J.; Wawrzak, Z.; Tereshko, V.; Du, Q.; Carrasco, N.; Huang, Z.; Egli, M. *Biochimie* **2002**, *84*, 849–858.

(44) Beld, J.; Woycechowsky, K. J.; Hilvert, D. *Biochemistry* **2007**, *46*, 5382–5390.

(45) Gowd, K. H.; Yarotskyy, V.; Elmslie, K. S.; Skalicky, J. J.; Olivera, B. M.; Bulaj, G. *Biochemistry* **2010**, *49*, 2741–2752.

(46) Cheng, G.; He, Y.; Xie, L.; Nie, Y.; He, B.; Zhang, Z.; Gu, Z. *Int. J. Nanomedicine* **2012**, *7*, 3991–4006.

(47) Schafer, F. Q.; Buettner, G. R. *Free Radic Biol Med* **2001**, *30*, 1191–1212.

(48) Ozaki, H.; Nakajima, K.; Tatsui, K.; Izumi, C.; Kuwahara, M.; Sawai, H. *Bioorganic Med. Chem. Lett.* **2003**, *13*, 2441–2443.

(49) Liczner, C.; Grenier, V.; Wilds, C. J. *Tetrahedron Lett.* **2018**, *59*, 38–41.

(50) Chaudhuri, N. C.; Kool, E. T. *J. Am. Chem. Soc.* **1995**, *117*, 10434–10442.

(51) Rumney, S.; Kool, E. T. *Angew. Chemie - Int. Ed.* **1992**, *31*, 1617–1619.

(52) Zheng, H.; Xiao, M.; Yan, Q.; Ma, Y.; Xiao, S. J. *J. Am. Chem. Soc.* **2014**, *136*, 10194–10197.

Appendix I: Supporting Information for Chapter 2

Contents Page

Supporting Figures and Schemes

Figure A2.1: 500 MHz ^{1}H NMR spectrum of compound **2.6** (in $CDCl_3$)

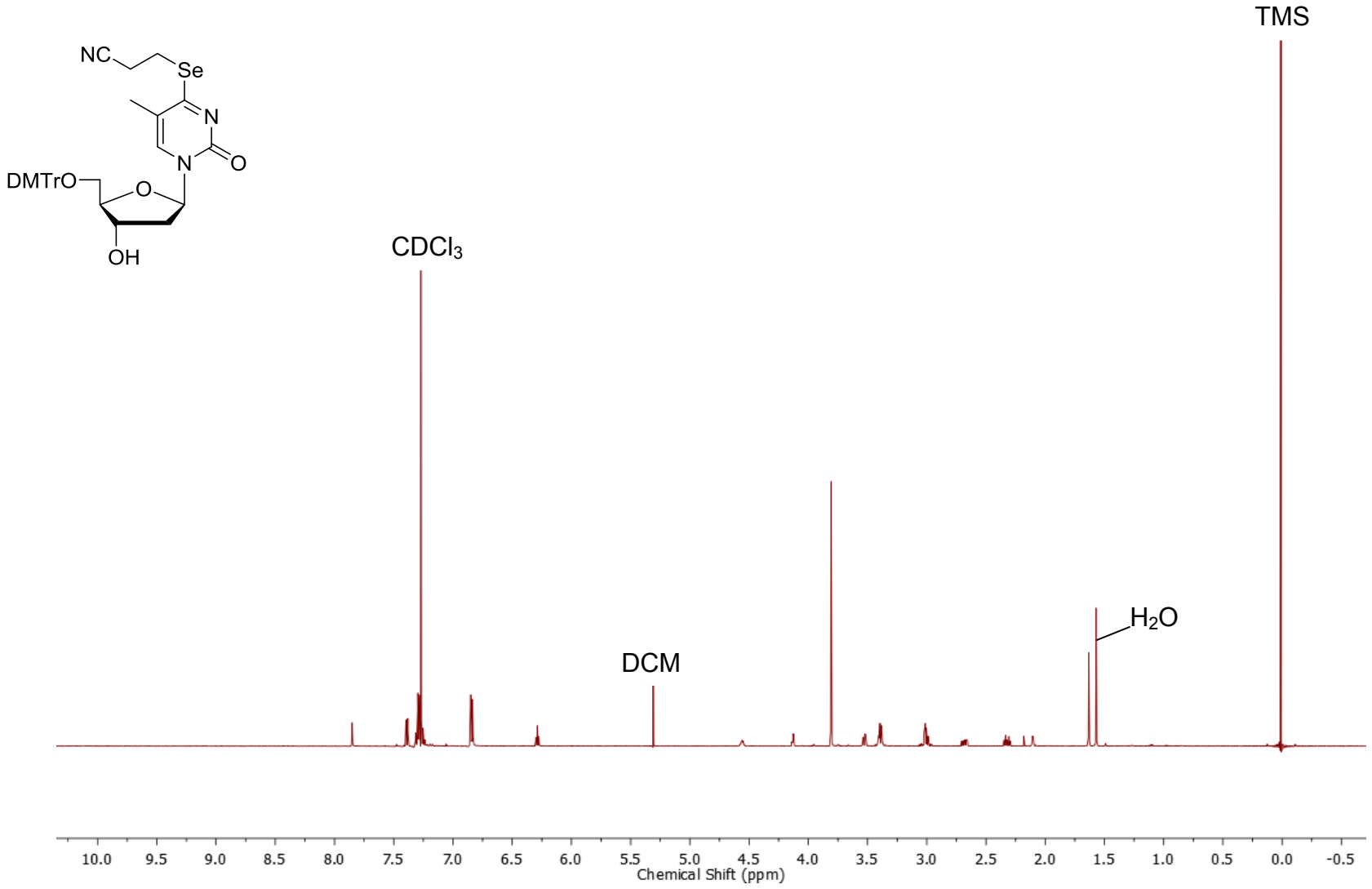


Figure A2.2: HR ESI-MS spectrum of compound **2.6** (expected mass: 684.1583)

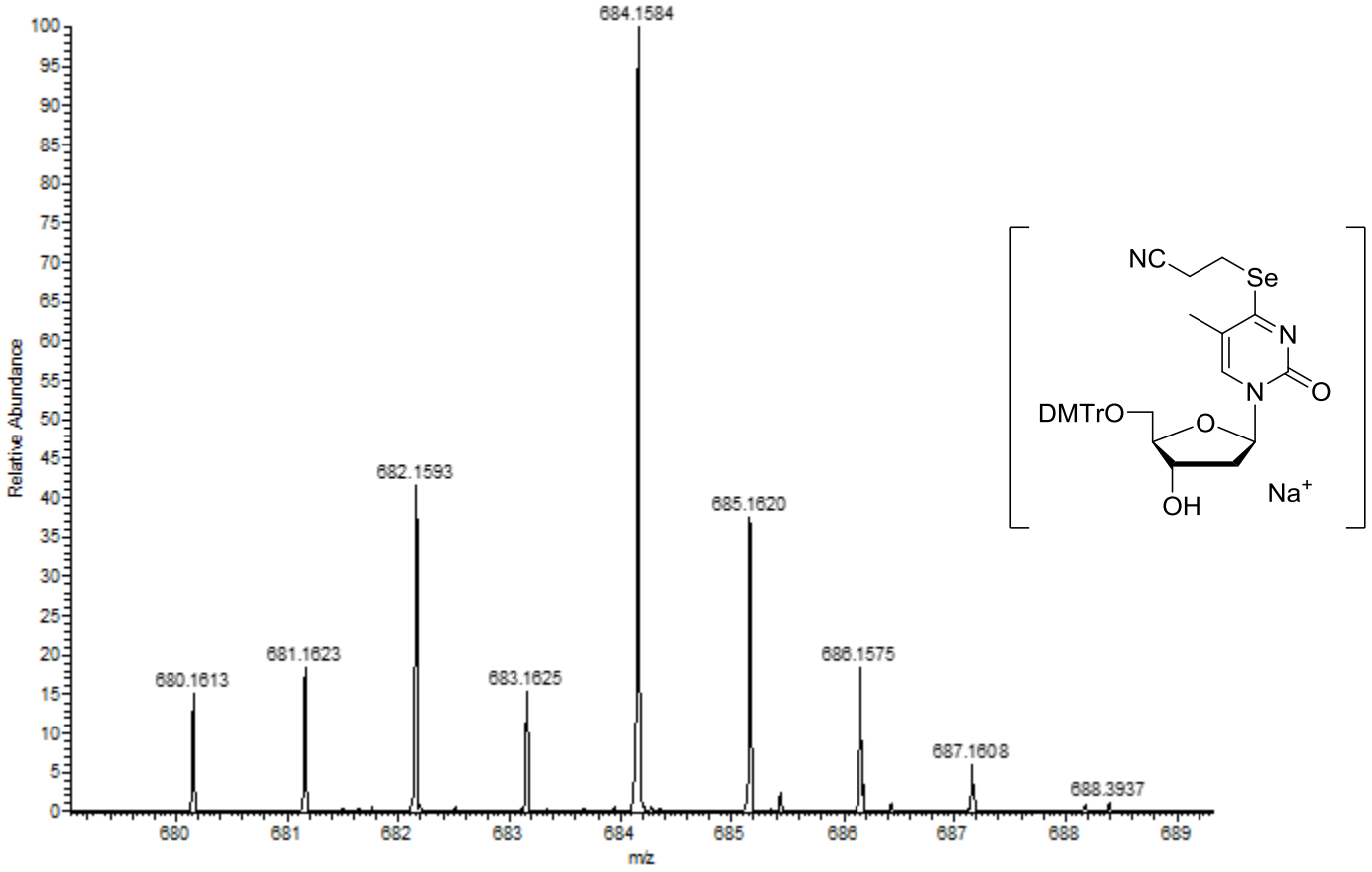


Figure A2.3: 500 MHz ^{1}H NMR spectrum of compound **2.7** (in d_6-acetone)

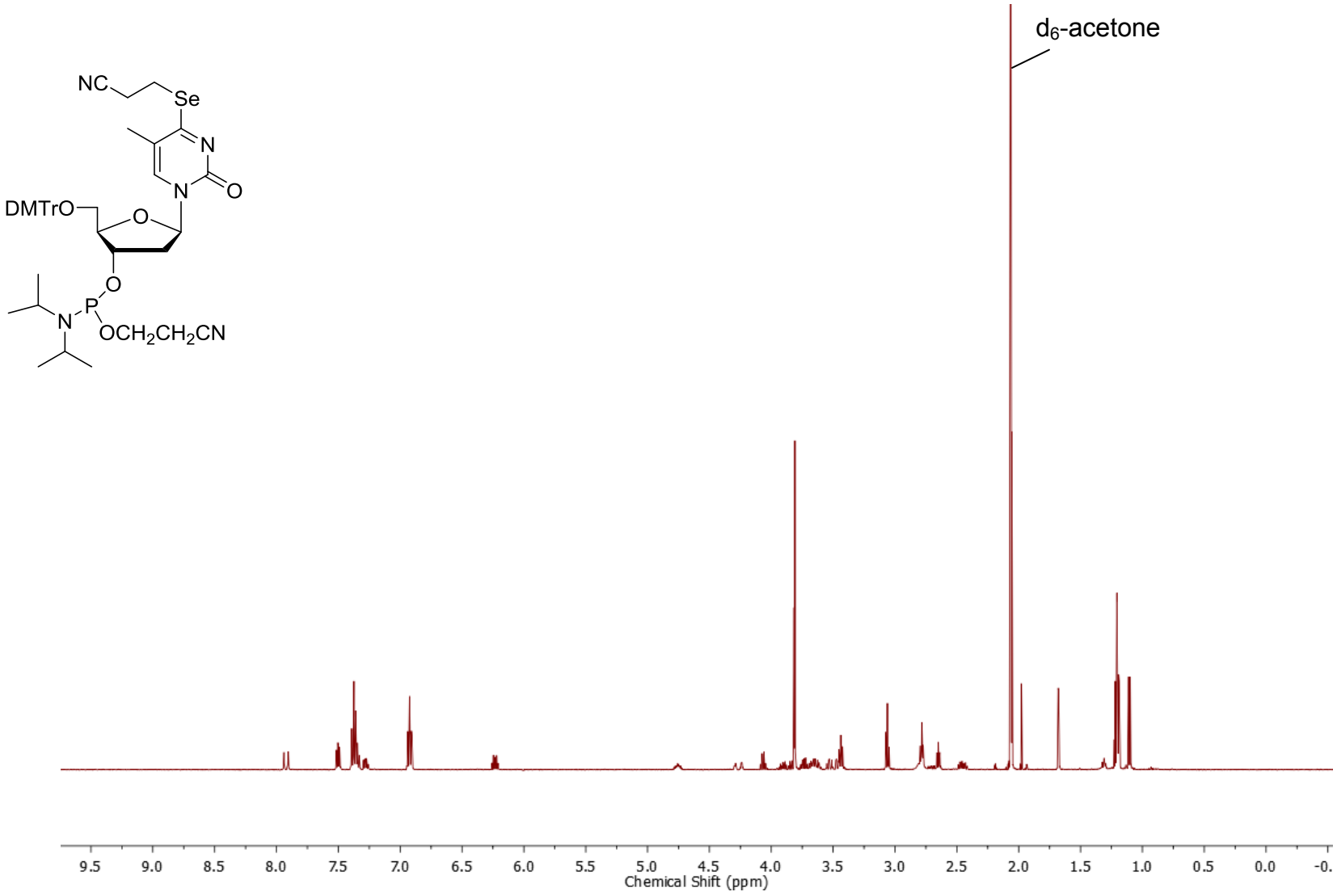


Figure A2.4: 202.3 MHz ^{31}P NMR spectrum of compound **2.7** (in d_6-acetone)

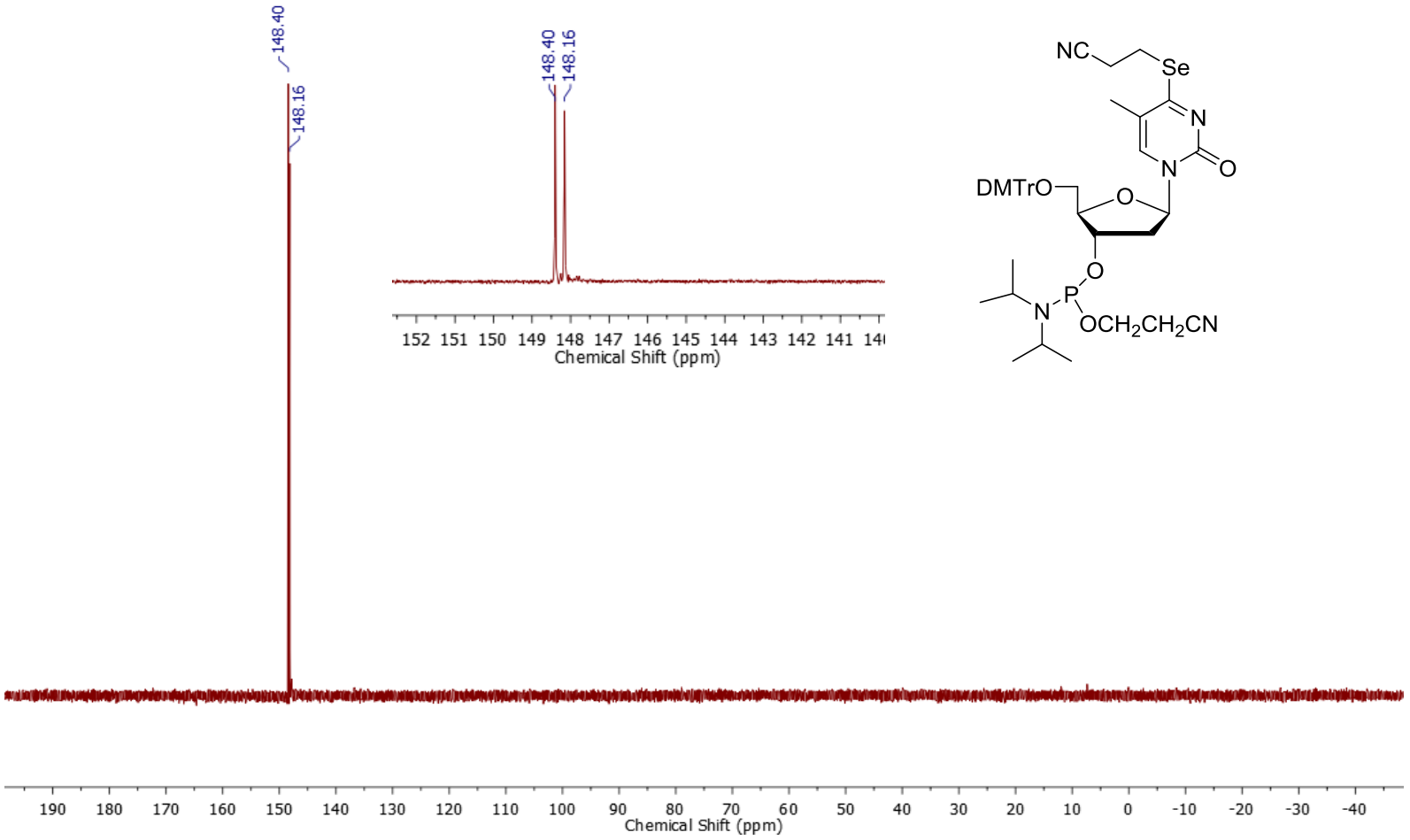


Figure A2.5: 500 MHz ^{1}H NMR spectrum of compound **2.9** (in d_6-acetone)

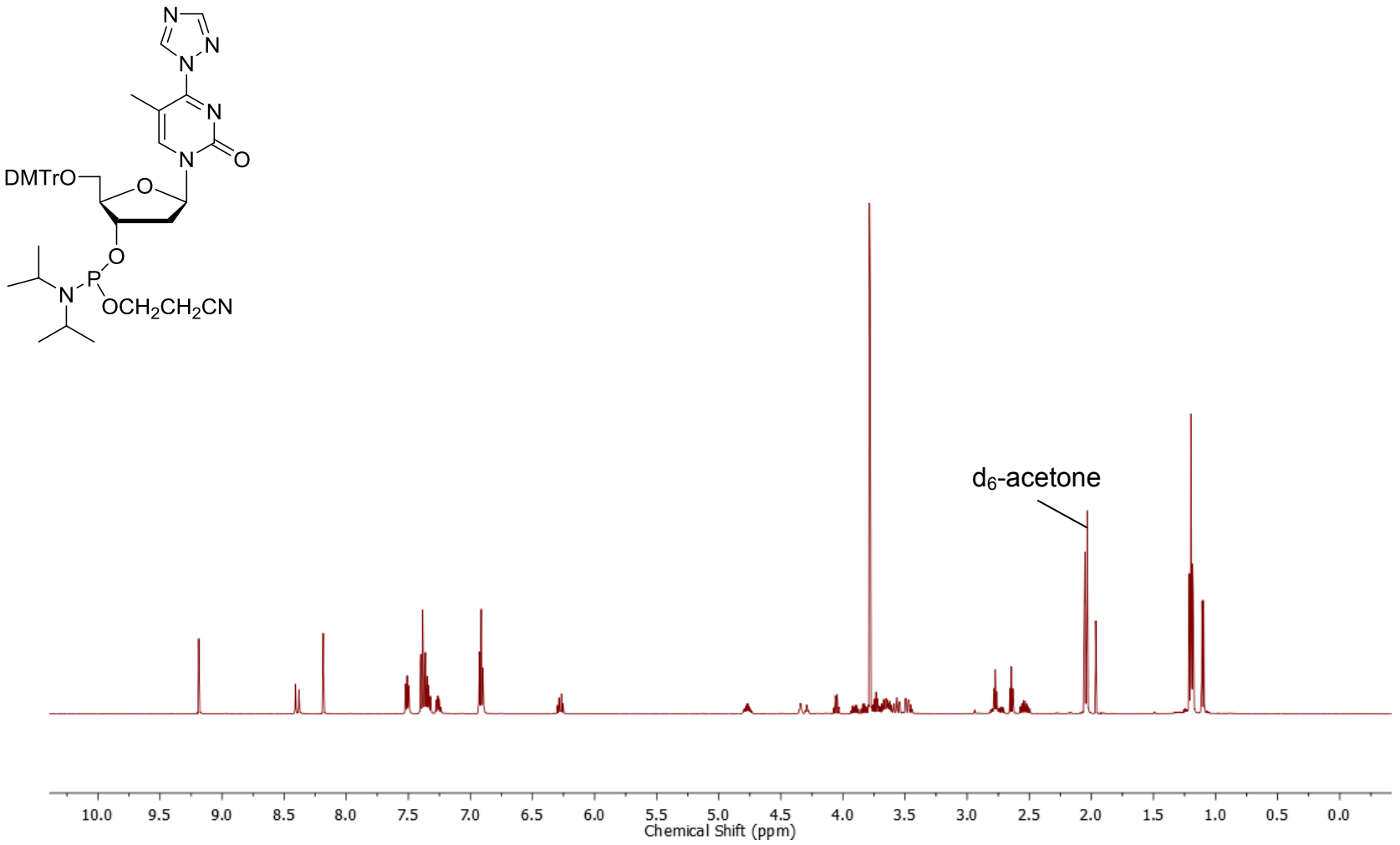


Figure A2.6: 202.3 MHz ^{31}P NMR spectrum of compound **2.9** (in d_6-acetone)

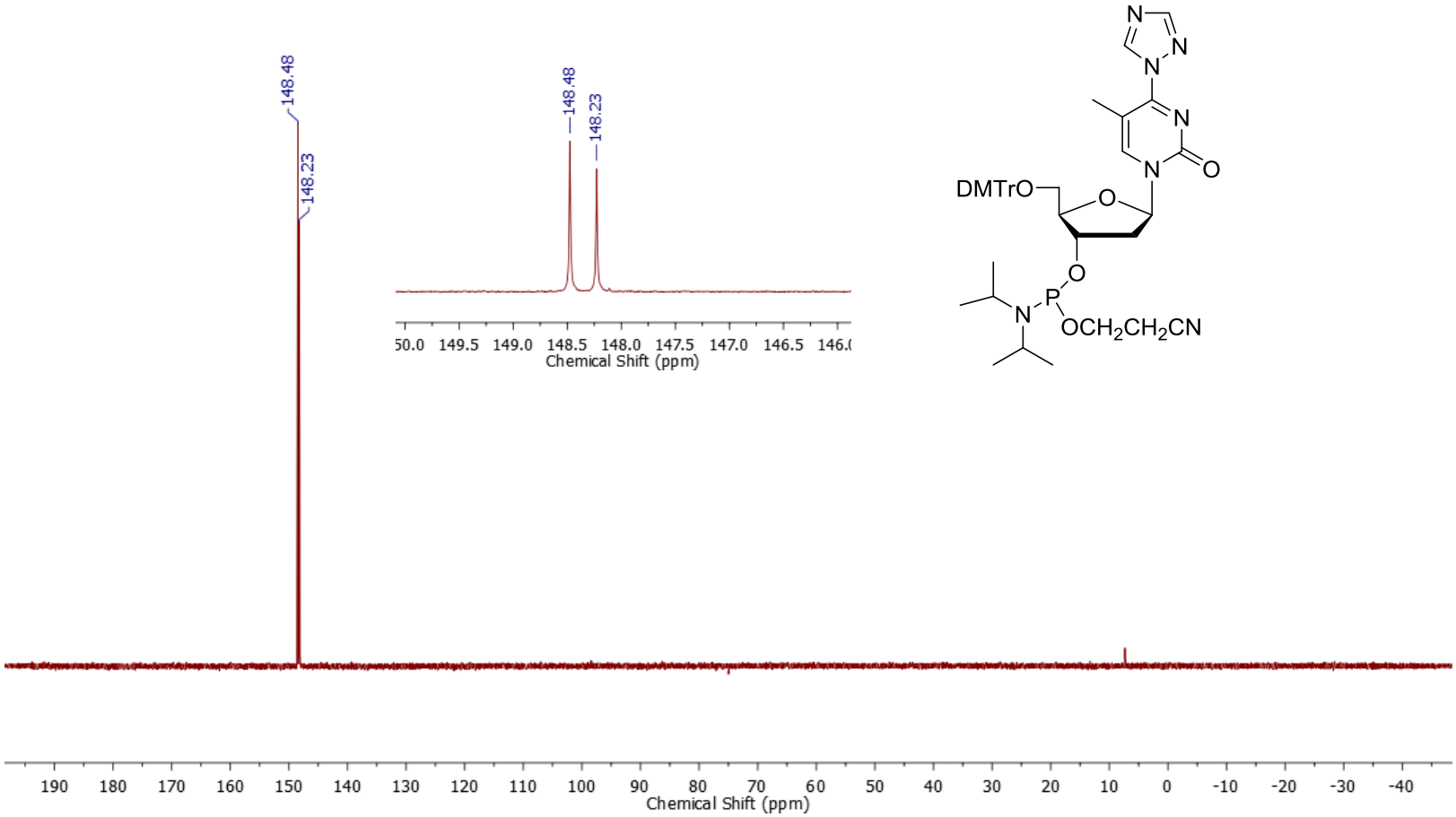


Figure A2.7: RP HPLC profiles of **ODN$_1$** deprotected with *t*-butylamine/H_2O + AcOH. The column was eluted with a linear gradient of 0-70 % buffer B over 30 min (buffer A: 50 mM sodium phosphate, pH 5.8, 2 % ACN and buffer B: 50 mM sodium phosphate, pH 5.8, 50 % ACN).

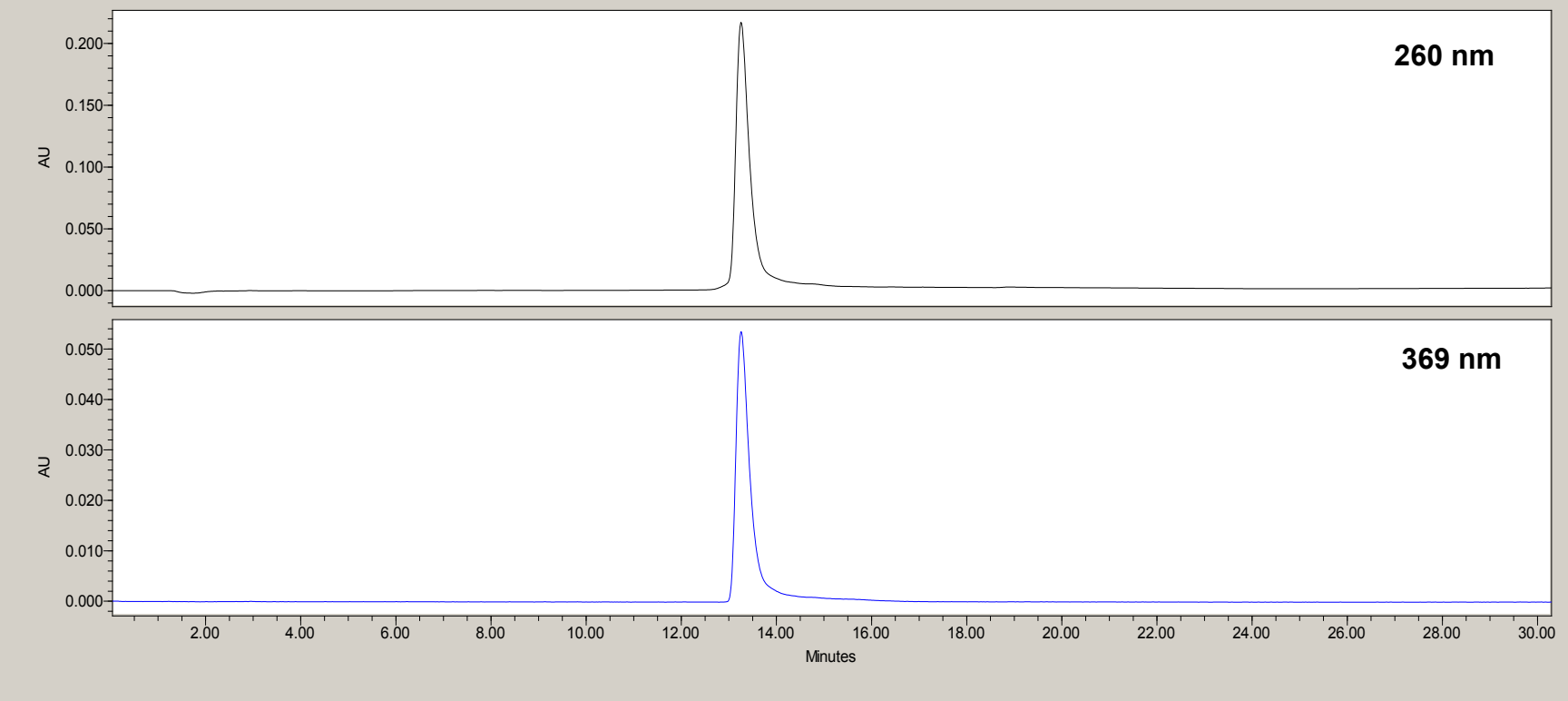


Figure A2.8: ESI-MS spectrum of **ODN_1** deprotected with *t*-butylamine/H_2O + AcOH (expected mass: 3042.9)

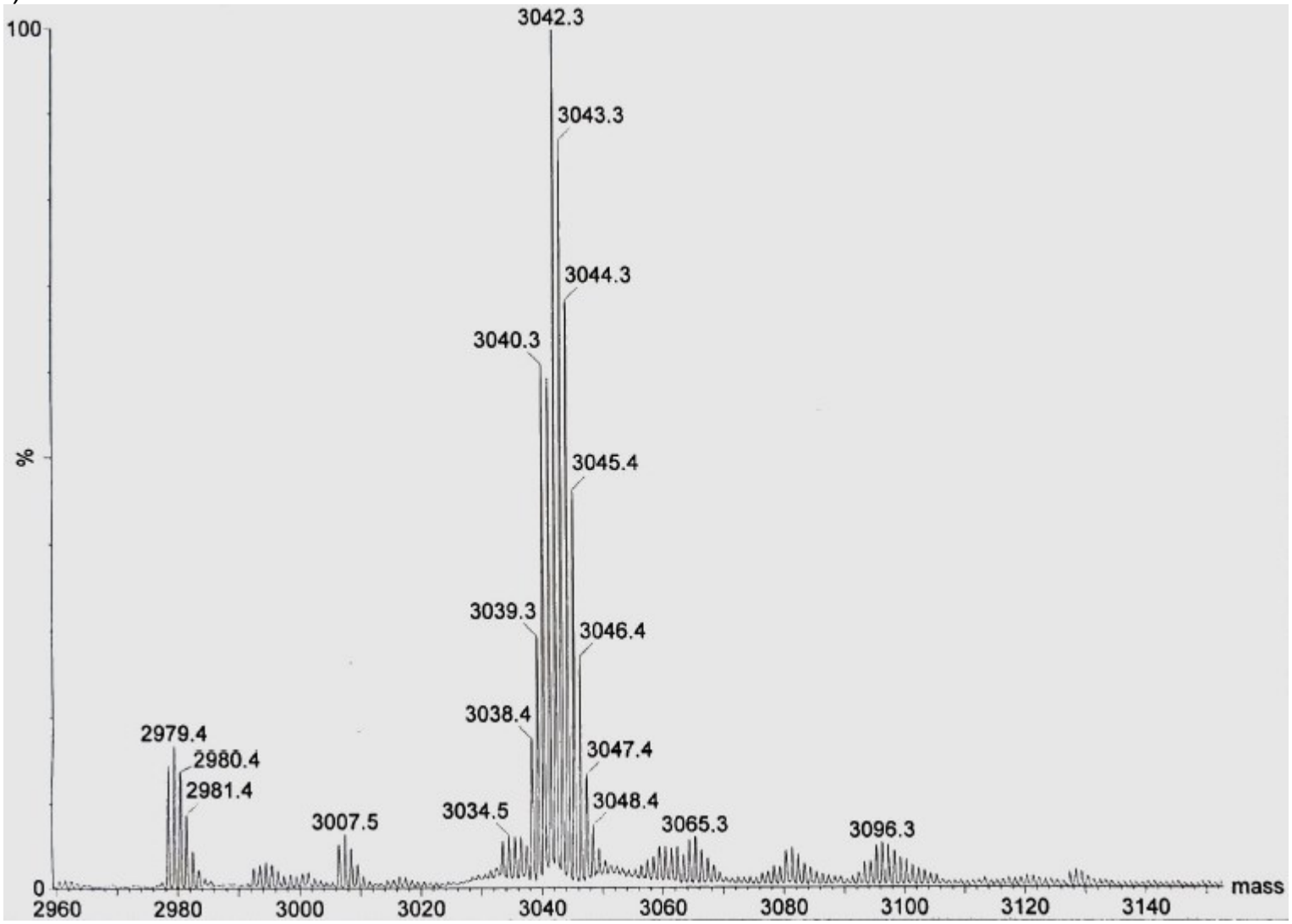


Figure A2.9: RP HPLC profiles of **ODN_2** deprotected with *t*-butylamine/H_2O + AcOH. The column was eluted with a linear gradient of 0-70 % buffer B over 30 min (buffer A: 50 mM sodium phosphate, pH 5.8, 2 % ACN and buffer B: 50 mM sodium phosphate, pH 5.8, 50 % ACN).

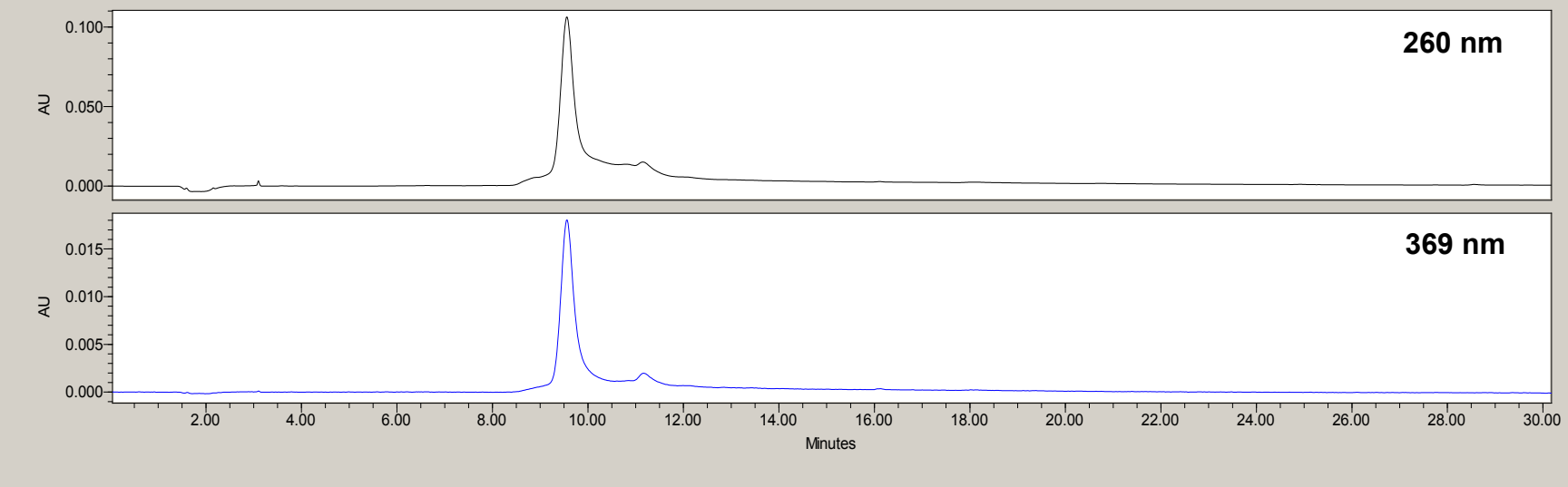


Figure A2.10: ESI-MS spectrum of **ODN_2** deprotected with *t*-butylamine/H_2O + AcOH (expected mass: 3709.3)

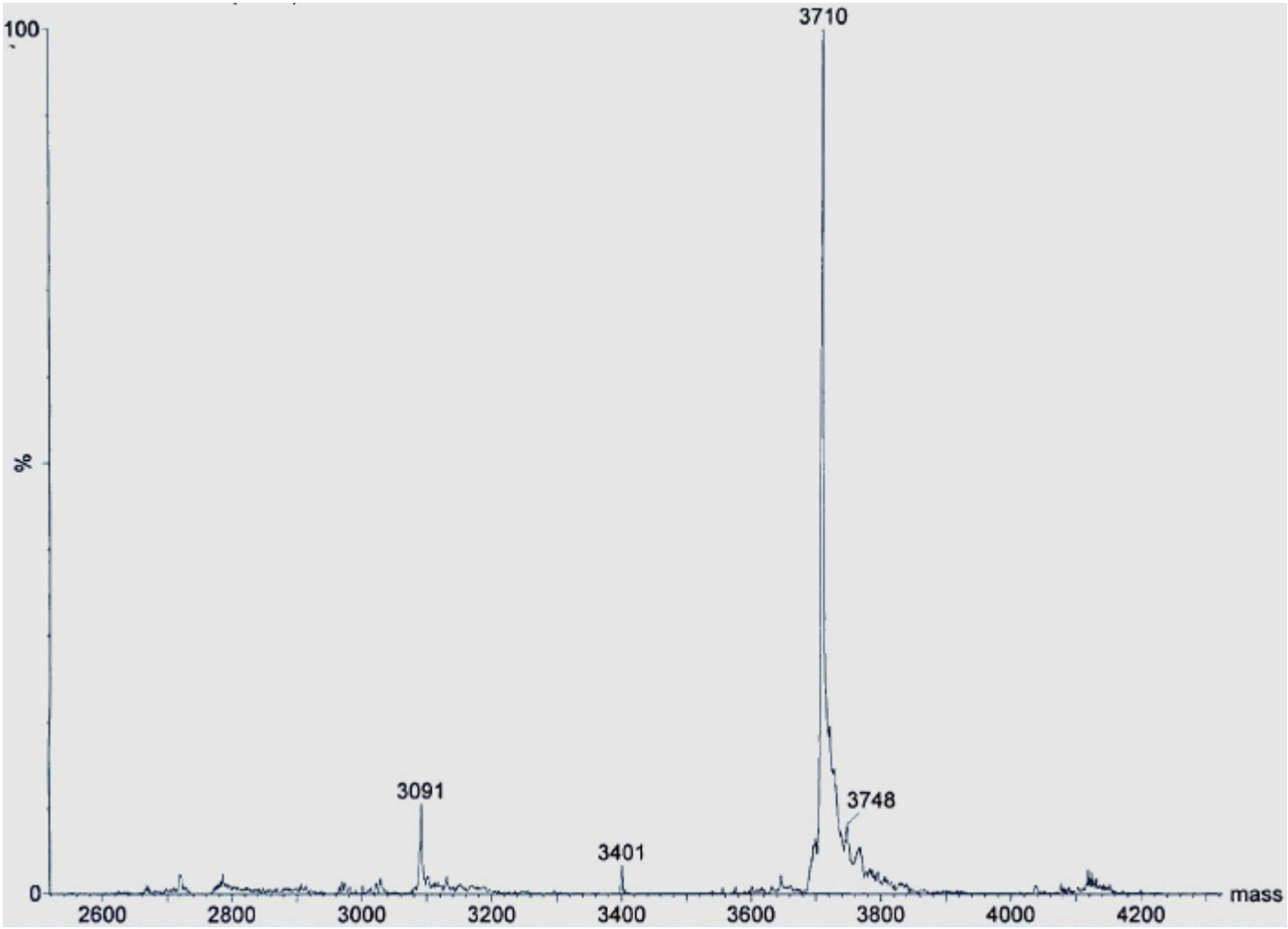


Figure A2.11: UV-Vis spectrum of d^{4Se}T oligonucleotide

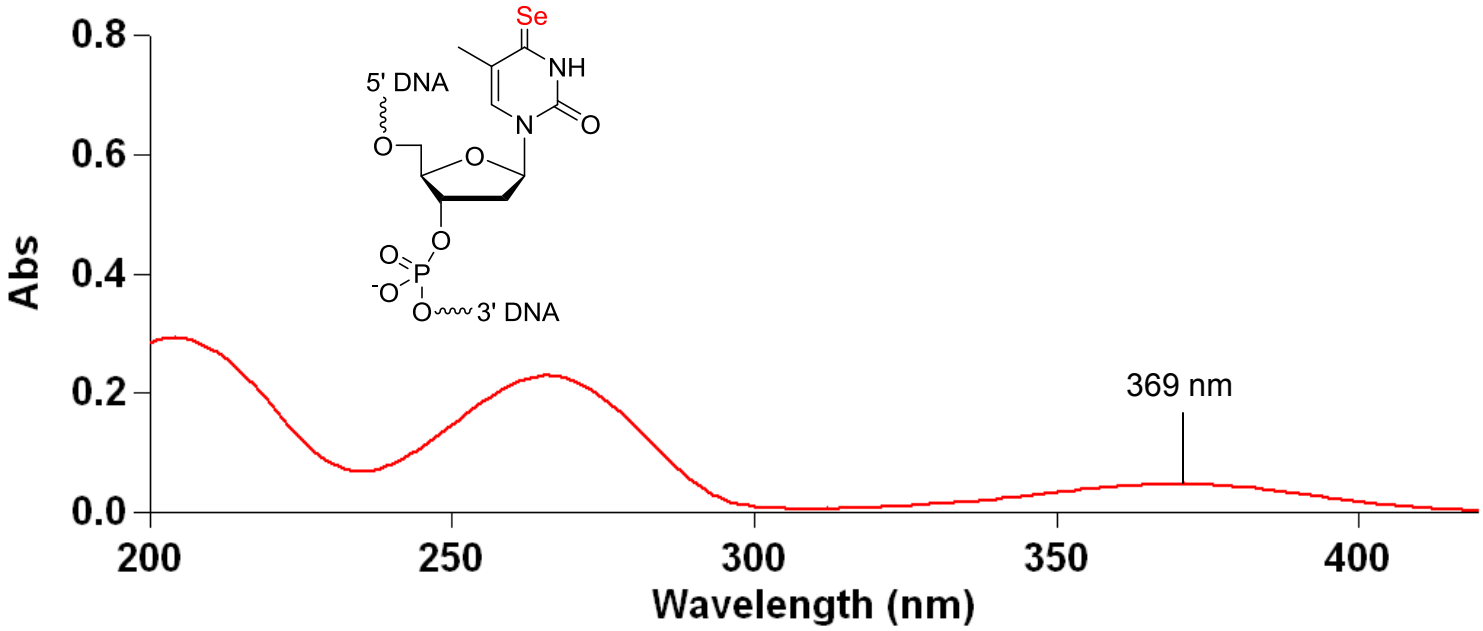


Figure A2.12: ESI-MS spectrum of crude **ODN$_3$** after on-column Se nucleophile treatment and deprotection (entry 5)

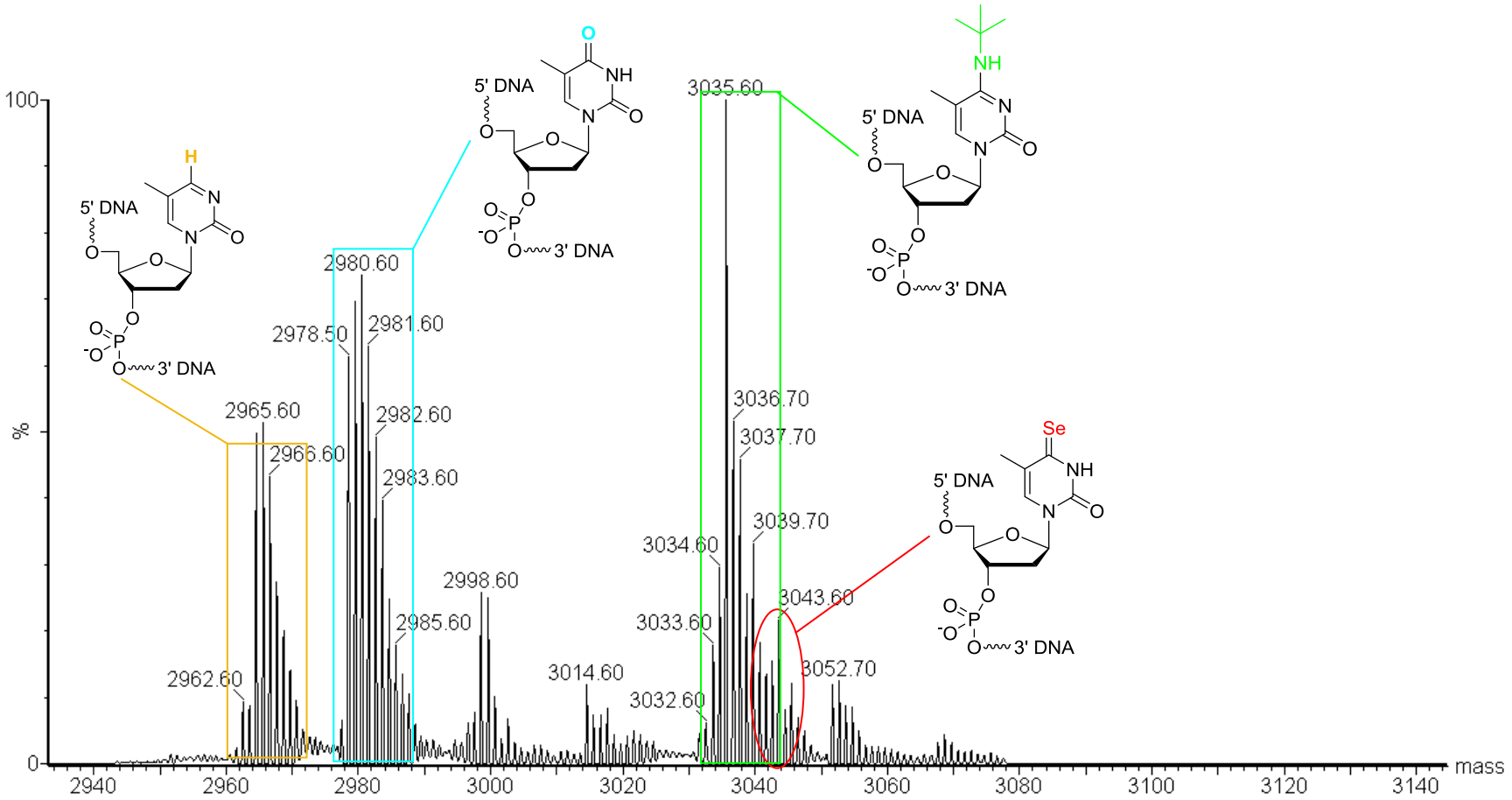
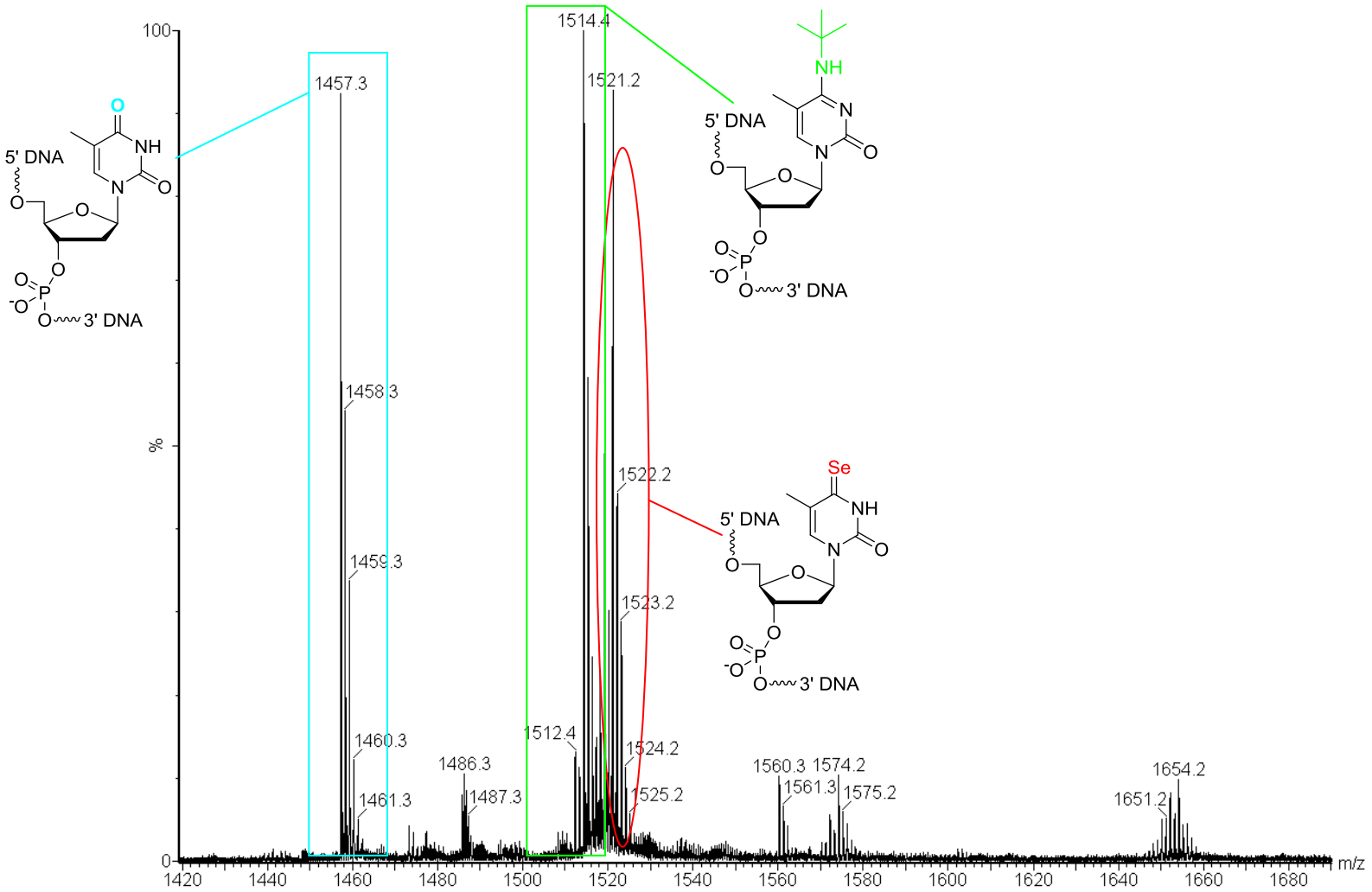
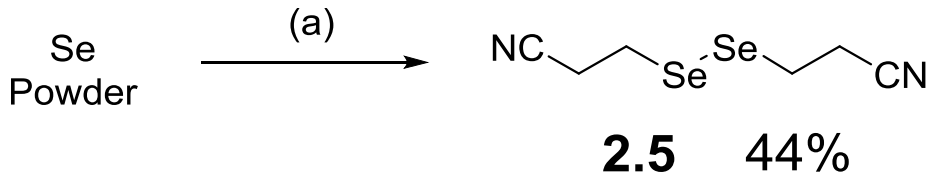


Figure A2.13: ESI-MS spectrum of crude **ODN_4** after on-column Se nucleophile treatment and deprotection (entry 19)



Scheme A2.1: Synthesis of diselenide **2.5**. Reagents and conditions: (a) (i) $NaBH_4$, 1,4-dioxane/EtOH (4:1, v/v), 1 h, rt (ii) 3-bromopropionitrile, 1 h, rt.



2.5 44%

Appendix II: Supporting Information for Chapter 3

Contents Page

Supporting Figures and Schemes

Figure A3.1: 500 MHz ^{1}H NMR spectrum of compound **3.4** (in $CDCl_3$)

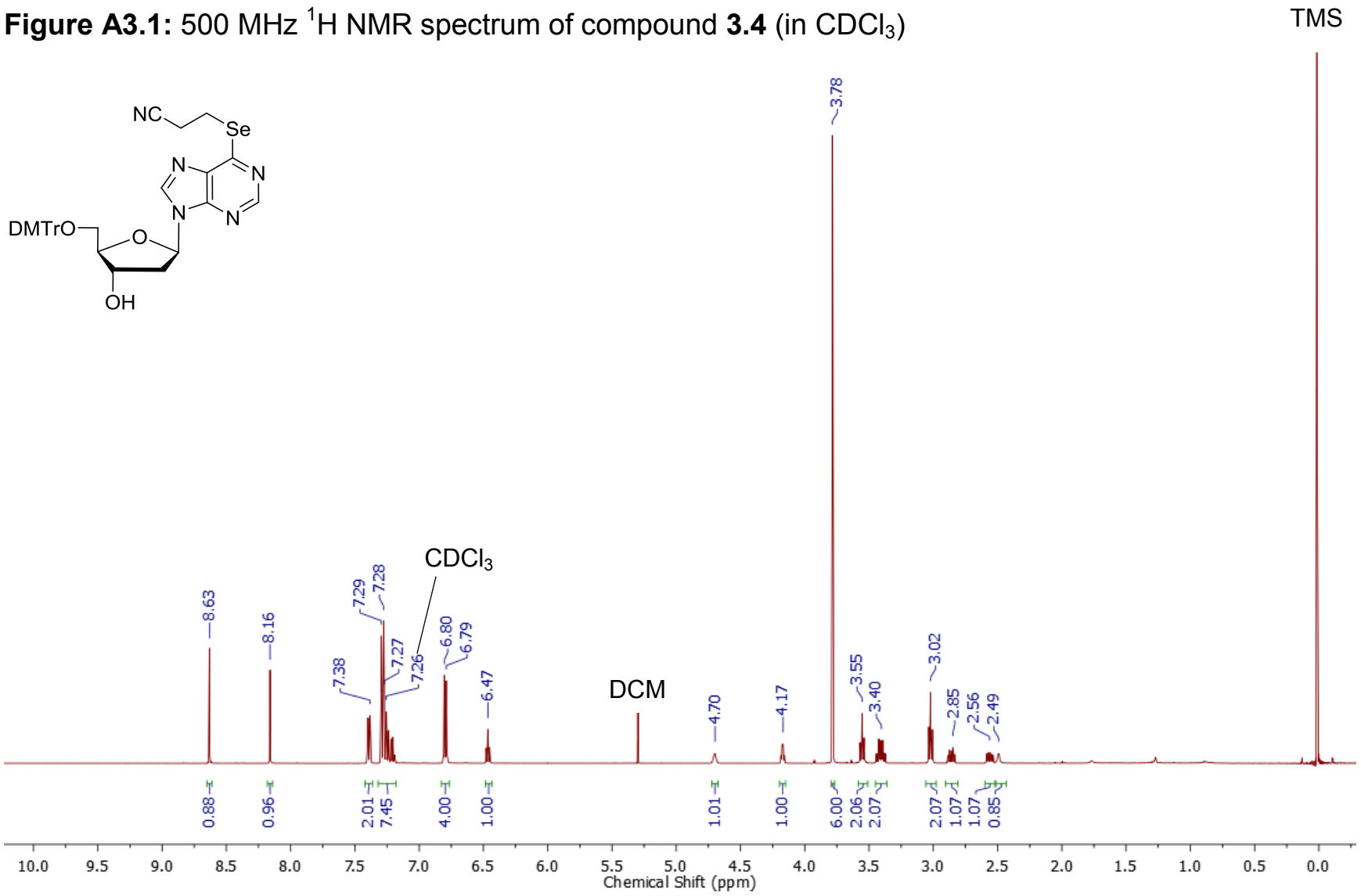


Figure A3.2: 125.7 MHz ^{13}C NMR spectrum of compound **3.4** (in $CDCl_3$)

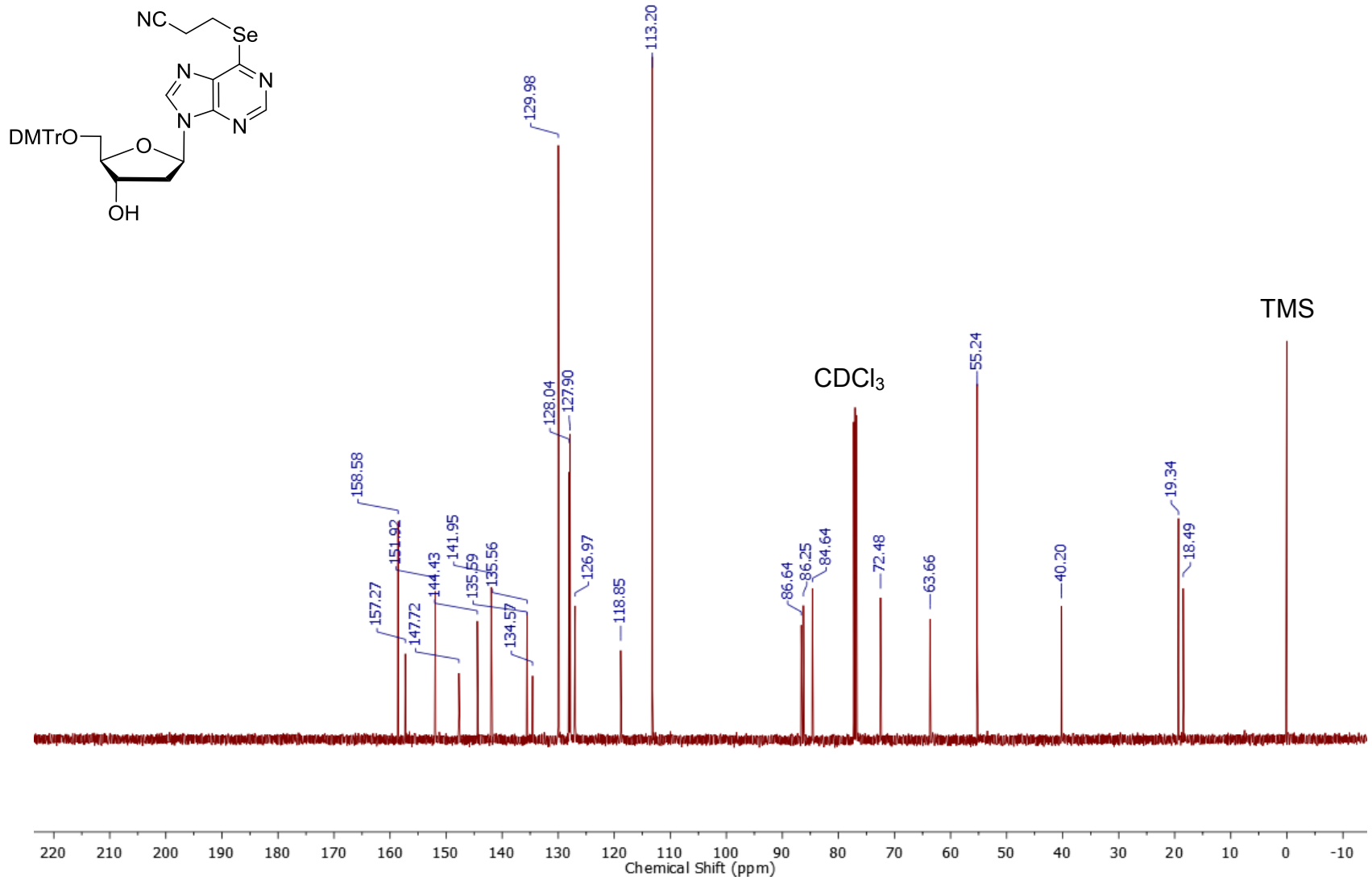


Figure A3.3: 95.3 MHz ^{77}Se NMR spectrum of compound **3.4** (in $CDCl_3$)

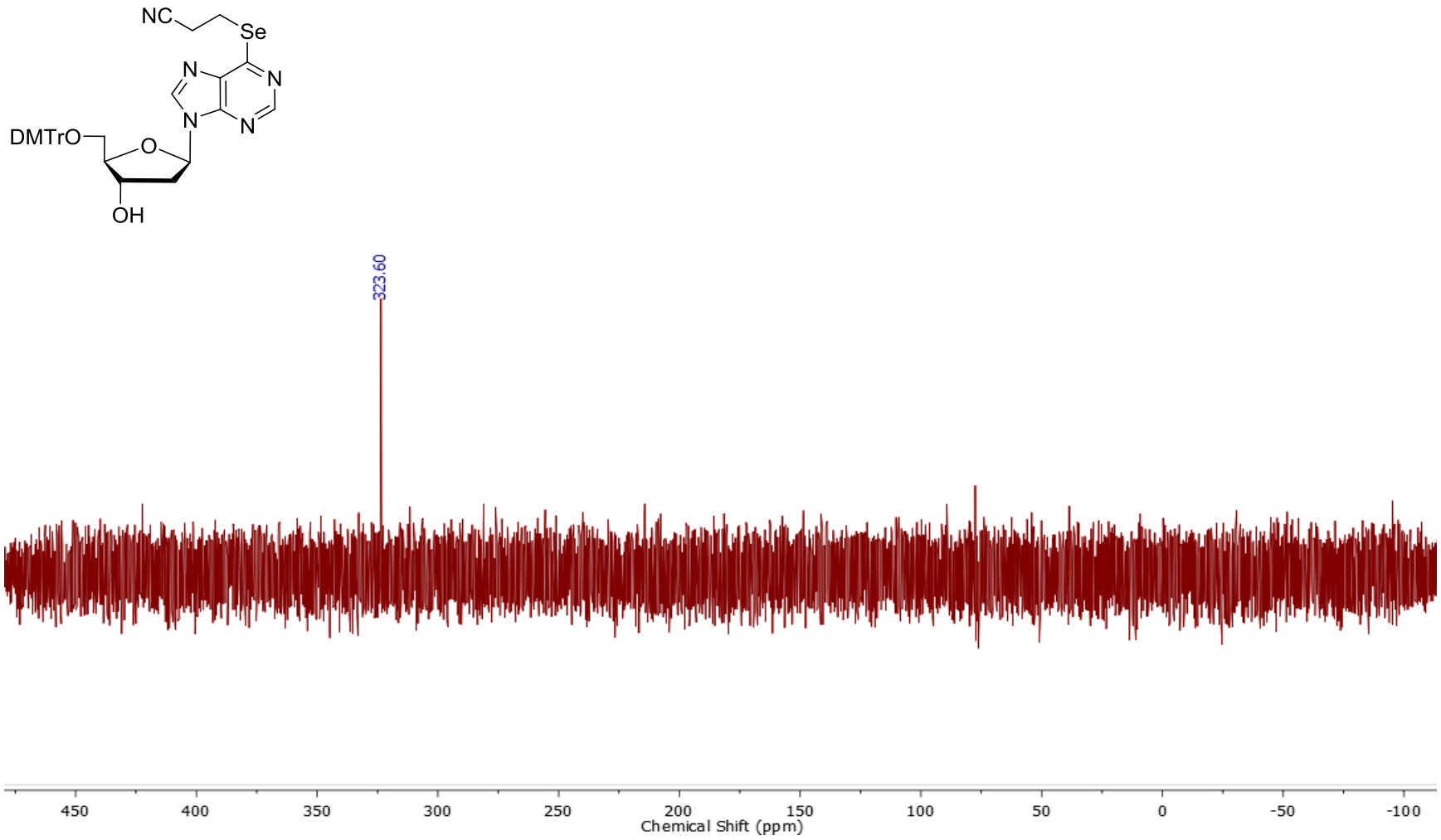


Figure A3.4: HR ESI-MS spectrum of compound **3.4** (expected mass: 694.1539)

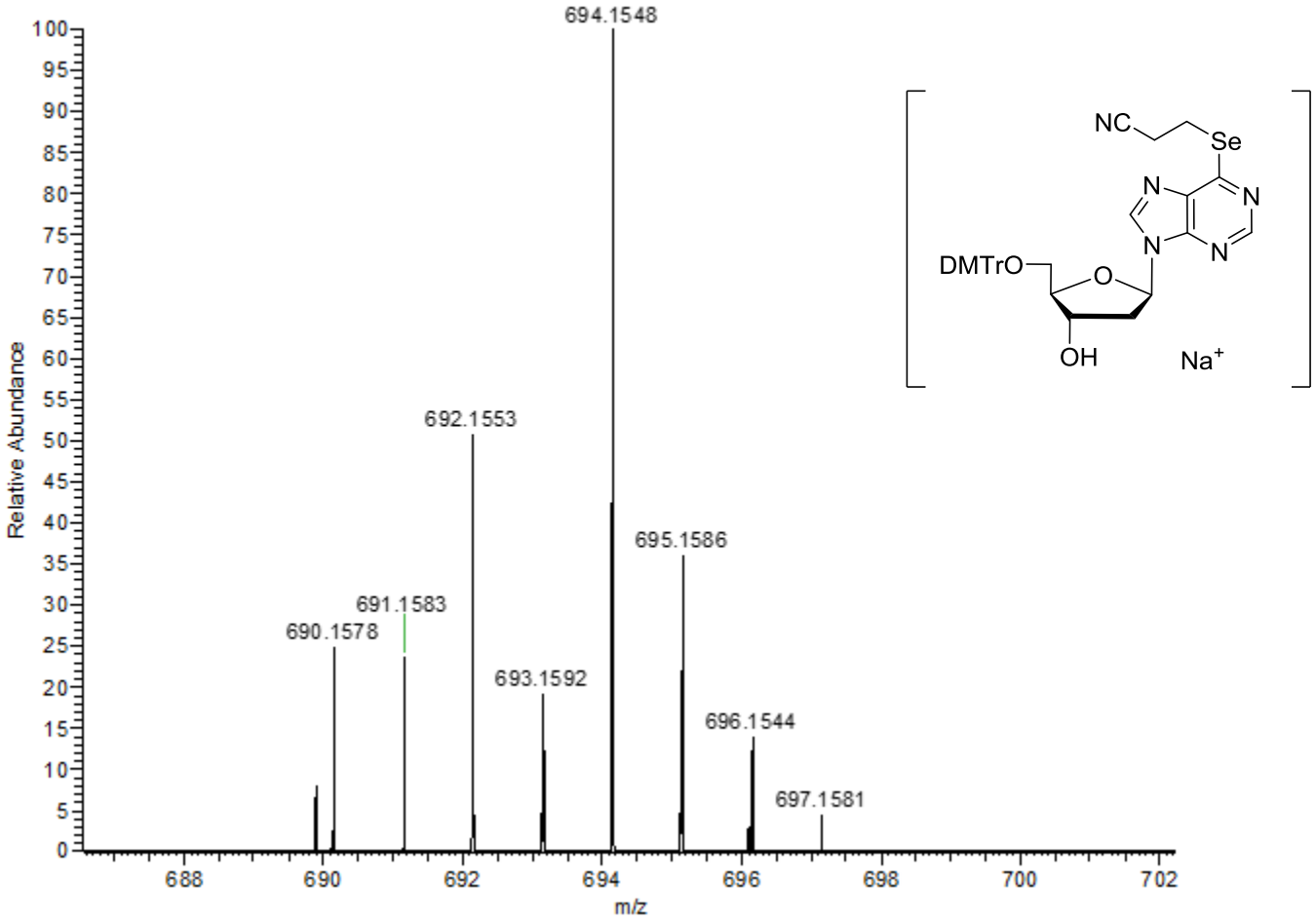


Figure A3.5: 500 MHz ^{1}H NMR spectrum of compound **3.5** (in d_6-acetone)

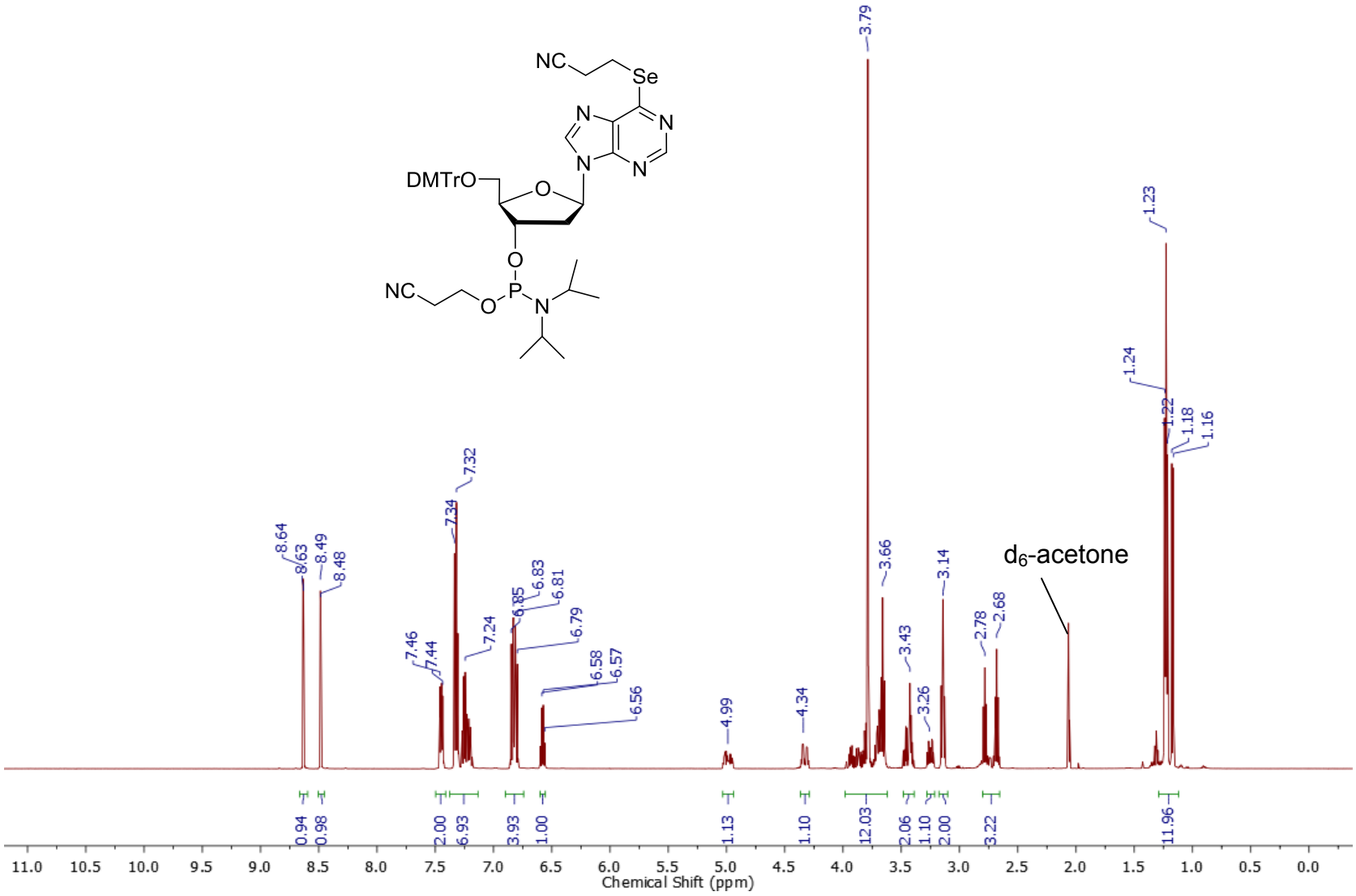


Figure A3.6: 125.7 MHz ^{13}C NMR spectrum of compound **3.5** (in d_6-acetone)

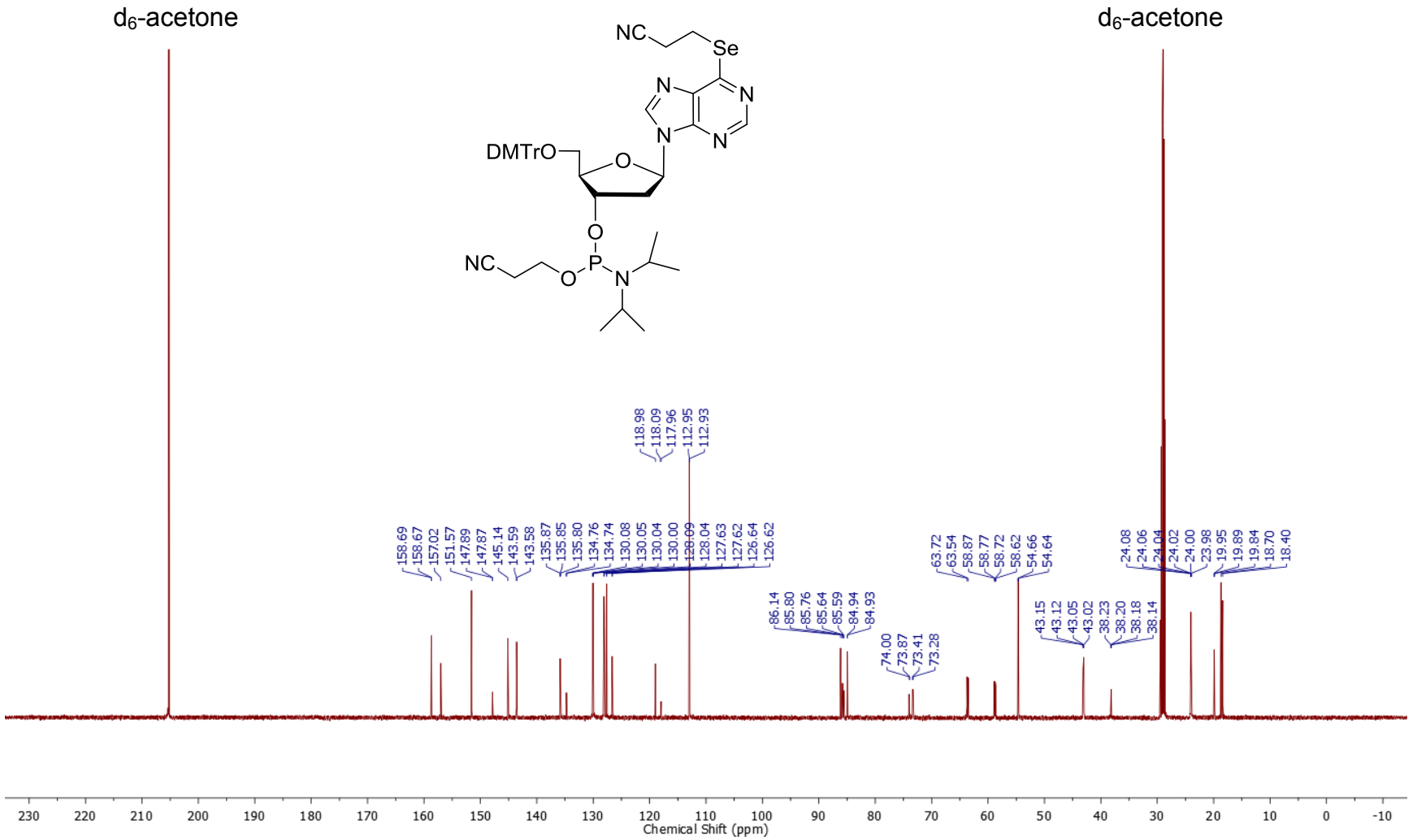


Figure A3.7: 202.3 MHz ^{31}P NMR spectrum of compound **3.5** (in d_6-acetone)

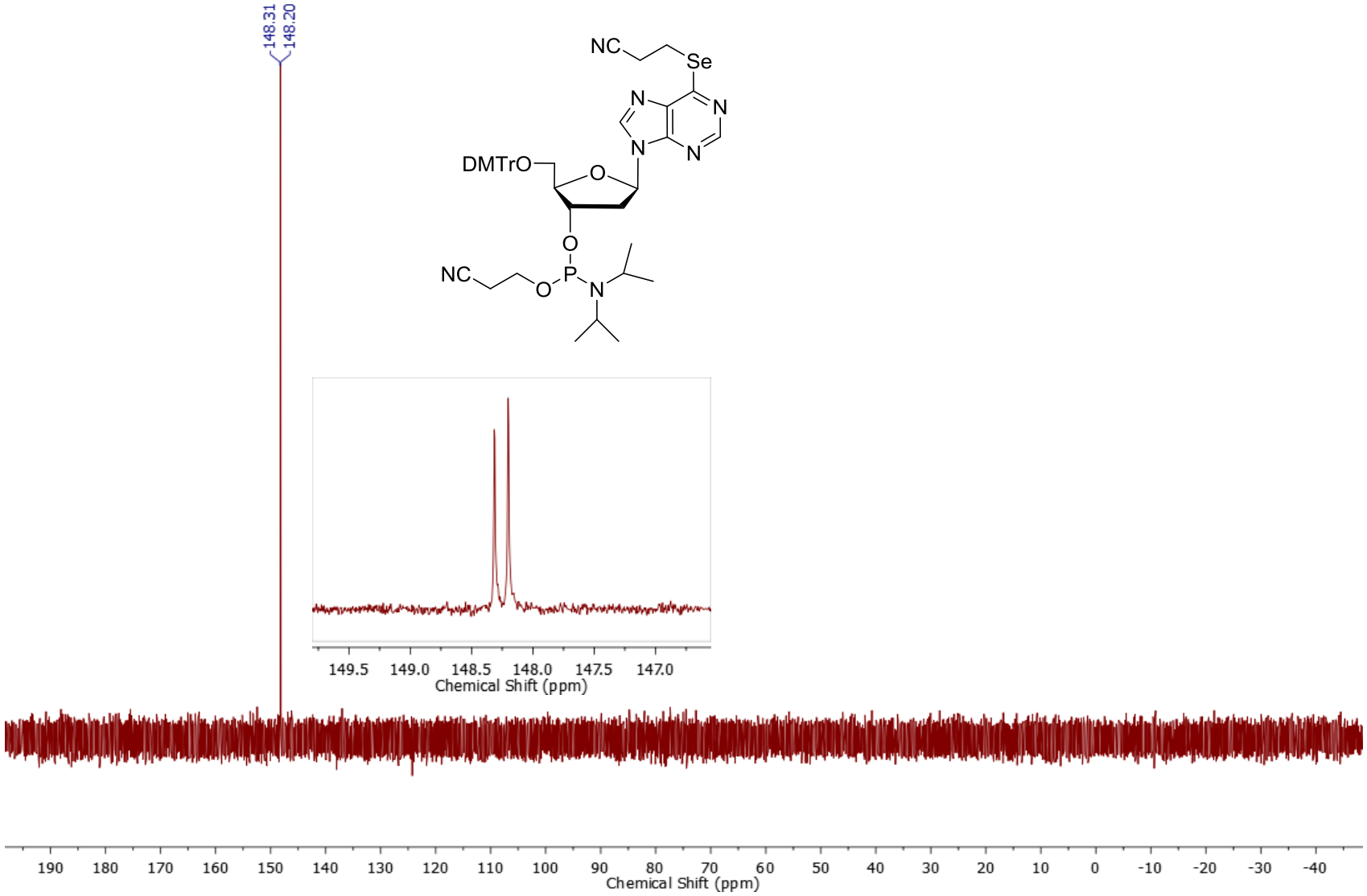


Figure A3.8: 95.3 MHz ^{77}Se NMR spectrum of compound **3.5** (in d_6-acetone)

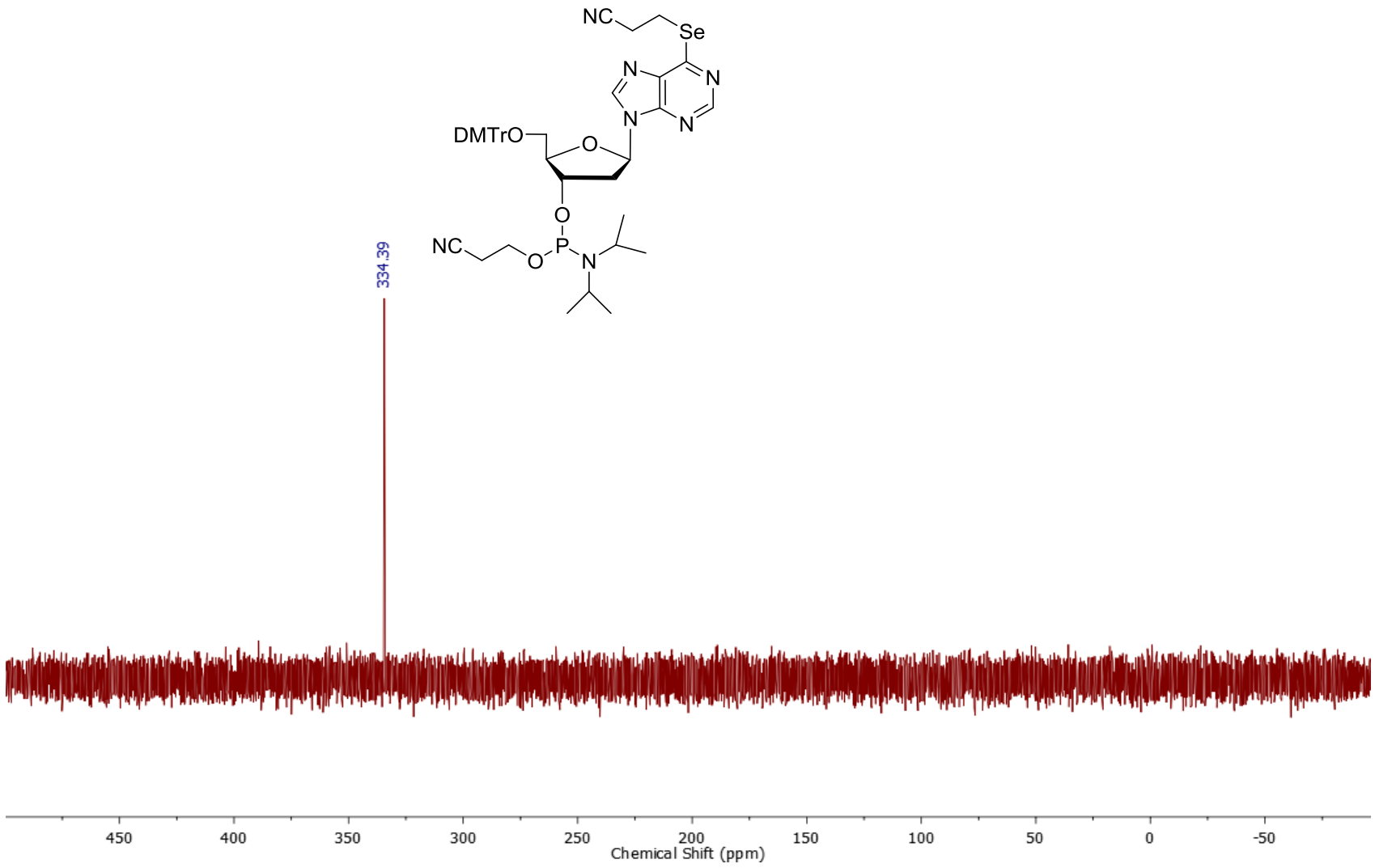


Figure A3.9: HR ESI-MS spectrum of compound **3.5** (expected mass: 872.2798)

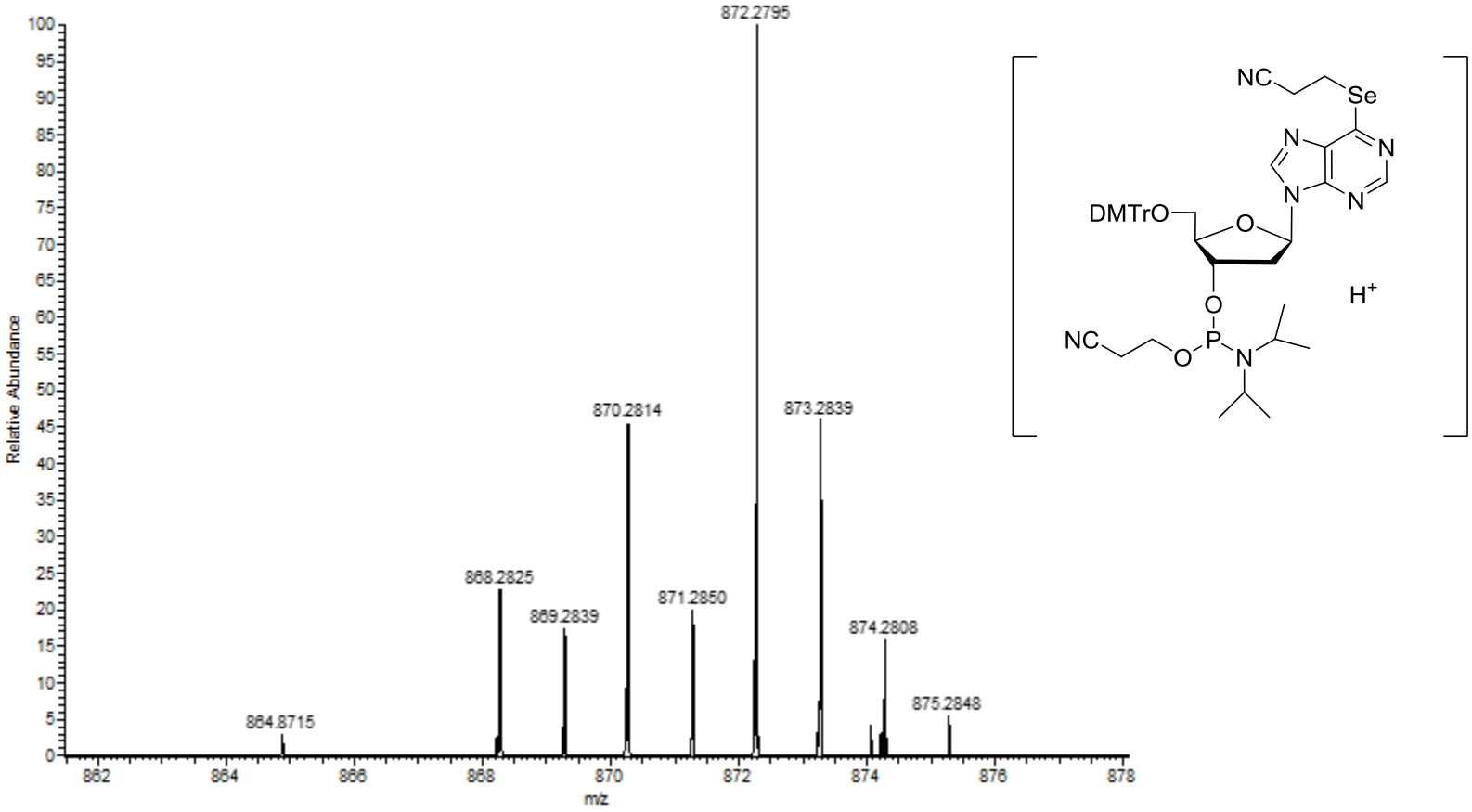


Figure A3.10: UV-Vis spectra of **dI-Se-ODN** and **dI-ODN**. The **dI-Se-ODN** spectrum is shown in blue, the **dI-ODN** spectrum is shown in red and the blank 18 MΩ H_2O spectrum is shown in black. Samples were dissolved in 1 ml 18 MΩ H_2O.

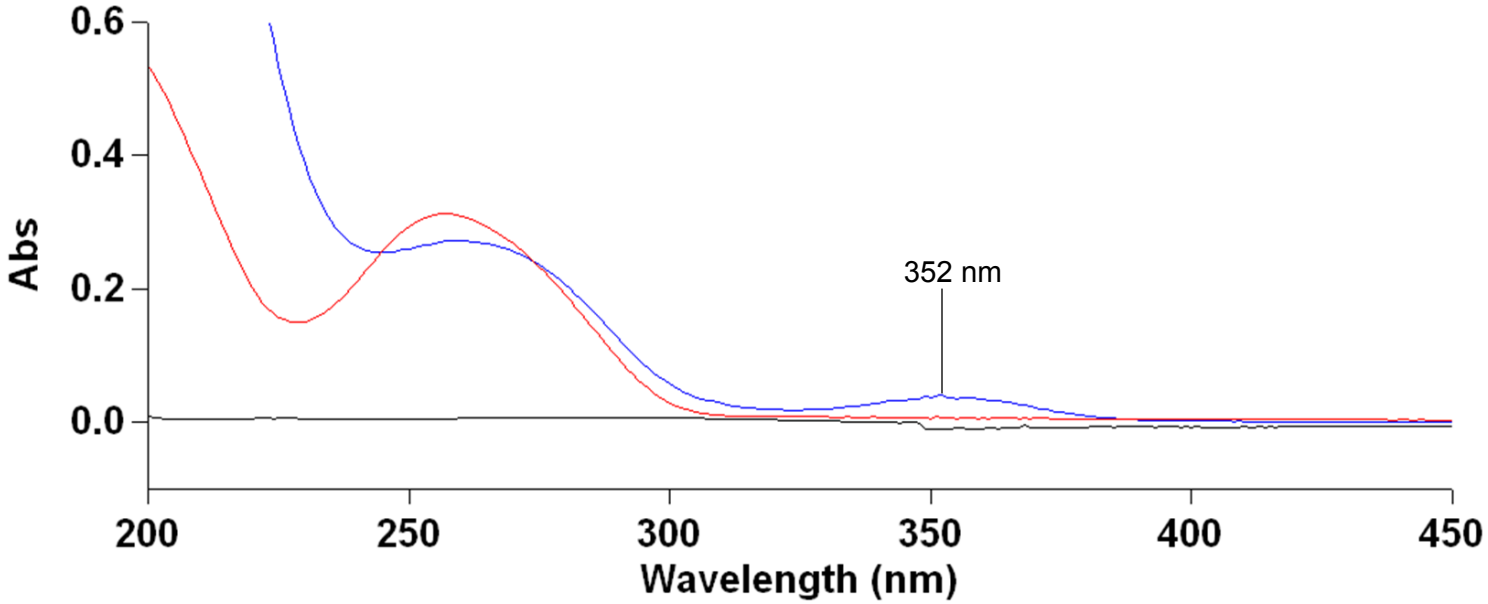


Figure A3.11: SAX HPLC profiles of purified **dI-Se-ODN$_2$**. (a) 260 nm profile. (b) 352 nm profile. (c) Expansion of 352 nm profile. The column was eluted at room temperature using a linear gradient of 0-52 % buffer B over 30 minutes (buffer A: 100 mM Tris HCl, pH 7.5, 10 % ACN; buffer B: 100 mM Tris HCl, pH 7.5, 10 % ACN, 1 M NaCl).

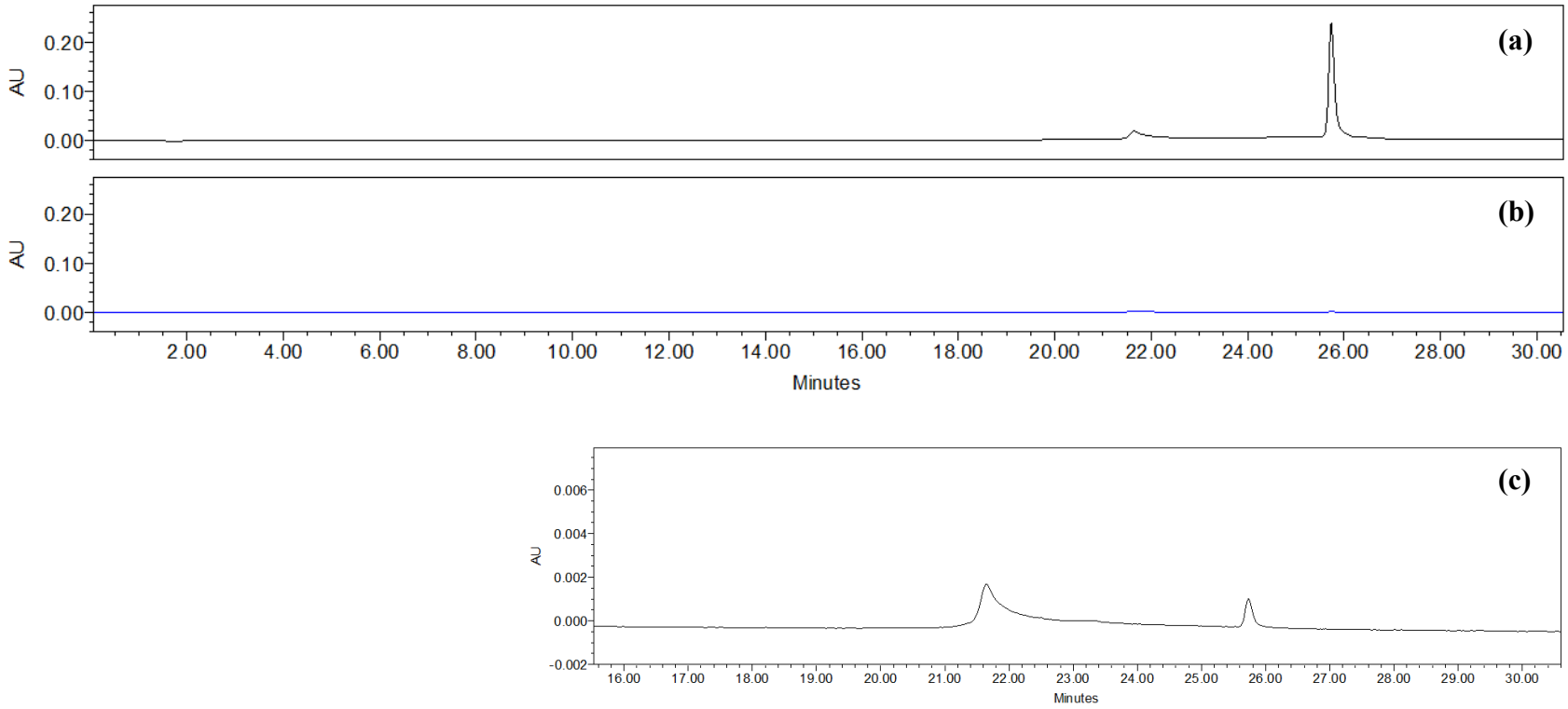


Figure A3.12: ESI-MS spectrum of **dI-Se-ODN$_2$** (expected mass: 8671.5)

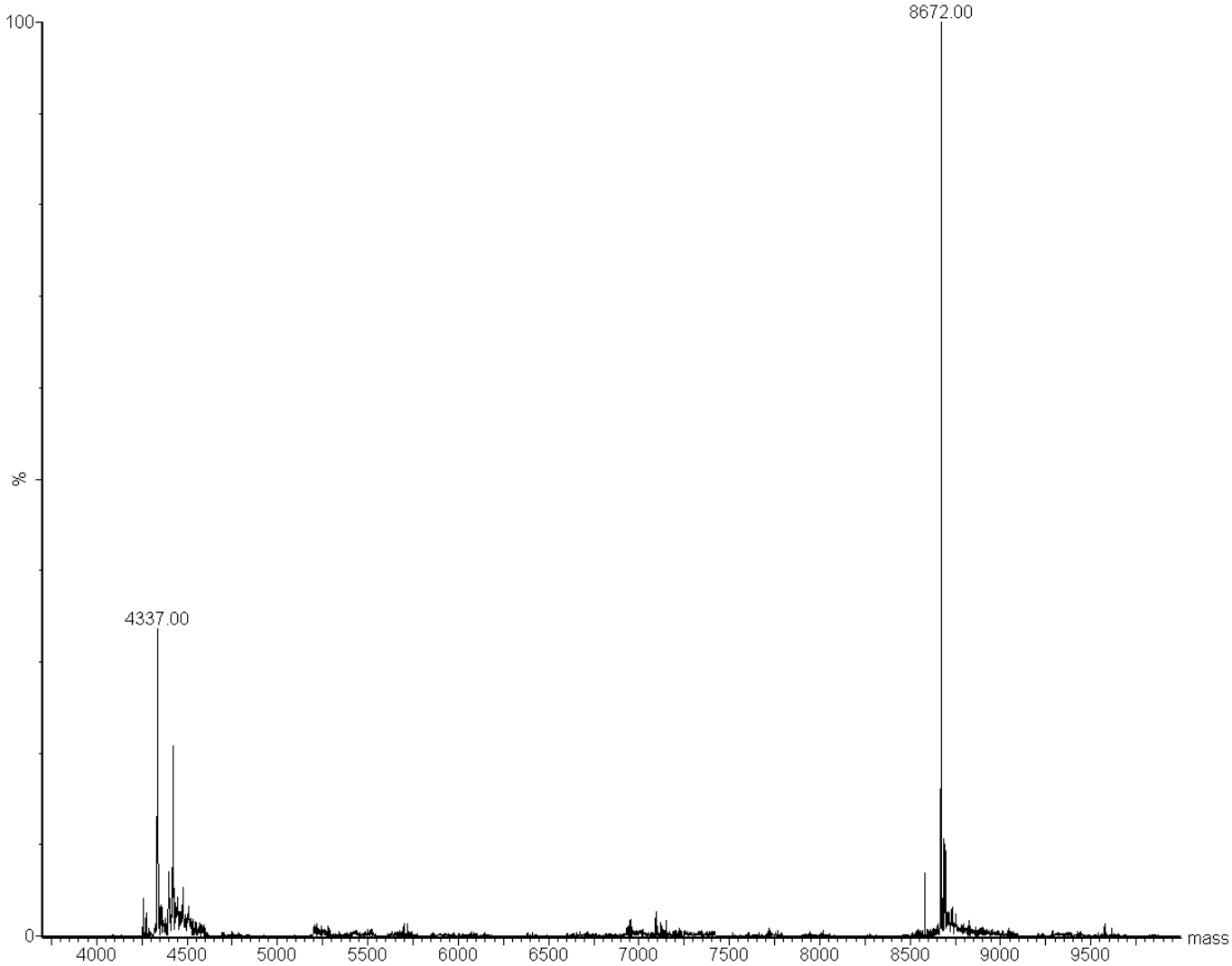


Figure A3.13: SAX-HPLC profiles of purified **dI-Se-ODN**. (a) 260 nm profile. (b) 352 nm profile. (c) Expansion of 352 nm profile. The column was eluted at room temperature using a linear gradient of 0-52 % buffer B over 30 minutes (buffer A: 100 mM Tris HCl, pH 7.5, 10 % ACN; buffer B: 100 mM Tris HCl, pH 7.5, 10 % ACN, 1 M NaCl).

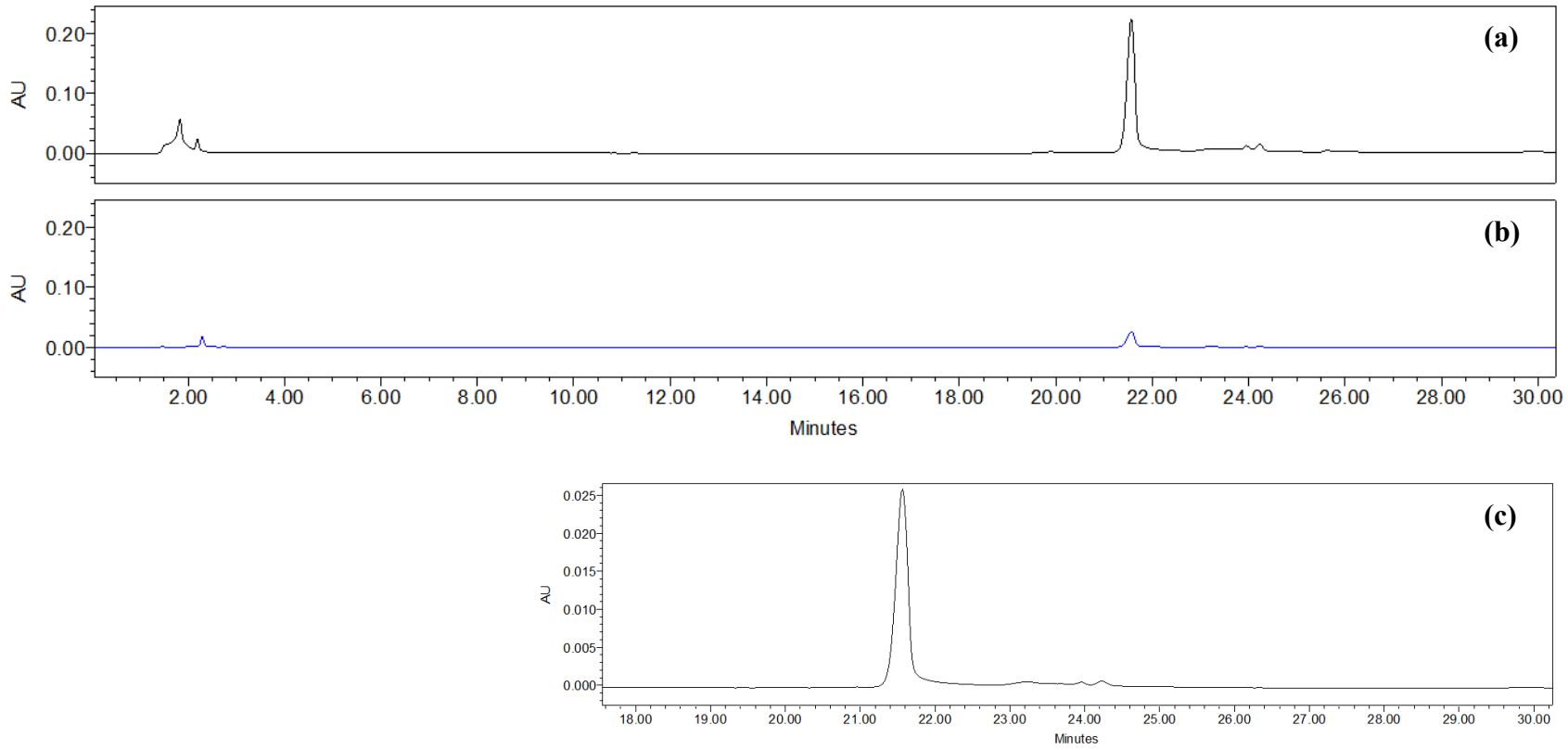


Figure A3.14: ESI-MS spectrum of **dI-Se-ODN** (expected mass: 4336.7)

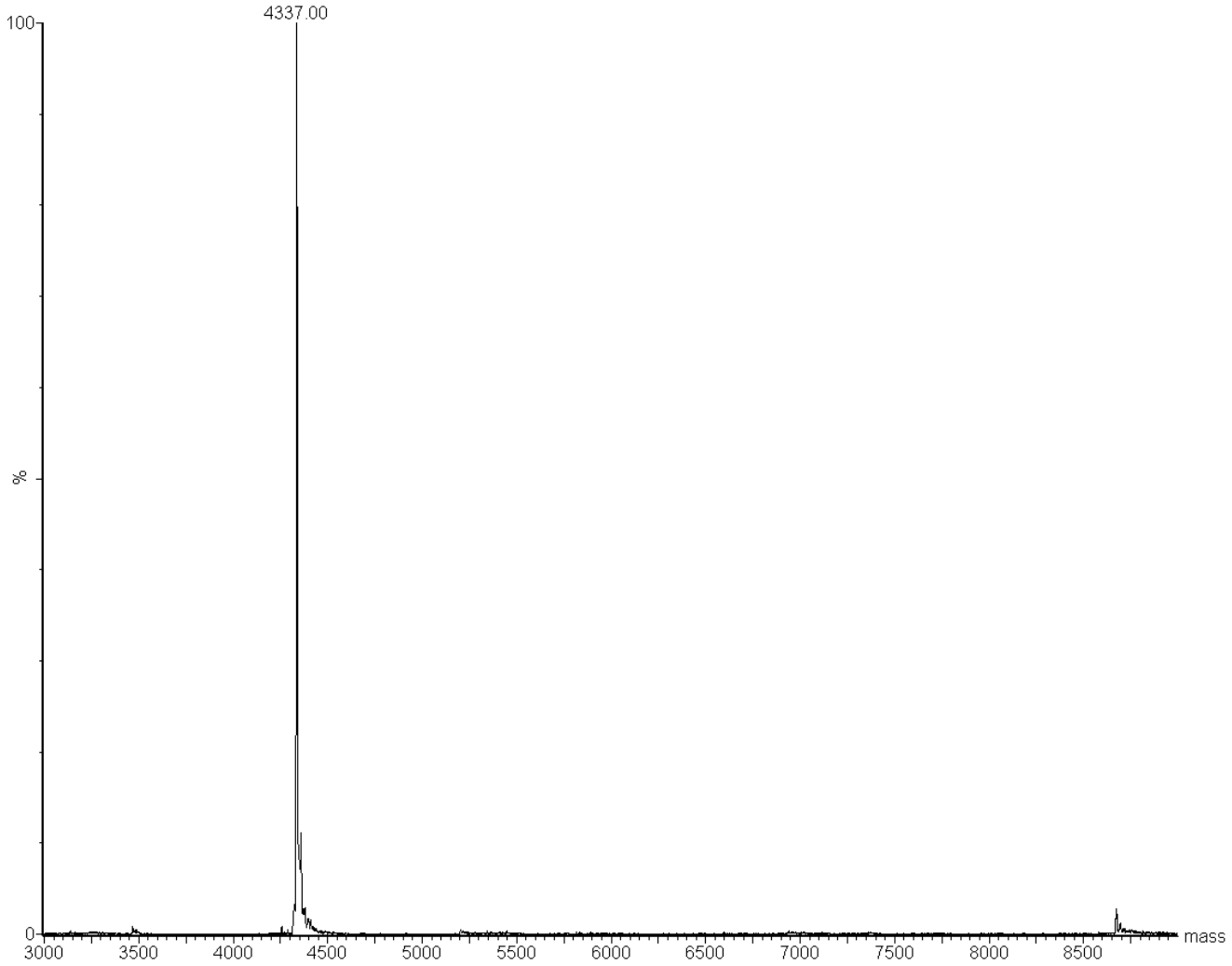


Figure A3.15: SAX-HPLC profiles of purified **dI-ODN**. (a) 260 nm profile. (b) 352 nm profile. (c) Expansion of 352 nm profile. The column was eluted at room temperature using a linear gradient of 0-52 % buffer B over 30 minutes (buffer A: 100 mM Tris HCl, pH 7.5, 10 % ACN; buffer B: 100 mM Tris HCl, pH 7.5, 10 % ACN, 1 M NaCl).

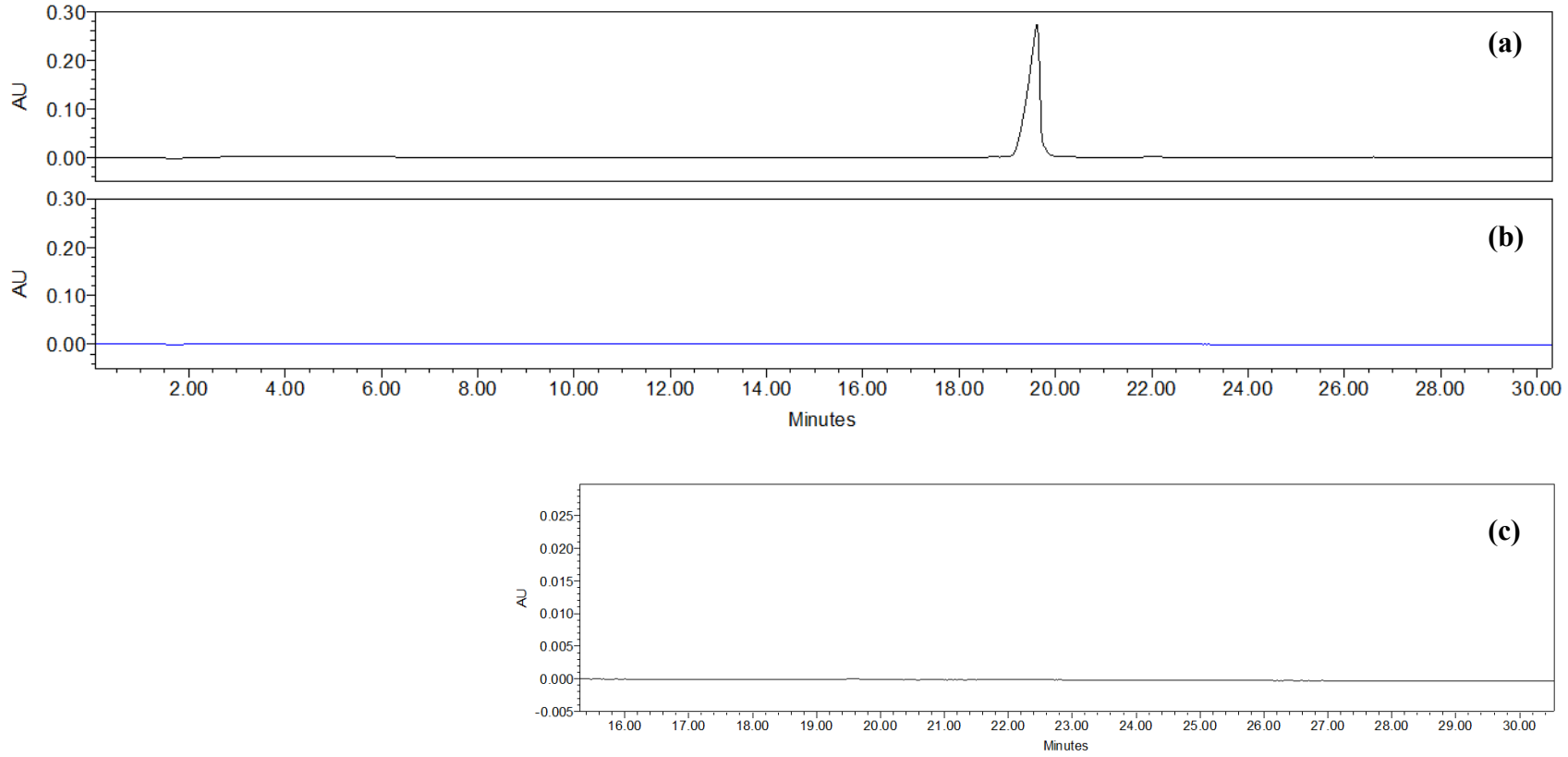


Figure A3.16: ESI-MS spectrum of **dI-ODN** (expected mass: 4273.8)

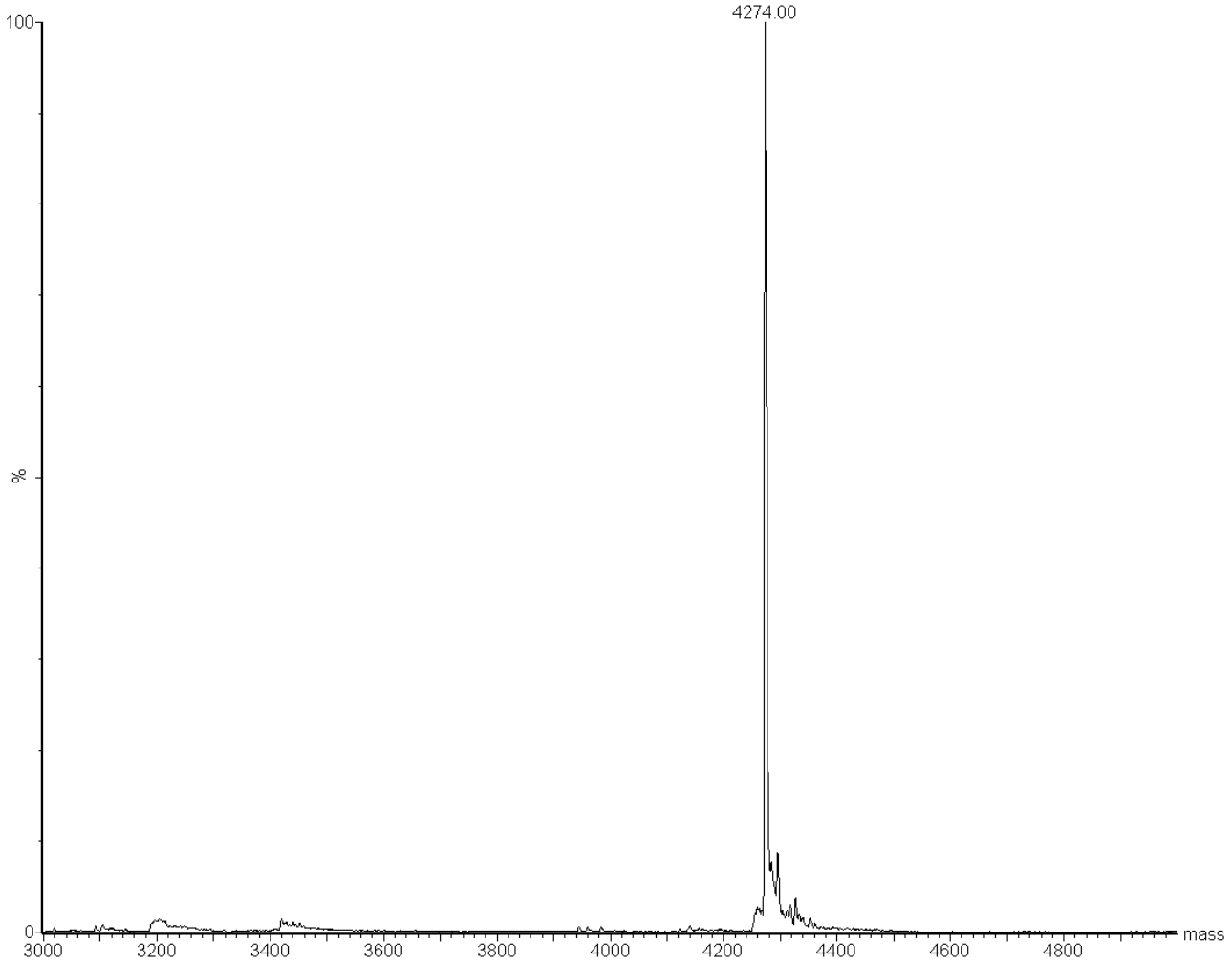


Figure A3.17: SAX HPLC traces of the crude d^{6Se}I containing oligonucleotide at different times after deprotection. The deprotection condition used was *t*-butylamine/H_2O (1:3, v/v) at 60^{o}C followed by the addition of an equimolar amount of glacial acetic acid. Red trace: 1 hour after deprotection; blue trace: 2 hours after deprotection; green trace: 3 hours after deprotection; black trace: 18 hours after deprotection followed by the addition of 50 mM DTT. The column was eluted at room temperature using a linear gradient of 0-52 % buffer B over 30 minutes (buffer A: 100 mM Tris HCl, pH 7.5, 10 % ACN; buffer B: 100 mM Tris HCl, pH 7.5, 10 % ACN, 1 M NaCl).

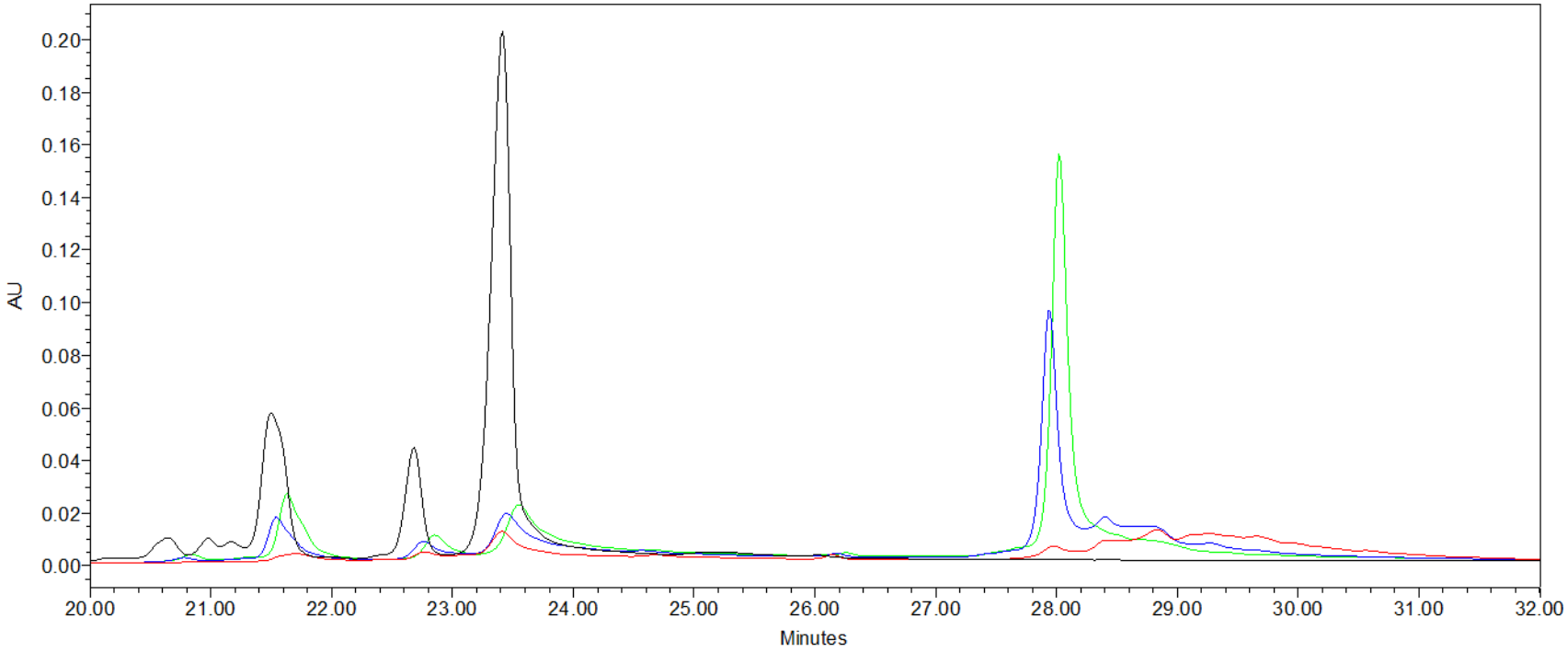


Figure A3.18: T_m values of duplexes containing $d^{6Se}I$ or dI across different base pairing partners. The sequence of the DNA duplex was 5′ dGGCT**X**GATCACCAG 3′ / 5′ dCTGGTGATC**Y**AGCC 3′ where X is $d^{6Se}I$ or dI and Y is dC, dA, dG or dT. The solutions contained a final duplex concentration of 3.7 μM in a pH 7.0 buffer consisting of 90 mM NaCl, pH 7.0, 10 mM sodium phosphate, 1 mM EDTA and 2.5 mM DTT. Data was acquired at 0.5 °C/min from 15 °C to 95 °C, monitoring UV absorption at 260 nm.

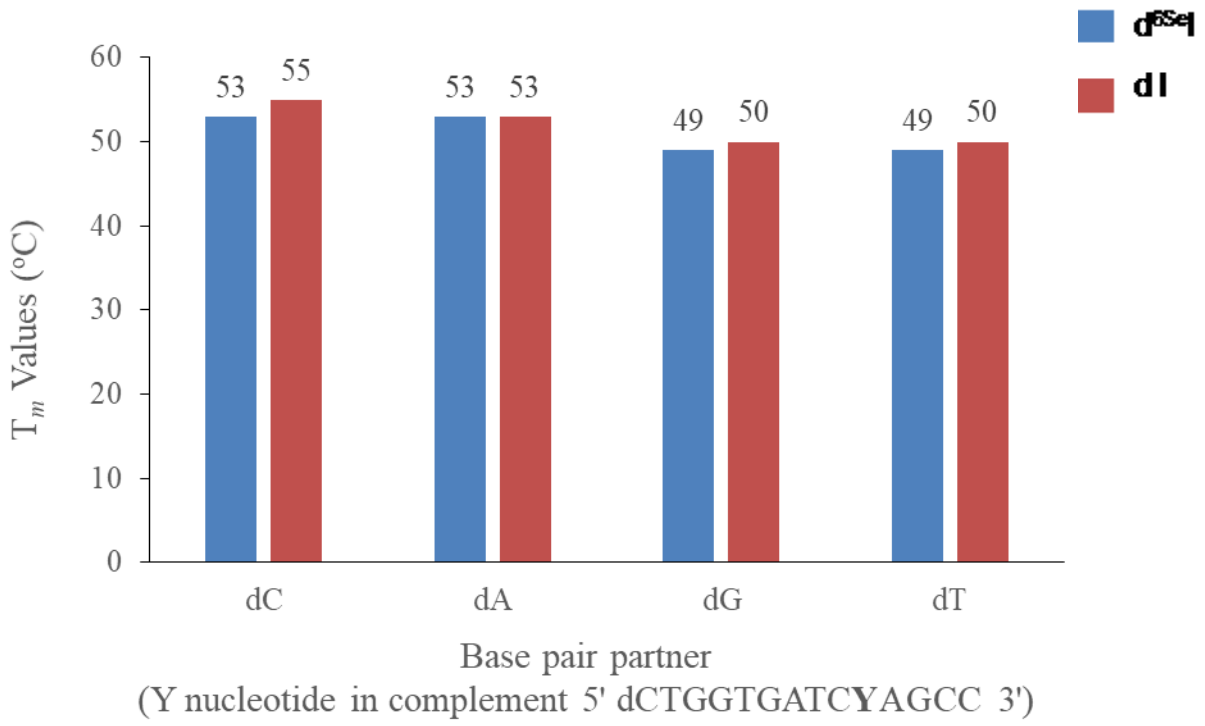
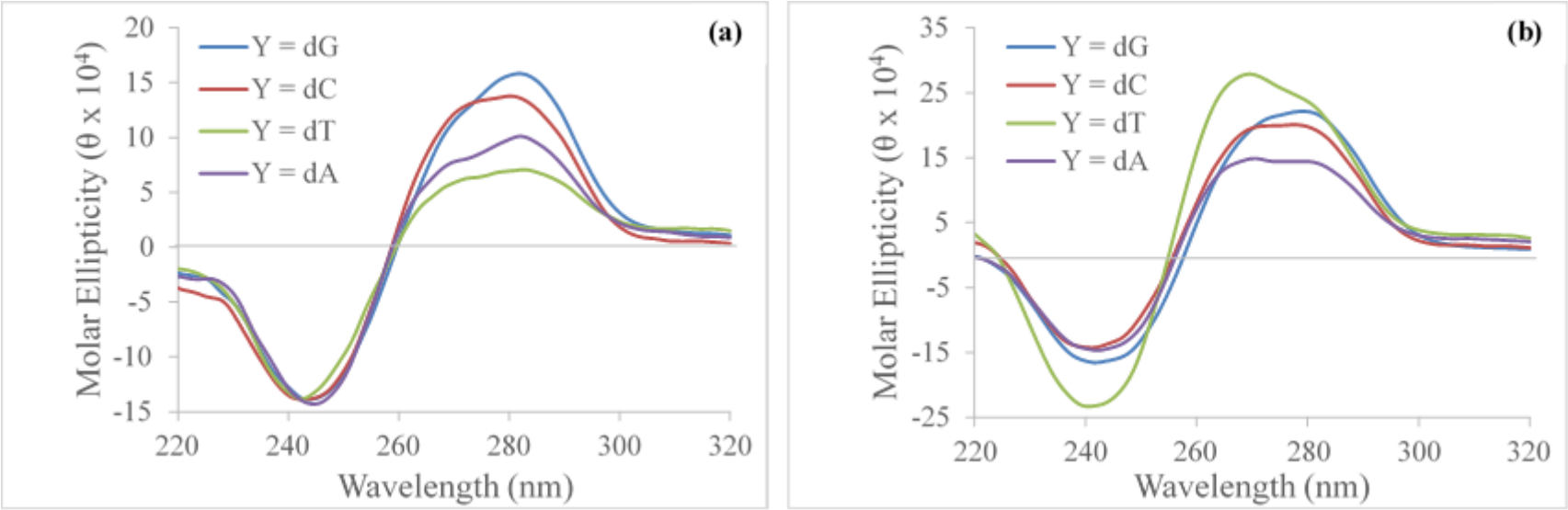
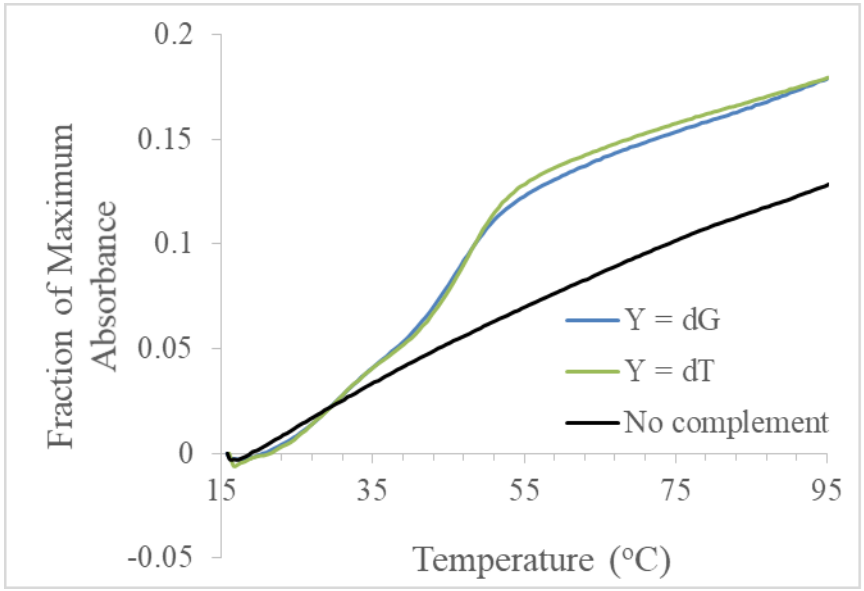
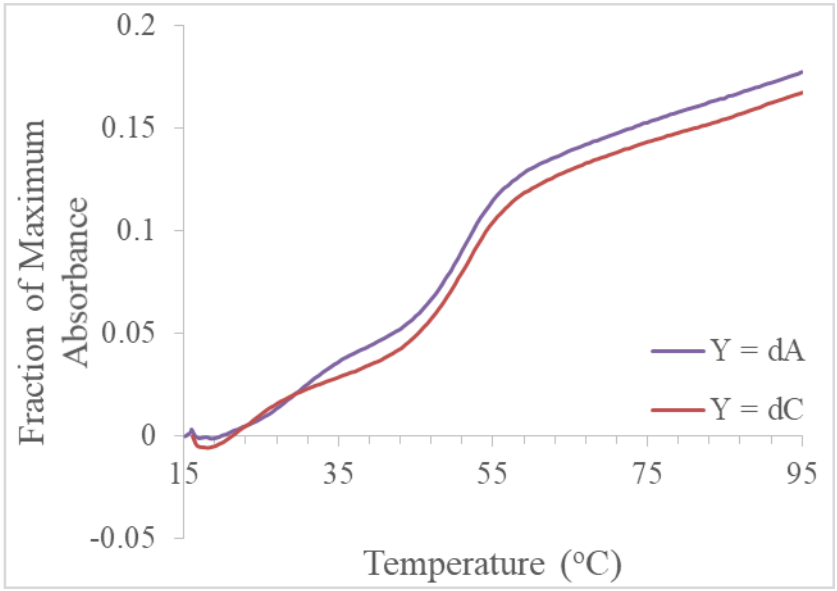


Figure A3.19: CD spectra of DNA duplexes. (a) 2'-Deoxy-6-selenoinosine nucleotide in duplexes 5' dGGCT^{6Se}**I**GATCACCAG 3' / 5' dCTGGTGATC**Y**AGCC 3'. (b) 2'-Deoxyinosine nucleotide in duplexes 5' dGGCT**I**GATCACCAG 3' / 5' dCTGGTGATC**Y**AGCC 3'. The solutions contained a final duplex concentration of 3.7 μM in a pH 7.0 buffer consisting of 90 mM NaCl, 10 mM sodium phosphate, 1 mM EDTA and 2.5 mM DTT.







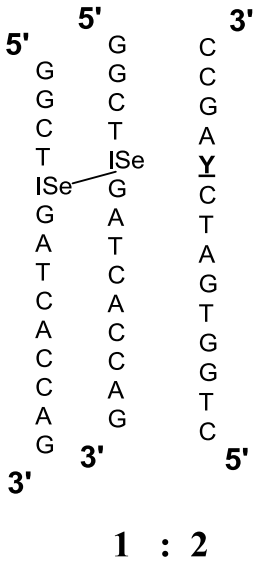


Figure A3.20: T_m curves of **dI-Se-ODN$_2$** with different complements. The solutions contained a final oligonucleotide dimer concentration of 1.0 μM in a pH 7.0 buffer consisting of 90 mM NaCl, pH 7.0, 10 mM sodium phosphate and 1 mM EDTA. Data was acquired at 0.5 °C/min from 15°C to 95°C, monitoring UV absorption at 260 nm. For clarity, the curves were separated into two graphs.

Figure A3.21: CD spectra of **dI-Se-ODN$_2$** with different complements. The solutions contained a final oligonucleotide dimer concentration of 1.0 μM in a pH 7.0 buffer consisting of 90 mM NaCl, 10 mM sodium phosphate and 1 mM EDTA.

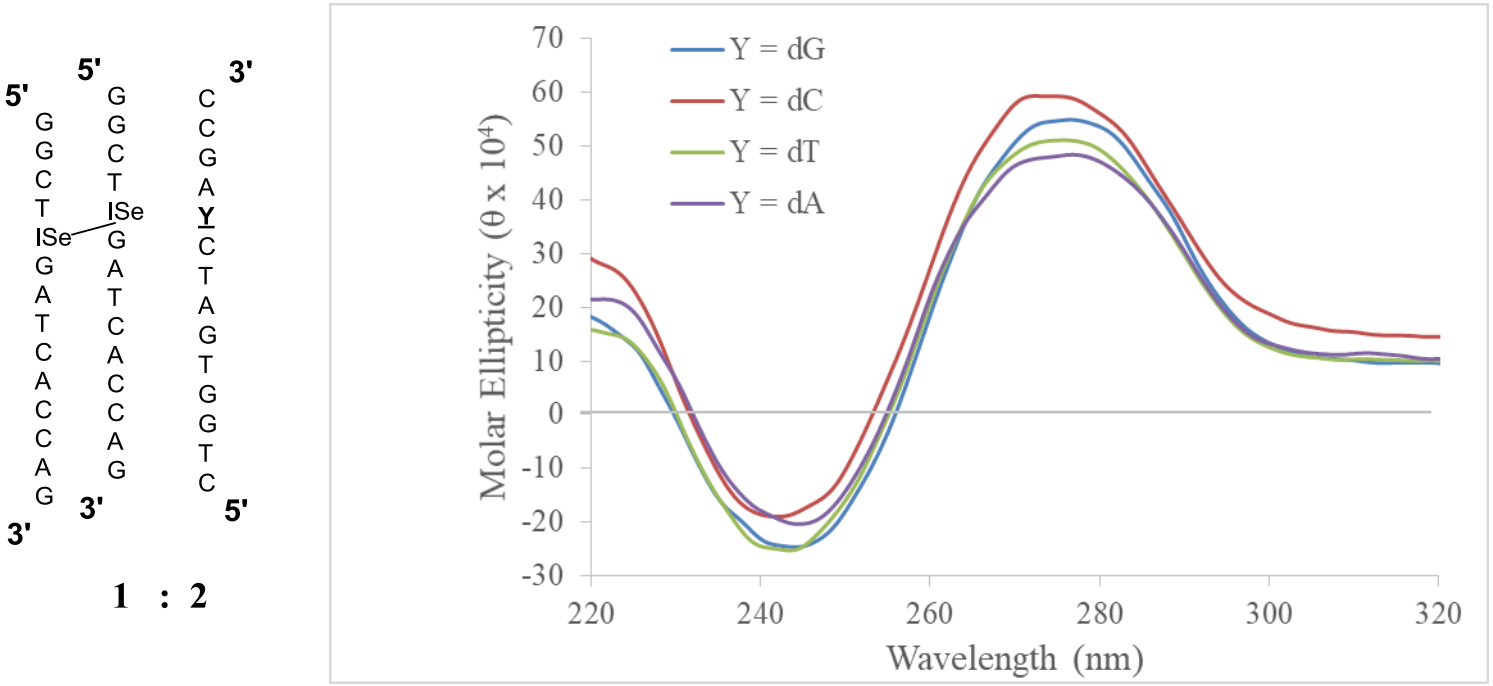
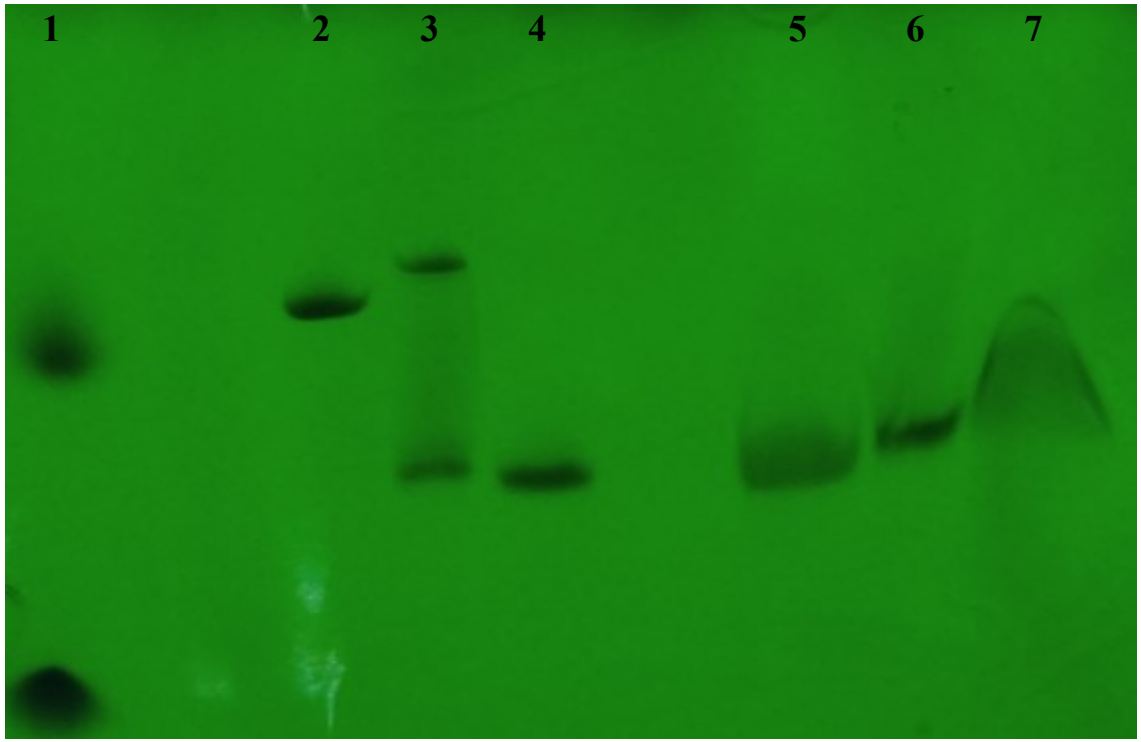
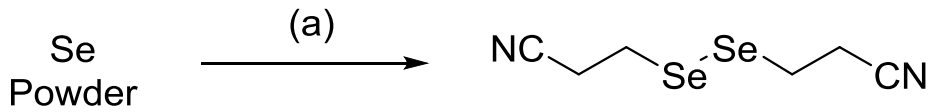


Figure A3.22: Denaturing PAGE of **dI-Se-ODN$_2$** treated with various reducing agents. Lane 1: dye markers (upper band: xylene cyanol FF; lower band: bromophenol blue); lane 2: 28-mer ssDNA (5' d(AT)$_{14}$ 3'); lane 3: **dI-Se-ODN$_2$**; lane 4: **dI-Se-ODN**; lane 5: **dI-Se-ODN$_2$** treated with 10 mM DTT; lane 6: **dI-Se-ODN$_2$** treated with 10 mM $NaBH_4$; lane 7: **dI-Se-ODN$_2$** treated with 10 mM GSH.



Scheme A3.1: Synthesis of diselenide **3.3**. Reagents and conditions: (a) (i) $NaBH_4$, 1,4-dioxane/EtOH (4:1, v/v), 1 h, rt (ii) 3-bromopropionitrile, 1 h, rt.



3.3 44%