UNIVERSITY OF BIRMINGHAM

THE OPTIMISATION OF TRANSFER CHEMISTRIES FOR DNA-TEMPLATED SYNTHESIS (DTS)

by

BETHANY CROW

A thesis submitted to
The University of Birmingham
for a degree of
DOCTOR OF PHILOSOPHY

School of Chemistry
College of Engineering and Physical Sciences
University of Birmingham
May 2025

Abstract

Inspired by the ribosome, researchers have developed strategies to synthesise programmable, sequence-specific polymers for material discovery. One method aiming to achieve this is DNA-templated synthesis (DTS). DTS co-localises reactive tags attached to complementary DNA strands, initiating a chemical reaction. Sequentially, spent adapters are replaced, through DNA's complementary base binding, to produce identifiable sequence-specific polymers. However, currently, the most successful chemistries used for monomer transfer are prone to degradation, limiting the reaction yield and polymer growth. Therefore, this thesis explores the optimisation of DTS transfer chemistries to limit the hydrolysis of reactive tags.

Chapter 1 provides a simplified introduction to ribosomal peptide synthesis and examples of how researchers have been inspired to develop similar strategies for artificial polymer synthesis, *i.e.* DTS. The chemistries and architectures that have demonstrated successful transfer of DTS monomers, as well as their limitations, are discussed. Susceptible to hydrolysis in aqueous conditions, Chapter 2 investigates the cause for thioester stabilisation in a DTS across-the-helix aminolysis mechanism, in order to harness its full capabilities. In addition, the effects nucleobases have on the architecture's DTS yields and thioester protection are explored. In order to improve reaction yields and rates, Chapter 3 introduces a new DTS electrophile, the selenoester, investigating its capabilities in a single-step DTS assay. Chapter 4 attempts to optimise Chapter 3′s selenoester-mediated DTS further, as well as incorporate the electrophile into a multistep DTS system.

For Jack, to our future.

Acknowledgements

Firstly, I would like to thank my supervisor, Prof. Rachel O'Reilly, for providing me with the opportunity to work on such an incredible project. Having always been interested in biochemistry, I'm grateful for your support to pursue a new research field. I would like to thank our University of Oxford collaborators, Prof. Andrew Turberfield and Dr. Jonathan Bath, for their expertise and guidance in DNA nanotechnology. I would like to thank EPSRC for their funding of the Artificial Ribosome project and the University of Birmingham for their funding of my research position.

A massive thank you to Ms Tessa Kintail, for her incredible kindness and generosity throughout my PhD. Thank you to the past and present O'Reilly and Dove group members for their continuous support. In particular, a special thank you to Dr. Stephen Fielden, Dr. Julia Rho, Dr. Josh Tibbetts, Dr. Sam Parkinson, Dr. Natalia Reis and Dr. Edward Maynard for their help within the lab and to my friends Alisha, David and Parvaneh. Notably, I cannot thank Dr. Christopher Williams enough for his expertise and mentoring in mass spectrometry. Jess and Jennifer, my DNA queens. Thank you to the moon and back for your love, advice and guidance in and outside the labs. Even at your busiest, you still gave me your time – I am forever in debt to you. You are both sensational scientists and friends.

Outside of the laboratory, I would like to thank Libby, Lottie and Eimear for welcoming me with open arms into Birmingham. To all of the instructors at Barefoot Yoga, thank you for teaching me how to remain grounded when times became difficult. Emma and Orla, thank you for some beautiful holidays, memories and for letting me stay over during my final weeks of lab work. Grandma and Grandad, I'm grateful for

your emotional support, reminding me to take breaks and to look after myself. Steve and Elsa, you are the best in-laws I could ever ask for, supporting me as if I were your third daughter. I wouldn't have completed this thesis without your love and encouragement.

To my best friends in the whole world, Josh, Paulina, Leena and Louisa, you made my time in Birmingham incredible. I love you all so much – your jokes, your hugs, your energy and the toothbrush parties! Josh, Leena and Louisa, you are, and will be forever, the best housemates in the world. I miss having you around all the time - even Josh's morning sighs. Especially to Louisa, without you, I wouldn't have got to this position. You let me be my unapologetic self; always listening to my random bs on our morning walks and tolerating my chaotic behaviour. The cackling, the gibberish, the bickering and the tears – I wouldn't change a second of it.

Jack, if I listed everything you have done for me during my PhD, it would be longer than this thesis. Thank you for your love, care, patience and everything in between. I'm grateful to have had you by my side for the past 11 and a half years – I love you. As this work comes to an end, we can finally start our life together, just the three of us (Vinnie included)!

Contents

Declaration of Authorship

This thesis is submitted to The University of Birmingham in support of my application for the degree of Doctor of Philosophy. The thesis has been composed by myself, with the work presented being carried out by myself between October 2021 to April 2025. The work in this thesis has not been submitted for any other degree or professional qualification.

The work presented herein has been conducted by me, except the following:

- Dr. Jennifer Frommer designed both “DNA-Templated Synthesis (DTS) assay” and “protection assay” described in Chapter 2, Chapter 3 and Chapter 4., which were published in J. Frommer, R. Oppenheimer, B. M. Allott, S. Núñez-Pertíñez, T. R. Wilks, L. R. Cox, J. Bath, R. K. O'Reilly, A. J. Turberfield, *Angew. Chem. Int. Ed.* 2024, **63**, e202317482.
- Dr. Jennifer Frommer synthesised the TAMRA amide product **2-4** used for the calibration of **2-4** in Chapter 2, Chapter 3 and Chapter 4.
- Deconvolution of DNA LC-MS spectra, using ProMass HR software, was completed with assistance from Dr Christopher Williams, Analytical Facility Staff, University of Birmingham.
- **Figure 2-4B** and **2-4C** were adapted, with permission, from A. K. Cunningham’s MSc dissertation “New Chemistries for DNA Templated Chemistries” from the University of Birmingham
- **Figure 3-1** was adapted, with permission, from J. Sayers, R. J. Payne and N. Winssinger, *Chem. Sci.*, 2018, **9**, 896.

List of Figures

List of Schemes

Chapter 1: Introduction

List of Tables

Chapter 3: Incorporation of selenoesters into a DTS-based system

Chapter 4: Optimisation of selenoester-mediated DTS towards a multistep synthesis strategy

List of Supplementary Figures

Chapter 2: Understanding the complexities of the TAMRA-thioester DTS across-the-helix architecture using absorbance and fluorescence spectroscopy

Chapter 4: Optimisation of selenoester-mediated DTS towards a multistep synthesis strategy

Abbreviations

°C	Degrees Celsius
A	Adenine
Abs	Absorbance
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
Boc	*Tert*-butyloxycarbonyl
Boc-Lys	Boc-lysine
BSOCOES	1,1′-[sulfonylbis(2,1-ethanediyloxycarbonyloxy)]di(2,5-pyrrolidinedione)
C	Cytosine
CAPS	3-(Cyclohexylamino)-1-propanesulfonic acid
Cys	Cysteine
d	Doublet splitting
DEAE	Diethylaminoethyl
DECLs	DNA-encoded libraries
DIPEA	*N,N*-diisopropylethylamine
DMF	*N,N*-dimethylformamide
DMSO	Dimethyl sulfoxide
DMT	Dimethoxytrityl
DMTMM	4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
DSL	Diselenide ligation
DTS	DNA-templated synthesis
DTT	Dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EF-Tu	Elongation factor Tu
eq.	Equivalents
ES	Electrospray
EtOAc	Ethyl acetate
EtOH	Ethanol

Fmoc	Fluorenylmethyloxycarbony
G	Guanine
g	Grams
Gly	Glycine
GTP	Guanosine triphosphate
h	Hours
HATU	Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium
HCR	Hairpin chain reaction
HPLC	High performance liquid chromatography
HRMS	High-Resolution Mass Spectrometry
HTS	High-throughput screening
IPA	Isopropyl alcohol
IR	Infrared
J	Coupling constant
kDa	Kilodaltons
L	Litre
LC-MS	Liquid chromatography-mass spectrometry
Lys	Lysine
m	Multiplet splitting
M	Molar
MeCN	Acetonitrile
MeOH	Methanol
min	Minutes
MMT	Monomethoxytrityl
mol	Moles
mRNA	Messenger RNA
ncAAs	Non-canonical amino acids
NCL	Native chemical ligation
NHS	*N*-hydroxysuccinimide
nm	Nanometre
NMR	Nuclear magnetic resonance
NVOC	*O*-nitroveratryloxycarbonyl
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PG	Protecting group
Phe	Phenylalanine

ppm	Parts per million
Pyl	Pyrrolysine
q	Quartet splitting
R_f	Retention factor
RNA	Ribonucleic acid
RP-HPLC	Reverse phase high-performance liquid chromatography
Rpm	Revolutions per minute
s	Singlet splitting
Sec	Selenocystine
Ser	Serine
S_NAr	Nucleophilic aromatic substitution
SPDP	*N*-succinimidyl 3-(2-pyridyldithio)propionate
SsDNA	Single-stranded DNA
t	Triplet splitting
T	Thymine
TAE	Tris-acetate-EDTA
TAMRA	6-Carboxytetramethylrhodamine
TAMRA-NHS	6-Carboxytetramethylrhodamine succinimidyl ester
TAPS	*N*-Tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid
TCEP	Tris(2-carboxyethyl)phosphine
TEAA	Triethylammonium acetate
THF	Tetrahydrofuran
TLC	Thin-layer chromatography
TRITC	Tetramethylrhodamine-5-isothiocyanate
Trna	Transfer RNA
U	Uracil
UV	Ultraviolet
UV-Vis	Ultraviolet-Visible
Val	Valine
Δ	Chemical shift
Λ	Wavelength
λ_{Em}	Emission wavelength
λ_{Ex}	Excitation wavelength
λ_{max}	Maximum wavelength

Chapter 1

Introduction

1.1 The Ribosome - nature's molecular machine

1.1.1 Protein synthesis

Molecular machines are small structures capable of performing tasks at a molecular scale, such as the synthesis or transportation of compounds.[1] Utilising nucleic acid chemistries, nature evolved the ribosome, a molecular machine that translates the genetic information to facilitate protein synthesis.[2, 3] Over billions of years, the ribosome gradually assembled a complex structure through improvements in efficiency and accuracy.[4, 5] In combination with other biological functions, this led to a large portfolio of proteins, each essential for the functions and regulations of sustaining life.[2] Proteins are large peptidyl molecules (macromolecules) consisting of a sequence of amino acid monomers, synthesised within the ribosome (Figure 1-1).[2]

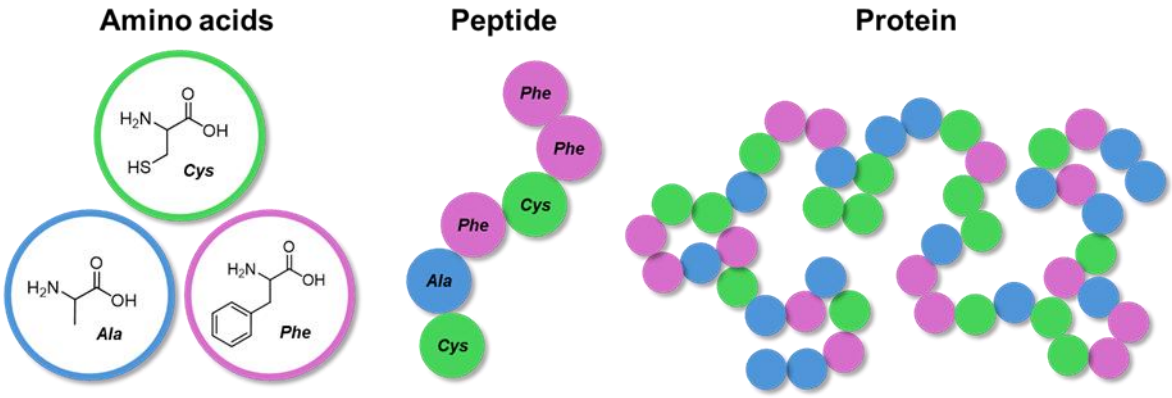


Figure 1-1: *Simplified model of the composition of proteins – constructed of peptides, which are sequences of amino acid monomers. Example monomers = alanine (Ala), cysteine (Cys) and phenylalanine (Phe).*

Proteins are formed through long chains of peptides; amino acids connected between their primary amine and carbonyl functional groups, *i.e.* biological polymers (Figure 1-2B). Proteins can take on multiple shapes (secondary, tertiary and quaternary structures) depending on the amino acid's side chain (Figure 1-2C).[2]

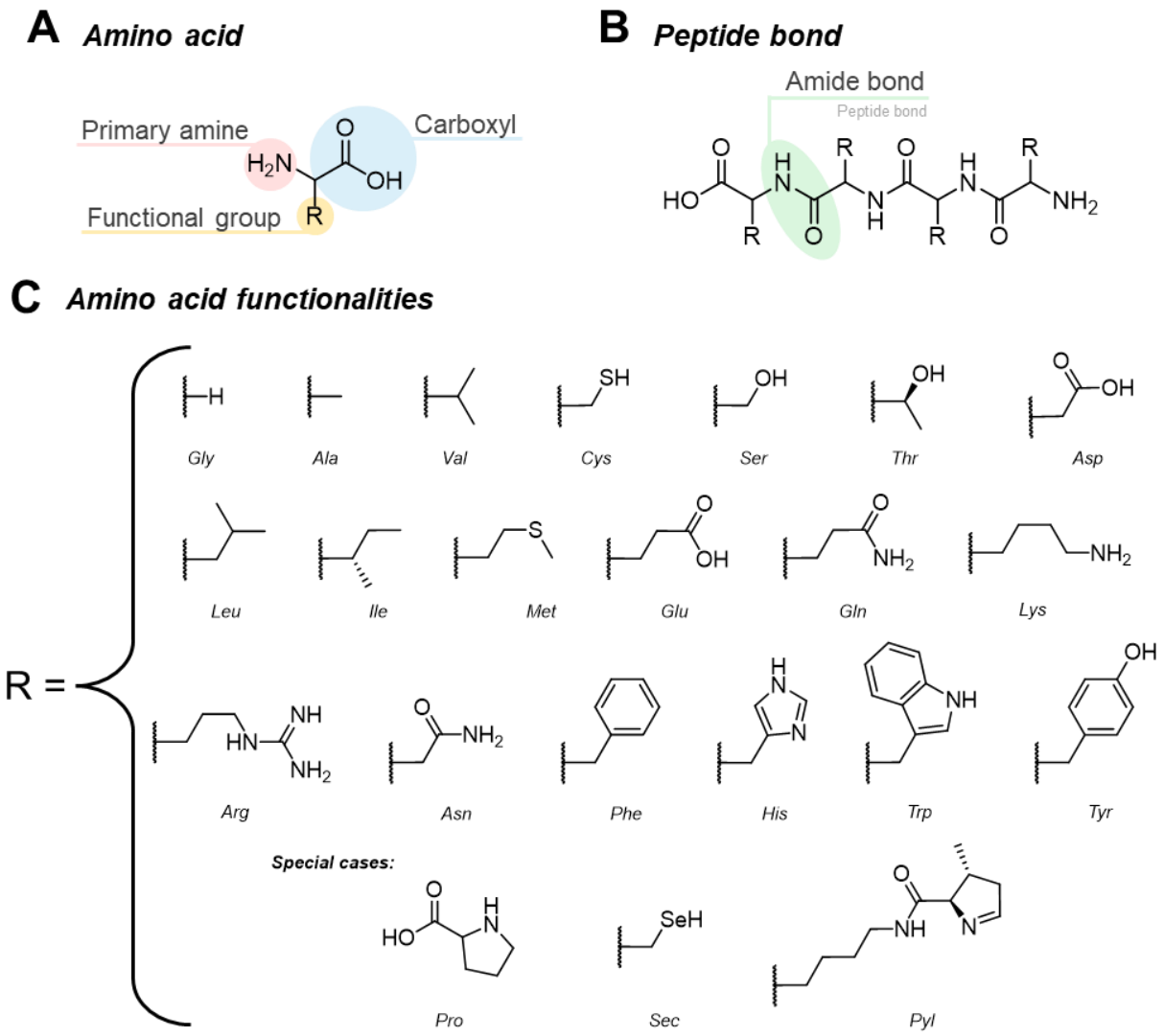


***Figure 1-2**: **A**) The general structure of an amino acid. **B**) Structure of a peptide/amide bond. **C**) Functionalities of the 22 proteinogenic amino acids. Special cases include proline (Pro) – cyclic functionality with the aminoacyl amine; selenocysteine (Sec) and pyrrolysine (Ply) – synthesised using STOP codons.*[6, 7]

With there being 22 proteinogenic amino acids, proteins exhibit considerable versatility in their conformations and functionalities.[8] Through mutations in genetic

sequence, the ribosome indirectly aids in the evolution of proteins, enabling the production of a vast combinatorial library of proteins essential for maintaining biological functions in living organisms.[9, 10] In doing so, it relies on nucleic acids, deoxyribonucleic acid (DNA, the genetic code) and ribonucleic acid (RNA), to code the required amino acids within the protein sequence.

1.1.2 DNA and RNA structure

DNA and RNA are the polynucleotide chains essential for the development and sustainment of life.[10] A phosphate group, a central five-membered ring sugar (deoxyribose or ribose) and a nucleobase construct the nucleotide monomers (Figure 1-3A & B). For DNA, four structures of nucleobases are known: adenine and guanine (purines), and thymine and cytosine (pyrimidines), creating a unique sequence of coded genetic information (Figure 1-3C).[2] As opposed to RNA, which replaces thymine with uracil, and bears an additional hydroxy group at the C2′ of the ribose sugar. Based upon the hydroxyl group positioning, the strands run in an antiparallel manner to produce 3′ and 5′ directionality.

Nucleobases can form base pairs with complementary nucleobases through hydrogen bonding on an adjacent nucleic acid strand (Figure 1-3D). Complementary base-pairing is predominantly observed between guanine with cytosine (G-C), and adenine with thymine (A-T) in DNA or with uracil (A-U) in RNA. Base pairing, combined with π-stacking with adjacent nucleobases, contributes to a helical structure. The non-covalent interactions between the nucleobases within the helix create a hydrophobic core, excluding water from the interior. In the more typical right-handed B-DNA helix

structure, the strands coil to form major and minor grooves (Figure 1-3D). Along the grooves, surrounding molecules (such as proteins) can bind to DNA.[11]

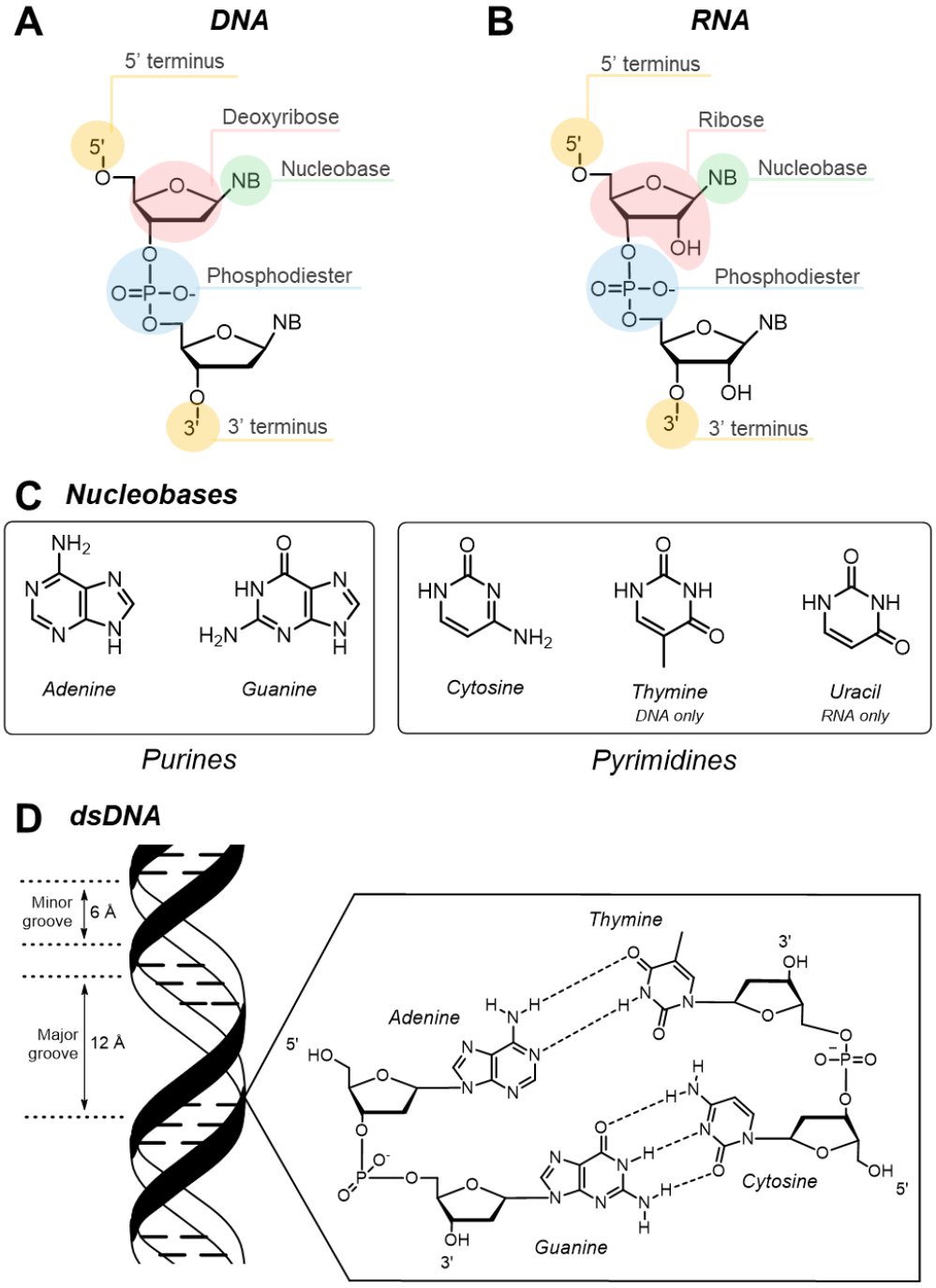


***Figure 1-3**: **A**) Structure of DNA. **B**) Structure of RNA. **C**) Structure of DNA and RNA nucleobases. Purines: adenine and guanine. Pyrimidines: cytosine, thymine (DNA only) and uracil (RNA only). **D**) Structure of hybridised double-stranded DNA (dsDNA). Hydrogen bonds are labelled between complementary base pairs: adenine-thymine and guanine-cytosine, connected to the sugar-phosphate backbone. Distances of the average major and minor grooves of DNA are annotated.*

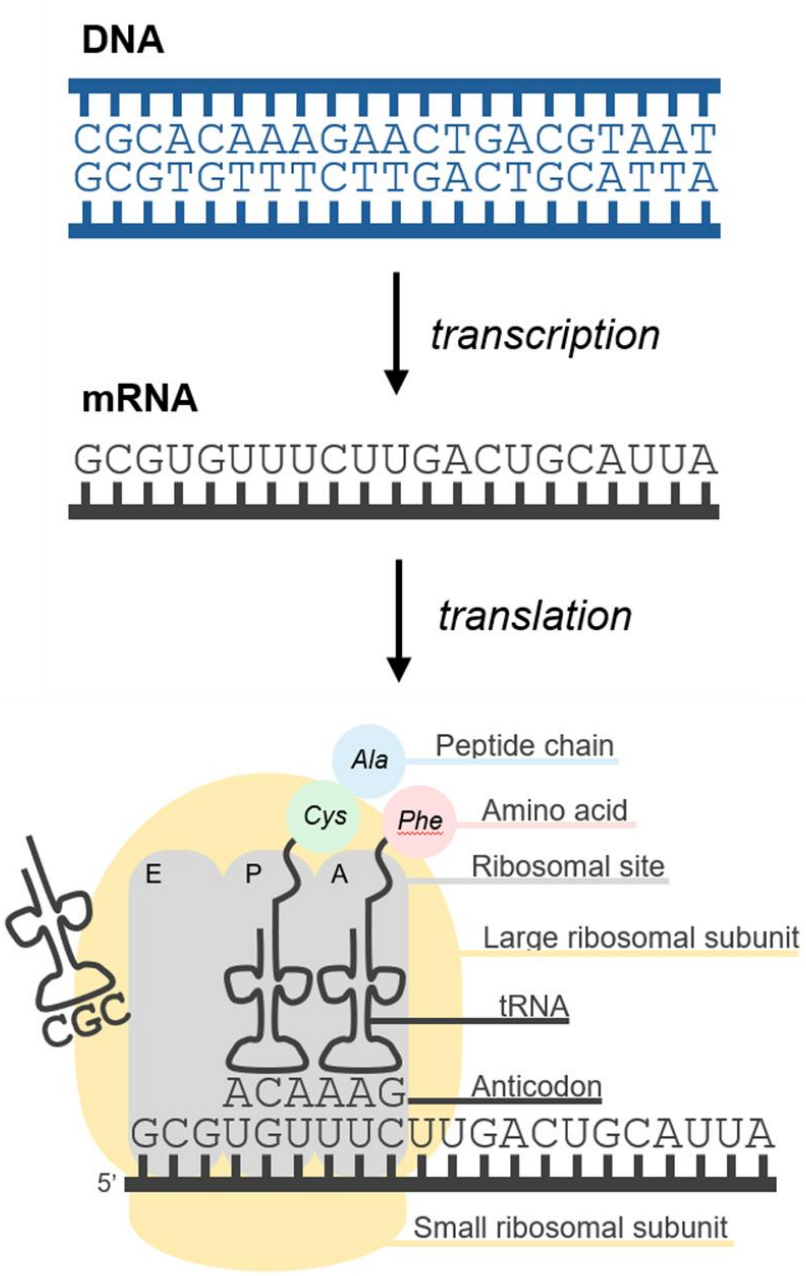
1.1.3 Ribosomal synthesis of peptides

The mechanisms involved in ribosome-catalysed peptide bond formation are extremely complex;[12-15] however, a simplified model will be described to demonstrate how it operates as a molecular machine.[16]

Proteins are coded, using adenine, thymine, guanine and cytosine, within the nucleobase sequence of DNA (Scheme 1-1). A biological reader, messenger RNA (mRNA), is the transcribed RNA version of the DNA template. mRNA contains three nucleobase segments, called codons. Each codon codes for the implementation of a specific amino acid. The mRNA strand is then decoded within the ribosome using transfer RNA (tRNA). tRNA is a single strand of RNA containing three complementary nucleobases to the mRNA's codon (an anticodon), and a covalently attached amino acid (Scheme 1-1). Therefore, the two subunits of the ribosome assemble around the mRNA, while the tRNA binds its specific anticodon to the codon of the mRNA sequence.[16] Here, protein synthesis commences at site A of the larger ribosomal subunit. Next, within the P-site, amino acids conjugated to their respective tRNA are brought into proximity, initiating peptide bond formation.[13, 16] Once used, the tRNA is ejected from the ribosome at the E-site.[17] Sequentially, aminoacyl tRNAs are introduced as substrates and increase the length of the peptide chain.

The specific sequence of a protein is what provides its defined function; therefore, maintaining the precise amino acid sequence is crucial.[9, 10] As a consequence of the mRNA codon-tRNA anticodon complementary binding, there are limited mistakes in protein synthesis, reducing the error frequency.[18, 19] Furthermore,

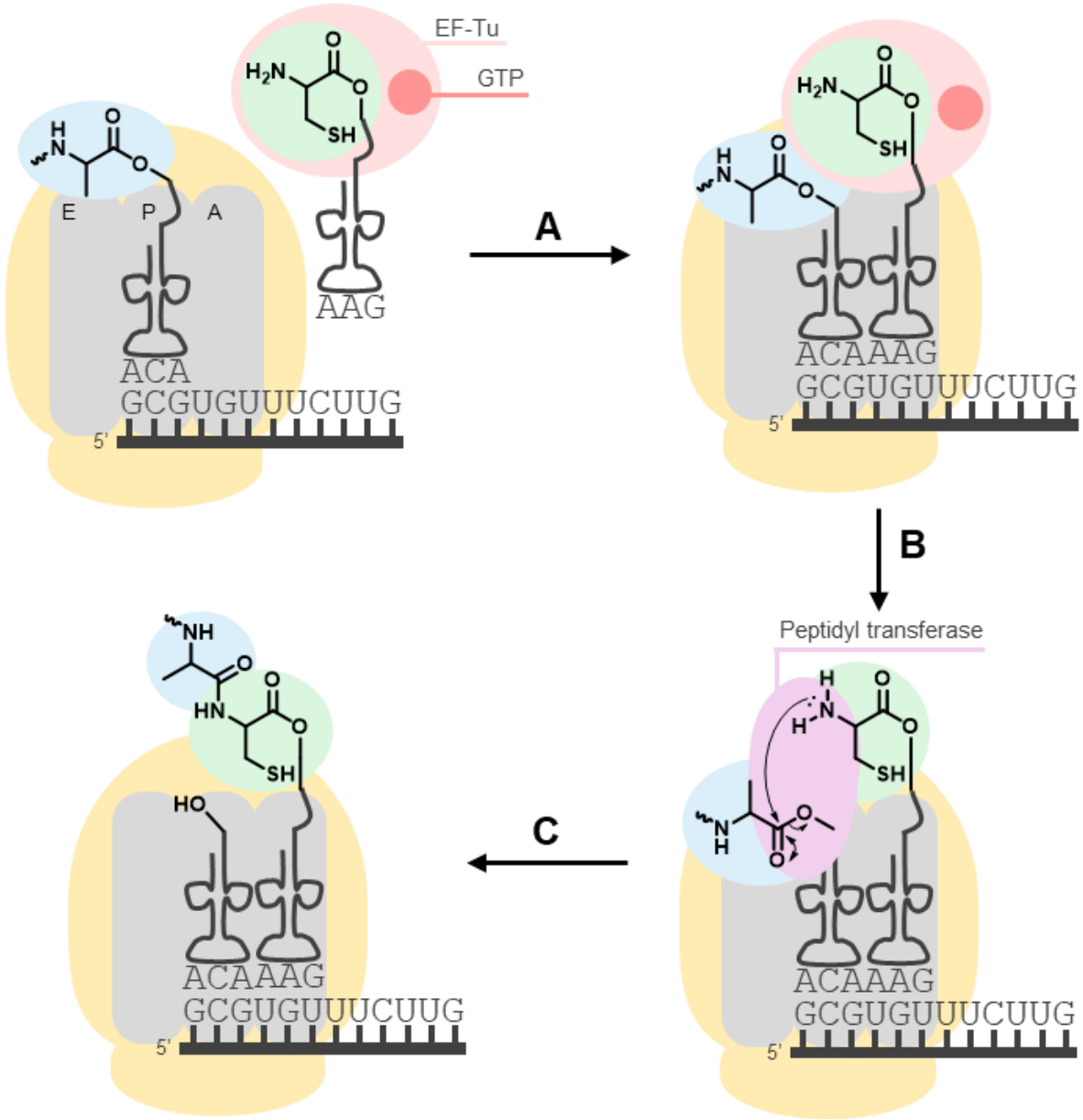
the addition of the following amino acid would not be completed until the previous had been completed, preventing peptide truncation further.[20]



Scheme 1-1: *Simplified schematic of components involved in peptide synthesis within the ribosome. Transcription of the DNA genetic code into mRNA, followed by the translation of mRNA by tRNA for peptide synthesis. Ribosomal sites, A, P and E are labelled in grey.*

Peptide bond formation, catalysed by the ribosome and peptidyl transferase, involves a primary amine and an aminoacyl-oxoester (Scheme 1-2).[2] Using the activated carboxylate of amino acids after entering the A site, the nucleophilic amino group, of the amino acid-tRNA complex located in the P site, reacts with the ester bond of tRNA to synthesise peptide bonds (Scheme 1-2C). Typically, esters are prone to degradation within aqueous conditions due to the nucleophilic nature of water, with their stability varying depending upon the conjugated amino acid functional group.[21] However, the ribosome utilises an enzyme, elongation factor Tu (EF-Tu), coupled with guanosine triphosphate (GTP), to stabilise the aminoacyl-tRNA ester during transportation to the A-site (Scheme 1-2A).[21, 22] Importantly, if the aminoacyl-tRNA were to hydrolyse, it would not participate in peptide synthesis, and an unhydrolyzed aminoacyl-tRNA would be introduced.[23] Another optimised feature of the ribosome is its proximity chemistries. Within peptidyl transferase, the primary amine and the oxoester are co-localised, increasing the effective local concentration and overcoming conformational challenges.[21] In addition, this further prevents hydrolysis of the *in-situ* ribose oxoester.

Overall, the ribosome aids in the synthesis of sequence-specific proteins according to the encoded instructions of messenger RNA (mRNA). It is a molecular machine which demonstrates high levels of precision in peptide synthesis. Through its reaction-dependent, hydrolytically stable, coded mechanism, there is a low error frequency in the proteins synthesised. However, mutations can occur within the genetic code (DNA) sequence. Through transcription and translation, these variations are incorporated into the protein sequence and, over time, aid in the evolution of protein functions.



***Scheme 1-2**: Simplified schematic of peptide synthesis within the ribosome. **A**) EF-Tu and (light pink) GTP (pink) stabilise the oxoester bond present on tRNA and transport it to the A-site of the large ribosomal subunit. **B**) EF-Tu and GTP leave the ribosome and are recycled. **C**) nucleophilic attack of aminoacyl tRNA at the A site with oxoester at the P site, catalysed by peptidyl transferase (purple) to produce a peptide bond.*

Because of its high efficiency, the ribosome has contributed to the controlled evolution of an array of biological functions, *e.g*. enhanced selectivity, stability and affinity of enzymes/antibodies.[8, 24] However, the ribosome is restricted to the 22 proteinogenic amino acids as molecular building blocks. If new, alternative chemistries

could be incorporated into a similar sequence-defined molecular machine, an abundance of therapeutic materials would be accessible. Enhancing the functional diversity would enable novel properties and improve bioavailability and ease of optimisation in peptides. Or, alternatively, introduce controlled synthetic polymerisation at smaller, cost-effective, scales. Therefore, researchers have been inspired to create artificial molecular machines to facilitate material synthesis[25-29] or have chosen to adapt preexisting ribosome syntheses to expand the scope of synthesisable proteins.[30-34]

1.2 Alternative molecular machines

1.2.1 Ribosomal engineering

Multiple strategies have been developed to increase the breadth of accessible chemistries in ribosomal peptide synthesis.[30-34] One example substitutes the amino acids connected to tRNA with non-canonical amino acids (ncAAs) in *E. coli*.[35] In order to introduce the ncAAs into the peptide sequence, mRNA and tRNA are engineered to recognise modified nucleobases[31, 36-38] or suppress STOP codons (which halt peptide synthesis)[39-43] while expanding the genetic code. Using the STOP codon strategy, the number of nucleobases within codons/anti-codons can also be increased from three bases to either four or six.[44-46]

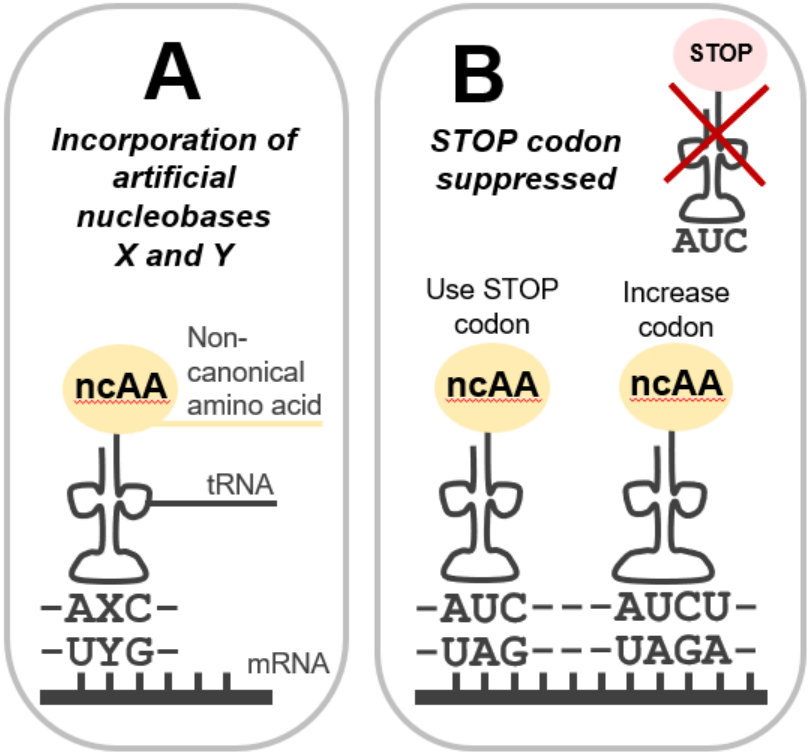


***Figure 1-4**: Incorporation of ncAA into peptide sequence by **A**) introducing artificial nucleobases (X and Y) into the genetic code or **B**) suppressing STOP codons (e.g. UAG) to allow use or to increase the codon size from three to four nucleobases.*

RNA is prone to degradation, which increases the complexity of RNA modification. The 2′-hydroxy group on the ribose sugar can cleave phosphodiester bonds in RNA's phosphate backbone under high pH conditions, making synthetic modifications laborious.[47] Therefore, applying traditional chemical synthetic strategies would be difficult and time-consuming. In addition, the introduction of ncAAs would require the modification of aminoacyl-tRNA synthetase.[48] As the protein which attaches the amino acid to the corresponding tRNA, it would need to recognise the ncAAs as a substrate. An alternative approach to the creation of a molecular machine would be the employment of a more stable DNA scaffold with modified chemical tags attached.[49] Using the complementary base binding of DNA, scientists can manipulate and program complex DNA architectures.[50, 51] Because of the unique structures DNA can make, it

has contributed to controlled polymer synthesis,[52-54] the production of chemical libraries,[28, 55, 56] material discovery[25, 57 58] and targeted drug delivery.[59-61]

1.2.2 DNA-encoded libraries

Inspired by the ribosome, research has commenced in engineering artificial DNA-based molecular machines for controlled syntheses.[62-64] One approach aims to apply molecular machines into high-throughput screening (HTS) libraries for drug discovery, diagnostics and therapeutics – DNA-encoded libraries (DECLs).[28, 65-69] Typical HTS assays require an excess of consumables and reagents in order to screen for target molecules (Figure 1-5). Whereas DECLs enable the large synthetic toolbox of traditional chemistries to be synthesised at pmol scales of DNA.[28, 70] DECLs operate at larger magnitudes than typical HTS assays. By working on the smaller scale of DNA and in a single pot, the cost of materials required is reduced, and the chemical space can be increased.[71]

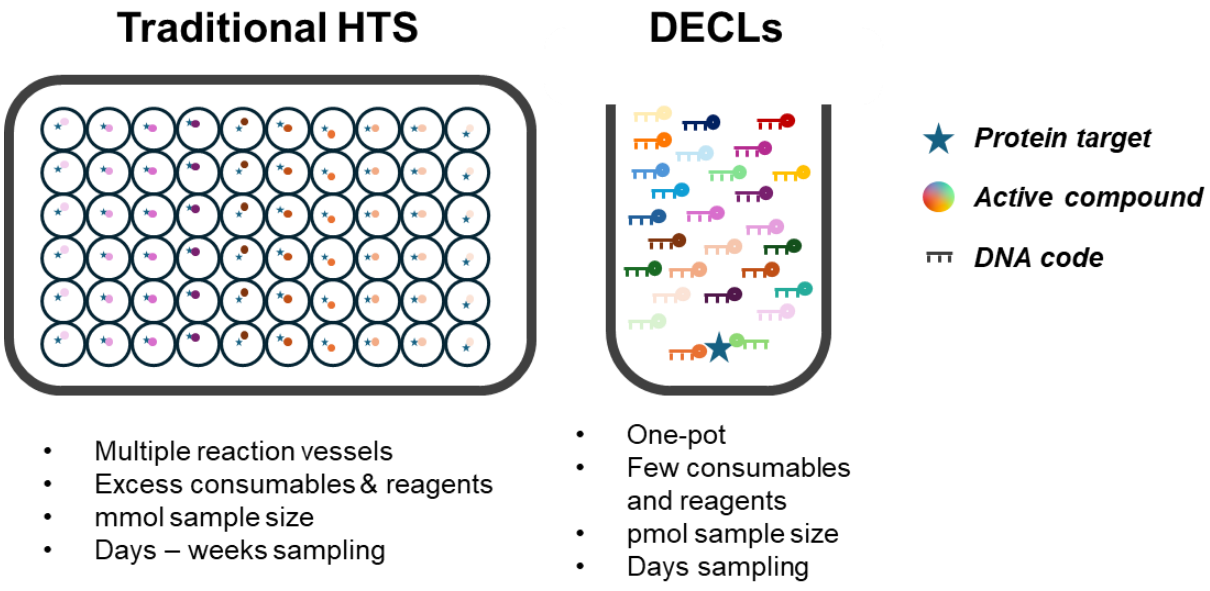
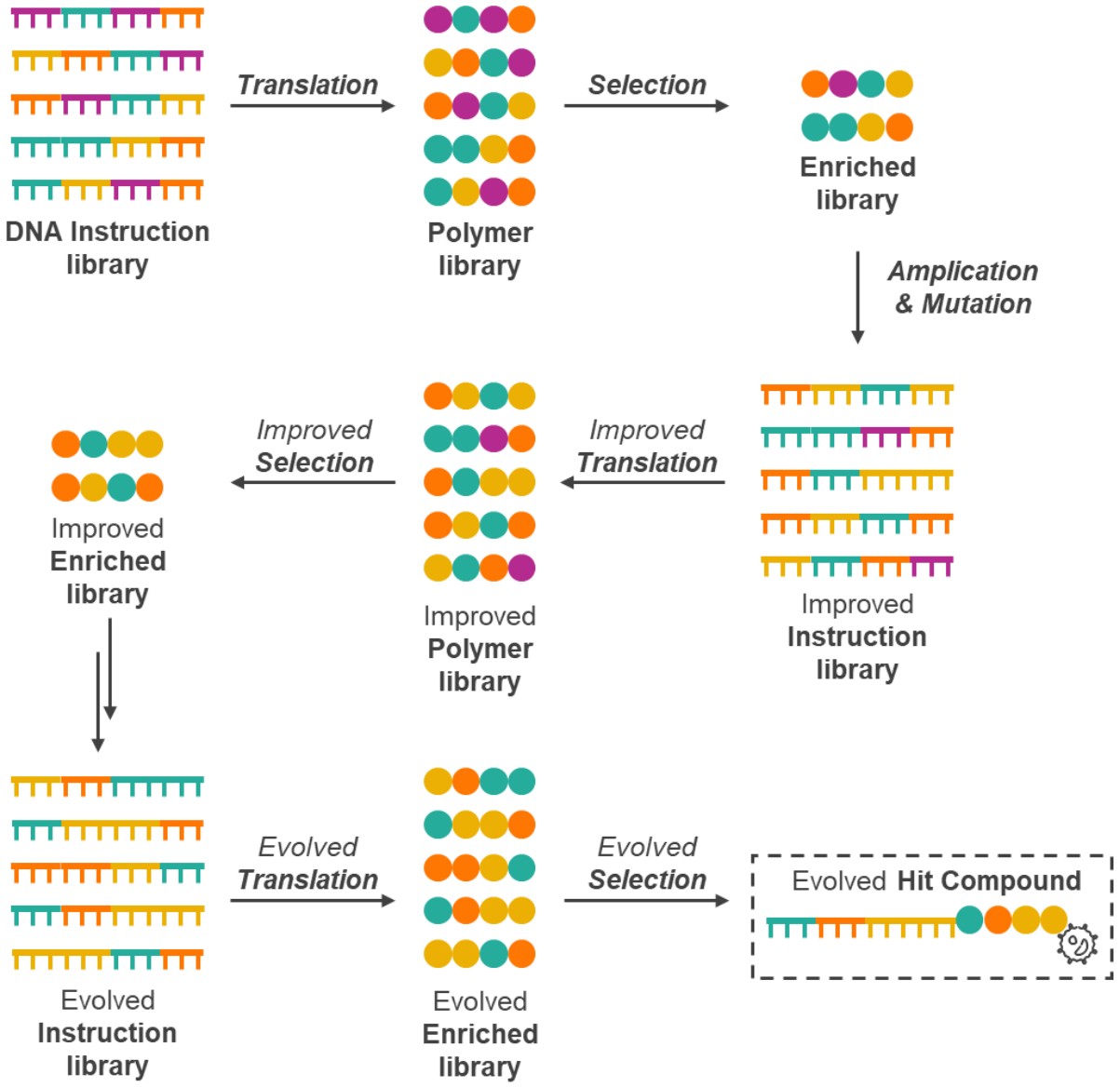


Figure 1-5: *Comparison between traditional HTS assays and DECL assays. HTS require multiple reaction vessels in comparison to DECLs, which can be performed in a single pot with the protein target.*[28]

As with mRNA for peptide sequences, DECLs use predetermined coded regions in the DNA sequence that correspond to specific chemistries of an attached compound. Initial candidates are attached or synthesised onto DNA and screened against a biological target (*i.e.* protein) within a single reactive vessel. (Scheme 1-3). Unbound/unreactive compounds are removed from the solution, leaving only the successful candidates.[70] The DNA "barcode" of successful compounds can then be amplified using a polymerase chain reaction to identify the structure of the successful molecule/polymer. This translation process aids in the diversification and evolution of target compounds. More sophisticated DECLs allow for the attachment chemistries to drive mutations in the DNA sequences, aiding in the molecular evolution over multiple cycles. Through this evolutionary process, not every chemistry needs to be sampled, allowing for quicker identification pathways. In order to synthesise the tagged DNA for DECLs, an array of DNA-bound chemical strategies have been constructed.[29, 55, 67, 72-75]



Scheme 1-3: *Generalised schematic of molecular evolution using DECLs. DNA strands are translated to produce a library of sequence-specific polymer products. The synthesised polymers are then selected based on the desired properties. Next, the DNA instructions of the enriched products are amplified and mutated to create an improved instruction library. The sequence is repeated until the optimum, evolved, hit molecule/polymer is synthesised. Figure adapted from O'Reilly et al.*[55]

1.3 Artificial DNA modifications

For DECLs to operate, chemical tags must be covalently attached to DNA. In order to achieve this, modification sites are incorporated during solid-phase DNA

synthesis (phosphoramidite chemistries) or post-synthesis.[76, 77] In the 3′ to 5′ direction, nucleotides are coupled together through the 5′ hydroxy group of one and the 3′ phosphoramidite of another (Figure 1-6C). Phosphoramidite chemistries utilise functional group-protected nucleotides to prevent unwanted side reactions during synthesis.[76, 78] The most common nucleotide protecting groups for chemically modified DNA include dimethoxytrityl (DMT) at the 5′ hydroxy site, 2-cyanoethyl at the phosphoramidite, and several nucleobase protection groups. (Figure 1-6A).[76, 78]

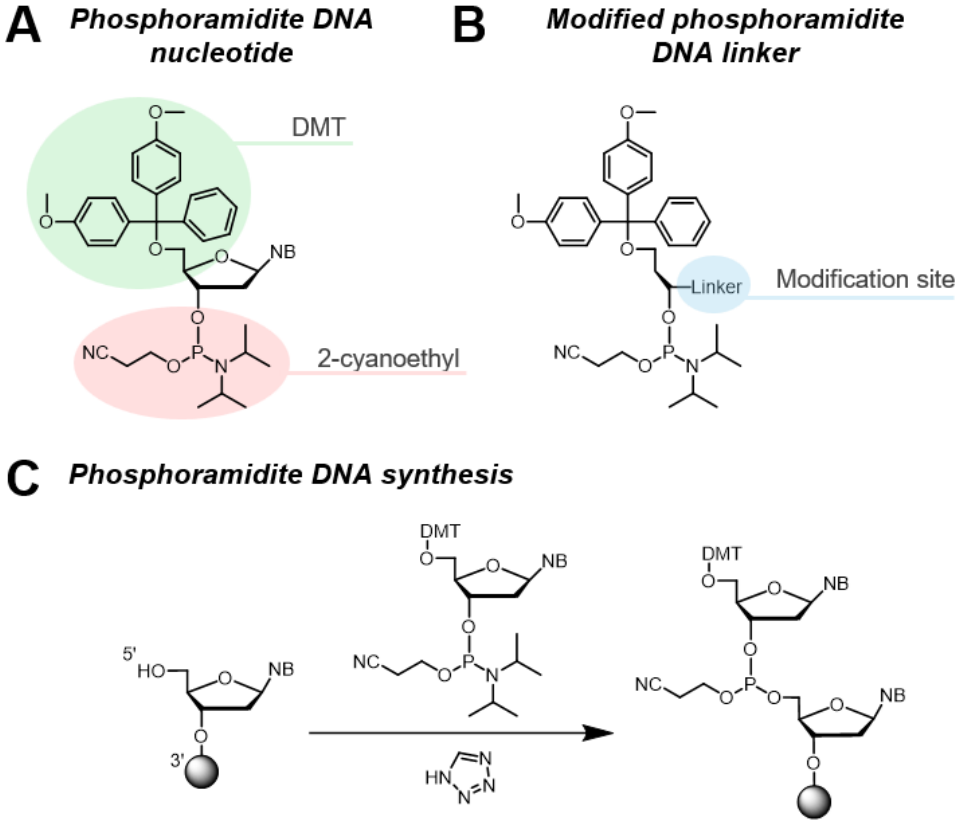


***Figure 1-6**: Phosphoramidite DNA synthesis. **A**) Common phosphoramidite DNA nucleotide with DMT and 2-cyanoethyl protecting groups. **B**) Modified phosphoramidite DNA nucleotide replacing the deoxyribose with a connecting linker (site for modification). **C**) Reaction of the 5' hydroxy group with an incoming phosphoramidite during solid-phase DNA synthesis.*

In the case of modifications, specific functionalities can be added, including deoxyribose-free linkers at the desired location (Figure 1-6B). Examples of common

DNA modifications include nucleobase-free spacers, fluorophores and linkers for biochemical analysis (Figure 1-7).

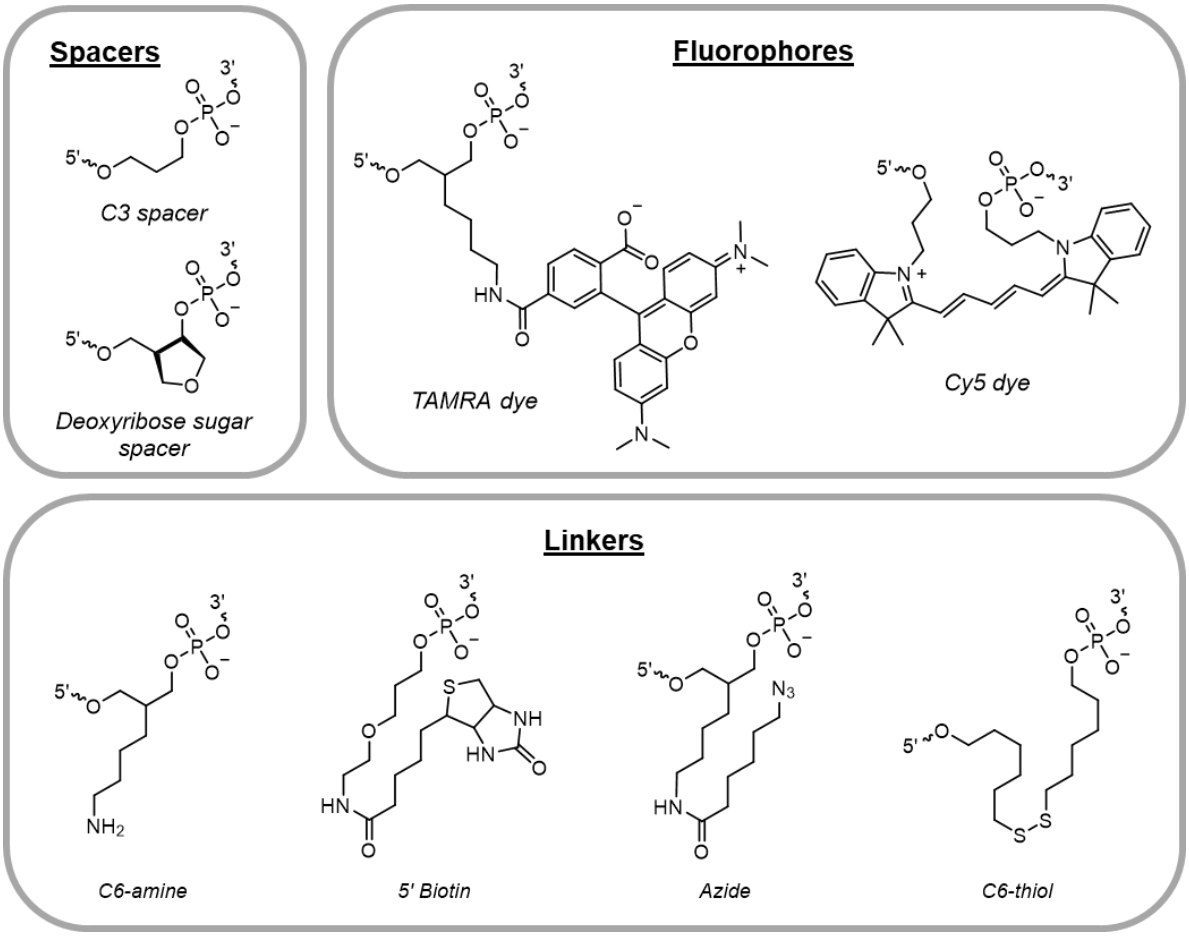


Figure 1-7: *Common synthetic DNA modifications: spacers (empty abasic sites), fluorophores (fluorescent trackers) and linkers (binding/modification sites). 5' and 3' denote the direction along the DNA phosphate backbone.*

The choice of DNA modifications depends upon the desired applications of the synthesised DNA. For example, fluorophores allow for the tracking of DNA complexes,[79-81] linkers provide sites of attachment for more complex chemistries[28, 82-84] and spacers mimic naturally occurring DNA abasic sites.[85, 86] Abasic sites are the vacant positions in a DNA sequence where a nucleobase is normally present, leaving the connecting sugar-phosphate backbone. Abasic sites are common DNA lesions in nature due to spontaneous depurination, depyrimidination or

deamination (Figure 1-8).[2, 87-89] If left unrepaired, abasic sites can lead to errors in DNA transcription and replication, resulting in mutagenesis or oncogenesis.[87, 89, 90] As for artificially modified DNA spacers, the deoxyribose sugar is removed, leaving only the spacer-phosphate backbone (Figure 1-7).

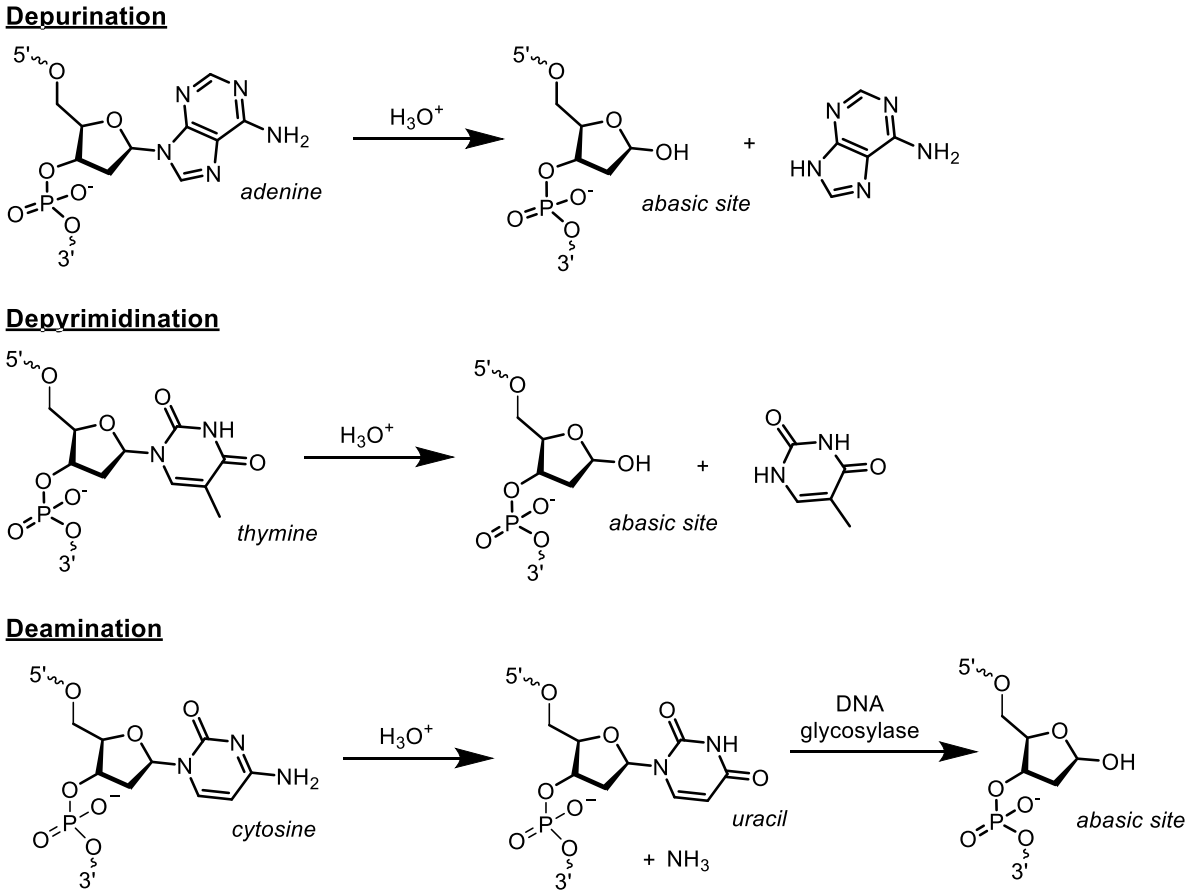


Figure 1-8: *Formation of naturally occurring abasic sites in DNA transcription and replication: depurination, depyrimidination and deamination in acidic conditions.*

1.3 DNA-templated synthesis

1.3.1 DNA-templated synthesis mechanism

One of the synthetic techniques utilised in DECLs for compound/polymer growth is DNA-templated synthesis (DTS). DTS uses short single strands of DNA (ssDNA)

with chemical tags attached to direct specific reactions in a sequence-defined manner.[91] Through the hybridisation of complementary DNA strands (dsDNA), the reactive tags are brought into proximity, increasing the local concentration and initiating a chemical reaction (Figure 1-9A). The reaction concentrations are significantly smaller than the mM to M scale of traditional chemistries, achieving successful bond formation at nM concentrations.[55] With every chemical tag specific to a unique segment in a DNA sequence, the DNA barcode can be used for DECL identification of the combined chemistries.[91-93]

Typically, subsequent strands can be incorporated *via* strand displacement to cleave "used" strands, allowing for new, unreacted, chemical tags to be introduced (Figure 1-9B).[92] By selecting the appropriate reactive tags, polymers can be produced and recorded through the sequence-specificity of DNA nucleobases. The order of reactive tags introduced to the polymer chain is programmed within the DNA sequence, as with the ribosome for amino acids with their corresponding tRNA codons. Following a similar mechanism to the ribosome, DTS aims to mimic the programmable sequential addition of reactive tags for an identifiable polymer, while expanding the pool of functional chemistries.[91-95] In addition, because of the complementary base binding of DNA, different DNA architectures can be orchestrated to promote sufficient proximity of the reactive tags (Figure 1-9C).[25, 62, 66, 92, 93, 96]

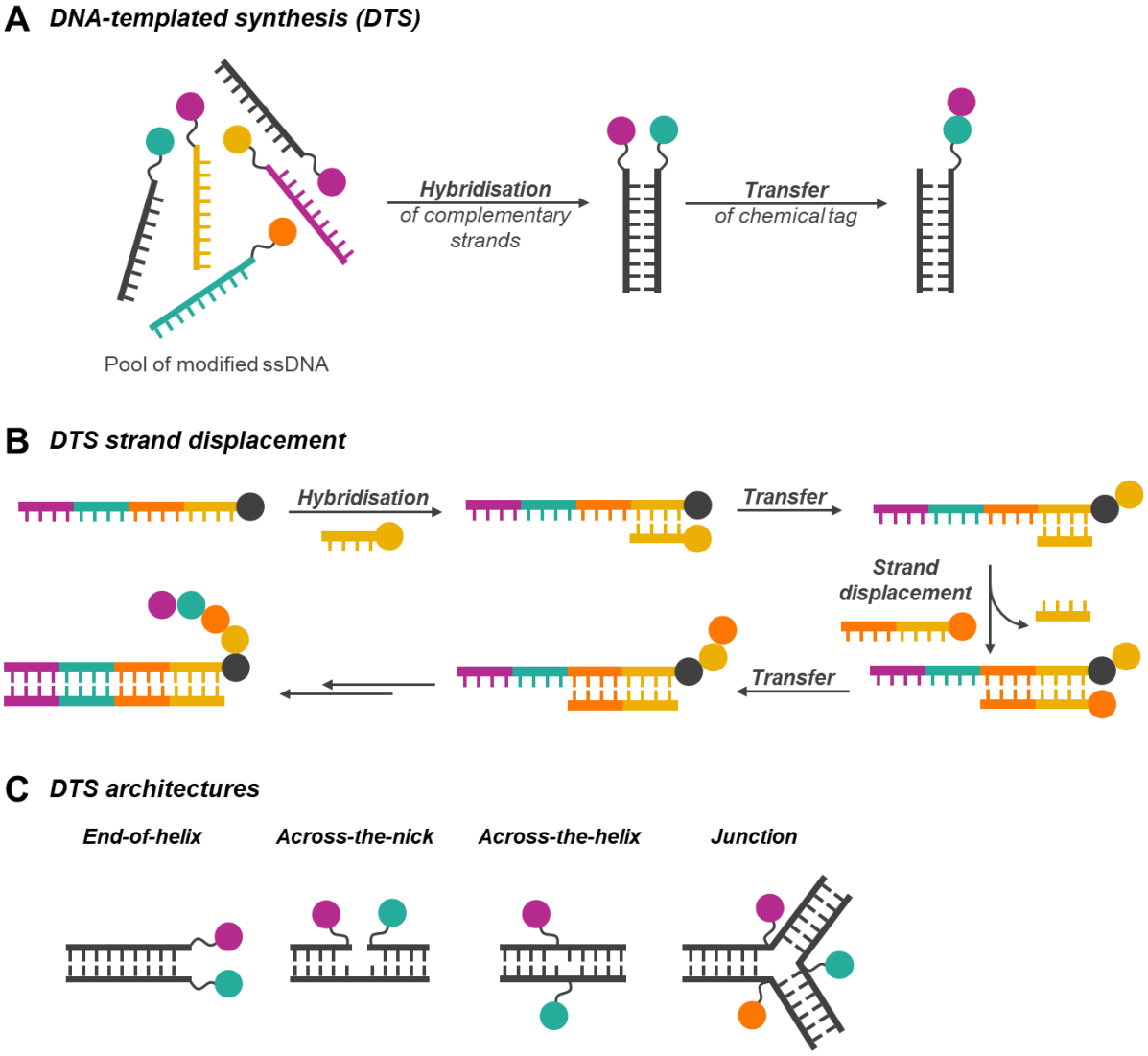


***Figure 1-9: A**) Simplified schematic of DTS. Within a pool of modified ssDNA, complementary strands hybridise, bringing reactants into proximity. Upon reaction completion, the chemical tag transfers from one DNA to another. **B**) DTS strand displacement mechanisms demonstrate how new chemical tags are introduced to the polymer chain. **C**) Common DTS DNA architectures to ensure localisation of reactants.*

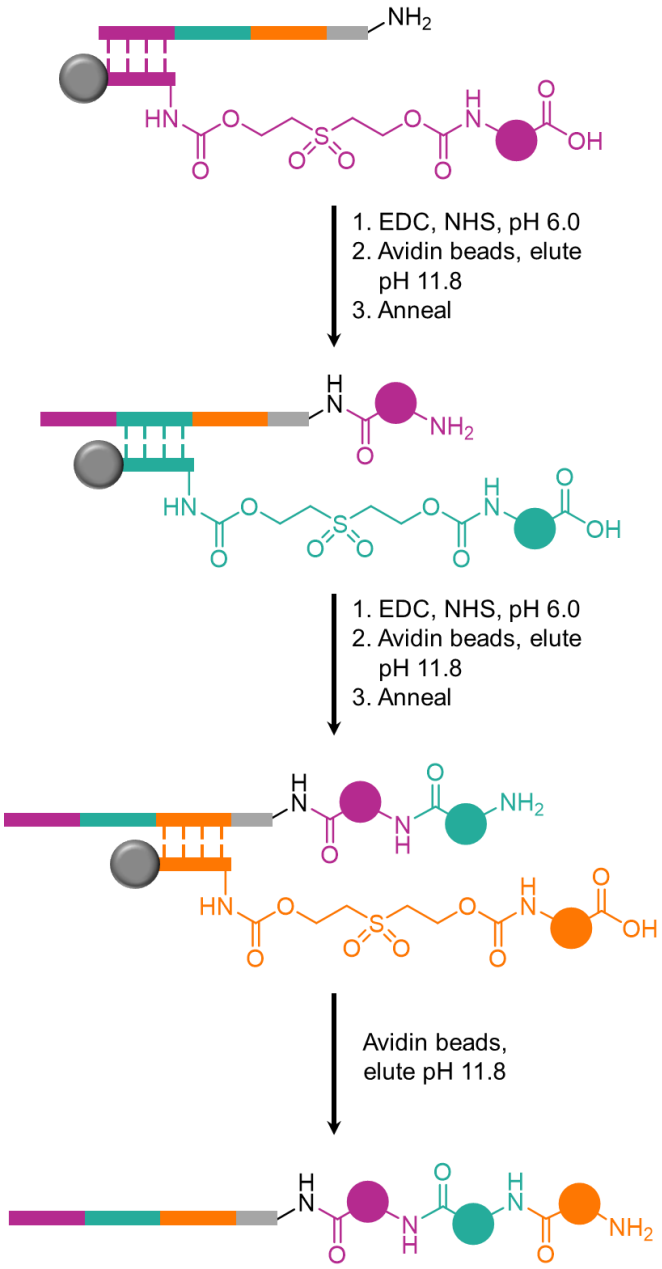
1.3.2 Multistep DNA-templated synthesis

Successful multistep DTS reactions have been performed by the Hansen group[93], Liu group[25, 62, 66, 92, 96] and the O'Reilly/Turberfield groups[58, 97, 98] to yield modified peptides and oligomers. Aminolysis and Wittig olefination have been the key chemistries in transferring reactive tags between DNA strands to synthesise sequence-specific oligomers.

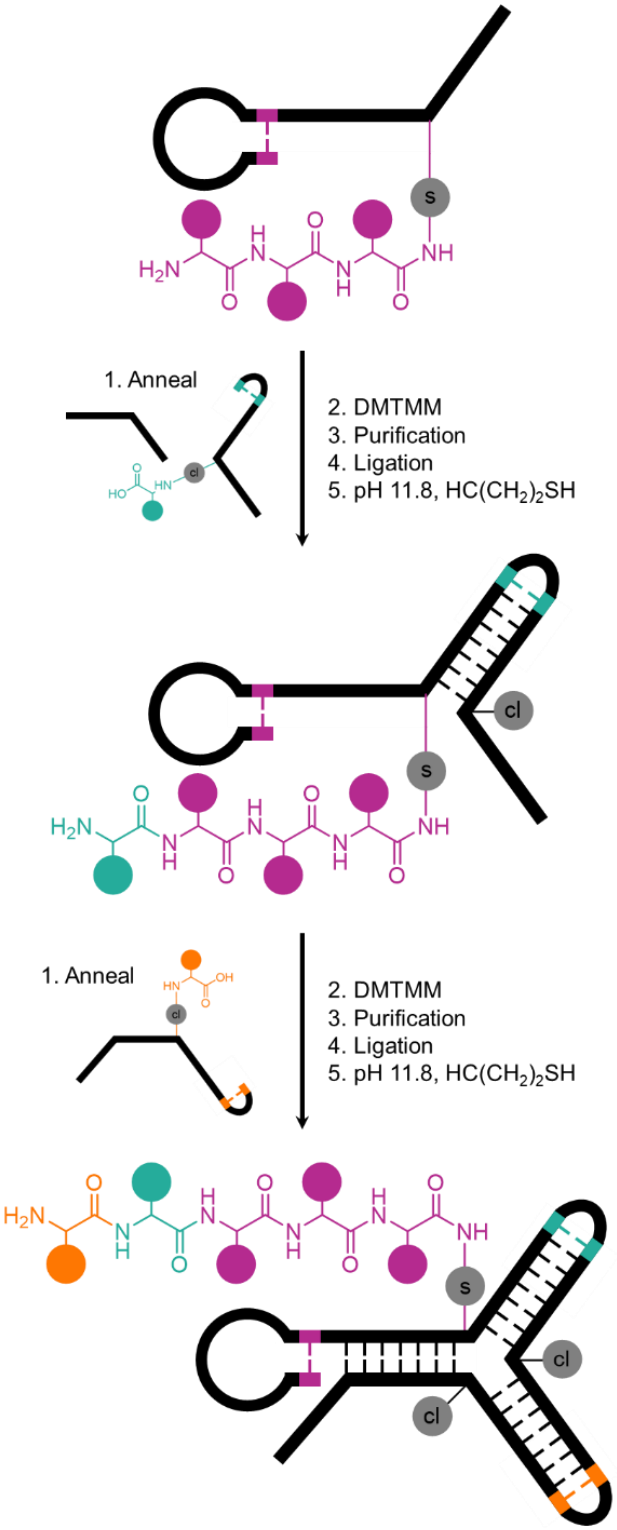
The first multistep DTS aminolysis mechanism was performed by Gartner *et al* over two decades ago (Scheme 1-4).[62] Using a 30-base DNA template stand with three coding regions, biotin-modified 10-base DNA building blocks were introduced sequentially and activated *in situ* using 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and *N*-Hydroxysuccinimide (NHS). The strong biotin-streptavidin binding of the shorter, biotinylated spent adapters allowed for sufficient removal from the reaction mixture using magnetic beads. Following three reactive steps and purification, the tripeptide was isolated in a 3% overall yield. The DTS strategy was applied later to synthesise a library of 13824 macromolecules,[66] and without periodic purification[96], by the same research group.

Following a similar *in situ* coupling strategy, Hansel *et al.* successfully performed a 3-step DTS reaction within the yocto-scale core of a three-way junction using ssDNA addition (Scheme 1-5).[93] In this architecture, an unmodified helper strand initiated the yocto-reactor formation, ensuring identical proximity of all chemical tags. As each leg of the junction was coded with a unique region (coloured), the peptide sequence could be identified, similar to the ribosome. 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM) was used as a carboxylic acid

coupling agent to perform aminolysis with the primary amine of amino acids. Following their addition and activation, the second and third adapters were cleaved from their initial DNA platform at pH 11.8, degrading an oxoester bond and exposing a terminal amine.

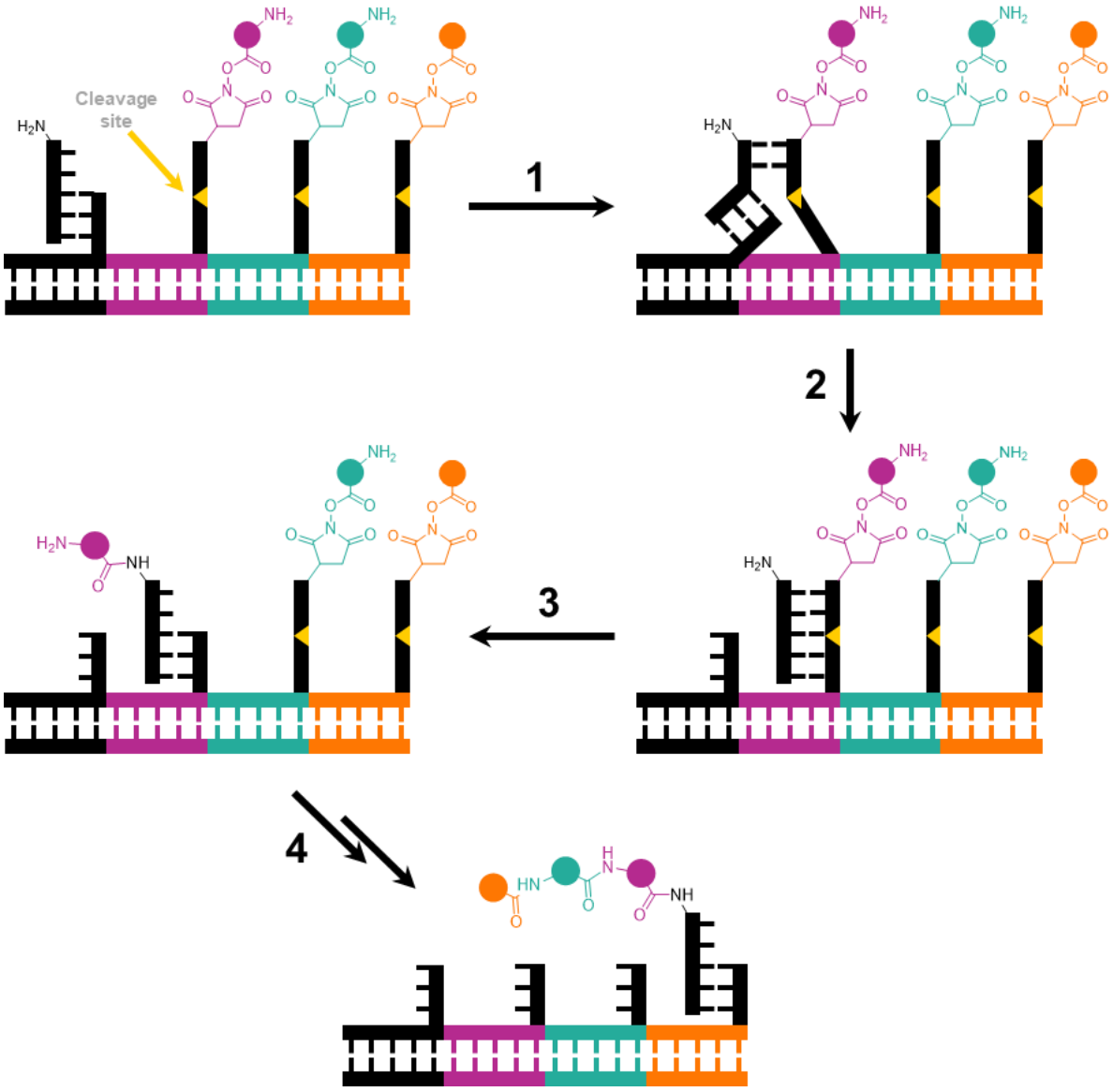


***Scheme 1-4**: Sequential activated aminolysis multistep DTS performed by Gartner et al. Coloured circles: varying monomer units. Grey sphere: biotin modification for strand cleavage.*[62]



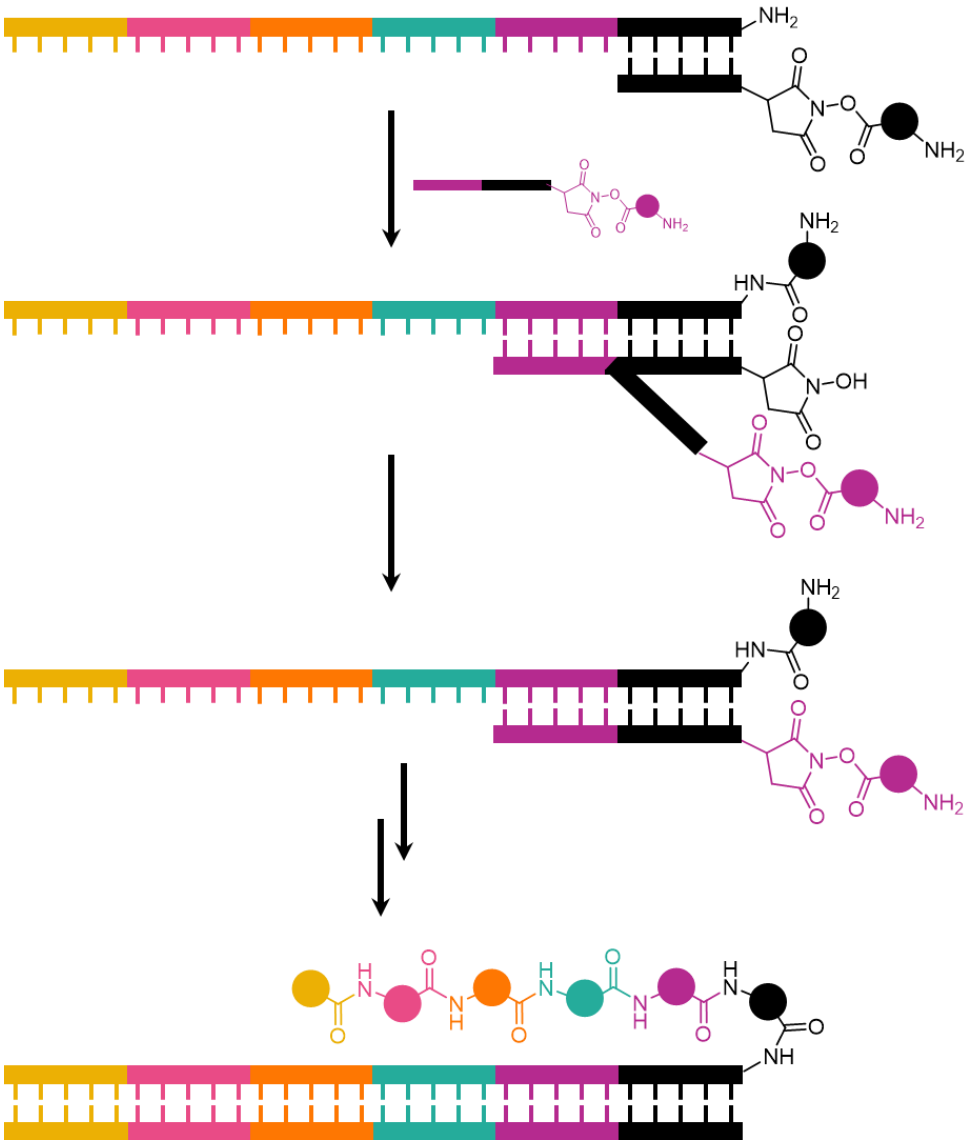
Scheme 1-5: *Yocto-scale aminolysis multistep DTS via three-way junction performed by Hansen et al.*[93] *Coloured circles: varying monomer units. Grey circle: cleavage (cl) or spacer (s) units connecting modifications to DNA strands.*

Building upon their DTS methods, Liu and coworkers developed an autonomous strand displacement DTS method: a DNA walker (Scheme 1-6).[25] Without external stimuli, a ssDNA is programmed to “walk” across a DNA template harbouring an initiation site and multiple codon sites, encoding for different modified DNA strands. The modified strands featured a primary amine linked by an NHS ester, two DNA docking sites and a complementary region to the template strand. The DNA walker consists of the complementary docking site sequences as well as an initiating primary amine. The movement of the walker is prompted by the DNAzyme-mediated cleavage of RNA nucleotides placed in the modified strand sequence, inducing instability in the dsDNA after successful aminolysis. In comparison to the prior multistep DTS, this mechanism operated autonomously and at an estimated yield of 45% over three steps. Outside of DECLs, the DNA walker has contributed to a range of nano-scale technologies.[99-101]



Scheme 1-6: *DNA "walker" aminolysis multistep performed by He et al.[25] Coloured circles: varying monomer units. Yellow triangle: RNA cleavage site using a cleaving DNAzyme.*

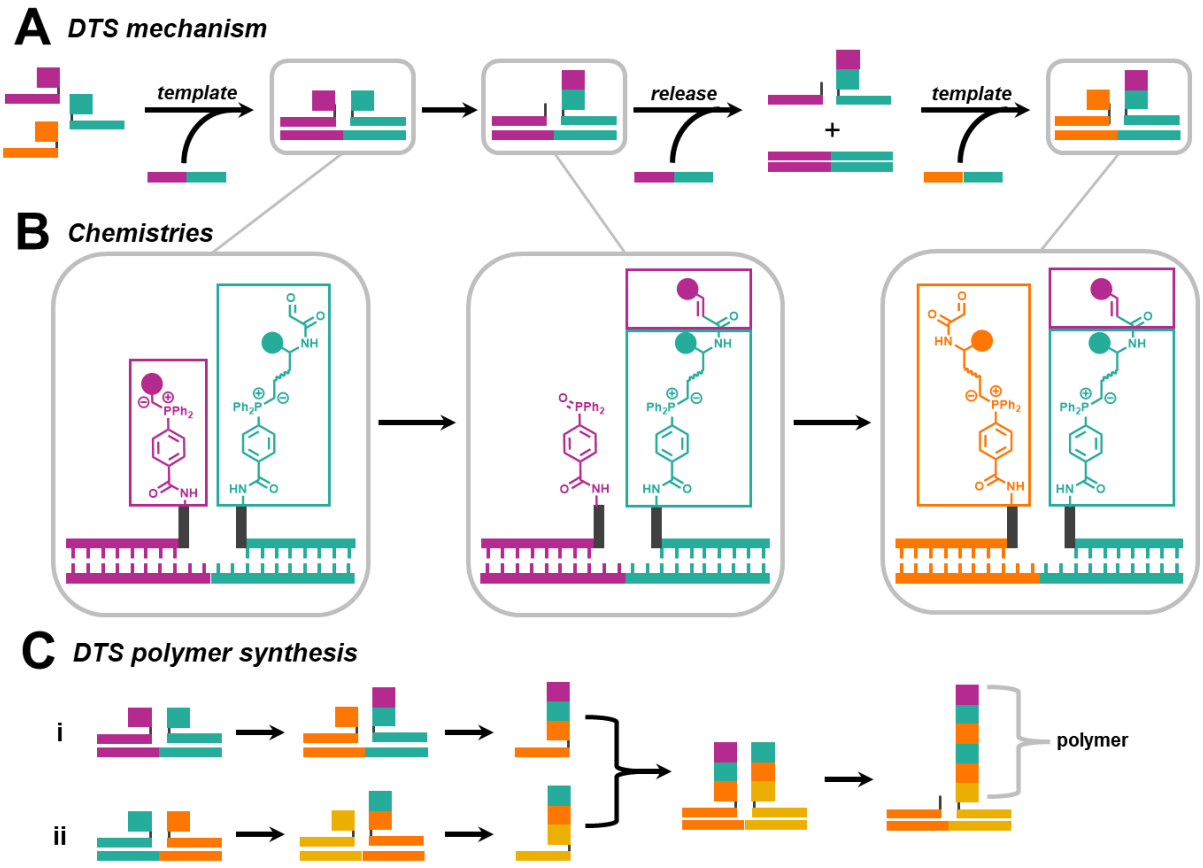
Improving strategies further, Liu and coworkers successfully synthesised a 6-mer using an end-of-helix strand displacement method (Scheme 1-7).[92] In a sequence, longer complementary strands were added to the reaction mixture, displacing spent adapters, and subsequently removing the need for intermittent purification steps. Despite not being autonomous, this mechanism yielded a 6-mer at 35% using NHS esters and primary amine chemical tags. However, the reaction was limited by the increase in distance between reactive tags with oligomer extension.



Scheme 1-7: *Strand displacement aminolysis multistep DTS performed by He et al. Coloured circles: varying monomer units based upon the added DNA sequence.*[92]

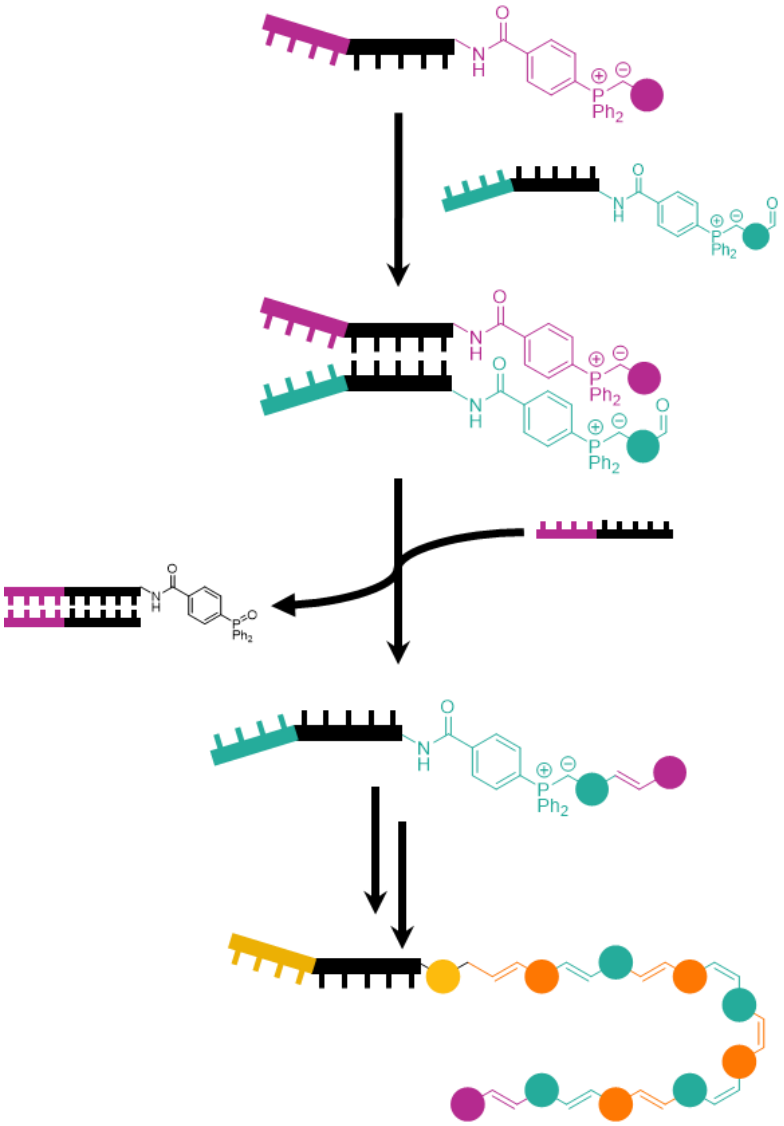
Within the same year, the O'Reilly and Turberfield groups published two DTS multistep mechanisms, transferring reactive tags *via* the Wittig reaction.[58, 98] The first mechanism required the use of a templating base to pair short complementary strands.[58] One of the reactive tags installed occupied two functionalities, aldehyde and ylide groups, to allow conjugation with two monomers. Upon transfer, the template was cleaved by strand displacement with a fully complementary strand. Two reactions were

performed in a single pot, in parallel, to produce a trimer. The trimeric intermediates were subsequently coupled in the same pot to afford a final hexamer, yielding at 14%. However, yields were limited by the oxidation of the phosphonium ylide moiety.



Scheme 1-8: *Templated Wittig multistep DTS performed by McKee et al.[58] Coloured circles: varying monomer units based upon the added DNA sequence.* ***A****) Mechanism to build short polymers.* ***B****) Wittig chemistries utilised for the transfer of reactive tags.* ***C****) Mechanism bringing short polymers together to yield a hexamer.*

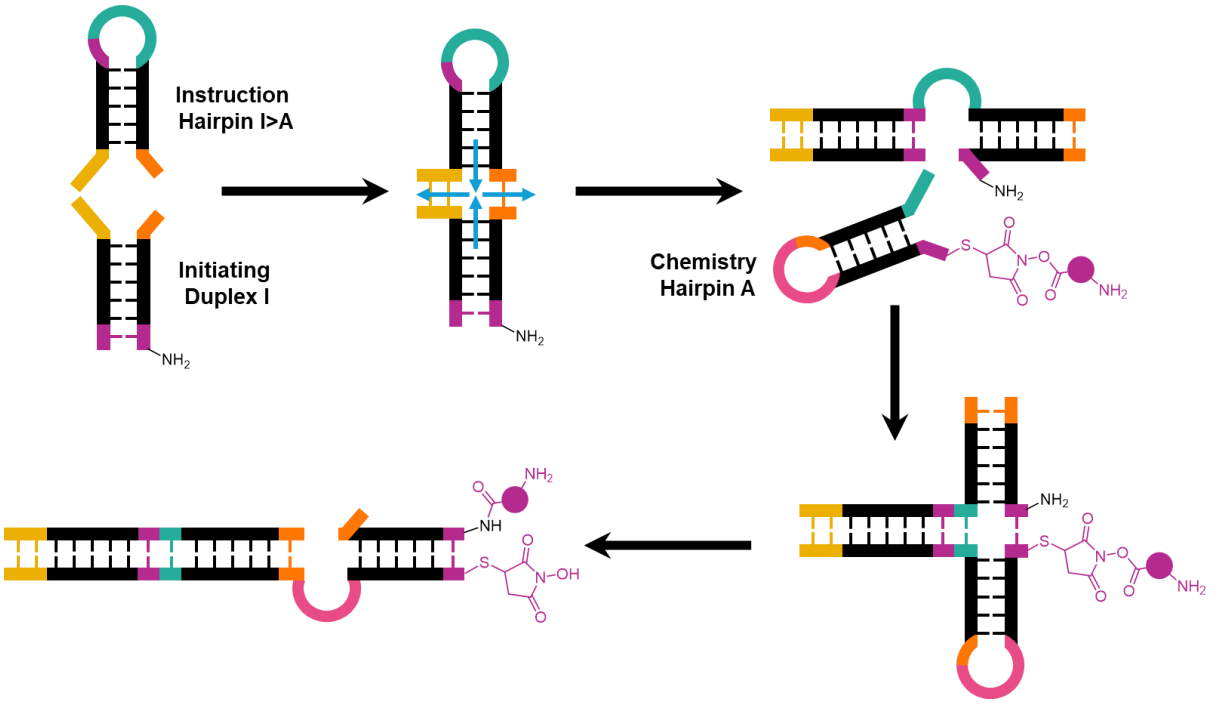
The second Wittig-mediated multistep mechanism synthesised the longest recorded polymer by DTS: a decamer (Scheme 1-9).[98] Using a simple strand-displacement mechanism, yields of 85% per reactive step were observed. However, as the order of chemical tags was not coded within the DNA sequence, the mechanism cannot be applied to evolutionary DECL experiments.



Scheme 1-9: *Strand displacement Wittig multistep DTS performed by Milnes et al. Coloured circles: varying monomer units based upon the added DNA sequence.*[98]

The most recent use of esters in multistep DTS was demonstrated by Meng *et al.* using a hairpin chain reaction (HCR) (Scheme 1-10).[97] Uncomplementary regions in a dsDNA form hairpin loops, which open using initiator strands, while becoming accessible to incoming functional DNA strands. Toehold regions located on the processed hairpin dictate the sequence of functional DNA hybridisation and, therefore, the order of reactive tags. In an autonomous fashion, an amine and an NHS ester are

brought into proximity. Over the course of 3 steps, 4 polypeptides were synthesised at 5-10% yield, of varying amino acid sequences.



Scheme 1-10: *HCR aminolysis multistep DTS performed by Meng et al. Coloured circles: varying monomer units based upon the added DNA sequence.*[97]

1.3.3 Limitations to DNA-templated synthesis

HCR and the DNA walker demonstrate DTS's capability to mimic ribosomal peptide synthesis; autonomously synthesising sequence-specific polymers in a single pot and with an identifiable DNA sequence.[25, 97] However, these mechanisms have only conjugated a few monomers in short sequences, unlike the ribosome, which is capable of producing long, complex proteins.

To date, the longest polymer synthesised by DTS was a decamer, and yields have not progressed beyond 45% overall.[25, 58] Typically, DTS requires aqueous conditions to support the hydrogen bonding necessary for DNA hybridisation. However, such conditions can lead to the degradation (esters) and oxidation

(phosphonium ylide) of chemical tags. In the case of activated esters, the electrophilic aminolysis chemical tags are prone to hydrolysis under the high pH conditions often required to successfully initiate a nucleophilic attack by the primary amine, limiting the desired polymer length.[55] Whereas for the ribosome, the activated ester is stabilised by the EF-Tu and GTP complex, preventing its degradation. However, Liu and coworkers demonstrated DTS in organic solvents when the DNA was pre-hybridised in aqueous solution.[102] Nonetheless, the mechanism and chemistry installed in this protocol lacked an autonomous nature, making it an unsuitable candidate when applied to a DTS molecular machine.

To overcome yield and polymer growth limitations, alternative transfer chemistries have been considered in single-step DTS mechanisms. For example, acyl transfer (aminolysis[102-106] and native chemical ligation[107-109]), Wittig olefination,[102, 110-114] nucleophilic aromatic substitution (S_NAr),[115, 116] and tetrazine transfer (Figure 1-10).[117] Other chemistries, such as photoligation,[118] reductive amination[119] and nitro-aldol[110], have been employed in DTS mechanisms. However, these coupling chemistries do not allow for the transfer and extension of chemical tags.[55]

Of the DTS transfer chemistries investigated, aminolysis has provided the greatest success in autonomous multistep DTS reactions[25, 97] and has produced the highest transfer yields.[25, 103, 106] By imitating the aminoacyl conjugation within the ribosome, aminolysis-based chemistries currently represent the most promising approach for facilitating the assembly of a DTS-based molecular machine. Nevertheless, optimisation of aminolysis-mediated DTS is required before this goal can be achieved.

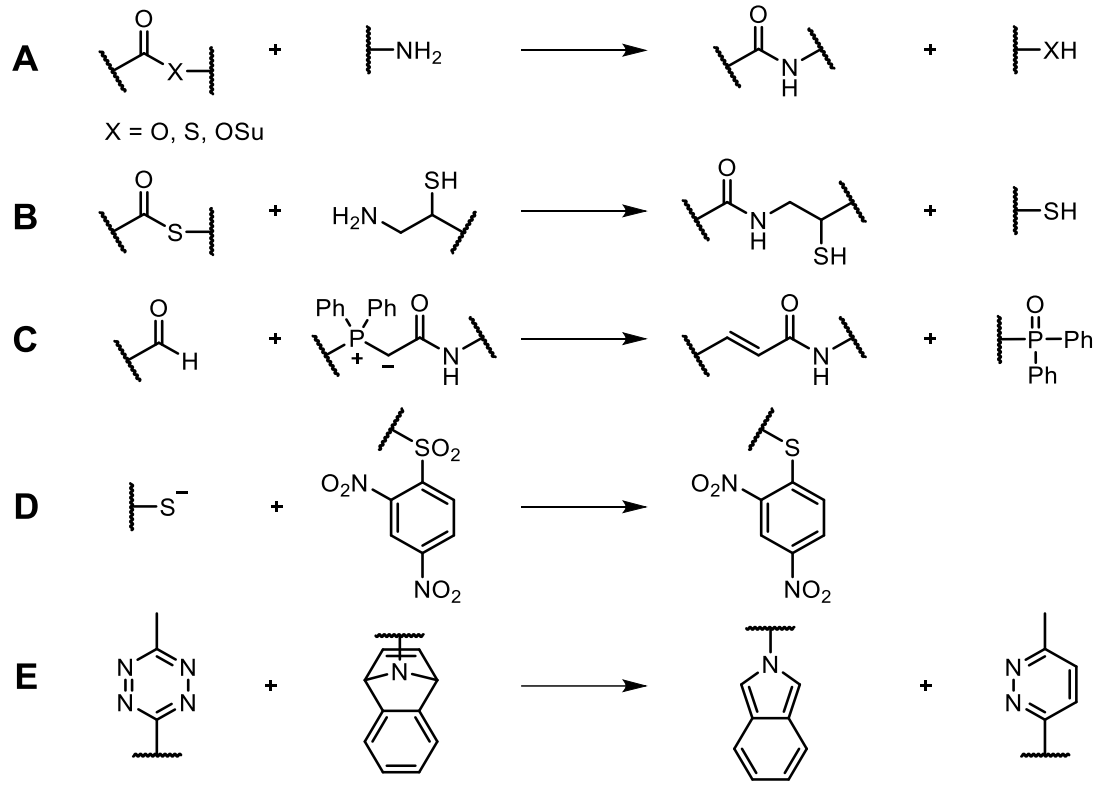


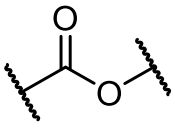
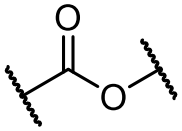
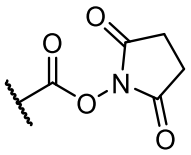
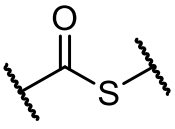
***Figure 1-10**: Types of transfer chemistries employed in DTS: acyl transfer: **A**) aminolysis (e.g. oxoester, thioester or activated ester),[102-106] **B**) native chemical ligation.[107-109] **C**) Wittig olefination.[102, 110-114] **D**) S_NAr.[115, 116] **E**) tetrazine transfer.[117]*

The most significant differences between ribosomal and DTS aminolysis are the electrophilic species used and the involvement of the peptidyl transferase enzyme in the ribosome (Table 1-1). As mentioned previously, peptidyl transferase catalyses the formation of peptide bonds between amino acids *via* aminolysis, using an oxoester and a primary amine. However, without the presence of an enzyme, aminolysis reaction rates using an oxoester can be sluggish, as previously demonstrated within the research group.[120]

To ensure successful reaction rates without the use of a complex enzymatic system, DTS reactions within the group have employed a stronger electrophile – a thioester.[103, 106] As sulfur is atomically larger than oxygen, the orbital overlap of the lone pair, within the sp^2 hybridised system, is reduced from the oxoesters *2p* orbital to

the thioesters *3p* orbital.[121] Thus, reducing the resonance delocalisation of the thioester and increasing the electrophilicity of its carbonyl group. In addition, thiols (pK_a = 10) are weaker bases than the oxoester's alcohols (pK_a = 15), improving the leaving group ability and allowing for conjugation with the incoming primary amine.[122] However, as the thioester increases the electrophilicity of the reactive system, hydrolysis increases. These rates are accelerated by the high pHs required for primary amine deprotonation in aqueous solution, to produce the nucleophilic $-NH_2$ species. By optimising additional parameters within the system, such as architecture, electrophiles and buffers, DTS yields would increase whilst limiting the competing hydrolysis reaction, aiding in the creation of a DTS molecular machine.

Table 1-1: *Comparison between ribosomal aminoacyl transfer and the most successful DTS aminolysis transfer reactions.*

Features	**Ribosomal aminoacyl transfer**[2, 21, 22]	**DTS aminolysis transfer**[25, 103, 106, 120]		
Ester(s) used for amide bond formation	 *oxoester*	 *oxoester*	 *NHS ester*	 *thioester*
Enzyme	Peptidyl transferase	None	None	None
Hydrolysis rate	Low	Low	High	High
Hydrolysis stabilisers	EF-Tu·GTP complex	None	None	None
Yields	High	Low	Mid-range	High
Multistep capabilities	Excellent	Not documented	Good	Not documented

1.4 Project aims

To date, hydrolysis of transfer chemistries has limited the progression of multi-step DTS; something the ribosome can control using EF-Tu and GTP with aminoacyl-tRNA. Because of this, the goal of synthesising an artificial ribosome for autonomous, controllable, polymer synthesis has been restricted. Previous methods in DTS optimisation have focused on alternative reactive linkers to improve reaction yields. However, aminolysis chemistries have proven to be the most effective method in autonomous, multistep DTS synthesis and improving product yields.[25, 97, 103, 106] Therefore, this thesis aimed to improve existing aminolysis chemistries within the O'Reilly and Turberfield groups. The routes of optimisation included DNA architecture-supported stabilisation and the replacement of the electrophilic species.

Chapter 2 investigated the thioester stabilisation observed by Frommer *et al.* in an across-the-helix DTS architecture.[106] Using absorbance and fluorescence spectroscopy, changes in the modification's environment could be hypothesised, developing a deeper understanding of how the architecture operates. In addition, the nucleobase sequence of the architecture was adjusted with the aim of improving thioester stabilisation and improving DTS yields.

Chapter 3 aimed to employ selenoester chemistries within a DTS framework, minimising hydrolysis and improving aminolysis yields. An initial small molecule scoping study was conducted to understand selenoester's capability in a DTS-compatible aqueous buffered solution. The chemistries employed were later incorporated into an across-the-helix architecture and compared to the groups' previously established DTS thioester analogue.[106]

The objective of Chapter 4 was to put Chapter 3′s results into practice. Aiming to optimise the selenoester's DTS reactive conditions, multiple parameters were investigated, such as pH/buffer conditions, incubation times and reactive nucleophile species. Building upon optimal conditions, the preparation of a selenoester-initiated multistep DTS mechanism was attempted.

1.5 References

1. I. Aprahamian, *ACS Cent. Sci.*, 2020, **6**, 347-358.
2. L. A. Moran, R. Horton, G. Scrimgeour and M. Perry, *Principles of Biochemistry: International Edition*, Pearson Education (US), United States, 5th edn., 2011.
3. M. J. Hammerling, B. R. Fritz, D. J. Yoesep, D. S. Kim, E. D. Carlson and M. C. Jewett, *Nature Commun.*, 2020, **11**, 1108.
4. A. S. Petrov, C. R. Bernier, C. Hsiao, A. M. Norris, N. A. Kovacs, C. C. Waterbury, V. G. Stepanov, S. C. Harvey, G. E. Fox, R. M. Wartell, N. V. Hud and L. D. Williams, *Proc. Natl. Acad. Sci.*, 2014, **111**, 10251-10256.
5. J. C. Bowman, A. S. Petrov, M. Frenkel-Pinter, P. I. Penev and L. D. Williams, *Chem. Rev.*, 2020, **120**, 4848-4878.
6. F. Zinoni, A. Birkmann, T. C. Stadtman and A. Böck, *Proc. Natl. Acad. Sci. U.S.A.*, 1986, **83**, 4650-4654.
7. B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki and M. K. Chan, *Science*, 2002, **296**, 1462-1466.
8. A. J. Bruce Alberts, Julian Lewis, Martin Raff, Keith Roberts, Peter Walter. , in *Molecular Biology of the Cell*, Garland Science, New York, 4th edition edn., 2002, ch. Chapter 3, pp. 134-169.
9. V. Jayaraman, S. Toledo-Patiño, L. Noda-García and P. Laurino, *Protein Sci.*, 2022, **31**, 1-18.
10. D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, W.H. Freeman, 5 edn., 2008.
11. S. Bhaduri, N. Ranjan and D. P. Arya, *Beilstein J Org Chem*, 2018, **14**, 1051-1086.
12. V. Ramakrishnan, *Cell*, 2014, **159**, 979-984.
13. D. Moazed and H. F. Noller, *Nature*, 1989, **342**, 142-148.
14. N. T. Ingolia, S. Ghaemmaghami, J. R. S. Newman and J. S. Weissman, *Science*, 2009, **324**, 218-223.
15. M. Beringer and M. V. Rodnina, *Mol. Cell*, 2007, **26**, 311-321.
16. M. Simonović and T. A. Steitz, *Biochim. Biophys. Acta*, 2009, **1789**, 612-623.
17. H. Stark, E. V. Orlova, J. Rinke-Appel, N. Jünke, F. Mueller, M. Rodnina, W. Wintermeyer, R. Brimacombe and M. van Heel, *Cell*, 1997, **88**, 19-28.
18. E. B. Kramer and P. J. Farabaugh, *RNA*, 2007, **13**, 87-96.
19. V. Ramakrishnan, *Cell*, 2002, **108**, 557-572.
20. H. S. Zaher and R. Green, *Cell*, 2009, **136**, 746-762.

21. J. R. Peacock, R. R. Walvoord, A. Y. Chang, M. C. Kozlowski, H. Gamper and Y.-M. Hou, *RNA*, 2014, **20**, 758-764.

22. T. Wagner and M. Sprinzl, *Eur. J. Biochem.*, 1980, **108**, 213-221.

23. J. R. Menninger, *J. Biol. Chem.*, 1976, **251**, 3392-3398.

24. A. Yonath, *Angew. Chem. Int. Ed.*, 2010, **49**, 4340-4354.

25. Y. He and D. R. Liu, *Nat. Nanotechnol.*, 2010, **5**, 778-782.

26. B. Lewandowski, G. De Bo, J. W. Ward, M. Papmeyer, S. Kuschel, M. J. Aldegunde, P. M. E. Gramlich, D. Heckmann, S. M. Goldup, D. M. D'Souza, A. E. Fernandes and D. A. Leigh, *Science*, 2013, **339**, 189-193.

27. G. De Bo, M. A. Y. Gall, M. O. Kitching, S. Kuschel, D. A. Leigh, D. J. Tetlow and J. W. Ward, *J. Am. Chem. Soc.*, 2017, **139**, 10875-10879.

28. A. A. Peterson and D. R. Liu, *Nat. Rev. Drug Discov.*, 2023, **22**, 699-722.

29. M. Dockerill, P. M. Sabale, F. Russo, S. Barluenga and N. Winssinger, *JACS Au*, 2024, **4**, 4013-4022.

30. C. Kofman, J. A. Willi, A. S. Karim and M. C. Jewett, *ACS Cent. Sci.*, 2024, **10**, 871-881.

31. E. C. Fischer, K. Hashimoto, Y. Zhang, A. W. Feldman, V. T. Dien, R. J. Karadeema, R. Adhikary, M. P. Ledbetter, R. Krishnamurthy and F. E. Romesberg, *Nat. Chem. Biol.*, 2020, **16**, 570-576.

32. J. D. Puglisi, *Nature*, 2015, **524**, 45-46.

33. J. Hanes and A. Plückthun, *Proc. Natl. Acad. Sci. U.S.A.*, 1997, **94**, 4937-4942.

34. T. Katoh, T. Sengoku, K. Hirata, K. Ogata and H. Suga, *Nat. Chem.*, 2020, **12**, 1081-1088.

35. F. E. Romesberg, *Philos Trans R Soc Lond B Biol Sci*, 2023, **378**, 20220030.

36. K. Hamashima, M. Kimoto and I. Hirao, *Curr. Opin. Chem. Biol.*, 2018, **46**, 108-114.

37. S. E. Morris, A. W. Feldman and F. E. Romesberg, *ACS Synth. Biol.*, 2017, **6**, 1834-1840.

38. Y. J. Seo, G. T. Hwang, P. Ordoukhanian and F. E. Romesberg, *J. Am. Chem. Soc.*, 2009, **131**, 14596-14596.

39. L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498-500.

40. J. S. Italia, P. S. Addy, S. B. Erickson, J. C. Peeler, E. Weerapana and A. Chatterjee, *J. Am. Chem. Soc.*, 2019, **141**, 6204-6212.

41. A. Chatterjee, S. B. Sun, J. L. Furman, H. Xiao and P. G. Schultz, *Biochemistry*, 2013, **52**, 1828-1837.

42. A. J. Burke, S. L. Lovelock, A. Frese, R. Crawshaw, M. Ortmayer, M. Dunstan, C. Levy and A. P. Green, *Nature*, 2019, **570**, 219-223.

43. D. L. Dunkelmann, J. C. W. Willis, A. T. Beattie and J. W. Chin, *Nat. Chem.*, 2020, **12**, 535-544.

44. B. Moore, B. C. Persson, C. C. Nelson, R. F. Gesteland and J. F. Atkins, *J. Mol. Biol.*, 2000, **298**, 195-209.

45. T. J. Magliery, J. C. Anderson and P. G. Schultz, *J. Mol. Biol.*, 2001, **307**, 755-769.

46. H. Neumann, K. Wang, L. Davis, M. Garcia-Alai and J. W. Chin, *Nature*, 2010, **464**, 441-444.

47. U. Chheda, S. Pradeepan, E. Esposito, S. Strezsak, O. Fernandez-Delgado and J. Kranz, *J. Pharm. Sci.*, 2024, **113**, 377-385.

48. E. S. Iqbal, K. K. Dods and M. C. T. Hartman, *Org. Biomol. Chem.*, 2018, **16**, 1073-1078.

49. J. J. Butzow and G. L. Eichhorn, *Nature*, 1975, **254**, 358-359.

50. Q. Jiang, C. Song, J. Nangreave, X. Liu, L. Lin, D. Qiu, Z. G. Wang, G. Zou, X. Liang, H. Yan and B. Ding, *J. Am. Chem. Soc.*, 2012, **134**, 13396-13403.

51. P. W. K. Rothemund, *Nature*, 2006, **440**, 297-302.

52. N. Seeman and H. Sleiman, *Nat. Rev. Mater.*, 2018, **3**, 1-20.

53. F. Zhang, S. Jiang, S. Wu, Y. Li, C. Mao, Y. Liu and H. Yan, *Nat. Nanotechnol.*, 2015, **10**, 779-784.

54. W. M. Shih and C. Lin, *Curr. Opin. Struct. Biol.* 2010, **20**, 276-282.

55. R. K. O'Reilly, A. J. Turberfield and T. R. Wilks, *Acc. Chem. Res.*, 2017, **50**, 2496-2509.

56. S. Brenner and R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.*, 1992, **89**, 5381-5383.

57. H. Sun, L. Yang, M. P. Thompson, S. Schara, W. Cao, W. Choi, Z. Hu, N. Zang, W. Tan and N. C. Gianneschi, *Bioconjugate Chem.*, 2019, **30**, 1889-1904.

58. M. L. McKee, P. J. Milnes, J. Bath, E. Stulz, R. K. O'Reilly and A. J. Turberfield, *J. Am. Chem. Soc.*, 2012, **134**, 1446-1449.

59. X. Xiong, C. Wu, C. Zhou, G. Zhu, Z. Chen and W. Tan, *Macromol. Rapid Commun.*, 2013, **34**, 1271-1283.

60. *United States of America Pat.,* US11332493B2, 2022.

61. H. V. P. Thelu, S. K. Albert, M. Golla, N. Krishnan, D. Ram, S. M. Srinivasula and R. Varghese, *Nanoscale*, 2018, **10**, 222-230.

62. Z. J. Gartner, M. W. Kanan and D. R. Liu, *J. Am. Chem. Soc.*, 2002, **124**, 10304-10306.

63. H. Gu, J. Chao, S. J. Xiao and N. C. Seeman, *Nature*, 2010, **465**, 202-205.

64. T. E. Tomov, R. Tsukanov, M. Liber, R. Masoud, N. Plavner and E. Nir, *J. Am. Chem. Soc.*, 2013, **135**, 11935-11941.

65. S. Iqbal, W. Jiang, E. Hansen, T. Aristotelous, S. Liu, A. Reidenbach, C. Raffier, A. Leed, C. Chen, L. Chung, E. Sigel, A. Burgin, S. Gould and H. H. Soutter, *npj Drug Discov.*, 2025, **2**, 5.

66. Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder and D. R. Liu, *Science*, 2004, **305**, 1601-1605.

67. D. R. Halpin and P. B. Harbury, *PLoS Biology*, 2004, **2**, 1015-1021.

68. F. Debaene, L. Mejias, J. L. Harris and N. Winssinger, *Tetrahedron*, 2004, **60**, 8677-8690.

69. S. Melkko, J. Scheuermann, C. E. Dumelin and D. Neri, *Nat. Biotechnol.*, 2004, **22**, 568-574.

70. Q. Nie, X. Fang, J. Huang, T. Xu, Y. Li, G. Zhang and Y. Li, *Small Methods*, 2025, 2401631.

71. X. Li and D. R. Liu, *Angew. Chem. Int. Ed.*, 2004, **43**, 4848-4870.

72. Y. Li, P. Zhao, M. Zhang, X. Zhao and X. Li, *J. Am. Chem. Soc.*, 2013, **135**, 17727-17730.

73. F. V. Reddavide, W. Lin, S. Lehnert and Y. Zhang, *Angew Chem Int Ed*, 2015, **54**, 7924-7928.

74. J. Scheuermann and D. Neri, *Curr. Opin. Chem. Biol.*, 2015, **26**, 99-103.

75. J. L. Harris and N. Winssinger, *Chem. Eur. J.*, 2005, **11**, 6792-6801.

76. M. H. Caruthers, A. D. Barone, S. L. Beaucage, D. R. Dodds, E. F. Fisher, L. J. McBride, M. Matteucci, Z. Stabinsky and J. Y. Tang, *Methods Enzymol.*, 1987, **154**, 287-313.

77. Y. Z. Xu, Q. Zheng and P. F. Swann, *J. Org. Chem.*, 1992, **57**, 3839-3845.

78. S. L. Beaucage and M. H. Caruthers, *Tetrahedron Lett.*, 1981, **22**, 1859-1862.

79. D. Egloff, I. A. Oleinich, M. Zhao, S. L. B. König, R. K. O. Sigel and E. Freisinger, *ACS Chem. Biol.*, 2016, **11**, 2558-2567.

80. R. Crawford, J. P. Torella, L. Aigrain, A. Plochowietz, K. Gryte, S. Uphoff and A. N. Kapanidis, *Biophys. J.*, 2013, **105**, 2439-2450.

81. M. Boonstra, N. Vesel and P. Kuipers Oscar, *mBio*, 2018, **9**.

82. N.-S. Hsu, C.-C. Lee, W.-C. Kuo, Y.-W. Chang, S.-Y. Lo and A. H. J. Wang, *Bioconjugate Chem.*, 2020, **31**, 1804-1811.

83. N. O. Kropacheva, A. A. Golyshkin, M. A. Vorobyeva and M. I. Meschaninova, *Molecules*, 2023, **28**, 1904.

84. D. Honcharenko, K. Druceikaite, M. Honcharenko, M. Bollmark, U. Tedebark and R. Strömberg, *ACS Omega*, 2021, **6**, 579-593.

85. J. Chen, F. Y. Dupradeau, D. A. Case, C. J. Turner and J. Stubbe, *Nucleic Acids Res.*, 2008, **36**, 253-262.

86. O. Potapova, N. D. F. Grindley and C. M. Joyce, *J. Biol. Chem.*, 2002, **277**, 28157-28166.

87. T. Lindahl, *Nature*, 1993, **362**, 709-715.

88. T. Lindahl and B. Nyberg, *Biochemistry*, 1972, **11**, 3610-3618.

89. S. Boiteux and M. Guillet, *DNA Repair*, 2004, **3**, 1-12.

90. J. H. J. Hoeijmakers, *Nature*, 2001, **411**, 366-374.

91. Z. J. Gartner and D. R. Liu, *J. Am. Chem. Soc.*, 2001, **123**, 6961-6963.

92. Y. He and D. R. Liu, *J. Am. Chem. Soc.*, 2011, **133**, 9972-9975.

93. M. H. Hansen, P. Blakskjær, L. K. Petersen, T. H. Hansen, J. W. Højfeldt, K. V. Gothelf and N. J. V. Hansen, *J. Am. Chem. Soc.*, 2009, **131**, 1322-1327.

94. R. E. Kleiner, Y. Brudno, M. E. Birnbaum and D. R. Liu, *J. Am. Chem. Soc.*, 2008, **130**, 4646-4659.

95. J. Niu, R. Hili and D. R. Liu, *Nat. Chem.*, 2013, **5**, 282-292.

96. T. M. Snyder and D. R. Liu, *Angew. Chem. Int. Ed.*, 2005, **44**, 7379-7382.

97. W. Meng, R. A. Muscat, M. L. McKee, P. J. Milnes, A. H. El-Sagheer, J. Bath, B. G. Davis, T. Brown, R. K. O'Reilly and A. J. Turberfield, *Nat. Chem.*, 2016, **8**, 542-548.

98. P. J. Milnes, M. L. McKee, J. Bath, L. Song, E. Stulz, A. J. Turberfield and R. K. O'Reilly, *Chem. Commun.*, 2012, **48**, 5614-5616.

99. G. Qiu, Y. Wang, W. Zhang, T. Bao, Z. Wu, X. Zhang, S. Wang and W. Wen, *Anal. Chem.*, 2024, **96**, 18464-18473.

100. M. Škugor, J. Valero, K. Murayama, M. Centola, H. Asanuma and M. Famulok, *Angew. Chem. Int. Ed.*, 2019, **58**, 6948-6951.

101. M. Chen, R. Duan, S. Xu, Z. Duan, Q. Yuan, F. Xia and F. Huang, *Anal. Chem.*, 2021, **93**, 16264-16272.

102. M. M. Rozenman and D. R. Liu, *ChemBioChem*, 2006, **7**, 253-256.

103. M. L. McKee, A. C. Evans, S. R. Gerrard, R. K. O'Reilly, A. J. Turberfield and E. Stulz, *Org. Biomol. Chem.*, 2011, **9**, 1661-1666.

104. A. Erben, T. N. Grossmann and O. Seitz, *Biorg. Med. Chem. Lett.*, 2011, **21**, 4993-4997.

105. S. Núñez-Pertíñez and T. R. Wilks, *Front. Chem.*, 2020, **8**, 1-11.

106. J. Frommer, R. Oppenheimer, B. M. Allott, S. Núñez-Pertíñez, T. R. Wilks, L. R. Cox, J. Bath, R. K. O'Reilly and A. J. Turberfield, *Angew. Chem. Int. Ed.*, 2024, **163**, e202317482.

107. R. Houska, M. B. Stutz and O. Seitz, *Chem. Sci.*, 2021, **12**, 13450-13457.

108. T. N. Grossmann and O. Seitz, *J. Am. Chem. Soc.*, 2006, **128**, 15596-15597.

109. J. Michaelis, A. Maruyama and O. Seitz, *Chem. Commun.*, 2013, **49**, 618-620.

110. Z. J. Gartner, M. W. Kanan and D. R. Liu, *Angew. Chem. Int. Ed.*, 2002, **41**, 1796-1800.

111. X.-H. Chen, A. Roloff and O. Seitz, *Angew. Chem. Int. Ed.*, 2012, **51**, 4479-4483.

112. M. L. McKee, P. J. Milnes, J. Bath, E. Stulz, A. J. Turberfield and R. K. O'Reilly, *Angew. Chem. Int. Ed.*, 2010, **49**, 7948-7951.

113. X. Li, Z. J. Gartner, B. N. Tse and D. R. Liu, *J. Am. Chem. Soc.*, 2004, **126**, 5090-5092.

114. B. N. Tse, T. M. Snyder, Y. Shen and D. R. Liu, *J. Am. Chem. Soc.*, 2008, **130**, 15611-15626.

115. A. Shibata, H. Abe, M. Ito, Y. Kondo, S. Shimizu, K. Aikawa and Y. Ito, *Chem. Commun.*, 2009, **43**, 6586-6588.

116. A. Shibata, T. Uzawa, Y. Nakashima, M. Ito, Y. Nakano, S. Shuto, Y. Ito and H. Abe, *J. Am. Chem. Soc.*, 2013, **135**, 14172-14178.

117. H. Wu, B. T. Cisneros, C. M. Cole and N. K. Devaraj, *J. Am. Chem. Soc.*, 2014, **136**, 17942-17945.

118. K. Gorska and N. Winssinger, *Angew. Chem. Int. Ed.*, 2013, **52**, 6820-6843.

119. Z. J. Gartner, R. Grubina, C. T. Calderone and D. R. Liu, *Angew. Chem. Int. Ed.*, 2003, **42**, 1370-1375.

120. S. Nunez Pertinez, PhD, University of Birmingham, 2020.

121. A. W. Woodhouse, A. Kocaarslan, J. A. Garden, H. Mutlu, *Macromol. Rapid Commun.*, 2024, **45**, 2400260.

122. Y. Shi, X. Liu, H. Cao, F. Bie, Y. Han, P. Yan, R. Szostak, M. Szostak and C. Liu, *Org. Biomol. Chem.*, 2021, **19**, 2991-2996.

Chapter 2

Understanding the complexities of the TAMRA-thioester DNA-templated synthesis across-the-helix architecture using absorbance and fluorescence spectroscopy

---

2.1 Introduction

2.1.1 Thioester DNA-templated synthesis protective effect

Thioesters have been used as acyl transfer electrophiles in multiple DTS architectures.[1, 2] As described in Introduction 1.3.3, thioesters are a suitable electrophile alternative to the ribosome's oxoester, due to their increased reactivity towards primary amines, when developing a DTS-induced programmable molecular machine. However, progress in developing such autonomous molecular machines has been limited by the hydrolysis of the electrophilic species due to the necessary aqueous environment for DNA hybridisation. Recently, a new DTS architecture was designed to protect aminolysis thioesters from hydrolysis by utilising an abasic site across the minor groove of DNA (Figure 2-1).[3] Abasic sites are positions in a DNA sequence where a nucleobase is missing, whilst the phosphate backbone remains intact. The across-the-helix architecture involves reactive tags being tethered at internal, abasic sites within a DNA helix. The reactive tags are occupied on opposite strands within the helix and can be synthesised at various locations along the DNA sequence (Figure 2-1A). One reactive tag has a permanent position on the first DNA strand (TAMRA thioester DNA **2-1**), whereas the second reactive tag (DNA-NH_2 **2-2**)

is positioned either across the minor groove (towards the 3′ end, **−*n***) or the major groove (towards to 5′ end, **+*n***) on the complementary DNA strand. This geometry allows for the distance between the tags to be altered, ensuring proximity between the reactive species, depending on the linker length and the major/minor groove of DNA. In comparison, the end-of-helix architecture occupies both reactive tags at one end of the DNA helix (either 3′ or 5′) at a fixed distance, whereas in the across-the-helix architecture, the distance between the reactive tags can be varied depending on the positioning within the sequence. In addition, the distance can be altered by secondary interactions between the functional groups installed on the reactive tags, which is described in this work. The work published by the O'Reilly and Turberfield groups showed an aminolysis reaction in the architecture, where a nucleophilic primary amine tag attacks an electrophilic thioester tag, transferring a fluorescent group to the amine acceptor, thus producing an amide bond (Figure 2-1B).[3]

The DTS aminolysis reaction installed the fluorophore 6-carboxytetramethylrhodamine (TAMRA) to DNA through a thioester linker (**2-1**, donor) opposite a butylamine-modified DNA strand (**2-2**, acceptor) (Figure 2-1B). However, as a consequence of the aqueous conditions of the DTS buffer, competing hydrolysis of the thioester bond can occur, releasing the small molecule TAMRA into the solution.

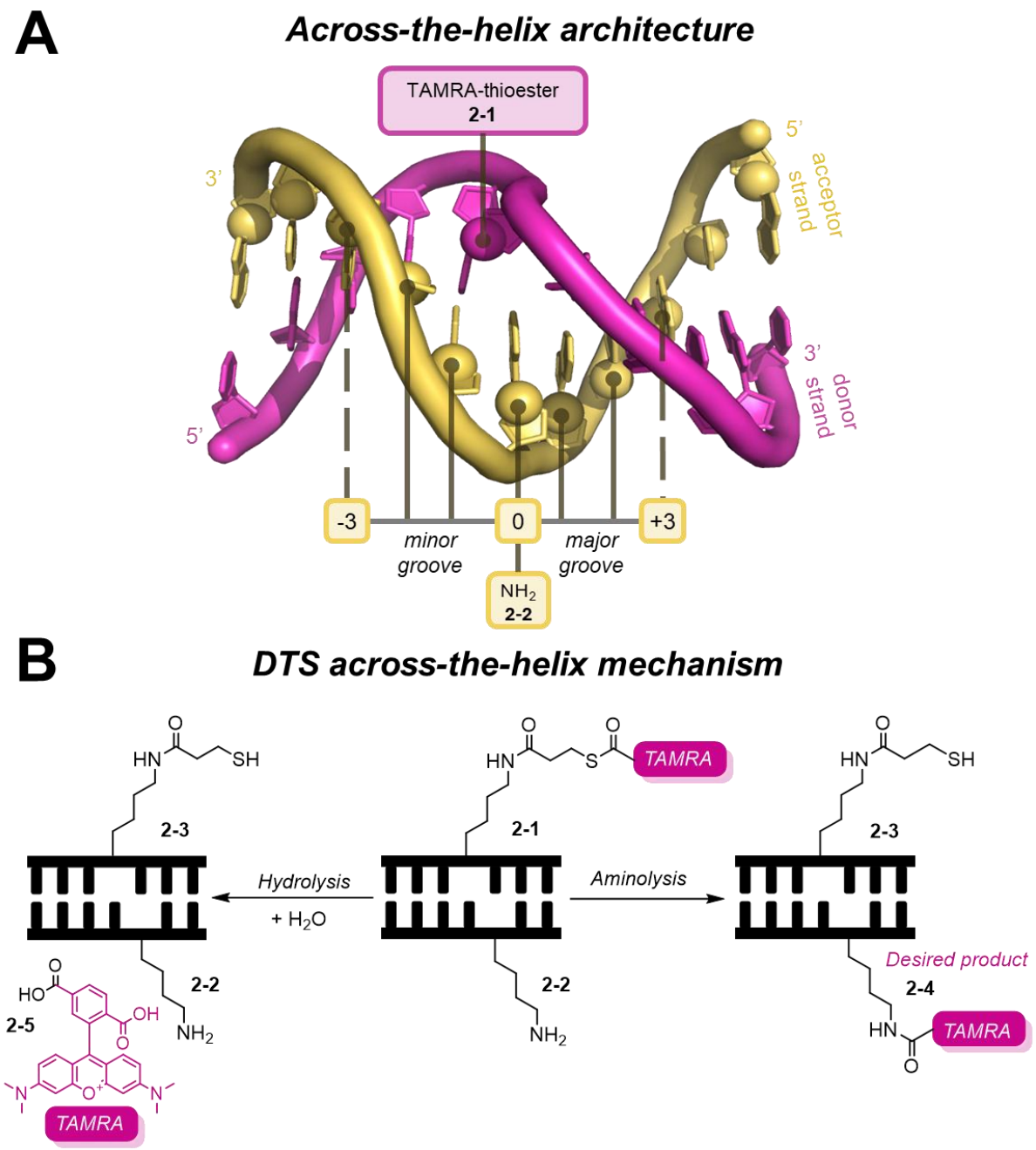


Figure 2-1: ***A****) 3D model of across-the-helix architecture depicting acceptor (TAMRA-thioester-DNA* ***2-1****, magenta) and donor (-NH_2* ***2-2****, yellow) reactive tag locations. Acceptor position can exist at various locations across the strand, either across the minor groove (−n) or major groove (+n) to TAMRA-thioester DNA* ***2-1*** *by moving the abasic site* ***2-2*** *across the strand. PyMOL Molecular Graphics System, Version 2.5.4 (PDB:1XCY).* ***B****) DTS across-the-helix mechanism: DTS aminolysis transfer reaction using modified TAMRA-thioester-DNA* ***2-1*** *and primary amine* ***2-2*** *(+1 position) modified DNA to form amide bond* ***2-4*** *(right) or hydrolysis by-product TAMRA* ***2-5*** *(left).*

Within the new design, the optimum DTS yield was observed at the +1 position across the major groove, with a decrease in yield as the distance between the acceptor and donor increased (Figure 2-2). However, when the DNA-NH_2 **2-2** was positioned across the minor groove from TAMRA-thioester DNA **2-1**, stabilisation of the thioester was observed, with both aminolysis and hydrolysis being prevented. This was determined by neither TAMRA-amide **2-4** nor free TAMRA **2-5** being observed *via* RP-HPLC.[3] The greatest degree of thioester-protection was observed at position −3 for DNA-NH_2 **2-2** across the minor groove; the shortest distance between the reactive tags (0.75 nm).[3] Because of the shorter distance, it was suggested that the planar TAMRA could easily access the open, nucleobase-absent, abasic site on the opposite DNA strand, protecting the thioester from nucleophilic attack. However, this hypothesis was not concluded. When the optimum DTS (+1) and protection (−3) positions were combined sequentially with TAMRA-thioester-DNA **2-1**, through strand displacement, a further increase in DTS yield was observed.[3]

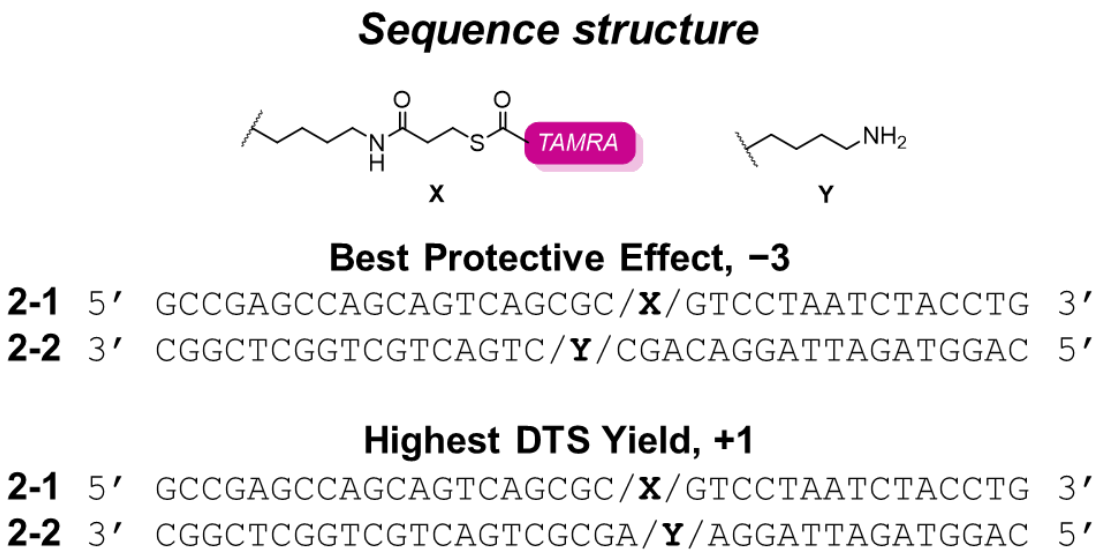


***Figure 2-2**: DNA duplex sequences for best protective effect (abasic −3) and highest DTS yield (abasic +1). **X** = TAMRA-thioester modification **2-1**, **Y** = amine modification **2-2**.*[3]

By utilising the position dependency, the across-the-helix architecture expands the potentially accessible chemistries, within an aminolysis transfer DTS reaction, by preventing the degradation of reactive tags. However, to exploit the strategy through rational design, the cause for the thioester-protection mechanism must be further investigated.

2.1.2 Absorbance and emission spectroscopy of the across-the-helix DNA-templated synthesis architecture

Rhodamine dyes are widely used within the nucleic acid community due to their accessibility in nucleotide attachment and compatibility with detection techniques (UV-Vis spectroscopy, fluorescence spectroscopy and circular dichroism spectroscopy). Common examples include Rhodamine B,[4-6] Rhodamine 6G,[7-9] tetramethylrhodamine-5-isothiocyanate (TRITC)[10, 11] and TAMRA[12-15] (Figure 2-3). By varying the modifications presented on the xanthene ring and/or carboxyphenyl group, the range of accessible molecular conformations changes, and thus, the wavelength at which light is absorbed and emitted shifts.

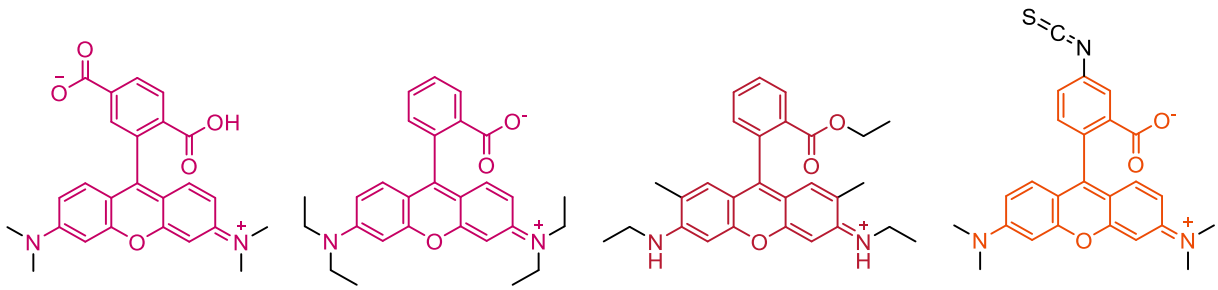


Figure 2-3: *Structure of common rhodamine dyes; TAMRA* ***2-5****, Rhodamine B, Rhodamine 6G and TRITC. The rhodamine core is highlighted with its corresponding fluorescence colour.*

As rhodamines are solvatochromic fluorophores, changes in the polarity of their environment are observable by absorbance and fluorescence spectroscopy through

spectral shifts,[16-18] *e.g.* hyperchromic (increase in absorbance), hypochromic (decrease in absorbance), bathochromic (increase in wavelength) and hypsochromic (decrease in wavelength) shifts (Figure 2-4). Such techniques have previously been used to estimate the conformation and interactions of rhodamine dyes whilst conjugated to DNA, as the hydrophobic core of DNA differs from the polar, aqueous environment surrounding the complex.[13, 19-25] However, current literature has not investigated the absorbance/fluorescence effects that the new across-the-helix architecture has on rhodamine dyes. To date, only the end-of-helix architecture[24-27], singular internal,[25, 27] and G-quadruplex[23, 28] modifications have been investigated. When compared to the literature findings of other DNA architectures, both techniques may provide insight into how the rhodamine fluorophore, TAMRA **2-5**, of TAMRA-thioester DNA **2-1** interacts to produce the position dependency and the protection effect described earlier.

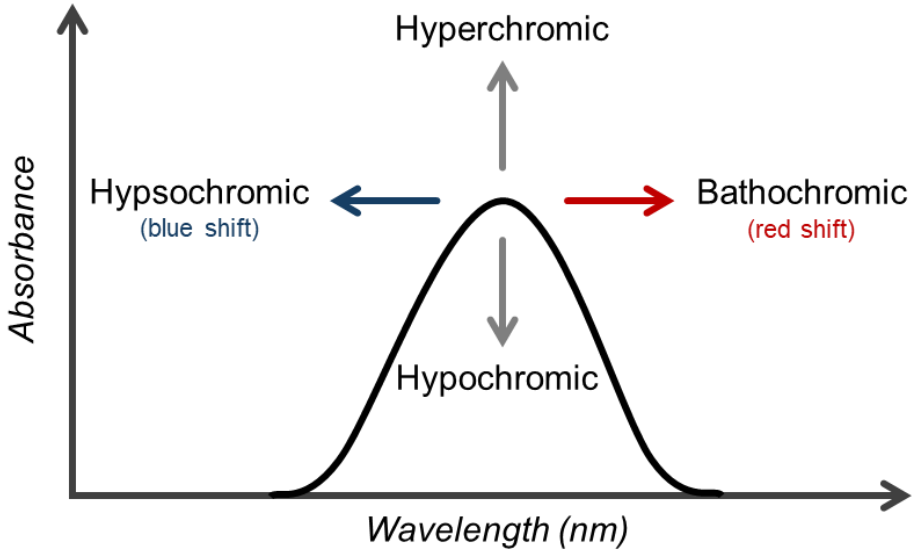


Figure 2-4: *Changes in absorption spectra of solvatochromic dyes. Hyperchromic shift (increase in absorbance), hypochromic shift (decrease in absorbance), bathochromic shift (increase in wavelength, red shift) and hypsochromic (decrease in wavelength, blue shift) shifts.*

Previous, unpublished, work completed within the group examined the changes in absorbance and emission spectra of a stable amide analogue within the across-the-helix architecture (Figure 2-5).[29] This work concluded that TAMRA, free in solution and tethered to DNA, interacted with a DNA helix containing an abasic site. Results observed a shift in absorbance and fluorescence depending on the location of the abasic site within the helices. Therefore, the distance and direction of the abasic position played a significant role in the binding capacity of the TAMRA modification (Figure 2-5B). Based upon the observations, it was hypothesised that the TAMRA moiety may interact with the smaller cytosine-surrounded abasic sites through π-π base stacking. If this theory were applied to the protection effect observed by Frommer *et al.*, the TAMRA-thioester DNA **2-1** modification would be stabilised through internal positioning in the hydrophobic DNA core, preventing thioester hydrolysis across the minor groove.[3]

However, the work completed by Cunningham replaced the unstable thioester linker with an amide, preventing unwanted hydrolysis during analysis.[29] As supported by the literature, differing the linker between the conjugated DNA and rhodamine-dyes has affected the emission and absorption spectra.[24, 30] Because Cunningham's work observed varying interactions between the amide-TAMRA modification and DNA at differing abasic positions, further investigation into how TAMRA provides a protective effect to the original thioester linker, during DTS, was conducted by UV/Vis and fluorescence spectroscopy. Any changes in the TAMRA environment would be observable by the photophysical spectroscopy techniques. In addition, a neighbouring nucleobase effect was noted by Cunningham; TAMRA fluorescence was quenched when exposed to guanine and cytosine.[29] Therefore, neighbouring abasic nucleobases

were changed in the TAMRA-thioester DNA **2-1** sequence and subject to DTS and protection effect assays to determine if a nucleobase dependency in the across-the-helix architecture was present.

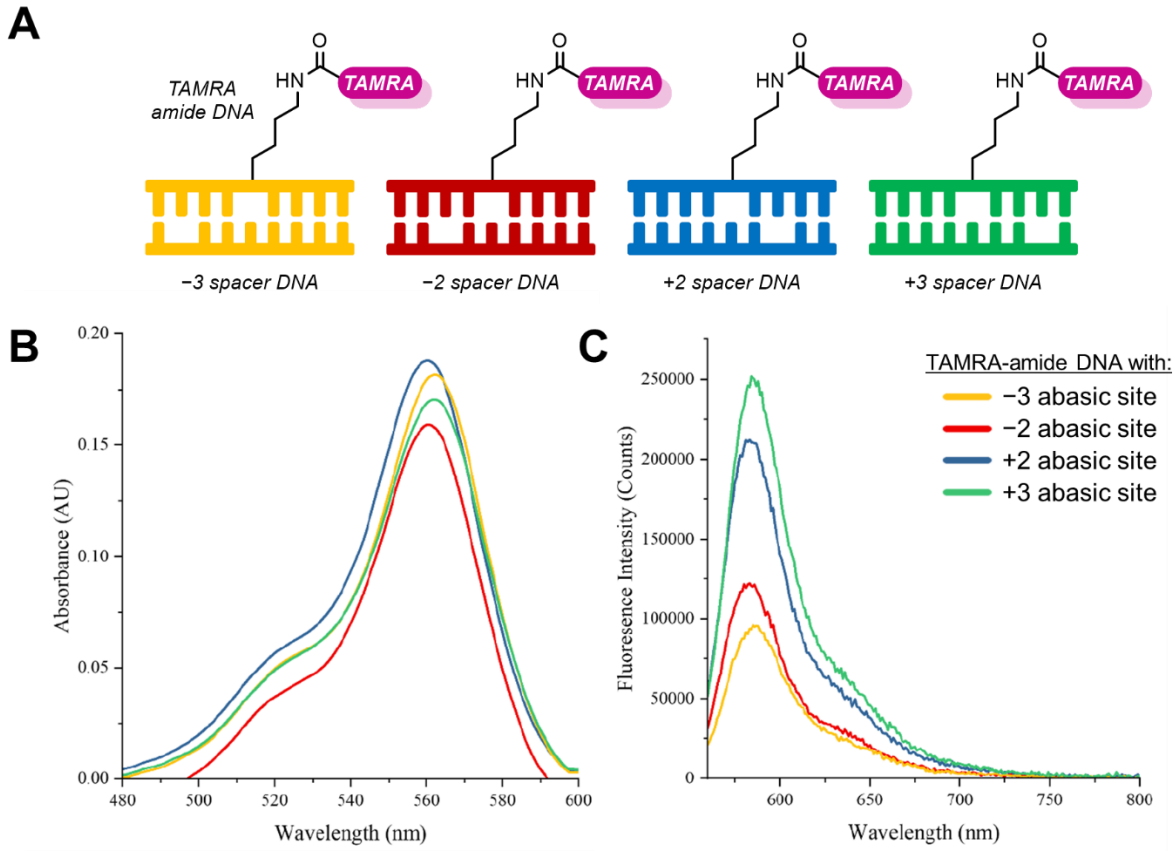


Figure 2-5: *Across-the-helix work completed by Cunningham.* ***A****) structure of TAMRA-amide DNA linkers used, at −3 (yellow), −2 (red), +2 (blue) and +3 (green) spacer sites.* ***B****) Absorbance spectra and* ***C****) fluorescence spectra (λ_{Ex} = 550 nm, bandwidth = 1) of the TAMRA-amide DNA linker (2.5 μM) with the complementary DNA strand accompanying a −2 (red), +2 (blue), −3 (yellow) or +3 (green) abasic site.*[29]

2.2 Results and Discussion

2.2.1 Understanding the TAMRA thioester DNA 2-1 position dependency

2.2.1.1 TAMRA vs TAMRA-NHS

To understand potential absorbance and fluorescence trends in the position dependency of the TAMRA-DNA architecture, similar studies had to be conducted

using the fluorescent small molecules, TAMRA **2-5** and 6-carboxy-tetramethylrhodamine *N*-succinimidyl ester (TAMRA-NHS) **2-6** (Figure 2-6). As TAMRA-NHS **2-6** contains an activated ester, it is a better candidate for comparison with our thioester-TAMRA **2-1** modification, compared to TAMRA **2-5** alone. However, there is limited literature on the spectroscopic properties of TAMRA-NHS **2-6**; therefore, both fluorophores were subject to examination.

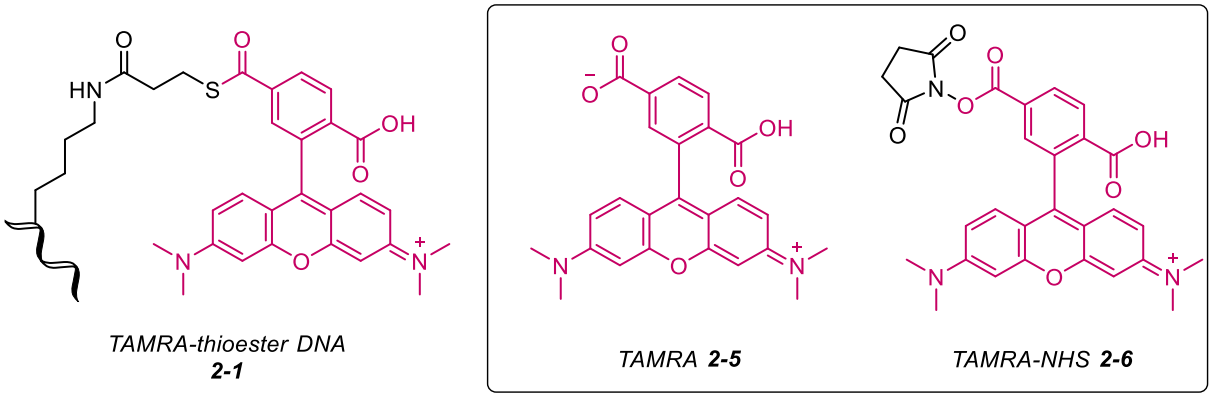


Figure 2-6: *Structure of TAMRA-thioester-DNA* ***2-1****, free TAMRA* ***2-5*** *and TAMRA-NHS* ***2-6****. TAMRA moiety is highlighted in magenta.*

By identifying the spectroscopic characteristics of the fluorophore, the relationship in the position dependency could be derived. To understand the photophysical properties of the TAMRA adapter, the fluorophores were dissolved in a range of protic and aprotic polar solvents, at 2.5 µM, to document trends in spectral shifts (Figure 2-7).

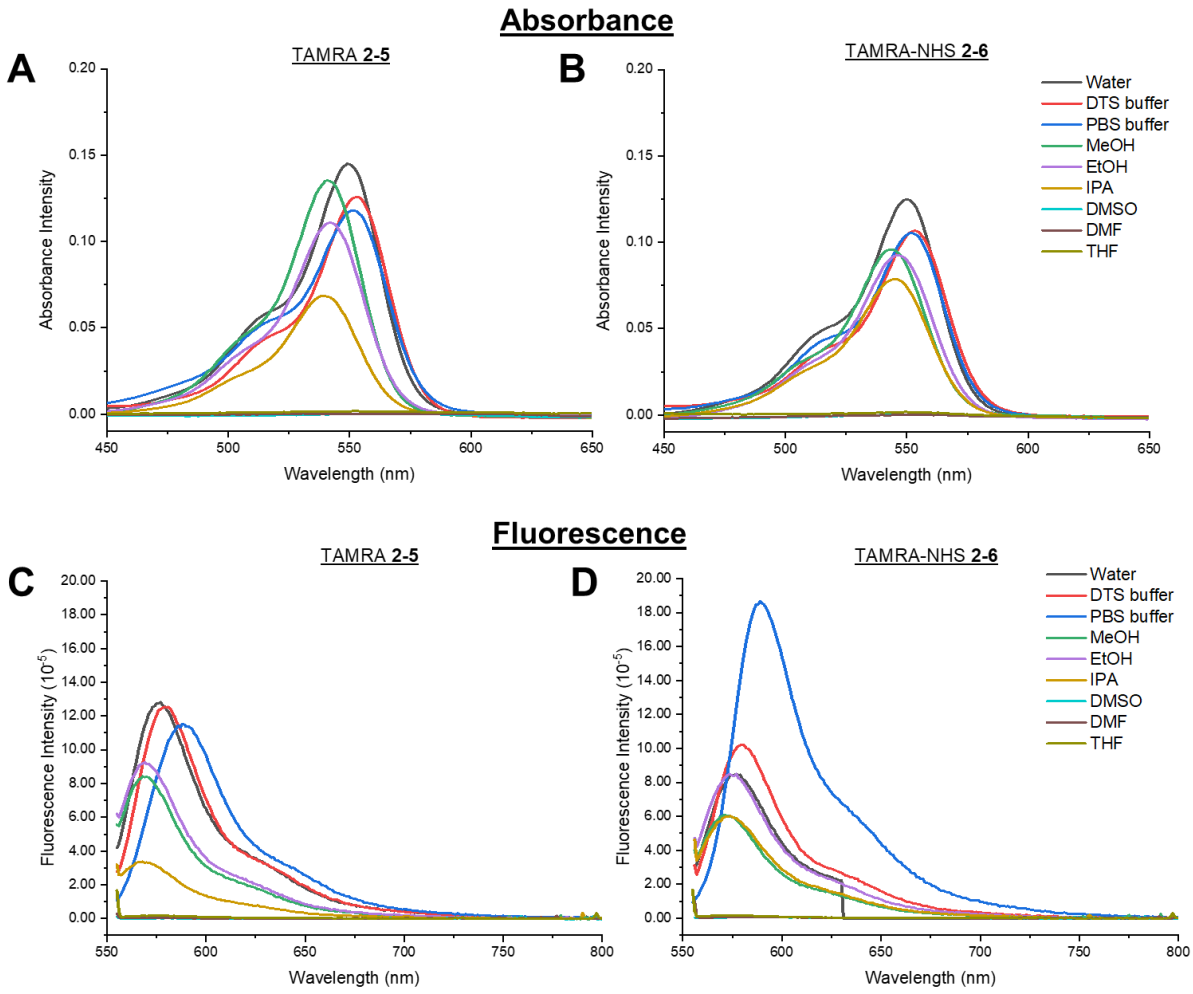


Figure 2-7: *UV-Vis spectrum (λ = 200 nm – 600 nm) of TAMRA* ***2-5*** *(**A**) and TAMRA-NHS* ***2-6*** *(**B**) fluorescence spectrum (λ_{Ex} = 550 nm, bandwidth = 1) of TAMRA* ***2-5*** *(**C**) and TAMRA-NHS* ***2-6*** *(**D**) in polar solvents (2.5 μM): water (black), DTS buffer (red, pH 11, 0.25 M NaCl, 0.25 M TAPS, 0.25 M CAPS, 0.25 M Na_3PO_4), PBS buffer (dark blue, pH 7.4, 0.1 M), MeOH (dark green), EtOH (purple), IPA (yellow), DMSO (light blue), DMF (brown) and THF (light green).*

Upon initial dilution, neither of the fluorophores were soluble in non-polar solvents, hexane and dichloromethane (CH_2Cl_2), due to the polar nature of rhodamine. Of the soluble solutions, the absorbance/emission spectrum was observed in the polar protic solutions: aqueous solutions, methanol (MeOH), ethanol (EtOH) and isopropyl alcohol (IPA) (Figure 2-7), due to hydrogen bonding with the fluorophores. In the solvents that possessed no hydrogen bonding capacity, dimethylformamide (DMF),

dimethyl sulfoxide (DMSO) and tetrahydrofuran (THF), solutions remained colourless, suggesting the quenching of TAMRA fluorescent excited states by aprotic solvents. An alternative hypothesis proposed the formation of a γ-lactone ring within the rhodamine core.[16]

Table 2-1: *Spectroscopic characteristics of TAMRA **2-5** and TAMRA-NHS **2-6**. Maximum (λ_{max}, nm), maxima differences ($\Delta\lambda_{max}$, nm) and Stokes shift (Abs λ_{max} - Em λ_{max}, nm) of absorbance and emission spectra in a range of polar solvents.*

	Absorbance (λ_{max} / nm)			**Emission** (λ_{max} / nm)			**Stokes Shift** (Abs λ_{max} - Em λ_{max} / nm)	
Solvent	**2-5**	**2-6**	**$\Delta\lambda_{max}$**	**2-5**	**2-6**	**$\Delta\lambda_{max}$**	**2-5**	**2-6**
IPA	539	546	7	568	574	6	29	28
MeOH	541	544	3	568	571	3	27	27
EtOH	542	546	4	569	577	8	27	31
Water	549	550	1	577	577	0	28	27
PBS buffer	551	552	1	588	589	1	37	37
DTS buffer	553	553	0	579	580	1	26	27

For the protic solvents, the absorbance spectra possessed a maximum band at 539 – 546 nm for organic solvents and 549 – 553 nm for aqueous solvents (Table 2-1). A smaller, secondary environment was observed at 513 nm, believed to be the formation of an excited, stacked TAMRA dimer; however, this could not be concluded without fluorescent lifetime calculations (Figure 2-7).[13] As for fluorescence, an emission band was identified between 568 – 577 nm for organic solvents and 577 – 589 nm for aqueous solvents. The spectral changes between the solvents were dictated by changes in the distance between the electronic ground state and the excited state of the TAMRA fluorophores. Known as solvatochromism, the energy absorbed by the fluorophore during excitation redistributes its internal electronic charge, altering the dipole moment and, thus, affecting how TAMRA interacts with a

polar solvent environment. For the positive solvatochromism effect displayed by both TAMRA **2-5** and TAMRA-NHS **2-6**, an increase in solvent polarity stabilises the excited state, reducing the energy gap, leading to an increase in maximum wavelength (λ_{max}) (Figure 2-8). Transition state stabilisation can be contributed by the solvation of TAMRA's carboxylate group *via* hydrogen bonding, or the reorientation of the phenyl moiety, depending on the surrounding solvent molecules.[16, 17, 31]

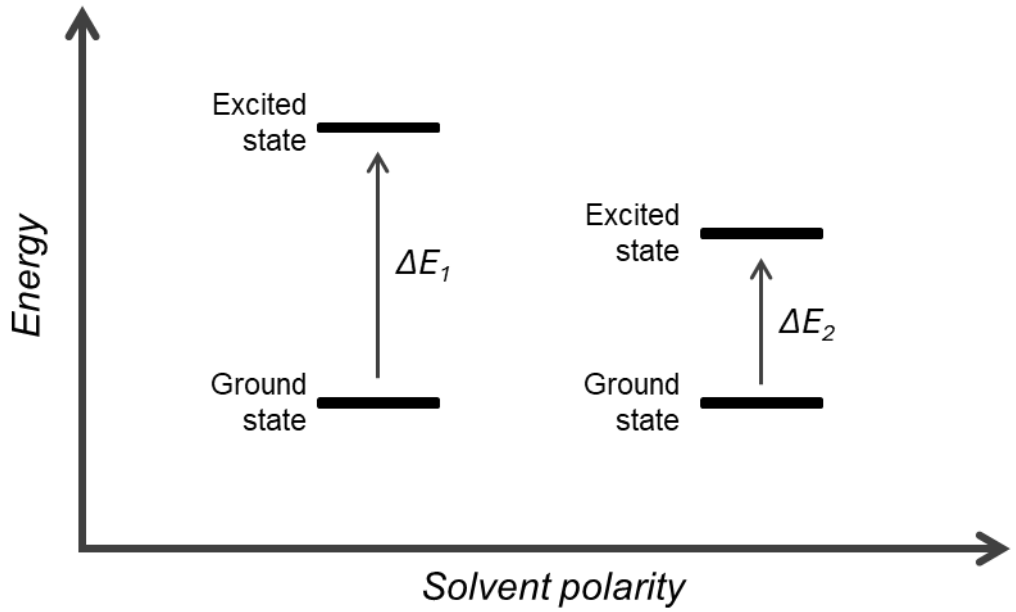


Figure 2-8: *Energy level diagram of positive solvatochromism. An increase in solvent polarity leads to the stabilisation of the excited state and, therefore, a decrease in the energy gap (ΔE).*

The highest absorption for both TAMRA **2-5** and TAMRA-NHS **2-6** was displayed in water (pH 6.5). For both molecules, the buffer conditions caused a bathochromic and hypochromic shift in comparison to water, due to the added polarity of salt-separated ions, changing the fluorophore's excited state. However, an opposite fluorescence effect was observed for TAMRA-NHS **2-6** in the aqueous solvents; an increase in fluorescence with a bathochromic shift was believed to be contributed by the buffer salts disrupting the weak intermolecular forces between the fluorophore

molecules, preventing aggregation of the dye (Figure 2-7).[32, 33] The increase in fluorescence for TAMRA-NHS **2-6** was not as significant in the DTS buffer, in comparison to the PBS buffer, likely due to hydrolysis of the ester bond at pH 11 in the DTS buffer.

When comparing protic organic solvents, a hypochromic shift and fluorescence quenching take place with the addition of methyl groups from MeOH < EtOH < IPA, reducing the polarity and solvent molecule's accessibility to the fluorophore.[16] A bathochromic shift in protic organic solvents was observed when the NHS ester of TAMRA-NHS **2-6** replaced the carboxylate of TAMRA **2-5** (Table 2-1). The carboxylate has a greater negative inductive effect (electron withdrawing) compared to the ester. Reducing the region in which the π-electron can be distributed across both the phenyl ring and the xanthene core causes an increase in the energy gap between ground and excited states, leading to spectral differences.[31, 34] The shift in wavelength is not observed in the aqueous solvents due to the increased hydrogen-bonding ability of water compared to the tested alcohols. Therefore, this leads to the protonation of the carboxylate species and reduces the negative inductive effect.

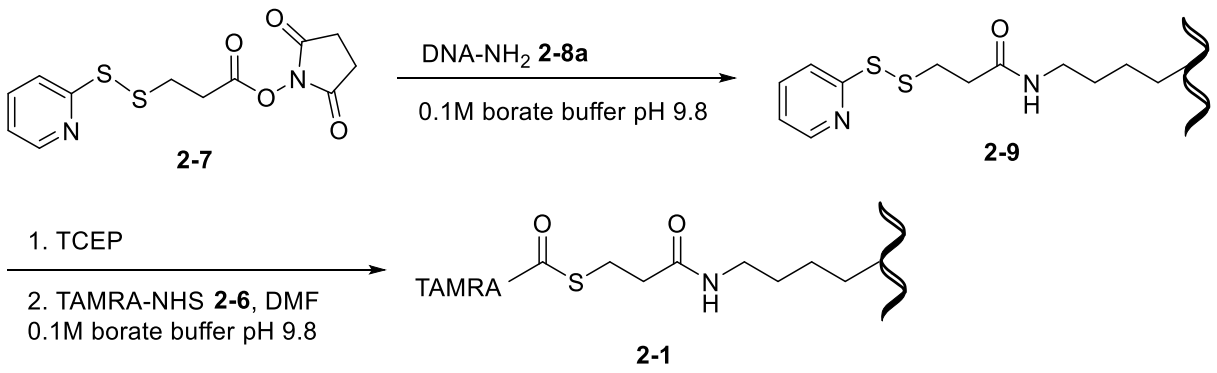
In summary, TAMRA **2-5** and TAMRA-NHS **2-6** are sensitive to their environment and display a positive solvatochromism effect with polar solvents. Thus, concluding that any increase in absorbance, fluorescence or wavelength in the future analysis of TAMRA-thioester-DNA **2-1** corresponds to TAMRA's exposure to a more polar, aqueous environment.

2.2.1.2 UV-Vis and fluorescence spectroscopy of the across-the-helix architecture

To understand the cause of the TAMRA-thioester DNA **2-1** position dependency in the across-the-helix architecture, differing abasic DNA duplexes were subject to UV-Vis and fluorescence spectroscopy, in the same conditions as before. Such techniques would provide insight into the changes in the fluorophore environment, aiding in the understanding of the binding mode between the TAMRA moiety and the DNA helix's abasic sites in the protection effect.

Following Frommer *et al.*'s synthetic procedure, the TAMRA-thioester DNA strand **2-1** was synthesised. An amine-modified DNA strand **2-8a** (See Table 2-6 for sequences) was incubated with succinimidyl 3-(2-pyridyldithio)propionate (SPDP) **2-7**, followed by a TCEP reduction and the addition of TAMRA-NHS **2-6** (Scheme 2-1).[3]

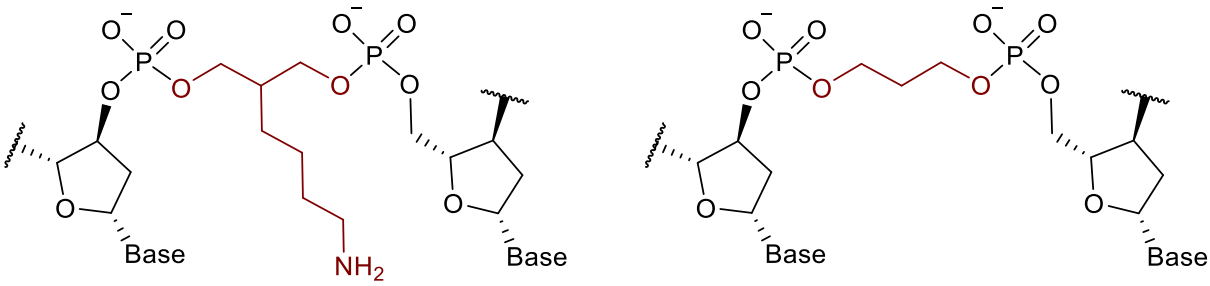
Thioester TAMRA DNA 2-1:



***Scheme 2-1**: Synthesis of TAMRA-thioester DNA **2-1**. Experimental section 2.4.2. Synthesis developed by Frommer et al.[3]*

Opposed to the original DNA-NH_2 **2-2** sequences used within the DTS aminolysis assay by Frommer *et al.* (Figure 2-1),[3] the $-NH_2$ modification was replaced with a spacer-abasic site (DNA-spacer **2-10**) (Figure 2-9). A spacer-abasic site removes the nucleobase within the DNA sequence, leaving an empty pocket. The

change prevented aminolysis from occurring during absorption and emission analysis, whilst still allowing the TAMRA-abasic interaction to be monitored. The DNA-spacer **2-10** was used by Frommer *et al.* to provide initial protection against hydrolysis, at position −3, followed by strand displacement with +1 DNA-NH_2 **2-2**, to initiate aminolysis transfer DTS.[3]



DNA-NH_2 ***2-2*** */iUniAmM/*

DNA-spacer ***2-10*** */iSpC3/*

Figure 2-9: *Structure of DNA-NH_2* ***2-2*** *used in the aminolysis transfer DTS reaction and DNA-spacer* ***2-10*** *used in the protection assay by Frommer et al.[3] DNA-spacer* ***2-10****, at varying positions, used for absorbance and emission spectroscopy.*

The interaction between TAMRA and the various abasic positions was investigated using absorbance and emission spectroscopy (Figure 2-10). TAMRA-thioester DNA **2-1** was incubated with DNA-spacer **2-10** (at various positions), its fully complementary strand (no abasic sites) or as a single strand in a PBS/$MgCl_2$ (0.1 M, pH 7.4)/0.3 M) mix for 30 minutes, prior to analysis, to ensure DNA hybridisation. Fluorescence analysis was conducted at λ_{Ex} = 550 nm to allow direct comparison with the earlier TAMRA-NHS **2-6** solvent analysis.

Position Dependency

Key	Description	Sequence **X** = /TAMRA-thioester/ **Y** = /iSpC3/	
	TAMRA thioester DNA **2-1**	5′	GCCGAGCCAGCAGTCAGCGC/**X**/GTCCTAATCTACCTG 3′
	2-1 & Complementary dsDNA	3′	CGGCTCGGTCGTCAGTCGCG A CAGGATTAGATGGAC 5′
	2-1 & +3 DNA-spacer **2-10a**	3′	CGGCTCGGTCGTCAGTCGCGACA/**Y**/GATTAGATGGAC 5′
	2-1 & +1 DNA-spacer **2-10b**	3′	CGGCTCGGTCGTCAGTCGCGA/**Y**/AGGATTAGATGGAC 5′
	2-1 & 0 DNA-spacer **2-10e**	3′	CGGCTCGGTCGTCAGTCGCG/**Y**/CAGGATTAGATGGAC 5′
	2-1 & −2 DNA-spacer **2-10f**	3′	CGGCTCGGTCGTCAGTCG/**Y**/GACAGGATTAGATGGAC 5′
	2-1 & −3 DNA-spacer **2-10h**	3′	CGGCTCGGTCGTCAGTC/**Y**/CGACAGGATTAGATGGAC 5′
	TAMRA-NHS **2-6**	–	

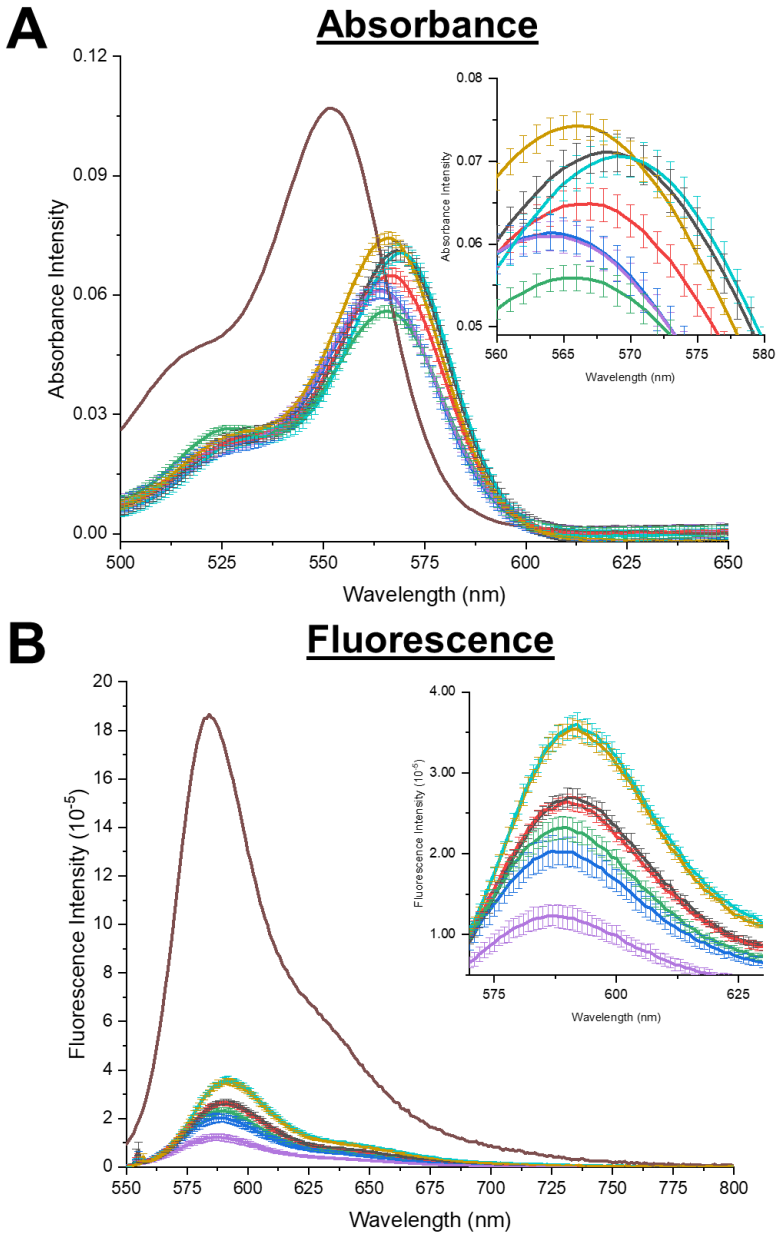


Figure 2-10: *(**A**) UV-Vis spectrum (λ = 200 nm – 600 nm) and (**B**) Fluorescence spectrum (λ_{Ex} = 550 nm, bandwidth = 1) of TAMRA-thioester DNA **2-1** (2.5 μM, 0.1 M PBS) within: alone ssDNA **2-1** (black), complementary dsDNA (red), DNA-spacer **2-10**: +3 (dark blue), +1 (green), 0 (purple), −2 (yellow), −3 (light blue) and TAMRA-NHS **2-6** alone (brown). Error bars show the standard deviation in intensity of three repeats.*

As observed in the solvent assays, when TAMRA was exposed to a polar environment, the absorbance maximum shifted to a longer wavelength. In the absorbance spectra of the abasic DNA systems, a significant bathochromic shift from 552 nm for TAMRA-NHS **2-6** in PBS buffer (0.1 M, pH 7.4) to 564 − 569 nm in the TAMRA-DNA conjugated system in PBS buffer (0.1 M, pH 7.4). Through association with the charged phosphate backbone of DNA, the polarity of the direct TAMRA environment increases, therefore causing the observed red shift. (Figure 2-10A).[13, 19, 35] The same effect was observed in the corresponding λ_{Ex} = 550 nm fluorescence spectra (Figure 2-10B), with TAMRA fluorescence shifting from 584 nm to 587 − 592 nm upon TAMRA's conjugation to the DNA strand. In addition to the red shift, a significant decrease in intensity was demonstrated for both spectral readings, with absorbance intensity decreasing from 0.101 AU to less than 0.074 AU and fluorescence intensity decreasing from 1867160 AU to less than 359784 AU for TAMRA-NHS **2-6** and the TAMRA thioester DNA **2-1** complexes. Whilst bound to the DNA architecture, the reduced mobility of the fluorophore may have contributed to the decrease in intensity.[13, 19]

In addition, the absorbance intensity was lower for the TAMRA-thioester DNA **2-1** modification (*e.g*. 0.071 AU for −3 DNA-spacer **2-10h**, Figure 2-10A, light blue) in comparison to Cunningham's TAMRA-amide modification (*e.g*. 0.175 AU for −3 DNA-spacer **2-10h,** Figure 2-5B, yellow).[29] Therefore, the additional ethyl-thioester linker also impacted the interactions between TAMRA whilst bound to DNA.

Only minor changes in the absorbance maximum of ssDNA TAMRA-thioester DNA **2-1** were observed, with a decrease from 0.071 AU to 0.065 AU upon complementary DNA hybridisation, due to π-π base-stacking interactions and reduced

flexibility (Figure 2-11A). This effect has also previously been observed for other rhodamine-DNA conjugates.[19, 36] However, there was no difference in emission spectra between the ssDNA and the full complementary dsDNA (Figure 2-11B), both displaying an intensity maximum at 591 nm. Because DNA absorbs light, the decrease in TAMRA absorbance, once conjugated to DNA, was likely contributed by the neighbouring base stacking and was not the result of a TAMRA conformation change. Therefore, any changes observed in the emission intensity of TAMRA thioester DNA **2-1** in the DTS architecture were a consequence of the abasic site.

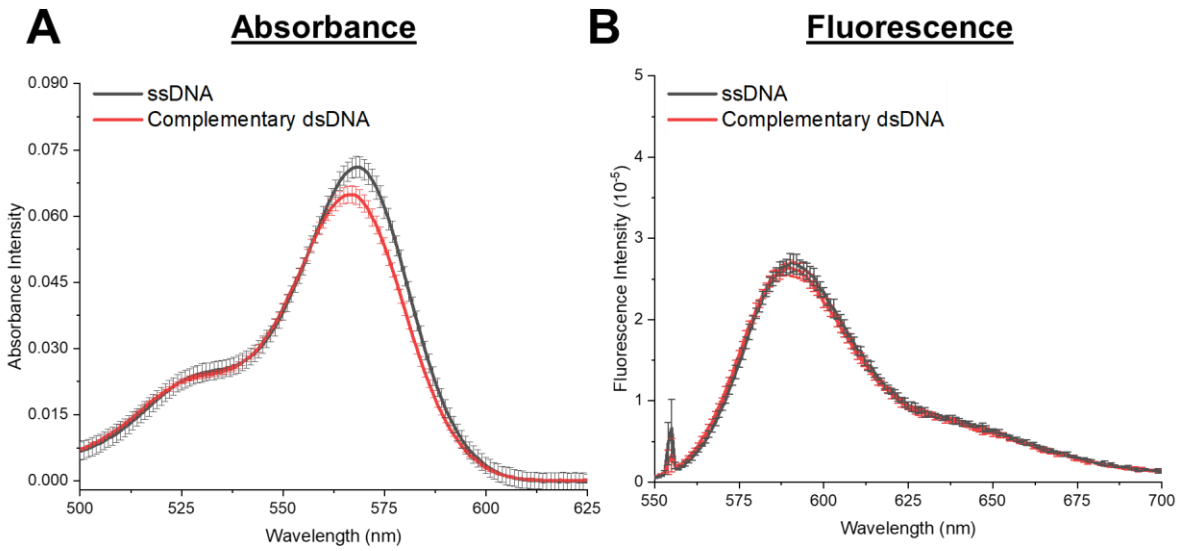


Figure 2-11: (A) *UV-Vis spectrum (λ = 200 nm – 600 nm) and* ***(B)*** *Fluorescence spectrum (λ_{Ex} = 550 nm, bandwidth = 1) of TAMRA-thioester DNA* ***2-1*** *(2.5 μM, 0.1 M PBS) as a single strand (ssDNA, black) or with the fully complementary strand as a dsDNA (red). Error bars show the standard deviation in intensity of three sample repeats.*

When analysing differences in position dependency, the abasic positions across the minor groove (−2, −3) displayed the highest absorbance intensity of the modifications (≥ 0.071 AU, Figure 2-12A). Such results suggested TAMRA's exposure to a polar environment, whilst also resembling the absorbance spectrum of ssDNA (0.071 AU). On the other hand, the positions across the major groove (+1, +3)

decreased in intensity ≤ 0.061 AU), indicating a change to a more non-polar fluorophore environment (Figure 2-12B).

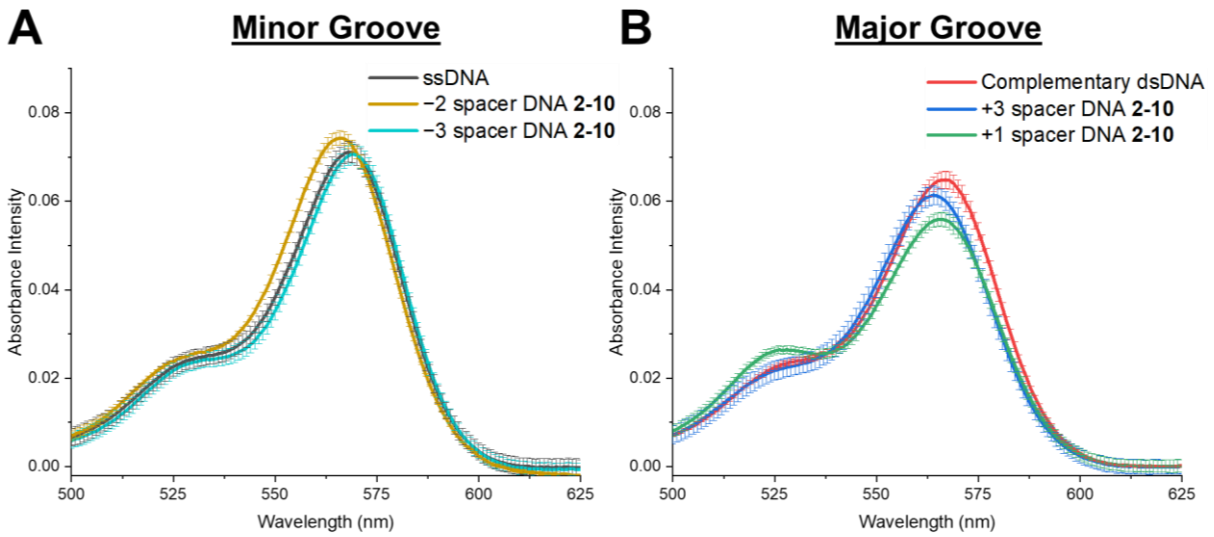


Figure 2-12: *UV-Vis spectrum (λ = 200 nm – 600 nm) of TAMRA-thioester DNA* ***2-1*** *(2.5 μM, 0.1 M PBS) with (**A**) DNA-spacer* ***2-10*** *minor groove positions −3 (light blue), −2 (yellow) and as a ssDNA (black) and (**B**) DNA-spacer* ***2-10*** *major groove positions +1 (green), +3 (blue) and with the fully complementary strand as a dsDNA (red). Error bars show the standard deviation in intensity of three sample repeats.*

Although no emission changes were observed between ssDNA and the fully complementary dsDNA, there was a significant difference in fluorescence between the major and minor groove positions (Figure 2-13). Following trends in the solvent emission spectra, the TAMRA modification exists in a more polar environment when a minor groove DNA-spacer is located on the opposite DNA strand (high emission intensity, ≥ 354195 AU). Whereas when presented with a DNA-spacer across the major groove position, TAMRA fluorescence is quenched and, hence, exists in a more non-polar environment (low emission intensity, ≤ 231988 AU).

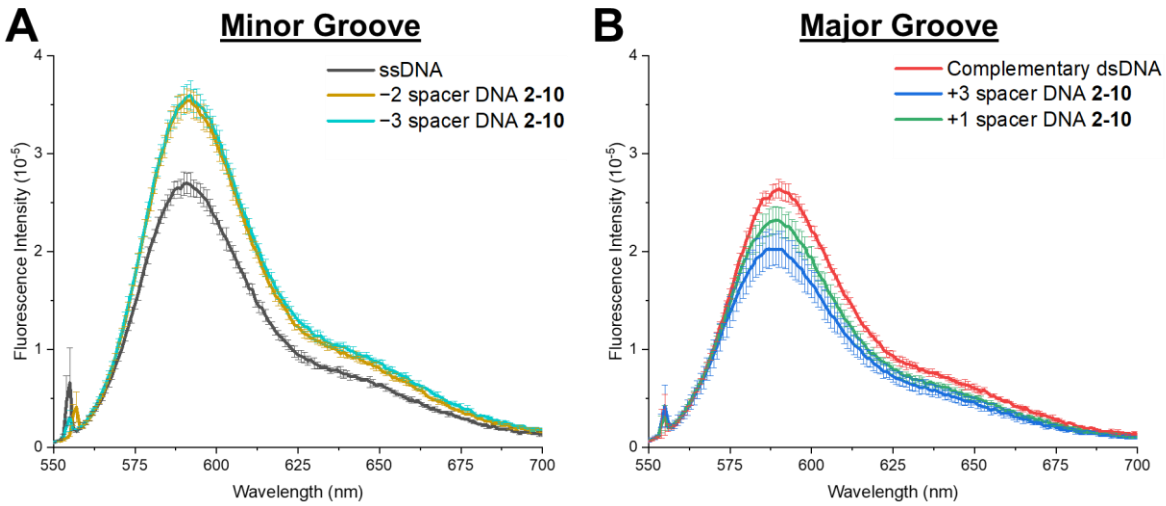


Figure 2-13: *Fluorescence spectrum (λ_{Ex} = 550 nm, bandwidth = 1) of TAMRA-thioester DNA* ***2-1*** *(2.5 μM, 0.1 M PBS) with (**A**) DNA-spacer* ***2-10*** *minor groove positions −3 (light blue), −2 (yellow) and as a ssDNA (black) and (**B**) DNA-spacer* ***2-10*** *major groove positions +1 (green), +3 (blue) and with the fully complementary strand as a dsDNA (red). Error bars show the standard deviation in intensity of three sample repeats.*

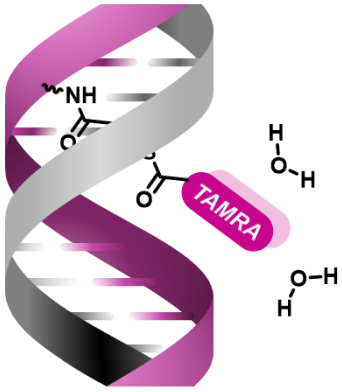
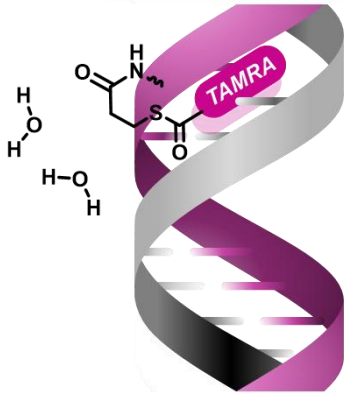
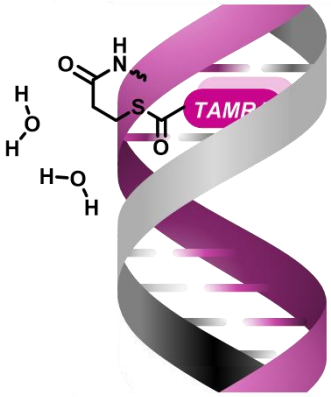
Following the spectroscopic changes of the varying abasic sites, two modes of interaction between TAMRA-thioester DNA **2-1** and the across-the-helix architecture were hypothesised. 1) At the minor groove (−2, −3) positions, TAMRA is exposed to the aqueous environment, whilst the thioester is stabilised in the hydrophobic abasic site, promoting the protection effect. In addition, the distance between the opposing modifications is at its shortest across the minor groove, increasing the probability of TAMRA-DNA interaction.[3] 2) At the major groove (+1, +3) positions, TAMRA is positioned partially within the DNA helices, exposing the thioester to the incoming nucleophile, and initiating an aminolysis reaction. Major and minor groove binding has been observed for fluorophores conjugated to DNA, and such interactions may be occurring in our system, contributing to the observed spectroscopic differences.[37-41]

Of the abasic sequences investigated, only position 0 DNA-spacer **2-10e** incorporates a double abasic site, which displayed the lowest fluorescence

(Figure 2-10B). Here, the two abasic sites are directly opposite, forming a larger abasic pocket in the DNA helices. The increased quenching demonstrated by 0 DNA-spacer **2-10e** suggests the TAMRA modification **2-1** was in a less polar environment. A possible interaction for this effect could be the TAMRA inserting into the larger abasic site, positioning into the DNA hydrophobic core, reducing water accessibility, and quenching the fluorescence.

In summary, the position dependency of abasic sites within the across-the-helix DTS architecture, documented by Frommer *et al.*,[3] was further supported by changes in absorbance and emission spectra of the tethered TAMRA fluorophore with varying abasic sites. For positions across the minor groove (−2, −3), where a thioester protection effect was observed, absorbance and emission intensity increased, suggesting TAMRA's exposure to a polar environment. Here, the thioester potentially occupies the abasic site due to the short distances between modifications, whereas the TAMRA is positioned externally to the DNA (Table 2-2, Protection effect). Alternatively, for positions across the minor groove (+1, +3), where aminolysis occurred, absorbance and emission intensities decreased, suggesting TAMRA's interaction with the DNA framework. In this structure, the thioester is exposed to nucleophilic attack, whilst the TAMRA occupies the non-polar DNA abasic site (Table 2-2, DTS aminolysis). Finally, a third structure was hypothesised for the abasic position 0, in which the TAMRA resides within the double-abasic site, demonstrated by a significant decrease in fluorescence and absorbance (Table 2-2, Double abasic site).

Table 2-2: *Table depicting the hypothesised TAMRA-thioester DNA* ***2-1*** *structures within the across-the-helix architecture, of varying DNA-spacer* ***2-10*** *sites across the major and minor groove. Hypothesised structures relating to DNA-spacer* ***2-10*** *position absorbance and emission spectra results.*

Protection effect	DTS aminolysis	Double abasic site
		
TAMRA exposed to solution Thioester protected	TAMRA partially in abasic site Thioester exposed to solution	TAMRA within two abasic sites
High absorbance	Low absorbance	Low absorbance
High fluorescence	Mid fluorescence	Low fluorescence
Minor groove (−*n*)	Major groove (+*n*)	0 abasic site

2.2.2 Modified nucleobase sequence of the across-the-helix architecture

As observed in the literature, the rhodamine-DNA complex's fluorescence varies depending on the neighbouring nucleobase sequence, with guanine residues in particular contributing to fluorescence quenching.[19, 24-26, 42-46] Therefore, to confirm the previously hypothesised TAMRA-abasic interactions were dictated by the abasic positioning, the spectroscopy studies were repeated for a set of DNA sequences in which the neighbouring nucleobases of abasic sites were adjusted. Because the −3 abasic site displayed the greatest protective effect, and the +1 abasic site demonstrated the greatest DTS aminolysis yield, these were the abasic positions

selected for initial investigation. The sequences chosen were the swapping of −3 and +1 abasic neighbouring nucleobases and the saturation of guanine (Table 2-3). As suggested by the literature, the neighbouring nucleobases impact TAMRA fluorescence more greatly than nucleobases away from the modification site.[25, 26] Therefore, only the opposite and adjacent nucleobases of the abasic site were substituted to understand the direct environment of the TAMRA modification. Swapping the −3 and +1 neighbouring nucleobases would determine whether the nucleobases influenced the different protection and aminolysis effects observed by Frommer *et al*.[3] On the other hand, guanine has been documented to display a quenching effect due to photoinduced electron transfer.[45, 47, 48]

Table 2-3: *Table displaying codes, description and corresponding sequence changes used for nucleobase sequence changes of +1 and −3 abasic sites. Neighbouring nucleobases to abasic sites are highlighted.* ***X*** *= TAMRA-thioester* ***2-1*** *site location* ***Y*** *= abasic site* ***2-10*** *location.*

Code	Description	DNA codes and sequences
+1 AGA	Original +1 sequence	**2-1a** 5′ -GC**X**G**T**CC- 3′ **2-10b** 3′ -CG**A**Y**A**GG- 5′
+1 CCC	+1 abasic sequence modified with −3 neighbouring bases	**2-1b** 5′ -GC**X**CGCC- 3′ **2-10c** 3′ -CG**C**Y**C**GG- 5′
+1 GGG	+1 abasic sequence modified with neighbouring guanine bases	**2-1c** 5′ -GC**X**GCCC- 3′ **2-10d** 3′ -CG**G**Y**G**GG- 5′
−3 CCC	Original −3 sequence	**2-1a** 5′ -CAGCGC**X**GT- 3′ **2-10h** 3′ -GT**C**Y**C**GACA- 5′
−3 AGA	−3 abasic sequence modified with +1 neighbouring bases	**2-1d** 5′ -CAT**G**TC**X**GT- 3′ **2-10i** 3′ -GT**A**Y**A**GACA- 5′
−3 GGG	−3 abasic sequence modified with neighbouring guanine bases	**2-1e** 5′ -CAT**G**TC**X**GT- 3′ **2-10j** 3′ -GT**G**Y**G**GACA- 5′

2.2.2.1 UV-Vis and fluorescence spectroscopy

−3 abasic site			+1 abasic site		
Key	Description	Sequence X = /TAMRA-thioester/ Y = /iSpC3/	Key	Description	Sequence X = /TAMRA-thioester/ Y = /iSpC3/
	−3 CCC	5′ -CAGCGCXGT- 3′ 3′ -GTCYCGACA- 5′		+1 AGA	5′ -GCXGTCC- 3′ 3′ -CGAYAGG- 5′
	−3 AGA	5′ -CATGTCXGT- 3′ 3′ -GTAYAGACA- 5′		+1 CCC	5′ -GCXCGCC- 3′ 3′ -CGCYCGG- 5′
	−3 GGG	5′ -CATGTCXGT- 3′ 3′ -GTGYGGACA- 5′		+1 GGG	5′ -GCXGGCC- 3′ 3′ -CGGYGGG- 5′

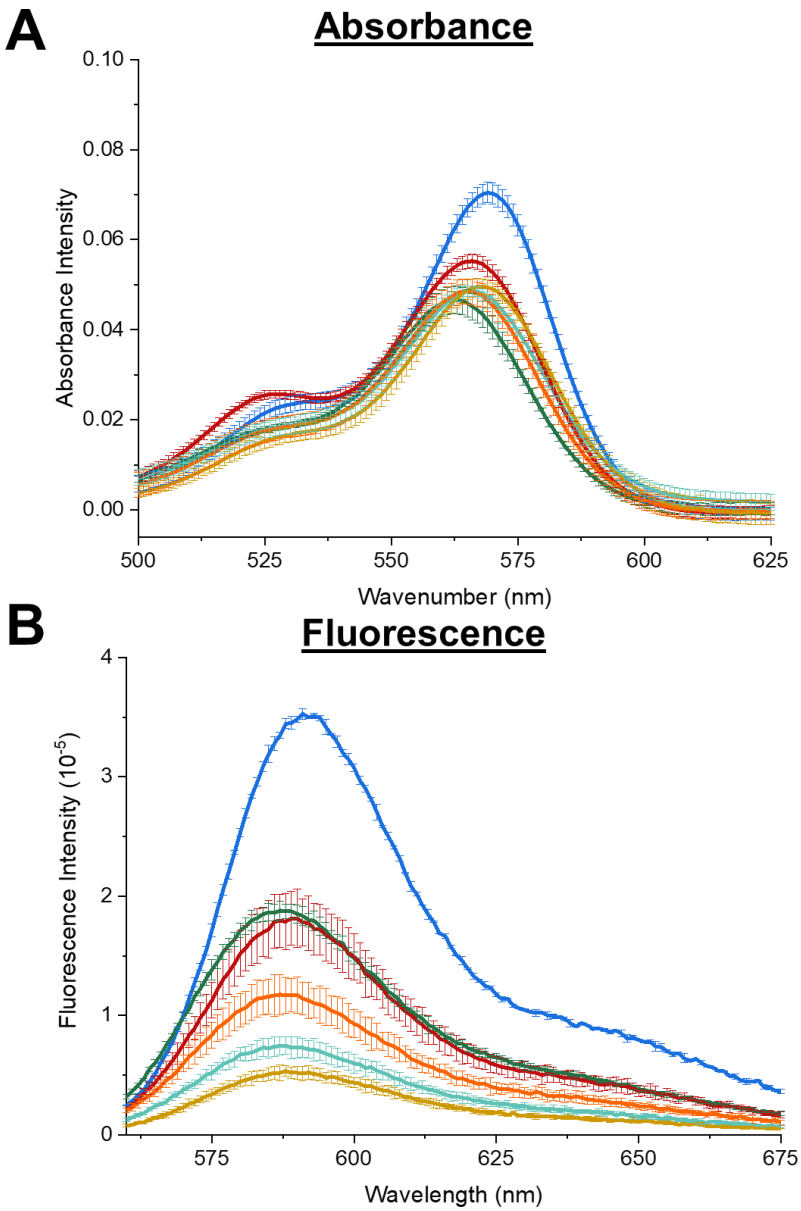


Figure 2-14: *(**A**) UV-Vis spectrum (λ = 200 nm – 600 nm) and (**B**) Fluorescence spectrum (λ_{Ex} = 550 nm, bandwidth = 1) of TAMRA-thioester DNA **2-1** (2.5 µM, 0.1 M PBS) of various sequences with various abasic DNA-spacers **2-10**: −3 CCC (original sequence, dark blue), −3 AGA (dark green), −3 GGG (turquoise), +1 AGA (original sequence, red), +1 CCC (orange) and +1 GGG (yellow). Error bars show the standard deviation in intensity of three repeats.*

For the absorbance measurements, no significant difference was observed between the various neighbouring nucleobases at position +1 (Figure 2-15A). However, a large hypochromic shift was observed from the original −3 CCC sequence to the new sequences at the same position, from 0.070 to 0.046 AU (Figure 2-15B). This suggests the −3 AGA sequence positions the TAMRA moiety in a less polar environment, similar to the original +1 AGA sequence (Figure 2-15A).

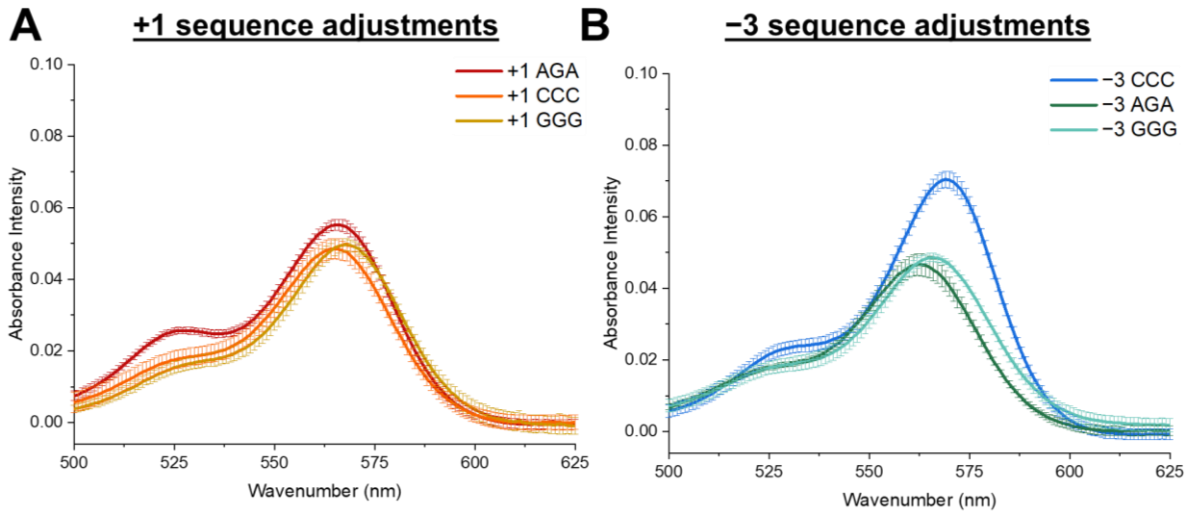


Figure 2-15: *UV-Vis spectrum (λ = 200 nm – 600 nm) of TAMRA-thioester DNA* ***2-1*** *(2.5 μM, 0.1 M PBS) with varying DNA-spacer* ***2-10*** *sequences (**A**) +1 position: +1 AGA (red), +1 CCC (orange) and +1 GGG (yellow) and (**B**) −3 position: −3 CCC (dark blue), −3 AGA (green) and −3 GGG (light blue). Error bars show the standard deviation in intensity of three repeats.*

The spectral changes of the −3 sequence adjustments may have been contributed by a TAMRA-nucleobase interaction or, following the earlier proposed hypothesis across the minor groove, the inability of the thioester to occupy the abasic site (Table 2-2). As DNA's nucleobases vary in size, cytosine is the smallest and would occupy a smaller surface area than adenine or guanine used in the −3 AGA modification.[49, 50] Therefore, the accessible space within an abasic site surrounded by cytosines may be larger than other nucleobases. To visualise this, a predefined −3 spacer site DNA model was downloaded from Protein Data Bank (PDB, 2L2V) and

adapted.[51] The sequence was modified on PyMOL (version 3.1.0) to accommodate the neighbouring nucleobases utilised in this project, in addition to a thymine-encompassed −3 abasic-spacer.[52] As thymine is the second smallest nucleobase, it was predicted that a thymine-rich environment would also provide accessibility to the unoccupied spacer. The shape of the solvent cavity within the spacer site was predicted using the apbs_gui plugin (Figure 2-16).[53] Thereby, the molecular model predicted the accessibility of the spacer site within various sequences.

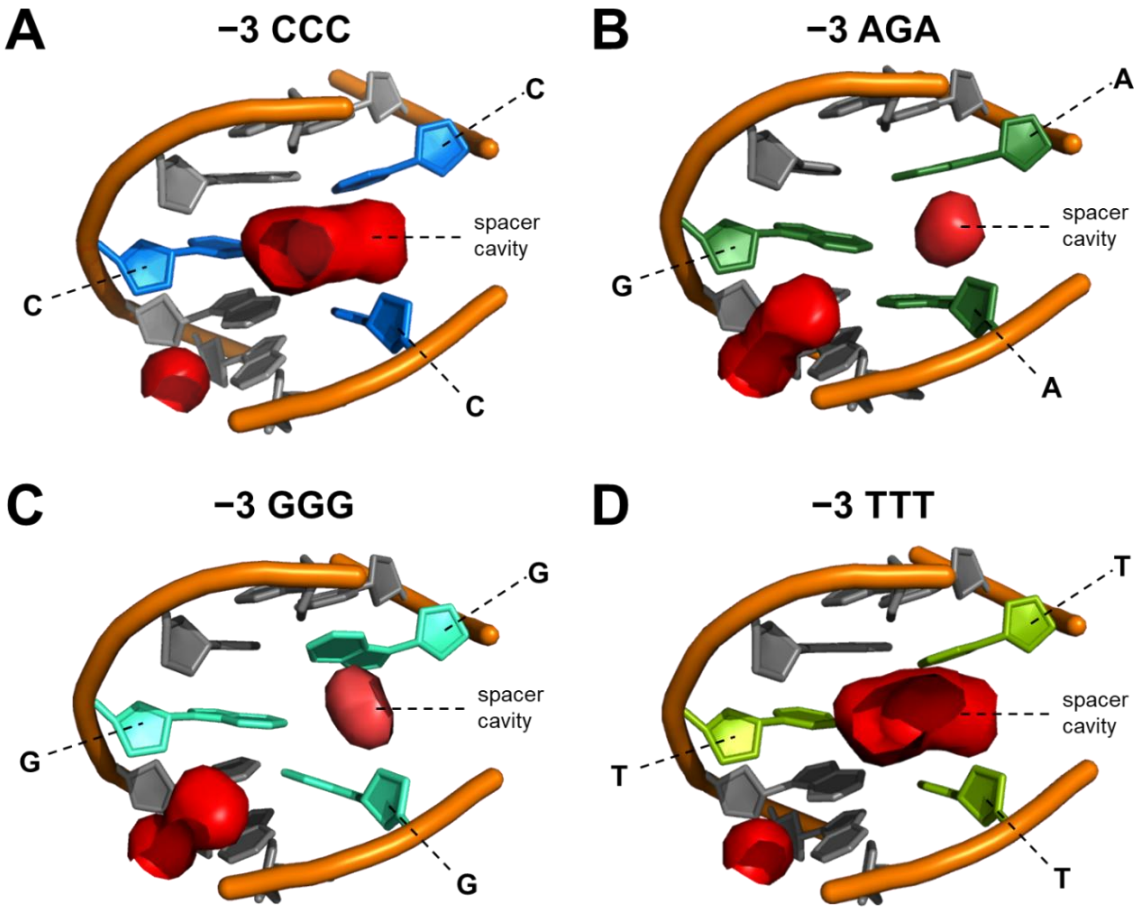


***Figure 2-16**: Predicted cavity of DNA-spacer **2-10** for various surrounding nucleobases at a −3 abasic site; **A**) −3 CCC (blue) **B**) −3 AGA (dark green) **C**) −3 GGG (light blue) **D**) −3 TTT (light green). Neighbouring nucleobases to the spacer site are highlighted in the corresponding colour. Unoccupied spacer cavity highlighted as a red spheroid. Modelled using PyMOL version 3.1.0, PDB: 2L2V, predicted by apbs_gui input (APBS Electrostatics).*

The predicted spacer cavity size varied according to whether purine or pyrimidine nucleobases surrounded the abasic site (Figure 2-16). For the purine nucleobases, a smaller cavity was observed (Figure 2-16B & C), as opposed to the larger cavity presented within the pyrimidine nucleobases (Figure 2-16A & D). The models suggest that the TAMRA thioester DNA **2-1** modification had increased accessibility to a pyrimidine-surrounded DNA-spacer **2-10** pocket on the opposing strand, potentially aiding in the protection of the thioester in −3 CCC. Following the accessibility hypothesis and the molecular models' prediction of a thymine-surrounded spacer's accessibility, the −3 TTT sequence adjustment may protect the thioester in the system.

On the other hand, the absorbance and emission spectra of +1 CCC (0.048 AU absorbance intensity, 117545 AU fluorescence intensity) did not resemble the spectra of −3 CCC (0.07 AU absorbance intensity, 352585 AU fluorescence intensity) (Figure 2-17). This suggests that the TAMRA thioester DNA **2-1** modification was not adaptable to a polar environment by adjusting the nucleobase sequence at the +1 abasic position. Therefore, proposing the protective effect would not be observed with +1 CCC.

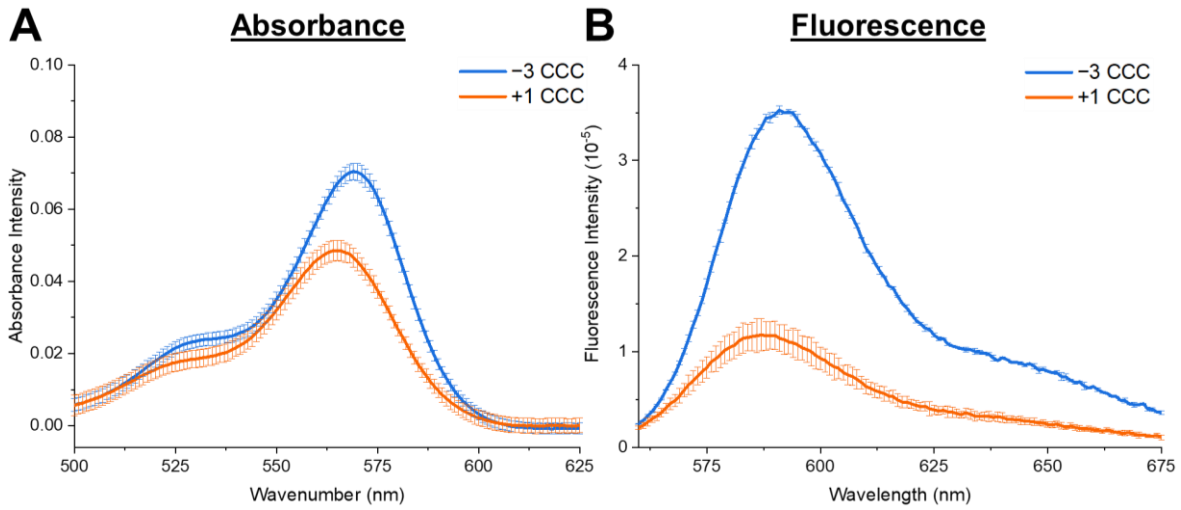


Figure 2-17: *(**A**) UV-Vis spectrum (λ = 200 nm – 600 nm) and (**B**) Fluorescence spectrum (λ_{Ex} = 550 nm, bandwidth = 1) of TAMRA-thioester DNA **2-1** (2.5 μM, 0.1 M PBS) with a cytosine environment at various abasic DNA-spacers **2-10**: −3 CCC (dark blue) and +1 CCC (orange). Error bars show the standard deviation in intensity of three repeats.*

Following literature trends, an expected decrease in fluorescence intensity was presented by both +1 GGG and −3 GGG, an attribute of the concentrated guanine environment (Figure 2-18).[19, 42, 47] This phenomenon arises from a photoinduced electron transfer between rhodamine (acceptor) and guanine (donor), when in proximity through xanthene-guanine stacking, preventing photon emission, and consequently, causing fluorescence quenching.[45, 47, 48, 54]

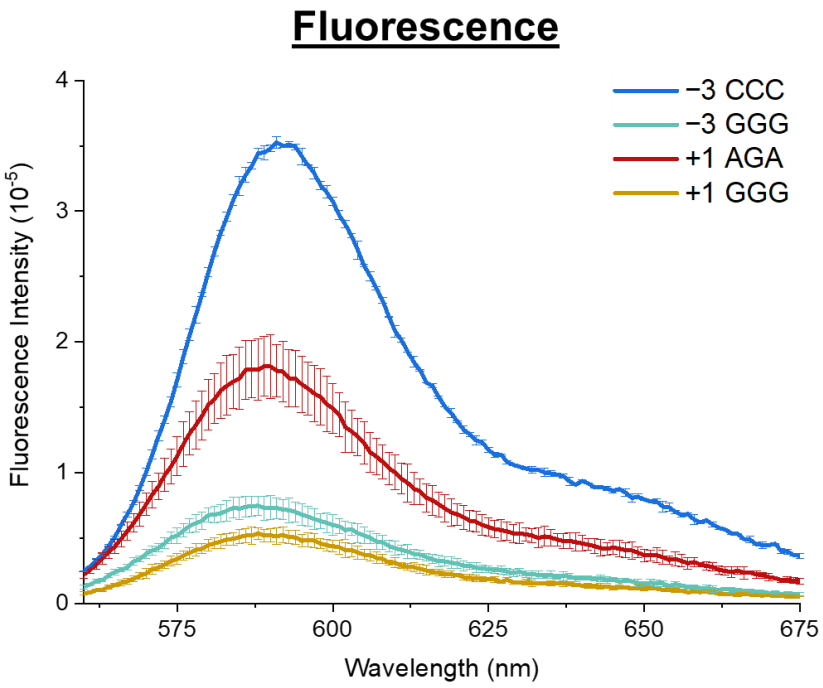


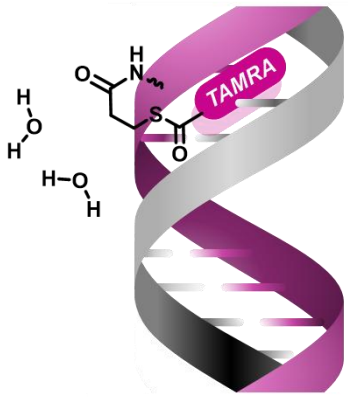
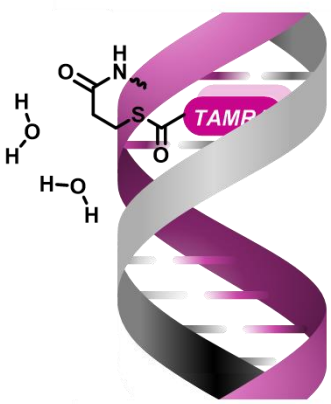
Figure 2-18: *Fluorescence spectrum (λ_{Ex} = 550 nm, bandwidth = 1) of TAMRA-thioester DNA* ***2-1*** *(2.5 μM, 0.1 M PBS) with various abasic DNA-spacers* ***2-10****: −3 CCC (original sequence, dark blue), −3 GGG (light blue), +1 AGA (original sequence, red) and +1 GGG (yellow). Error bars show the standard deviation in intensity of three repeats.*

In addition, the decrease in fluorescence intensity displayed by +1 CCC may have been contributed by the new opposing Gs in the TAMRA-thioester DNA **2-1** sequence (Figure 2-17). Supporting this theory, the literature states that an opposite-guanine fluorescence quenching effect can be observed within an end-of-helix architecture.[19, 24] However, −3 CCC occupies guanine on the opposing abasic strand but displays the highest fluorescence intensity of the sequences at 352585 AU. This observation may be a characteristic of the protection effect, in accordance with the TAMRA-DNA interaction suggested earlier in this work (Table 2-2).

The literature suggests that for the structurally similar tetramethylrhodamine, the fluorophore exists in three states when bound to DNA: A) exposure to solvent, B) exposure to the DNA environment and C) quenched by guanine.[42] Reviewing both DNA-TAMRA absorbance/fluorescence, the same hypotheses can be made for the

TAMRA thioester DNA **2-1** system. Here, the −3 CCC displayed a high intensity (state A), +1 AGA, +1 CCC and −3 AGA displayed a low intensity (state B), and both +1 GGG and −3 GGG demonstrated guanine quenching (state C). The proposed states resemble the hypothesised structures in Table 2-2, however, with nucleobase sequence as an additional parameter. Therefore, the previous hypothesised structures were revised to accommodate the adjustments (Table 2-4).

Table 2-4: *Table depicting the revised hypothesised TAMRA-thioester DNA* ***2-1*** *structures within the across-the-helix architecture.*

Protection effect	**DTS aminolysis**	**Fluorescence Quenching**
		
TAMRA exposed to solution Thioester protected State A High absorbance High fluorescence Minor groove (−*n*) with pyrimidine bases	TAMRA partially in abasic site Thioester exposed to solution State B Low absorbance Mid fluorescence Major groove (+*n*) and minor groove (−n) with purine bases	Photoinduced electron transfer/two abasic sites State C Low absorbance Low fluorescence Guanine-rich environment and double-abasic site

Furthermore, because the nucleobase sequence had a significant effect on emission and absorbance spectra at position −3, it was hypothesised that the minor groove protection effect would not be as effective for the new −3 sequences when

applied to a protection assay. If applicable, the nucleobase sequence would introduce a new parameter into determining the effectiveness of the protection assay observed by Frommer *et al*.[3] Therefore, the sequences were applied in a protection assay and a DTS assay to determine if the nucleobases affected aminolysis yields.

2.2.2.2 Protection assay and DNA-templated synthesis assay

To measure the DTS reactivity and protection ability of the sequence changes in the across-the-helix architecture, the protocols designed by Frommer *et al.* were followed (Figure 2-1B).[3] The DTS assay measures the formation of desirable amide **2-4** between the TAMRA thioester DNA **2-1** (1 μM) and the DNA-NH_2 **2-2** (2 μM) by aminolysis (Figure 2-1B). The process was monitored by RP-HPLC *via* the fluorescence channel (λ_{ex}: 550 nm, λ_{em}: 580 nm), following the corresponding signal change in the TAMRA environment (thioester to amide) (Figure 2-19). The product yield was quantified by integrating the corresponding RP-HPLC signals and compared to a predetermined calibration curve of the synthesised amide **2-4** and the hydrolysis by-product TAMRA **2-5**. The single strands of DNA-NH_2 **2-2** and TAMRA thioester DNA **2-1** were incubated together to induce hybridisation and injected periodically over 12 hours into the RP-HPLC. The reactions took place at 5 ºC and in a pH 11 buffer, named DTS buffer (see Experimental 2.4.1 for the formulation), to ensure sufficient deprotonation of the amine to produce the $-NH_2$ nucleophile.

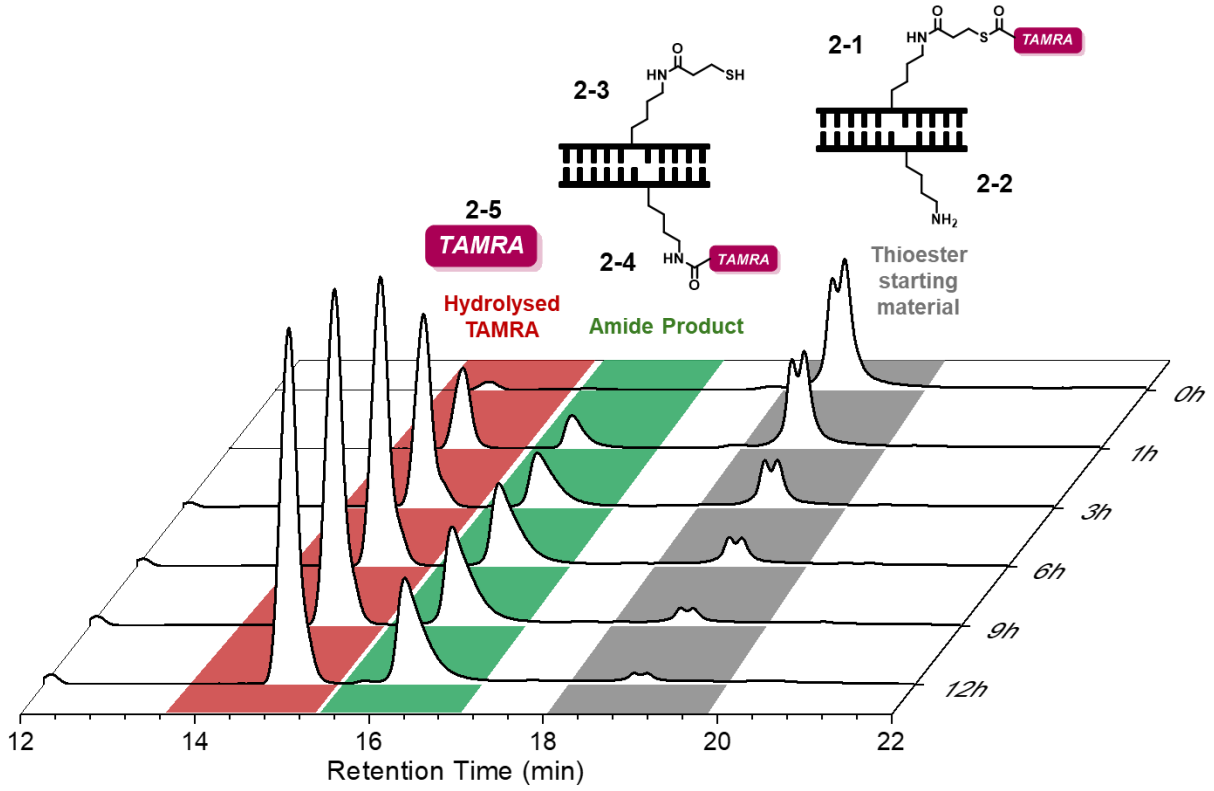


Figure 2-19: *Example RP-HPLC waterfall plot of TAMRA thioester DNA **2-1** (1 μM) and DNA-NH_2 **2-2** +1 AGA (2 μM) DTS reaction in DTS buffer, pH 11.0 (15 μL), over the course of 12 hours. Fluorescence channel: (λ_{ex} 550 nm, λ_{em} 580 nm). 2 μL injection volume. Red = free TAMRA **2-5** (hydrolysis of **2-1**). Green = successful aminolysis (amide product **2-4** with spent adapter **2-3**). Grey = starting material (TAMRA thioester DNA **2-1** with DNA-NH_2 +1 AGA **2-2**).*

In the basic conditions required for aminolysis to occur, the thioester of TAMRA thioester DNA **2-1** is prone to hydrolysis. However, as documented by Frommer *et al*., a thioester stabilisation effect occurs across the minor groove of DNA within the TAMRA thioester DNA **2-1** across-the-helix system.[3] Adapting the DTS assay, the TAMRA thioester DNA **2-1** was incubated with the DNA-spacer **2-10** strand used in the absorbance/fluorescence studies, rather than DNA-NH_2 **2-2**. In doing so, the thioester protection effect can be monitored independently from the aminolysis reaction. The protection assay measures the degree of hydrolysis, through the production of free TAMRA **2-5**, in which a low percentage of free TAMRA **2-5** corresponds to successful thioester protection (Figure 2-20). As for the DTS assay, the degree of hydrolysis was

monitored by TAMRA **2-5** RP-HPLC signal intensity, over 24 hours, and calculated using a calibration curve.

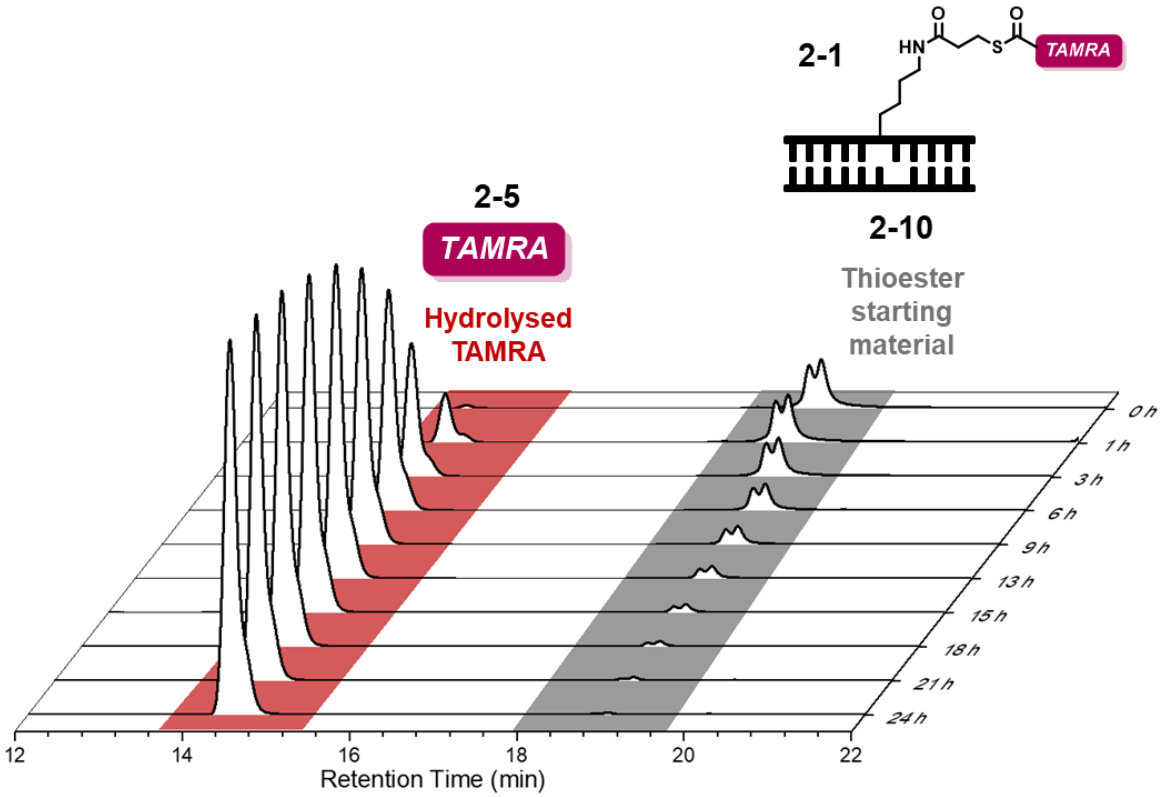


Figure 2-20: *Example RP-HPLC waterfall plot of TAMRA thioester DNA **2-1** (1 μM) and DNA-spacer **2-10** +1 AGA (2 μM) protection assay in DTS buffer, pH 11.0 (15 μL), over the course of 24 hours. Fluorescence channel: (λ_{ex} 550 nm, λ_{em} 580 nm). 2 μL injection volume. Red = free TAMRA **2-5** (hydrolysis of **2-1**). Grey = starting material (TAMRA thioester DNA **2-1** with DNA-spacer +1 AGA **2-10**).*

For the protection assay, no significant difference was observed in thioester protection of TAMRA thioester DNA **2-1** at the various sequences surrounding spacer position +1 **2-10** (Figure 2-21B). As predicted by the absorbance and emission spectra intensities of the DNA-spacer **2-10** complexes (Figure 2-14), the neighbouring nucleobase sequence did not improve the protection effect of the major groove position. However, when applied to the DTS assay, a dramatic decrease in TAMRA amide **2-3** yield was observed, from 50% **2-4** yield to less than 12%, suggesting that the nucleobase sequence affects the degree of TAMRA thioester DNA **2-1** DTS

aminolysis (Figure 2-21C). Therefore, the protective effect cannot be combined with successful +1 abasic DTS aminolysis by varying the nucleobase sequence surrounding the abasic site. Herein, the previously proposed TAMRA thioester DNA **2-1** DTS structural interaction did not apply to all sequences of NH_2-bound +1 abasic sites **2-2** (Table 2-4, DTS aminolysis).

Interestingly, the corresponding DTS hydrolysis data did not reflect the TAMRA **2-5** formation observed in the protection assay (Figure 2-21D vs Figure 2-21B) for the +1 CCC sequence, at 21% and 45%, respectively, after 12 hours. This was expected of +1 AGA and +1 GGG due to the formation of TAMRA amide product **2-4**, but not for +1 CCC where no aminolysis occurred (0%). Therefore, a thioester protective effect was possible at the +1 abasic position when a primary amine modification was installed and the abasic position was surrounded by cytosines. It was hypothesised that the primary amine also interacts with the DNA helices, as does the TAMRA moiety, potentially aiding in thioester protection of TAMRA thioester DNA **2-1**. Although not possible due to potential side reactions, if the absorbance and fluorescence were to be measured, it was hypothesised that +1 CCC DNA-NH_2 **2-2** would have a high intensity, resembling the proposed TAMRA-DNA interactions of the protection assay (Table 2-4).

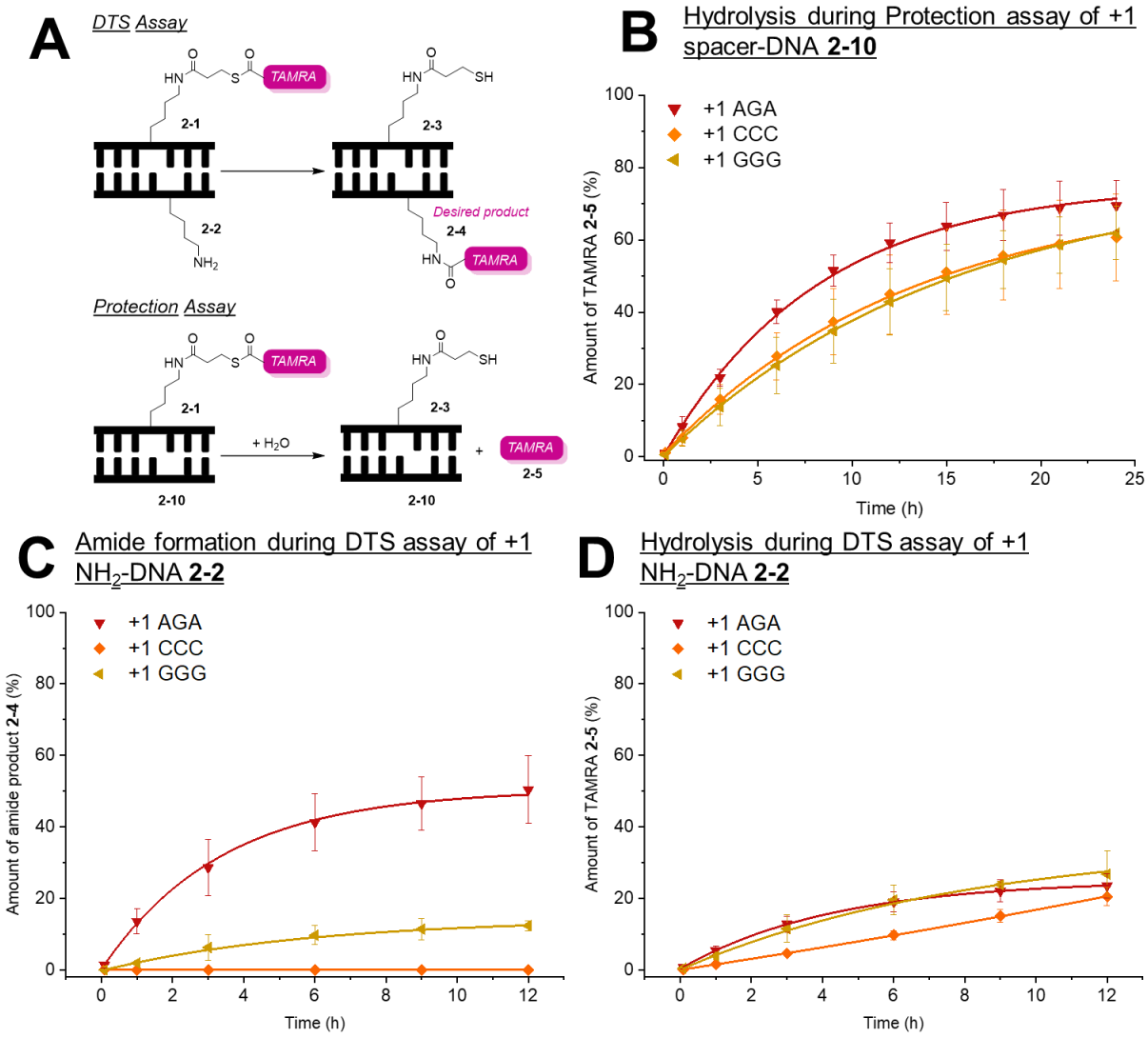


Figure 2-21: *Monitored by RP-HPLC fluorescence channel (λ_{ex} 550 nm, λ_{em} 580 nm); varying nucleobase sequences of TAMRA thioester DNA **2-1** with various abasic site +1 sequences (+1 AGA – red triangle, +1 CCC orange diamond, +1 GGG yellow triangle), in DTS buffer pH 11, applied in **A**) a Protection assay (DNA-spacer **2-10**): hydrolysed TAMRA **2-5** formation (%) over 24 hours. And DTS assay (DNA-NH_2 **2-2**) **B**) TAMRA amide **2-4** formation (%) over 12 hours and **C**) hydrolysed TAMRA **2-5** formation (%) over 12 hours. Error bars show the standard deviation in the amount (%) of three repeats.*

On the other hand, a significant decrease in thioester protection of TAMRA thioester DNA **2-1** was observed when the sequence was adjusted to AGA or GGG (57% and 44% hydrolysis after 12 hours) compared to the original CCC (14% hydrolysis) neighbouring bases of the −3 abasic site (Figure 2-22B). As predicted by the change in absorbance/emission results between sequence adjustments

(Figure 2-14), a higher absorbance and emission intensity correlated to a higher protection effect observed in the protection assay. As previously mentioned, the decrease in protection of the thioester when the −3 sequence was modified from CCC to AGA or GGG may have been a steric contribution. The small cytosine in the original sequence may increase the accessibility of the modification into the abasic site, as opposed to the larger purine nucleobases (Figure 2-16). In doing so, the thioester can occupy the empty space, allowing for the protective effect to occur.

The contrasting effects in the protection of +1 and −3 with nucleobase adjustments demonstrate that the position dependency documented by Frommer *et al.*, and supported earlier in this chapter, still existed when the DNA sequence was adjusted.[3] However, the nucleobase sequence is an additional parameter affecting the system design and would need to be considered when designing an across-the-helix mechanism. Therefore, modifying the previously proposed TAMRA-DNA interaction hypothesis further (Table 2-4, protection effect): a high absorbance/emission is not minor groove dependent, but abasic modification, abasic position and nucleobase interdependent.

Unlike the +1 amine position, the sequence did not impact the DTS aminolysis yield of the −3 abasic position, and the yields were limited to 6% after 12 hours (Figure 2-22C). Supporting previous claims at the +1 position, the neighbouring nucleobase adjustments do not allow for a combined protective effect and successful aminolysis.

Similar to the +1 position, the primary amine modification provided an additional level of thioester protection, as opposed to the spacer site, for the −3 AGA sequence, but not for −3 GGG (Figure 2-22D). This may be a consequence of secondary

interaction between the modification and the concentrated guanine environment, as observed in the absorbance and fluorescence analysis (Figure 2-14), in addition to the smaller accessible cavity of GGG (Figure 2-16C).

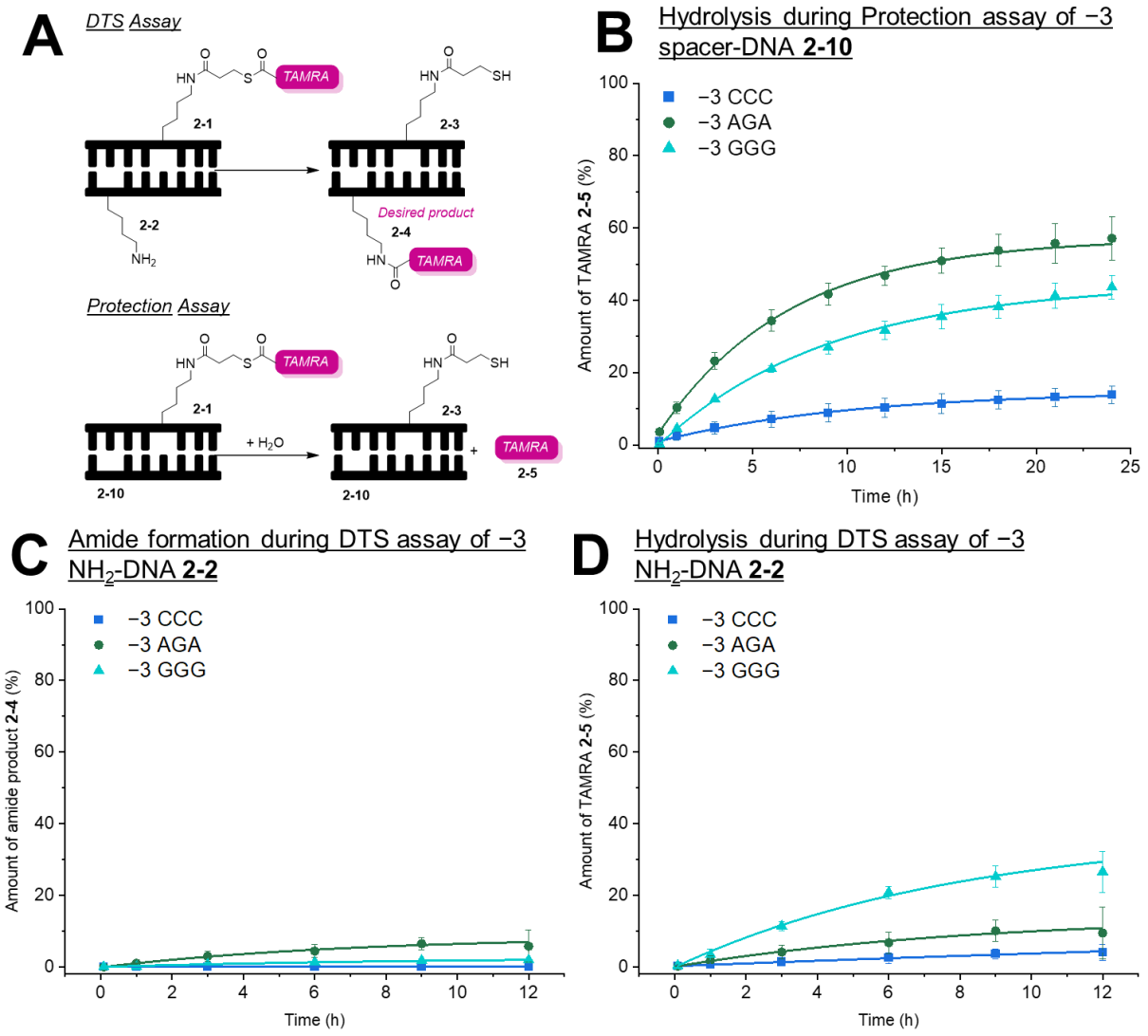


Figure 2-22: *Monitored by RP-HPLC fluorescence channel (λ_{ex} 550 nm, λ_{em} 580 nm); varying nucleobase sequences of TAMRA thioester DNA **2-1** with various abasic site −3 sequences (−3 CCC dark blue square, −3 AGA green circle, −3 GGG light blue triangle), in DTS buffer pH 11, applied in a **A**) Protection assay (DNA-spacer **2-10**): hydrolysed TAMRA **2-5** formation (%) over 24 hours. And DTS assay (DNA-NH_2 **2-2**) **B**) TAMRA amide **2-4** formation (%) over 12 hours and **C**) hydrolysed TAMRA **2-5** formation (%) over 12 hours. Error bars show the standard deviation in the amount (%) of three repeats.*

Overall, the protection assay and DTS results of the −3 and +1 abasic sites highlighted that both the abasic position and the surrounding nucleobase sequence significantly impact the reaction yields, interdependently.

2.2.2.3 Additional sequence variants

Following the predicted spacer-size molecular models (Figure 2-16), the alternative pyrimidine, thymine, was investigated surrounding the −3 abasic position. In addition to this, another abasic site across the minor groove, −2, was also subject to investigation in sequence dependency. In the original sequence, this position occupies a concentrated guanine environment surrounding the abasic site; however, it also displays a protection effect. Already contradicting the sequence dependency observed for the −3 abasic position, the guanine nucleobases were replaced by cytosines, to determine whether an opposite nucleobase-stabilisation effect to −3 was present at the −2 position (Table 2-5).

Table 2-5: *Table displaying codes, description and corresponding sequence changes used for nucleobase sequence changes of −3 and −2 abasic sites. Neighbouring nucleobases to abasic sites are highlighted.* ***X*** *= TAMRA-thioester* ***2-1*** *site location* ***Y*** *= abasic site* ***2-10*** *location.*

Code	Description	DNA codes and sequences
−3 TTT	−3 abasic sequence modified with neighbouring thymine bases	**2-1f** 5′ −CAG**T**GC**X**GT− 3′ **2-10k** 3′ −GT**T**Y**T**GACA− 5′
−2 GGG	Original −2 sequence	**2-1a** 5′ − AGC**G**C**X**GT − 3′ **2-10f** 3′ − TC**G**Y**G**ACA − 5′
−2 CCC	−2 abasic sequence modified with −3 neighbouring bases	**2-1g** 5′ −AGG**C**G**X**GT− 3′ **2-10g** 3′ −TC**C**Y**C**ACA− 5′

As before, the sequences were subject to UV-Vis and fluorescence spectroscopy in a PBS (0.1 M, pH 7.4) buffer to analyse the TAMRA environment (Figure 2-23).

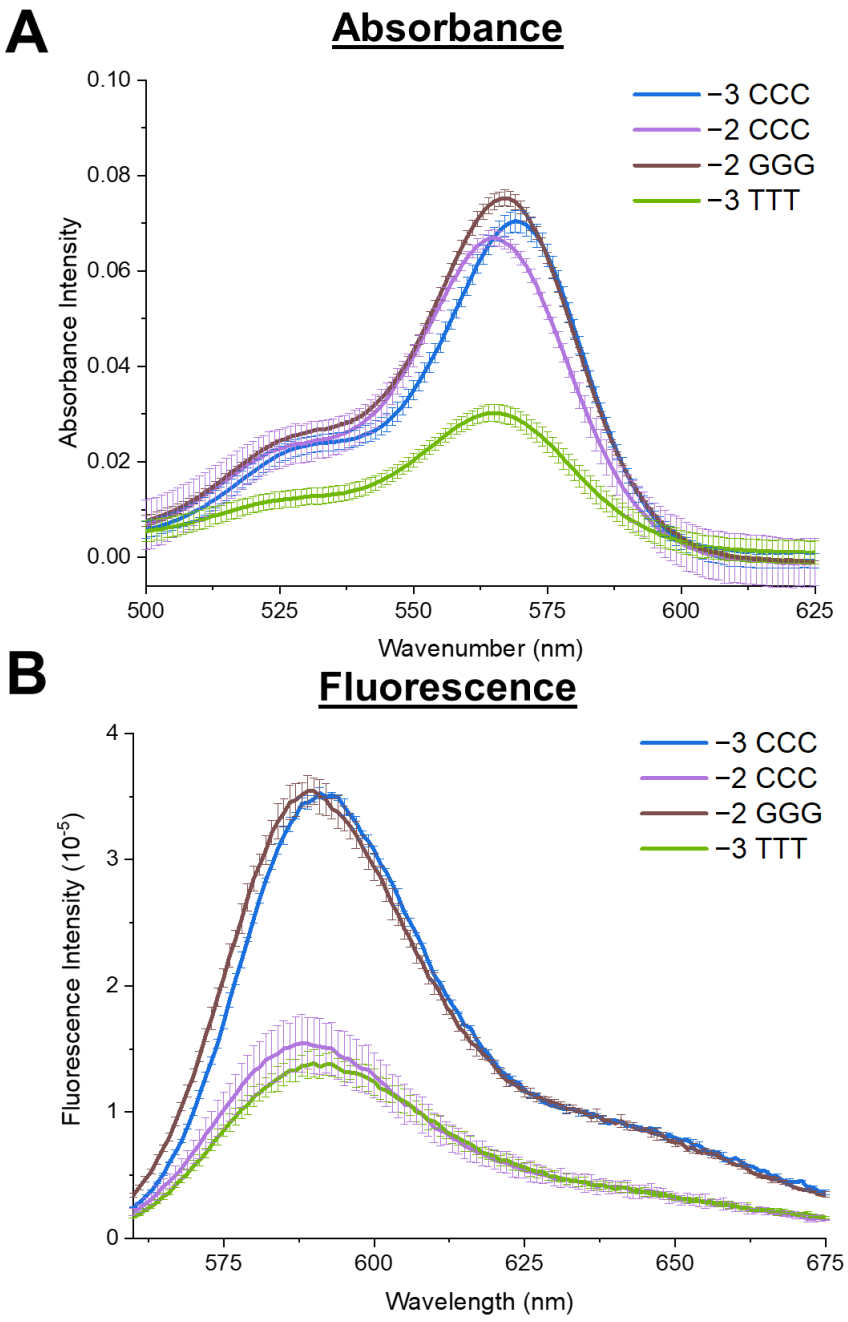


Figure 2-23: *(**A**) UV-Vis spectrum (λ = 200 nm – 600 nm) and (**B**) Fluorescence spectrum (λ_{Ex} = 550 nm, bandwidth = 1) of TAMRA-thioester DNA **2-1** (2.5 µM, 0.1 M PBS) of various sequences with various abasic DNA-spacers **2-10**: −3 CCC (original sequence, dark blue), −2 CCC (purple), −2 GGG (original sequence, brown) and −3 TTT (light green). Error bars show the standard deviation in intensity of three repeats.*

The −2 CCC sequence adjustment displayed a high absorbance, at 0.067 AU, but a low fluorescence intensity at 154589 AU, contradicting the previous assumption of lower absorbance intensities (Figure 2-23). As low fluorescence intensities had previously correlated with poor thioester protection, it was predicted that the protective effect would not be observed for the −2 CCC sequence adjustment. Also, despite previous support for fluorescence quenching in guanine-rich environments, −2 GGG deviated from this trend and displayed a high fluorescence at 354195 AU. It was hypothesised that the interactions within the DNA-TAMRA complex, which contributed to successful thioester protection, outweighed/prevented guanine-induced quenching. In addition, guanine-induced quenching has not yet been documented for the across-the-helix, and the effect may not be consistent at every abasic position.

On the other hand, for the −3 TTT modification, a decrease in absorbance and emission was observed in comparison to the original −3 CCC sequence, from 0.070 AU to 0.030 AU in absorbance intensity and 354195 AU to 138672 AU in fluorescence intensity. As thymine is slightly larger than cytosine, the size increase may reduce the accessibility of the modification into the spacer cavity and, thus, cause a decrease in absorbance/emission. Alternatively, cytosine may have a unique interaction with the TAMRA thioester DNA **2-1** modification at the −3 abasic position, promoting the protection effect. To confirm that the decrease in fluorescence intensity of −3 TTT and −2 CCC was a consequence of the lack of thioester-abasic insertion, the sequence adjustments were subject to a protection assay.

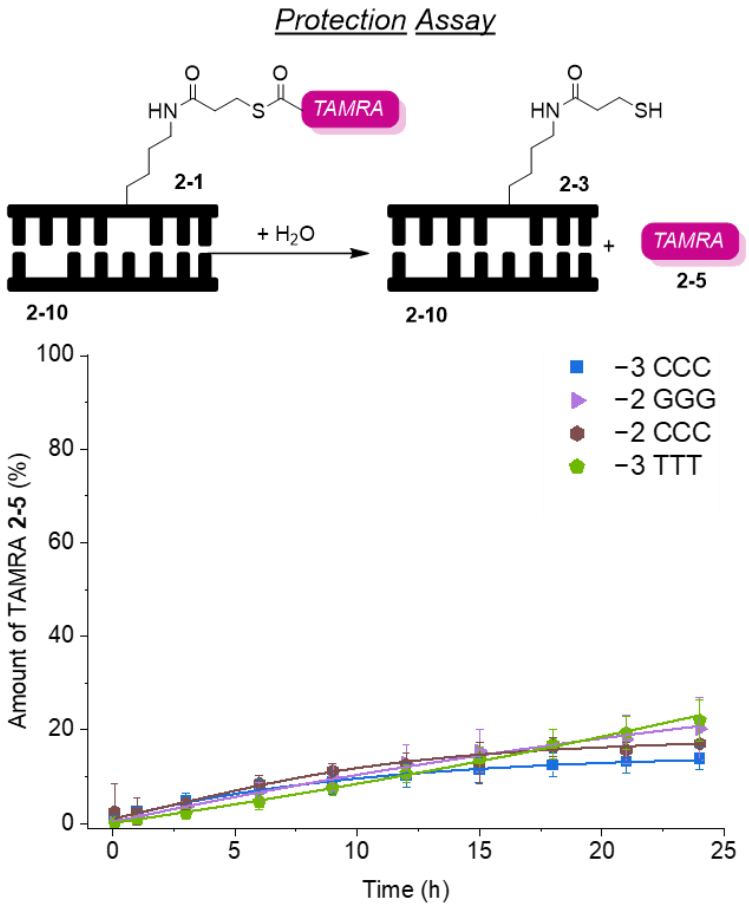


Figure 2-24: *Hydrolysed TAMRA* ***2-5*** *formation (%) within a protection assay (DNA-spacer* ***2-10****) of TAMRA thioester DNA* ***2-1*** *with various abasic −3 and −2 sequences (−3 CCC dark blue square, −3 TTT – green pentagon, −2 GGG purple triangle, −2 CCC brown hexagon), in DTS buffer pH 11, over 24 hours.*

Based upon the emission and absorbance trends, the protection effect unexpectedly occurred for sequence −3 TTT, in which 22% of TAMRA thioester DNA **2-1** hydrolysed after 12 hours at pH 11. However, the previous hypothesis that the size of the spacer aided in protection was confirmed for the −3 abasic site. As both concentrated thymine and concentrated cytosine environments stabilised the thioester of TAMRA thioester DNA **2-1**, it demonstrated a nucleobase size-dependent protective effect at the −3 position in the across-the-helix architecture. Supported by the molecular modelling of the −3 abasic sequence adjustments (Figure 2-16), the

accessibility of a pyrimidine-surrounded abasic site likely stabilises TAMRA thioester DNA **2-1**, preventing hydrolysis or aminolysis from occurring.

In addition, as the −3 TTT spectrometry readings deviated from the protection effect trends (high fluorescence = successful thioester protection), it was concluded that there was no correlation between absorbance/fluorescence intensity and the protection effect of the adjusted sequences. Therefore, UV-Vis and fluorescence spectroscopy cannot predict the outcome of either aminolysis or protection yields of the sequence-adjusted TAMRA thioester DNA **2-1** DTS system.

For the −2 DNA spacer sequences, both −2 GGG and −2 CCC demonstrated successful thioester protection, at 20% and 17% degradation, respectively, after 24 hours of incubation. Contrary to the previous hypothesis, it was concluded that the protection effect at the −2 position was independent of the size dependency observed at the −3 position, as both the smallest and largest nucleobases protected the thioester at the −2 position. The observation at the −3 abasic position may be a consequence of the short distance between the spacer site and the TAMRA thioester DNA **2-1** modification site (0.75 nm for −3 vs 0.90 nm for −2).[3] Thus, utilisation of the smaller nucleobases provided the optimum distance for thioester protection to occur.

Overall, both the protection effect and DTS aminolysis displayed a unique set of dependencies within the TAMRA thioester DNA **2-1** across-the-helix architecture. The absorbance, the fluorescence, the protection effect and the DTS yield were all interdependent on both the abasic position and its surrounding nucleobases. Each of the abasic positions subject to investigation demonstrated varying degrees of dependence on its nucleobase sequence, making trends difficult to predict.

2.3 Conclusions

The photophysical properties of TAMRA **2-5**, TAMRA-NHS **2-6** and the TAMRA thioester DNA **2-1** modification, in an across-the-helix DTS architecture, were explored to identify TAMRA's interaction with DNA in the thioester protection effect documented by Frommer *et al*.[3]

Initial investigation into the absorbance and fluorescence of both TAMRA **2-5** and TAMRA-NHS **2-6** demonstrated a positive solvatochromism effect. Here, polar protic solvents affected the photophysical properties of the fluorophores, shifting spectra to a higher intensity and a longer wavelength due to the stabilisation of excitation transition states.

The abasic position dependency of the TAMRA thioester DNA **2-1** modification in an across-the-helix architecture was further supported by the changes in absorbance and emission spectra. Here, the thioester-protecting minor groove positions demonstrated high absorbance/fluorescence intensities, indicating TAMRA's access to a polar environment. In contrast, the low absorbance/fluorescence intensities of the major groove positions demonstrated TAMRA's positioning in a non-polar environment, *i.e*. the hydrophobic DNA core. Therefore, it was concluded that TAMRA was interacting with the connected DNA environment and aiding in the stabilisation of the thioester across the minor groove.

As suggested by rhodamine-DNA literature, the order of nucleobases surrounding the thioester-TAMRA modification affected the absorbance and emission spectra, and also when further applied to DTS and protection assays. The −3 and +1 abasic positions demonstrated contrasting DTS/protective effects when presented with

similar nucleobase environments. Nucleobase sequence adjustments did not promote aminolysis/stabilisation of the opposite position, concluding that the position dependency was a dominant parameter in the across-the-helix architecture's DTS ability, in addition to the newly established sequence dependency. Therefore, the DNA-TAMRA photophysical trends proposed for the position dependency were expanded to encompass abasic-surrounding nucleobases. However, these were later disproved when applied to additional abasic sites, the −2 position and a thymine-rich −3 position, due to contradictory trends between protection and spectral intensities. It was concluded that UV/Vis and fluorescence spectroscopy were not suitable prediction methods for aminolysis or protection effect efficiency. However, the techniques did provide evidence of structural changes of TAMRA in the modified DNA architecture.

Despite the complexities in concluding TAMRA's protective mechanism, one hypothesis could be drawn for the −3 abasic position – a nucleobase-size dependency. As suggested by predictive molecular models, the protection effect was possible at this position due to an increase in accessibility to the spacer cavity, contributed by neighbouring pyrimidine bases. Here, it was believed that the TAMRA-thioester adapter inserts into the larger abasic site, preventing hydrolysis and aminolysis. However, this theory could not be applied to other positions across the minor groove due to a potential distance dependency.

In summary, a nucleobase dependency, a position dependency and a potential distance dependency were identified for the TAMRA-thioester DTS mechanism established by the O'Reilly and Turberfield groups. Multiple, interdependent interactions occur between the nucleobases, TAMRA and the thioester linker, all in the

presence of a varying abasic site position, in the across-the-helix architecture. Such interactions were observable by the changes in absorbance and fluorescence of varying sequences and spacer locations, concluding the importance of TAMRA in the modified architecture. However, due to the inhomogeneous nature of DNA, no trends were identified between the parameters, the photophysical properties or the thioester protection effect. Supported by the literature, the interaction between rhodamine and DNA is complex and involves multiple parameters, as demonstrated within this project.[35, 36, 42] In order to develop an understanding of TAMRA's exact mechanism with DNA, a comprehensive 2D NMR spectroscopy study would be required. However, the findings of this work emphasise the importance of parameter considerations when designing an across-the-helix DTS mechanism and TAMRA's role in the thioester protection effect. Proceeding on, any future work involving TAMRA with the DNA architecture requires careful consideration, and parameter variation should be minimised.

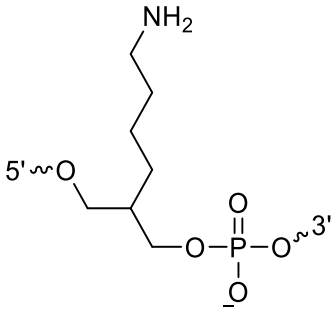
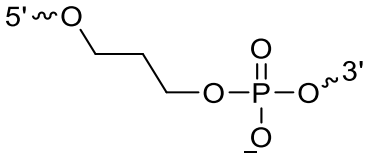
2.4 Experimental

For details of the RP-HPLC, LC-MS, UV-Vis spectroscopy and fluorescence spectroscopy, see Appendix 1: Experimental Methods.

2.4.1 Materials

TCEP and SPDP **2-7** were purchased from Sigma and (6-)TAMRA-NHS **2-6** and (6-)TAMRA **2-5** were purchased from Cambridge Bioscience. Phosphate-buffered saline sachets (PBS) were purchased from Sigma and made up using 18.2 MΩ H_2O. For the synthesis of DNA modifications, borate buffer was formulated using boric acid (Sigma), diluted with 18.2 MΩ H_2O and adjusted to pH 9.8 with 1 M NaOH. DTS buffer was formulated with NaCl (0.25 M), *N*-Tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid (0.25 M, TAPS), 3-(Cyclohexylamino)-1-propanesulfonic acid (0.25 M, CAPS) and Na_3PO_4 (0.25 M). The pH was adjusted to pH 11 using 1 M NaOH. DTS buffer components were purchased from Sigma, other than CAPS, which was purchased from Thermo Fischer. DNA sequences were purchased from Integrated DNA Technologies (IDT), with either standard desalting or HPLC purification. All samples were suspended in water (18.2 MΩ) at a concentration of 1 mM.

***Table 2-6**: DNA modifications and sequences, purchased from IDT.*

Key	Description	Modification
/iUniAmM/ **2-2**	internal Uni-Link™ Amino Modifier	 ** /iUniAmM/
/iSpC3/ **2-10**	Int C3 Spacer	 ** /iSpC3/
/**X**/	/iUniAmM/ **2-2** or /iSpC3/ **2-10**	

Identification		Sequence 5′ → 3′ (differences from original are highlighted)
Donor (TAMRA) **2-8**	Original Donor (ssDNA-NH_2 **2-8a**)	GCCGAGCCAGCAGTCAGCGC/**iUniAmM**/GTCCTAATCTACCTG
	+1 CCC 2-8b	GCCGAGCCAGCAGTCAGCGC/**iUniAmM**/CGCCTAATCTACCTG
	+1 GGG 2-8c	GCCGAGCCAGCAGTCAGCGC/**iUniAmM**/GCCCTAATCTACCTG
	−3 AGA 2-8d	GCCGAGCCAGCAGTCATGTC/**iUniAmM**/GTCCTAATCTACCTG
	−3 GGG 2-8e	GCCGAGCCAGCAGTCACGCC/**iUniAmM**/GTCCTAATCTACCTG
	−3 TTT 2-8f	GCCGAGCCAGCAGTCAATAC/**iUniAmM**/GTCCTAATCTACCTG
	−2 CCC 2-8g	GCCGAGCCAGCAGTCAGGCG/**iUniAmM**/GTCCTAATCTACCTG
	Complementary	CAGGTAGATTAGGACAGCGCTGACTGCTGGCTCGGC
Acceptor (spacer or NH_2) **2-2/2-10**	**+3** (a)	CAGGTAGATTAG/ **iSpC3** /ACAGCGCTGACTGCTGGCTCGGC
	+1 AGA (b)	CAGGTAGATTAGGA/ **X** /AGCGCTGACTGCTGGCTCGGC
	+1 CCC (c)	CAGGTAGATTAGGC/ **X** /CGCGCTGACTGCTGGCTCGGC
	+1 GGG (d)	CAGGTAGATTAGGG/ **X** /GGCGCTGACTGCTGGCTCGGC
	0 (e)	CAGGTAGATTAGGAC/ **iSpC3** /GCGCTGACTGCTGGCTCGGC
	−2 GGG (f)	CAGGTAGATTAGGACAG/ **iSpC3** /GCTGACTGCTGGCTCGGC
	−2 CCC (g)	CAGGTAGATTAGGACAC/ **iSpC3** /CCTGACTGCTGGCTCGGC
	−3 CCC (h)	CAGGTAGATTAGGACAGC/ **X** /CTGACTGCTGGCTCGGC
	−3 AGA (i)	CAGGTAGATTAGGACAGA/ **X** /ATGACTGCTGGCTCGGC
	−3 GGG (j)	CAGGTAGATTAGGACAGG/ **X** /GTGACTGCTGGCTCGGC
	−3 TTT (k)	CAGGTAGATTAGGACAGT/ **X** /TTGACTGCTGGCTCGGC

2.4.2 TAMRA thioester DNA synthesis

Stock solutions: SPDP **2-7** (0.1 M, DMF), TCEP (0.1 M, H_2O), 6-TAMRA-NHS **2-6** (0.1 M, DMF)

In a microcentrifuge tube, 2 µL of NH_2-labelled DNA (**2-8**, 1 mM, H_2O) was added to 7 µL borate-buffered saline (100 mM, pH 9.8), followed by 2 µL SPDP **2-7** (DMF, 0.1 M). The solution was mixed for 18 h at 19 ºC and 1000 rpm. Afterwards, the solution was dissolved in 300 µL ethanol with 20 µL sodium acetate (NaOAc) (0.3 M, pH 5.2) and placed in the freezer overnight. The modified DNA pellet was centrifuged (15000 rpm), emptied and washed twice with ice-cold ethanol (70%, H_2O). The modified DNA pellet was dried, dissolved in 500 µL H_2O and spin-filtered (Amicon-Ultra 3 kDa 0.5 mL centrifugal filter). The remaining 100 µL of solution was dried to a pellet and resuspended in 2 µL H_2O along with 4 µL of TCEP (0.1 M, H_2O). The solution was mixed for 30 minutes at 19 ºC, 1000 rpm, before adding 1 µL of TAMRA-NHS **2-6** (0.1 M, DMF) and 7 µL of borate-buffered saline (100 mM, pH 9.8). The magenta solution was mixed for 18 h, at 19 ºC, 1000 rpm. After 18 h, 300 µL ethanol and 20 µL NaOAc (0.3 M, pH 5.2) were added, and the tube was placed in the freezer overnight. As before, the tubes were centrifuged, washed with ice-cold ethanol, dried and spin-filtered. The sample was then purified using fluorescence RP-HPLC (see *Appendix: Experimental Methods A.2.2*), eluting at 18.5 mins, and washed with H_2O in Amicon-Ultra 3 kDa 0.5 mL centrifugal filters, removing HPLC buffer salts. Resuspended in nanopure H_2O, the modified DNA was identified using LC-MS and the concentration was determined using a NanoDrop One/One UV-Vis Spectrophotometer absorbance

at λ = 260 nm. **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 11382.02 m/z, found: 11382.51 m/z. See Supplementary Figure 2-1 to 2-7 for characterisation spectra.

2.4.3 Fluorescence and absorbance analysis

2.4.3.1 TAMRA 2-5 and TAMRA-NHS 2-6

In a microcentrifuge vial, 1.25 µL of small molecule (100 µM, DMF) was added to 1.67 µL of $MgCl_2$ (0.3 M) and 47.08 µL of PBS (0.1 M, pH 7.4), DTS buffer (0.25 M NaCl, 0.25 M TAPS, 0.25 M CAPS and 0.25 M Na_3PO_4, pH 11), water or organic solvent (DMF, DMSO, EtOH, IPA, MeOH or THF) and mixed for 1 minute. Samples were analysed by UV/Vis and fluorescence spectroscopy (see *Appendix: Experimental Methods A.6 and A.7*).

2.4.3.2 Varying abasic positions, original sequence

In a primary microcentrifuge vial, 1.25 µL of TAMRA-thioester DNA **2-1** (100 µM) was added to 23.75 µL of PBS (0.1 M, pH 7.4). In a secondary vial, 1.67 µL of $MgCl_2$ (0.3 M), 1.25 µL of ssDNA (See Table 2-7) and 22.08 µL of PBS (0.1 M, pH 7.4) were mixed. Both microcentrifuge vials were combined and incubated (thermocycler, 21 ºC) for 30 minutes before being analysed.

Table 2-7: *ssDNA used for original across-the-helix duplex fluorescence and UV-Vis spectroscopy analysis. See Table 2-6 for sequences.*

Code	**DNA (100 µM)**
+3 DNA-spacer	+3 ssDNA-spacer acceptor **2-10a**
+1 DNA-spacer	+1 AGA ssDNA-spacer acceptor **2-10b**
0 DNA-spacer	0 ssDNA-spacer acceptor **2-10e**
−2 DNA-spacer	−2 GGG ssDNA-spacer acceptor **2-10f**
−3 DNA-spacer	−3 CCC ssDNA-spacer acceptor **2-10h**
Complementary	Complementary ssDNA
ssDNA TAMRA-thioester DNA **2-1**	PBS (0.1 M, pH 7.4)

2.4.3.3 Varying sequences and varying abasic positions

Adapted method of Experimental 2.4.3.2.: 1.25 µL of TAMRA-thioester DNA **2-1** (donor) and 1.25 µL ssDNA (acceptor, **2-1** and **2-10**) were replaced with a new set of sequences, see Table 2-8.

Table 2-8: *TAMRA thioester DNA **2-1** and ssDNA used for sequence-modified across-the-helix duplex fluorescence and UV-Vis spectroscopy analysis. See Table 2-6 for sequences.*

Code		**DNA (100 µM)**
+1 CCC	Donor	TAMRA-thioester DNA **2-1** modified on +1 CCC **2-8b**
	Acceptor	+1 CCC ssDNA-spacer **2-10c**
+1 GGG	Donor	TAMRA-thioester DNA **2-1** modified on +1 GGG **2-8c**
	Acceptor	+1 GGG ssDNA-spacer **2-10d**
−2 CCC	Donor	TAMRA-thioester DNA **2-1** modified on −2 CCC **2-8d**
	Acceptor	−2 CCC ssDNA-spacer **2-10g**
−3 AGA	Donor	TAMRA-thioester DNA **2-1** modified on −3 AGA **2-8e**
	Acceptor	−3 AGA ssDNA-spacer **2-10i**
−3 GGG	Donor	TAMRA-thioester DNA **2-1** modified on −3 GGG **2-8f**
	Acceptor	−3 GGG ssDNA-spacer **2-10j**
−3 TTT	Donor	TAMRA-thioester DNA **2-1** modified on −3 TTT **2-8g**
	Acceptor	−3 TTT ssDNA-spacer **2-10k**

2.4.4 Molecular modelling

Molecular modelling studies of the unoccupied cavity in the −3 DNA-spacer **2-10** sequences with TAMRA thioester DNA **2-1** across-the-helix architecture were predicted using PyMOL Version 3.1.0, with the built-in plugin APBS Electrostatics (apbs_gui).[52] The NMR structure of a −3 abasic site in an across-the-helix architecture was used as a model (PDB ID: 2L2V).[51, 53] The model's sequence was adjusted to accommodate the surrounding nucleobases of our TAMRA thioester DNA **2-1** and −3 DNA-spacer **2-10** sequence adjustments using the mutagenesis function. Structure optimisation and molecular surface were calculated using the built-in APBS

Electrostatics plugin. The surface layer was adjusted to visualise only structural cavities using the PyMOL display settings.

***Table 2-9**: DNA sequences used for molecular modelling of −3 DNA-spacer **2-10** compared to original PDB ID: 2L2V. X = /iSpC3/. Sequence adjustments are highlighted in blue.*

	Sequence X = /iSpC3/
Original sequence (2L2V)	5′-CAGCTATG**X**GAAGC 3′-GTCGA**X**ACACTTCG
−3 CCC	5′-CAGAGCGC**X**GAAGC 3′-GTCTC**X**CGACTTCG
−3 AGA	5′-CAGATGTC**X**GAAGC 3′-GTCTA**X**AGACTTCG
−3 GGG	5′-CAGACGCC**X**GAAGC 3′-GTCTG**X**GGACTTCG
−3 TTT	5′-CAGAATAC**X**GAAGC 3′-GTCTT**X**TGACTTCG

2.4.5 Protection assay

Stock solutions: TAMRA thioester DNA **2-1** (9 µM, H_2O, various sequences, see Table 2-10), ssDNA-spacer **2-10** (9 µM, H_2O, various positions and sequences, see Table 2-10), DTS buffer (pH 11, 0.25 M NaCl, 0.25 M TAPS, 0.25 M CAPS, 0.25 M Na_3PO_4).

In a PCR tube (200 µL), 0.6 µL $MgCl_2$ (0.5 M), 4.4 µL H_2O (18.2 MΩ), 3.33 µL TAMRA thioester DNA **2-1** (9 µM, donor, see Table 2-6) and 6.67 µL ssDNA-spacer **2-10** (9 µM, acceptor, see Table 2-6) were added. In a Mastercycler, the tube was incubated for 20 minutes at 15 ºC. The solution was then diluted in a spring vial insert (100 µL) containing 15 µL of DTS buffer (pH 11), mixing three times with the transferred pipette tip. The solution was injected (2 µL) directly for RP-HPLC analysis, analysing

the fluorescence channel (λ_{ex} 550 nm, λ_{em} 580 nm). Samples were injected at 0 h, 1 h, 3 h, 6 h, 9 h, 12 h, 15 h, 18 h, 21 h and 24 h intervals and completed in triplicate (minimum). The injection inlet was set to purge and flush prior to every sample injection to prevent salt buildup (2 minutes). The RP-HPLC sample tray was incubated at 15 ºC for the duration of the assay. Yields were calculated using calibration curves of free TAMRA **2-5** and normalised to 30 pmol (maximum reaction pmol).

***Table 2-10**: ssDNA used for sequence-modified across-the-helix duplex RP-HPLC protection assay. See Table 2-6 for sequences.*

Code	**DNA**	
+1 AGA	Donor	TAMRA-thioester DNA **2-1** modified on Original Donor **2-8a**
	Acceptor	+1 AGA ssDNA-spacer **2-10b**
+1 CCC	Donor	TAMRA-thioester DNA **2-1** modified on +1 CCC **2-8b**
	Acceptor	+1 CCC ssDNA-spacer **2-10c**
+1 GGG	Donor	TAMRA-thioester DNA **2-1** modified on +1 GGG **2-8c**
	Acceptor	+1 GGG ssDNA-spacer **2-10d**
−2 GGG	Donor	TAMRA-thioester DNA **2-1** modified on Original Donor **2-8a**
	Acceptor	−2 GGG ssDNA-spacer **2-10f**
−2 CCC	Donor	TAMRA-thioester DNA **2-1** modified on −2 CCC **2-8d**
	Acceptor	−2 CCC ssDNA-spacer **2-10g**
−3 CCC	Donor	TAMRA-thioester DNA **2-1** modified on Original Donor **2-8a**
	Acceptor	−3 CCC ssDNA-spacer **2-10h**
−3 AGA	Donor	TAMRA-thioester DNA **2-1** modified on −3 AGA **2-8e**
	Acceptor	−3 AGA ssDNA-spacer **2-10i**
−3 GGG	Donor	TAMRA-thioester DNA **2-1** modified on −3 GGG **2-8f**
	Acceptor	−3 GGG ssDNA-spacer **2-10j**
−3 TTT	Donor	TAMRA-thioester DNA **2-1** modified on −3 TTT **2-8g**
	Acceptor	−3 TTT ssDNA-spacer **2-10k**

2.4.6 Single-step DNA-templated synthesis assay

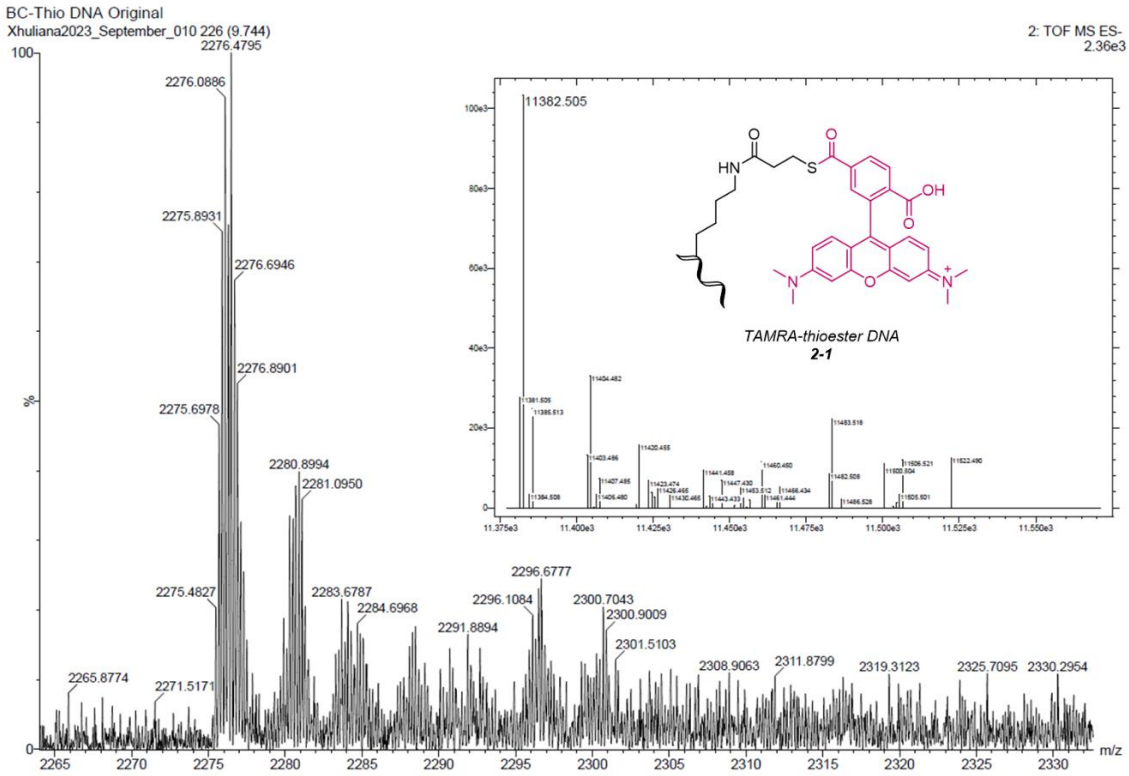
Stock solutions: TAMRA thioester DNA **2-1** (9 μM, various sequences, H_2O), ssDNA-NH_2 **2-2** (9 μM, various positions and sequences, H_2O), DTS buffer (pH 11, 0.25 M NaCl, 0.25 M TAPS, 0.25 M CAPS, 0.25 M Na_3PO_4).

Following a modified procedure to *Experimental 2.4.4:* 6.67 μL ssDNA-spacer **2-10** (9 μM, acceptor) was replaced with 6.67 μL ssDNA-NH_2 **2-2** (9 μM, acceptor, see Table 2-11) and 3.33 μL of the corresponding TAMRA-thioester DNA **2-1** (9 μM, donor, see Table 2-11). Samples were injected at 0 h, 1 h, 3 h, 6 h, 9 h, and 12 h intervals and completed in triplicate (minimum). Yields were calculated using calibration curves of free TAMRA **2-5** and TAMRA amide product **2-4** and normalised to 30 pmol (maximum reaction pmol).

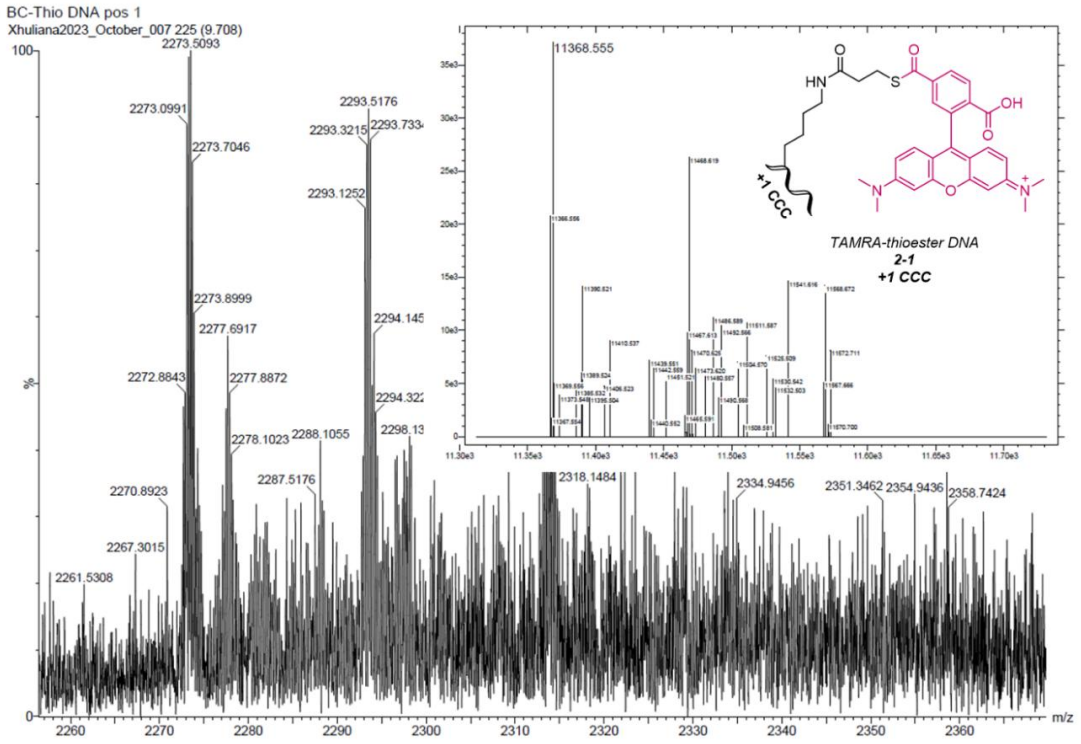
Table 2-11: *ssDNA used for sequence-modified across-the-helix duplex RP-HPLC DTS assay. See Table 2-6 for sequences.*

Code	DNA	
+1 AGA	Donor	TAMRA-thioester DNA **2-1** modified on Original Donor **2-8a**
	Acceptor	+1 AGA ssDNA-NH_2 **2-2b**
+1 CCC	Donor	TAMRA-thioester DNA **2-1** modified on +1 CCC **2-8b**
	Acceptor	+1 CCC ssDNA-NH_2 **2-2c**
+1 GGG	Donor	TAMRA-thioester DNA **2-1** modified on +1 GGG **2-8c**
	Acceptor	+1 GGG ssDNA-NH_2 **2-2d**
−2 GGG	Donor	TAMRA-thioester DNA **2-1** modified on Original Donor **2-8a**
	Acceptor	−2 GGG ssDNA-NH_2 **2-2f**
−2 CCC	Donor	TAMRA-thioester DNA **2-1** modified on −2 CCC **2-8d**
	Acceptor	−2 CCC ssDNA-NH_2 **2-2g**
−3 CCC	Donor	TAMRA-thioester DNA **2-1** modified on Original Donor **2-8a**
	Acceptor	−3 CCC ssDNA-NH_2 **2-2h**
−3 AGA	Donor	TAMRA-thioester DNA **2-1** modified on −3 AGA **2-8e**
	Acceptor	−3 AGA ssDNA-NH_2 **2-2i**
−3 GGG	Donor	TAMRA-thioester DNA **2-1** modified on −3 GGG **2-8f**
	Acceptor	−3 GGG ssDNA-NH_2 **2-2j**
−3 TTT	Donor	TAMRA-thioester DNA **2-1** modified on −3 TTT **2-8g**
	Acceptor	−3 TTT ssDNA-NH_2 **2-2k**

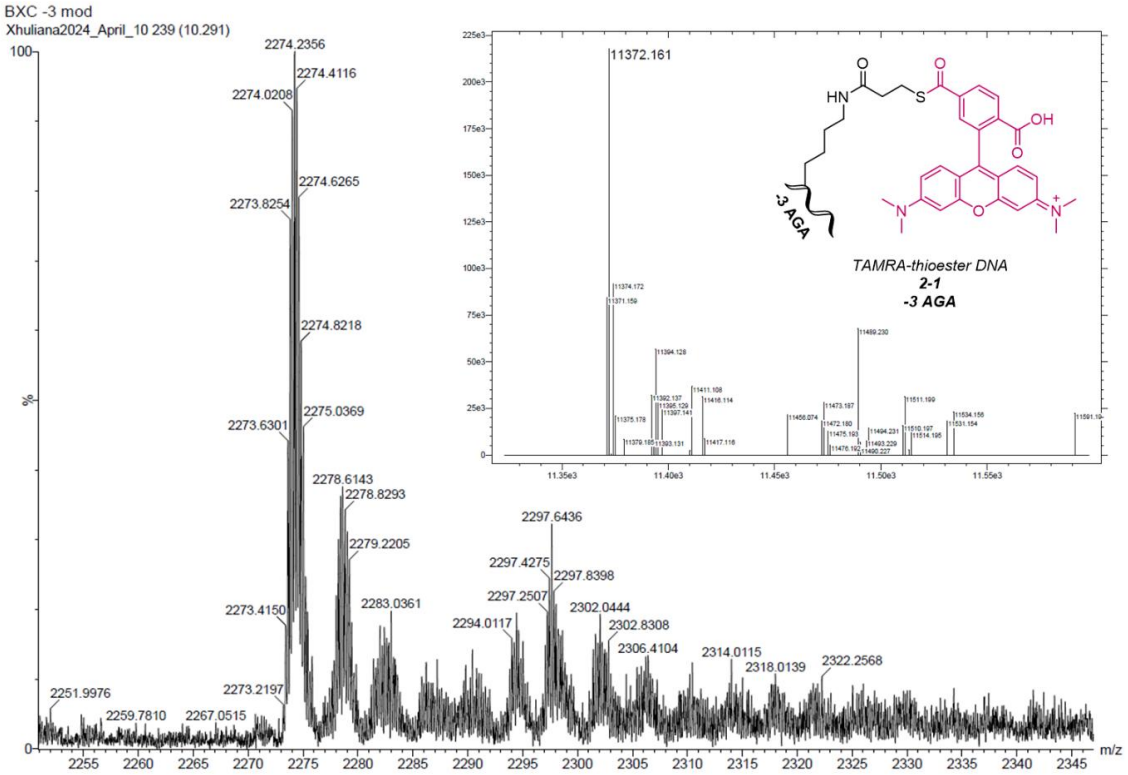
2.5 Supplementary data



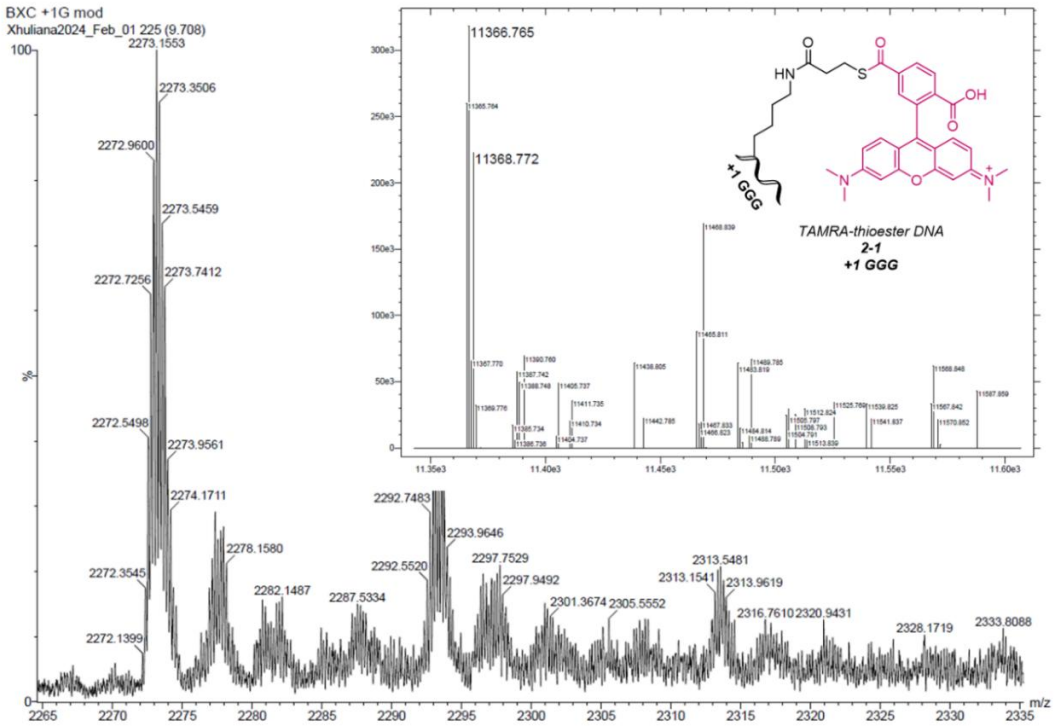
Supplementary Figure 2-1: *TOF MS ES− and deconvolution (ProMass HR) of TAMRA thioester DNA* ***2-1*** *on the original DNA strand. Expected Mass: 11383.03 m/z, found: 11382.51 m/z.*



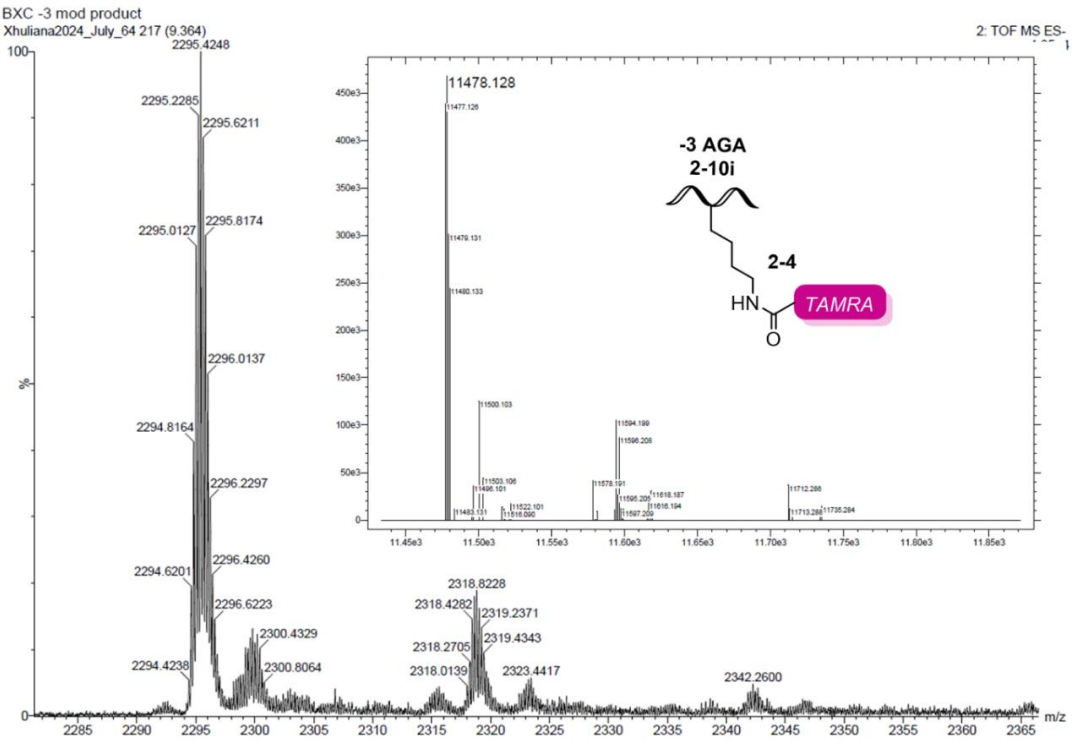
Supplementary Figure 2-2: *TOF MS ES− and deconvolution (ProMass HR) of TAMRA thioester DNA* ***2-1*** *+1 CCC. Expected Mass: 11368.03 m/z, found: 11368.56 m/z.*



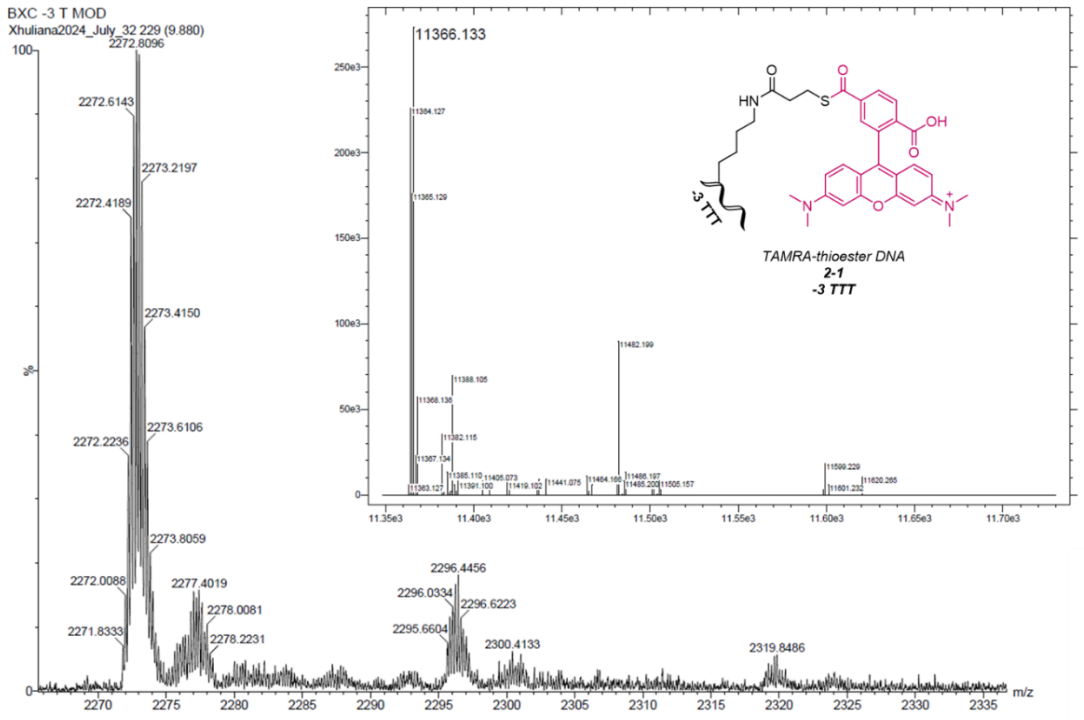
Supplementary Figure 2-3: *TOF MS ES− and deconvolution (ProMass HR) of TAMRA thioester DNA* ***2-1*** *−3 AGA. Expected Mass: 11373.02 m/z, found: 11371.16 m/z.*



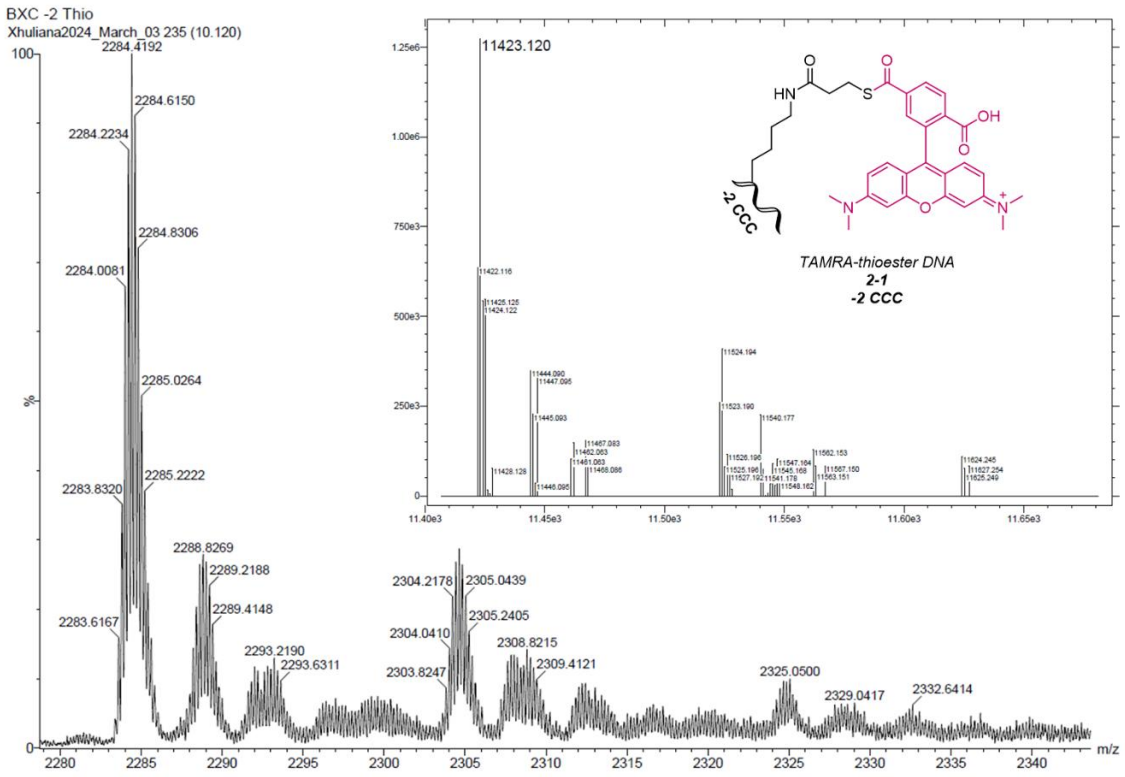
Supplementary Figure 2-4: *TOF MS ES− and deconvolution (ProMass HR) of TAMRA thioester DNA* ***2-1*** *+1 GGG. Expected Mass: 11368.03 m/z, found: 11368.77 m/z.*



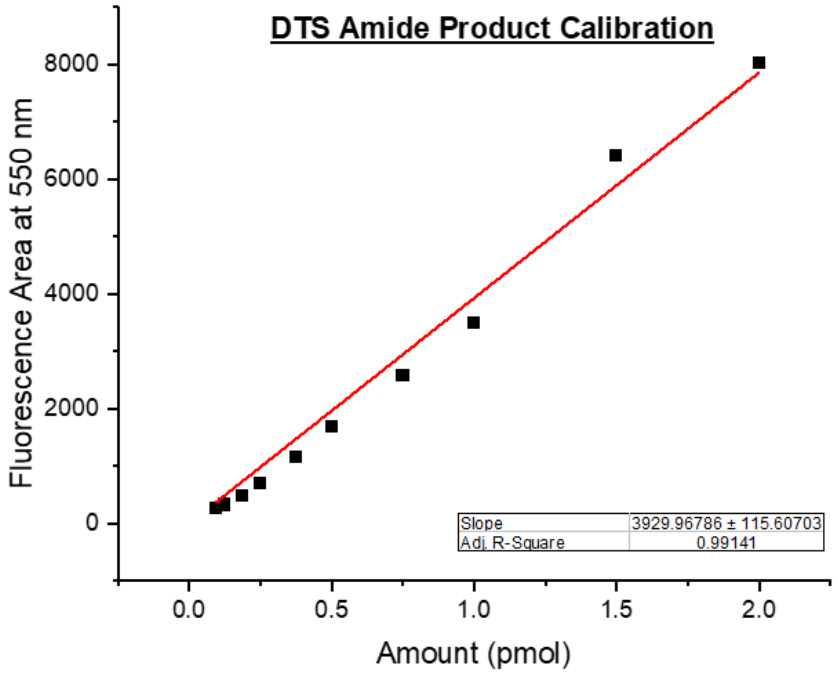
Supplementary Figure 2-5: *TOF MS ES− and deconvolution (ProMass HR) of TAMRA amide product* ***2-4*** *for the −3 AGA DNA sequence. Expected Mass: 11477.05 m/z, found: 11478.13 m/z.*



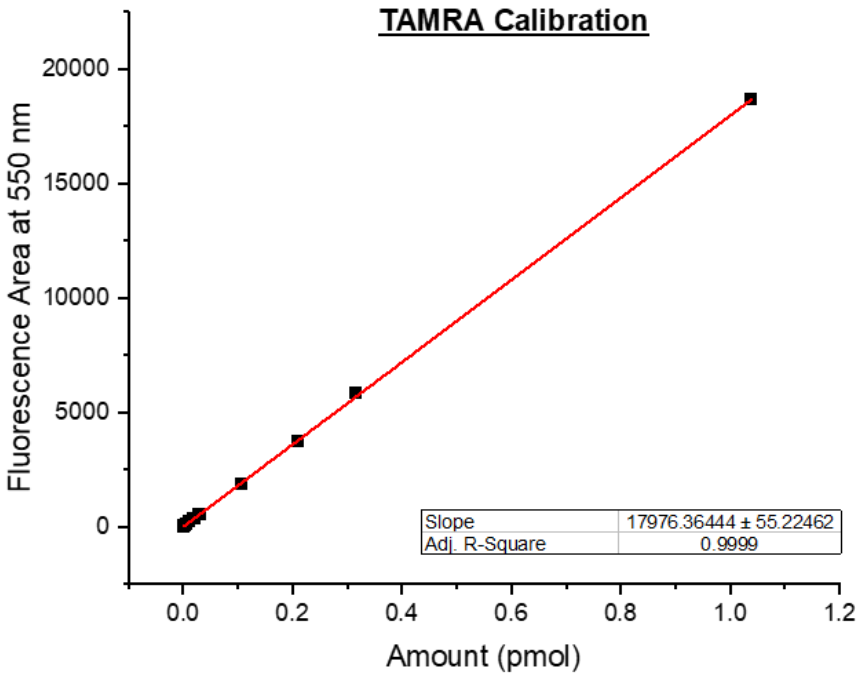
Supplementary Figure 2-6: *TOF MS ES− and deconvolution (ProMass HR) of TAMRA thioester DNA* ***2-1*** *−3 TTT. Expected Mass: 11366.04 m/z, found: 11366.13 m/z.*



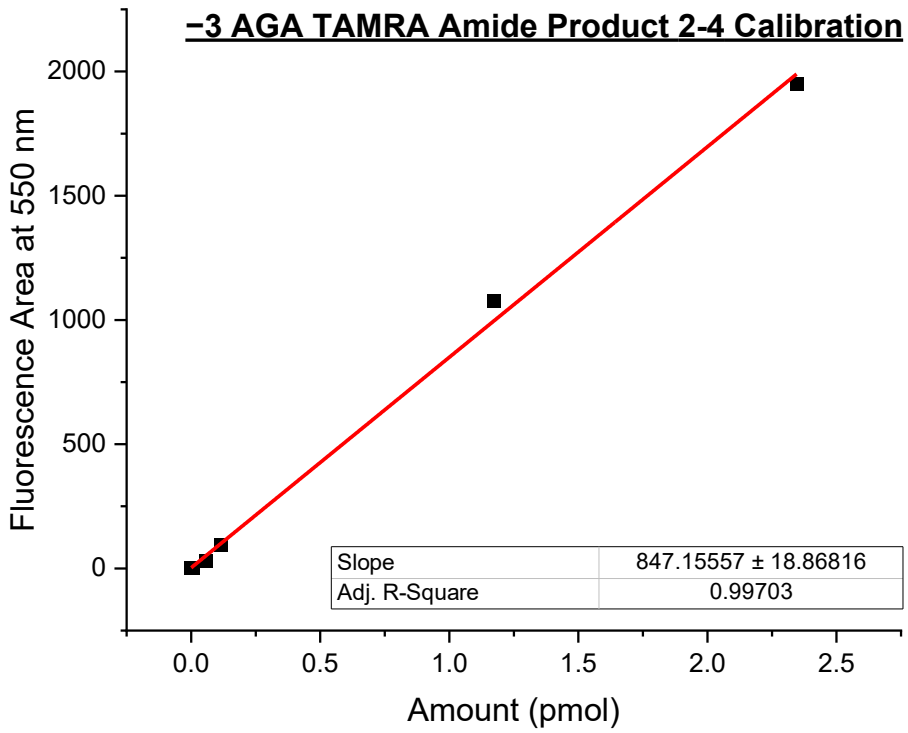
Supplementary Figure 2-7: *TOF MS ES− and deconvolution (ProMass HR) of TAMRA thioester DNA* ***2-1*** *−2 CCC. Expected Mass: 11423.03 m/z, found: 11423.12 m/z.*



***Supplementary Figure 2-8**: Calibration curve of DTS amide product **2-4**, original sequence, with slope and R^2 annotated.*



***Supplementary Figure 2-9:** Calibration curve of free TAMRA 2-5 with slope and R^2 annotated.*



***Supplementary Figure 2-10**: Calibration curve of DTS amide product **2-4** for −3 AGA sequence (2-10i) with slope and R^2 annotated.*

2.6 References

1. C. Zhang, Y. Li, M. Zhang and X. Li, *Tetrahedron*, 2012, **68**, 5152-5156.
2. M. L. McKee, A. C. Evans, S. R. Gerrard, R. K. O'Reilly, A. J. Turberfield and E. Stulz, *Org. Biomol. Chem.*, 2011, **9**, 1661-1666.
3. J. Frommer, R. Oppenheimer, B. M. Allott, S. Núñez-Pertíñez, T. R. Wilks, L. R. Cox, J. Bath, R. K. O'Reilly and A. J. Turberfield, *Angew. Chem. Int. Ed.*, 2024, **163**, e202317482.
4. A. J. Sami, S. Bilal, N.-u.-A. Ahsan, N. Hameed and S. Saleem, *Environ. Monit. Assess.*, 2023, **195**, 1442.
5. S. Zhan, H. Xu, X. Zhan, Y. Wu, L. Wang, J. Lv and P. Zhou, *Microchimica Acta*, 2015, **182**, 1411-1419.
6. S. Dehghani, N. M. Danesh, M. Ramezani, M. Alibolandi, P. Lavaee, M. Nejabat, K. Abnous and S. M. Taghdisi, *Anal. Chim. Acta*, 2018, **1030**, 142-147.
7. C. Li, X. Feng, S. Yang, H. Xu, X. Yin and Y. Yu, *ACS Appl. Mater. Interfaces*, 2021, **13**, 52406-52416.
8. Y. Fu, H. Jin, X. Bu and R. Gui, *J. Agric. Food. Chem.*, 2018, **66**, 9819-9827.
9. X. Wang, S. Zhong, Y. He and G. Song, *Anal. Methods*, 2012, **4**, 360-362.
10. K. Kaneko, M. Hara, T. Nishino and T. Maruyama, *Anal. Chem.*, 2020, **92**, 1978-1987.
11. W. Bao, J. Wang, Q. Wang, D. O'Hare and Y. Wan, *Sci. Rep.*, 2016, **6**, 26738.
12. T. Oh, T. Takahashi, S. Kim and M. J. Heller, *J. Biophotonics*, 2016, **9**, 49-54.
13. T. Oh, S. Kim, J.-Y. Choi, H. Chang and M. J. Heller, *RSC Advances*, 2017, **7**, 14902-14909.
14. P. Picchetti, S. Volpi, M. Sancho-Albero, M. Rossetti, M. D. Dore, T. Trinh, F. Biedermann, M. Neri, A. Bertucci, A. Porchetta, R. Corradini, H. Sleiman and L. De Cola, *J. Am. Chem. Soc.*, 2023, **145**, 22903-22912.
15. K. Kawai, E. Matsutani, A. Maruyama and T. Majima, *J. Am. Chem. Soc.*, 2011, **133**, 15568-15577.
16. I. L. Arbeloa and K. K. Rohatgi-Mukherjee, *Chem. Phys. Lett.*, 1986, **128**, 474-479.
17. F. Lopez Arbeloa, T. Lopez Arbeloa, M. J. Tapia Estevez and I. Lopez Arbeloa, *J. Phys. Chem.* 1991, **95**, 2203-2208.
18. C. Bindhu, S. Harilal, V. P. N. Nampoori and C. Vallabhan, *Mod. Phys. Lett. B*, 1999, **10**, 563-576.
19. T. Chen, L. Fu and L. Zu, *Luminescence*, 2013, **28**, 860-864.

20. S. Sun, X. Ning, G. Zhang, Y. C. Wang, C. Peng and J. Zheng, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 2421-2424.

21. C. J. Tredwell and A. D. Osborne, *J. Chem. Soc., Faraday Transactions 2: Molec. Chem. Phys.*, 1980, **76**, 1627-1637.

22. M. M. Wong and Z. A. Schelly, *J. Phys. Chem.*, 1974, **78**, 1891-1895.

23. S. Nagatoishi, T. Nojima, E. Galezowska, A. Gluszynska, B. Juskowiak and S. Takenaka, *Anal. Chim. Acta*, 2007, **581**, 125-131.

24. K. Yabuki, T. Ouchi and Y. Ohya, *Supramol. Chem.*, 2003, **15**, 149-154.

25. L. Wang, A. K. Gaigalas, J. Blasic and M. J. Holden, *Spectrochim. Acta A.* 2004, **60**, 2741-2750.

26. Y. Ohya, K. Yabuki, M. Tokuyama and T. Ouchi, *Supramol. Chem.*, 2003, **15**, 45-54.

27. I. Nazarenko, R. Pires, B. Lowe, M. Obaidy and A. Rashtchian, *Nucleic Acids Res.*, 2002, **30**, 2089-2195.

28. S. Nagatoishi, T. Nojima, E. Galezowska, A. Gluszynska, B. Juskowiak and S. Takenaka, *Anal. Chim. Acta*, 2007, **581**, 125-131.

29. A. K. Cunningham, MSc Dissertation, University of Birmingham, 2022.

30. L. Wang, A. K. Gaigalas, F. Abbasi, G. E. Marti, R. F. Vogt and A. Schwartz, *J. Res. NIST*, 2002, **107**, 339-353.

31. F. López Arbeloa, T. López Arbeloa, E. Gil Lage, I. López Arbeloa and F. C. De Schryver, *J. Photochem. Photobiol. A: Chem.*, 1991, **56**, 313-321.

32. A. Shiba, E. Kinoshita-Kikuta, E. Kinoshita and T. Koike, *Sensors*, 2017, **17**, 1877.

33. R. J. Christie, C. J. Tadiello, L. M. Chamberlain and D. W. Grainger, *Bioconjugate Chem.*, 2009, **20**, 476-480.

34. J. M. Drake, R. I. Morse, R. N. Steppel and D. Young, *Chem. Phys. Lett.*, 1975, **35**, 181-188.

35. J. R. Unruh, G. Gokulrangan, G. S. Wilson and C. K. Johnson, *Photochem. Photobiol.*, 2005, **81**, 682-690.

36. J. R. Unruh, G. Gokulrangan, G. H. Lushington, C. K. Johnson and G. S. Wilson, *Biophys. J.*, 2005, **88**, 3455-3465.

37. D. A. Barawkar and K. N. Ganesh, *Nucleic Acids Res.*, 1995, **23**, 159-164.

38. J. Bordello, M. I. Sánchez, M. E. Vázquez, J. L. Mascareñas, W. Al-Soufi and M. Novo, *Chem. Eur. J.*, 2015, **21**, 1609-1619.

39. X.-X. Zhang, S. L. Brantley, S. A. Corcelli and A. Tokmakoff, *Commun. Biol.*, 2020, **3**, 525.

40. M. Kubista, B. Akerman and B. Nordén, *Biochemistry*, 1987, **26**, 4545-4553.

41. N. Narayanaswamy, S. Das, P. K. Samanta, K. Banu, G. P. Sharma, N. Mondal, S. K. Dhar, S. K. Pati and T. Govindaraju, *Nucleic Acids Res.*, 2015, **43**, 8651-8663.

42. C. Eggeling, J. R. Fries, L. Brand, R. Günther and C. A. M. Seidel, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 1556-1561.

43. S. Wennmalm, L. Edman and R. Rigler, *Proc. Natl. Acad. Sci. U.S.A.*, 1997, **94**, 10641-10646.

44. J. Widengren, J. Dapprich and R. Rigler, *Chem. Phys.*, 1997, **216**, 417-426.

45. T. Heinlein, J.-P. Knemeyer, O. Piestert and M. Sauer, *J. Phys. Chem. B*, 2003, **107**, 7957-7964.

46. E. M. Peterson, E. J. Reece, W. Li and J. M. Harris, *J. Phys. Chem. B*, 2019, **123**, 10746-10756.

47. L. Edman, U. Mets and R. Rigler, *Proc. Natl. Acad. Sci. U.S.A.*, 1996, **93**, 6710-6715.

48. P. Qu, X. Chen, X. Zhou, X. Li and X. Zhao, *Sci. China Ser. B: Chemistry*, 2009, **52**, 1653-1659.

49. L. Clowney, S. C. Jain, A. R. Srinivasan, J. Westbrook, W. K. Olson and H. M. Berman, *J. Am. Chem. Soc.*, 1996, **118**, 509-518.

50. R. Heyrovska, *Open Struct. Biol. J.*, 2008, **2**, 1-7.

51. R. D. Hazel and C. de los Santos, *Biochemistry*, 2010, **49**, 8978-8987.

52. Schrödinger and LLC, *The PyMOL Molecular Graphics System (Version 3.1.0)*, 2024.

53. E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G.-W. Wei, M. J. Holst, J. A. McCammon and N. A. Baker, *Protein Sci.*, 2018, **27**, 112-128.

54. M. Kabeláč, F. Zimandl, T. Fessl, Z. Chval and F. Lankaš, *Phys. Chem. Chem. Phys.*, 2010, **12**, 9677-9684.

Chapter 3

Incorporation of selenoesters into a DNA-templated synthesis-based system

---

3.1 Introduction

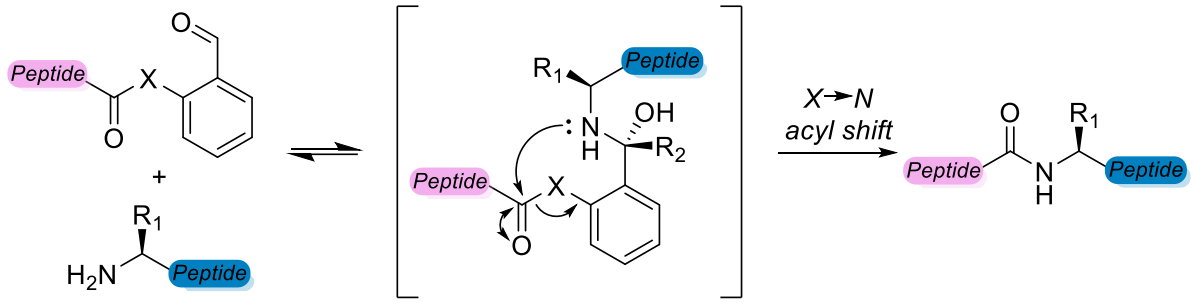
3.1.1 Selenoester's compatibility

DTS has the potential to revolutionise material discovery by autonomously synthesising novel sequence-specific polymers at low concentrations. However, acyl transfer-initiated DTS faces degradation of the reactive moieties, *via* hydrolysis, in the required basic aqueous environment.[1-3] Thus, limiting the overall yield and contributing to missed transfer steps (truncation) in a multi-step reaction progression.

Outside the field of DTS, methods to increase the efficiency of acyl transfer chemistries have involved the electrophile replacement of thioesters with selenoesters.[4-8] Reaction rates differ between the chalcogen analogues due to the enhanced leaving group ability of the large selenol atomic radius (103 pm) versus the smaller thiol (88 pm). Because of selenium's increase in size, the C−Se bond within a selenoester (582 kJ mol^{-}1) is weaker than the C−S of a thioester (699 kJ mol^{-}1);[9] the resonance donation from the selenium to the carbonyl carbon is reduced, increasing the electrophilicity of the carbonyl carbon.

For example, *via* aldehyde capture ligation, Raj *et al.* demonstrated that the X → N acyl shift proceeded at a faster rate for the selenoester, compared to thioester and oxoester analogues, in less than 2 minutes, 60 minutes, and 5 hours, respectively

(Scheme 3-1).[10] However, organic solvents were used to perform such acyl transfer reactions, *i.e.* dimethylformamide (DMF), limiting the competing hydrolysis reaction. As DTS takes place in aqueous conditions, the stability of chalcogen esters would need to be compared to determine their suitability in a DTS-compatible solution.



Scheme 3-1: *Aldehyde capture ligation. X = O, S or Se. Adapted from Raj et al.*[10]

Furthermore, Sayers *et al.* utilised a selenoester in peptide nucleic acid ligation, in a PBS-buffered system of pH 7.0 (Figure 3-1).[11] The kinetics studied in this work compared diselenide-selenoester ligation (DSL) to the sulfur-equivalent technique, native chemical ligation (NCL), resulting in reaction rates of 60 seconds for the seleno-compounds versus 600 seconds for the thiol-compounds (Figure 3-1B). Thus, demonstrating that selenoesters can perform as an effective electrophile at pH 7 and limit hydrolysis during an acyl transfer reaction. However, these conditions were successful due to the presence of the highly nucleophilic selenol. In the same work, Sayers *et al.* showed that selenoesters can be used in templated chemistries involving nucleobase hybridisation on peptide nucleic acids, similar to DTS (Figure 3-1A).[11]

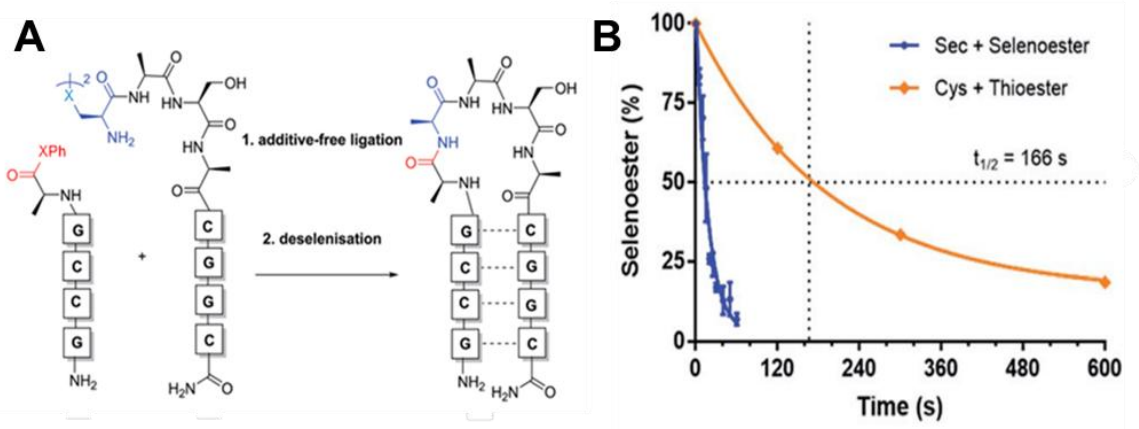
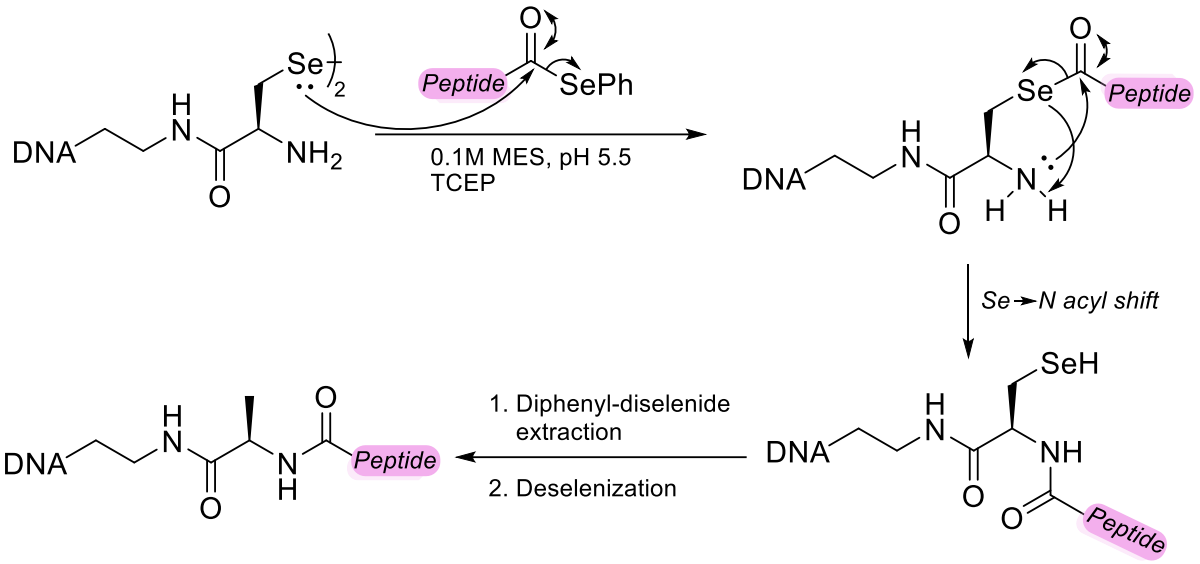


Figure 3-1: *Comparison between selenoesters (DSL, X = Se) and thioesters (NCL, X = S) reaction rates of peptide nucleic acid ligation of complementary strands.* ***A****) Reaction.* ***B****) Reaction rates. Sec = selenocysteine, Cys = cystine. Adapted from Sayers et al.*[11]

In addition, Liczner *et al.* illustrated the possibilities of selenoester ligation with DNA involvement (Scheme 3-2).[12] They demonstrated amide bond formation between a selenoester-activated peptide and a primary amine, connected to a DNA moiety, in a pH 5.5, 0.1 M 4-(2-sulfonatoethyl)morpholin-4-ium (MES) buffer. Although this work utilised a transfer reaction on DNA, the work cannot be considered a DTS reaction as it did not utilise the co-localisation of reactants through DNA hybridisation. In comparison, this DSL-DNA assay operated at a concentration of 0.7 mM selenoester-DNA, as opposed to DTS's less than 1 μM thioester-DNA concentration.[1] However, this DSL research provides an initial platform for the development and compatibility of a selenoester-based transfer reaction on a DNA framework. Therefore, making them a suitable candidate for the improvement of DTS transfer reactions.



***Scheme 3-2**: DSL reaction to form peptide-oligonucleotide conjugates, using a DNA fragment on the diselenide reactant. Adapted from Liczner et al.*[12]

Because of the enhanced reaction rates of selenoesters and their DNA compatibility, the replacement of thioesters with selenoesters in an aminolysis reaction has the potential to increase the efficiency of DTS. However, high pH conditions are often required in aminolysis reactions to deprotonate a primary ammonium salt, from its ammonium conjugate ($R\text{-}NH_3^+$), to produce a nucleophilic species ($R\text{-}NH_2$). The high pH conditions may accelerate the rate of hydrolysis of selenoesters, as a consequence of the increase in nucleophilic OH^- species, even more so than thioesters.[13] The DSL reactions documented above utilised the highly nucleophilic selenol as an intermediate, allowing for lower pHs to be used. DSL co-localises the selenoester with the primary amine, promoting successful Se→N acyl transfer.[13] However, NCL (including DSL) techniques have not yet been a successful candidate in multistep DTS reactions.[14-16] As DSL requires three reactive species (amine, selenol and selenoester), numerous protecting strategies would be required in intermittent

steps to prevent product truncation, making it unsuitable as an autonomous transfer reaction in DTS.[17]

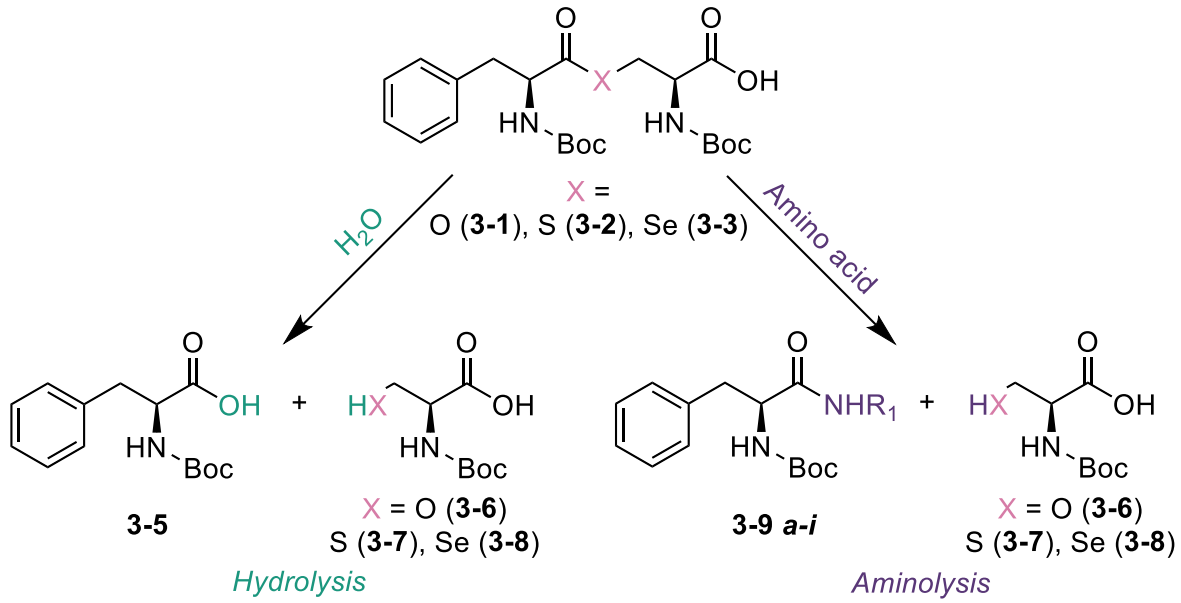
Nonetheless, selenoesters have demonstrated that if the optimum conditions are determined by balancing the sensitivity of both selenoesters and nucleophile reactivity, programmable amide bond formation can occur whilst reducing the rate of hydrolysis.[11] Therefore, this chapter set to investigate the compatibility of selenoesters in an aminolysis DTS mechanism, whilst comparing it to its thioester analogue. Initial aminolysis experiments were conducted with chalcogen-ester small molecules before proceeding to nucleic acid-conjugation chemistries in an attempt to improve DTS yields.

3.2 Results and discussion

3.2.1 Synthetic design

To determine whether selenoesters would make a viable candidate for the improvement in DTS transfer rates, under conditions required for DTS, a comparison with other acyl transfer electrophiles, oxo- and thioesters, was needed. As a preliminary strategy to compare the chalcogen analogues, the three ester types were incorporated into an amino acid framework (Scheme 3-3).

Amino acids were chosen because of their adaptability to a DTS system (established DNA attachment strategies,[18] and water compatibility) and their direct comparison between the chalcogen elements: serine (Ser) (oxygen **3-1**), cysteine (Cys) (sulfur **3-2**) and selenocysteine (Sec) (selenium **3-3**). The ester reactivity was studied in a peptide formation using phenylalanine (Phe) to facilitate monitoring *via* ultraviolet (UV) absorbance of RP-HPLC. By incubating the esters with an unprotected amino acid in a PBS-buffered solution (pH 7.4, 0.1 M), aminolysis would be observed (monitored *via* RP-HPLC and characterised *via* LC-MS) through the formation of a peptide bond within **3-9**, and hydrolysis by the formation of Boc-Phe-OH **3-5**.



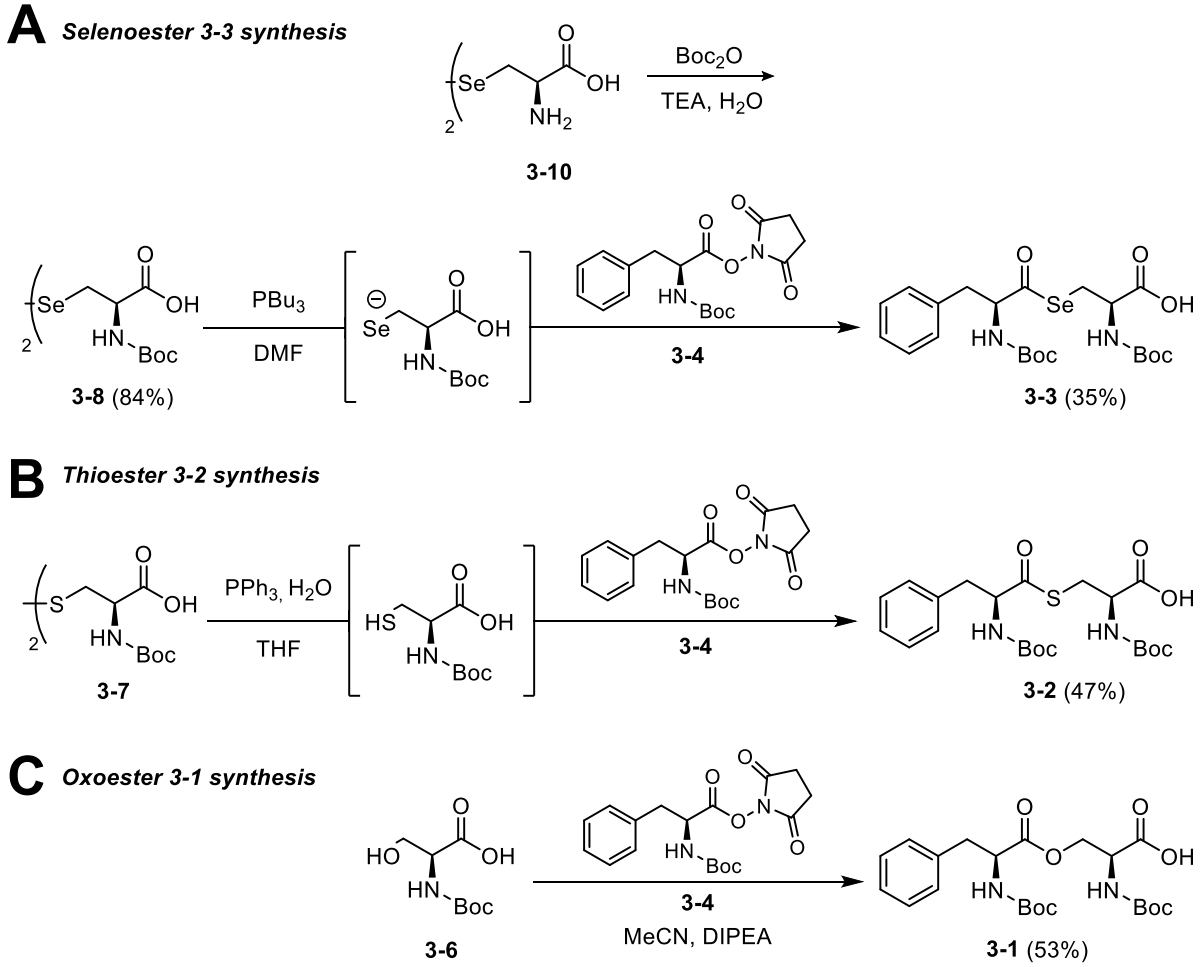
***Scheme 3-3**: Proposed Synthetic strategy to measure reactive behaviour of all three ester types **3-1** to **3-3** relevant to DTS transfer reactions: acyl transfer (aminolysis - purple) and hydrolysis (green). Built upon an amino acid framework.*

Following literature trends, there was an expected reactivity of the chalcogen ester analogues **3-1** to **3-3** in aqueous conditions.[10, 11] Down Group-16 of the Periodic Table, the reactivity of the chalcogen-containing esters is expected to increase. The increased reactivity of selenoesters and thioesters relates to their lower bond dissociation energy within the ester moiety (Se-C = 582 kJ mol^{-1}, S-C = 699 kJ mol^{-1} and O-C = 1076 kJ mol^{-1})[9]. This is an attribute of the decrease in resonance delocalisation down Group 16 of the periodic table. As selenium is atomically larger than the other chalcogen elements, the *4p* orbital overlap of the lone pair is reduced in a sp^2 hybridised selenium, leading to a weak bond between the elements and improving the leaving group's ability.

3.2.2 Synthetic optimisation

In order to synthesise the proposed analogous ester compounds **3-1** to **3-3**, chalcogenated amino acids were amine Boc-protected to prevent a side aminolysis

reaction. Boc-*l*-phenylalanine (Boc-Phe-OH) **3-5** was introduced as an *N*-hydroxysuccinimide ester (Boc-Phe-OSu, **3-4**) to induce an addition-elimination reaction with the nucleophilic chalcogen -ol moiety (Scheme 3-4). As selenoester synthesis literature is limited, determining the optimal reactive conditions for the small molecules was critical to ensure a smooth progression when transitioning to DNA-based chemistries for a DTS transfer assay.



***Scheme 3-4**: Synthetic route for amino acid esterification to produce the chalcogen esters: oxoester (C: **3-1**), thioester (B: **3-2**) and selenoester (A: **3-3**). For the synthetic procedure, see Experimental sections 3.4.2 to 3.4.4.*

Oxoester **3-1** synthesis commenced with the incubation of 1 eq. Boc-Ser-OH **3-6** and 2 eq. Boc-Phe-OSu **3-4** in acetonitrile (MeCN), to which no

reaction occurred (Table 3-1, entries 1-2). Therefore, acting as a proton scavenger, *N,N*-Diisopropylethylamine (DIPEA), was introduced.[19] The equivalents of DIPEA were adjusted accordingly, with 4 eq. providing the highest yield (53%), as opposed to 1 and 2 eq. (12% and 25%, respectively) (Table 3-1, entries 3-5). As expected, instability was observed when reaction temperatures were increased for oxoester **3-1** (Table 3-1, entries 6-8); therefore, the reaction temperature was set to 21 ºC for all proceeding syntheses.

Table 3-1: *Optimisation of chalcogen ester synthesis: oxoester **3-1**, thioester **3-2** and selenoester **3-3**. Boc-amino acid **3-6/7/8** (1 equivalence) with Boc-Phe-OSu **3-4** (2 equivalence) with additives. All DMF and THF were dry and degassed before use. * yield 0% due to solubility issues.*

	Entry	Boc-amino acid (1 eq.)	Boc-Phe-OSu 3-4	Additive (eq.)	Solvent	Temp. (ºC)	Time (h)	Yield (%)
Oxoester 3-1	***1***	**3-6**	2 eq	-	MeCN (20 mL)	21	72	0
	2	**3-6**	2 eq	-	MeCN (20 mL)	40	72	0
	3	**3-6**	2 eq	DIPEA (1 eq)	MeCN (20 mL)	21	72	12
	4	**3-6**	2 eq	DIPEA (2 eq)	MeCN (20 mL)	21	72	25
	5	**3-6**	2 eq	DIPEA (4 eq)	MeCN (20 mL)	21	72	**53**
	6	**3-6**	2 eq	DIPEA (4 eq)	MeCN (20 mL)	21	24	32
	7	**3-6**	2 eq	DIPEA (4 eq)	MeCN (20 mL)	40	72	22
	8	**3-6**	2 eq	DIPEA (4 eq)	MeCN (20 mL)	80	72	0
Thioester 3-2	***9***	**3-7**	2 eq	PPh_3 (1.2 eq) H_2O (40 eq)	MeCN (15 mL)	21	24	0*
	10	**3-7**	2 eq	PPh_3 (1.2 eq) H_2O (40 eq)	DMF (15 mL)	21	24	18
	11	**3-7**	2 eq	PPh_3 (1.2 eq) H_2O (40 eq)	THF (15 mL)	21	24	23
	12	**3-7**	2 eq	PPh_3 (1.2 eq) H_2O (20 eq)	THF (15 mL)	21	24	8
	13	**3-7**	2 eq	PPh_3 (1.2 eq) H_2O (80 eq)	THF (15 mL)	21	24	**47**
Selenoester 3-3	***14***	**3-8**	2 eq	PPh_3 (1.2 eq) H_2O (80 eq)	THF (4 mL)	21	24	0*
	15	**3-8**	2 eq	PPh_3 (1.2 eq) H_2O (80 eq)	DMF (4 mL)	21	24	0
	16	**3-8**	2 eq	PPh_3 (2 eq) H_2O (80 eq)	DMF (4 mL)	21	24	0
	17	**3-8**	2 eq	PBu_3 (1 eq)	DMF (4 mL)	21	24	26
	18	**3-8**	2 eq	PBu_3 (2 eq)	DMF (4 mL)	21	24	**35**

As optimal conditions were established for oxoester **3-1**, the same reactive conditions were initially used in the synthesis of thioester **3-2** (Table 3-1, entry 9). However, solubility issues with $(Boc\text{-}Cys\text{-}OH)_2$ **3-7** were identified, and MeCN was replaced with dry tetrahydrofuran (THF) (Table 3-1, entries 12-13). In addition,

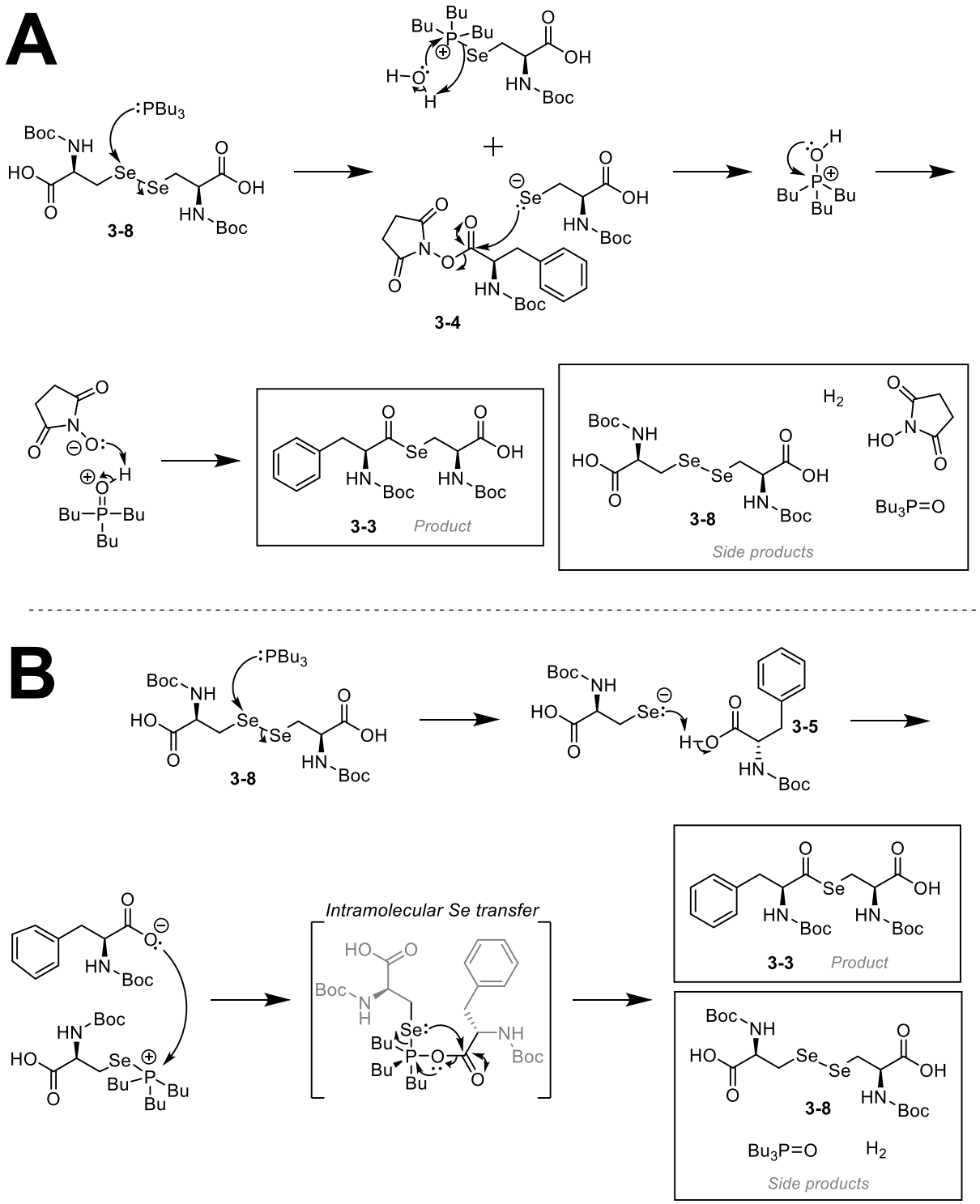
(Boc-Cys-OH)$_2$ **3-7** and (Boc-Sec-OH)$_2$ **3-8** dimerise under oxygenated conditions;[20, 21] therefore, the reducing agent triphenylphosphine (PPh_3) was added to expose the thiolate nucleophile.[22] Because water is required as a proton source in a phosphine-initiated reduction, the equivalence of water was modified with the motive of reducing potential degradation of thioester **3-2** during synthesis. However, yields were impaired when the equivalence of water was decreased, and the optimum conditions were determined at 80 eq. of water (47%) for thioester **3-2** (Table 3-1, entries 11-13).

Because of their similarities in starting materials, thioester **3-2**'s reactive conditions were initially applied to the synthesis of selenoester **3-3** (Table 3-1, entry 14). However, (Boc-Sec-OH)$_2$ **3-8** was insoluble in THF and was replaced with dry dimethylformamide (DMF) (Table 3-1, entry 15). In addition, PPh_3 proved to be unsuccessful at reducing (Boc-Sec-OH)$_2$ **3-8,** and a stronger reducing agent, tributylphosphine (PBu_3), was selected, yielding selenoester **3-3** at 35% for 2 eq. PBu_3 (Table 3-1, entries 16-18).[23]

Despite the successful product formation of selenoester **3-3** (Table 3-1, entries 17-18, Supplementary Figures 3-13:3-18), upon reflection of the mechanism, the reaction would not have proceeded without the presence of an available oxygen, quenching the reaction (tributylphosphine oxide). As PBu_3 is pyrophoric, the reaction was performed in an air/water-free environment. Glassware was dried at 180 °C overnight, and all solvents were dried and degassed before use. However, possible residual water (on the glassware or starting materials) may have allowed for the reaction to proceed (Scheme 3-5A). Alternatively, a possible theory proposed by Singh

et al. suggests the formation of a penta-coordinated seleno-oxy-phosphonium intermediate (Scheme 3-5B).[24] This approach requires the selenol-initiated deprotonation of a carboxylic acid, which may have been present as an impurity of Boc-Phe-OSu **3-4** (prior degradation/synthetic impurity). Here, a carboxylate ion binds with the positively charged selenophosphonium to form a penta-coordinated intermediate. Through the displacement of the weaker P-Se bond (363 kJ mol^{-1})[9] with a stronger P-O bond (596 kJ mol^{-1})[9], combined with the high nucleophilicity of selenols, an intramolecular O→Se acyl transfer reaction occurs. Thus, producing the stable by-product tributylphosphine oxide and the desired selenoester **3-3**.

The compounds were characterised by 2D HMBC NMR spectroscopy through the signal attributed to the $-CH_2$ proton between both carboxyl groups (ester and carboxylic acid) (Figure 3-2A, Supplementary Figures 3-4, 3-10, and 3-16). The change in the selenium environment was confirmed using ^{77}Se NMR spectroscopy, changing from the diselenide environment (^{77}Se-Se) of **3-8** to the selenoester environment (^{77}Se-C=O) of **3-3** (Figure 3-2B & Figure 3-2C).



Scheme 3-5: *Possible mechanisms of action in the reduction of the diselenide (Boc-Sec-OH)$_2$* ***3-8*** *for the synthesis of selenoester* ***3-3****.* ***A****) Proceeding in the presence of residual water.* ***B****) Proceeding in the presence of impurity Boc-Phe-OH* ***3-5****.*[24]

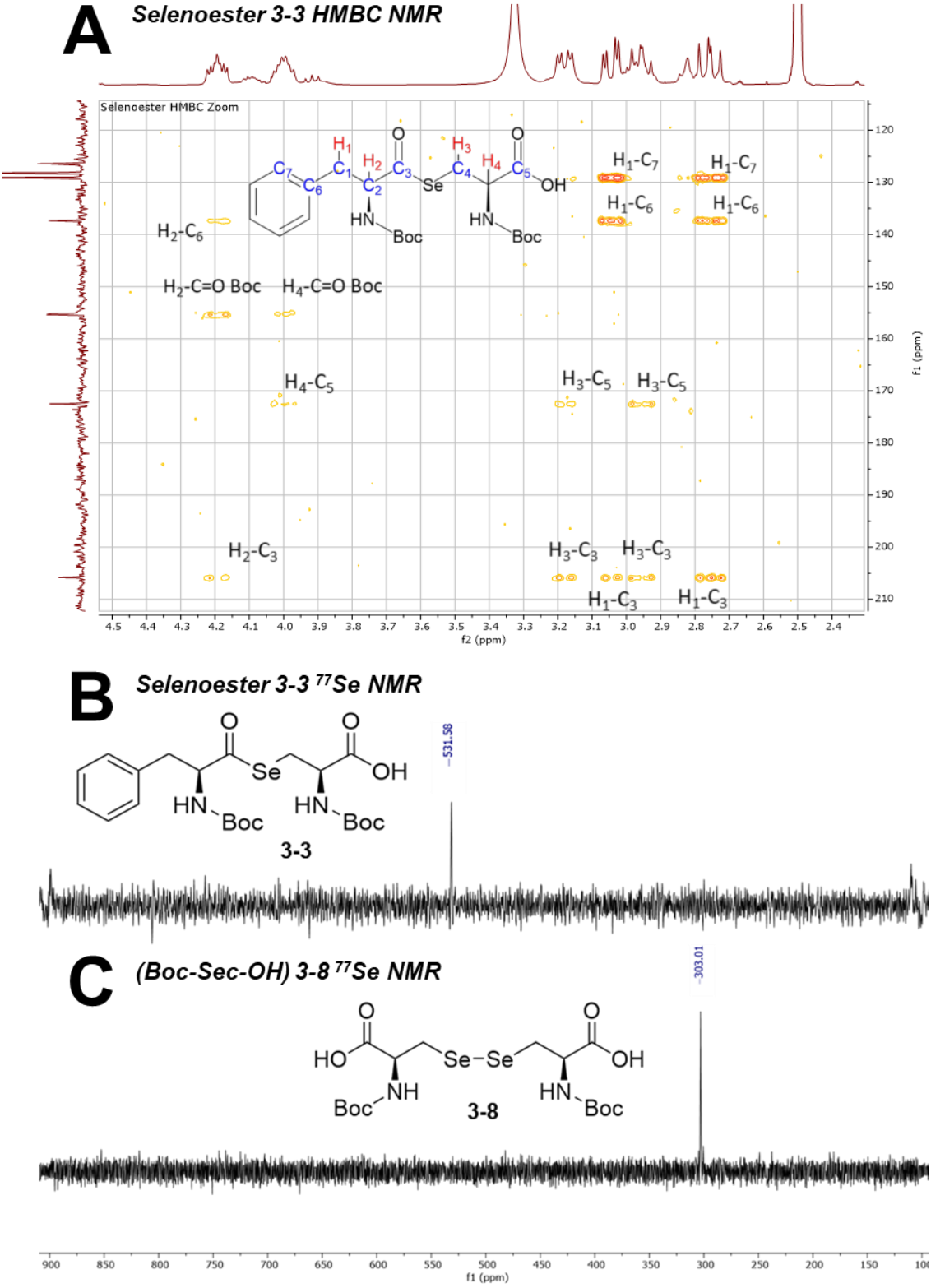


***Figure 3-2**: **A**) Characterisation of key identifier in HMBC of selenoester **3-3,** 400MHz/101 MHz, DMSO-d_6 **B**) ^{77}Se NMR spectrum of selenoester **3-3** ($\delta_{Se/ppm}$ = 531) and **C**) Boc-l-selenocystine **3-8** ($\delta_{Se/ppm}$ = 303), 76 MHz, DMSO-d_6.*

3.2.3 Hydrolysis and aminolysis

To compare the compatibility of chalcogen compounds within a buffered solution (1 mL), the rate of acyl transfer (aminolysis) of the ester (2 mM) with an unprotected amino acid (4 mM) and the rate of hydrolysis were compared. Experimental conditions were chosen in an attempt to replicate a DTS system without the use of DNA. *i.e*., smallest possible operational concentrations for small molecule synthesis whilst being RP-HPLC detectable and pH range accessible for aminolysis of reactants. The esters were diluted in a 10 mM phosphate-buffered saline (PBS) solution (the most common buffer used for DTS transfer reactions[25]), at varying pH values. pHs of 6.5, 9.5 and 10.5 were chosen to determine the optimal conditions required for amide bond formation.

To understand the stability of the three ester species and their compatibility within the selected aqueous system, the hydrolysis rate was measured without any amino acids present in the solution (Figure 3-3). In agreement with the literature, at high pH conditions after 10 days, selenoester **3-3** showed the fastest degree of hydrolysis (14% selenoester **3-3** remained) of the chalcogen esters, followed by thioester **3-2** (81%) and then oxoester **3-1** (95%) (Figure 3-3), correlating with bond dissociation energies. Oxoester **3-1** remained consistently stable at both pH 9.5 and pH 10.5, whereas selenoester **3-3** and thioester **3-2** had accelerated rates of degradation at pH 10.5. Subsequently, as selenoester **3-3** did not completely degrade upon exposure to an aqueous environment, the study progressed to observe reactivity with other nucleophilic species.

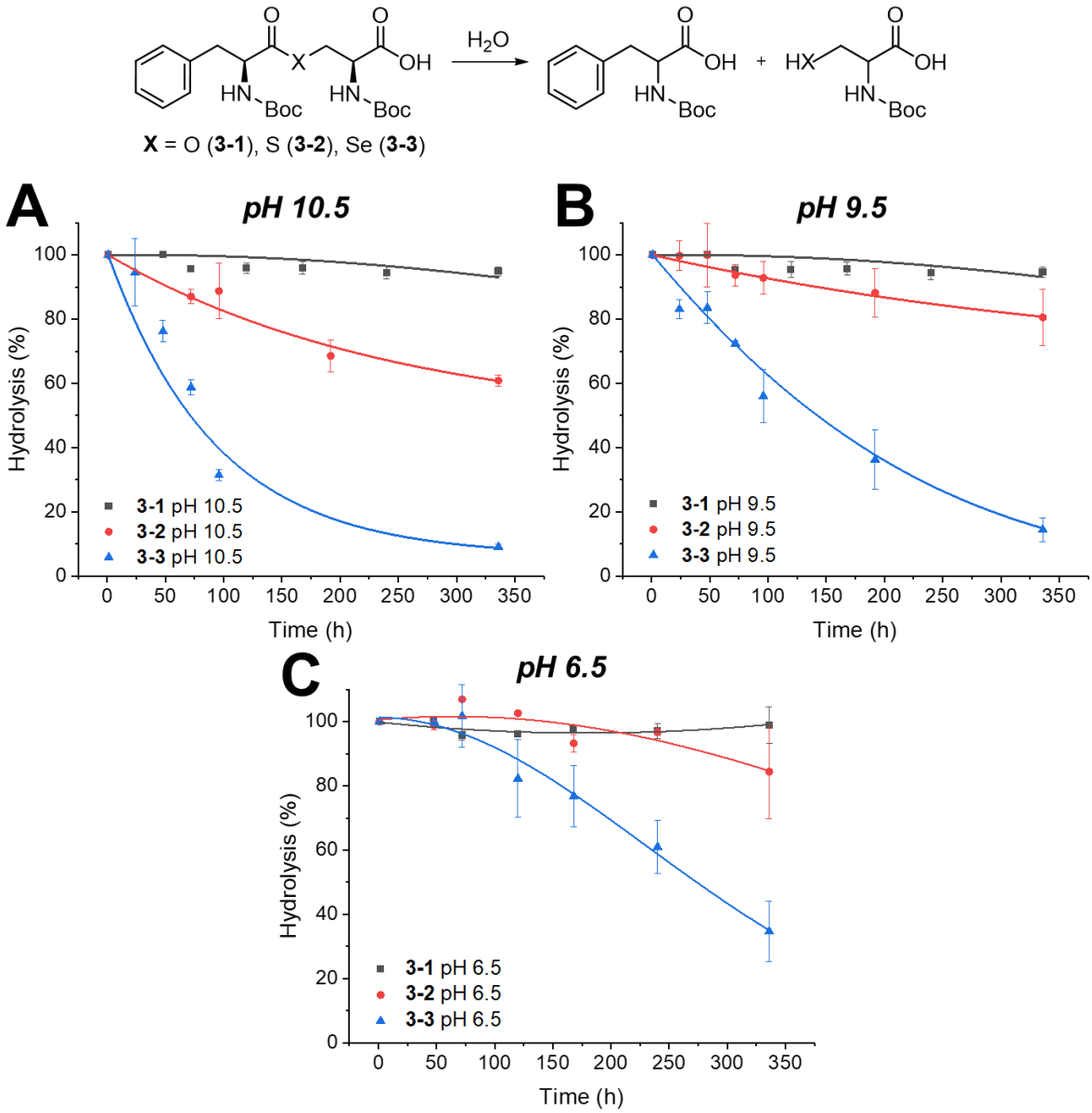


Figure 3-3: *Measured hydrolysis of chalcogen esters (2 mM)* ***3-1*** *to* ***3-3*** *(%) in PBS pH 10.5 (**A**), pH 9.5 (**B**) and pH 6.5 (**C**), via a change in RP-HPLC UV-absorbance. Error bars show the standard deviation in the hydrolysis of three repeats. See Experimental section 3.4.5.*

Once the hydrolysis of the esters was established, the rate of aminolysis between the ester and the α-NH_2 of multiple amino acids was monitored by RP-HPLC. (Scheme 3-3). The aminolysis study measured peptide formation (Boc-Phe-Amino Acid **3-9**) between an ester (**3-1** to **3-3**) and multiple amino acids: glycine (Gly), alanine (Ala), phenylalanine (Phe), asparagine (Asn), aspartic acid (Asp), lysine (Lys), arginine (Arg) and *Boc*-lysine (Boc-Lys). The amino acids were selected based on their diversity

between functional groups and amino acid type. Simultaneously, the degree of hydrolysis was measured based on the formation of Boc-Phe-OH **3-5**.

When subjected to aminolysis, oxoester **3-1** did not react with the *α*-NH_2 of any amino acids, within any of the tested pH values, over 2 weeks. This was expected due to the stability observed during the hydrolysis assay (Figure 3-3) and was further supported by literature trends in low oxoester-mediated aminolysis yields with primary amines.[26, 27] As a consequence of the high bond dissociation energy of the C-O bond and leaving group ability ($^-$OR), oxoesters are poor electrophiles and despite oxoester **3-1** being in the presence of a nucleophilic *α*-NH_2, aminolysis does not occur. Successful examples of oxoester-mediated aminolysis have had to utilise adjacent electron-withdrawing groups (to increase destabilisation of the carbonyl)[28], coupling agents[29] or organic base/metal catalysts[30-33] to prompt product formation.

Surprisingly, thioester **3-2** did not demonstrate successful aminolysis with the investigated amino acids, after 2 weeks of incubation. However, unlike oxoester **3-1**, thioester **3-2** hydrolysed during the aminolysis and hydrolysis assays at pH ≥ 9.5, observed by the formation of Boc-Phe-OH **3-5** (Figure 3-4, Figure 3-3). It was hypothesised that the bulky phenyl modification prevented access of the amino acid nucleophile to the electrophilic ester carbonyl. However, the smaller water nucleophile can readily access the thioester, leading to hydrolysis. Supported by the literature, previous research has documented ester stabilisation when bulky substituents are positioned within proximity.[34-37] Whilst previous DTS systems have demonstrated successful aminolysis between primary amines and thioesters,[1, 3, 38] this work did not

employ reactant attachment to DNA, removing the benefits displayed by DTS's proximity chemistries.

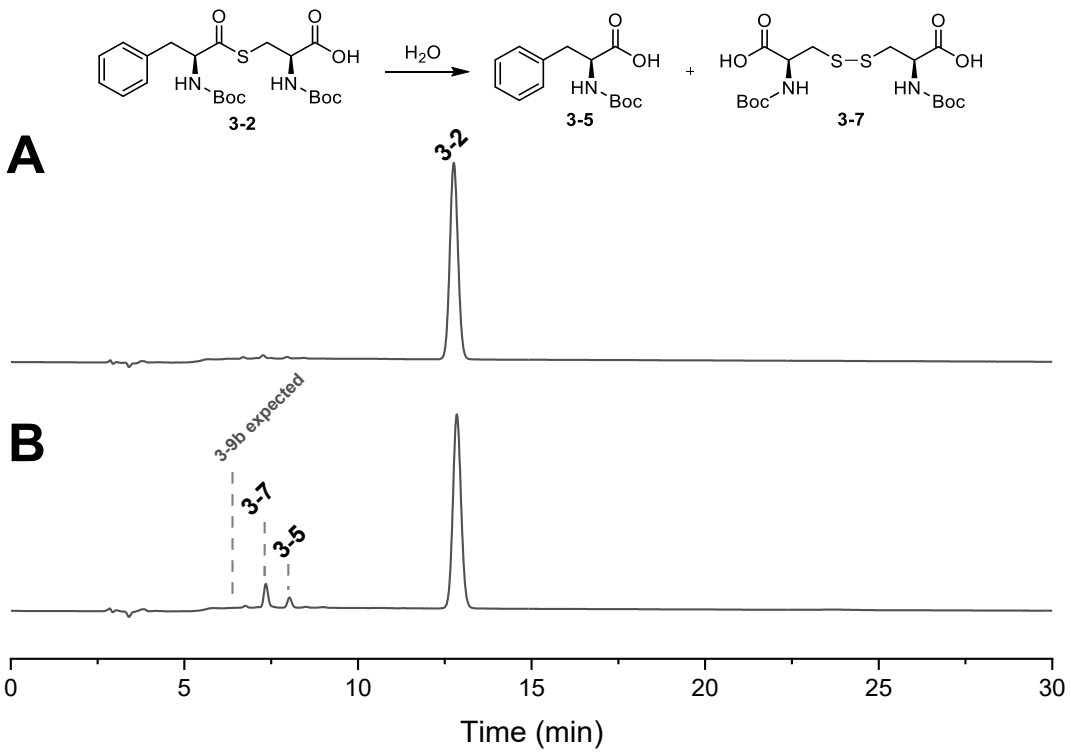


Figure 3-4: *RP-HPLC spectra of thioester **3-2** (2 mM) with Ala (4 mM) aminolysis assay at pH 9.5 after 24 hours (**A**) and 840 hours (**B**) of incubation. Hydrolysis by-product signals highlighted (Boc-Phe-OH **3-5** and (Boc-Cys-OH)$_2$ **3-7**). Expected signal of the desired product Boc-Phe-Ala-OH **3-9b** indicated in grey. λ = 254 nm, 20 µL injection volume.*

To test the phenyl-obstructing aminolysis hypothesis, the Ala-analogues of both thioester **3-2** and selenoester **3-3** were synthesised, in which Boc-Phe-OSu **3-5** was replaced with Boc-Ala-OSu **3-11**. However, despite successfully synthesising both Ala-selenoester **3-12** and Ala-thioester **3-13** (Supplementary Figure 3-23:3-32), both sterically smaller esters degraded during purification. It was hypothesised that the increase in accessibility, utilised by the Ala functionality, improved the ester reactivity, leading to its degradation.

Conversely, selenoester **3-3** showed observable aminolysis for five of the eight amino acids investigated, following the expected trend of the chalcogen esters

(Table 3-2). The aminolysis yield was calculated by dividing the moles of peptide **3-9** by the sum of peptide **3-9** moles and the hydrolysis by-product Boc-Phe-OH **3-5**. The moles of peptide **3-9** and Boc-Phe-OH **3-5** were calculated by integrating the corresponding RP-HPLC signal and comparing it to a calibration curve of either the aminolysis product or the hydrolysis product (Figure 3-5). The percentage yield represents the proportion in which the primary amine was successful in aminolysis, as both the primary amine of the amino acid and water from the solution are competing nucleophiles.

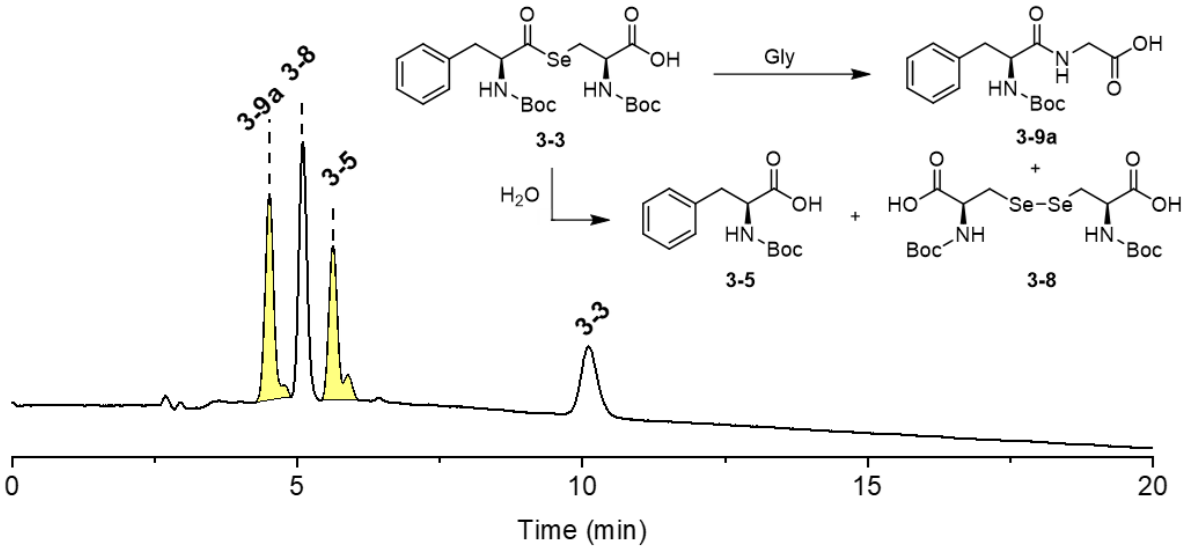


Figure 3-5: *Example RP-HPLC spectrum used to calculate aminolysis yield. Selenoester **3-3** (2 mM) with Gly (4 mM) aminolysis assay at pH 9.5 after 192 hours of incubation (λ = 254 nm, 20 µL injection volume). Signals integrated to determine the aminolysis yield are highlighted in yellow (Boc-Phe-Gly-OH **3-9a** Boc-Phe-OH **3-5**).*

Despite pK_a being a predictor of the nucleophilicity of primary amines, no correlation was observed between the α-NH_2 pK_a and the degree of aminolysis (Table 3-2). For example, of the investigated amino acids, Asn should have displayed the highest peptide yield at the lowest pK_a of 8.72; however, it was outperformed by three other amino acids, all of which had higher pK_a values. Upon investigation, steric

hinderance was identified as a contributing factor: the smaller the amino acid side chain, the greater the peptide **3-9** yield. The highest-yielding reactions involved Gly and Ala, the smallest of the proteinogenic amino acids, with 79% and 68% yield, respectively, at pH 10.5 after 336 hours. Opposed to the bulky phenyl substituent of Phe, which yielded 15% peptide **3-9e** after 336 hours at pH 10.5.

***Table 3-2**: Percentage (%) yield of aminolysis reaction between selenoester **3-3** (2 mM) and unprotected amino acid (4 mM) at varying pH values, after 336 hours. Percentage calculated by moles of peptide product **3-9** divided by the sum of peptide product **3-9** moles and hydrolysis Boc-Phe-OH **3-5** moles. Amino acids are identified by their R group. Experimental section 3.4.6.*

$$\frac{Moles\ of\ Peptide\ \mathbf{3-9}}{Moles\ of\ Peptide\ \mathbf{3-9} + Moles\ of\ BocPheOH\ \mathbf{3-5}} \times 100 = Aminolysis\ \%$$

	pH 6.5	pH 7.5	pH 8.5	pH 9.5	pH 10.5	R group	pK_a	Comments
Gly 3-9a	40%	66%	71%	74%	79%	None	α-NH_2: 9.78	
Ala 3-9b	27%	58%	61%	64%	68%	-CH_3	α-NH_2: 9.87	
Boc-Lys 3-9c	16%	56%	63%	65%	67%	α-NHBoc-$CH_2CH_2CH_2$-$CHNH_2$	ε-NH_2: 10.54	
Asn 3-9d	12%	16%	19%	25%	26%	-$CHOCNH_2$	α-NH_2: 8.72	
Phe 3-9e	3%	8%	8%	14%	15%	-CH_2Ph	α-NH_2: 9.13	
Asp 3-9f	0%	0%	0%	0%	0%	-CHOCOH	α-NH_2: 9.90	Acidic side group
Arg 3-9g	0%	0%	0%	0%	0%	-CHCHCHNH-C(NH)NH_2	α-NH_2: 8.99 ζ-NH_2: 12.48	Basic side group
Lys 3-9h	0%	0%	0%	0%	0%	-$CH_2CH_2CH_2$-$CHNH_2$	α-NH_2: 9.06 ε-NH_2: 10.54	Product not soluble in water

In addition, aminolysis with Gly was observed at the lowest pH of 6.5 with selenoester **3-3**. As Gly has an isoelectric point of 6.06, α-NH_2 nucleophiles are present at pH 6.5 despite the protonation equilibrium favouring α-NH_3^+ as the dominant form. As a result of the reactive nature of selenoesters, the small amount of available nucleophilic α-NH_2 of the amino acid still performs peptide bond formation, but at a slower rate than the higher pH conditions. As selenoester **3-3** is relatively stable in low pHs, whilst also performing aminolysis, it shows potential for incorporation into a DNA-mediated acyl transfer reaction, whilst also limiting hydrolysis of the starting material. Being able to incorporate selenoesters into such a setting has the potential to broaden the synthetic horizon of modified oligomers. For example, Liu *et al.* demonstrated DTS was possible at pH 7.5 when a strong electrophile is chosen, whilst also limiting the hydrolysis of the ester.[39]

Of the three amino acids where no product formation could be detected, Lys's results did not reflect the degree of hydrolysis the selenoester **3-3** should have displayed if aminolysis had not occurred (Figure 3-6). No selenoester **3-3** remained in the solution after 48 hours, despite remaining after one week in hydrolysis conditions at pH 10.5 and the amount of degradation compound, Boc-Phe-OH **3-5** (green line), in solution did not correlate to the expected amount (grey line) (Figure 3-6). It was hypothesised that the Lys-based product, Boc-Phe-Lys-OH **3-9h**, was insoluble in water due to its aliphatic side chain combined with the primary amine and, therefore, undetectable by RP-HPLC. Whereas for Arg, the Boc-Phe-OH **3-5** formed during the aminolysis studies (red line), matched that of the Boc-Phe-OH **3-5** formed in the

hydrolysis studies (grey line), indicating that no peptide product **3-9g** was formed between Arg and selenoester **3-3** (Figure 3-6).

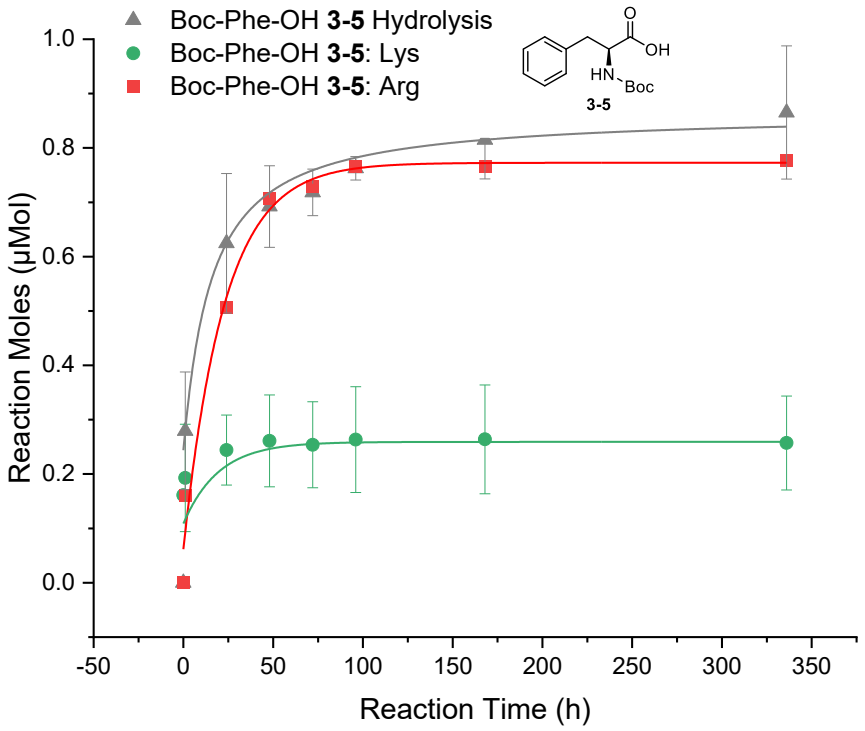


***Figure 3-6**: Amount (µMol) of Boc-Phe-OH **3-5** formed during aminolysis studies for Lys (green circles) and Arg (red squares) at pH 10.5 over the course of 336 h with selenoester **3-3** (2 mM). Boc-Phe-OH **3-5** formed during selenoester **3-3** hydrolysis studies (grey triangles). Error bars show the standard deviation in reaction moles of three repeats. Exponential fitting. Experimental section 3.4.6.*

Boc-Lys-OH was investigated to understand which primary amine of Lys would have initiated aminolysis: α-NH_2 or ε-NH_2 (pK_a = 9.06 vs 10.54, respectively). Following previous pK_a assumptions, α-NH_2 was the assumed nucleophile for Lys; however, Boc-Lys-OH presented successful aminolysis with the ε-NH_2 (Table 3-2, Supplementary Figure 3-20). The ε-NH_2 functionality is sterically available for electrophilic attack, providing successful aminolysis with selenoester **3-3**, and hence yielded peptide **3-9c**.

As for the unsuccessful nucleophiles, the charged side chains of Arg (basic) and Asp (acidic) prevented aminolysis from occurring. Accelerated rates of selenoester **3-3** hydrolysis was observed in the presence of Arg, faster than the controlled hydrolysis reactions (Figure 3-7). It was hypothesised that the highly basic guanidinium group may have disrupted the equilibrium balance of the buffered system for Arg. To provide a conclusion, a pH reading was taken at the end of the aminolysis trial: the pH had risen from 6.5-8.5 pH to 8.5-9.5 pH. As Arg has an isoelectric point of 10.7, the H^+ protons within the investigated buffered system are accepted by the double basic functionality of Arg (Figure 3-7), increasing the net concentration of OH^-. Thus, disrupting the buffer equilibrium, changing the pH and accelerating the hydrolysis of selenoester **3-3**.

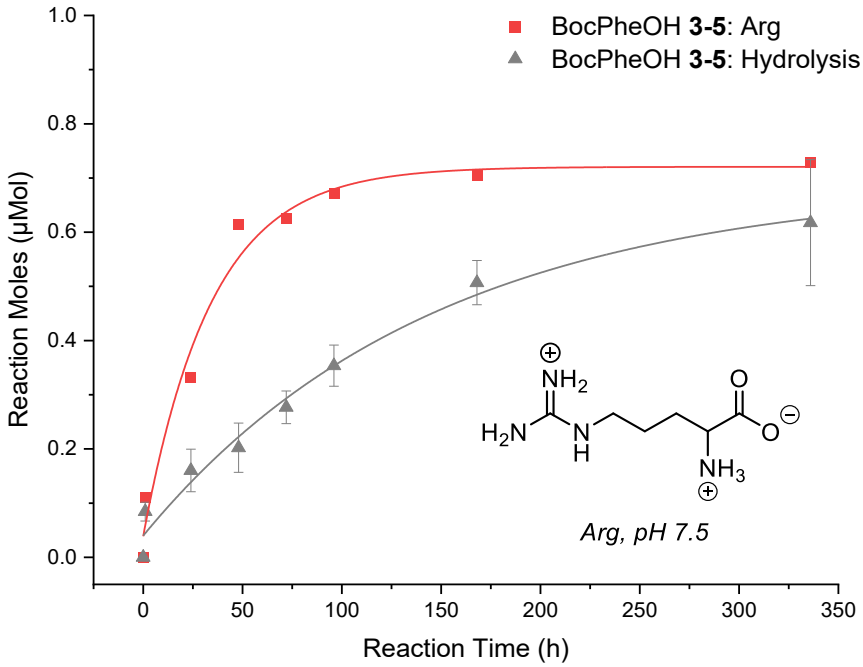


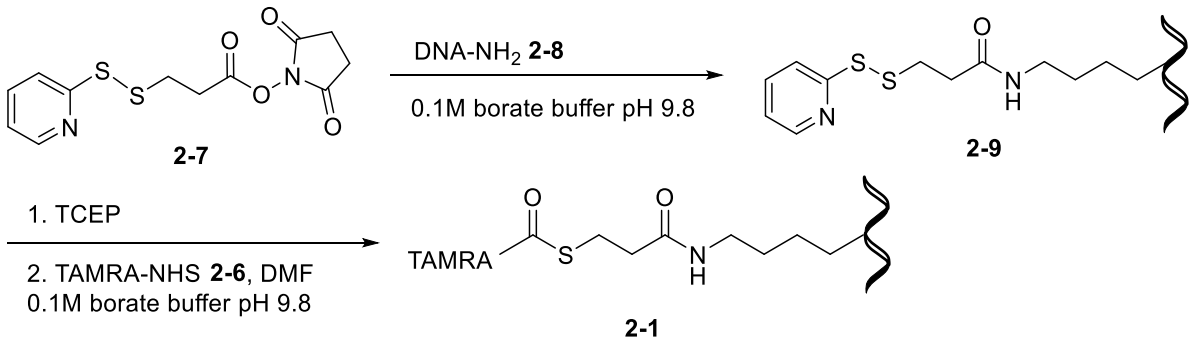
***Figure 3-7**: Amount (µMol) of Boc-Phe-OH **3-5** formed during aminolysis studies for Arg (red squares) with selenoester **3-3** (2 mM) at pH 7.5 over the course of 336 h. Boc-Phe-OH **3-5** formed during selenoester **3-3** hydrolysis studies (grey triangles). Structure of Arg at pH 7.5 is displayed. Error bars show the standard deviation in reaction moles of three repeats. Exponential fitting. Experimental section 3.4.6.*

Following the selenoester's success, selenoester **3-3** (400 μM) was subjected to the same reaction conditions, but with a modified DNA nucleophilic amine: ss-DNA-C_6-NH_2 **3-14** (4 μM), to determine whether the chemistries would be viable in a DNA-occupied system. However, the reaction yield was deemed negligible after 336 hours of incubation (3 %, HPLC integration). Because the selenoester moiety was not modified to a complementary ssDNA, the proximity chemistries of DTS were not utilised. Thus, indicating the importance of DTS architectures to access small reaction concentrations.

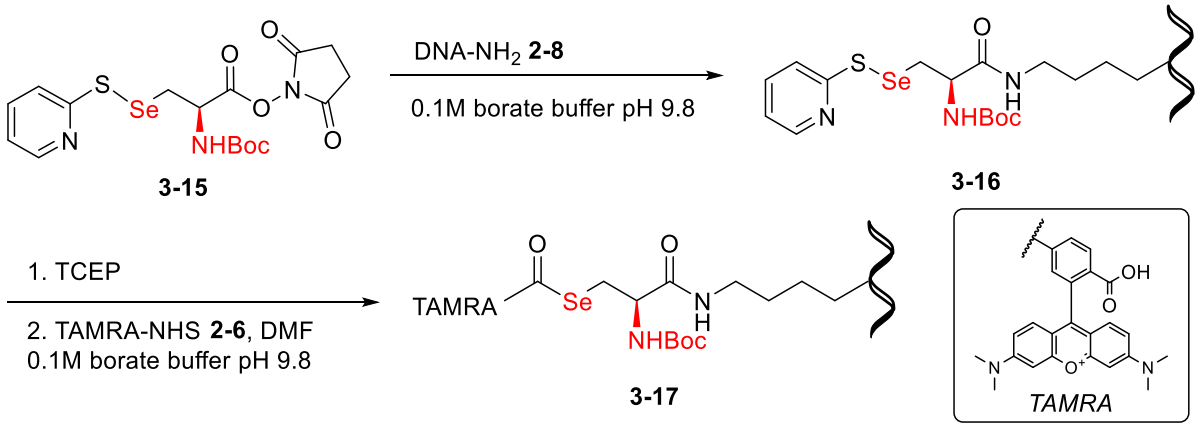
3.2.5 Synthetic design of selenoester-DNA complex

As selenoester **3-3** performed successful aminolysis in a DTS-compatible aqueous environment, it was hypothesised that single-step DTS aminolysis would be possible using selenoester-based chemistries. Following a similar synthesis to the group's established thioester-modified DTS adapter (Chapter 2), a synthetic procedure was drafted to create a DTS-feasible selenoester DNA adapter (Scheme 3-6).[1]

Thioester TAMRA DNA 2-1 (Chapter 2):



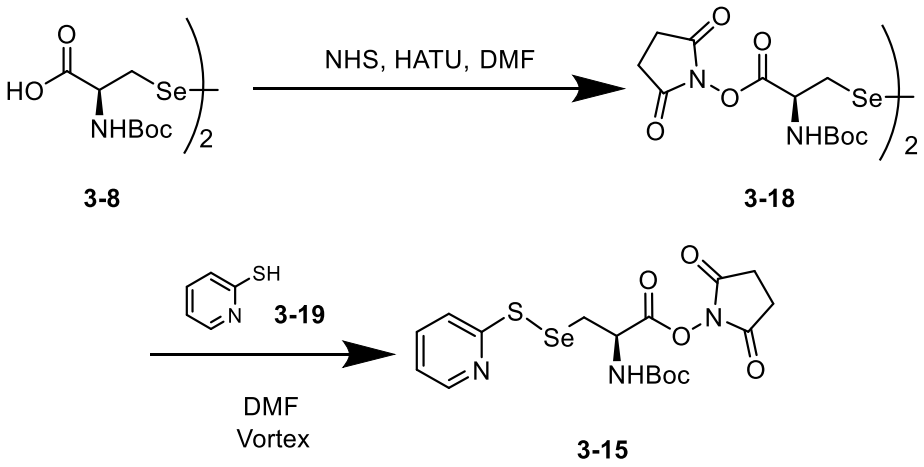
Selenoester TAMRA DNA 3-17:



***Scheme 3-6**: Initial synthetic procedure for TAMRA selenoester DNA **3-17** alongside established synthesis for TAMRA thioester DNA **2-1**. Structural differences are highlighted in red. TAMRA thioester DNA synthesis established by Frommer et al.*[1]

Having faced complexities previously in selenium-based chemistries, a simplified synthetic approach was drafted using the previously synthesised $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** (Scheme 3-6), leading to the incorporation of an NH-Boc group in **3-15** (Scheme 3-6 - red). Boc-protection was selected over Fmoc-protection due to potential deprotection in the highly basic DTS conditions used in the later assay pH 11.

In order to synthesise the initial S-Se bound small molecule **3-15**, $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** was activated using an NHS ester and the coupling agent HATU, followed by vortexing with 2-mercaptopyridine **3-19** (Scheme 3-7). The S-Se small molecule **3-15** was successfully synthesised and characterised using LC-MS.



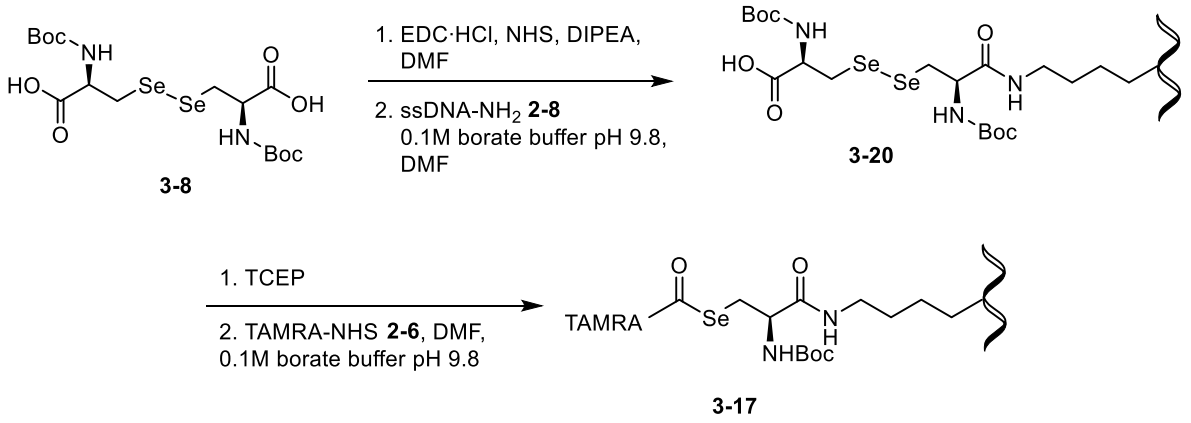
***Scheme 3-7**: Synthesis of S-Se small molecule **3-15**. Experimental section 3.4.10.*

Once purified, the small molecule **3-15** was subject to conjugation with DNA using the modified DNA-NH_2 **2-8** (Scheme 3-6). Following the same synthetic route as TAMRA thioester DNA **2-1**, the small molecule **3-15** (100 mM, DMF, 2 µL) was incubated with DNA-NH_2 **2-8** (1 mM, H_2O, 2 µL) in a pH 9.8 borate buffer (0.1 M) for 18 hours. However, upon LC-MS analysis, the desired product **3-16** was not formed, and the equivalence of **3-15** was increased (5 µL and 10 µL) to encourage product formation. Reassessing the synthetic procedure, the pH of the borate buffer was reduced to pH 8.5 and pH 7.5, to prevent degradation of **3-15′**s activated ester. Again, no product formation was observed, and syntheses were revised. It was hypothesised that the addition of the bulky Boc substituent next to the activated ester blocked the nucleophilic attack.

Subsequently, the direct attachment of (Boc-Sec-OSu)$_2$ **3-18** (500 mM, DMF, 2 µL) to DNA-NH_2 **2-8** (1 mM, H_2O, 2 µL) was attempted, again in the various borate buffers (0.1 M, pH 7.5, 8.5 and 9.8). In using (Boc-Sec-OSu)$_2$ **3-18**, the number of potential electrophilic sites increases from 1 to 2, therefore, increasing the possibility of DNA conjugation. However, this also proved unsuccessful. It was hypothesised that

the basic aqueous environment required for aminolysis caused the degradation of the activated ester. Therefore, an *in situ* activation and acyl transfer were proposed. A concentrated solution of $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** was subjected to 10 minutes of mixing with EDC, NHS and DIPEA, before being added to DNA-NH_2 **2-8** in a borate-buffered solution. After 18 hours, LC-MS analysis confirmed $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** attachment to DNA *via* an amide linkage within a borate buffer solution at pH 9.8, yielding TAMRA selenoester DNA **3-17** (Scheme 3-8, Supplementary Figure 3-35).

TAMRA selenoester DNA 3-17:



***Scheme 3-8**: Synthesis of TAMRA selenoester DNA **3-17**. Experimental section 3.4.11.*

As a diselenide was already present on Boc-Sec DNA **3-20**, rather than continuing with the originally proposed synthesis of S-Se bond formation with 2-mercaptopyridine **3-19** (Scheme 3-7), the diselenide bond of DNA modification **3-20** was reduced independently, and TAMRA-NHS **2-6** was introduced (Scheme 3-8). Introducing 2-mercaptopyridine **3-19** into the system, to create the S-Se bond, would have created unnecessary complexity in the synthesis of TAMRA selenoester DNA **3-17**. Following the second step in the synthesis of TAMRA thioester DNA **2-1**, the Boc-Sec DNA **3-20** was incubated with TCEP (0.1 M, H_2O) for 10 minutes to reduce

the diselenide bond, followed by the addition of TAMRA-NHS **2-6**. After 18 hours of mixing, the desired TAMRA selenoester DNA **3-17** product was successfully synthesised, purified and characterised using LC-MS and UV-Vis spectroscopy (Figure 3-8). Because of the small reaction scale used during DNA-modification synthesis (10 nmol reaction scale), LC-MS was the most suitable characterisation technique to detect successful product formation at such low abundance. Complementing this, UV-Vis analysis on a RP-HPLC indicated both a DNA absorbance band (258 nm) and a TAMRA absorbance band (560 nm) at the corresponding TAMRA selenoester DNA **3-17** signal (Figure 3-8B/C).

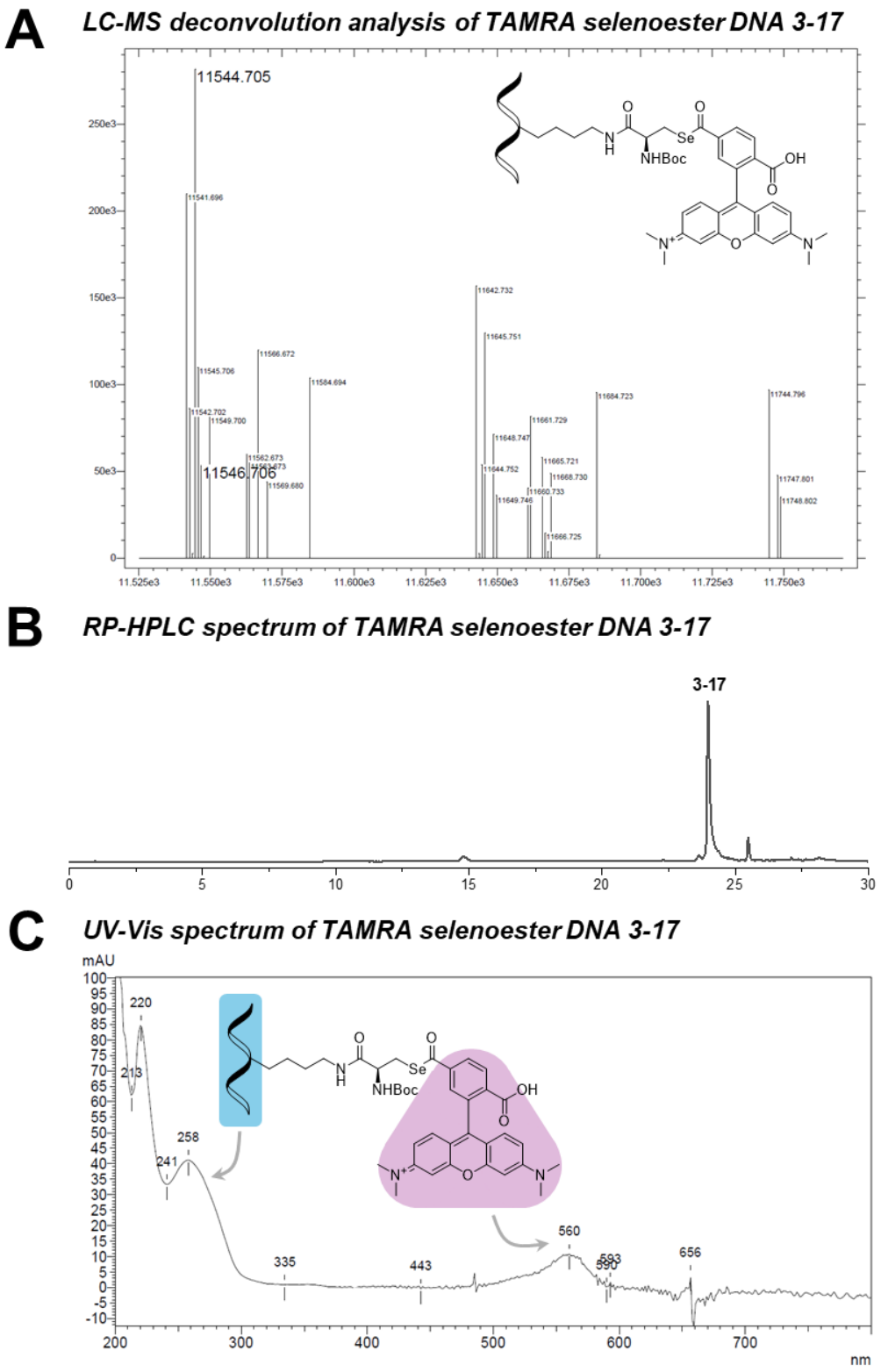


Figure 3-8: *Characterisation of TAMRA selenoester DNA* ***3-17*** ***A****) LC-MS deconvolution analysis (Waters Xevo-G2-XS). [M]^0 m/z calculated: 11546.06 found: 11544.71. Additional peaks are due to the six natural selenium isotopes.* ***B****) RP-HPLC spectrum* ***C****) UV-Vis spectrum collected from TAMRA selenoester DNA* ***3-17*** *corresponding signal on RP-HPLC, 23.91 min. DNA region corresponds to 258 nm signal, the TAMRA region corresponds to 560 nm signal.*

3.2.6 DNA-templated synthesis assay

Following the protocol described in Chapter 2, TAMRA selenoester DNA **3-17** was installed in the newly developed across-the-helix DTS architecture, replacing TAMRA thioester DNA **2-1**.[1] Successful aminolysis between the electrophilic selenoester **3-17** and nucleophilic primary amine **2-2** would promote the transfer of the TAMRA moiety by forming an amide bond. Because of its fluorescent capabilities, the change in the TAMRA environment can be monitored *via* the RP-HPLC fluorescence channel of λ = 550 nm. In addition, competing hydrolysis would be observed by the production of the small molecule, free TAMRA **2-5** (Figure 3-9).

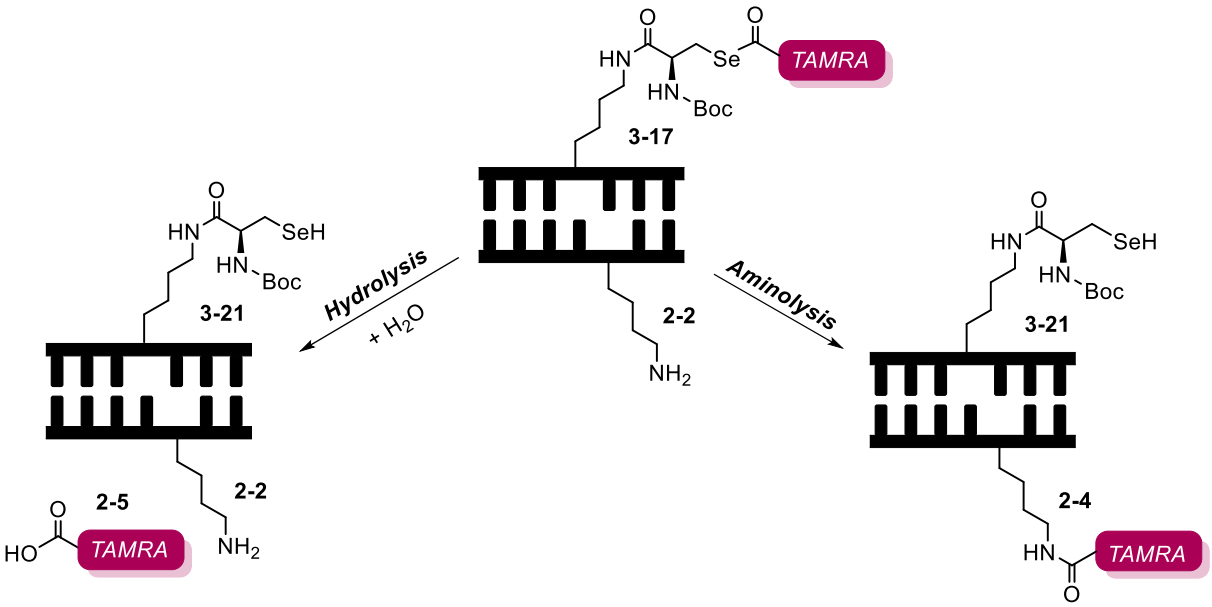


***Figure 3-9**: Mechanisms of action: DTS aminolysis forming an amide bond **2-4** (right) or hydrolysis (left) of TAMRA selenoester DNA **3-17** forming **2-5** at the +1 NH_2 position **2-2**. Additional spent selenol adapter **3-21** was undetectable by RP-HPLC.*

By utilising an across-the-helix architecture, the primary amine **2-2** position can vary across the complementary DNA strand, replacing the corresponding nucleobase. For example, the +1 NH_2 **2-2** position is one nucleobase apart, across the major groove, from the opposing TAMRA selenoester DNA **3-17** position.

To test the hypothesis that selenoesters would improve DTS aminolysis yields, TAMRA selenoester **DNA 3-17** was subject to the DTS assay designed by Frommer *et al.* and described in Chapter 2.[1] Both ssDNA strands **3-17** and **2-2** at the +1 position were incubated together for 20 minutes, in magnesium chloride buffered solution, to induce dsDNA hybridisation. The mixture was diluted in the DTS buffered solution, (pH 11) to a total volume of 30 µL, and injected directly into the RP-HPLC, in 2 µL aliquots over the course of 12 hours. The newly modified TAMRA selenoester DNA **3-17** aminolysis ability was compared to that of the previously established TAMRA thioester DNA **2-1**.[1]

As predicted, the increase in selenoester's electrophilicity improved aminolysis DTS yields (Figure 3-10). Following previous trends in chalcogen ester performance, the TAMRA selenoester DNA **3-17** successfully increased the amide **2-4** yield by 9% in comparison to TAMRA thioester DNA **2-1** (68% to 77%), after 12 hours of incubation. Interestingly, faster reaction rates were observed for TAMRA selenoester DNA **3-17** – with a 62% amide **2-4** yield after 3 hours of incubation, in comparison to 39% amide **2-4** yield for TAMRA thioester DNA **2-1** (Figure 3-10). The improved reaction rates demonstrated here reduce the chance of hydrolysis of reactive species by reducing necessary incubation times. The increase in electrophilic capabilities of selenoesters provides an advantage in multistep DTS systems by reducing the chance of product truncation, which would lead to undesirable conjugation in DTS polymer synthesis.

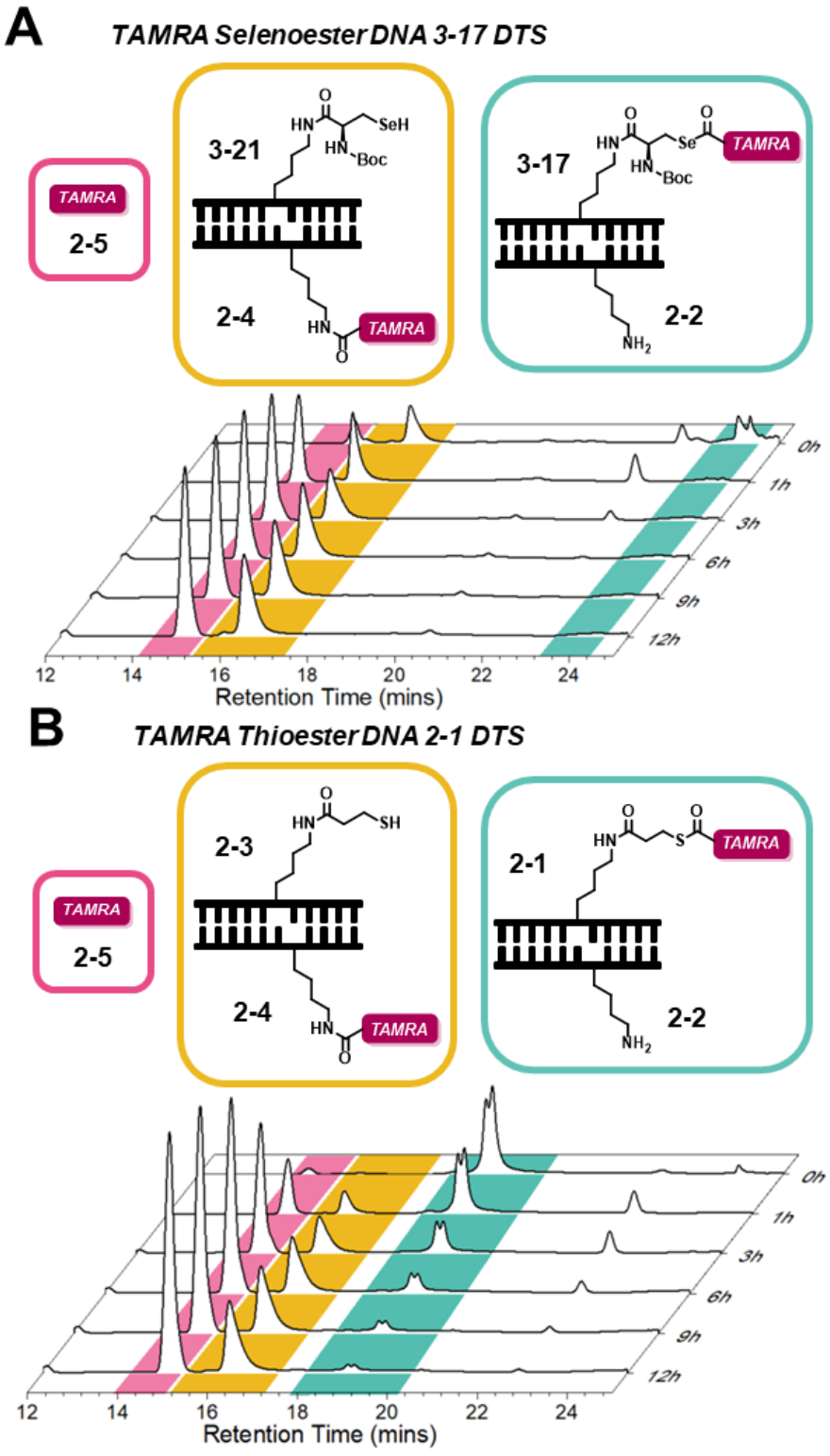


Figure 3-10: *Waterfall plot of identical RP-HPLC conditions of **A**) TAMRA selenoester DNA **3-17** and **B**) TAMRA thioester DNA **2-1** initiated DTS in DTS buffer, pH 11.0, over the course of 12 hours. Fluorescence channel: (λ_{ex} 550 nm, λ_{em} 580 nm). 2 µL injection volume. Pink = free TAMRA **2-5** (hydrolysis of **3-17** or **2-1**). Yellow = successful aminolysis (amide product **2-4** with spent adapter **3-21** or **2-3**). Green = starting material (TAMRA ester DNA **3-17** or **2-1** with primary amine **2-2**). Experimental section 3.4.12.*

Previous work within the group investigated the different amine positions **2-2** across the opposing DNA strand in the across-the-helix architecture (Table 3-3).[1] As described in Chapter 2, the stabilisation of the thioester moiety in **2-1** was observed across the minor groove (−*n*), with successful aminolysis across the major groove (+*n*).

Table 3-3: *Example of primary amine location across DNA helix (**2-2**) in the Across-the-Helix architecture. **X** = site of NH_2 modification DNA sequence (IDT code: iUniAmM)*

Position/Code	Groove	DNA Sequence
−3 **2-2**	Minor (−)	CAGGTAGATTAGGACAGC/**X**/CTGACTGCTGGCTCGGC
+3 **2-2**	Major (+)	CAGGTAGATTAG/**X**/ACAGCGCTGACTGCTGGCTCGGC

Following the same DTS assay as before, positions +5 to −3 of DNA-NH_2 **2-2** were investigated as aminolysis sites of the primary amine with TAMRA selenoester DNA **3-17**. It was hypothesised that an increase in yield could be observed across the different sites due to the increase in selenoester's electrophilicity. Table 3-4 displays the aminolysis yields of TAMRA selenoester DNA **3-17** with varying positions of the primary amine nucleophile **2-2**. Analysed on the fluorescence channel (λ_{ex} 550 nm, λ_{em} 580 nm), TAMRA **2-5** (hydrolysis) and amide product **2-4** (aminolysis) yields were calculated using the integrated intensity of the corresponding RP-HPLC signal and compared against calibration curves.

Table 3-4: *Yield (30 pmol / %) of free TAMRA* ***2-5*** *and amide product* ***2-4*** *from aminolysis DTS reaction of TAMRA selenoester DNA* ***3-17*** *and DNA-NH_2* ***2-2*** *at varying site locations. Comparison with TAMRA thioester DNA* ***2-1*** *under the same conditions, position +1* ***2-2****. Values taken at 12 hours, pH 11 DTS buffer, 2 µL injection volume. Error equates to the standard deviation of three repeats. Experimental section 3.4.12.*

	TAMRA 2-5 produced at 12 h		***Amide 2-4 produced at 12 h***	
Position of 2-2	***pmol (yield %)***	***σ / pmol***	***pmol (yield %)***	***σ / pmol***
− 3	11.7 (39)	2.1	18.4 (61)	4.0
− 2	11.2 (37)	3.3	18.8 (63)	3.9
− 1	10.3 (34)	0.5	19.7 (66)	2.2
0	9.6 (32)	2.7	20.4 (68)	4.5
+ 1	6.9 (23)	0.8	23.1 (77)	3.2
+ 2	12.9 (43)	4.4	17.1 (57)	4.4
+ 3	11.2 (37)	1.7	18.8 (63)	1.4
+ 5	13.1 (44)	3.3	16.9 (56)	1.9
2-1 *with + 1*	9.6 (32)	1.4	20.4 (68)	3.9

Statistically similar aminolysis yields were observed across all investigated positions of **2-2** (56 – 77 % yield), along both the major and minor grooves of DNA in the Across-the-Helix architecture for TAMRA selenoester DNA **3-17**. Despite sterics impacting the performance of selenoesters in previous work, the major/minor groove position of the primary amine **2-2** with **3-17** in the DNA helix did not impact **2-4** product yields (Figure 3-11). However, a small position dependency was observed when the distance between nucleophile and electrophile increased (*i.e.* position +5 **2-2** yielded 56% amide **2-4**, as opposed to position 0 **2-2** yielded 68% amide **2-4**, Table 3-4). Although their yields differed across the minor groove of DNA, both TAMRA thioester DNA **2-1** and TAMRA selenoester DNA **3-17** presented the highest yield at the **2-2** +1 position. Indicating the optimum distance between the reactive tags was 1.75 nm.[1] At all investigated positions, the reaction proceeded to completion within 6 hours, as

evidenced by the consumption of starting material (Figure 3-10). In comparison to the thioester analogue, the incorporation of a selenoester reduced reaction times by half (12 h with **2-1** vs 6 h with **3-17**, Figure 3-10). As explained earlier, fast reaction times would prevent truncation and aid in sequence-controlled DTS polymer growth.[40]

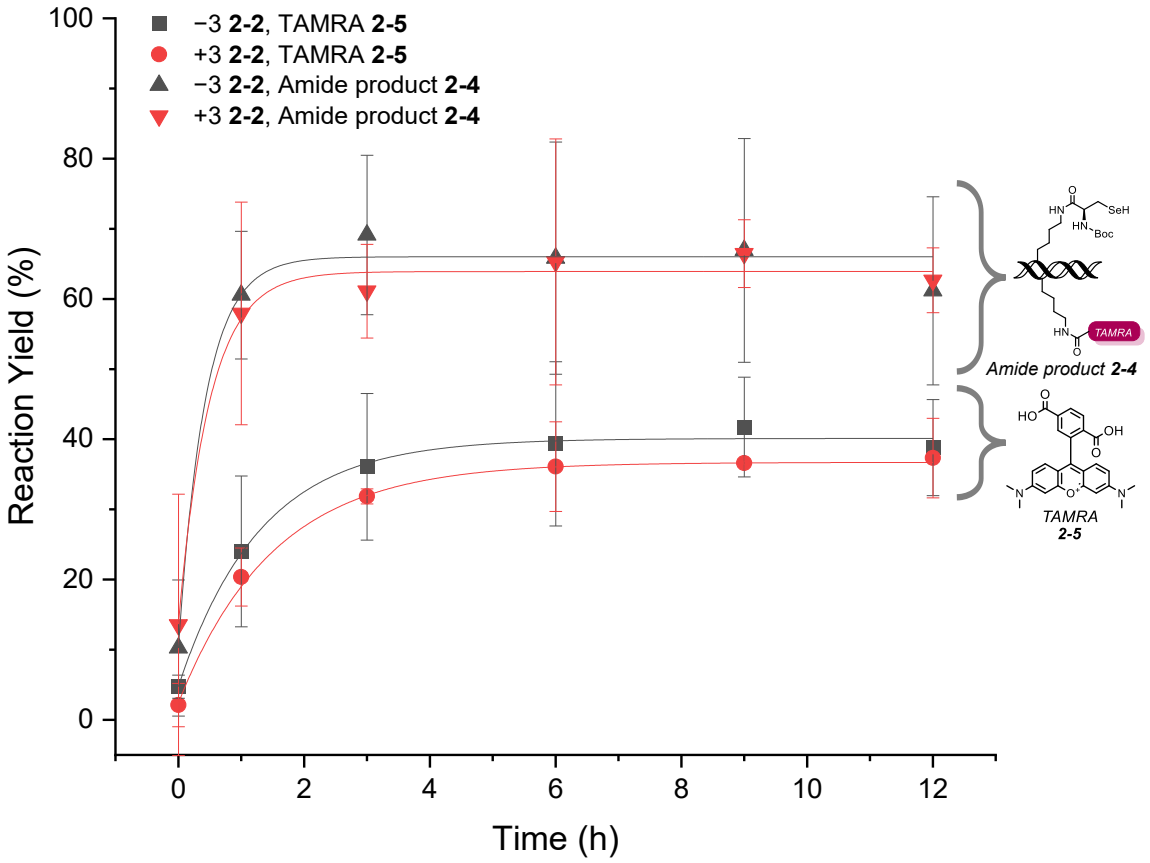


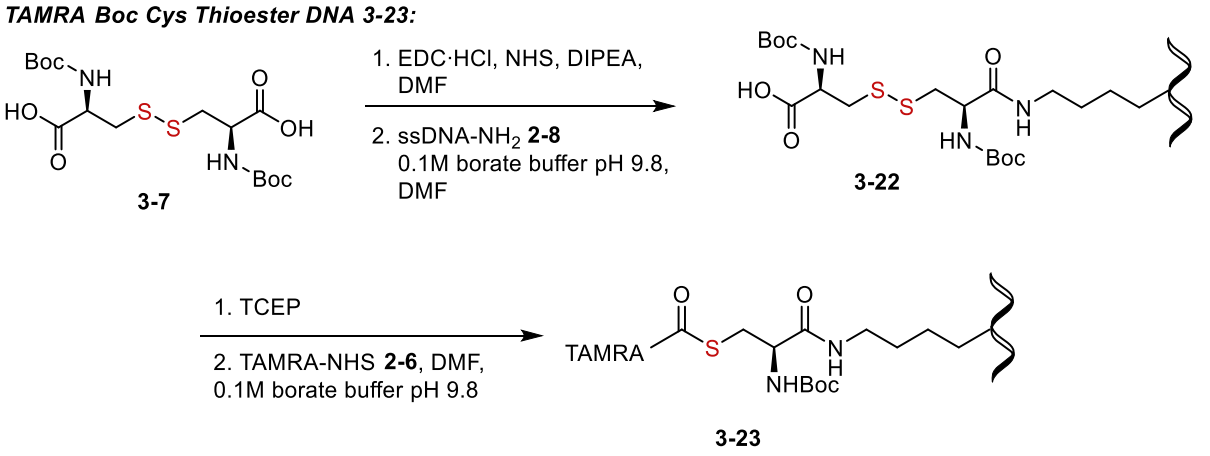
***Figure 3-11**: Example DTS aminolysis yield (%) progression over 12 hours of TAMRA selenoester DNA **3-17** with DNA-NH_2 **2-2** positions −3 (black) and +3 (red). Circle and square = amide product yield **2-4**. Triangles = free TAMRA **2-5** (hydrolysis). Error bars show the standard deviation in the reaction yield of three repeats. Experimental section 3.4.12.*

Conversely, the nucleophile **2-2** position impacted product formation using TAMRA thioester DNA **2-1**, which did not display successful aminolysis across the minor groove positions, but a thioester stabilisation, as described in Chapter 2.[1] Thus, indicating that selenoesters as electrophiles are less sensitive to the DNA architecture.

3.2.7 DNA-templated synthesis confirmation assay

A Boc-protected amine was incorporated into the syntheses of TAMRA selenoester DNA **3-17** to simplify the synthetic protocol. To confirm that the increase in DTS yield was initiated by the selenoester of **3-17**, rather than the steric change of the Boc-protected amine, a Boc group was incorporated into the original TAMRA thioester DNA **2-1** modification. Concerned by the impact of the Boc functionality on the DTS yield, a new TAMRA Boc-Cys thioester DNA **3-23** was synthesised.

Following the same synthetic procedure for TAMRA selenoester DNA **3-17**, $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** was replaced with $(Boc\text{-}Cys\text{-}OH)_2$ **3-7** to yield TAMRA Boc-Cys thioester DNA **3-23** (Scheme 3-9).



***Scheme 3-9**: Synthesis of TAMRA Boc-Cys thioester DNA **3-23**. Structural changes from TAMRA selenoester DNA **3-17** highlighted in red. Experimental section 3.4.13.*

TAMRA Boc-Cys thioester DNA **3-23** was subjected to the same DTS conditions as previously described, with $DNA\text{-}NH_2$ **2-2** at the +1 position. After 12 hours of incubation at pH 11.0, DTS proved to be successful, however, at a significantly smaller yield than previous designs (Figure 3-12). Such an assay proved that the fast reactivity observed for TAMRA selenoester DNA **3-17** was a consequence of the selenoester

moiety rather than the additional Boc-protected amine. The involvement of the Boc-protected amine decreased aminolysis yields from 68% to 20% for the thioester (**2-1** and **3-23**, respectively). It was hypothesised that the removal of the bulky Boc-protected amine residing on TAMRA selenoester DNA **3-17** may have increased yields further.

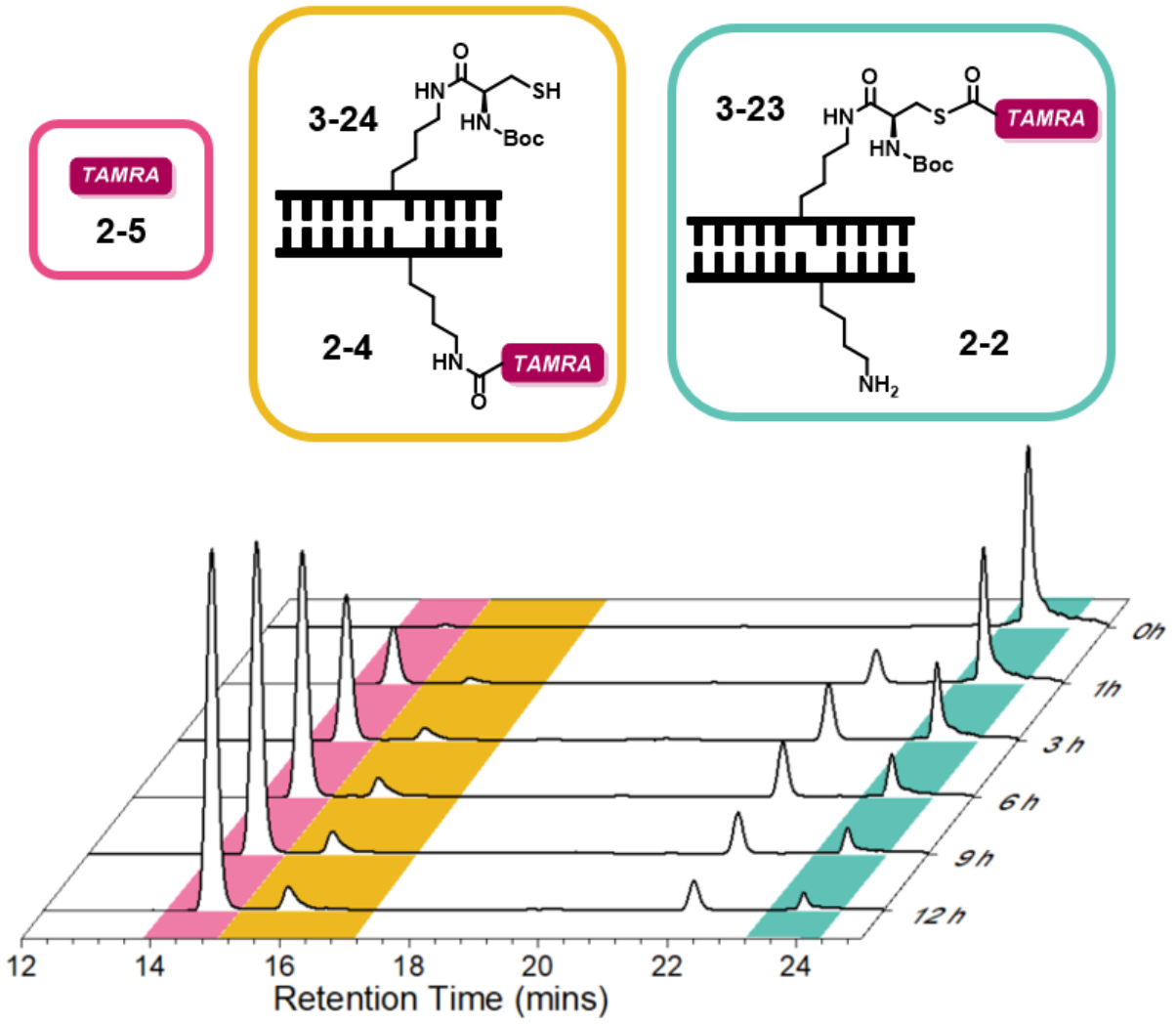


***Figure 3-12**: Waterfall plot of RP-HPLC spectrum of TAMRA Boc-Cys thioester DNA **3-23** initiated DTS in DTS buffer, pH 11.0, over the course of 12 hours. Fluorescence channel: (λ_{ex} 550 nm, λ_{em} 580 nm). 2 µL injection volume. Pink = free TAMRA **2-5** (hydrolysis of **3-23**). Yellow = successful aminolysis (amide product **2-4** with spent adapter **3-24**). Green = starting material (TAMRA Boc-Cys thioester DNA **3-23** with primary amine **2-2**). Experimental section 3.4.12.*

3.2.8 DNA-templated synthesis protection assay

As mentioned previously, TAMRA selenoester DNA **3-17** demonstrated successful aminolysis across the minor groove, which was not observed for TAMRA thioester DNA **2-1**, but rather a stabilisation (protective) effect, limiting both aminolysis and hydrolysis. Previous work hypothesised that combining an abasic site with a hydrophobic, planar TAMRA modification aided in the insertion of the TAMRA thioester DNA **2-1** modification into the hydrophobic dsDNA core of the across-the-helix architecture (Chapter 2).[1] This observed protective effect displayed the greatest stabilisation at DNA-NH_2 **2-2** position −3. To understand why aminolysis occurred at this position for TAMRA selenoester DNA **3-17**, the novel DNA modifications **3-17** and **3-23** were subject to an 18 h incubation protection assay. Herein, the DTS nucleophile, NH_2 (IDT code = iUniAmM) **2-2** at position −3, was replaced with a spacer abasic site **2-10** (IDT code = iSpC3) (Figure 3-13). The slight change in modification allowed for the protective effect to be investigated in isolation from the aminolysis reaction.

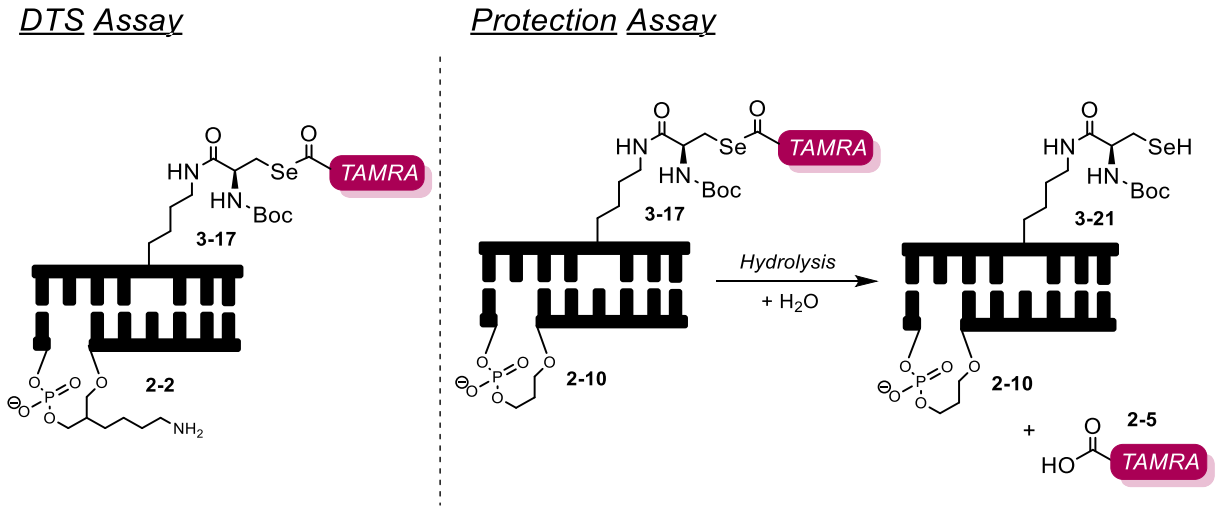


***Figure 3-13**: Structural differences between DTS assay (NH_2 **2-2**, left) and protection assay (spacer abasic site **2-10**, right). The degree of protection was calculated by the amount of free TAMRA **2-5** produced (better protection = less free TAMRA **2-5**).*

The protection assay followed the same protocol as the DTS assay (Experimental section 3.4.14). However, reaction times were extended to 18 hours, and DNA-NH_2 **2-2** was replaced with spacer-DNA **2-10** as hydrolysis rates were slower than aminolysis rates (see Chapter 2). Modifications **2-1, 3-17** and **3-23** were incubated with DNA-spacer **2-10** at position −3 independently and compared against ssDNA TAMRA selenoester DNA **3-17** in the DTS buffer (Figure 3-14). Any differences observed in stability would determine the dictating moiety of the protection observed by Frommer *et al.* and in Chapter 2.[1]

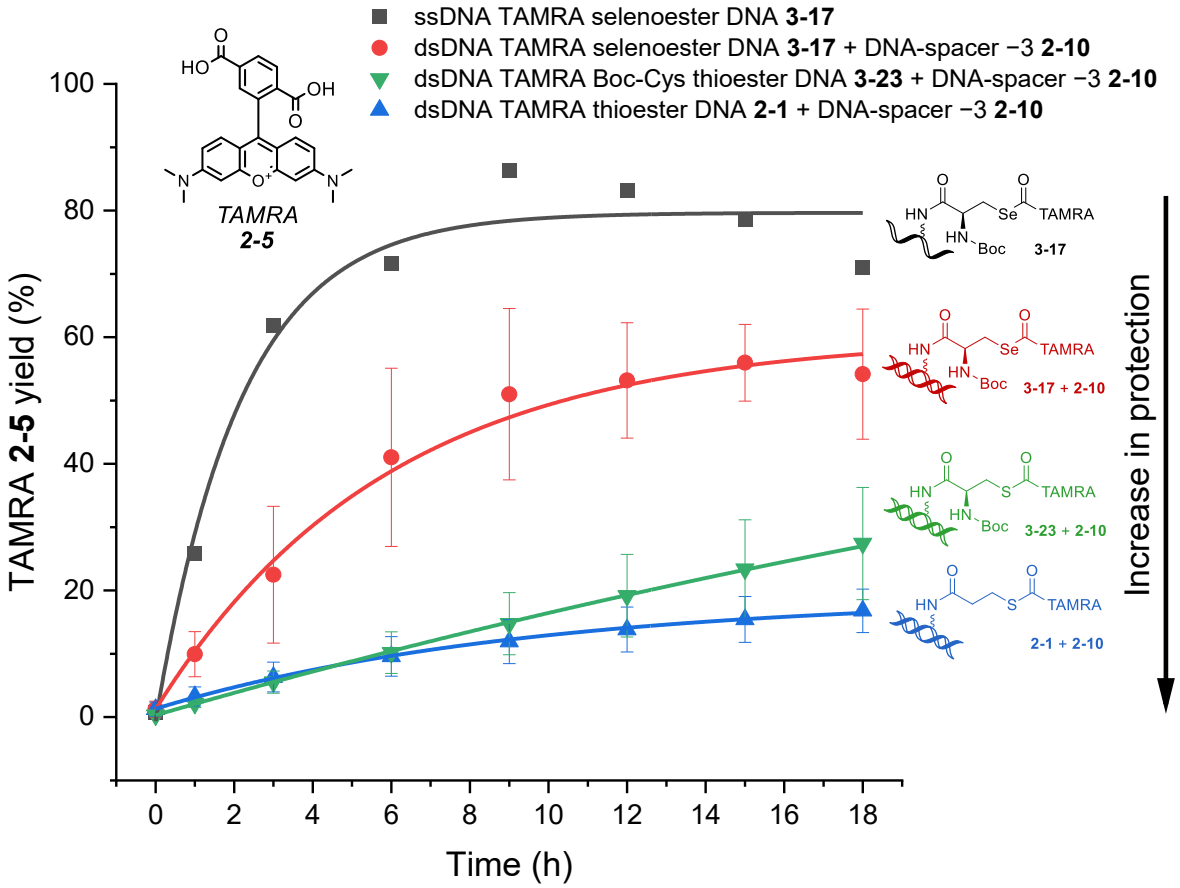


***Figure 3-14**: Free TAMRA **2-5** production during the protection assays (pH 11) of TAMRA selenoester DNA **3-17** with spacer-DNA **2-10** −3 position (red circle), TAMRA thioester DNA **2-1** + spacer-DNA −3 **2-10** (blue up triangle) and TAMRA Boc-Cys thioester DNA **3-23** + spacer-DNA −3 **2-10** (green down triangle) over 18 hours. Stability assay of ssDNA strand TAMRA selenoester DNA **3-17** (black square). Error bars show the standard deviation in TAMRA **2-5** yield of three repeats. Experimental section 3.4.14.*

Similar levels of stabilisation were observed for both TAMRA thioester DNA **2-1** and TAMRA Boc-Cys thioester DNA **3-23** in the protection assay (Figure 3-14 – blue and red). However, a slight increase in hydrolysis was observed with the addition of the NH-Boc modification in TAMRA Boc-Cys thioester DNA **3-23,** but it did not limit the stabilisation of the thioester. Such findings suggest that the protection effect is contributed by the thioester moiety, by interacting with either the TAMRA functionality or the DNA helix.

On the other hand, TAMRA selenoester DNA **3-17** did not display a protective effect and proceeded to degrade in the basic conditions (54% TAMRA **2-5**, 18 h, Figure 3-14 - red). Such results suggest the atomic differences between sulfur and selenium interrupt the stabilisation presented by the spacer-abasic site across the minor groove. Following previous steric assumptions and conclusions drawn in Chapter 2, selenium is atomically larger than sulfur and insertion within the hydrophobic DNA core may have been prevented. However, this effect would have been observed with TAMRA Boc-Cys thioester DNA **2-23** due to the bulky NH-Boc group. Alternatively, selenium may disrupt the intermolecular forces between DNA and the TAMRA modification, preventing the crucial positioning that aids in stabilisation of the thioester analogue. Supported by the limited literature, it is unlikely that selenium interferes with the stacking of the DNA duplex.[41, 42]. Therefore, it was concluded that the thioester within TAMRA thioester DNA **2-1** is crucial in promoting the stabilisation effect presented by Frommer *et al.*, whilst also contributing to the findings discussed in Chapter 2.[1] However, to date, there is currently limited literature on thioester-TAMRA and thioester-DNA specific interactions.

Furthermore, the stabilisation effect observed between ssDNA and dsDNA TAMRA selenoester DNA **3-17** is induced by the formation of a double helix (Figure 3-14 – black and red).[43] ssDNA is thermodynamically less stable than dsDNA, due to its high flexibility, and any residing modifications are more vulnerable to hydrolysis.[44] Therefore, unfortunately, the protection effect cannot be utilised in any selenoester-mediated DTS reactions using the across-the-helix architecture.

3.3 Conclusions

This chapter of work set to investigate the compatibility of selenoesters as electrophiles within an acyl transfer DTS system. In accordance with literature reports, an initial set of small-molecule aminolysis assays of chalcogen esters proved that selenoesters excel as electrophiles in aminolysis reactions. Within DTS-compatible aqueous conditions, a screening of chalcogen esters with various amino acid nucleophiles demonstrated the superiority of selenoesters over thio- and oxo- esters. Because of the enhanced leaving group ability of selenolates, dictated by poor orbital overlap and reduced resonance delocalisation, selenoesters provided a suitable substitute to the previously established thioester-mediated DTS reaction within the group.

After the promising results during the monomer assays, a selenoester was successfully conjugated to a modified ssDNA. In comparison to the established thioester analogue, the selenoester modification improved DTS aminolysis yields by 9% and reduced reaction times from 12 hours to < 3 hours, in an across-the-helix architecture at pH 11. In contrast to the thioester analogue, aminolysis was achieved across both the major and minor grooves of the across-the-helix architecture, demonstrating selenoesters' superiority as an electrophile in this architecture. However, a small trend in position dependency was observed by increasing the number of nucleobases between the chemical tags.

Confirming the improvement in DTS yield was initiated by the selenoester moiety rather than the additional Boc-protected amine, DTS yields decreased from 68% to 20% when a Boc group was adjacent to the DNA-thioester modification. Future

work may progress acyl transfer yields further by removing the Boc-protected amine of the investigated selenoester modification, supported by the hinderance in thioester reactivity.

Concluded by a stability assay, the cause for the selenoester's ability to perform DTS across both the major and minor grooves was due to the selenium atom itself. Aiding in conclusions to work performed in Chapter 2, the addition of an NH-Boc group to the thioester DNA modification did not impact the thioester stabilisation effect. Because of the increase in modification size, it was concluded that the sterics did not impact the protection effect as significantly as previously believed, but were rather dictated by intermolecular forces between the TAMRA, thioester and DNA.

Overall, this chapter of work documented the first use of selenoesters as a transfer-DTS electrophile. As the selenoester proves to be the most promising candidate in improving reaction kinetics and yields in a single-step DTS mechanism, multi-step DTS for peptide growth may be achieved, in the future, if the electrophilic moiety is incorporated. Improving reactivity and limiting hydrolysis within the sequence-controlled mechanism would aid in the directed evolution of oligomers when parameters are controlled and optimised.

3.4 Experimental

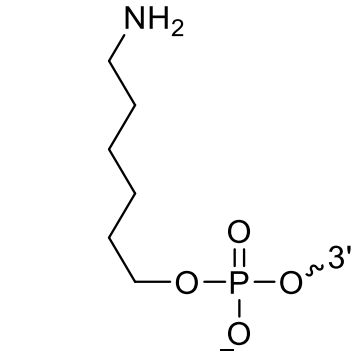
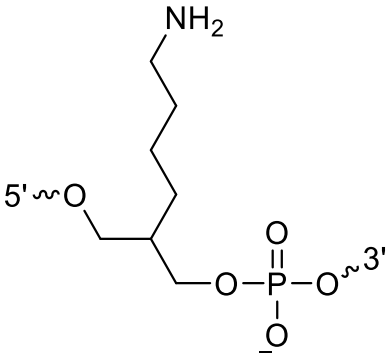
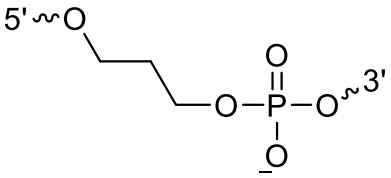
For details of the RP-HPLC, LC-MS, NMR and IR techniques, see Appendix 1: Experimental Methods.

3.4.1 Materials

All small molecules were purchased from Sigma, other than Selenocysteine **3-10**, TAMRA-NHS (6-TAMRA-SE) **2-6** or 6-TAMRA **2-5**, which were purchased from Cambridge Bioscience. PBS sachets were purchased from Sigma and made up using 18.2 MΩ H_2O. For the preliminary aminolysis reactions, the pH of PBS was adjusted to 6.5-10.5 pH using either 1 M HCl or 1 M NaOH. For the synthesis of DNA modifications, borate buffer was formulated using boric acid (Sigma), diluted with 18.2 MΩ H_2O and adjusted to pH 9.8 with 1 M NaOH. See *Chapter 2 Experimental 2.4.1 Materials* for DTS buffer formulation.

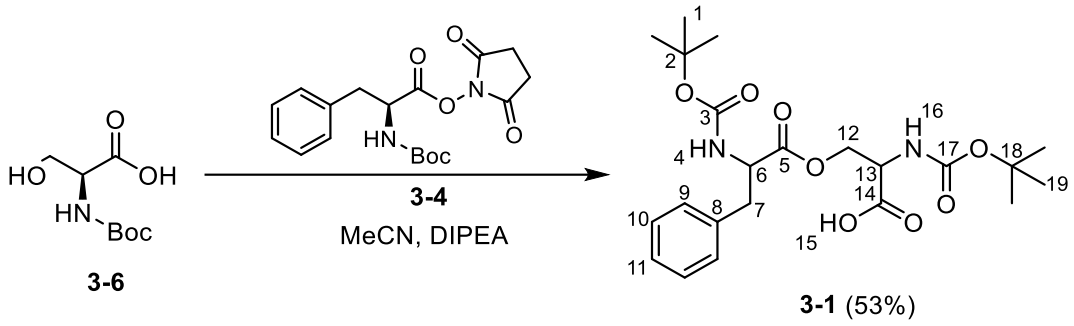
DNA sequences were purchased from Integrated DNA Technologies (IDT), with either standard desalting or HPLC purification. All samples were suspended in water (18.2 MΩ) at a concentration of 1 mM, 0.2 mM or 0.1 mM.

***Table 3-5**: DNA modifications and sequences*

IDT Key	Description	Modification
/5AmMC6/	5′ Amino Modifier C6	 */5AmMC6/*
/iUniAmM/	internal Uni-Link™ Amino Modifier	 */iUniAmM/*
/iSpC3/	Int C3 Spacer	 */iSpC3/*

Identification		Sequence 5′ → 3′
ssDNA-C6-NH_2 **3-14**		/**5AmMC6**/TCAACTTCAACACGCCCTTT
Original Donor (ssDNA-NH_2 **2-8**)		GCCGAGCCAGCAGTCAGCGC/**iUniAmM**/GTCCTAATCTACCTG
Acceptor	**2-2** Position -3	CAGGTAGATTAGGACAGC/**iUniAmM**/CTGACTGCTGGCTCGGC
	2-2 Position -2	CAGGTAGATTAGGACAG/**iUniAmM**/GCTGACTGCTGGCTCGGC
	2-2 Position -1	CAGGTAGATTAGGACA/**iUniAmM**/CGCTGACTGCTGGCTCGGC
	2-2 Position 0	CAGGTAGATTAGGAC/**iUniAmM**/GCGCTGACTGCTGGCTCGGC
	2-2 Position +1	CAGGTAGATTAGGA/**iUniAmM**/AGCGCTGACTGCTGGCTCGGC
	2-2 Position +2	CAGGTAGATTAGG/**iUniAmM**/CAGCGCTGACTGCTGGCTCGGC
	2-2 Position +3	CAGGTAGATTAG/**iUniAmM**/ACAGCGCTGACTGCTGGCTCGGC
	2-2 Position +5	CAGGTAGATT/**iUniAmM**/GGACAGCGCTGACTGCTGGCTCGGC
	2-10 Spacer -3	CAGGTAGATTAGGACAGC/**iSpC3**/CTGACTGCTGGCTCGGC

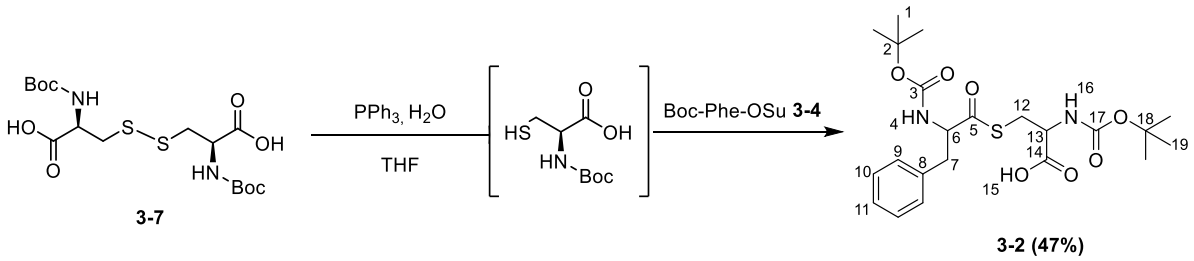
3.4.2 Synthesis of oxoester 3-1; *N*-(Boc)-*O*-((Boc)-*l*-phenylalanyl)-*l*-serine



Under an inert atmosphere, Boc-Ser-OH **3-6** (5.00 g, 2.44 mmol) and Boc-Phe-OSu **3-4** (4.41 g, 1.22 mmol) were added to acetonitrile (20 mL), and the reaction mixture was stirred. DIPEA (12 mL) was added, and the reaction was stirred at room temperature. After 72 hours, the mixture was dissolved in EtOAc (150 mL) and washed with 1 M hydrochloric acid (200 mL x 3). The organic layer was dried over anhydrous $MgSO_4$. The solution was filtered and concentrated under reduced pressure. The crude mixture was subjected to Prep-HPLC with a mobile phase of 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B) (Solvent A: 35 – 45%) at a flow rate of 0.5 mL min^{-1}, eluting at 8.0 minutes on a Shim-pack GISS C18, 4.6 x 125 mm, 5 µm column. Oxoester **3-1** was isolated as a sticky, viscous, pale yellow oil (0.29 g, 53%). **TLC** (EtOAc / 100%) R_f = 0.10. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.93 (s, 1H, H_{15}), 7.32-7.18 (m, 5H, $H_{9\text{-}11}$), 7.11 (d, 1H, *J* = 8.7 Hz, H_4), 7.09 (d, 1H, *J* = 8.3 Hz, H_{16}), 4.37-4.09 (m, 4H, H_6, H_{12}, H_{13}), 3.01 (dd, 1H, *J* = 13.9, 4.8 Hz, H_7), 2.81 (dd, 1H, *J* = 14.0,10.4 Hz, H_7), 1.38 (s, 9H, H_1), 1.32 (s, 9H, H_{19}). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 172.1 (C_5), 171.5 (C_{14}), 155.9 (C_3, C_{17}), 138.5 (C_8), 129.5 (C_9), 128.7 (C_{10}), 126.92 (C_{11}), 78.9 (C_2), 78.8 (C_{18}), 64.4 (C_{12}), 55.7 (C_6), 52.9 (C_{13}), 36.9 (C_7), 28.6 (C_1, C_{19}). **TOF MS**: (ES-) *m/z* $[M\text{-}H]^-$ calculated: 451.22, found 451.21. **IR** (ν_{max} / cm^{-1}) 3300 (O-H), 2950-2920 (N-H stretch, C-H

phenyl), 1691 (C=O stretch, ester), 1500 (C=O stretch, Boc). See Supplementary Figure 3-1:3-6 for characterisation spectra.

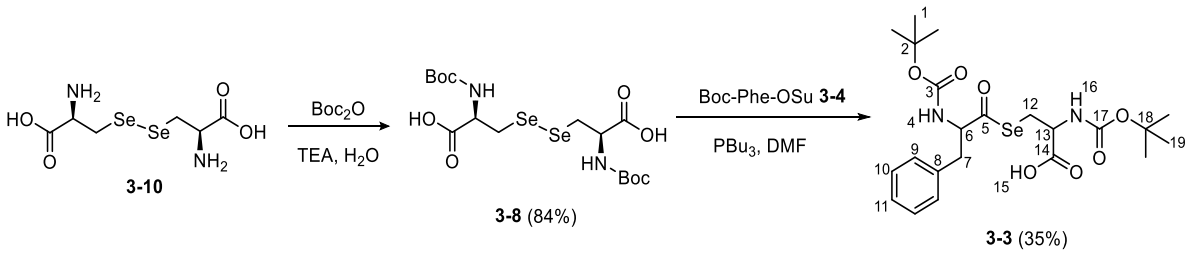
3.4.3 Synthesis of thioester 3-2; *N*-(Boc)-*S*-((Boc)-*l*-phenylalanyl)-*l*-cysteine



Under an inert atmosphere and in a brown reaction vessel, $(Boc\text{-}Cys\text{-}OH)_2$ **3-7** (1.00 g, 2.27 mmol) was added to degassed THF (15 mL), and the mixture was stirred until dissolved. The mixture was cooled to 0 °C in an ice bath, and triphenylphosphine (0.7 g, 2.5 mmol) was added. The ice bath was removed, and the mixture was left to return to room temperature. Water (4 mL) was added. The mixture was heated at 50 °C for 14 hours. After 14 hours, the mixture was cooled to room temperature and Boc-Phe-OSu **3-4** (1.65 g, 4.54 mmol) in degassed THF (10 mL) was added with DIPEA (4 mL). After 24 hours, the mixture was dissolved in EtOAc (150 mL) and washed with 1 M HCl (200 mL x 3). The organic layer was extracted and dried over anhydrous $MgSO_4$. The solution was filtered and concentrated under reduced pressure. The crude mixture was purified by silica column chromatography, using a gradient of 100% CH_2Cl_2 → 10% MeOH/90% CH_2Cl_2. Thioester **3-2** was isolated as a sticky, viscous, yellow oil (1.00 g, 47%). **TLC:** (CH_2Cl_2:MeOH (9:1) R_f = 0.12. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.81 (s, 1H, H_{15}), 7.65 (d, 1H, J = 8.4 Hz, H_{16}), 7.43 – 7.18 (m, 5H, $H_{9\text{-}11}$), 7.16 (d, 1H, J = 8.4 Hz, H_4), 4.18 (m, 1H, H_6), 3.98 (m, 2H, H_{13}), 3.14 – 2.72 (m, 3H, H_{12}, H_7), 1.38 (s, 9H, H_1), 1.31 (s, 9H, H_{19}). **^{13}C NMR**

(101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 204.7 (C_5), 172.4 (C_{14}), 155.3 (C_3, C_{17}), 137.5 (C_8), 129.1 (C_9), 128.2 (C_{10}), 126.4 (C_{11}), 78.6 (C_2, C_{18}), 62.2 (C_6), 52.7 (C_{13}), 36.4 (C_7), 29.7 (C_{12}), 28.1 (C_1, C_{19}). **TOF MS**: (ES-) *m/z* $[M-H]^-$ calculated: 467.19, found 467.18. **IR** (ν_{max} / cm^{-1}) 3300 (O-H), 2950-2920 (N-H stretch, C-H phenyl), 1675 (C=O stretch, ester), 1500 (C=O stretch, Boc). See Supplementary Figure 3-7:3-12.

3.4.4 Synthesis of selenoester 3-3; *N*-(Boc)-*Se*-((Boc)-*l*-phenylalanyl)-*l*-selenocysteine



Seleno-*l*-cystine **3-3** (1.00 g, 3.0 mmol) was dissolved in water (7.2 mL). The mixture was cooled to 0 °C in an ice bath, and triethylamine (0.91 g, 9.0 mmol) was added. The temperature was returned to room temperature and di-tert-butyl dicarbonate (2.0 g,

9.0 mmol) was added to the mixture. The mixture was stirred for 18 hours at room temperature. Upon completion, 0.1 M hydrochloric acid (7.2 mL) was added, and the mixture was stirred to form a white precipitate. The mixture was dissolved in water (150 mL) and washed with EtOAc (200 mL x 3). The organic layers were combined and dried with anhydrous $MgSO_4$. The solution was filtered and concentrated under reduced pressure. The crude product was recrystallised using hexane and filtered to produce the pale-yellow solid $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** (1.35 g, 84%). **^{1}H NMR** (400 MHz,

$(CD_3)_2SO$) $\delta_{H/ppm}$: 12.78 (s, 2H), 7.18 (d, *J* = 8.2 Hz, 2H), 4.14 (td, *J* = 9.2, 4.8 Hz, 2H), 3.27 (dd, *J* = 12.6, 4.6 Hz, 2H), 3.10 (dd, *J* = 12.3, 9.8 Hz, 2H), 1.38 (s, 18H).

^{13}C NMR (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: δ 173.0 (2C), 155.8 (2C), 78.7 (2C), 54.7 (2C), 31.4 (2C), 28.6 (6C). **^{77}Se NMR** (76 MHz, $(CD_3)_2SO$) $\delta_{Se/ppm}$: 303 (1Se). **TOF MS**: (ES-) *m/z* $[M-H]^-$ calculated: 535.02, found 535.01.

In a brown reaction vessel, *Boc*-seleno-*l*-cystine ($(Boc\text{-}Sec\text{-}OH)_2$) **3-8** (0.5 g, 0.94 mmol) and Boc-Phe-OSu **3-4** (0.68 g, 1.87 mmol) were purged under a nitrogen atmosphere and dissolved in anhydrous DMF (4 mL). The mixture was cooled to 0 °C in an ice bath, and tributyl phosphine (460 µL, 1.87 mmol) was added. The ice bath was removed, and the mixture was left for 18 hours. The reaction mixture was then concentrated under reduced pressure. The crude mixture was purified by silica column chromatography, using a gradient of 100% hexane → 100% EtOAc → 50% EtOAc/50% MeOH. Selenoester **3-3** was isolated as a sticky, viscous, orange oil (0.17 g, 38%). **TLC:** (EtOAc 100%) R_f = 0.10. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.83 (s, 1H, H_{15}), 7.85 (d, 1H, *J* = 8.2 Hz, H_{16}), 7.34-7.25 (m, 5H, $H_{9\text{-}11}$), 7.22 (d, 1H, *J* = 8.8 Hz, H_4), 4.28-4.22 (m, 1H, H_6), 4.09-4.03 (m, 1H, H_{13}), 3.23 (dd, 1H, *J* = 12.4, 5.0 Hz, H_{12}), 3.10 (dd, 1H, *J* = 13.9, 4.1 Hz, H_7), 3.01 (dd, 1H, *J* = 12.6, 9.6 Hz, H_{12}), 2.81 (dd, 1H, *J* = 12.9, 9.9 Hz, H_7) 1.43 (s, 9H, H_1), 1.38 (s, 9H, H_{19}). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 205.4 (C_5), 172.8 (C_{14}), 155.8 (C_3, C_{17}), 136.9 (C_8), 129.1 (C_9), 128.4 (C_{10}), 126.7 (C_{11}), 79.6 (C_2, C_{18}), 64.8 (C_6), 53.5 (C_{13}), 36.5 (C_7), 28.4 (C_1, C_{19}), 25.5 (C_{12}). **^{77}Se NMR** (76 MHz, $(CD_3)_2SO$) $\delta_{Se/ppm}$: 531 (1Se). **HRMS**: (ES+) *m/z* $[M+Na]^+$ calculated: 539.1375, found 539.1304 m/z. **IR** (ν_{max} / cm^{-1})

3300 (O-H), 2960-2920 (N-H stretch, C-H phenyl), 1681 (C=O stretch, ester), 1500 (C=O stretch, Boc). See Supplementary Figure 3-13:3-18 for characterisation spectra.

3.4.5 Small molecule hydrolysis

Each ester (2 mM, **3-1** to **3-3**) was diluted in a PBS-buffered solution, at varying pHs (6.5 pH, 9.5 pH and 10.5 pH) in a black Eppendorf vial (1 mL). Samples were added to a thermomixer and mixed at 1000 rpm, 21 °C for 336 h. Aliquots (30 µL) of the sample were taken at various time points and directly injected (20 µL) for RP-HPLC analysis. The percentage of hydrolysis was calculated by the integral of the ester HPLC signal at time point (n), *t = n,* and divided by the starting integral of the ester (*t = 0)*.

$$\frac{\int ester\ t_n}{\int ester\ t_0} \times 100 = Hydrolysis\ \% \qquad \text{Eq. 3-1}$$

3.4.6 Small molecule aminolysis

Each amino acid was dissolved in acetonitrile, dimethylformamide or water to produce a stock solution (100 mM), which was further diluted into the reaction vial. A ratio of 1:2 ester (2 mM, **3-1** to **3-3**) to amino acid (4 mM) was mixed in a PBS-buffered solution, at varying pHs (6.5 pH, 7.5 pH, 8.5 pH, 9.5 pH and 10.5 pH) in a black Eppendorf vial (1 mL). Samples were added to a thermomixer and mixed at 1000 rpm, 21 °C. Aliquots (30 µL) were taken at various time points and directly injected (20 µL) for RP-HPLC analysis. Calibration curves were performed for the aminolysis peptide products **3-9** (Experimental 3.4.7) and the hydrolysis product Boc-Phe-OH **3-5** to determine the mole amount and, thus, the percentage of aminolysis/hydrolysis (Eq. 3-2). Calibrations ranged from 10 mM to 0.001 mM.

$$\frac{Moles\ of\ Peptide\ \mathbf{3-9}}{Moles\ of\ Peptide\ \mathbf{3-9} + Moles\ of\ BocPheOH\ \mathbf{3-5}} \times 100 = Aminolysis\ \% \qquad \text{Eq. 3-2}$$

3.4.7 Peptide synthesis Boc-Phe-Amino Acid 3-9 for calibrations

For peptides **3-9a:d**, followed the procedure by Popovic *et al.*[45]:

Boc-Phe-OSu **3-4** (0.36 g, 1 mmol) was dissolved in 1 mL EtOH:Acetone (1:4) mix and mixed with a solution of amino acid (1.1 mmol) with $NaHCO_3$ (0.21 g, 2 mmol) in H_2O (1 mL) for 18 h. The reaction mixture was then concentrated under reduced pressure. The mixture was acidified with 1 M $KHSO_4$ to pH 2-3 and dissolved in EtOAc (100 mL) and washed with 1 M $KHSO_4$ (100 mL x 3). Followed by a final wash with brine (30 mL). The organic layer was extracted and dried over anhydrous $MgSO_4$. The solution was filtered and concentrated under reduced pressure. If impure, the crude mixture was purified by silica column chromatography using 95:5 EtOAc:CH_2Cl_2.

Boc-Phe-Gly-OH* 3-9*a: l-Gly (0.083 g). White solid. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.68 (s, 1H), 8.30 (d, *J* = 5.8 Hz, 1H), 7.39 – 7.17 (m, 5H), 6.99 (d, *J* = 8.8 Hz, 1H), 4.26 (t, *J* = 7.2 Hz, 1H), 3.85 (dd, *J* = 16.2, 5.8 Hz, 2H), 3.07 (dd, *J* = 14.0, 3.9 Hz, 1H), 2.76 (dd, *J* = 13.8, 11.4 Hz, 1H), 1.34 (s, 9H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 172.6 (1C), 171.6 (1C), 155.7 (1C), 138.7 (1C), 129.7 (2C), 128.4 (2C), 126.6 (1C), 78.4 (1C), 56.0 (1C), 41.2 (1C), 37.9 (1C), 28.6 (3C). **TOF MS**: (ES+) *m/z* $[M+Na]^+$ calculated: 345.15, found 345.14 m/z.

Boc-Phe-Ala-OH* 3-9*b: l-Ala (0.098 g). White solid. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.74 (s, 1H), 7.99 (d, *J* = 7.9 Hz, 1H), 7.37 – 7.28 (m, 5H), 6.97 (d, *J* = 7.9 Hz, 1H), 4.51 (td, *J* = 8.0, 5.1 Hz, 1H), 4.04 (td, *J* = 14.5, 7.4, 1H), 3.14 (dd, *J* = 13.8, 5.2 Hz, 1H), 3.01 (dd, *J* = 14.1, 8.4, 1H), 1.45 (s, 9H), 1.21 (d, *J* = 7.2 Hz, 3H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 173.2 (1C), 173.1 (1C), 155.5 (1C), 137.8

(1C), 129.7 (2C), 128.6 (2C), 126.9 (1C), 78.5 (1C), 53.7 (1C), 50.1 (1C), 37.2 (1C), 28.6 (3C), 18.6 (1C). **TOF MS**: (ES+) *m/z* $[M+Na]^+$ calculated: 359.17, found 359.16 m/z.

BocPheBocLysOH* 3-9*c: l-BocLys (0.27 g). Pale yellow solid. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.42 (s, 1H), 7.86 (t, *J* = 5.7 Hz, 1H), 7.32 – 7.13 (m, 5H), 7.02 (d, *J* = 7.9 Hz, 1H), 6.85 (d, *J* = 8.5 Hz, 1H), 4.15 – 4.01 (m, 1H), 3.86 – 3.72 (m, 1H), 3.13 – 3.04 (m, 1H), 3.03 – 2.86 (m, 2H), 2.74 (dd, *J* = 13.6, 9.7 Hz, 1H), 1.57 (d, *J* = 27.8 Hz, 2H), 1.37 (s, 9H), 1.30 (s, 9H), 1.23 (s, 2H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 174.8 (1C), 171.8 (1C), 156.1 (1C), 155.6 (1C), 138.6 (1C), 129.6 (2C), 128.4 (2C), 126.6 (1C), 78.4 (2C), 56.2 (1C), 53.9 (1C), 38.6 (1C), 38.3 (1C), 30.8 (1C), 29.1 (1C), 28.7 (3C), 23.4 (1C). **TOF MS**: (ES+) *m/z* $[M+Na]^+$ calculated: 516.28, found 516.27 m/z.

Boc-Phe-Asn-OH* 3-9*d: l-Asn (0.15 g). White solid. Replaced acid $KHSO_4$ wash with H_2O. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.61 (s, 1H), 8.14 (d, *J* = 7.9 Hz, 1H), 7.42 (s, 1H), 7.26 (m, 5H), 6.94 (s, 1H), 6.86 (d, *J* = 8.9 Hz, 1H), 4.58 (dd, *J* = 13.6, 5.2 Hz, 1H), 4.22 (ddd, *J* = 14.9, 9.4, 3.8 Hz, 1H), 3.05 (dd, *J* = 13.9, 4.0 Hz, 1H), 2.74 (dd, *J* = 13.4, 10.7 Hz, 1H), 2.60 (d, *J* = 6.0 Hz, 2H), 1.29 (s, 9H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 172.9 (1C), 171.7 (1C), 171.6 (1C), 155.3 (1C), 138.3 (1C), 129.4 (2C), 128.1 (2C), 126.2 (1C), 78.2 (1C), 55.8 (1C), 48.8 (1C), 37.6 (1C), 36.8 (1C), 28.2 (3C). **TOF MS**: (ES+) *m/z* $[M+Na]^+$ calculated: 402.17, found 402.16 m/z.

Boc-Phe-Phe-OH* 3-9*e, followed the procedure by Sohora *et al.*[46]: *l*-Phe (0.83 g, 5 mmol) was dissolved in a THF-H2O (1:1, 20 mL) mix with $NaHCO_3$ (1.05 g,

10 mmol). Once dissolved, Boc-Phe-OSu **3-4** (1.98 g, 5.5 mmol) was added dropwise and stirred for 72 hours. The reaction mixture was then concentrated under reduced pressure. The mixture was acidified with 1M HCl to pH 2-3, dissolved in EtOAc (50 mL) and washed with 0.1 M HCl (50 mL x 3). Followed by a final wash with brine (30 mL). The organic layer was extracted and dried over anhydrous $MgSO_4$. The solution was filtered and concentrated under reduced pressure. The crude mixture was purified by silica column chromatography using 95:5 EtOAc:CH_2Cl_2. White solid 65% yield. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.79 (s, 1H), 8.12 (d, *J* = 7.9 Hz, 1H), 7.29 – 7.19 (m, 10H), 6.86 (d, *J* = 8.8 Hz, 1H), 4.49 (td, *J* = 8.1, 5.2 Hz, 1H), 4.18 (td, *J* = 10.3, 4.0 Hz, 1H), 3.09 (dd, *J* = 13.9, 5.2 Hz, 1H), 2.99 – 2.89 (m, 2H), 2.68 (dd, *J* = 13.9, 10.5 Hz, 1H), 1.28 (s, 9H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 172.8 (1C), 171.7 (1C), 155.1 (1C), 138.1 (1C), 137.3 (2C), 129.6 (2C), 129.2 (2C), 128.2 (2C), 128.0 (1C), 126.5 (1C), 126.2 (1C), 78.0 (1C), 55.7 (1C), 53.3 (1C), 39.5 (1C), 37.5 (1C), 36.8 (1C), 28.1 (3C). **TOF MS**: (ES+) *m/z* $[M+Na]^+$ calculated: 435.20, found 435.19 m/z.

3.4.8 Synthesis of chalcogen esters alanine 3-12 and 3-13

3.4.8.1 Synthesis of alanine-selenoester 3-12 (*N*-(Boc)-*Se*-((Boc)-*l*-alanyl)-*l*-selenocysteine)

The synthetic procedure follows that of *Experimental 3.4.4* in which Boc-Phe-OSu **3-4** was replaced by Boc-Ala-OSu **3-11** (0.17 g, 1.87 mmol). **CRUDE ^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.18 (s, 1H), 7.77 (d, *J* = 7.2 Hz, 1H), 7.05 (d, *J* = 8.8 Hz, 1H), 4.00 (m, 1H), 3.92 (m, 1H), 3.15 – 2.93 (m, 1H), 1.41 (s, 9H), 1.39 (s, 9H), 1.18 (d, *J* = 7.3 Hz, 3H). **CRUDE ^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 207.7 (1C), 172.9 (1C), 155.7 (2C), 79.2 (1C), 78.6 (1C), 59.2 (1C), 49.1 (1C), 28.6 (6C),

25.2 (1C), 17.1 (1C). **TOF MS**: (ES+) *m/z* $[M+Na]^+$ calculated: 463.11, found 463.10 m/z. See Supplementary Figure 3-23:3-27 for characterisation spectra.

3.4.8.2 Synthesis of alanine-thioester 3-13 (*N*-(Boc)-*S*-((Boc)-*l*-alanyl)-*l*-cysteine)

The synthetic procedure follows that of *Experimental 3.4.3,* in which Boc-Phe-OSu **3-4** was replaced by Boc-Ala-OSu **3-11** (0.17 g, 1.87 mmol). **CRUDE ^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.84 (s, 1H), 7.52 (d, *J* = 7.0, 1H), 7.12 (d, *J* = 8.5 Hz, 1H), 4.16 – 4.04 (m, 1H), 4.01 – 3.90 (m, 1H), 3.27 (dd, *J* = 13.4, 4.8 Hz) and 2.94 (dd, *J* = 13.5, 9.6 Hz, 1H), 1.38 (d, *J* = 11.2 Hz, 18H), 1.21 (d, *J* = 7.3 Hz, 3H). **CRUDE ^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 203.1 (1C), 172.6 (1C), 155.7 (2C), 79.1 (1C), 78.7 (1C), 56.7 (1C), 53.5 (1C), 29.9 (1C), 28.7 (6C), 17.6 (1C). **TOF MS**: (ES+) *m/z* $[M+Na]^+$ calculated: 415.16, found 415.15 m/z. See Supplementary Figure 3-28:3-32 for characterisation spectra.

3.4.9 Selenoester 3-3 with DNA assay

In a microcentrifuge tube, 5 µL of selenoester **3-3** (8 mM, DMSO) was added to 2 µL of ssDNA-NH_2 **3-14** (200 µM) and diluted in 93 µL PBS (pH 6.5 or 9.5, 10 mM) to give a 400 µM:4 µM ester:DNA solution. The solution was mixed at 1000 rpm, 19 ºC and aliquots (10 µL) were taken at various time points. The aliquots were dissolved in 490 µL of H_2O and spin-filtered using Amicon-Ultra 3 kDa 0.5 mL centrifugal filters to remove buffer salts. This process was repeated twice, diluting with 400 µL of solution. The remaining filtered solution was injected (50 µL) directly for RP-HPLC analysis, analysing at λ = 260 nm UV/Vis channel.

3.4.10 Se-S small molecule 3-15 (2,5-dioxopyrrolidin-1-yl (*R*)-2-((*tert*-butoxycarbonyl)amino)-3-((pyridin-2-ylthio)selanyl)propanoate) synthesis

Stock solutions: $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** (0.0537 g, 1 M, DMF), HATU (0.0385 g, 1 M, DMF), NHS (0.0461 g, 4 M, DMF), 2-mercaptopyridine **3-19** (0.1116 g, 10 M, DMF).

In a microcentrifuge tube, 10 µL $(Boc\text{-}Sec\text{-}OH)_2$ (1 M, 1 eq) and 10 µL NHS (4 M, 4 eq) were mixed with 49 µL of DMF. The reaction was cooled to 0 ºC and 21 µL HATU (1 M, 2.1 eq) was added. The reaction was brought back to room temperature (21 ºC) and mixed overnight. After 18 hours, 25 µL of 2-mercaptopyridine **3-19** (10 M, 25 eq) was added to the reaction mixture and vortexed for 2 minutes. **TOF MS**: (ES+) *m/z* $[M+Na]^+$ $(Boc\text{-}Sec\text{-}OSu)_2$ **3-18** - expected: 753.05, found: 753.04 m/z. $[M+H]^+$ Se-S small molecule **3-15** – expected: 476.03, found: 476.04 m/z. See Supplementary Figure 3-35 for the characterisation spectrum.

3.4.11 TAMRA selenoester DNA 3-17 synthesis

Stock solutions: $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** (300 mM, DMF), EDC-HCl (300 mM, DMF), NHS (300 mM, DMF), TCEP (0.1 M, H_2O), TAMRA-NHS **2-6** (0.1 M, DMF)

In a primary microcentrifuge tube, 67 µL of $(Boc\text{-}Sec\text{-}OH)_2$ **3-8** stock (300 mM), 67 µL of EDC-HCl stock (300 mM), 67 µL of NHS (300 mM) and 5 µL of DIPEA were added and mixed for 20 minutes, 19 ºC at 1000 rpm. In a secondary microcentrifuge tube, 20 µL of NH_2-labelled DNA (**2-8**, 1 mM, H_2O), 180 µL borate-buffered saline (100 mM, pH 9.8) and 200 µL DMF were added along with all contents of the primary microcentrifuge tube. The combined solution was mixed for 18 h at 19 ºC and 1000 rpm. The solution was separated into three microcentrifuge tubes (~202 µL),

dissolved in 300 µL ethanol with 20 µL sodium acetate (NaOAc) (0.3 M, pH 5.2) and placed in the freezer overnight to remove unreacted small molecules. The modified DNA pellets were centrifuged (15000 rpm), emptied and washed twice with ice-cold ethanol (70%, H_2O). The modified DNA pellet was dried, dissolved in 500 µL H_2O and spin-filtered (Amicon-Ultra 3 kDa 0.5 mL centrifugal filter) to remove any remaining small molecules. The remaining 100 µL of solution was dried to a pellet and resuspended in 20 µL H_2O. In a microcentrifuge vial, 2 µL of the modified DNA **3-20** solution was added to 4 µL of TCEP stock and mixed for 10 minutes at 19 ºC, 1000 rpm, before adding 1 µL of TAMRA-NHS **2-6** (0.1 M, DMF) and 7 µL of borate buffered saline (100 mM, pH 9.8) and mixed for 18 h, at 19 ºC, 1000 rpm. After 18 h, 300 µL ethanol and 20 µL NaOAc (0.3 M, pH 5.2) were added, and the tube was placed in the freezer overnight. As before, the tubes were centrifuged, washed with ice-cold ethanol, dried and spin-filtered. The sample was then purified using fluorescence RP-HPLC (see *Appendix: Experimental Methods A.2.2*), eluting at 23.91 mins, and washed with H_2O in Amicon-Ultra 3 kDa 0.5 mL centrifugal filters, removing HPLC buffer salts. Resuspended in nanopure H_2O, the modified DNA was identified using LC-MS and the concentration was determined using a NanoDrop One/One UV-Vis Spectrophotometer absorbance at λ = 260 nm. **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 11546.04 m/z, found: 11546.71 m/z. See Supplementary Figure 3-36 for the characterisation spectrum.

3.4.12 Single-step DNA-templated synthesis assay

Stock solutions: TAMRA selenoester DNA **3-17** (9 µM, H_2O), ssDNA-NH_2 **2-2** (various sites, 9 µM, H_2O), DTS buffer (pH 11, 0.25 M NaCl, 0.25 M TAPS, 0.25 M CAPS, 0.25 M Na_3PO_4).

The synthetic procedure follows that of *Chapter 2 Experimental 2.4.5* in which TAMRA thioester DNA **2-1** (9 μM) was replaced by TAMRA selenoester DNA **3-17** (9 μM) and ssDNA-NH_2 **2-2** (various nucleobase sequences) were replaced with ssDNA-NH_2 **2-2** (varying NH_2 locations, see Table 2-6).

3.4.13 TAMRA Boc-Cys thioester 3-23 synthesis

The synthetic procedure follows that of *Experimental 3.4.11,* in which the (Boc-Sec-OH)$_2$ **3-8** 300 mM stock was replaced with a (Boc-Cys-OH)$_2$ **3-7** 300 mM stock. **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 11498.09 m/z, found: 11498.21 m/z. See Supplementary Figure 3-39 for the characterisation spectrum.

3.4.14 Protection assay

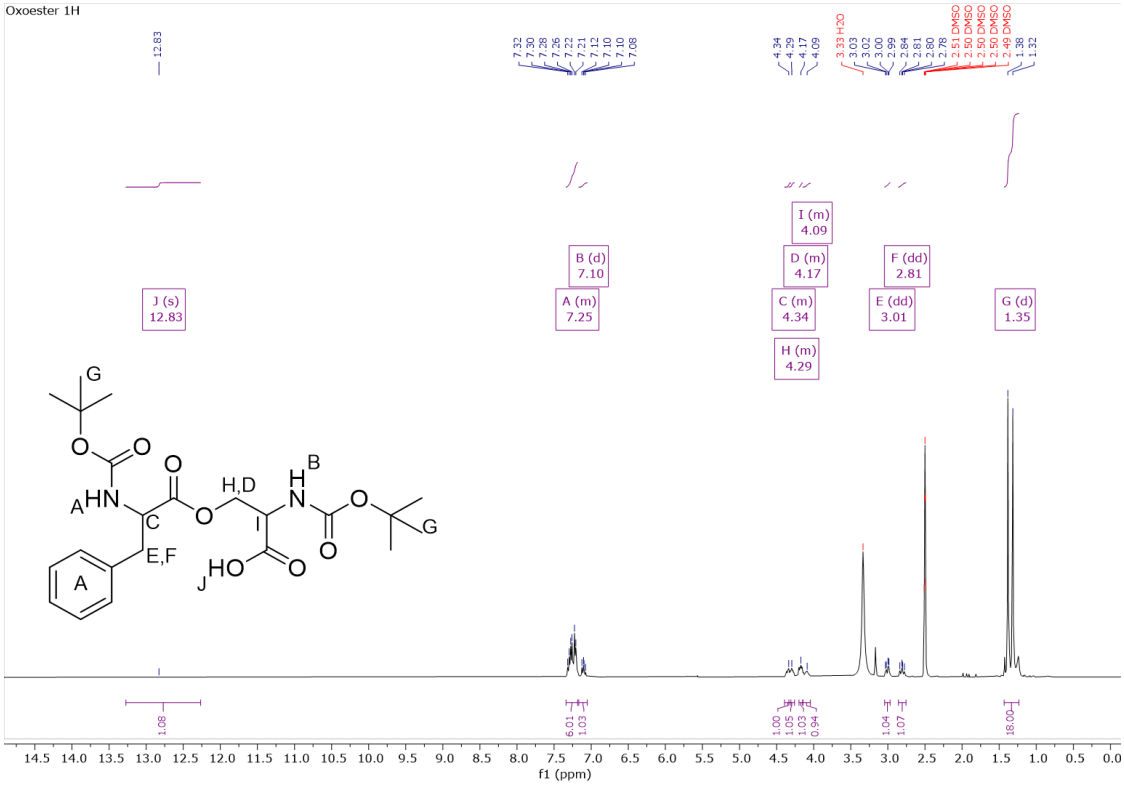
dsDNA protection assay for 2-1, 3-17 and 3-23:

The assay follows the procedure of *Experimental 3.4.12* in which the 6.67 μL ssDNA-NH_2 **2-2** (9 μM) was replaced with 6.67 μL DNA-spacer −3 **2-10** (9 μM). A sample interval of 18 h was added.

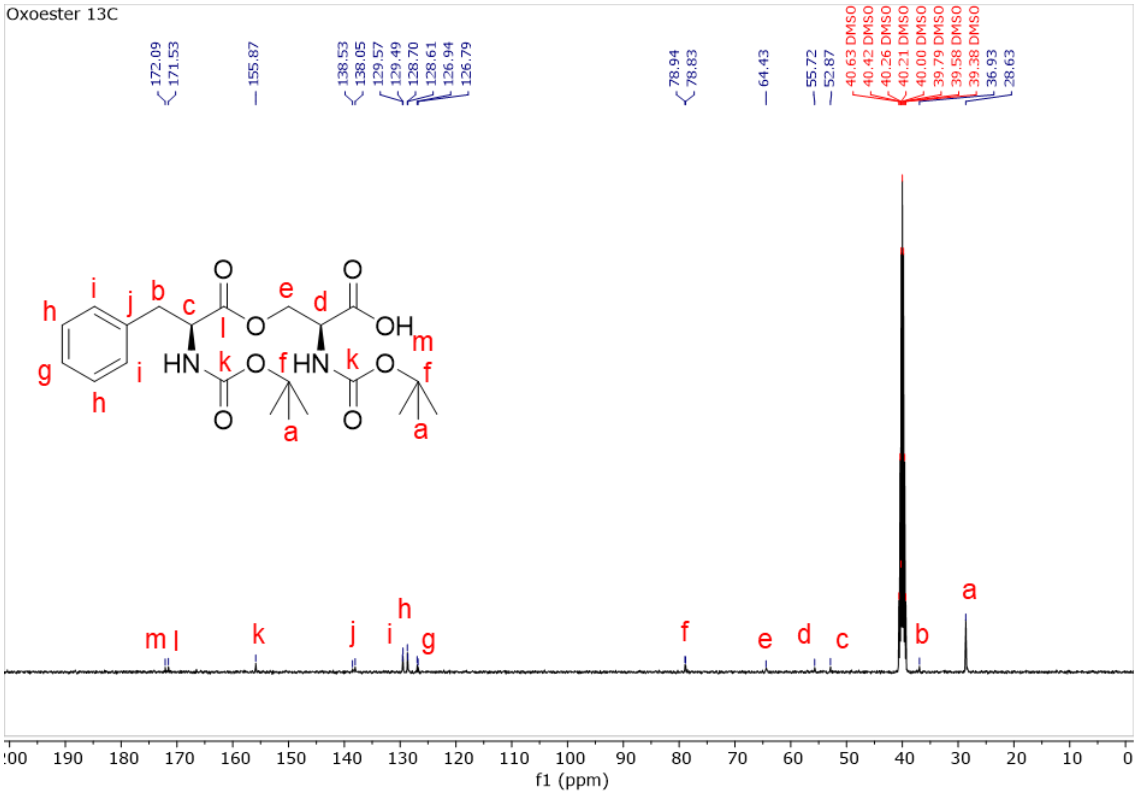
ssDNA stability assay of TAMRA selenoester DNA 3-17:

The assay follows the procedure of *Experimental 3.4.12* in which the 6.67 μL ssDNA-NH_2 **2-2** (9 μM) was replaced with 6.67 μL of Nanopure H_2O (18.2 MΩ). A sample interval of 18 h was added.

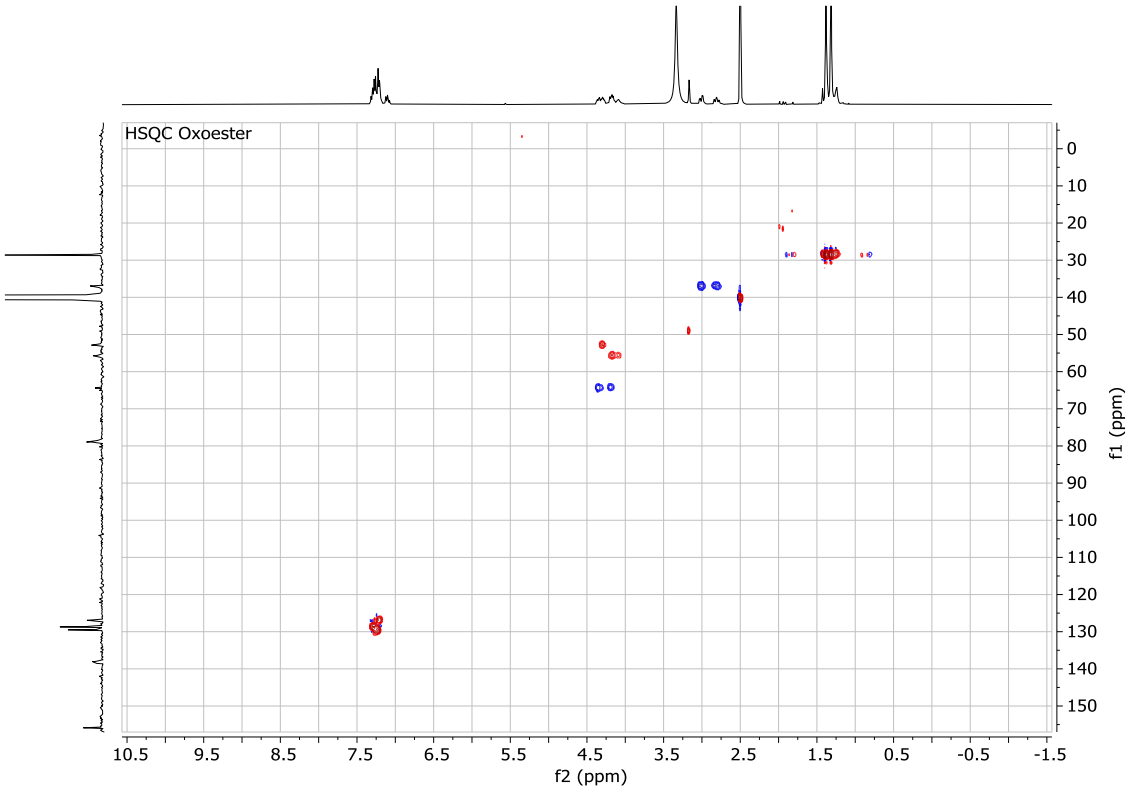
3.5 Supplementary data



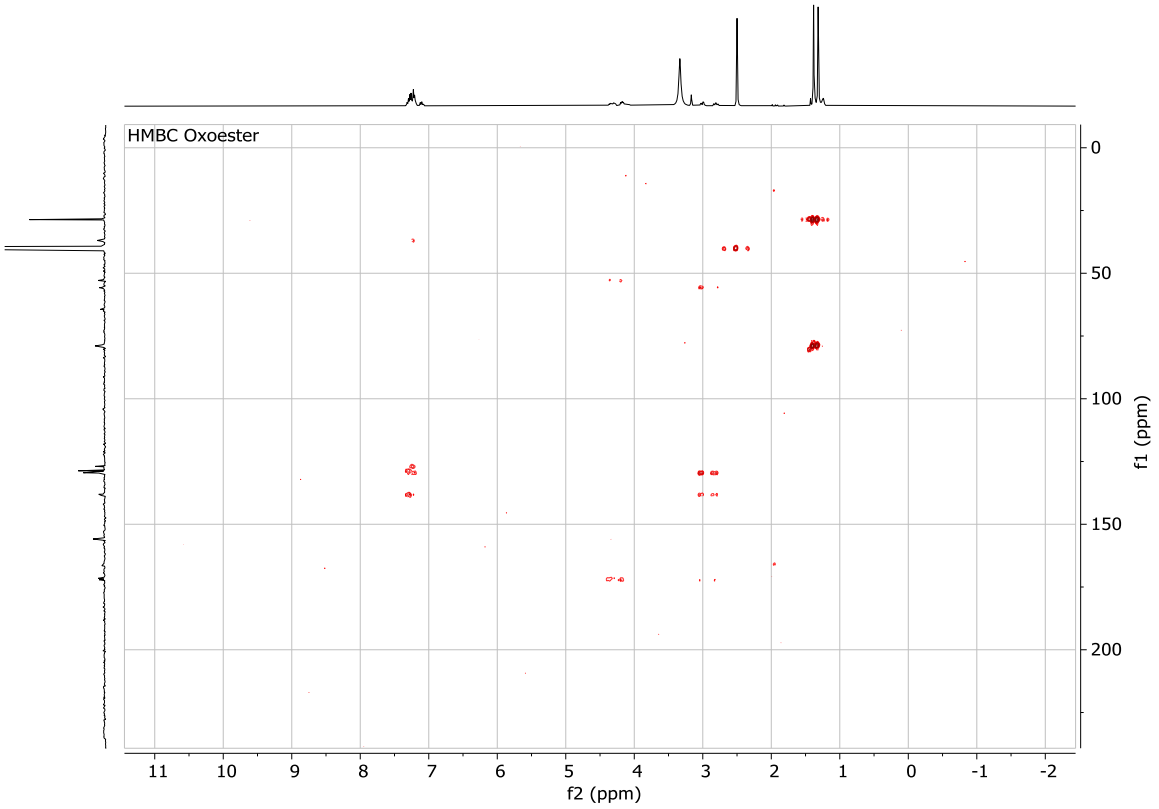
***Supplementary Figure 3-1**: ^{1}H NMR spectrum of oxoester **3-1**, 400 MHz, DMSO-d_6*



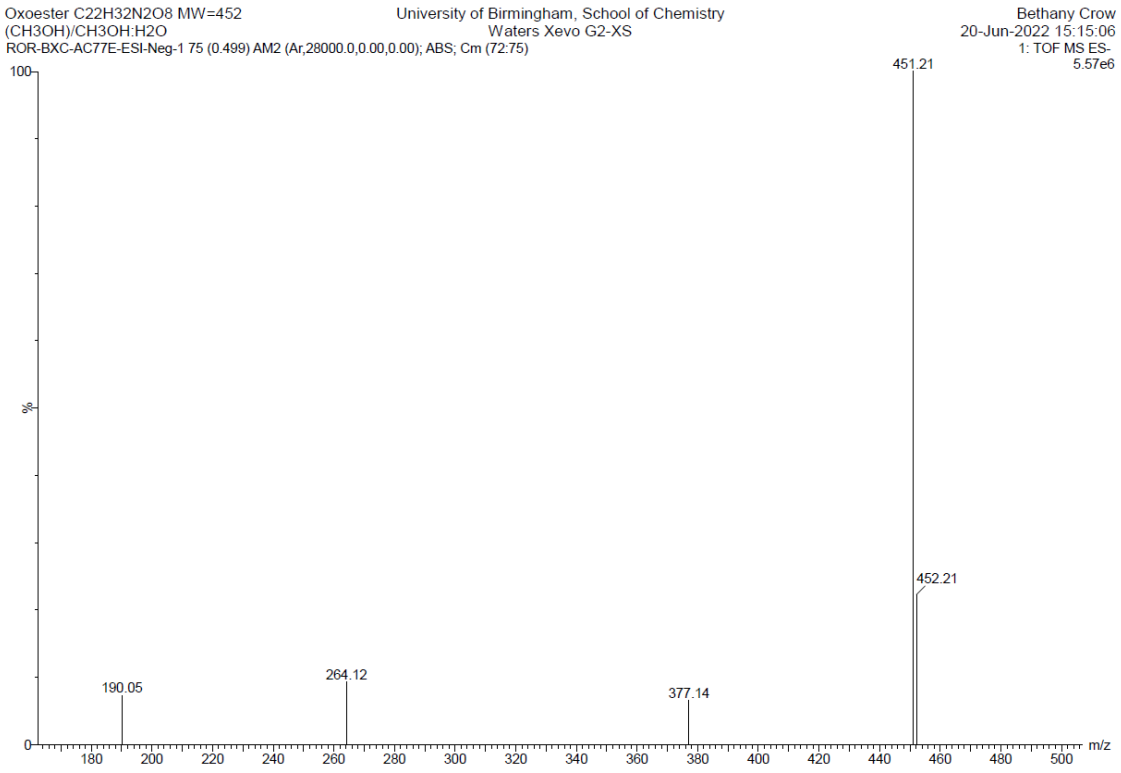
***Supplementary Figure 3-2:** ^{13}C NMR spectrum of oxoester **3-1**, 101 MHz, DMSO-d_6*



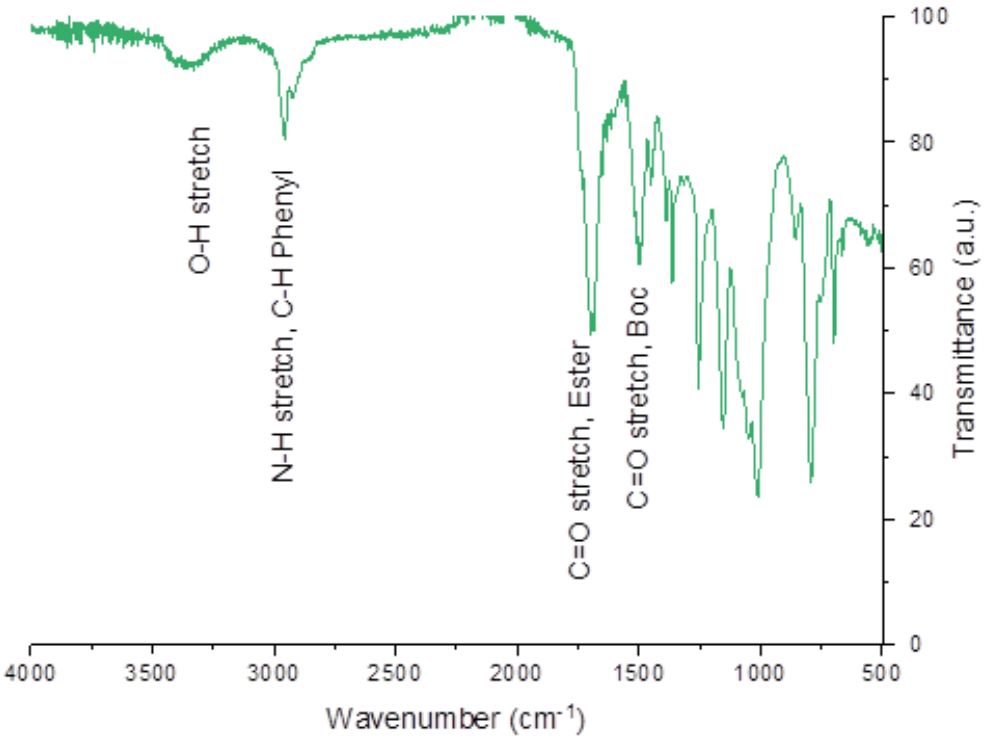
Supplementary Figure 3-3: *HSQC spectrum of oxoester* ***3-1****, 400/101 MHz, DMSO-d_6*



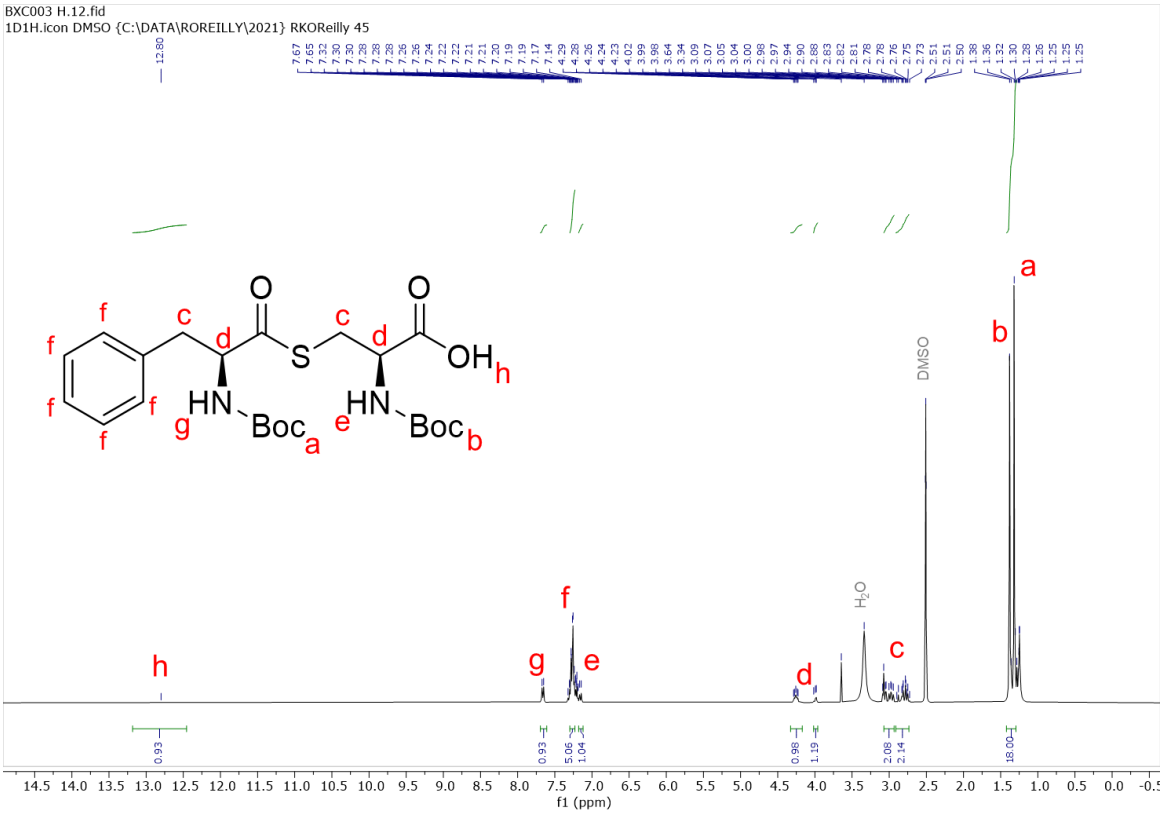
Supplementary Figure 3-4: *HMBC spectrum of oxoester* ***3-1****, 400MHz/101 MHz, DMSO-d_6.*



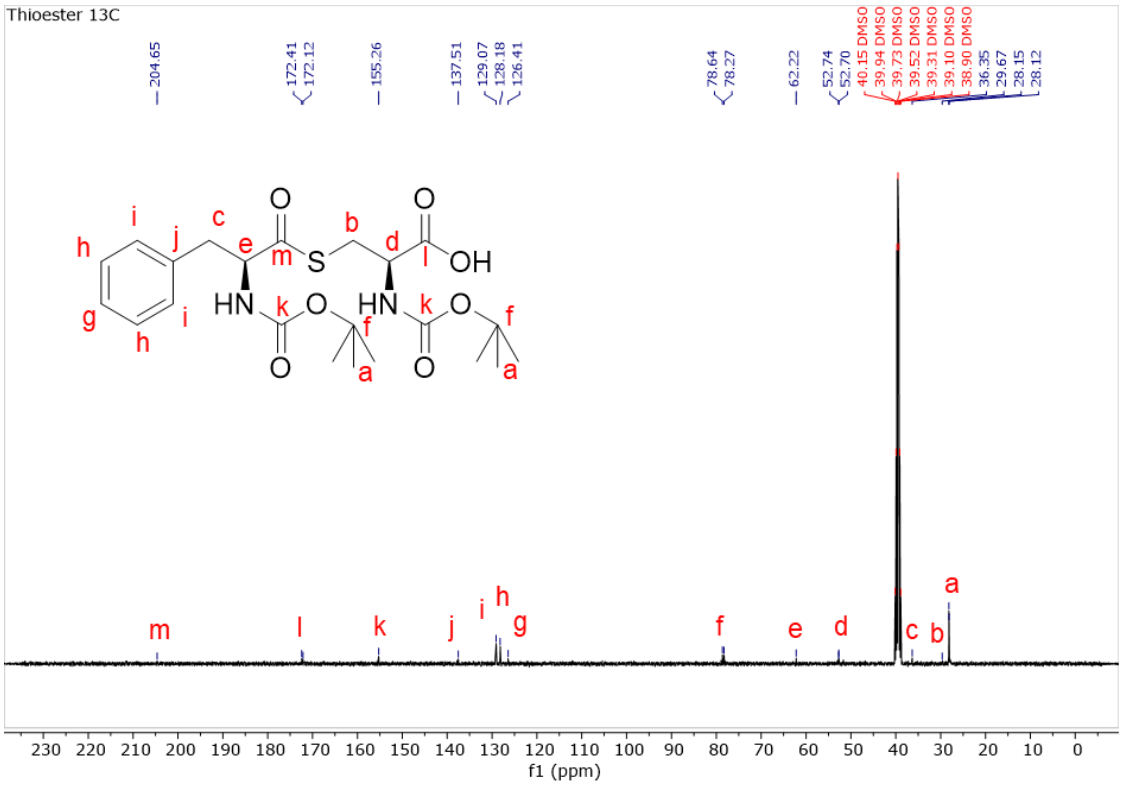
Supplementary Figure 3-5: *TOF MS ESI$^-$ of oxoester **3-1** [M-H]$^-$ expected: 451.22 m/z, found: 451.21 m/z.*



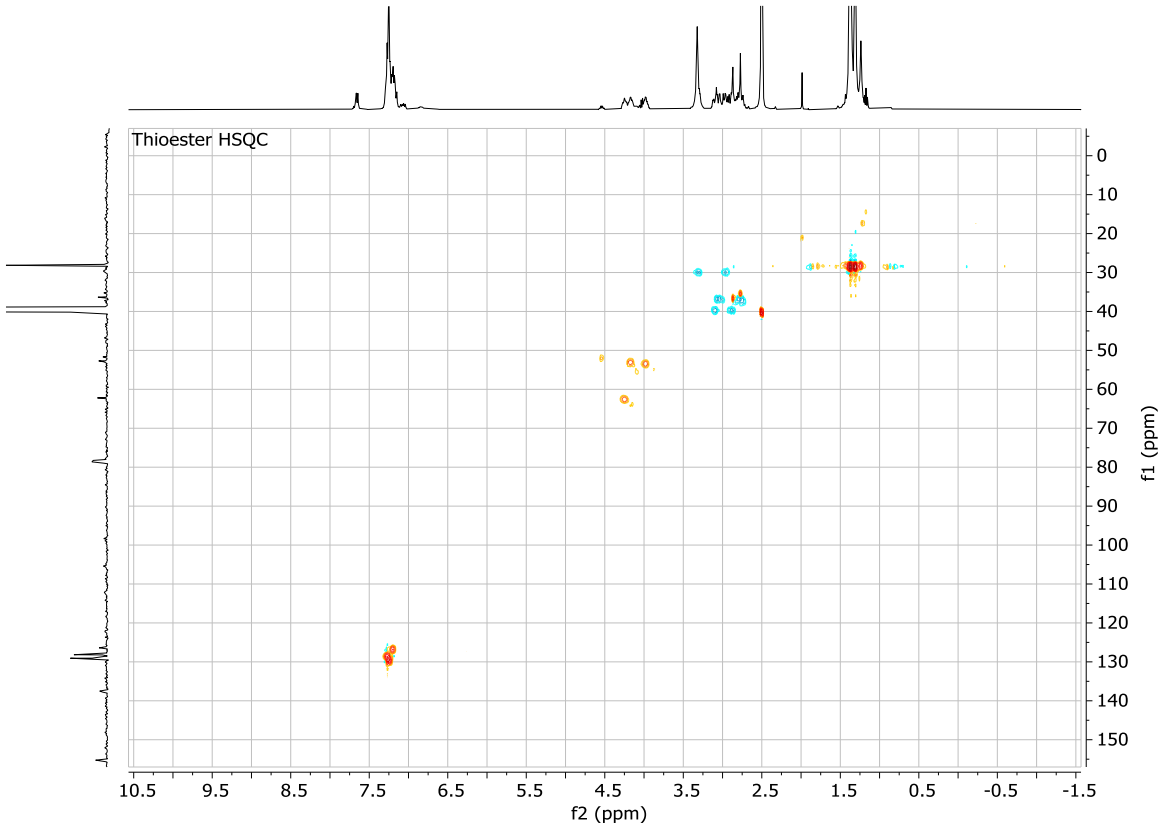
Supplementary Figure 3-6: *FT-IR spectrum of oxoester **3-1**.*



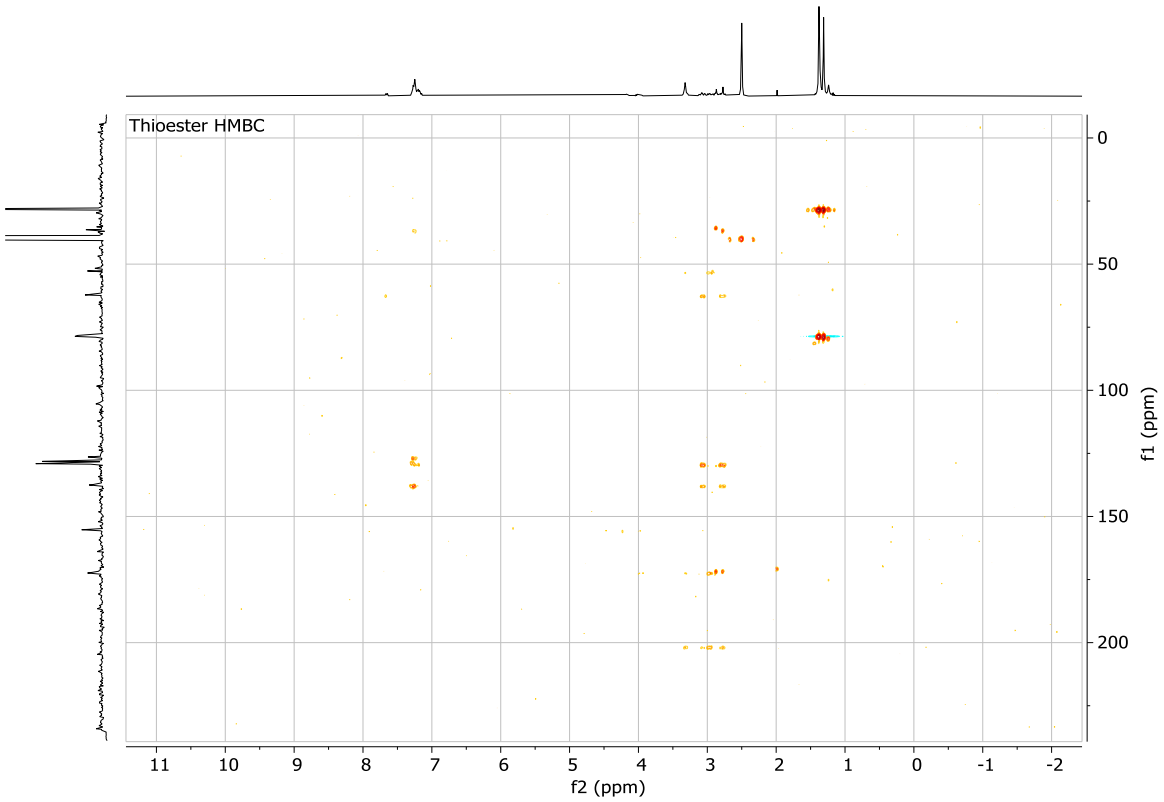
Supplementary Figure 3-7: *^{1}H NMR spectrum of thioester **3-2**, 400 MHz, DMSO-d_6*



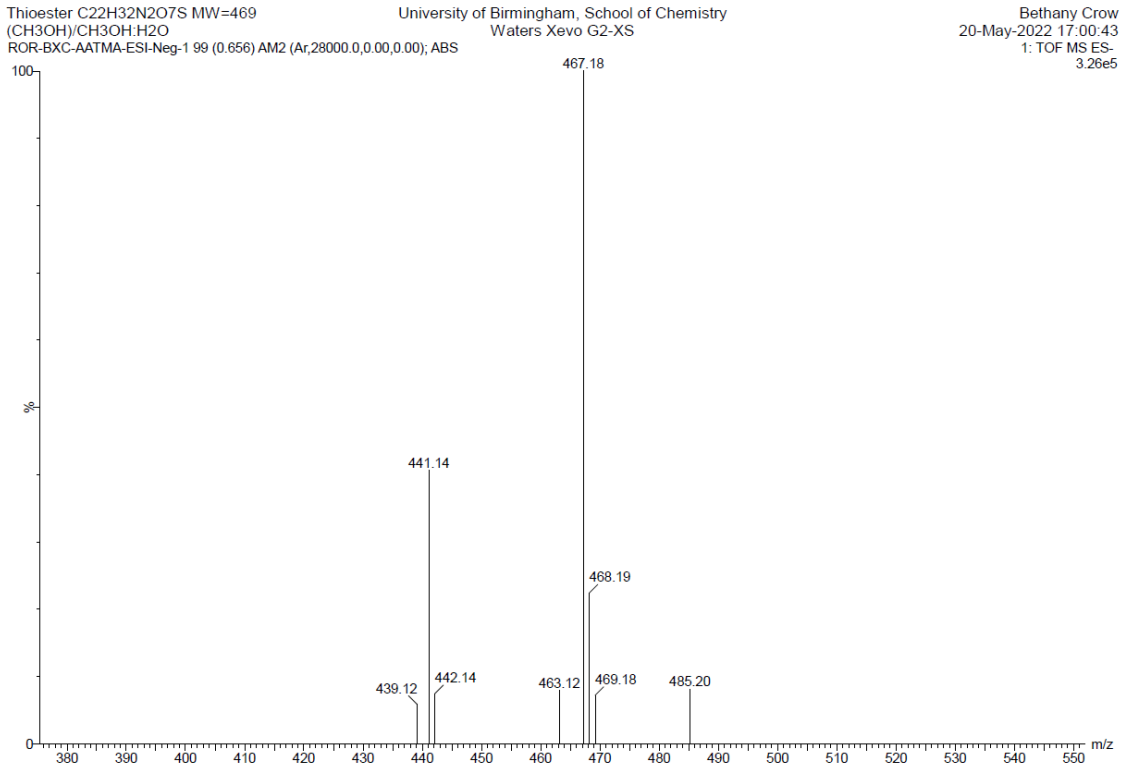
***Supplementary Figure 3-8**: ^{13}C NMR spectrum of thioester **3-2**, 101 MHz, DMSO-d_6*



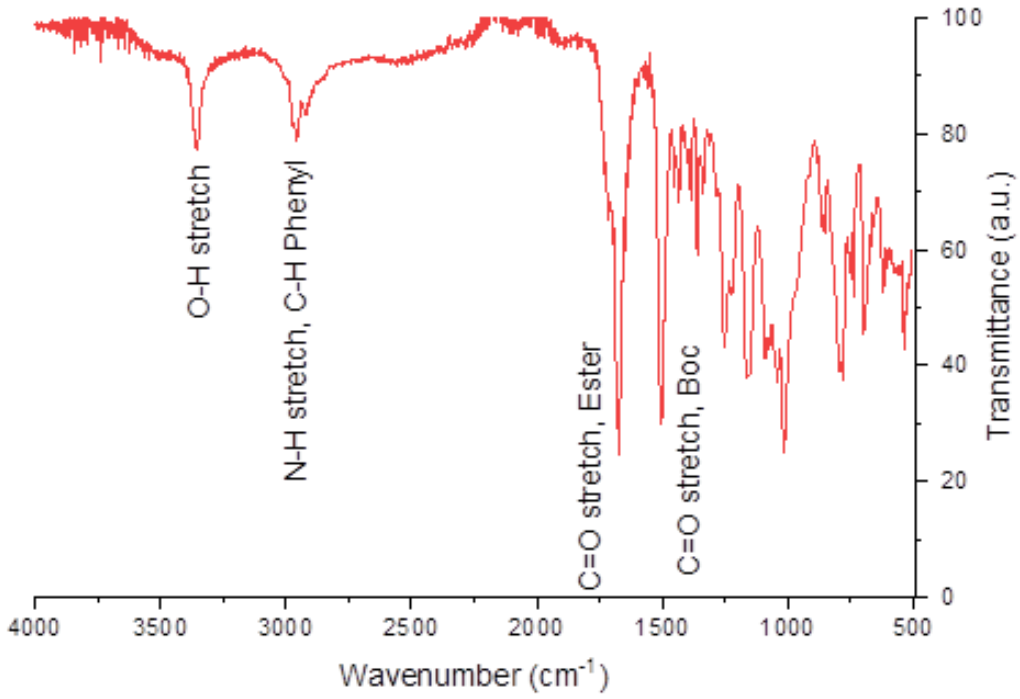
Supplementary Figure 3-9: *HSQC spectrum of thioester* ***3-2****, 400MHz/101 MHz, DMSO-d*$_6$



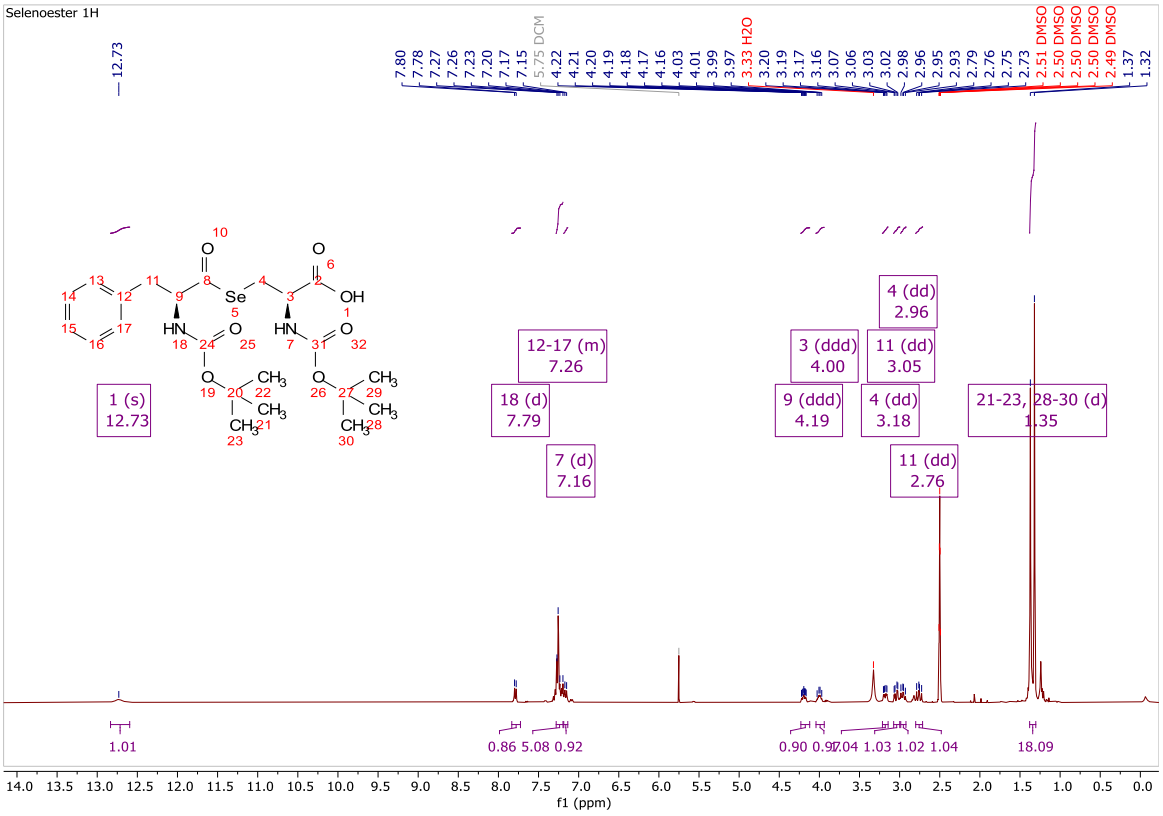
Supplementary Figure 3-10: *HMBC spectrum of thioester* ***3-2****, 400MHz/101 MHz, DMSO-d*$_6$.



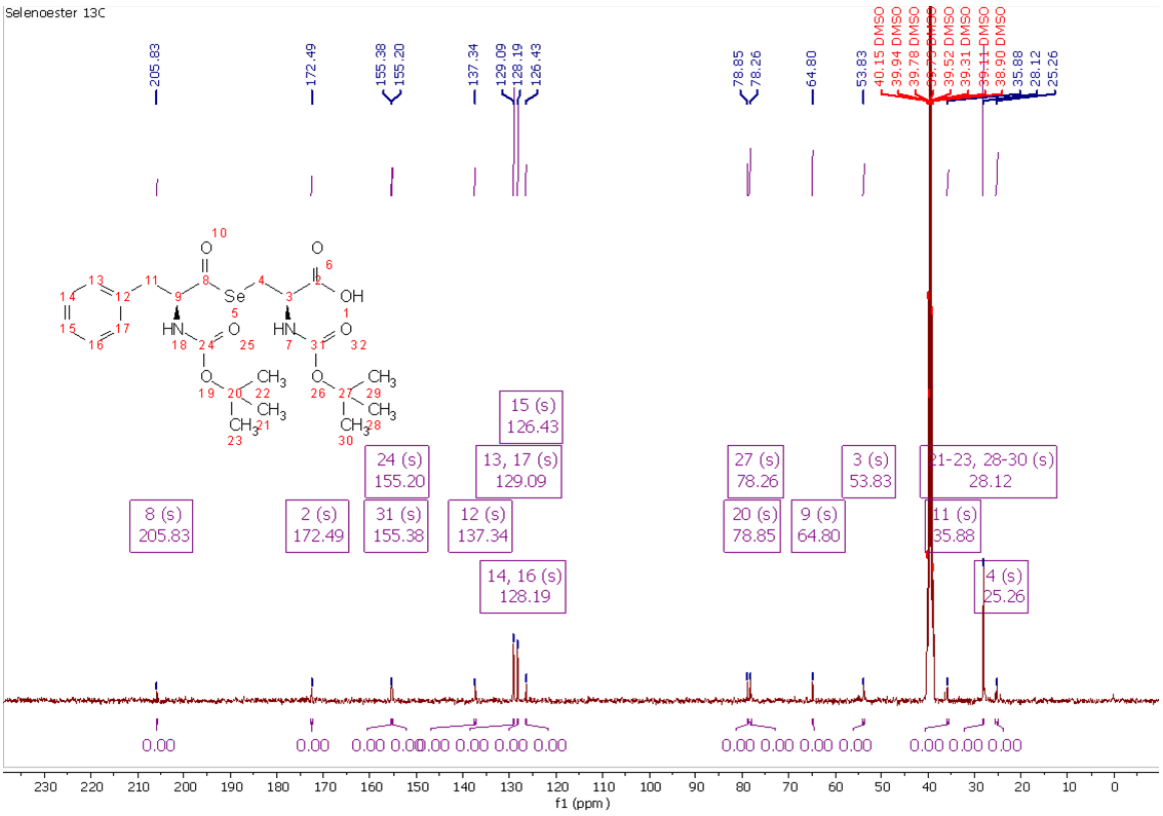
Supplementary Figure 3-11: *TOF MS ESI$^-$ of thioester **3-2**. [M-H]$^-$ expected: 467.19, found: 467.18 m/z.*



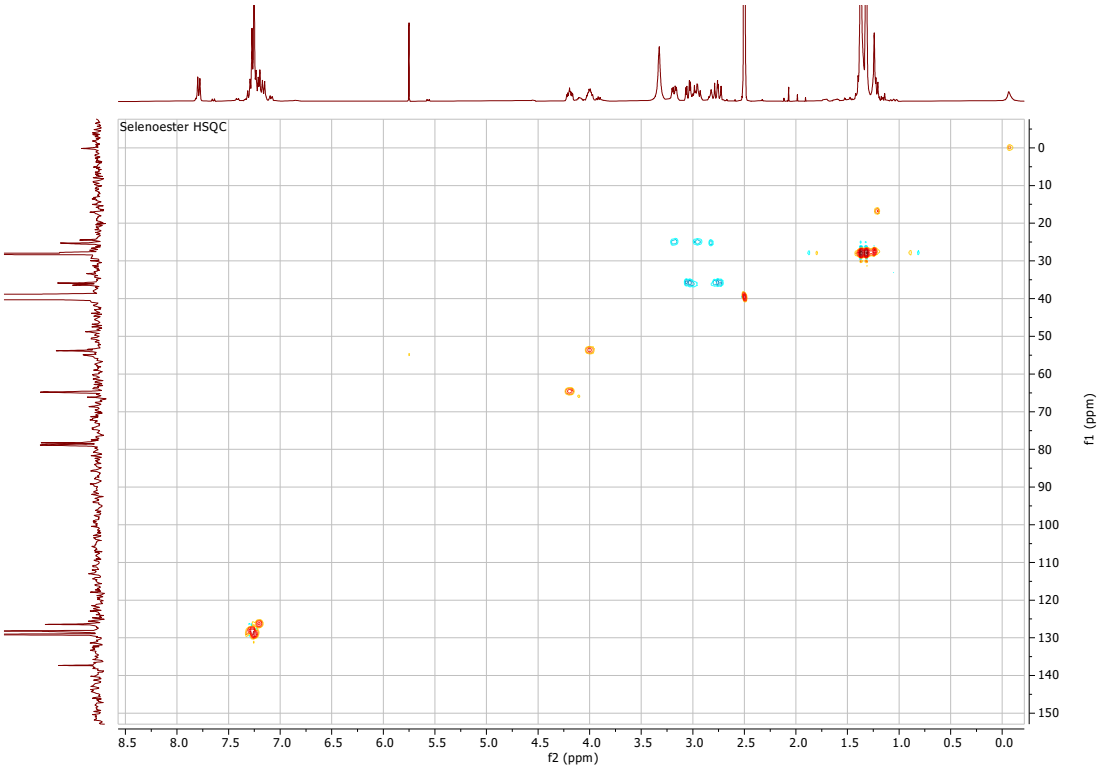
Supplementary Figure 3-12*: FT-IR spectrum of thioester **3-2**.*



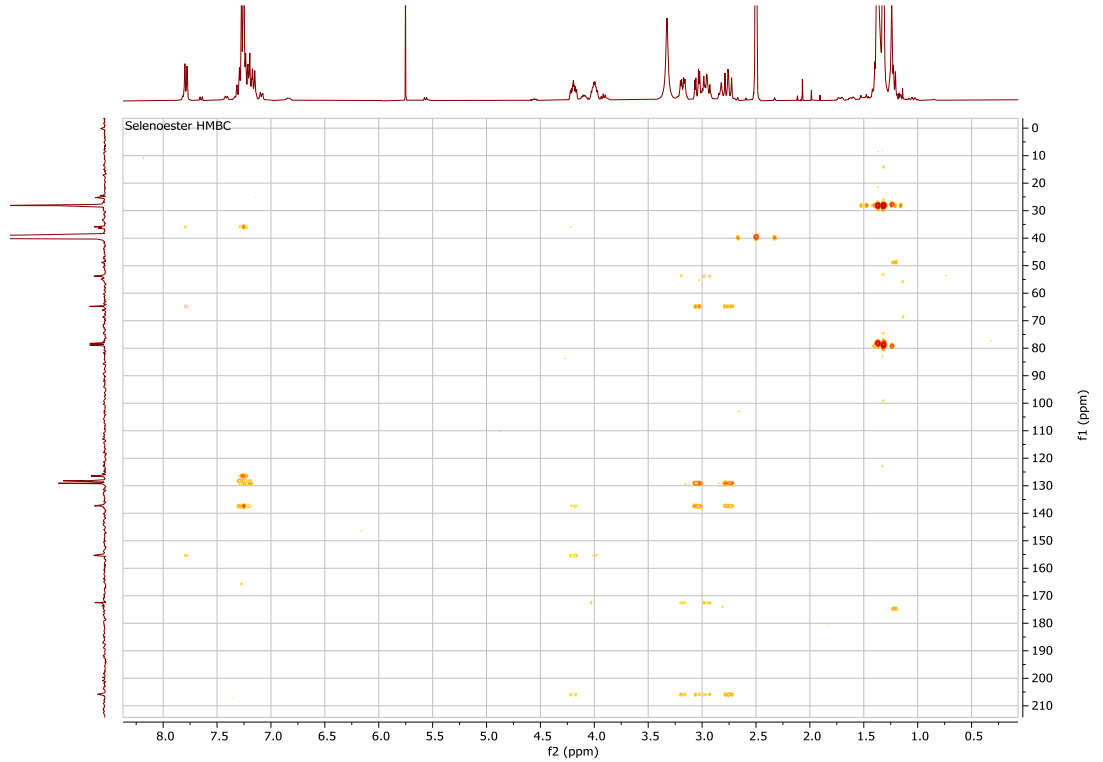
Supplementary Figure 3-13: *^{1}H NMR spectrum of selenoester **3-3**, 400 MHz, DMSO-d_6*



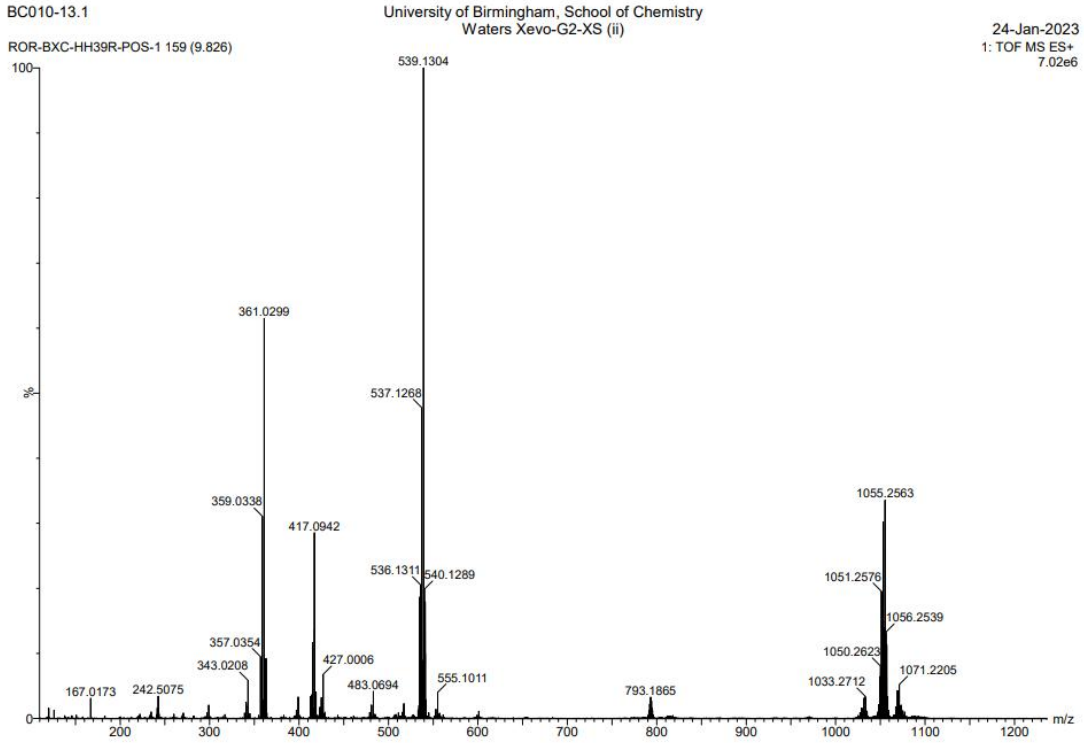
Supplementary Figure 3-14*: ^{13}C NMR spectrum of selenoester **3-3**, 101 MHz, DMSO-d_6*



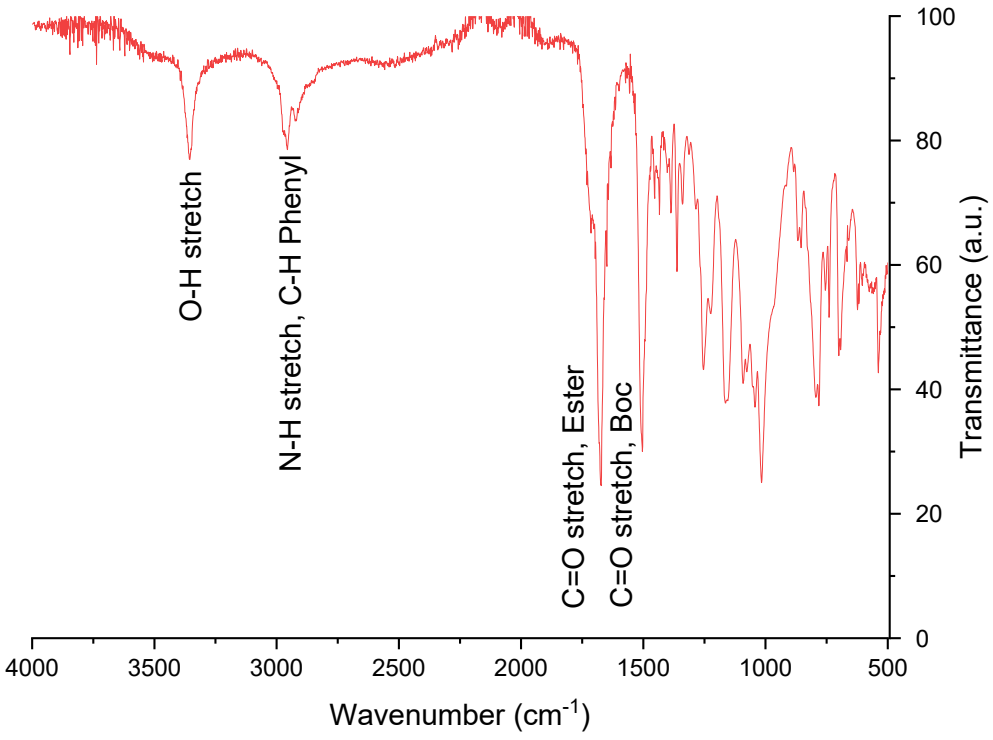
***Supplementary Figure 3-15**: HSQC spectrum of selenoester **3-3**, 400MHz/101 MHz, DMSO-d_6*



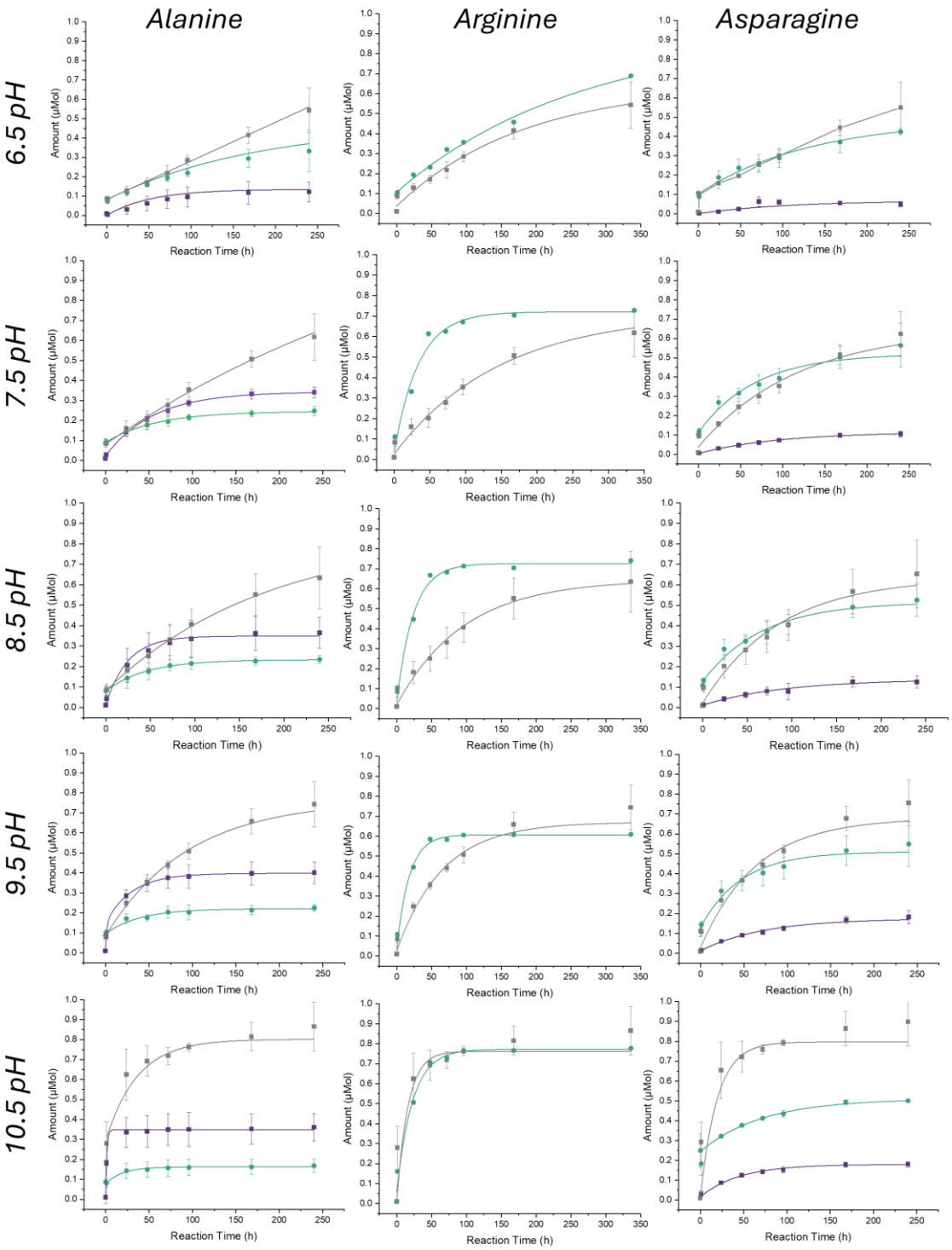
***Supplementary Figure 3-16**: HMBC spectrum of selenoester **3-3**, 400MHz/101 MHz, DMSO-d_6*



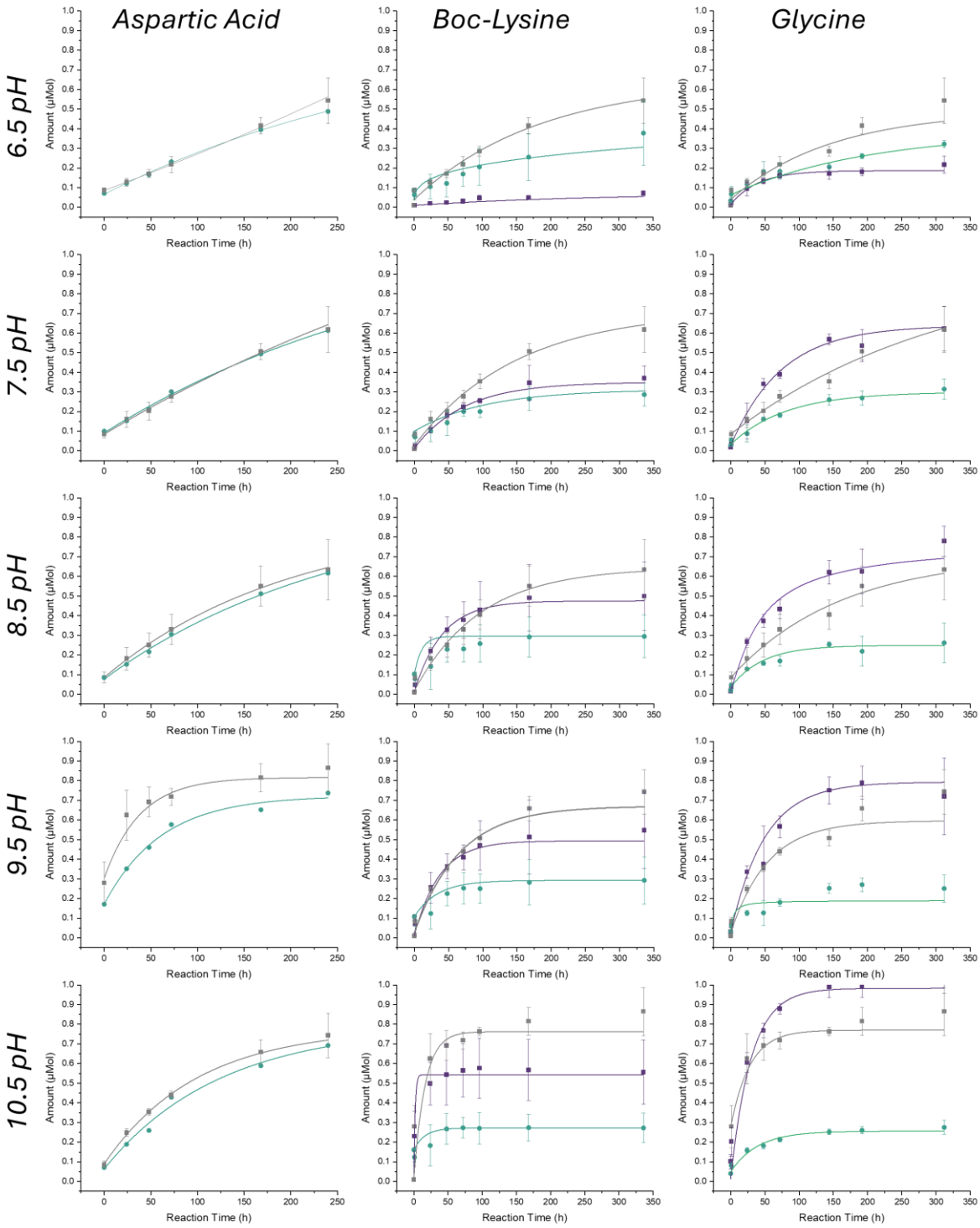
Supplementary Figure 3-17: *TOF MS ES+ of selenoester* ***3-3*** *$[M+Na]^+$ expected: 539.14, found: 539.13 m/z. 1055 m/z = $[2M+Na]^+$, 361 & 417 m/z = amino acid fragments.*



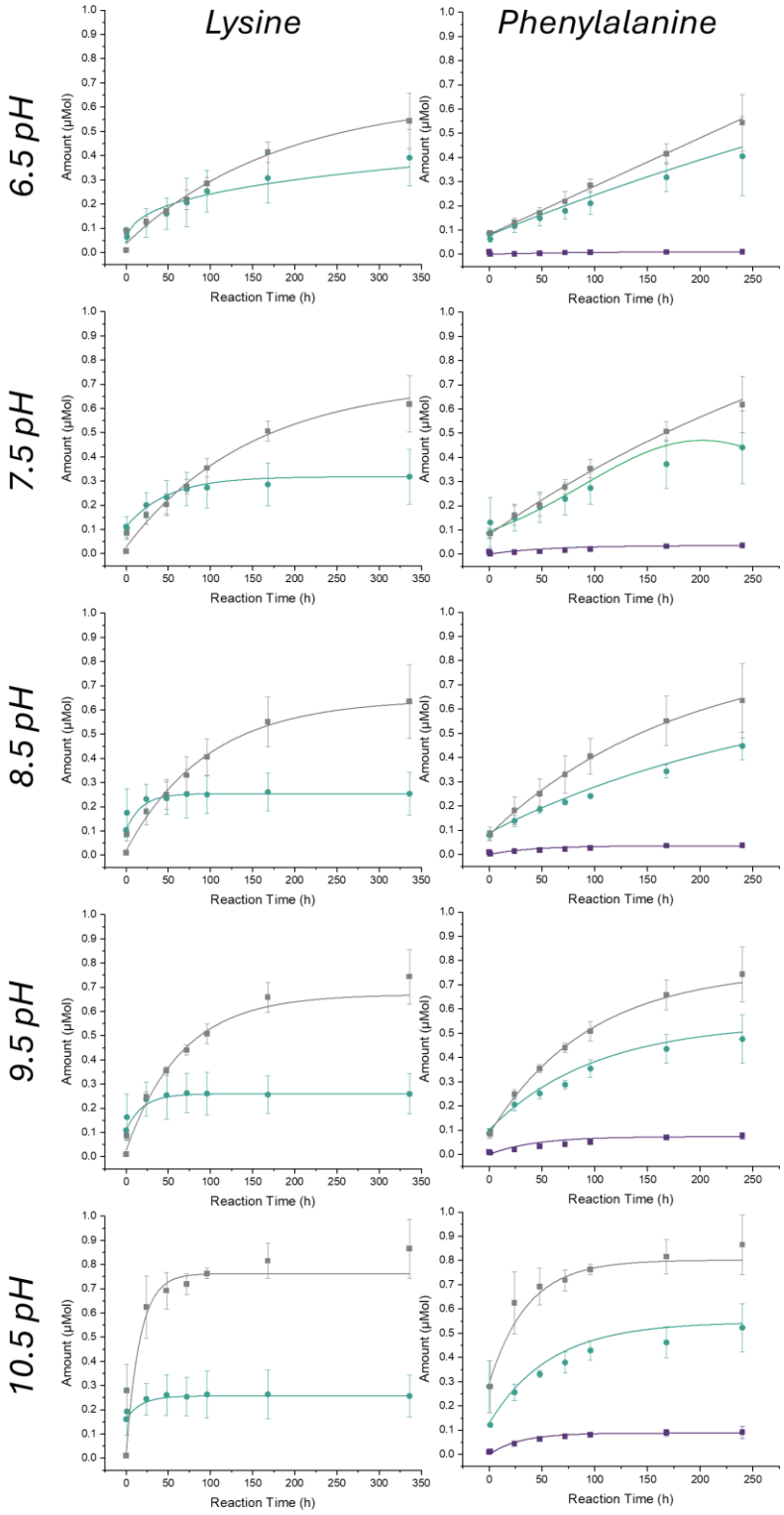
Supplementary Figure 3-18*: FT-IR spectrum of selenoester* ***3-3****.*



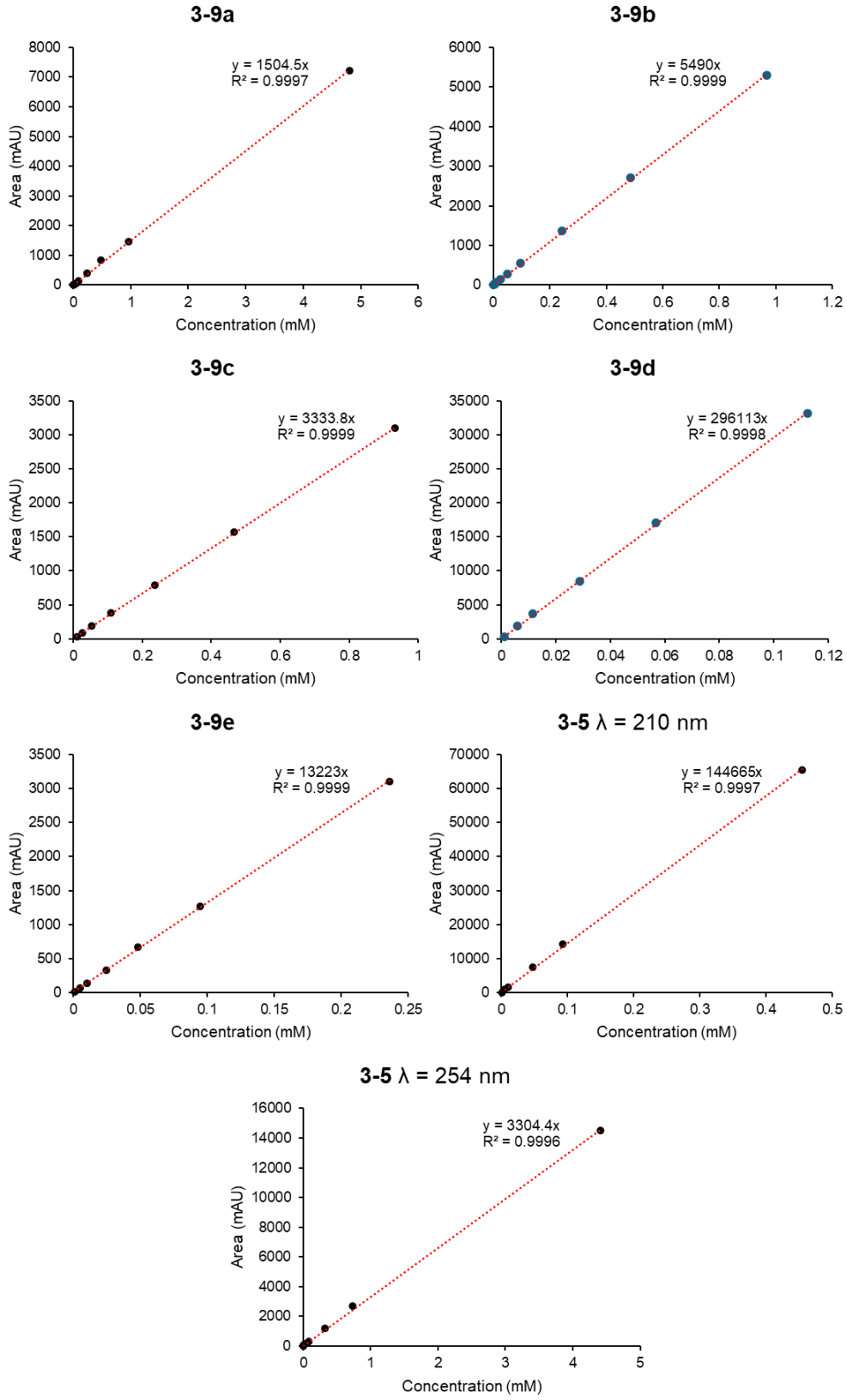
***Supplementary Figure 3-19**: Aminolysis of selenoester **3-3** with Ala (**3-9b**), Arg (**3-9g**) and Asn (**3-9d**) at various pHs over 2 weeks. Purple square = peptide aminolysis product (**3-9**). Green diamond = Boc-Phe-OH **3-5** hydrolysis product in aminolysis reaction. Grey square = Boc-Phe-OH **3-5** formed during selenoester **3-3** hydrolysis studies.*



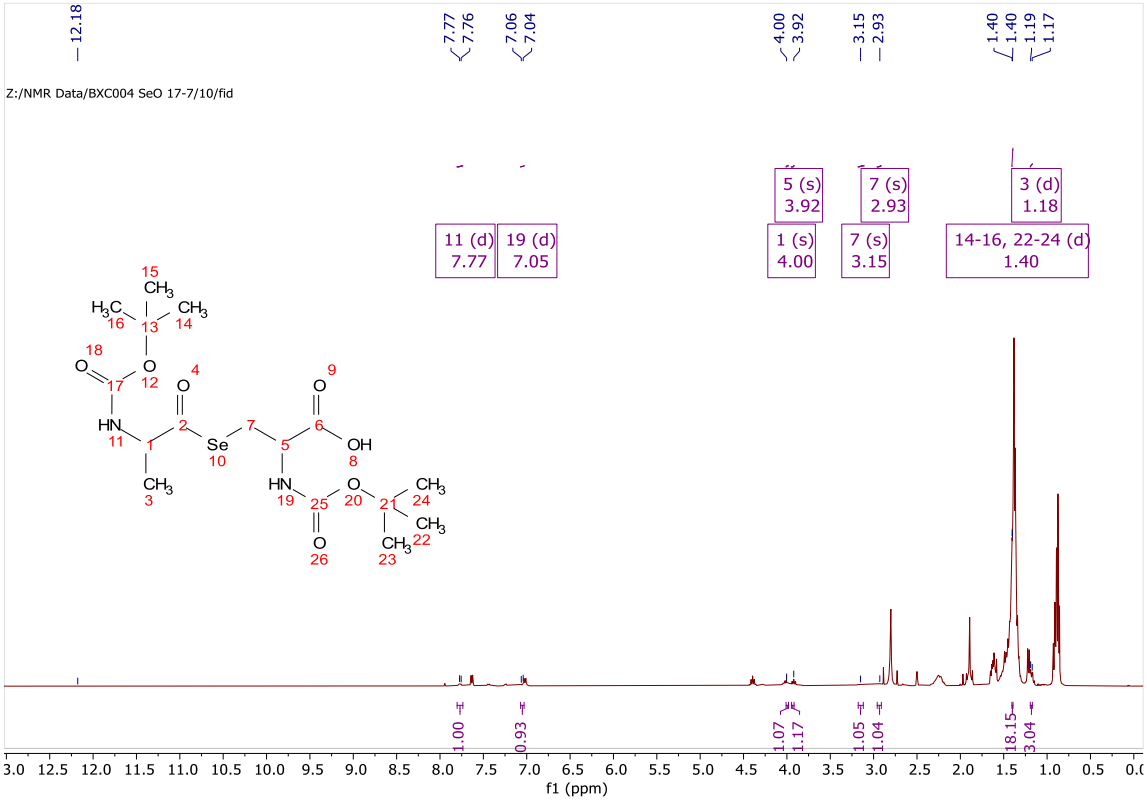
***Supplementary Figure 3-20**: Aminolysis of selenoester **3-3** with Asp (**3-9f**), Boc-Lys (**3-9c**) and Gly (**3-9a**) at various pHs over 2 weeks. Purple square = peptide aminolysis product (**3-9**). Green diamond = Boc-Phe-OH **3-5** hydrolysis product in aminolysis reaction. Grey square = Boc-Phe-OH **3-5** formed during selenoester **3-3** hydrolysis studies.*



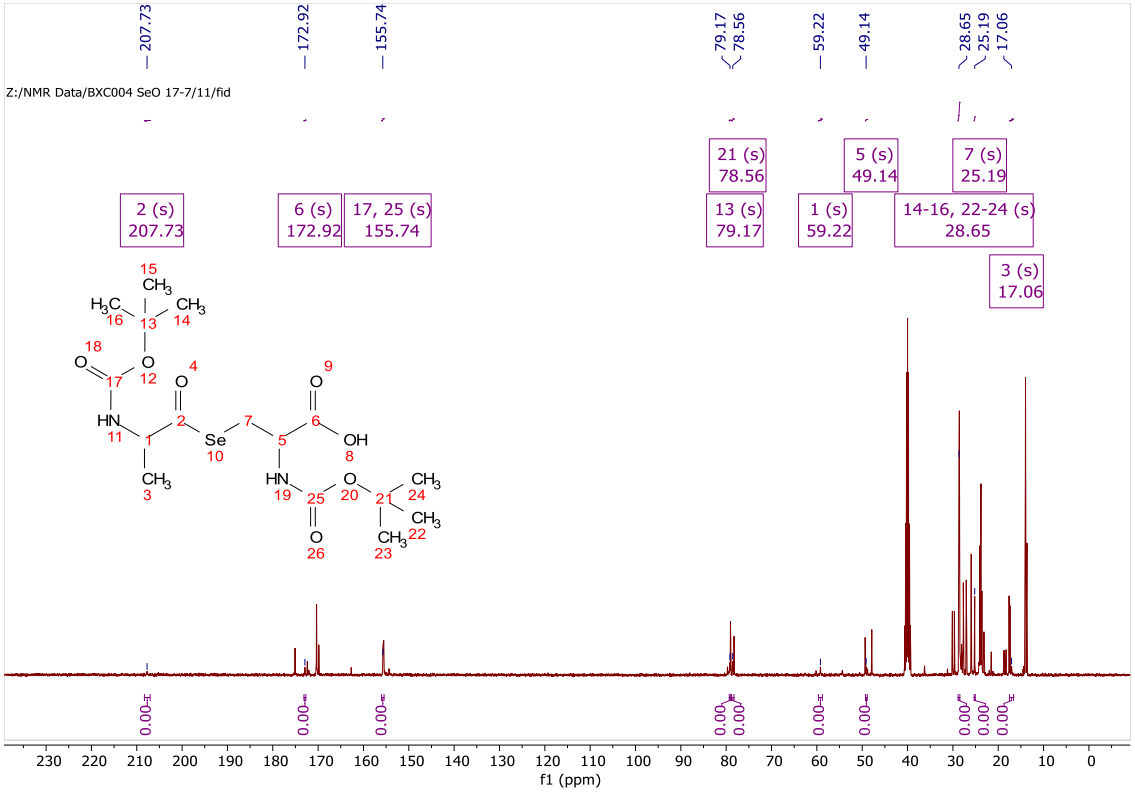
Supplementary Figure 3-21*: Aminolysis of selenoester* ***3-3*** *with Lys (****3-9h****) and Phe (****3-9e****) at various pHs over 2 weeks. Purple square = peptide aminolysis product (****3-9****). Green diamond = Boc-Phe-OH* ***3-5*** *hydrolysis product in aminolysis reaction. Grey square = Boc-Phe-OH* ***3-5*** *formed during selenoester* ***3-3*** *hydrolysis studies.*



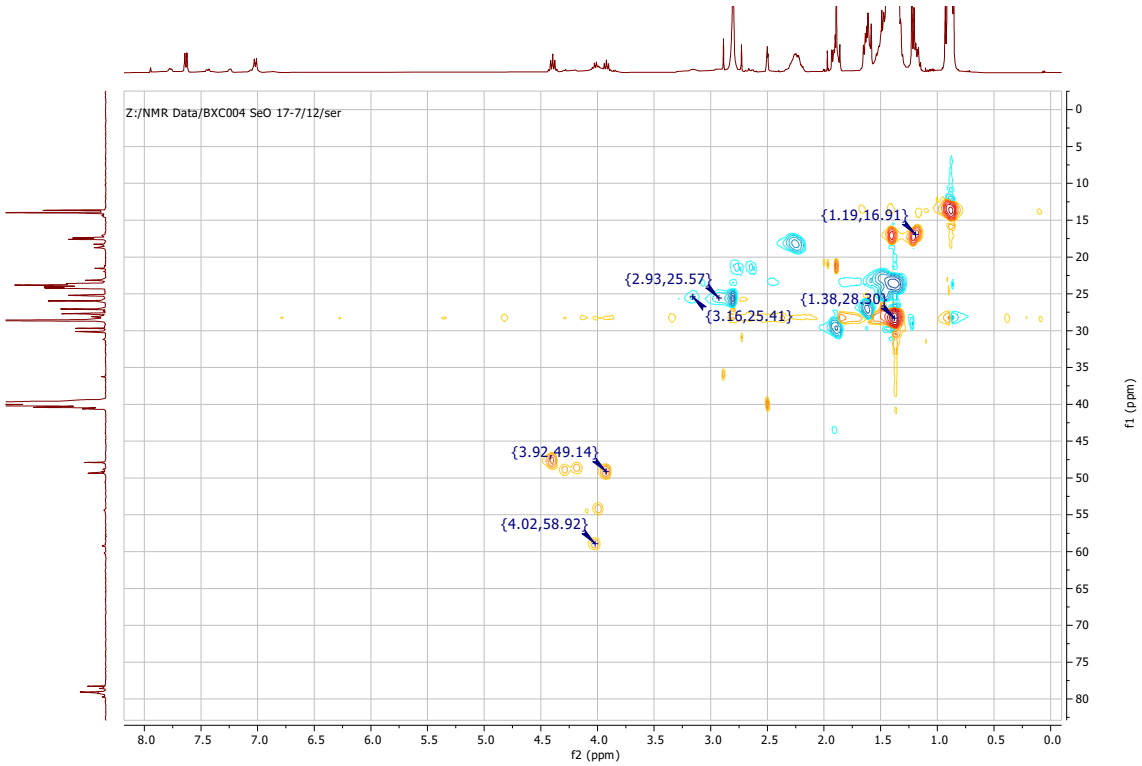
***Supplementary Figure 3-22**: Calibration curves for aminolysis peptide products **3-9** and hydrolysis byproduct **3-5** with slope and R^2 annotated. **3-9a**, **b** and **c** at λ = 254 nm and **3-9d** and **e** at λ = 210 nm.*



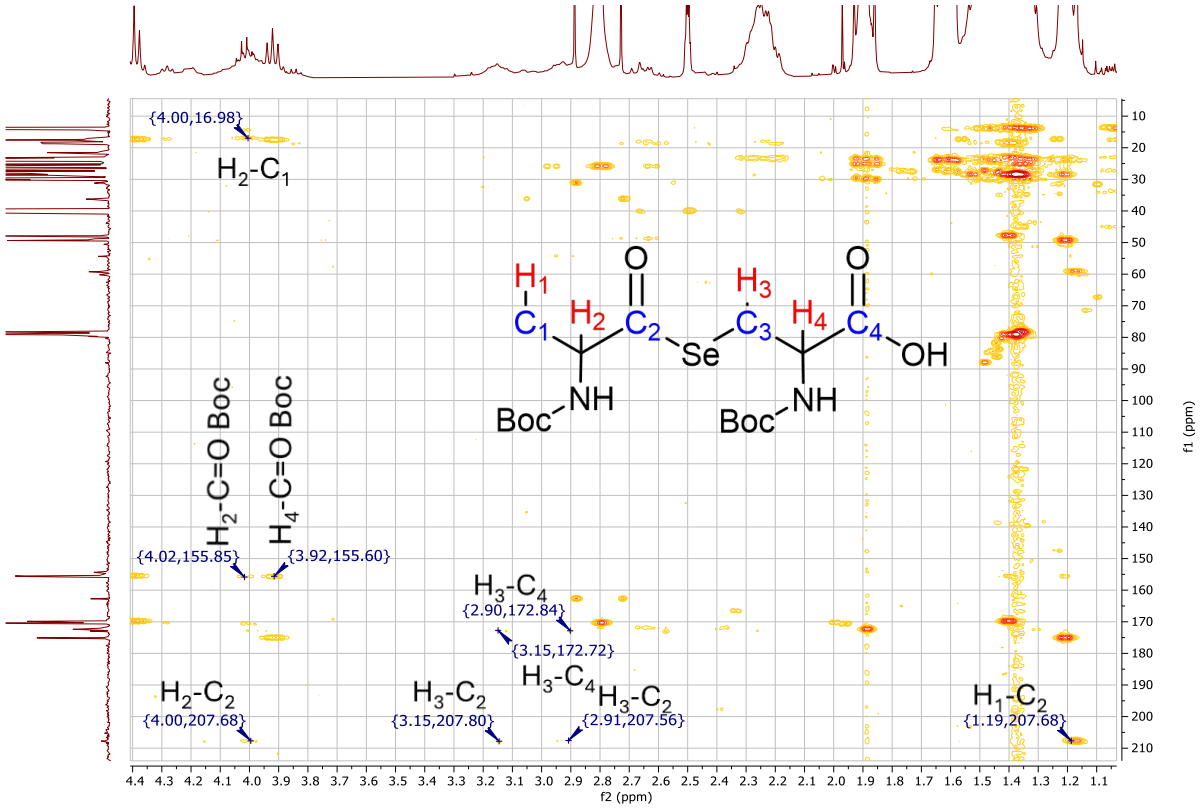
Supplementary Figure 3-23: *^{1}H NMR spectrum of selenoester **3-12**, 400 MHz, DMSO-d_6*



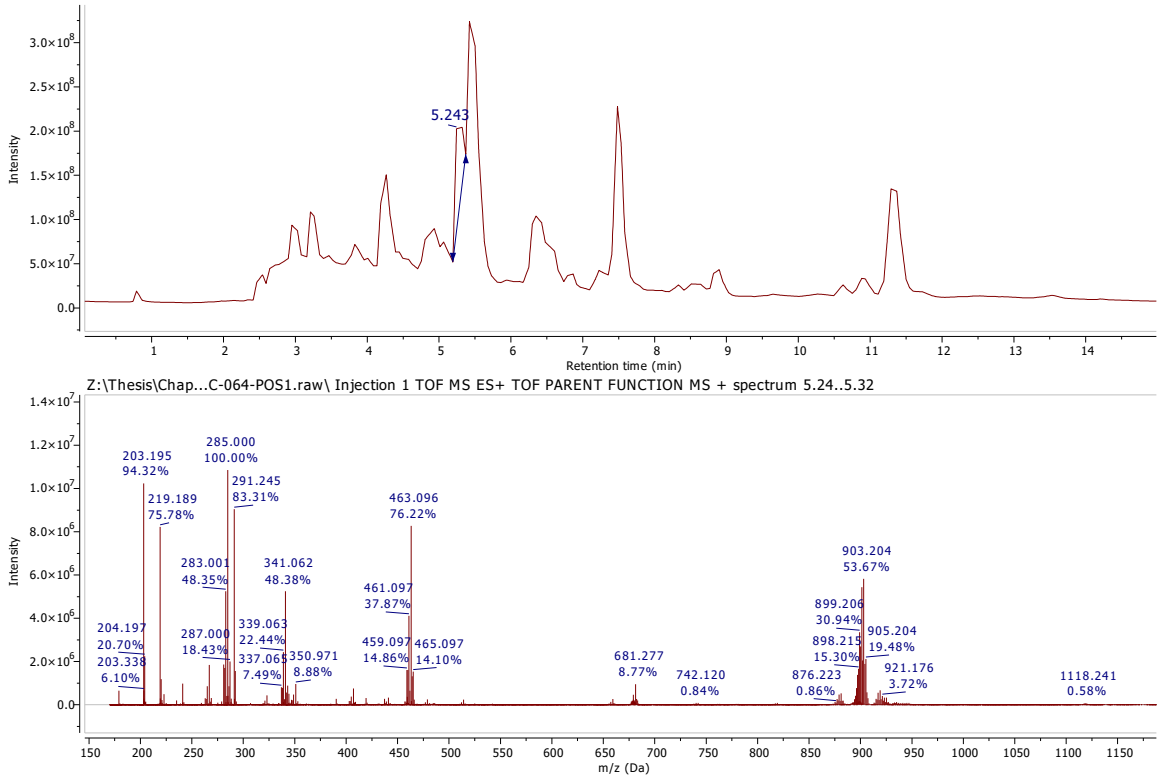
Supplementary Figure 3-24*: ^{13}C NMR spectrum of selenoester **3-12**, 101 MHz, DMSO-d_6*



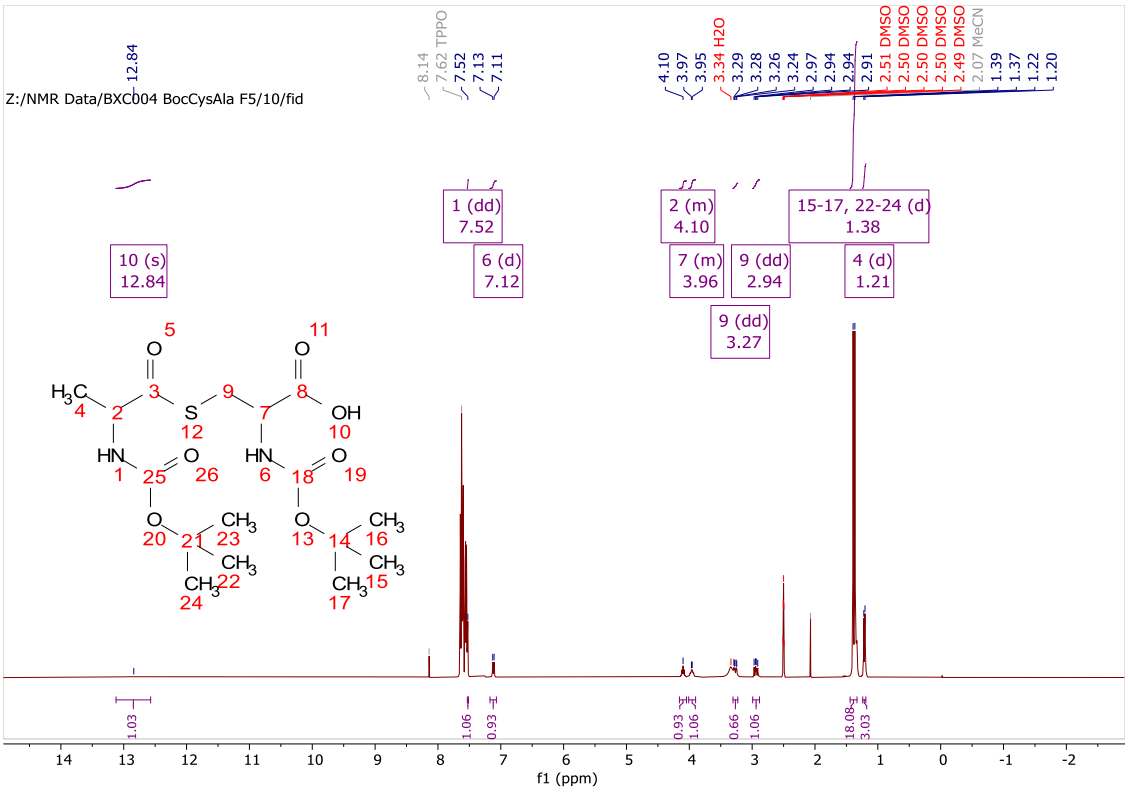
***Supplementary Figure 3-25**: HSQC spectrum of selenoester **3-12**, 400MHz/101 MHz, DMSO-d_6*



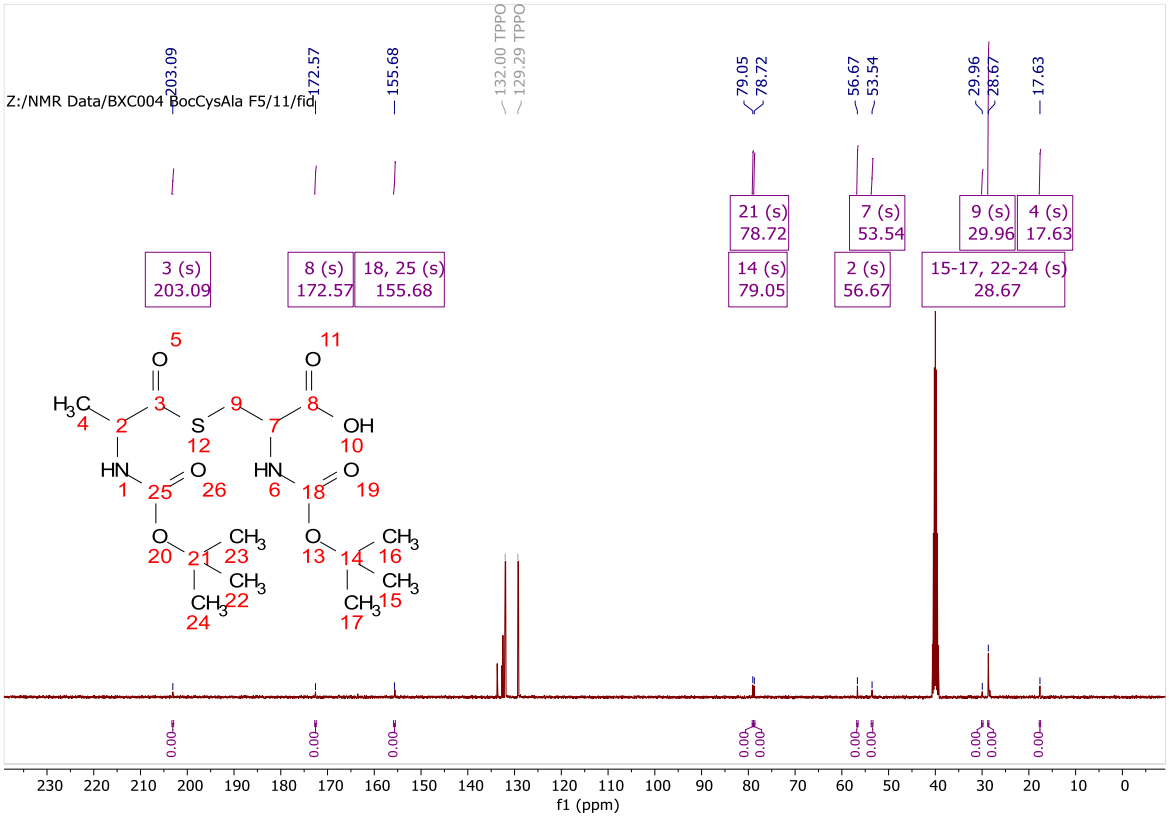
***Supplementary Figure 3-26**: Characterisation of key identifier in HMBC of selenoester **3-12**, 400MHz/101 MHz, DMSO-d_6*



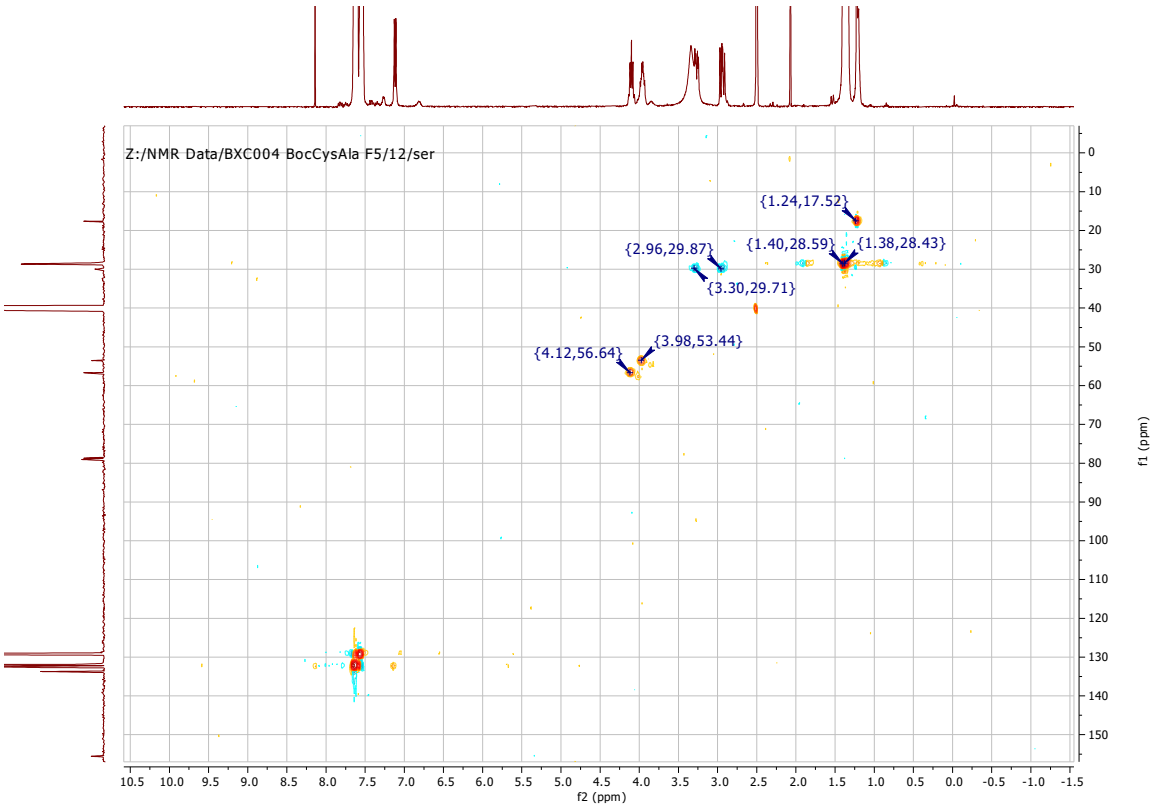
Supplementary Figure 3-27: *LC (above) TOF MS ES+ (below) of selenoester* ***3-12*** *$[M+Na]^+$ expected: 463.11, found: 463.10 m/z. 903.20 m/z = $[2M+Na]^+$*



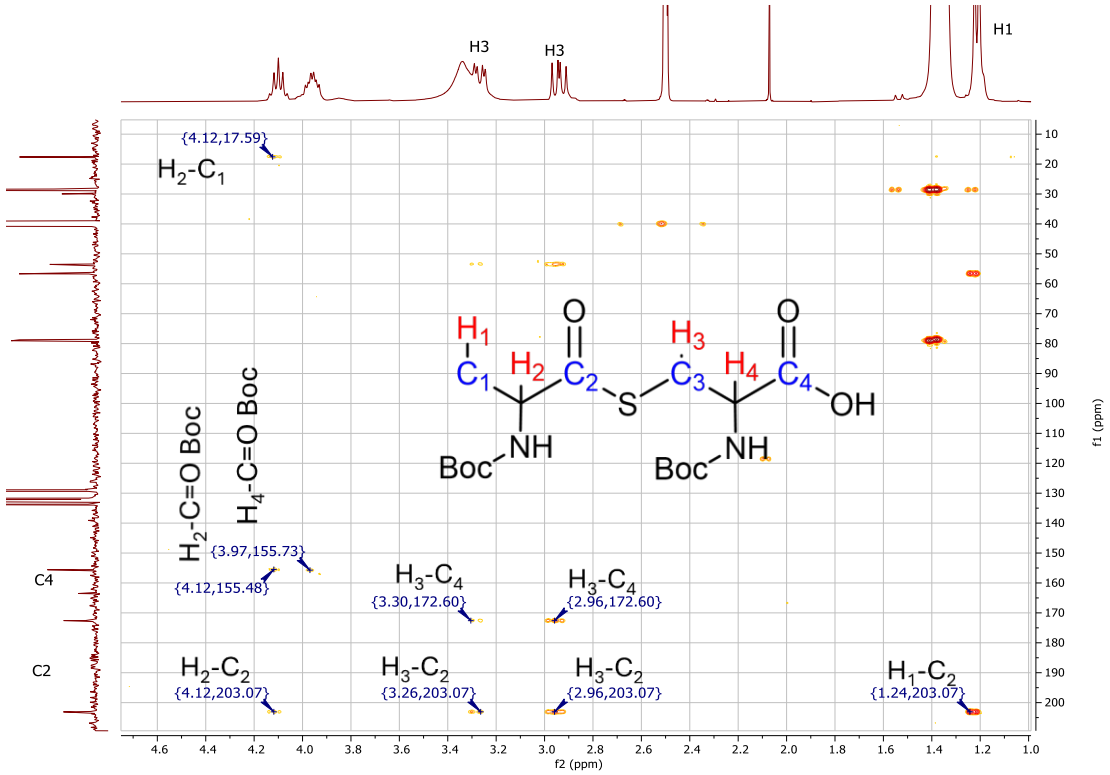
Supplementary Figure 3-28*: ^{1}H NMR spectrum of thioester* ***3-13****, 400 MHz, DMSO-d_6*



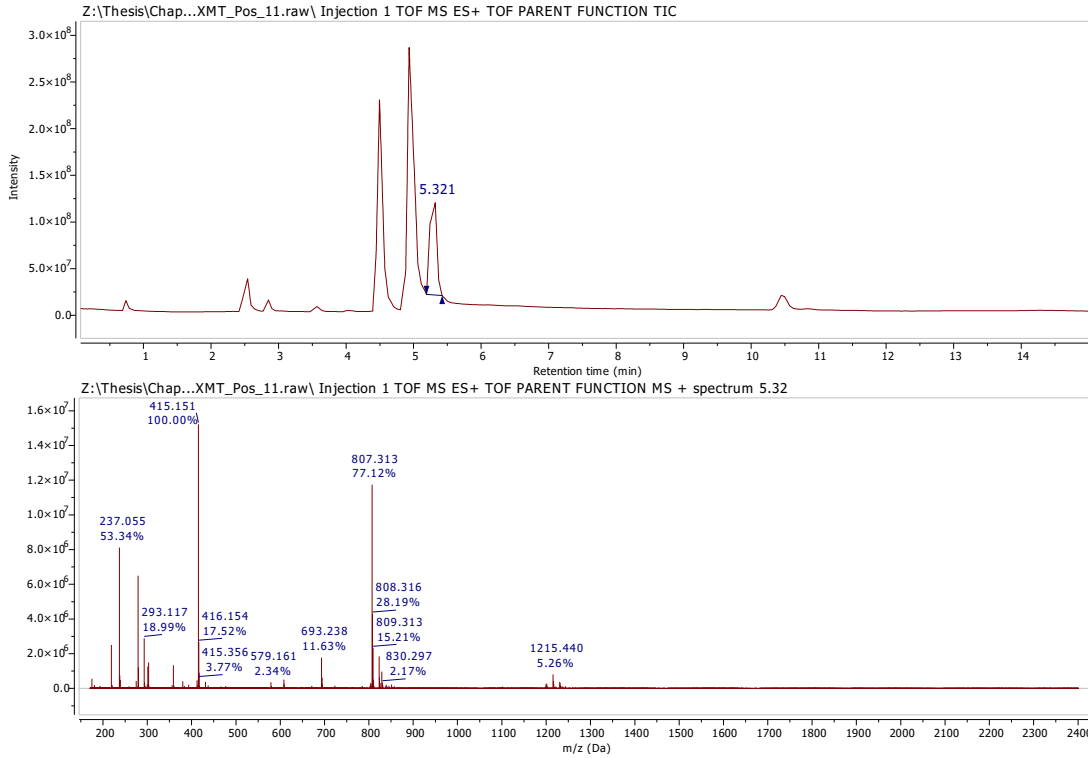
***Supplementary Figure 3-29**: ^{13}C NMR spectrum of thioester **3-13**, 101 MHz, DMSO-d_6*



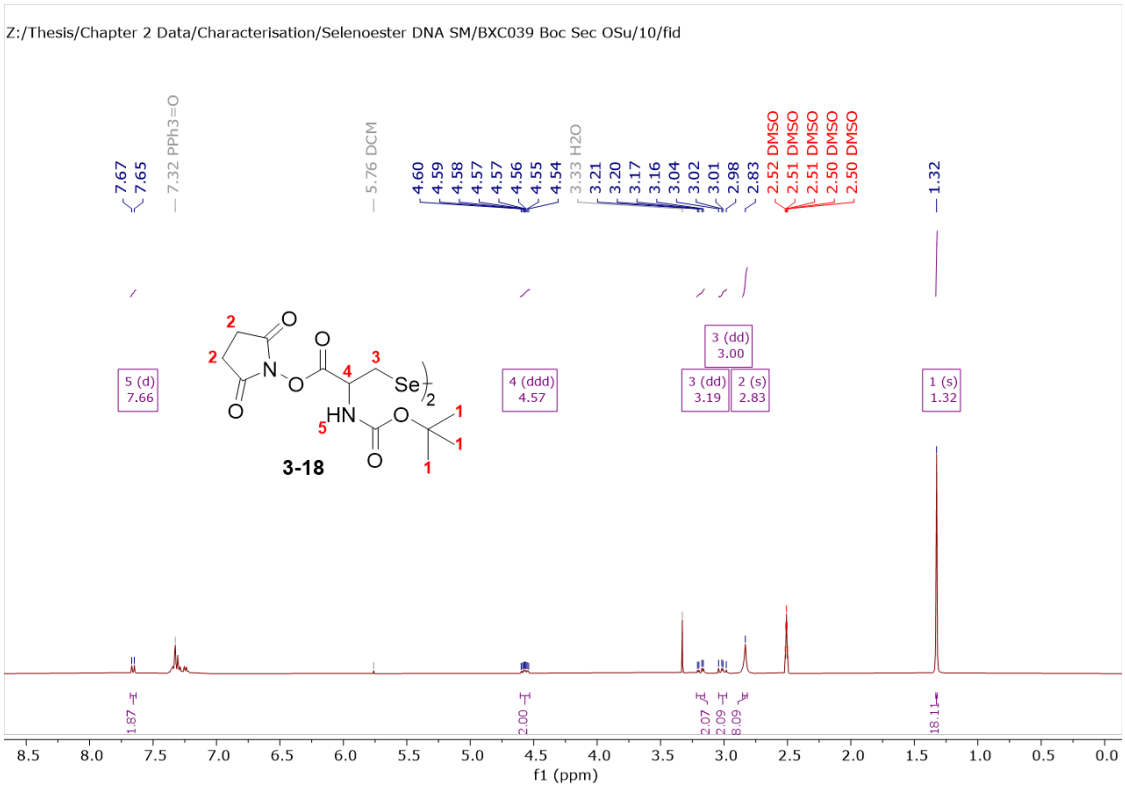
***Supplementary Figure 3-30**: HSQC spectrum of thioester **3-13**, 400MHz/101 MHz, DMSO-d_6*



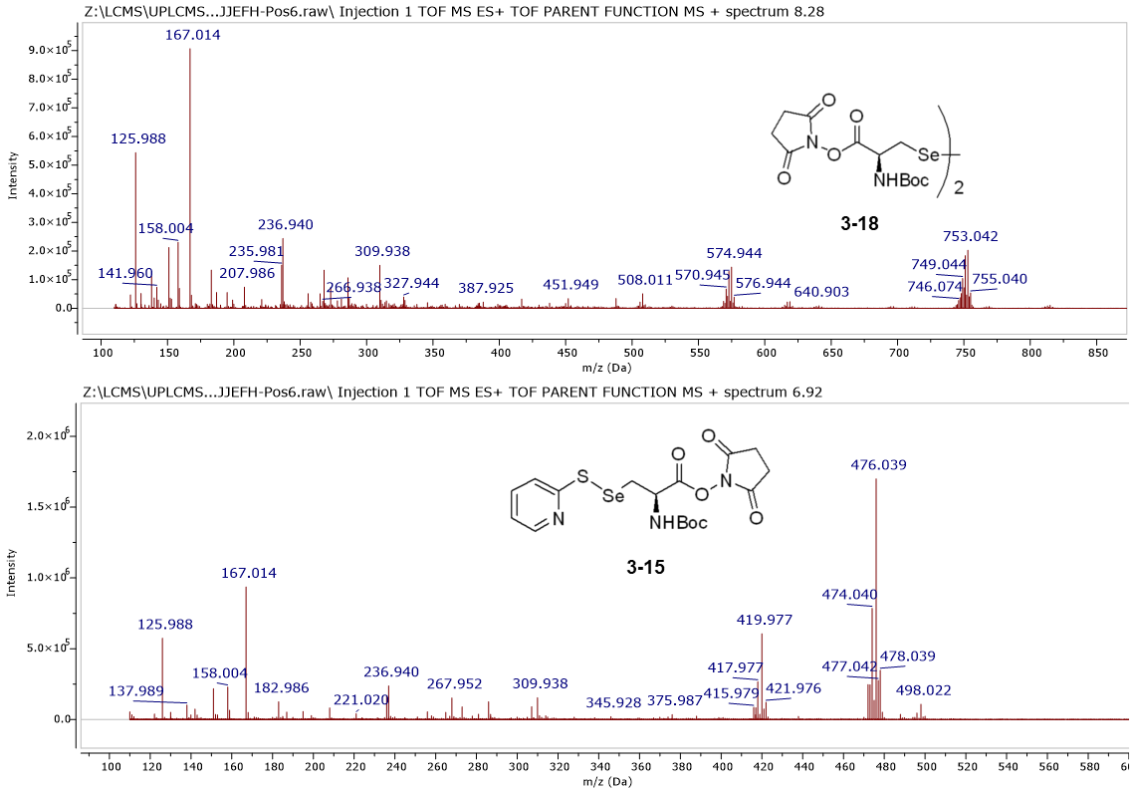
***Supplementary Figure 3-31**: Characterisation of key identifier in HMBC of thioester **3-13**, 400MHz/101 MHz, DMSO-d_6*



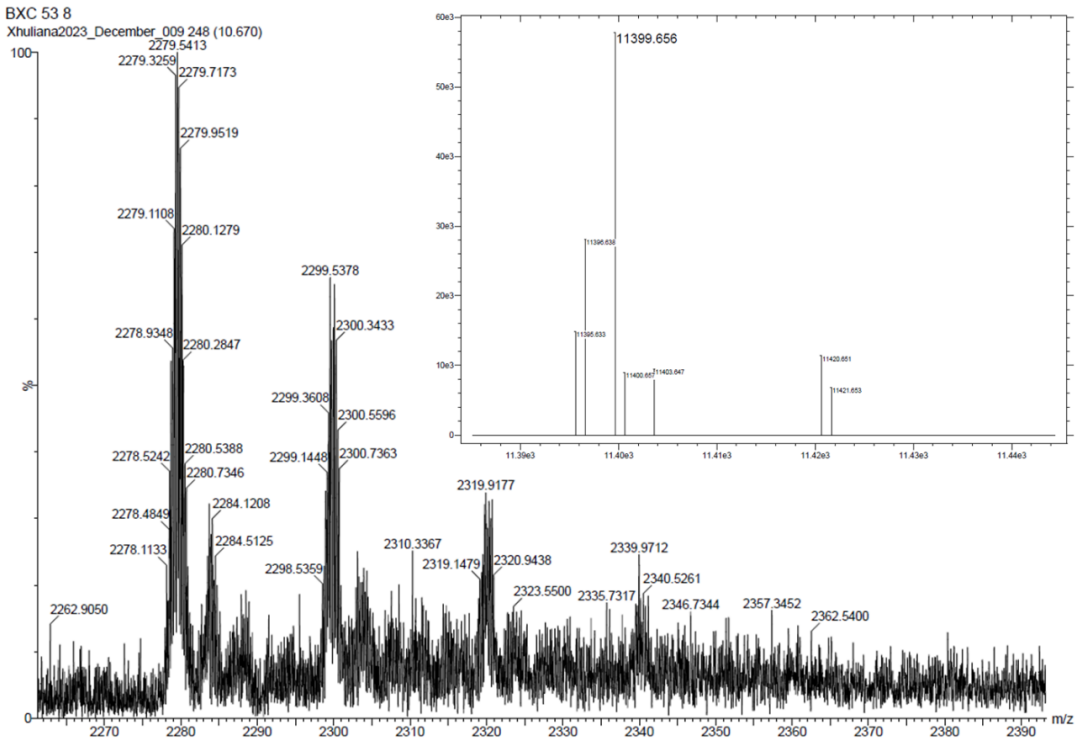
***Supplementary Figure 3-32:** LC (above) TOF MS ES+ (below) of thioester **3-13** $[M+Na]^+$ expected: 415.16, found: 415.15 m/z. 807.31 m/z = $[2M+Na]^+$*



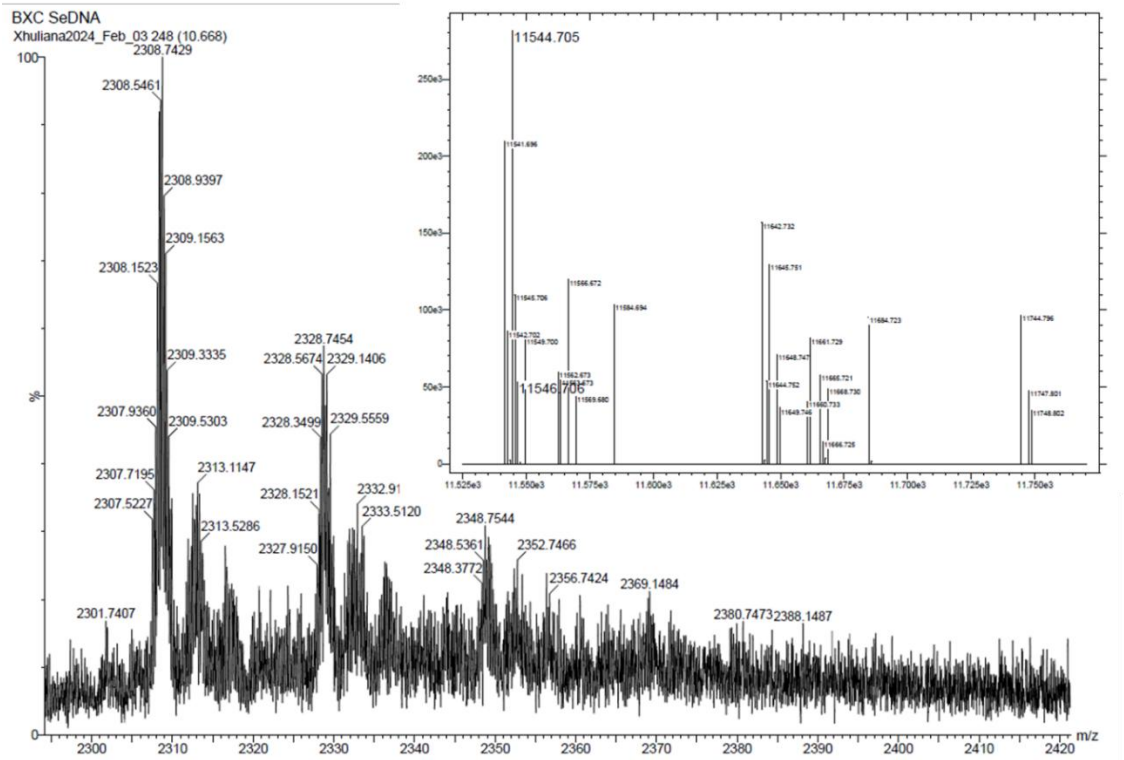
***Supplementary Figure 3-33**: [1]H NMR spectrum of **3-18**, 400 MHz, DMSO-d6. Impurity of triphenylphosphine oxide (PPh3=O) indicated due to silica column contamination.*



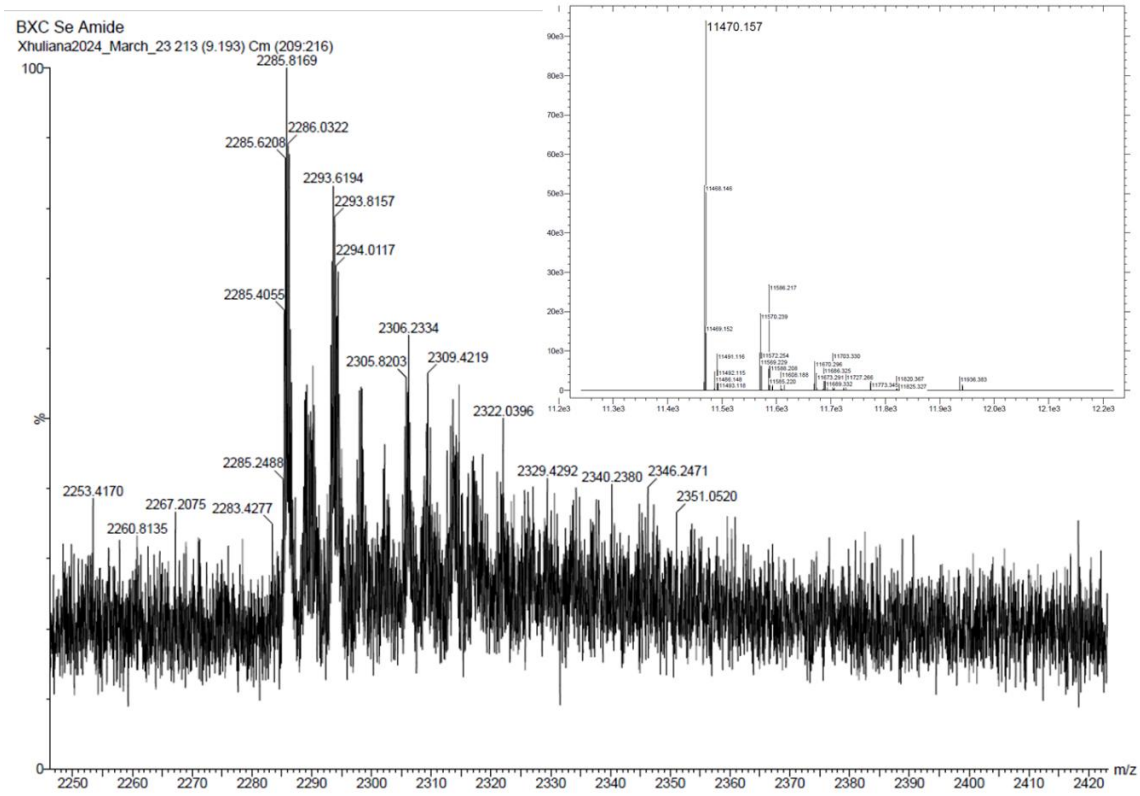
***Supplementary Figure 3-34**: TOF MS ES+ of **3-18** [M+Na]+ expected: 753.05, found: 753.04 m/z. and **3-15**: [M+H]+ expected: 476.03, found: 476.04 m/z.*



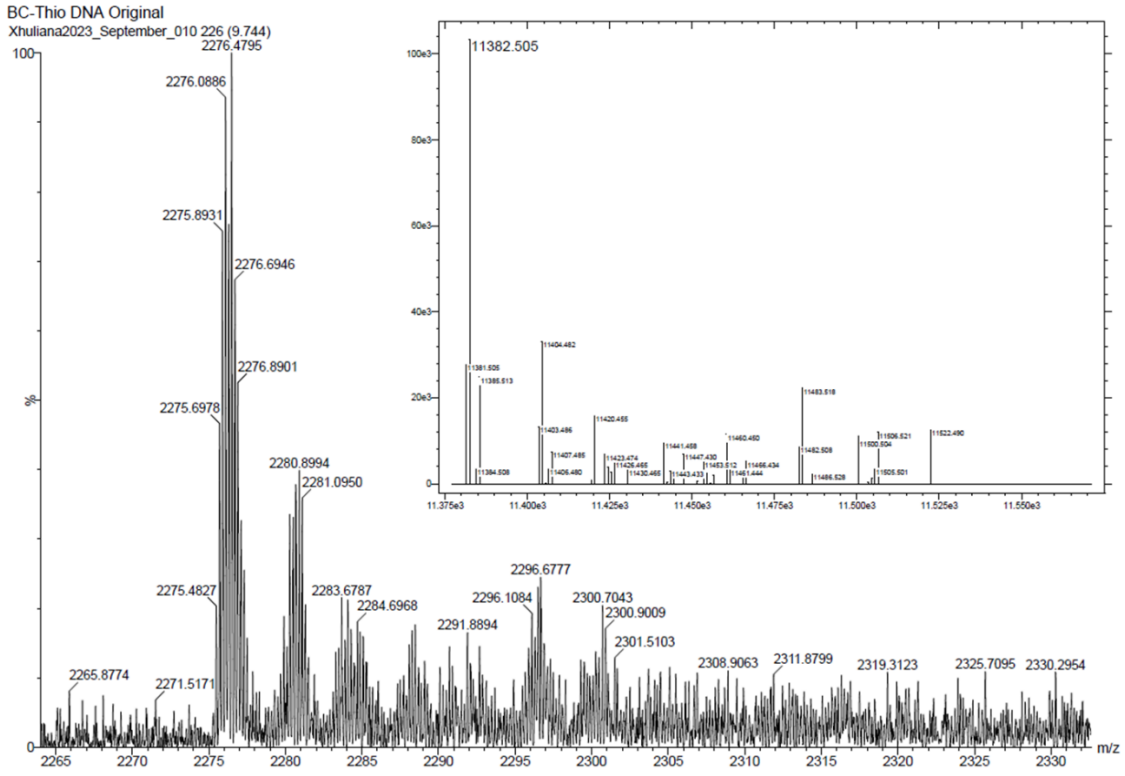
Supplementary Figure 3-35: *TOF MS ES− and deconvolution (ProMass HR) of Boc-Sec-DNA* ***3-20****. Expected Mass: 11399.89 m/z, found: 11399.66 m/z.*



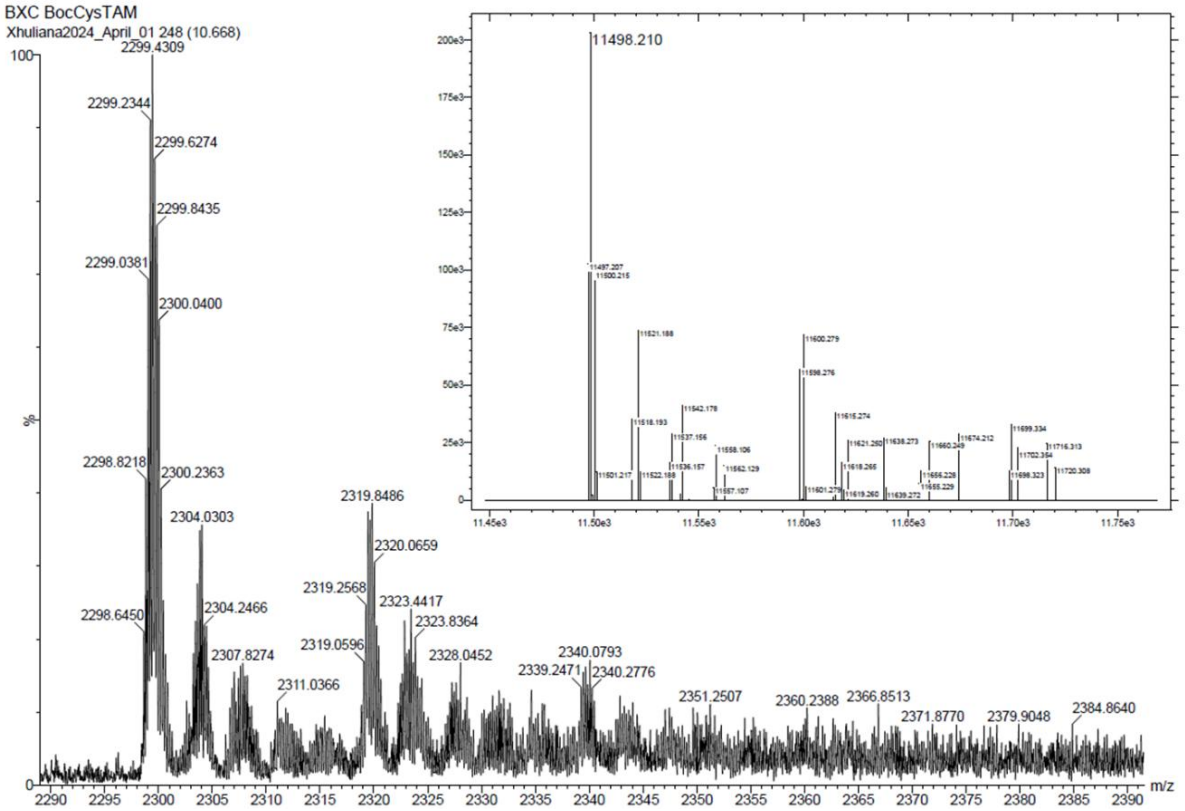
Supplementary Figure 3-36*: TOF MS ES− and deconvolution (ProMass HR) of TAMRA selenoester DNA* ***3-17****. Expected Mass: 11546.04 m/z, found: 11546.71 m/z.*



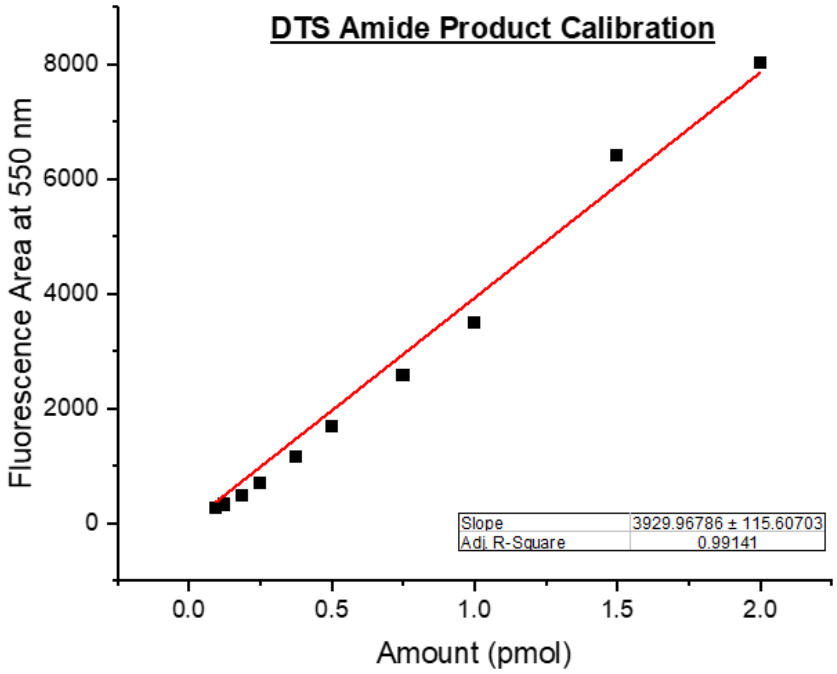
Supplementary Figure 3-37: *TOF MS ES− and deconvolution (ProMass HR) of DTS amide product* ***2-4****. Expected Mass: 11470.05 m/z, found: 11470.16 m/z.*



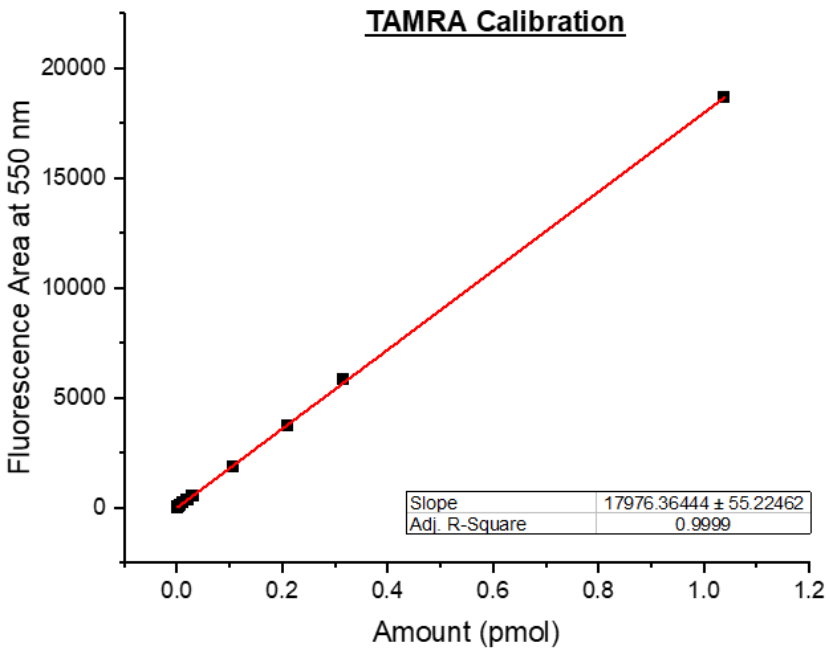
Supplementary Figure 3-38*: TOF MS ES− and deconvolution (ProMass HR) of TAMRA thioester DNA* ***2-1****. Expected Mass: 11383.03 m/z, found: 11382.51 m/z.*



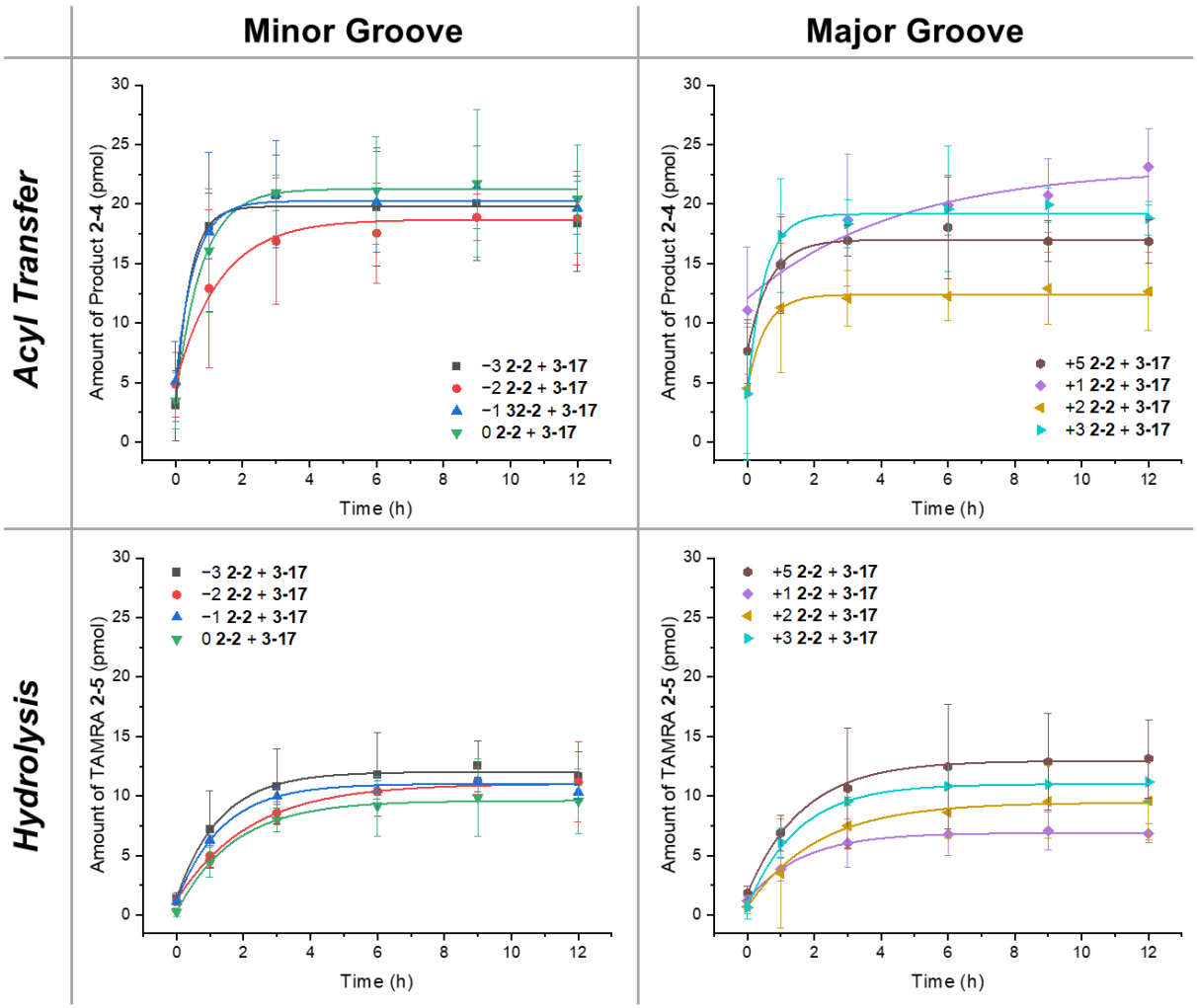
***Supplementary Figure 3-39**: TOF MS ES− and deconvolution (ProMass HR) of TAMRA Boc-Cys thioester DNA **3-23**. Expected Mass: 11498.09 m/z, found: 11498.21 m/z.*



***Supplementary Figure 3-40**: Calibration curve of DTS amide product **2-4** with slope and R^2 annotated.*



Supplementary Figure 3-41: *Calibration curve of free TAMRA* ***2-5*** *with slope and* R^2 *annotated.*



***Supplementary Figure 3-42**: DTS aminolysis progression over 12 hours of TAMRA selenoester DNA **3-17** with DNA-NH_2 **2-2**. Left = minor groove (**2-2** −n position), right = minor groove (**2-2** +n position), top = aminolysis product **2-4** progression (pmol), bottom = hydrolysis progression via TAMRA **2-5** production (pmol).*

3.6 References

1. J. Frommer, R. Oppenheimer, B. M. Allott, S. Núñez-Pertíñez, T. R. Wilks, L. R. Cox, J. Bath, R. K. O'Reilly and A. J. Turberfield, *Angew. Chem. Int. Ed.*, 2024, **163**, e202317482.
2. X. Li and D. R. Liu, *Angew. Chem. Int. Ed.*, 2004, **43**, 4848-4870.
3. M. L. McKee, A. C. Evans, S. R. Gerrard, R. K. O'Reilly, A. J. Turberfield and E. Stulz, *Org. Biomol. Chem.*, 2011, **9**, 1661-1666.
4. M. D. Gieselman, L. Xie and W. A. Van Der Donk, *Org. Lett.*, 2001, **3**, 1331-1334.
5. R. Quaderer, A. Sewing and D. Hilvert, *Helv. Chim. Acta*, 2001, **84**, 1197–1206.
6. R. J. Hondal, B. L. Nilsson and R. T. Raines, *J. Am. Chem. Soc.*, 2001, **123**, 5140-5141.
7. T. Durek and P. F. Alewood, *Angew. Chem. Int. Ed.*, 2011, **50**, 12042-12045.
8. M. Schrems, A. V. Kravchuk, G. Niederacher, F. Exler, C. Bello and C. F. W. Becker, *Chemistry*, 2023, **29**, e202301253.
9. J. A. Dean, *Lange's Handbook of Chemistry (15th Edition)*, McGraw-Hill Education, 15 edn., 1999.
10. M. Raj, H. Wu, S. L. Blosser, M. A. Vittoria and P. S. Arora, *J. Am. Chem. Soc.*, 2015, **137**, 6932-6940.
11. J. Sayers, R. J. Payne and N. Winssinger, *Chem. Sci.*, 2018, **9**, 896-903.
12. C. Liczner, C. C. Hanna, R. J. Payne and C. J. Wilds, *Chem. Sci.*, 2022, **13**, 410-420.
13. M. J. Bedding, S. S. Kulkarni and R. J. Payne, in *Selenoprotein Structure and Function*, ed. E. Weerapana, Academic Press, Cambridge, U.S., 2022, vol. 662, ch. 15, pp. 363-399.
14. R. Houska, M. B. Stutz and O. Seitz, *Chem. Sci.*, 2021, **12**, 13450-13457.
15. T. N. Grossmann and O. Seitz, *J. Am. Chem. Soc.*, 2006, **128**, 15596-15597.
16. J. Michaelis, A. Maruyama and O. Seitz, *Chem. Commun.*, 2013, **49**, 618-620.
17. G. Nakamura, K. Nakatsu and G. Hayashi, in *Methods Enzymol.*, ed. D. Chenoweth, Academic Press, 2024, vol. 698, pp. 169-194.
18. P. Čapek, H. Cahová, R. Pohl, M. Hocek, C. Gloeckner and A. Marx, *Chem. Eur. J.*, 2007, **13**, 6196-6203.
19. M. El Khatib, L. Jauregui, S. R. Tala, L. Khelashvili and A. R. Katritzky, *MedChemComm*, 2011, **2**, 1087.
20. T. C. Stadtman, *Annu. Rev. Biochem*, 1990, **59**, 111-127.

21. T. C. Stadtman, *Annu. Rev. Biochem*, 1996, **65**, 83-100.

22. *Japan Pat.,* US2012282386A1, 2012.

23. S. Trunschke, E. Piemontese, O. Fuchs, S. Abboud and O. Seitz, *Chem. Eur. J.*, 2022, **28**, e202202065.

24. U. Singh, S. K. Ghosh, M. S. Chadha and V. R. Mamdapur, *Tetrahedron Lett.*, 1991, **32**, 255-258.

25. W. Meng, R. A. Muscat, M. L. McKee, P. J. Milnes, A. H. El-Sagheer, J. Bath, B. G. Davis, T. Brown, R. K. O'Reilly and A. J. Turberfield, *Nat. Chem.*, 2016, **8**, 542-548.

26. W. Yang and D. G. Drueckhammer, *J. Am. Chem. Soc.*, 2001, **123**, 11004-11009.

27. W. Yang and D. G. Drueckhammer, *Org. Lett.*, 2000, **2**, 4133-4136.

28. T. Yamada, Y. Watanabe and S. Okamoto, *RSC Advances*, 2021, **11**, 24588-24593.

29. L. C. Chan and B. G. Cox, *J. Org. Chem.*, 2007, **72**, 8863-8869.

30. N. Basilio, L. García-Río, J. C. Mejuto and M. Pérez-Lorenzo, *J. Org. Chem.*, 2006, **71**, 4280-4285.

31. N. Caldwell, P. S. Campbell, C. Jamieson, F. Potjewyd, I. Simpson and A. J. B. Watson, *J. Org. Chem.*, 2014, **79**, 9347-9354.

32. C. Sabot, K. A. Kumar, S. Meunier and C. Mioskowski, *Tetrahedron Lett.*, 2007, **48**, 3863-3866.

33. D. T. Nguyen, D. C. Lenstra and J. Mecinović, *RSC Adv.*, 2015, **5**, 77658-77661.

34. B. S. Vig, P. J. Lorenzi, S. Mittal, C. P. Landowski, H.-C. Shin, H. I. Mosberg, J. M. Hilfinger and G. L. Amidon, *Pharm. Res.*, 2003, **20**, 1381-1388.

35. J. R. Peacock, R. R. Walvoord, A. Y. Chang, M. C. Kozlowski, H. Gamper and Y.-M. Hou, *RNA*, 2014, **20**, 758-764.

36. M. Takahashi, I. Hirota, T. Nakano, T. Kotani, D. Takani, K. Shiratori, Y. Choi, M. Haba and M. Hosokawa, *Drug Metab. Pharmacokinet.*, 2021, **38**, 100391.

37. M. D. Pol, R. Thomann, Y. Thomann and C. G. Pappas, *J. Am. Chem. Soc.*, 2024, **146**, 29621-29629.

38. A. Erben, T. N. Grossmann and O. Seitz, *Biorg. Med. Chem. Lett.*, 2011, **21**, 4993-4997.

39. Y. He and D. R. Liu, *J. Am. Chem. Soc.*, 2011, **133**, 9972-9975.

40. R. K. O'Reilly, A. J. Turberfield and T. R. Wilks, *Acc. Chem. Res.*, 2017, **50**, 2496-2509.

41. Z. Fang, Y. Dantsu, C. Chen, W. Zhang and Z. Huang, *ChemistrySelect*, 2023, **8**, e202302253.

42. A. Christofferson, L. Zhao, H. Sun, Z. Huang and N. Huang, *J. Phys. Chem. B*, 2011, **115**, 10041-10048.

43. D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, W.H. Freeman, 5 edn., 2008.

44. Dustin B. McIntosh, G. Duggan, Q. Gouil and Omar A. Saleh, *Biophys. J.*, 2014, **106**, 659-666.

45. S. Popovic, H. Bieräugel, R. J. Detz, A. M. Kluwer, J. A. A. Koole, D. E. Streefkerk, H. Hiemstra and J. H. Van Maarseveen, *Chem. Eur. J.*, 2013, **19**, 16934-16937.

46. M. Sohora, N. Vidović, K. Mlinarić-Majerski and N. Basarić, *Res. Chem. Intermed.*, 2017, **43**, 5305-5320.

Chapter 4

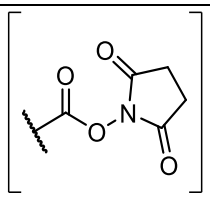
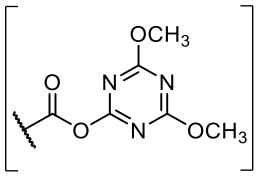
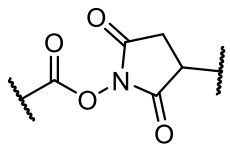
Optimisation of selenoester-mediated DNA-templated synthesis towards a multistep synthesis strategy

4.1 Introduction

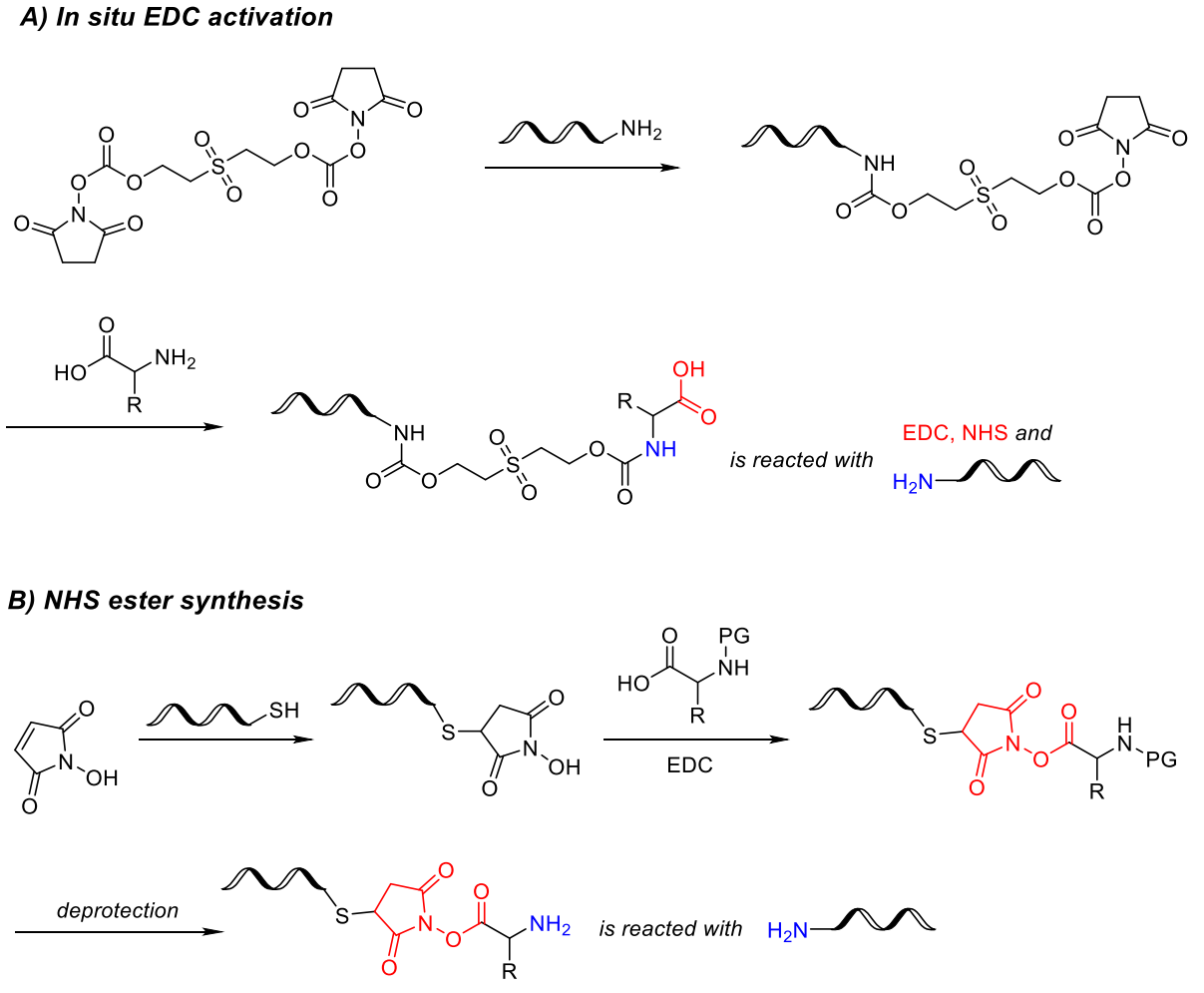
4.1.1 Aminolysis multistep DNA-templated synthesis

As described in Chapter 1, activated esters are successful electrophiles in performing aminolysis-mediated multistep DTS to yield modified oligomers. Using different DTS multistep architectures, active esters have transferred multiple functionalities such as alcohols, alkenes, amides and phenyl groups. (Table 4-1).[1-7]

Table 4-1: *Summary table of activated esters used in aminolysis-mediated multistep DTS.*

Electrophile	Nucleophile	Functionalities	Architecture	Product
 *In situ*, EDC/NHS	1º Amine	Amino acids	Strand addition and removal (biotin), end of helix	Peptide (3 %)[1, 2]
		Amino acids, alkenes		Macromolecules (trimer)[3]
 *In situ*, DMTMM	1º Amine	Amino acids	3-way junction	Peptide[4]
	1º Amine	Phenyl groups, alkenes	Strand displacement, end-of-helix	Hexamer (35%)[5]
		Phenyl groups	DNA walker	Trimer (~45 %)[6]
			Hairpin chain reaction (HCR)	Polypeptides (5-10%)[7]

To prepare the multistep activated esters, either NHS esters were utilised on 1,1′-[Sulfonylbis(2,1-ethanediyloxycarbonyloxy)]di(2,5-pyrrolidinedione) (BSOCOES) (Scheme 4-1A)[1-3] or through a bioconjugation between thiol-modified DNA and a maleimide (Scheme 4-1B).[4-6] Chemical tags functionalised with an amine and carboxylic acid were coupled to the DNA modifications to act as a transferable group. For the synthesis of NHS ester reactive tags, primary amines required protective groups, such as *O*-nitroveratryloxycarbonyl (NVOC), to prevent side reactions with the active ester during multistep DTS.



***Scheme 4-1**: Synthesis of aminolysis multistep DTS electrophiles: **A**) in situ EDC coupling[1-4] or **B**) NHS ester synthesis[5-7]. Electrophilic activated ester (red) and nucleophile (blue) are highlighted. R = transferable chemical tag. PG = protecting group.*

For the *in situ* activation method, the BSOCOES linker provides an oxoester cleavage site to expose a second primary amine in order to perform subsequent transfer steps in the multistep mechanism (Scheme 4-1A).[1-3] Despite the strategy limiting hydrolysis through *in situ* EDC activation, the breaking of the ester was completed on avidin beads, limiting the strategy's capabilities as an autonomous multi-step mechanism. On the other hand, the NHS ester synthesis provides a simplified, multistep approach to inserting an activated ester within the reactive tag through a Michael addition (Scheme 4-1B).[4-7] However, it would be complex to replace the NHS ester with more superior electrophiles, such as thio- and selenoesters, in this synthetic route due to undesirable Michael addition with thiols and selenols during synthesis.

Despite their success, product truncation and ester hydrolysis have generated low yields in aminolysis multistep DTS.[6] It is hypothesised that improvements in reaction kinetics and limitations to ester hydrolysis would aid in a higher product yield. Following results documented in Chapter 3.2.6, selenoesters increased the initial and overall yield of a single-step aminolysis DTS reaction, as opposed to its thioester analogue. In addition, selenoesters have provided access to complex synthetic targets that are unachievable by NHS- and thio-esters due to their decreased reactivity.[8, 9] Therefore, this chapter of work aimed to explore the incorporation of selenoesters into a multistep DTS mechanism to improve aminolysis yields, limit the formation of undesirable truncated products, and potentially access new polymer functionalities. Furthermore, the across-the-helix architecture previously described in this thesis was utilised to investigate its capabilities in a multistep synthesis strategy. However, selenoesters still provide limitations due to their accelerated degradation in highly basic

conditions. To maximise the potential, single-step DTS optimisation of TAMRA selenoester DNA **3-17** was investigated before adapting to a multistep strategy.

4.2 Results and discussion

4.2.1 Optimisation of selenoester single-step DNA-templated synthesis

Performing the single-step DTS assay presented in Chapter 3, parameters were altered to determine the optimum conditions for the TAMRA selenoester DNA **3-17** (Figure 4-1). Parameters such as pH, incubation times and new nucleophiles were investigated before the development of a selenoester-mediated multistep DTS assay.

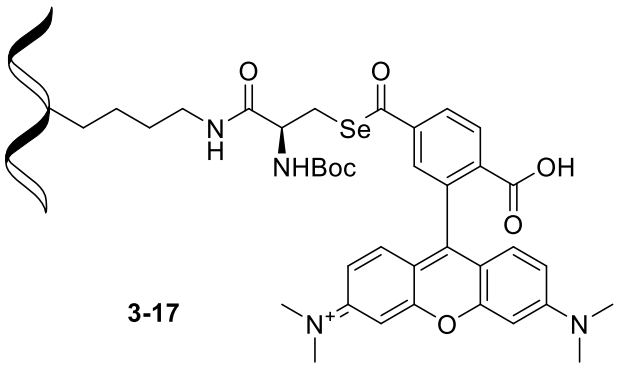
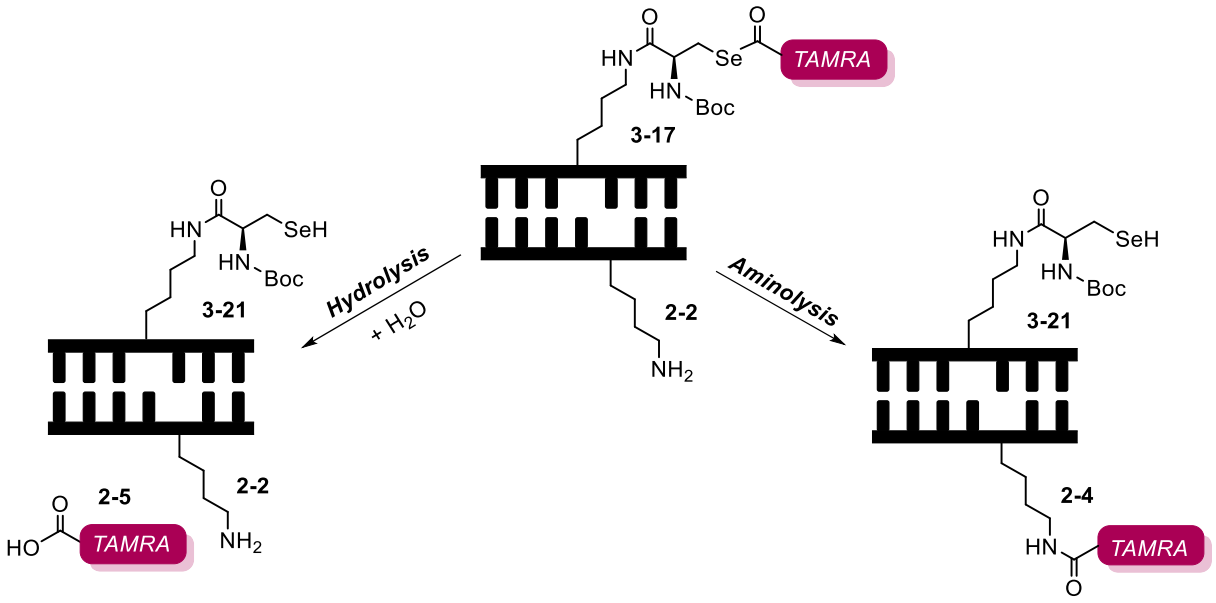


Figure 4-1: *Structure of TAMRA selenoester DNA* ***3-17.***

4.2.1.1 Optimisation of DTS buffer pH

As previously mentioned, basic conditions can lead to the degradation of the selenoester and the release of free TAMRA **2-5**. To limit selenoester hydrolysis whilst ensuring sufficient aminolysis, various pHs were investigated. To reduce the electrostatic repulsion of the negatively charged ssDNA, additives (*e.g*. NaCl) are required for DNA hybridisation.[10] The highly basic DTS buffer used for the initial DTS assay consisted of four salts: NaCl (0.25 M), TAPS (0.25 M), CAPS (0.25 M) and Na_3PO_4 (0.25 M), to ensure sufficient nucleophilic NH_2-DNA +1 **2-2** species. To not significantly impact the rate of hybridisation, a new buffer consisting of the same four

DTS salts was formulated at pH 10.2 by reducing the Na_3PO_4 concentration to 0.025 M. In addition to a pH 10.2 buffer, a salt solution was prepared of $MgCl_2$ (0.5 M) – the salt solution used for DTS hybridisation. As work presented in Chapter 3 showed successful aminolysis of a selenoester small molecule and amino acid as low as pH 6.5, the rate of aminolysis was investigated at pH 7.0 (Scheme 4-2). By selecting a neutral pH, the rate of TAMRA selenoester DNA **3-17** hydrolysis could potentially be controlled, whilst exploiting the few deprotonated $-NH_2$ molecules in solution.



Scheme 4-2: *Mechanisms of action: DTS aminolysis forming an amide bond* ***2-4*** *(right) of TAMRA selenoester DNA* ***3-17*** *forming* ***2-5*** *at the +1 NH_2 position* ***2-2****. Or free TAMRA* ***2-5*** *produced if subject to hydrolysis (left).*

Following the single-step DTS assay, TAMRA selenoester DNA **3-17** with DNA-NH_2 +1 **2-2** was subject to incubation at pH 10.2 (diluted DTS buffer) and pH 7.0 (0.5 M $MgCl_2$) (Figure 4-2).

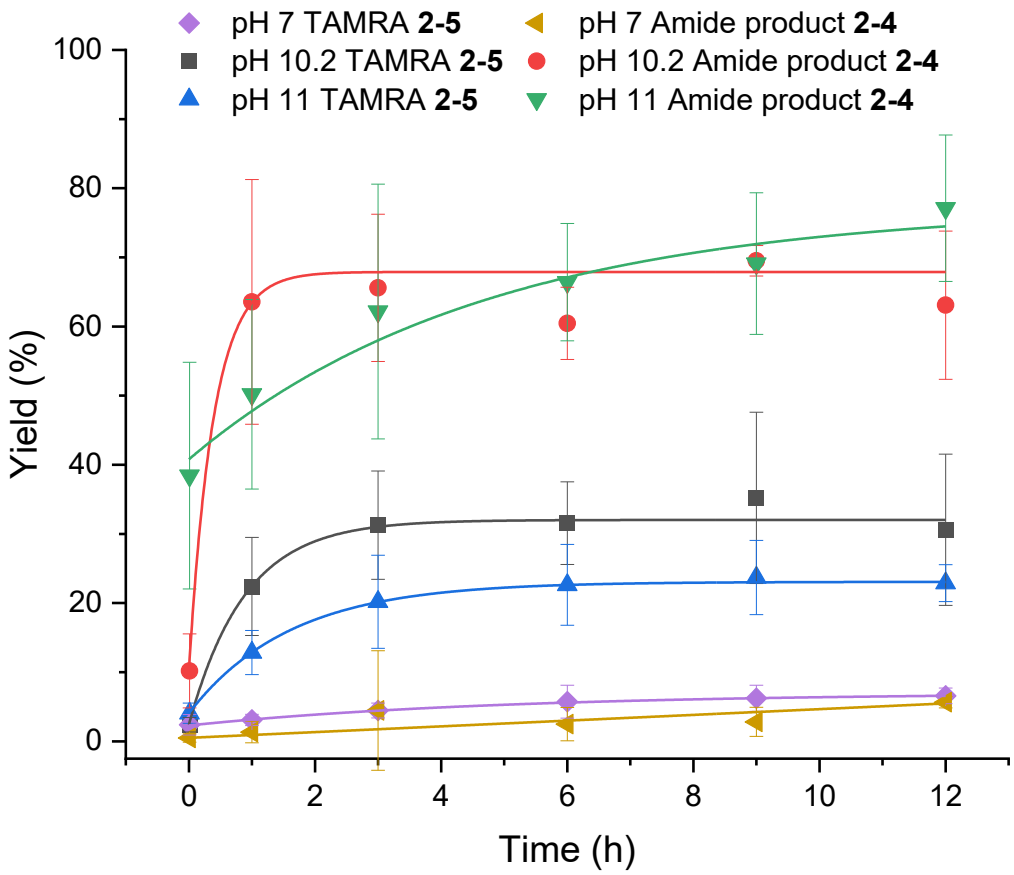


***Figure 4-2**: DTS aminolysis yield (%) progression (amide product **2-4**) and hydrolysis (free TAMRA **2-5**) over 12 hours between TAMRA selenoester DNA **3-17** with DNA-NH_2 +1 **2-2** in pH 7.0 $MgCl_2$ salt solution, pH 10.2 DTS diluted buffer and pH 11.0 DTS buffer. Error bars show the standard deviation in product **2-4** and free TAMRA **2-5** amount (%) of three repeats.*

After 12 hours of incubation, a yield of 6% amide product **2-4** was observed at pH 7.0. The slow reaction kinetics were due to the low concentration of deprotonated amine species in the neutral conditions, which presented an undesirable reaction time for a multistep system. On the other hand, statistically similar aminolysis yields at 63 ± 10% for pH 10.2 and 77 ± 10% for pH 11.0 were observed in both basic pH conditions at 12 hours of incubation. However, the amount of product formed at initial incubation differed between the two conditions. Upon instant dilution in the buffer, 10% of amide product **2-4** was formed at pH 10.2 compared to 38% at pH 11.0. As stated in Chapter 2.2.2.2, a high pH is necessary to ensure the availability of the $-NH_2$ nucleophile. Because fast reaction kinetics can minimise the degree of undesired

truncation within a multistep system, future work would continue to operate at a pH of 11.0.

4.2.1.2 Optimisation of DTS incubation

In the single-step DTS assay, TAMRA selenoester DNA **3-17** and DNA-NH_2 **2-2** at the +1 position were incubated together for 20 minutes at 15 ºC, before dissolution in the DTS buffer to ensure hybridisation. However, to reduce reaction times, the incubation time was varied and analysed *via* RP-HPLC to determine the minimum required time to ensure complete hybridisation for position +1 (Figure 4-3).

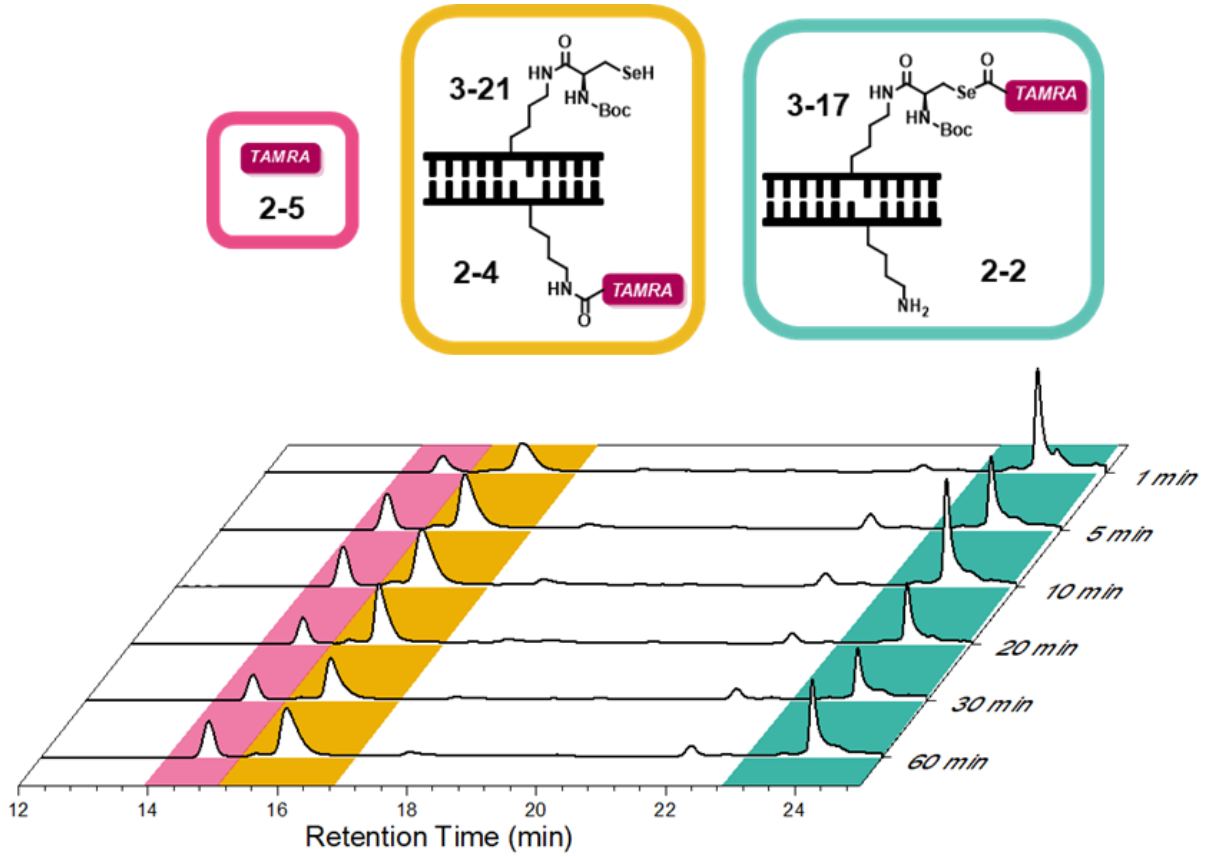
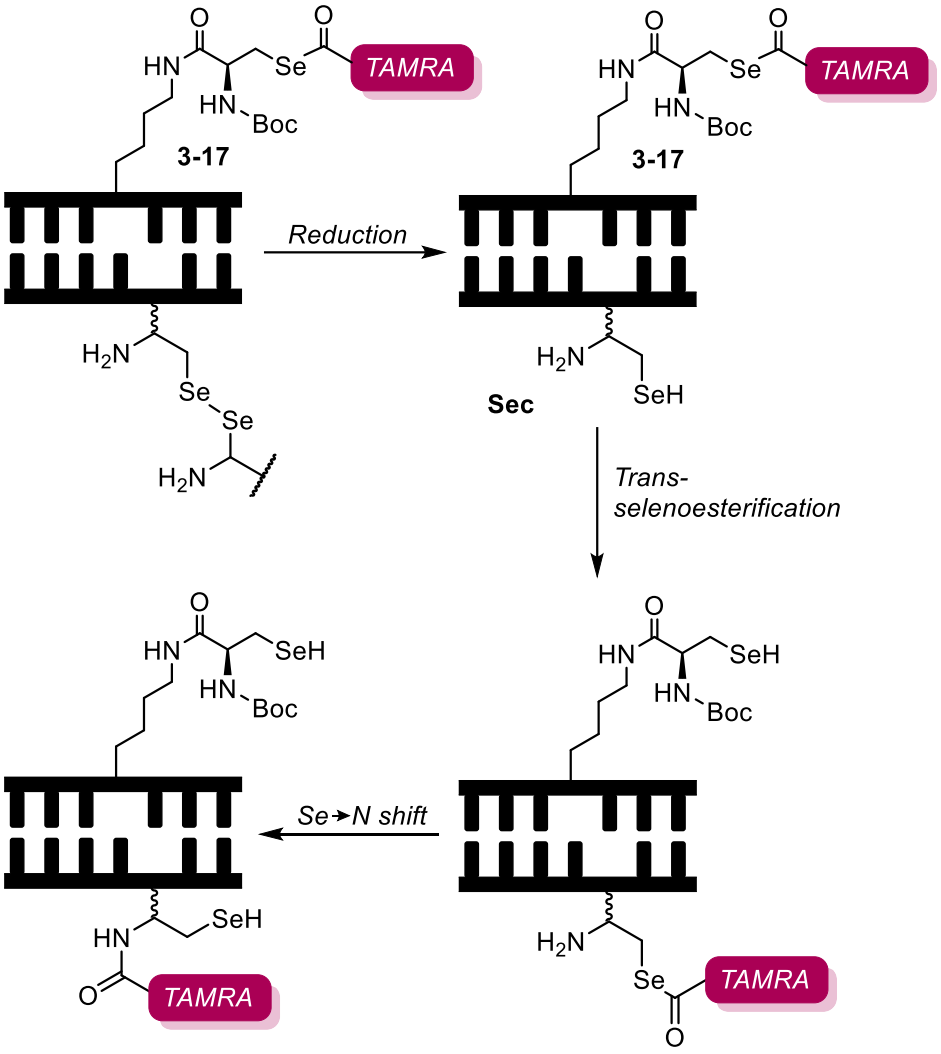


***Figure 4-3**: Waterfall plot of RP-HPLC spectrum of TAMRA selenoester DNA **3-17** initiated DTS in DTS buffer, pH 11.0, at various incubation times. Fluorescence channel: (λ_{ex} 550 nm, λ_{em} 580 nm). 2 µL injection volume. Pink = free TAMRA **2-5** (hydrolysis of **3-17**). Yellow = successful aminolysis (amide product **2-4** with spent adapter **3-21**). Green = starting material (TAMRA selenoester DNA **3-17** with primary amine, +1 position **2-2**).*

The incubation time before the addition of DTS buffer did not impact the rate of aminolysis, at 27-34% amide product **2-4** yield between 5 to 60 minutes. Considering previous buffer results (Figure 4-2), the data highlights the importance of a high pH to ensure nucleophile activation combined with a highly electrophilic localised selenoester to enable fast reaction kinetics – something observed across all incubation times. Proceeding on, initial incubation times would be reduced to decrease reaction times.

4.2.1.3 DTS nucleophile adjustment

To increase the reactive scope of TAMRA selenoester DNA **3-17**, an alternative nucleophile was investigated; selenocysteine (Sec) (Scheme 4-3). Differing from the primary amine DNA-NH_2 **2-2**, Sec would proceed *via* diselenide-selenoester ligation (DSL) (rather than aminolysis), one of the most effective synthetic strategies for short-chain peptide synthesis.[11] Therefore, if successful acyl transfer occurs between Sec and TAMRA selenoester DNA **3-17**, the catalogue of accessible attachment chemistries would increase.



Scheme 4-3: *Schematic of DSL of Sec bound to DNA with TAMRA selenoester DNA* ***3-17****.*

Implementation of Sec as a nucleophile in a single-step DTS reaction required its attachment to a modified ssDNA (Figure 4-4A). The small molecule selenocystine **3-10** was Fmoc-protected, followed by an *in situ* activation and coupling with the DNA-NH_2 +1 **2-2** strand used in the previous DTS assays. The Fmoc-Sec-DNA **4-2** was deprotected on a solid support, using piperidine **4-3**, to yield Sec-DNA **4-4a** (Figure 4-4B).

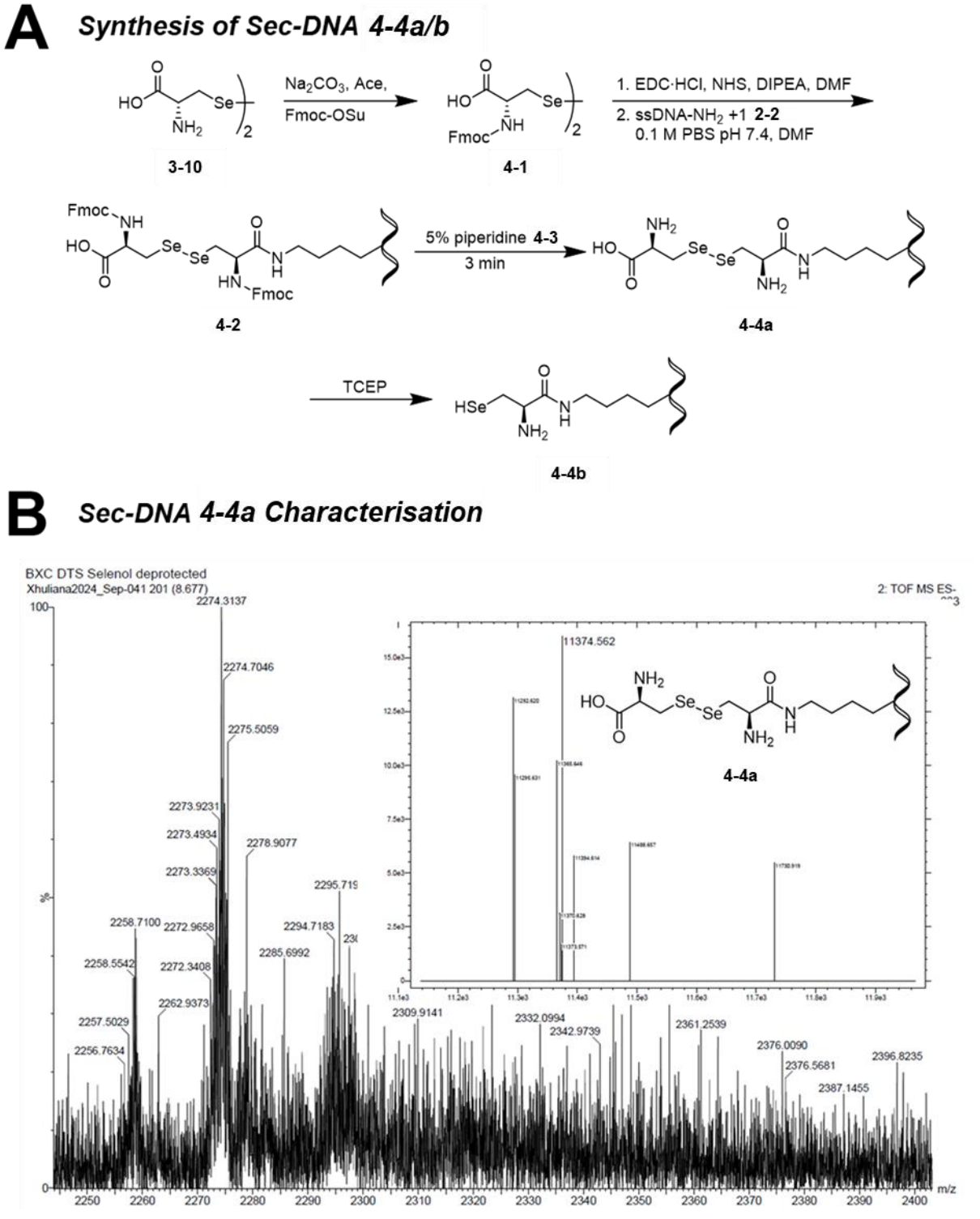


Figure 4-4: A) *Synthesis of Sec-DNA* ***4-4a/b.B)*** *TOF MS ES− and deconvolution (ProMass HR) of Sec-DNA* ***4-4a****. Expected Mass: 11374.81, found: 11374.56. Additional peaks correspond to selenium isotopes. See Experimental section 4.4.4.*

Sec-DNA **4-4a** was applied within the single-step DTS assay; however, an *in situ* diselenide reduction was incorporated to reveal the nucleophilic selenol of Sec. As

TCEP successfully reduced the diselenide in the esterification of TAMRA selenoester DNA **3-17** synthesis, it was selected for the reduction of Sec-DNA **4-4a** to **4-4b**. During incubation of the newly modified single-step DTS assay, TCEP (200:1 ratio with **4-4a**) was added to the solution containing Sec-DNA **4-4a** and TAMRA selenoester DNA **3-17**. Upon dilution with DTS buffer (15 µL), the reaction was tracked *via* RP-HPLC; however, no observable acyl transfer had occurred in the fluorescence channel (λ_{ex} 550 nm, λ_{em} 580 nm) (Figure 4-5).

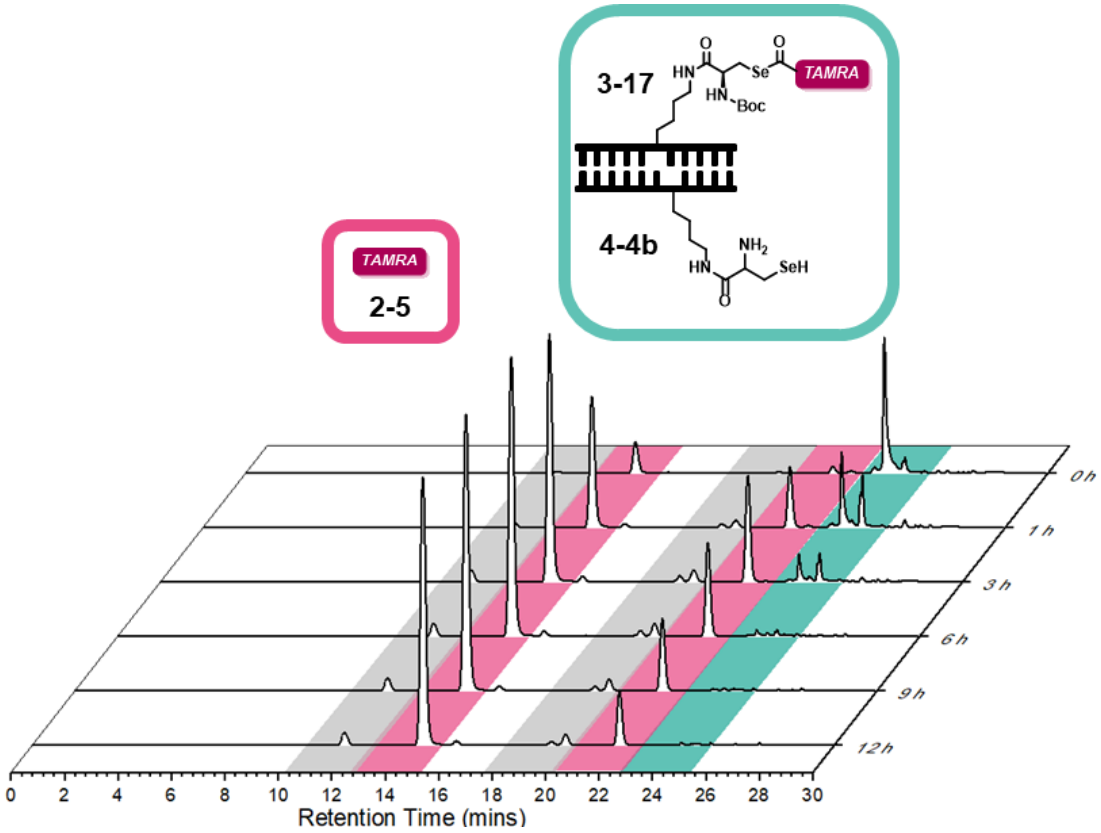


***Figure 4-5**: Waterfall plot of RP-HPLC spectrum of Sec-DNA **4-4b** with TAMRA selenoester DNA **3-17** in DTS buffer, pH 11.0, over the course of 12 hours. Fluorescence channel: (λ_{ex} 550 nm, λ_{em} 580 nm). 2 µL injection volume. Pink = free TAMRA **2-5** (hydrolysis of **3-17**). Green = starting material (TAMRA selenoester DNA **3-17** with Sec-DNA **4-4b**). Grey = DTS buffer salts.*

As for the UV-Vis DNA channel (λ = 262 nm), peak formation occurred at 18.4 minutes. The peak likely corresponded to both the degraded selenol adapter **3-21** of TAMRA selenoester DNA **3-17** and the newly modified Sec-DNA **4-4** due to their

matching functionalities (Figure 4-6). In addition, as no fluorescence (λ = 550 nm) was detected at this time, it proved difficult to distinguish between the modifications. Accordingly, it was concluded that acyl transfer between the DNA adapters was unsuccessful.

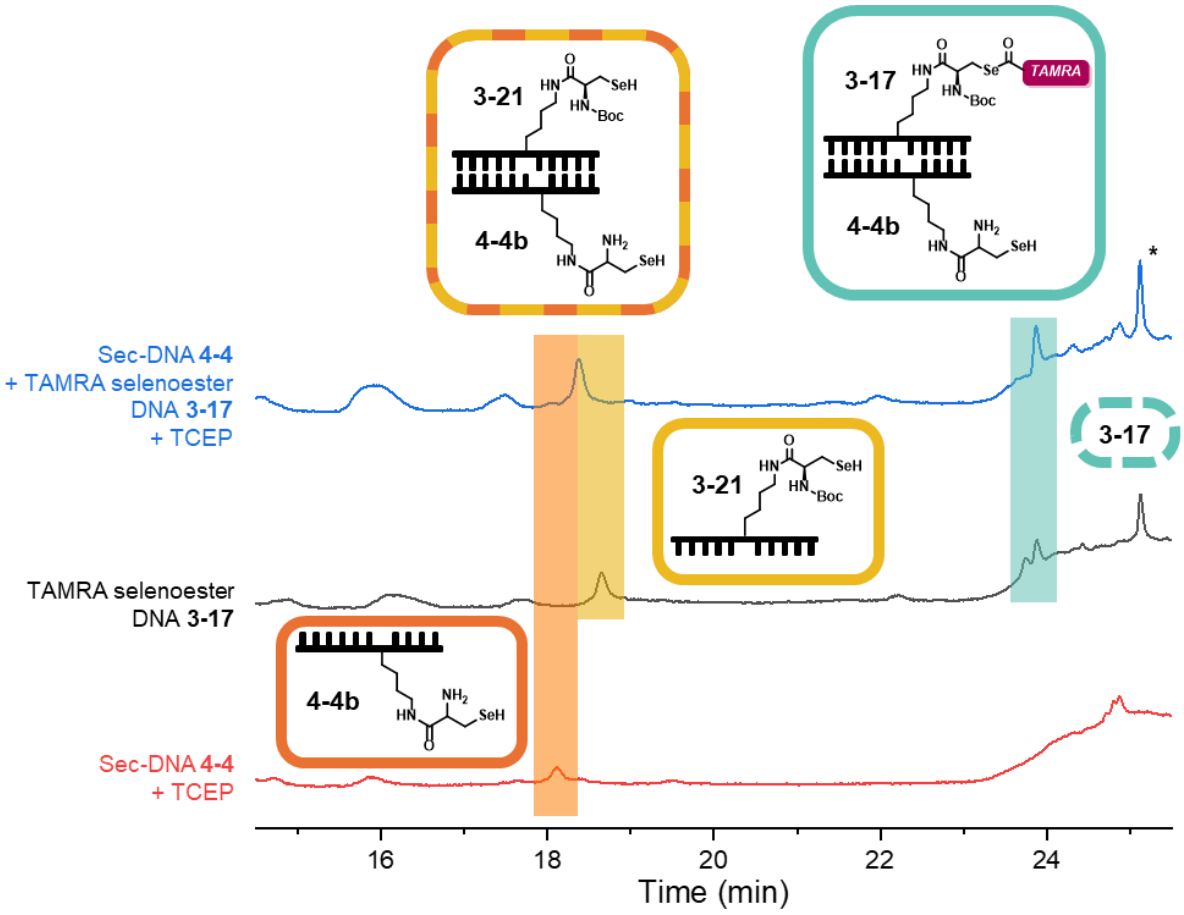
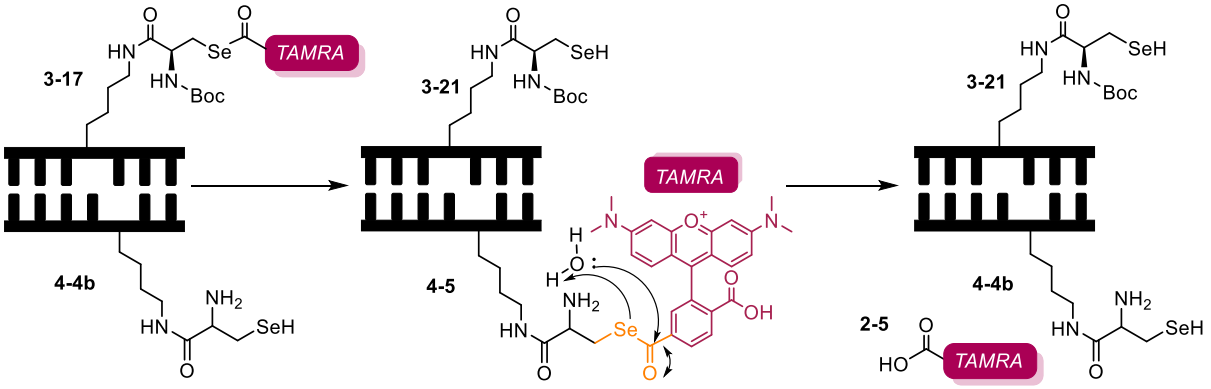


***Figure 4-6**: RP-HPLC analysis (λ = 262 nm) of Sec-DNA **4-4** & TAMRA selenoester DNA **3-17** DTS assay (blue line), compared to stability assay of ssDNA TAMRA selenoester DNA **3-17** (black line) and ssDNA TCEP reduction of Sec-DNA **4-4a** to **4-4b** (red line). Structures highlighted: dsDNA TAMRA selenoester DNA **3-17** with Sec-DNA **4-4b** (solid turquoise), ssDNA TAMRA selenoester DNA **3-17** (dashed turquoise), ssDNA Sec-DNA **4-4b** (solid orange), ssDNA hydrolysed selenol **3-21** (solid yellow – degradation of **3-17**), dsDNA hydrolysed selenol **3-21** with Sec-DNA **4-4b** (dashed yellow & orange).*

Based on the results, a hypothesis was constructed that the second selenoester **4-5** of the DSL mechanism degraded within the highly basic conditions, in conjunction with a sterically hindered primary amine, thus preventing peptide bond formation (Scheme 4-4). Typically, DSL is performed at a pH range of 3-7, as opposed to our

DTS system at pH 11.0, contributing to the degradation of selenoesters.[11] As a DSL intermediate selenoester **4-5**, its presence in a concentrated solution is transient (< 60 minutes, 1.4 mM)[12], causing difficulty in monitoring within this dilute, basic, DTS assay (2 μM). As for sterics, the primary amine of Sec-DNA **4-4b** is more sterically hindered to the amino acid framework, as opposed to the flexible primary amine of DNA-NH_2 **2-2**. In addition, the bulky TAMRA of TAMRA selenoester DNA **3-17** is transferred, as opposed to the smaller amino acid functionalities documented in DSL literature.[13]



***Scheme 4-4**: Hydrolytic route of DSL mechanism in DTS. Degradation of selenoester intermediate **4-5** of Sec-DNA **4-4b** transfer with TAMRA selenoester DNA **3-17**. TAMRA (pink) and selenoester (orange) highlighted.*

In conclusion, DSL would not be a suitable reactive pathway when designing a multistep DTS system using a TAMRA moiety and in basic conditions. However, the mechanism may show potential if buffers were modified to accommodate a neutral pH.[11]

4.2.2 Multistep system design

As optimal single-step DTS conditions were determined within the newly developed across-the-helix architecture, a mechanism of action was developed in

order to perform selenoester-mediated multistep DTS. The system was built upon the existing and optimised selenoester chemistries, including the incorporation of the TAMRA moiety used for fluorescence tracking *via* RP-HPLC (Figure 4-7). Amino acid-based chemical tags were selected to mimic the peptide synthesis of the ribosome. DNA strands would follow the same DNA sequence of the original TAMRA selenoester DNA **3-17** modification with the addition of an uncomplementary 3′ toehold region (10 bases). The across-the-helix architecture was employed, with the **Strand 2 4-8** modification positioned at the +1 site (Figure 4-7).

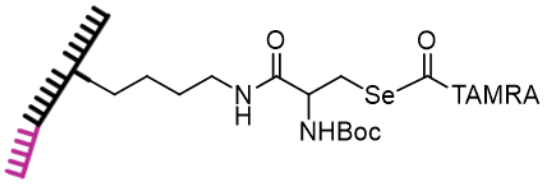
A

Identification	Sequence 5' → 3' X = modification
Strand 1	GCCGAGCCAGCAGTCAGCGC/**X**/GTCCTAATCTACCTGTTGGTTGGTT
Strand 2	CAGGTAGATTAGGA/**X**/AGCGCTGACTGCTGGCTCGGCGGGGATTGGT

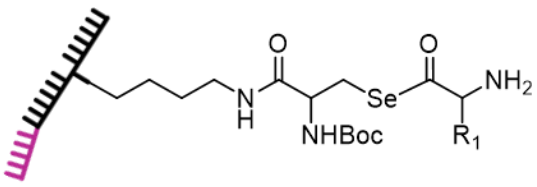
B

Strand 1:

R_1 = amino acid side chain

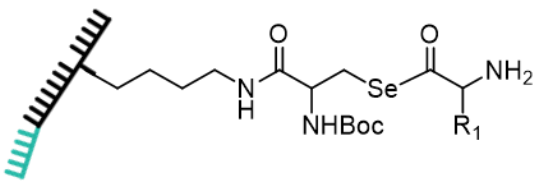


Strand 1-TAMRA-selenoester-DNA
4-6

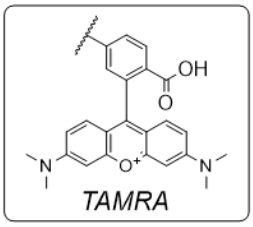


Strand 1-selenoester-amino acid-DNA
4-7

Strand 2:



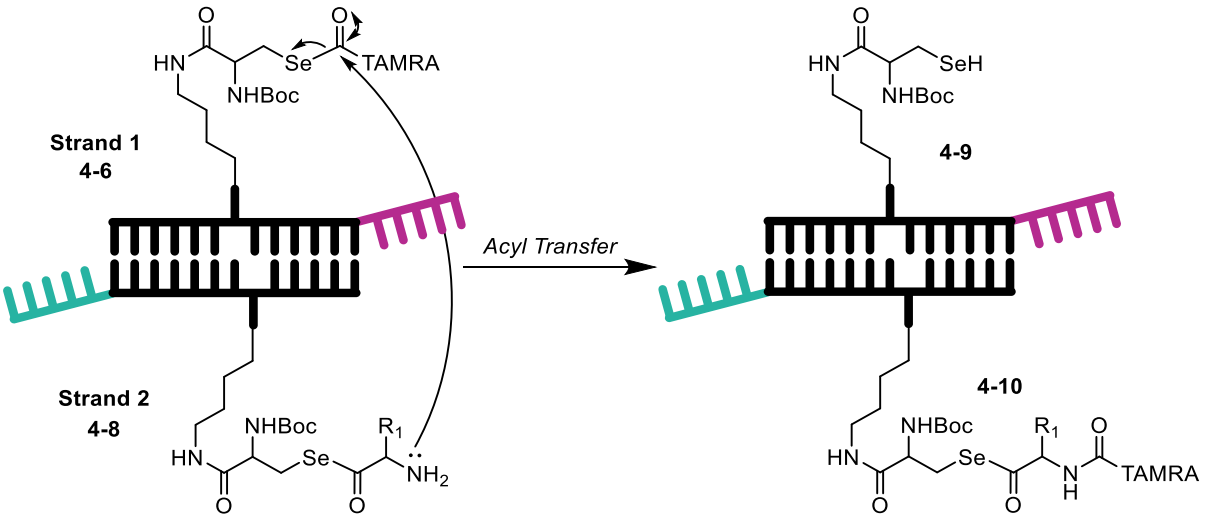
Strand 2-selenoester-amino acid-DNA
4-8



TAMRA

***Figure 4-7**: **A**) Sequences **of Strand 1** and **Strand 2** with the additional toe hold regions (magenta and turquoise) in comparison to the original sequences of **2-2** and **3-17**. X = location of modification within the sequence. **B**) Structure of proposed multistep chemical tags. **Strand 1** (magenta) and **Strand 2** (turquoise) are complementary DNA strands with uncomplementary toehold 3' regions (highlighted). R_1 = side group of the chosen amino acid incorporated into the oligomer.*

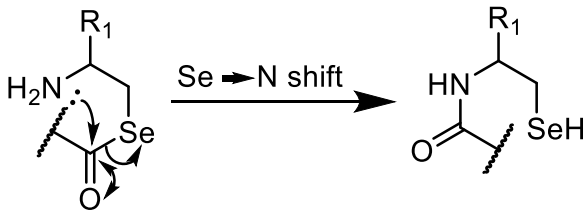
Initial designs executed oligomer extension at the external site of the growing chain. However, as previously reported, the elongation of the oligomer broadens the distance between the chemical reactants, hindering growth.[14] Therefore, an internal selenoester attachment site was implemented following across-the-helix architecture (Scheme 4-5).



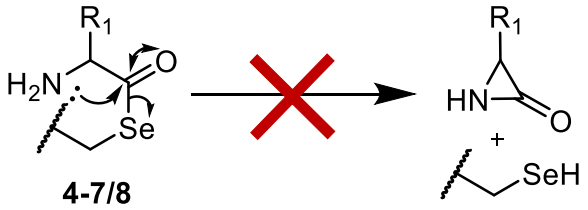
***Scheme 4-5**: The proposed first step of a multistep aminolysis mechanism, involving **Strand 1**-TAMRA-selenoester-DNA **4-6** and **Strand 2**-amino-acid-DNA **4-7**. Strand growth occurs at the **4-10**'s selenoester site. Toehold regions are labelled in magenta and turquoise.*

The aminolysis would commence with the transfer of TAMRA from **Strand 1** to **Strand 2** using the TAMRA-selenoester electrophile **4-6** and a newly synthesised nucleophile: an amino acid attached to secondary selenoester **4-8**. Here, the **Strand 2** amino acid is strategically positioned so that an intermolecular Se→N shift of a DSL reaction does not occur between the primary amine and selenoester (Scheme 4-6).[6] Thus, allowing for the adjacent selenoester to perform as the preceding electrophile.

Diselenide-Selenoester Ligation



Proposed Multistep Mechanism

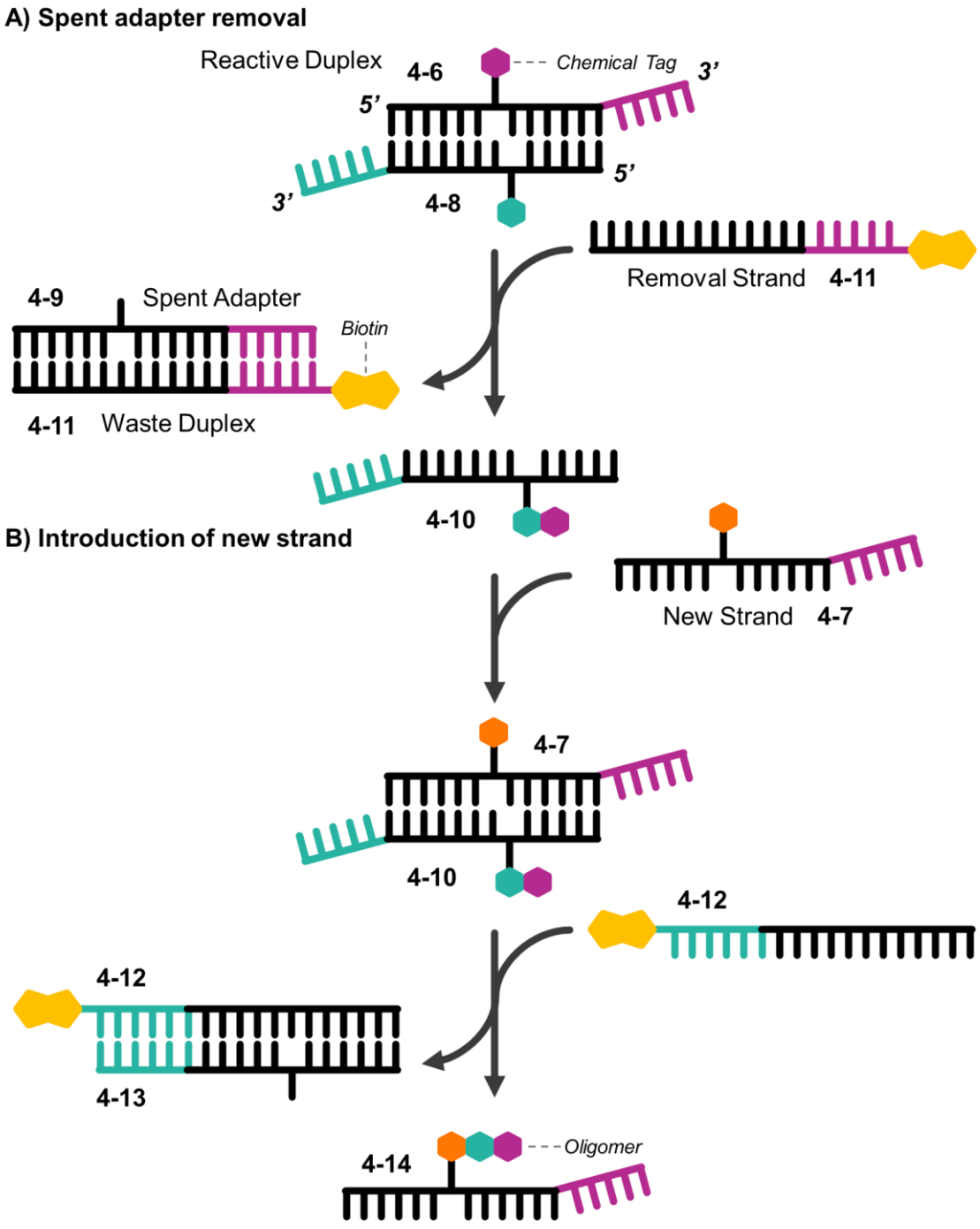


4-7/8

Scheme 4-6*: Se→N shift of a DSL reaction (above). Incomplete Se→N shift of the amino acid bi-functional nucleophile/electrophiles* ***4-7*** *and* ***4-8*** *due to functionality placement (below); intermolecular cyclisation.*

Toehold regions present on reactive strands act as a complementary extension to a "removal strand": a longer, more stable ssDNA which binds to the spent DNA strand **4-9** within the double helix (Scheme 4-7A). The removal **Strand 1 4-11** is modified with a 5′ biotin, which binds to streptavidin-coated magnetic beads (DynaBeads), and can be cleared from the reaction mixture, leaving the transferred TAMRA on **Strand 2 4-10**. Although previous multistep DTS strand displacement mechanisms installed toehold regions on opposing 3′-5′ ends of a helix (see Chapter 1), this mechanism installed the toehold regions on opposite 3′-3′ ends of the helix. This ensured accessibility of the bulky biotin-modified removal strands **4-11** and **4-12** to the toehold regions. A new amino acid-selenoester modified ssDNA **4-7** is introduced, of the same sequence as **Strand 1**, creating a new reactive aminolysis duplex with **4-10** (Scheme 4-7B). Following reaction incubation, a new removal strand **4-12** (complementary to **Strand 2**) is introduced, and the spent adapter strand **4-13** is removed from the synthetic pot. The process is repeated through multiple cycles of

Strand1/2 (**4-7**/**4-8**) aminolysis and displacement, forming a peptide product attached to the remaining ssDNA **4-14**.



Scheme 4-7*: Proposed selenoester-mediated multistep DTS mechanism: strand exchange.* ***A)*** *Introduction of removal strand* ***4-11*** *displaces spent adapter* ***4-9*** *following successful aminolysis of* ***4-6*** *to* ***4-8,*** *forming* ***4-10****. The waste duplex* ***4-9:4-11*** *is removed from the reaction using DynaBeads* ***B)*** *Introduction of a new strand* ***4-7*** *to initiate a second aminolysis reaction. Coloured hexagons: chemical tag. Yellow pentagons: biotin.*

4.2.3 Dynabead adaptation

In theory, following the successful strand displacement of the spent adapter strand **4-9** in the reactive double helix **4-6**/**4-8** to the removal strand **4-11** (Scheme 4-7A), the fully complementary double helix is cleansed from the reaction mixture using DynaBeads.[15] DynaBeads are superparamagnetic nanoparticles coated in streptavidin. The protein has a high affinity to biotin, wherein the strong interaction can be utilised in the immobilisation of biotinylated nucleic acids onto the magnetic bead surface.[15] Within the proposed multistep reaction, spent DNA adapters from the reaction mixture are removed through a binding buffer dilution and the addition of the DynaBeads. A magnet is externally introduced to the reaction vessel and the ssDNA occupying the site of peptide growth (**4-10**/**4-14**) is extracted, dried and resuspended in the DTS buffer.

To confirm that strand displacement occurs with the reactive duplex and that the spent biotin duplex binds to the DynaBeads, a control native polyacrylamide gel electrophoresis (PAGE) was performed using the starting primary amine DNA strands **4-15** and **4-16** (Figure 4-8). Results displayed successful reactive duplex formation and strand displacement with 2 equiv. of the first removal strand **4-11**. Sufficient removal of the excess ssDNA **4-11** was observed, due to its smaller occupied space on the magnetic beads. However, the incomplete removal of the waste double helix occurred. As the spent duplex would be unreactive, its presence in the reaction mixture would not impede the multistep synthesis. Therefore, the DynaBeads were used to ensure clearance of the excess of the removal strand, once sufficient strand displacement had taken place.

Lane	1	2	3	4	5	6
4-15 (−NH_2)	+	-	-	+	+	+
4-16 (−NH_2)	-	+	-	+	+	+
4-11 (Removal)	-	-	+	-	+	+
T1 Dynabeads	-	-	-	-	-	+

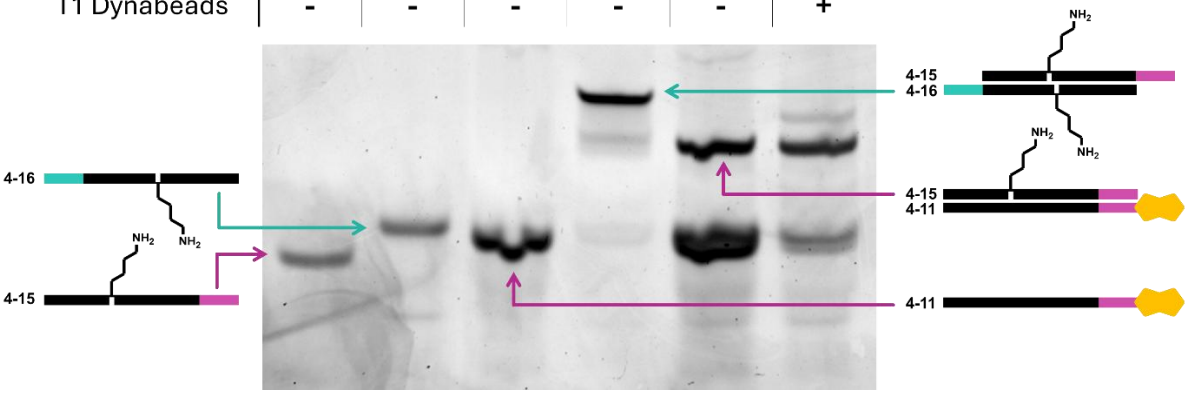


Figure 4-8: *Native PAGE (10%) of strand exchange DTS involving initial modified amine strands* ***4-15*** *and* ***4-16*** *and T1 DynaBead removal of biotin duplex* ***4-11****:****4-15****.*

As for optimisation, the incubation time of the reactive duplex **4-15**:**4-16** with the removal strand **4-11** was varied at 20 minutes and 60 minutes (Figure 4-9). No difference was observed in incubation times, indicating complete strand displacement within 20 minutes. As for spent duplex removal, four types of streptavidin DynaBeads were investigated: MyOne C1, MyOne T1, M-270 and M-280. The magnetic beads vary in size (1.0 – 2.8 μm) and binding capacity (200 – 500 pmol). However, minimal differences were observed in efficiency between the DynaBead types, with successful clearance of the excess removal strand **4-11** (Figure 4-9). However, DynaBeads MyOne T1 were selected for application into the multistep strategy due to their lower sedimentation rate during sample preparation, compared to the other DynaBeads investigated. By having a low sedimentation rate, separation of the reactive solution from the magnetic beads was easier, preventing product loss during preparative steps.

Lane	1	2	3	4	5	6	7	8	9	10
4-15 (−NH_2)	+	-	+	+	+	+	+	+	+	+
4-16 (−NH_2)	-	+	+	+	+	+	+	+	+	+
4-11 (Removal)	-	-	-	+	+	+	+	+	20 min	60 min
Dynabeads	-	-	-	-	C1	T1	M-270	M-280	-	-

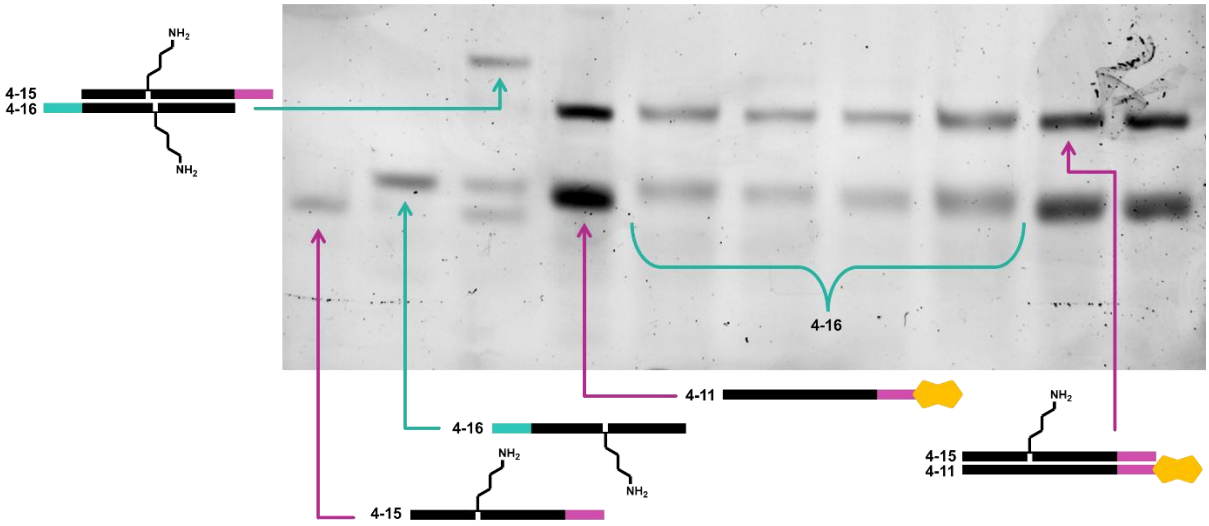


***Figure 4-9**: Optimisation: native PAGE (10%) of strand exchange DTS involving initially modified amine strands **4-15** and **4-16**. Varying DynaBead types (lanes 5-8) and strand exchange incubation time (lanes 9-10).*

4.2.3 DNA adapter synthesis

4.2.3.1 Strand 1: TAMRA selenoester DNA 4-6

As established in Chapter 2, the synthetic attachment of TAMRA to DNA *via* a selenoester bond was achievable. The TAMRA modification was installed in the multistep strategy as the synthetic route had been established, as well as the ease of analysis *via* RP-HPLC's fluorescence channel. The new, longer, modified DNA-NH_2 **4-15** was subject to an *in situ* activation and coupling with $(Boc\text{-}Sec\text{-}OH)_2$ **3-8**, followed by a TCEP reduction and esterification with TAMRA-NHS **2-6** (Figure 4-10). However, in comparison to the original TAMRA selenoester DNA **3-17**, the reaction yield had decreased from 18% to 13% (LC-MS peak area, see Supplementary Data 4-7) despite identical conditions.

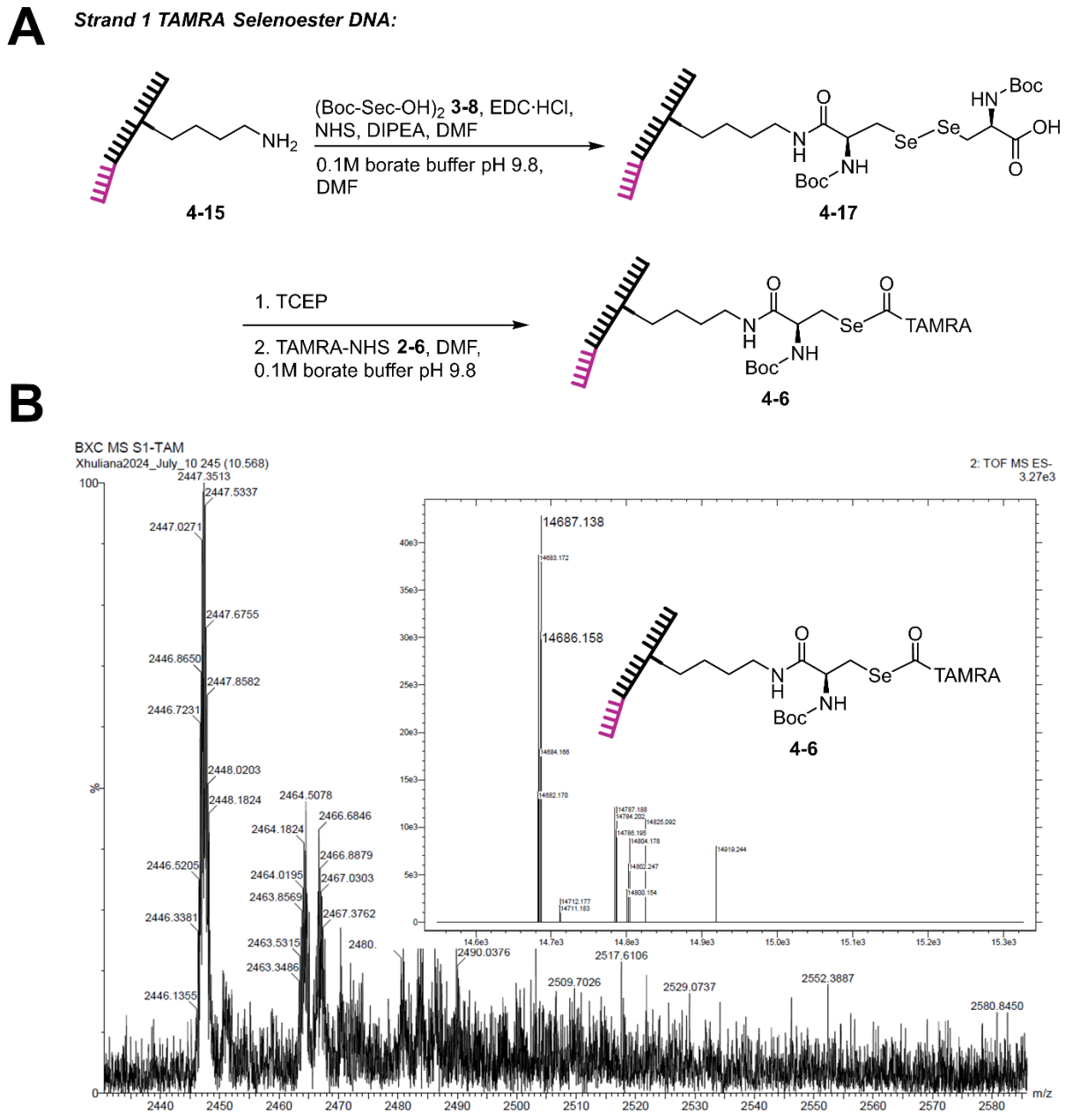


***Figure 4-10: A**) Synthesis of **Strand 1** TAMRA selenoester DNA **4-6**. **B**) TOF MS ES− and deconvolution (ProMass HR) of **Strand 1** TAMRA selenoester DNA **4-6**. Expected Mass: 14686.52 m/z, found: 14687.14 m/z. See Experimental 4.4.7 for synthesis.*

4.2.3.1 Strand 2: amino acid selenoester DNA 4-8

Following similar synthetic steps, an analogue of the amino acid selenoester DNA **4-8** was constructed using the activated Fmoc-protected valine (Fmoc-Val-OSu **4-19**). Boc-Sec DNA **4-18** using **Strand 2** DNA-NH_2 **4-16** was synthesised, and the

crude DNA solution was mixed with the activated ester Fmoc-Val-OSu **4-19** to yield the amine-protected Fmoc-Val selenoester DNA **4-20** (Figure 4-11).

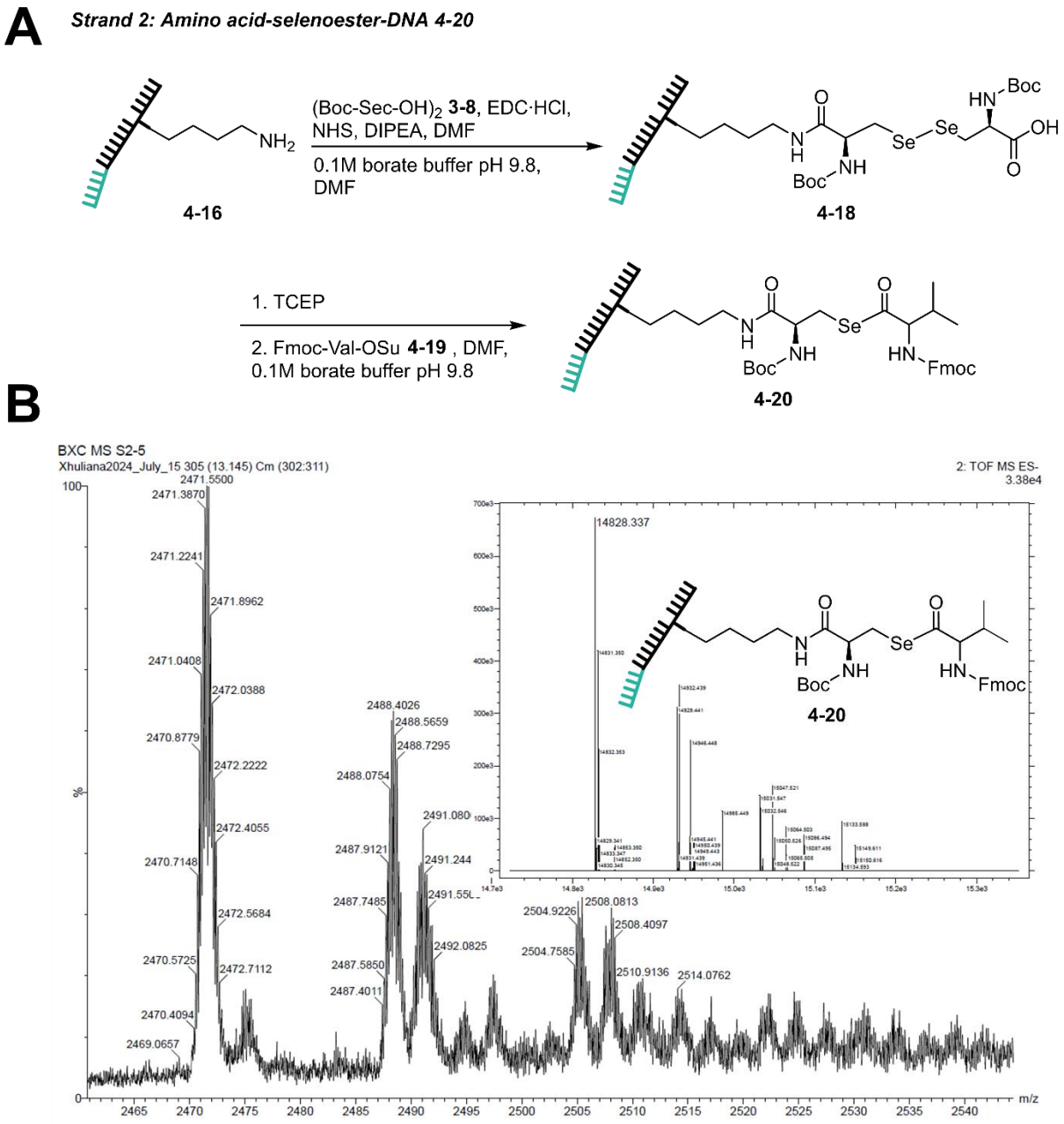


***Figure 4-11**: **A**) Synthesis of **Strand 2** Fmoc-Val selenoester DNA **4-20**. **B**) TOF MS ES− and deconvolution (ProMass HR) of **Strand 2** amino acid-selenoester DNA **4-20**. Expected Mass: 14828.82 m/z, found: 14828.34 m/z. See Experimental 4.4.8 for synthesis.*

The Fmoc protecting group was selected as the deprotecting conditions are DNA compatible (basic conditions) and allow Boc protection to remain on the primary amine of Sec. To expose the nucleophilic primary amine, Fmoc-Val selenoester DNA **4-20** was subject to Fmoc deprotection, following the same protocol as the previously investigated Fmoc-Sec-DNA **4-2** deprotection, at 5% and also 1% piperidine in DMF. However, the nucleophilic piperidine caused degradation of the selenoester, reverting to the starting material (Figure 4-12). As the small molecule is localised on a solid support, the reaction was conducted in the solution phase at the same concentration. Opposed to the solid support, Fmoc-Val-DNA **4-20** remained in the solution. As Fmoc deprotection was not observed without the degradation of the selenoester, a new amine protection chemistry was investigated.

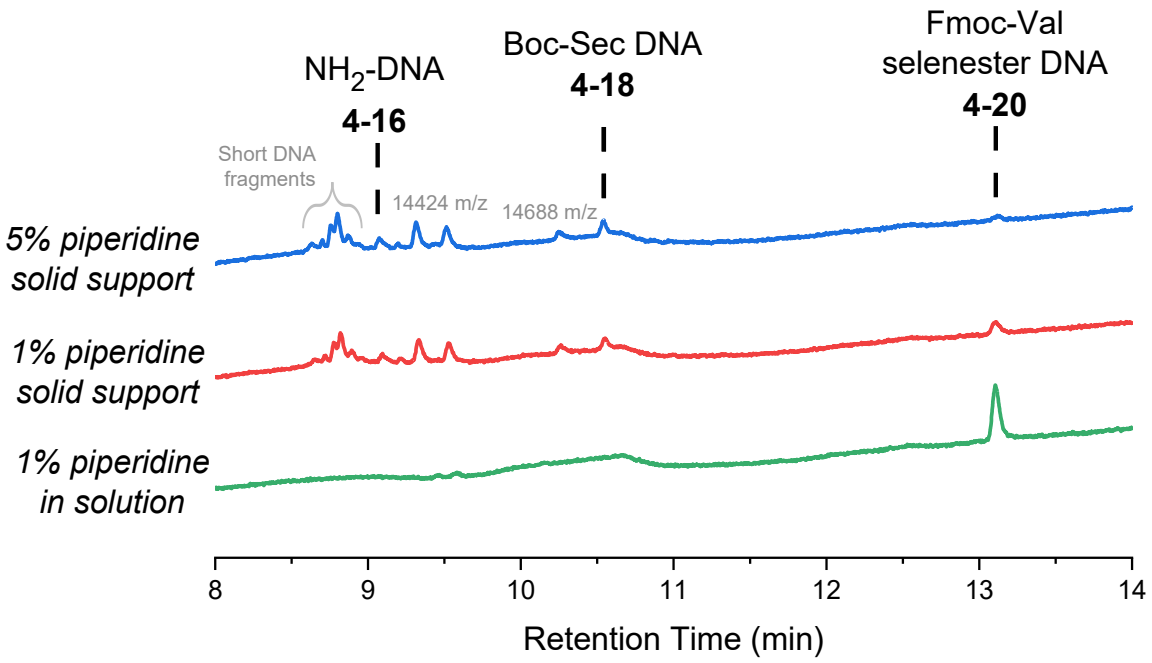
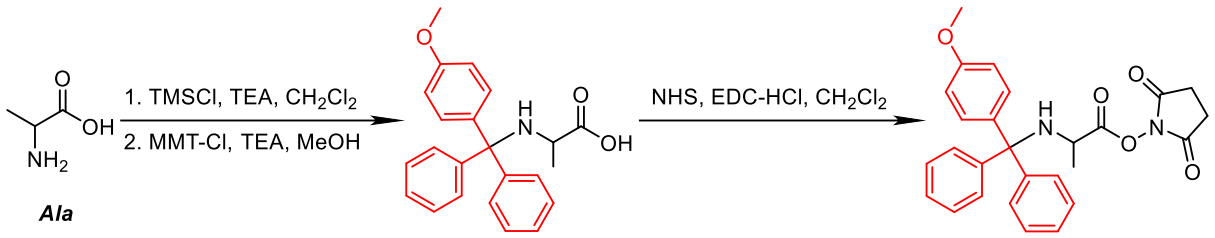


***Figure 4-12**: LC-MS plots of Fmoc-Val selenoester DNA **4-20** piperidine reaction conditions. Degradation products observed: **Strand 2** NH_2-DNA **4-16** and Boc-Sec DNA **4-18**. Solid support: 5% piperidine (blue) and 1% piperidine (red), Solution phase: 1% piperidine (green). Additional peaks observed: 14424 and 14688 m/z did not correspond to the desired deprotected DNA modification at 14606 m/z.*

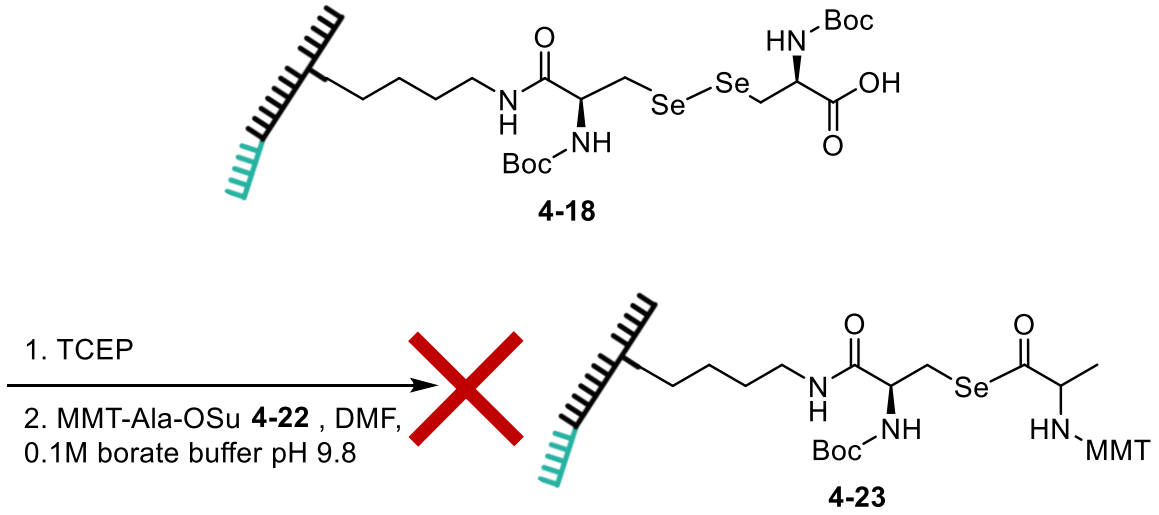
Monomethoxytrityl (MMT) was selected as the Fmoc protection replacement because of its mildly acidic deprotection conditions, which were envisaged to prevent the degradation of the selenoester linker. As highly acidic conditions lead to depurination of DNA, the diluted acid, 20% acetic acid, cleaves the MMT group without interfering with the DNA framework.[16] Because MMT is an uncommon primary amine protecting group, the protected analogue of Ala was synthesised, followed by carbodiimide esterification to accommodate DNA connection (Scheme 4-8).



***Scheme 4-8**: Synthesis of MMT-Ala-OSu **4-22**. MMT is highlighted in red. See Experimental 4.4.10 for synthesis.*

Following the documented procedure of Brandsen *et al.*,[17] MMT-Ala-OH **4-21** was synthesised, activated and purified to yield MMT-Ala-OSu **4-22**. As a consequence of the MMT group's sensitive nature, in combination with the high MS temperatures, MS characterisation indicated deprotection by a high intensity of the MMT small molecule. However, ^{1}H and ^{13}C NMR spectroscopy concluded the successful isolation of MMT-Ala-OSu **4-22** (Sup. Figures 4.8 – 4.9).[17] Adapting the synthetic route of Fmoc-Val selenoester DNA **4-20**, Fmoc-Val-OSu **4-19** was replaced with newly synthesised MMT-Ala-OSu **4-22** in order to yield MMT-Ala selenoester DNA **4-23** (Scheme 4-9). Despite previous success in diselenide reduction using TCEP in borate buffer (pH 9.8), the formation of MMT-Ala selenoester DNA **4-23** did not occur.

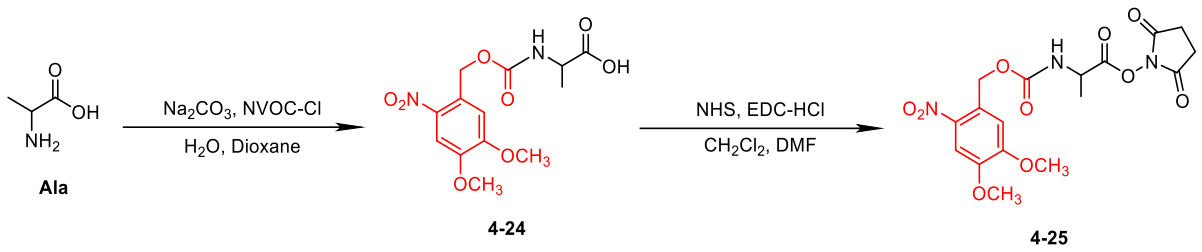
Strand 2: MMT-Ala selenoester-DNA 4-31



Scheme 4-9: *Attempted synthesis of* ***Strand 2*** *MMT-Ala selenoester DNA* ***4-23****. See Experimental 4.4.11 for synthesis.*

It was theorised that the acidic TCEP led to the degradation of the MMT group in the aqueous conditions, preventing the production of MMT-Ala selenoester DNA **4-23**. Therefore, TCEP was substituted with dithiothreitol (DTT), an alternative disulfide-reducing agent compatible with the basic conditions.[18] Upon MS analysis, an alternative DNA modification was detected, at 29 m/z higher than the desired m/z of MMT-Ala selenoester DNA **4-23** (Supplementary Figure 4-15). Nonetheless, the attachment of MMT-Ala was deemed unsuccessful, and a final protection strategy was proposed.

As both acidic and basic-labile groups had presented difficulties during their synthesis, a light-cleavable group was drafted for *in situ* exposure of the primary amine during the DTS assay. Inspired by Liu *et al*.'s NHS ester multistep DTS mechanism, the light-cleavable protecting group NVOC was synthesised on Ala to yield NVOC-Ala-OSu **4-25** (Scheme 4-10).[5, 6]



Scheme 4-10: *Synthesis of NVOC-Ala-OSu* ***4-25****. NVOC is highlighted in red. See Experimental 4.4.12 for synthesis.*

To ensure NVOC-Ala attachment to form the **Strand 2** amino acid selenoester DNA **4-8**, multiple strategies were derived. Strategies implemented either TCEP or DTT reducing agents with *in situ* NHS coupling or pre-activated NVOC-Ala-OSu **4-25** (Figure 4-13). However, none of the investigated synthetic routes were deemed suitable for DNA attachment *via* the desired selenoester linker. Despite light-sensitive conditions being installed throughout the synthesis, peptide formation was observed in multiple assays, suggesting NVOC cleavage.

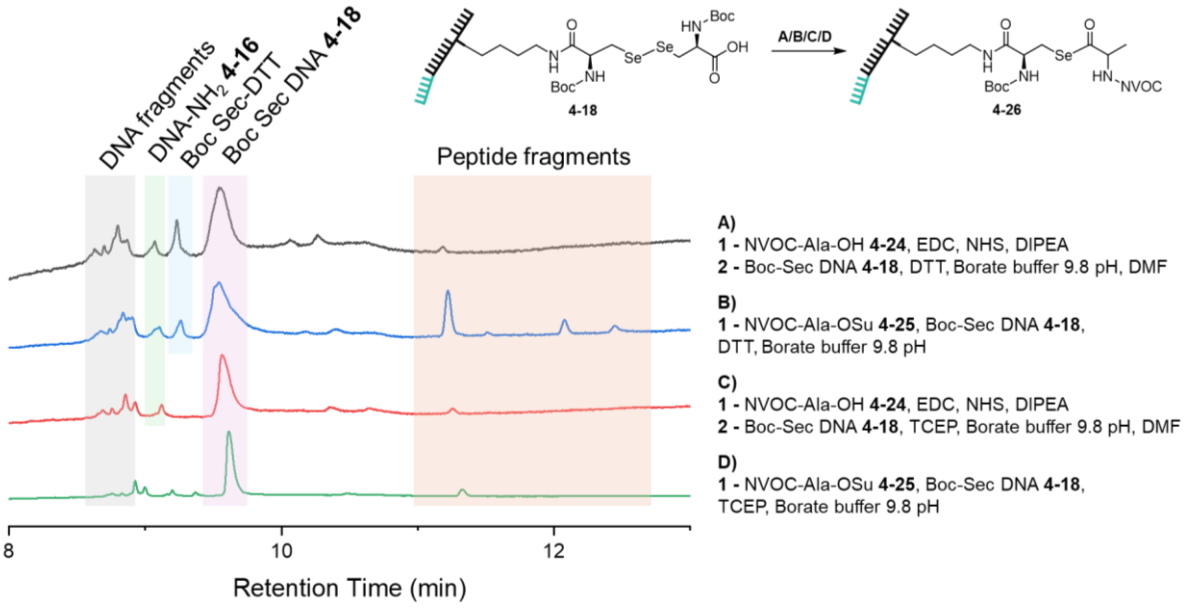


Figure 4-13*: LC-MS plots of* ***Strand 2*** *NVOC-Alal selenoester DNA* ***4-26*** *synthetic routes. Additional peaks observed: DNA fragments (< 12000 m/z, grey), starting material DNA-NH_2 (14256 m/z, green), Boc-Sec DNA bound Se-S to DTT (14659 m/z, blue), Boc-Sec DNA* ***4-18*** *(14774 m/z, pink) and peptide fragments (orange). 14424 m/z did not correspond to the desired deprotected DNA modification.*

One hypothesis proposed for this observation, as well as the lack of successful MMT attachment, was due to the increase in DNA length. With an additional 10 nucleobases to the original DNA sequence, the flexibility of the DNA increases and may restrict access to the modification site. Supporting this hypothesis, there was a 5% yield reduction when synthesising the longer **Strand 1** TAMRA selenoester DNA **4-6** in comparison to the short, original TAMRA selenoester DNA **3-17**, despite identical synthetic steps.

To test this, the original shorter strands: DNA-NH_2 **2-8** and DNA-NH_2 +1 **2-2**, were modified with Boc-Sec moieties, followed by incubation with NVOC Ala OSu **4-25** and TCEP. Confirming the hypothesis, the shorter NVOC-Ala selenoester modifications **4-27** and **4-28** were successfully synthesised (Figure 4-14).

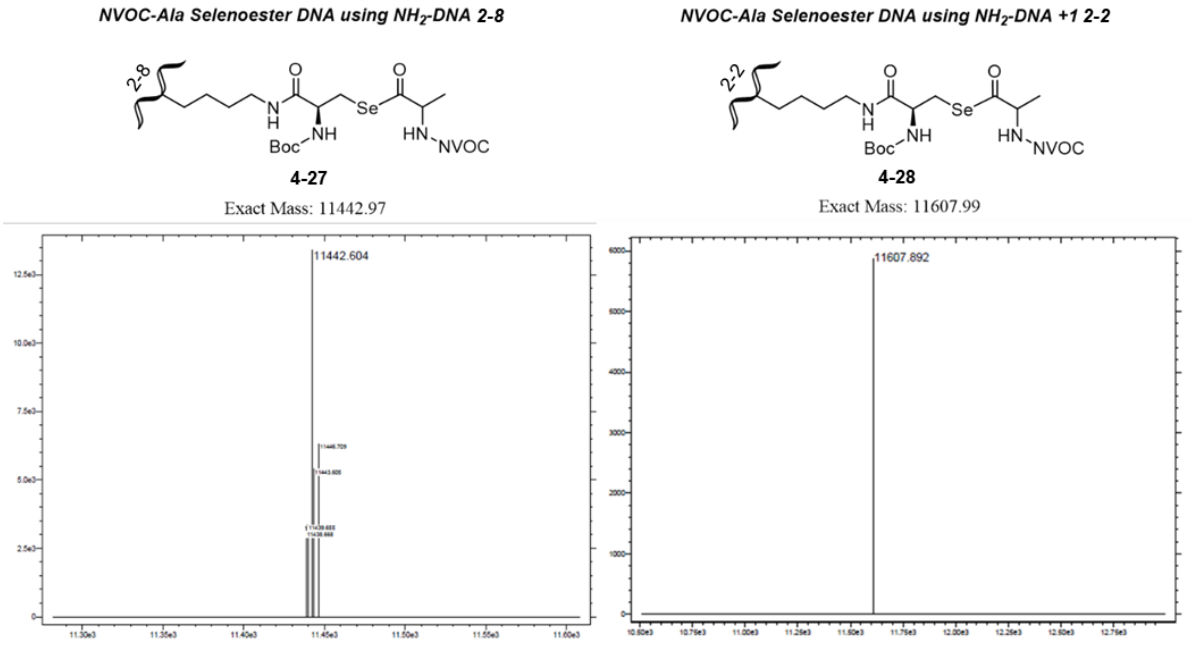


***Figure 4-14**: Deconvoluted (ProMass HR) TOF MS ES− of **left**: NVOC-Ala selenoester DNA **4-27** (using DNA-NH_2 **2-8**, Expected Mass: 11442.97, found: 11442.60 m/z) and **right**: NVOC-Ala selenoester DNA **4-28** (using DNA-NH_2 +1 **2-2**, Expected Mass: 11607.99, found: 11607.89 m/z).*

In order to predict the success of NVOC Ala as a nucleophile prior to synthetic optimisation of the longer modification, the shorter NVOC-Ala selenoester +1 DNA **4-28** was issued in a single-step DTS assay with TAMRA selenoester DNA **3-17**. During incubation at 19 ºC, the sample was exposed to a UV light at λ = 365 nm, to initiate NVOC deprotection to Ala selenoester DNA **4-29** (Figure 4-15).[5]

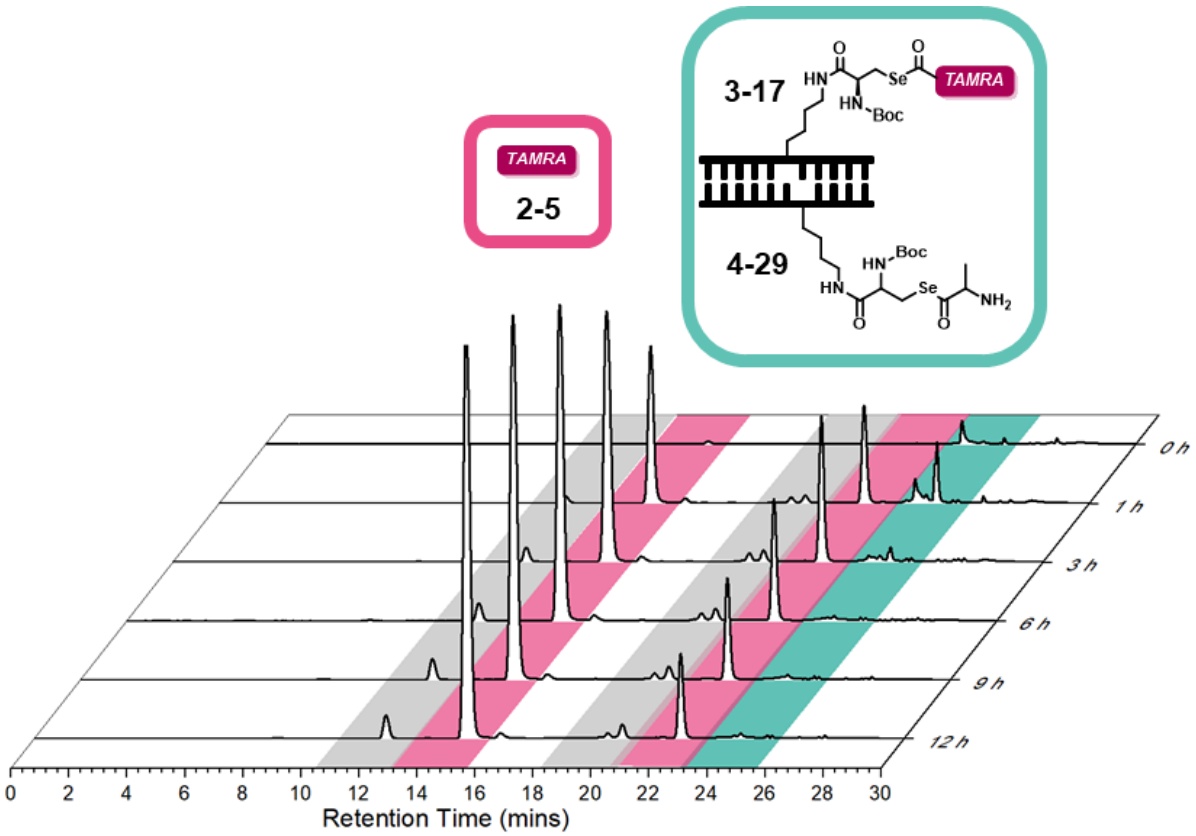


Figure 4-15: *Waterfall plot of RP-HPLC spectrum of Ala selenoester DNA* ***4-29*** *with TAMRA selenoester DNA* ***3-17*** *in DTS buffer, pH 11.0, over the course of 12 hours. Fluorescence channel: (λ_{ex} 550 nm, λ_{em} 580 nm). 2 µL injection volume. Pink = free TAMRA* ***2-5*** *(hydrolysis of* ***3-17****). Green = starting material (TAMRA selenoester DNA* ***3-17*** *with Ala selenoester DNA* ***3-36****). Grey = DTS buffer salts.*

Despite efforts to attach the selenoester functionality, aminolysis was not observed between Ala selenoester DNA **4-29** and the original TAMRA selenoester DNA **3-17**. The HPLC assay displayed consistent hydrolysis of the TAMRA selenoester DNA **3-17**, indicated by the increase in free TAMRA **2-5**, with no additional peak formation, over the course of 12 hours.

Reviewing data presented in Chapter 3.2.3, the use of Ala as a chemical tag was not a suitable candidate in terms of stability; the functional group adjacent to the selenoester plays a significant role in the rate of hydrolysis as well as aminolysis. As Ala is one of the smallest amino acids, the space surrounding the selenoester is more accessible by water and, thus, hydrolysis can occur when presented as an ester (See Chapter 3). Supporting these claims, the successful synthesis of **Strand 2** Fmoc-Val selenoester DNA **4-20** to the longer DNA strand may have been due to the bulky Val group (−$CH(CH_3)_2$) aiding in the stabilisation of the selenoester. Additionally, the literature has shown increased stability in Val-adjacent acyl linkages compared to other amino acids, including Ala.[19, 20] Such observations emphasise the precise chemistries required to accommodate the reactive nature of selenoesters.

Based upon the evidence presented in this body of work, it is hypothesised that selenoester-mediated DTS can be potentially optimised through the selection of a bulky transfer group (*e.g*. TAMRA) to minimise hydrolysis and the selection of a sterical small nucleophile (*e.g*. aliphatic primary amine) to accommodate the restricted access to the ester. To complement the ester steric hypothesis, current multistep aminolysis DTS methods have predominantly accommodated the transfer of larger functional groups, *e.g*. cyclic, phenyl compounds.[1-3, 5-7]

4.3 Conclusions

Building on the success of selenoesters in a single-step DTS assay in Chapter 3, this body of work explored the reactive chemistries further to aid in the implementation of selenoesters in a multistep DTS assay. By performing a set of optimisation assays, pH 11.0 was confirmed as the superior reactive condition for selenoester-mediated DTS within the across-the-helix architecture, compared to pH 10.2 and pH 7.0, concluding pH 11.0 provided a balance between $-NH_2$ (nucleophile) exposure and hydrolysis limitation. Secondly, the initial incubation period of the DTS assay between reactive ssDNA did not impact hybridisation, and fast reaction kinetics were observed at pH 11.0 for all incubation times investigated. Such findings would aid in a multistep DTS system by preventing the truncation of desired products.

Despite success in aminolysis DTS, no DSL DTS was observed between the TAMRA selenoester and a Sec-derived selenol nucleophile, concluding DSL would not be a suitable candidate in the construction of a multistep DTS system. It was hypothesised that the selenoester intermediate of DSL was degrading in the basic conditions, preventing reaction completion. Therefore, aminolysis mechanisms utilising primary amines and esters were a more promising candidate in the transfer of reactive tags, as demonstrated in the current multistep DTS literature.

Building upon Chapter 3′s synthetic strategies, a TAMRA-selenoester was successfully attached to a longer DNA strand, acting as the initial ester in the multistep DTS. As for the initiating nucleophile, multiple attempts were made to synthesise an amino acid conjugated to a secondary selenoester on a complementary DNA strand. Despite successful attachment to DNA of Fmoc-Val *via* a selenoester linker, Fmoc

deprotection conditions caused degradation of the modification. Two alternative protection strategies, MMT and NVOC on Ala, were attempted; however, neither of the small molecules successfully attached to the desired DNA strand due to the increase in complexity of the longer DNA strand in combination with the hydrolytically accessible Ala. On the other hand, NVOC-Ala was successfully attached to a shorter DNA strand; however, aminolysis of TAMRA was not observed during a DTS assay. Such results exemplified the importance of sterics and accessibility in ester chemistries.

Despite the unaccomplished multistep reaction and unsuccessful single-step aminolysis of the NVOC Ala selenoester, the electrophilic moiety still possesses great potential to improve DTS yields. For example, the ribosome has elapsed billions of years of evolution in order to perform oxoester-mediated peptide synthesis.[19] However, outside the specific conditions required by the ribosome, oxoesters are a limiting electrophile for aminolysis.[21] The conditions reviewed in this chapter cover a fraction of the investigable possibilities. Alternative DTS architectures, fluorophores, attachment groups (amino acids) and attachment parameters (buffer/pH) are yet to be investigated. Only one DNA architecture was investigated in this body of work, whereas there are others (end-of-helix, across-the-nick, G-quadruplex) that are still to be exploited. The across-the-helix architecture has only recently been established within our research group and may not be the most viable route for successful multistep synthesis.[22]

4.4 Experimental

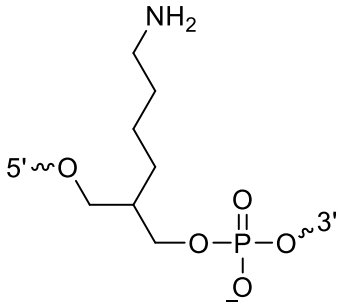
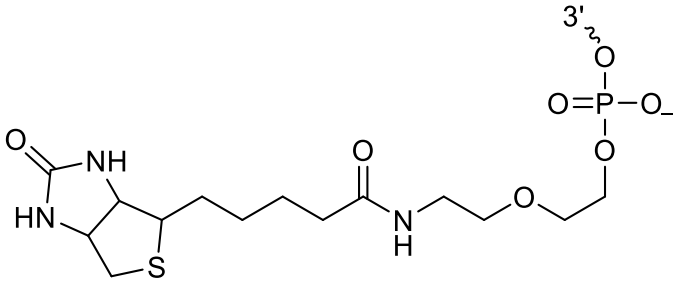
For details of the RP-HPLC, LC-MS, NMR and IR techniques, see Appendix 1: Experimental Methods.

4.4.1 Materials

See *Chapter 3 Experimental 3.4.1 Materials* for specialised small molecule suppliers and buffer formulations. MMT-Cl and NVOC-Cl were purchased from Sigma-Aldrich. Solid support DNA synthesis was performed using a Glen Research Empty Synthesis Columns-TWIST (40 nm, 0.2 or 1 μm) and DEAE Sepharose Fast Flow resin, purchased from Glen Research and Avantor, respectively. DynaBeads were purchased as a Dynabeads™ Streptavidin Trial Kit (65801D)[15] from ThermoFischer Scientific, containing the bead types: Dynabeads M-280 Streptavidin, Dynabeads M-270 Streptavidin, Dynabeads MyOne Streptavidin C1, and Dynabeads MyOne Streptavidin T1 (all 10 mg mL^{-1}). SYBR™ Gold Nucleic Acid Gel Stain (10,000X) was purchased from ThermoFisher Scientific. All PAGE-related chemicals were purchased from Sigma-Aldrich, other than Gel Loading Dye, Purple (6X), which was purchased from New England Biolabs.

DNA sequences were purchased from Integrated DNA Technologies (IDT), with either standard desalting or HPLC purification. All samples were suspended in water (18.2 MΩ) at a concentration of 1 mM, 0.2 mM or 0.1 mM.

***Table 4-2**: DNA modifications and sequences*

IDT Key	Description	Modification
/iUniAmM/	internal Uni-Link™ Amino Modifier	 */iUniAmM/*
/5Biosg/	5′ Biotin	 */5Biosg/*

Identification	Sequence 5′ → 3′
Original Donor (ssDNA-NH_2 for **3-17**)	GCCGAGCCAGCAGTCAGCGC/**iUniAmM**/GTCCTAATCTACCTG
ssDNA-NH_2 +1 **2-2**	CAGGTAGATTAGGA/**iUniAmM**/AGCGCTGACTGCTGGCTCGGC
Strand 1 removal DNA-Biotin **4-11**	/**5Biosg**/AACCAACCAACAGGTAGATTAGGACAGCGCTGACTGCTGGCTCGGC
Strand 1 removal DNA-Biotin **4-12**	/**5Biosg**/ACCAATCCCCGCCGAGCCAGCAGTCAGCGCTGTCCTAATCTACCTG
ssDNA-NH_2 **4-15**	GCCGAGCCAGCAGTCAGCGC/**iUniAmM**/GTCCTAATCTACCTGTTGGTTGGTT
ssDNA-NH_2 **4-16**	CAGGTAGATTAGGA/**iUniAmM**/AGCGCTGACTGCTGGCTCGGCGGGGATTGGT

4.4.2 TAMRA selenoester DNA 3-17 synthesis

See *Chapter 3 Experimental 3.4.11* for TAMRA selenoester DNA **3-17** synthesis.

4.4.3 Single-step DNA-templated synthesis assay optimisation

pH optimisation:

Modified assay of Chapter 3 Experimental 3.4.12: 15 µL of DTS buffer (pH 11) was replaced with 15 µL of pH 10.2 buffer (0.25 M NaCl, 0.25 M TAPS, 0.25 M CAPS, 0.025 M Na_3PO_4, pH 10.2 assay) or pH 7.0 salt solution (0.5 M $MgCl_2$, pH 7.0 assay).

Incubation optimisation:

Modified assay of *Chapter 3 Experimental 3.4.12*: Mastercycler incubation time was varied at 1, 5, 10, 20, 30 and 60 minutes before being mixed in a DTS buffer (pH 11, 15 µL) and injected directly into an RP-HPLC (2 µL).

4.4.4 Synthesis of Sec-DNA 4-4

4.4.4.1 Synthesis of $(Fmoc\text{-}Sec\text{-}OH)_2$ 4-1

Following the procedure presented by Flemer *et al*.[23], in a brown round-bottom flask, Na_2CO_3 (0.20 g, 1.91 mmol) was dissolved in nanopure H_2O (15 mL). Acetone (17.5 mL) was added to the reaction mixture, followed by *l*-selenocystine **3-10** (0.32 g, 0.96 mmol). The reaction temperature was increased to 50 ºC and Fmoc-OSu (0.97 g, 2.87 mmol) was added and dissolved. The reaction was brought back to room temperature (19 ºC) and left for 18 h. Once completed, HCl (1M, 20 mL) and EtOAc (25 mL) were added, and an acidified aqueous workup was performed. The acidic layer was washed with EtOAc (3 × 50 mL). The organic layer was extracted, dried with

$MgSO_4$ and reduced under vacuum. The crude solid was recrystallised with chloroform, producing a pure pale-yellow powder of $(Fmoc\text{-}Sec\text{-}OH)_2$ (0.35 g, 47%). **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.94 (s, 2H), 7.90 (d, *J* = 8.2 Hz, 4H), 7.79 (d, *J* = 8.2 Hz, 2H), 7.70 (d, *J* = 7.7 Hz, 4H), 7.40 (t, *J* = 7.3 Hz, 4H), 7.31 (t, *J* = 7.5 Hz, 4H), 4.26 (m, 8H), 3.34 (m, 2H), 3.17 (m, 2H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: δ 172.3 (2C), 155.8 (2C), 143.7 (4C), 140.7 (4C), 127.6 (4C), 127.1 (4C), 125.2 (4C), 120.1 (4C), 65.7 (2C), 54.6 (2C), 46.6 (2C), 39.5 (2C). **^{77}Se NMR** (76 MHz, $(CD_3)_2SO$) $\delta_{Se/ppm}$: 302 (2Se). **TOF MS**: (ES-) *m/z* $[M\text{-}H]^-$ calculated: 781.05, found 781.06. See Supplementary Figure 4-1 to 4-4 for characterisation.

4.4.4.2 DNA synthesis of Fmoc-Sec-DNA 4-2

Stock solutions: $(Fmoc\text{-}Sec\text{-}OH)_2$ **4-1** (300 mM, DMF), NHS (300 mM, DMF), EDC-HCl (300 mM, DMF), piperidine **4-3** (5%, DMF), DEAE Binding Buffer (10 mM acetic acid, 0.005% Triton X-100) and DEAE Elution Buffer (1.5 M NaCl, 50 mM tris-HCl, 0.005% Triton X-100).

In a primary microcentrifuge tube, 67 µL of $(Fmoc\text{-}Sec\text{-}OH)_2$ **4-1** stock (300 mM), 67 µL of NHS (300 mM), 67 µL of EDC-HCl stock (300 mM) and 3.5 µL of DIPEA were added and mixed for 10 minutes, 19 ºC at 1000 rpm. In a secondary microcentrifuge tube, 20 µL of NH_2-labelled DNA (**2-8**, 1 mM, H_2O), 180 µL PBS buffer (100 mM, pH 7.4) and 200 µL DMF were added along with all contents of the primary microcentrifuge tube. The combined solution was mixed for 18 h at 19 ºC and 1000 rpm. The solution was separated into three microcentrifuge tubes (~202 µL), dissolved in 300 µL ethanol with 20 µL sodium acetate (NaOAc) (0.3 M, pH 5.2) and placed in the freezer overnight to remove unreacted small molecules. The modified

DNA pellets were centrifuged (15000 rpm), emptied and washed twice with ice-cold ethanol (70%, H_2O). The modified DNA pellet was dried, dissolved in 500 µL H_2O and spin-filtered (Amicon-Ultra 3 kDa 0.5 mL centrifugal filter) to remove any remaining small molecules. The sample was then purified using fluorescence RP-HPLC (see *Appendix: Experimental Methods A.2.2*) and washed with H_2O in Amicon-Ultra 3 kDa 0.5 mL centrifugal filters, leaving a 50 µL solution. Fmoc-Sec-DNA **4-2** was characterised by LC-MS: **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 11818.94, found: 11817.54 m/z. See Supplementary Figure 4-5 for characterisation.

4.4.4.3 Synthesis of Sec-DNA 4-4

Following procedures documented in Halpin *et al.*[24], an empty Glen Research TWIST synthesis column housing was loaded with 250 µL DEAE Sepharose resin. The column was flushed with H_2O (20 mL) followed by 12 mL DEAE binding buffer. The resuspended Fmoc-Sec-DNA **4-2** solution was diluted in DEAE binding buffer (450 µL, 500 µL volume) and slowly washed through the column for 3 minutes. Next, DEAE binding buffer (3 mL), H_2O (500 µL) and DMF (3 mL) were individually flushed through the column. The Fmoc-deprotection proceeded by flushing with 5 wt% piperidine **4-3** solution (DMF, 2 mL), washed slowly and incubated with a fresh 5 wt% piperidine **4-3** solution (1 mL) for 3 minutes, and finally flushed with more 5% piperidine **4-3** solution (1 mL). To stop the reaction, DMF (20 mL, minimum) was flushed through the column, followed by DEAE binding buffer (5 mL). DEAE elution buffer (4 × 500 µL) was flushed through, collecting each sample from the column. The combined fractions were washed and desalted using Amicon-Ultra 3 kDa 0.5 mL centrifugal filters. The sample was then purified using fluorescence RP-HPLC (see *Appendix: Experimental Methods A.2.2*) and washed with H_2O in Amicon-Ultra 3 kDa 0.5 mL centrifugal filters, removing

HPLC buffer salts. Resuspended in nanopure H_2O, the modified Sec-DNA **4-4** was identified using LC-MS and the concentration was determined using a NanoDrop One/One UV-Vis Spectrophotometer absorbance at λ = 260 nm. **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 11374.81, found: 11374.56.

4.4.5 Selenol DNA-templated synthesis assay

Stock solutions: TAMRA selenoester DNA **3-17** (9 μM, H_2O), Sec-DNA **4-4** (9 μM, H_2O), TCEP (10 mM, H_2O) and DTS buffer (pH 11, 0.25 M NaCl, 0.25 M TAPS, 0.25 M CAPS, 0.25 M Na_3PO_4).

In a PCR tube (200 μL), 0.6 μL $MgCl_2$ (0.5 M), 3.2 μL H_2O (18.2 MΩ), 3.33 μL TAMRA selenoester DNA **3-17** (9 μM), 6.67 μL Sec-DNA **4-4** (9 μM) and 1.2 μL TCEP (10 mM) were added. In a Mastercycler, the tube was incubated for 5 minutes at 15 ºC. The solution was then diluted in a spring vial insert (100 μL) containing 15 μL of DTS buffer (pH 11), mixing three times with the transferred pipette tip. The solution was injected (2 μL) directly for RP-HPLC analysis, analysing the fluorescence channel (λ_{ex} 550 nm, λ_{em} 580 nm) and UV-Vis channel (λ = 262 nm). Samples were injected at 0 h, 1 h, 3 h, 6 h, 9 h, 12 h and 15 h intervals and completed in triplicate. The RP-HPLC sample tray was incubated at 15 ºC for the duration of the assay.

4.4.6 DynaBead assay

Modified assay of *Chapter 3 Experimental 3.4.12*: replaced 3.33 μL TAMRA selenoester DNA **3-17** (9 μM) with **Strand 1** ssDNA-NH_2 **4-15** (9 μM) and 6.67 μL DNA-NH_2 +1 **2-2** (9 μM) with **Strand 2** ssDNA-NH_2 **4-16**. Next, 6.67 μL **Strand 1** removal

DNA-Biotin **4-11** (9 μM) was added and incubated for 20 minutes, without spin filtering and HPLC analysis. DynaBead addition and PAGE analysis followed.

Assay followed by the Dynabeads™ Streptavidin Trial Kit (Publication No. MAN0015762) protocol.[15] Stock solutions: Binding buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M NaCl) and Washing buffer (1:1 Binding buffer:H_2O). In a microcentrifuge tube, 10 μL of DynaBead solution (varying models, 10 mg mL^{-1}) was added to 1 mL of washing buffer and mixed with a pipette tip. The vial was added to a MagJET Separation Rack for 1 minute and the solution was removed, avoiding the magnet. Repeating on two more occasions, 1 mL of washing buffer was added, mixed, magnetised and removed. Following the washing cycle, 20 μL of binding buffer was added to the DynaBeads along with 8 μL of reaction mixture and mixed for 10 minutes. The vial was added to the magnet again for 1 minute, the solution was removed and washed twice with washing buffer. The collected washes were concentrated and desalted using Amicon-Ultra 3 kDa 0.5 mL centrifugal filters

10% native gel polyacrylamide gel electrophoresis (PAGE) data were performed in a 1× TAE (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA, pH 8.3) running buffer. 10% native solutions were formulated with acrylamide (1.6 mL), 10× TAE (0.5 mL) and H_2O (2.9 mL) and polymerised with the addition of tetramethylethylenediamine (5 μL) and 10 wt% ammonium persulfate solution (50 μL). Quickly, the solution was poured into a 1 mm BioRad gel plate casting system, an 8/10-well comb was inserted and left for 30 minutes to polymerise. DynaBead samples were diluted to 150 nM in water and further diluted at 5:1 with Gel Loading Dye, Purple (6X). Samples (1 μL) were loaded onto the gel and ran using a BioRad vertical nucleic acid electrophoresis cell connected to PowerPack basic power supply. Following electrophoresis, gels were stained with

SybrGold gel stain solution (1 µL) in water (10 mL), washed and imaged using a BioRad ChemiDoc MP transilluminator.

4.4.7 Synthesis of Strand 1 TAMRA selenoester DNA 4-6

The synthetic procedure follows that of *Chapter 3 Experimental 3.4.11* in which the NH_2-labelled DNA **2-8** was replaced with **Strand 1** NH_2-labelled DNA **4-15** (1 mM, H_2O). **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 14686.52 m/z, found: 14687.14 m/z. See Figure 4-10 for the characterisation spectrum.

4.4.8 Synthesis of Strand 2 Fmoc Val selenoester DNA 4-20

The synthetic procedure follows that of *Chapter 3 Experimental 3.4.11* in which the NH_2-labelled DNA **2-8** was replaced with **Strand 2** NH_2-labelled DNA **4-16** (1 mM, H_2O) and TAMRA-NHS **2-6** was replaced with Fmoc-Val-OSu **4-19** (0.1 M, DMF). **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 14828.82 m/z, found: 14828.34 m/z. See Figure 4-11 for characterisation spectrum.

4.4.9 Strand 2 Fmoc Val selenoester DNA 4-20 deprotection assay

Solid support: Modified assay of *Experimental 4.4.4.3*: Fmoc-Sec-DNA **4-2** was replaced with **Strand 2** Fmoc Val selenoester DNA **4-20** (20 µL, 6.4 µM). Piperidine wt% varied from 5 wt% to 1 wt% and 5 wt%. Solutions were analysed using LC-MS.

Solution phase: in a PCR tube, 3 µL pure **Strand 2** Fmoc Val selenoester DNA **4-20** (45 µM) was added to 7 µL 1 wt% piperidine and mixed for 20 minutes at 1000 rpm. The solution was diluted in 490 µL H_2O and washed using an Amicon-Ultra 3 kDa 0.5 mL centrifugal filter. Solutions were analysed using LC-MS.

4.4.10 Synthesis of MMT-Ala-OSu 4-22

Following the procedure presented by Brandsen *et al*.[17]. In a round-bottom flask, Ala (0.267 g, 3.0 mmol) was dissolved in CH_2Cl_2 (25 mL) until 80% dissolved. Trimethylsilyl chloride (1.4 mL, 11 mmol) was added, followed by triethylamine (1.6 mL, 11.4 mmol). The reaction was heated under reflux (70 ºC) for 2 hours. The heat was removed and cooled to room temperature (21 ºC). Once achieved, 4-methoxytriphenylmethyl chloride (0.92 g, 3.0 mmol) was added and refluxed again for 2 hours. Afterwards, the reaction was cooled to room temperature and, again, in an ice bath. Triethylamine (4 mL) and MeOH (4 mL) were added to quench the reaction. Performing an aqueous workup, the reaction was dissolved in CH_2Cl_2 (150 mL), washed with 5% citric acid (30 mL) and concentrated NaCl (30 mL). The organic layer was extracted, dried with $MgSO_4$ and reduced under vacuum to produce a pale-yellow crude mixture of MMT-Ala-OH **4-21**. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 7.42 – 7.39 (m, 4H), 7.29 – 7.14 (m, 10H), 6.83 (d, *J* = 8.9 Hz, 2H), 3.71 (s, 3H), 3.01 (q, *J* = 6.9 Hz, 1H), 1.00 (d, *J* = 7.0 Hz, 3H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 177.4 (1C), 157.5 (1C), 146.9 (1C), 146.7 (1C), 138.2 (1C), 129.8 (2C), 128.4 (3C), 128.3 (2C), 127.7 (2C), 126.1 (1C), 112.9 (2C), 70.4 (1C), 54.9 (1C), 51.9 (1C), 21.5 (1C). **TOF MS**: (ES−) *m/z* $[M]^-$ expected: 360.17, found: 360.16 m/z. See Supplementary Figures 4-8 to 4-10 for characterisation.

In a round-bottom flask, the crude mixture of MMT-Ala-OH **4-21** (0.43 g, estimated 1.19 mmol) was suspended in CH_2Cl_2 (5 mL), with N-hydroxysuccinimide (0.28 g, 2.4 mmol), until dissolved. The reaction was cooled in an ice bath, and EDC·HCl (0.46 g, 2.4 mmol) was added. Removing the ice bath, DIPEA (150 µL) was added, and the reaction was stirred for 18 hours. Performing an aqueous wash, the

reaction was diluted in CH_2Cl_2 (20 mL) and washed with 5% citric acid (3 × 20 mL) and brine (20 mL). The crude mixture was purified by silica column chromatography, using a gradient of 50:50 hexane/EtOAc → 100% EtOAc → 20:80 MeOH/EtOAc. MMT-Ala-OSu **4-22** was isolated as a pale-yellow powder (0.12 g, 23%). **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 7.45 (m, 4H), 7.35 – 7.16 (m, 9H), 6.86 (d, *J* = 9.0 Hz, 2H), 3.72 (s, 3H), 3.45 (q, *J* = 8.2, 7.0 Hz, 1H), 2.80 (s, 4H), 0.85 (d, *J* = 7.1 Hz, 3H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 171.5 (1C), 170.5 (1C), 158.1 (1C), 146.7 (2C), 138.1 (1C), 130.0 (2C), 128.7 (2C), 128.6 (2C), 128.4 (4C), 126.7 (2C), 113.8 (2C), 71.2 (1C), 55.4 (1C), 51.3 (1C), 25.9 (2C), 20.3 (1C). **TOF MS**: (ES+) m/z $[M+Na]^+$ expected: 481.18, found: 481.18 m/z. **TLC:** (EtOAc 100%) R_f = 0.80. See Supplementary Figures 4-11 to 4-12 for characterisation.

4.4.11 Strand 2 MMT Ala selenoester DNA 4-23 assay

TCEP: The synthetic procedure follows that of *Chapter 3 Experimental 3.4.11* in which the NH_2-labelled DNA **2-8** was replaced with **Strand 2** NH_2-labelled DNA **4-16** (1 mM, H_2O) and TAMRA-NHS **2-6** was replaced with MMT-Ala-OSu **4-22** (0.1 M, DMF).

DTT: The synthetic procedure follows that as above (TCEP), in addition to 4 µL TCEP (0.1 M, H_2O) being replaced with 4 µL DTT (0.1 M, H_2O).

4.4.12 Synthesis of NVOC-Ala-OSu 4-25

Following the procedure presented by Robertson *et al.*[25], in a brown round-bottom flask, Ala (0.053 g, 0.6 mmol), Na_2CO_3 (0.063 g, 0.6 mmol) was dissolved in nanopure H_2O (4 mL) and dioxane (4 mL). In a separate flask,

4,5-Dimethoxy-2-nitrobenzyl chloroformate (0.1652, 0.6 mmol) was dissolved in dioxane (8 mL). The 4,5-dimethoxy-2-nitrobenzyl chloroformate mixture was then slowly, with stirring, added to a brown round-bottom flask. After 5 hours of mixing, $NaHSO_4$ (1 M, 5 mL) was added to quench the reaction, followed by CH_2Cl_2 (7.5 mL) and an aqueous wash was performed. The acidic layer was washed with CH_2Cl_2 (3 × 50 mL). The organic layer was extracted, dried with $MgSO_4$ and reduced under vacuum. **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) $\delta_{H/ppm}$: 12.44 (s, 1H), 7.86 (d, *J* = 7.7 Hz, 1H), 7.70 (s, 1H), 7.19 (s, 1H), 5.35 (q, *J* = 15.1 Hz, 2H), 4.02 (q, *J* = 8.5 Hz, 1H), 3.93 (s, 3H), 3.87 (s, 3H), 1.29 (d, *J* = 7.4 Hz, 3H). **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) $\delta_{C/ppm}$: 174.8 (1C), 155.9 (1C), 154.0 (1C), 148.1 (1C), 139.5 (1C), 128.6 (1C), 110.5 (1C), 108.6 (1C), 62.9 (1C), 56.6 (2C), 49.7 (1C), 17.5 (1C). Characterisation support from Kessler *et al*. (NVOC-Leu-OH compared to NVOC-Ala-OH **4-24**).[26] **TOF MS**: (ES+) *m/z* $[M+Na]^+$ expected: 351.09, found: 351.08 m/z. See Supplementary Figures 4-15 to 4-17 for characterisation.

Without purification, NVOC-Ala-OH **4-24** (0.036 g, 0.1 mmol) was dissolved in a 1:5 DMF:CH_2Cl_2 mix (500 µL). EDC-HCl (0.025 g, 0.13 mmol) and NHS (0.015 g, 0.13 mmol) were added to the reaction mixture, followed by DIPEA (20 µL). After mixing for 18 h, the crude solution was analysed by LC-MS (concentration too low for NMR spectroscopy). **TOF MS**: (ES+) *m/z* $[M+Na]^+$ expected: 448.11, found: 488.10 m/z. See Supplementary Figure 4-18 for characterisation.

4.4.13 Synthesis of NVOC-Ala selenoester DNA

Stock solutions: impure **Strand 2** Boc-Sec DNA **4-18** (~0.5 mM, H_2O), NVOC-Ala-OH **4-24** (300 mM, DMF), EDC-HCl (300 mM, DMF), NHS (300 mM, DMF), NVOC-Ala-OSu **4-25** (0.1 M, DMF).

Route A: In a primary microcentrifuge tube, 6.7 µL of NVOC-Ala-OH **4-24** stock (300 mM), 6.7 µL of EDC-HCl stock (300 mM), 6.7 µL of NHS (300 mM) and 0.5 µL of DIPEA were added and mixed for 20 minutes, 19 ºC at 1000 rpm. In a secondary microcentrifuge tube, 2 µL of impure **Strand 2** Boc-Sec DNA **4-18** (~0.5 mM, H_2O) was mixed with 4 µL of DTT (0.1 M, H_2O) for 10 minutes before adding 18 µL borate buffered saline (100 mM, pH 9.8), 20 µL DMF and the contents of the primary microcentrifuge tube. The combined solution was mixed for 18 h at 19 ºC and 1000 rpm. The solution was dissolved in 300 µL ethanol with 20 µL sodium acetate (NaOAc) (0.3 M, pH 5.2) and placed in the freezer overnight. The modified DNA pellets were centrifuged (15000 rpm), emptied and washed twice with ice-cold ethanol (70%, H_2O). The DNA pellet was dried, dissolved in 500 µL H_2O and spin-filtered (Amicon-Ultra 3 kDa 0.5 mL centrifugal filter) to remove any remaining small molecules. The sample was then analysed by LC-MS.

Route B: In a microcentrifuge vial, 2 µL of **Strand 2** Boc-Sec DNA **4-18** (~0.5 mM, H_2O) solution was added to 4 µL of DTT (0.1 M, H_2O) and mixed for 10 minutes at 19 ºC, 1000 rpm, before adding 1 µL of NVOC-Ala-OSu **4-25** (0.1 M, DMF) and 7 µL of borate buffered saline (100 mM, pH 9.8) and mixed for 18 h, at 19 ºC, 1000 rpm. After 18 h, 300 µL ethanol and 20 µL NaOAc (0.3 M, pH 5.2) were

added, and the tube was placed in the freezer overnight. The tubes were centrifuged, washed with ice-cold ethanol, dried and spin-filtered, followed by LC-MS analysis.

Route C: Modified synthesis of Route A in which DTT (0.1 M, H_2O) was replaced with TCEP (0.1 M, H_2O).

Route D: Modified synthesis of Route B in which DTT (0.1 M, H_2O) was replaced with TCEP (0.1 M, H_2O).

4.4.14 Short-strand NVOC-Ala selenoester DNAs 4-27 and 4-28

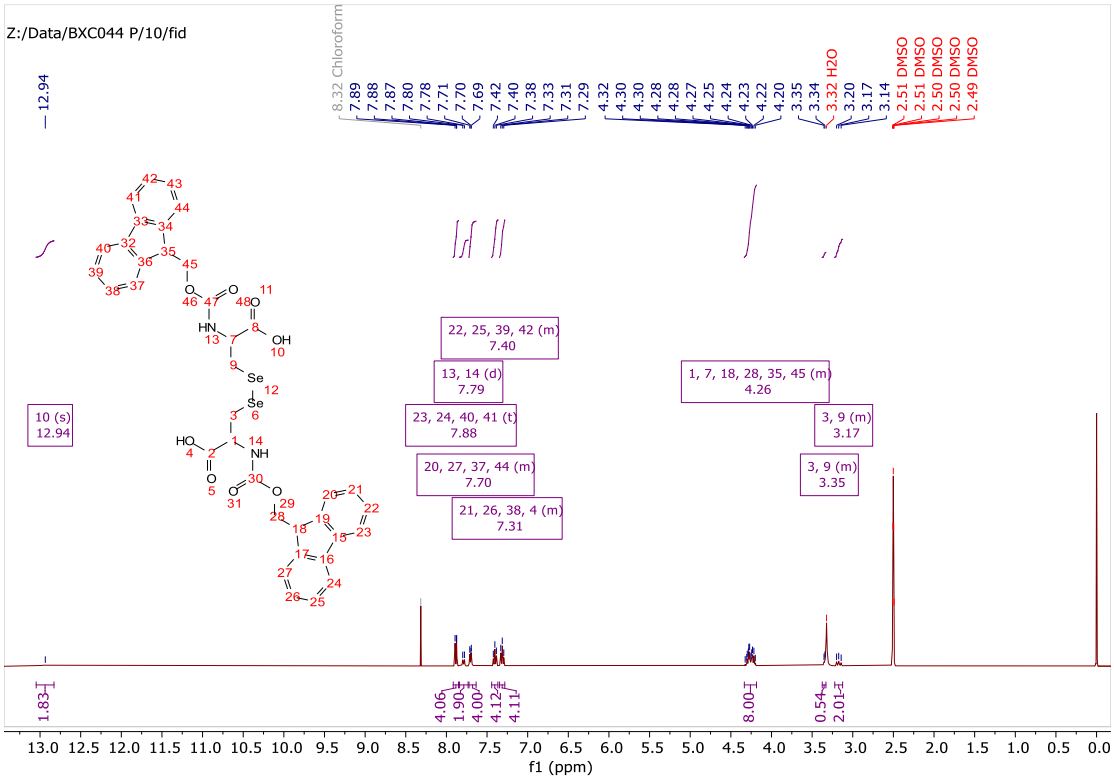
4-27: The synthetic procedure follows that of *Experimental 4.4.13* in which **Strand 2** Boc-Sec DNA **4-18** (~0.5 mM, H_2O) was replaced with the shorter strand Boc-Sec DNA **3-20** (~0.5 mM, H_2O). **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 11442.97 m/z, found: 11442.60 m/z. See Figure 4-13 for the characterisation spectrum.

4-28: The synthetic procedure follows that of *Chapter 3 Experimental 3.4.11* in which the NH_2-labelled DNA **2-8** was replaced with ssDNA-NH_2 +1 position **2-2**, TAMRA-NHS **2-6** with NVOC-Ala-OSu **4-25** and TCEP (0.1 M, H_2O) was replaced with DTT (0.1M, H_2O). **TOF MS**: (ES−) *m/z* $[M]^0$ calculated: 11607.99 m/z, found: 11607.89 m/z. See Figure 4-13 for the characterisation spectrum.

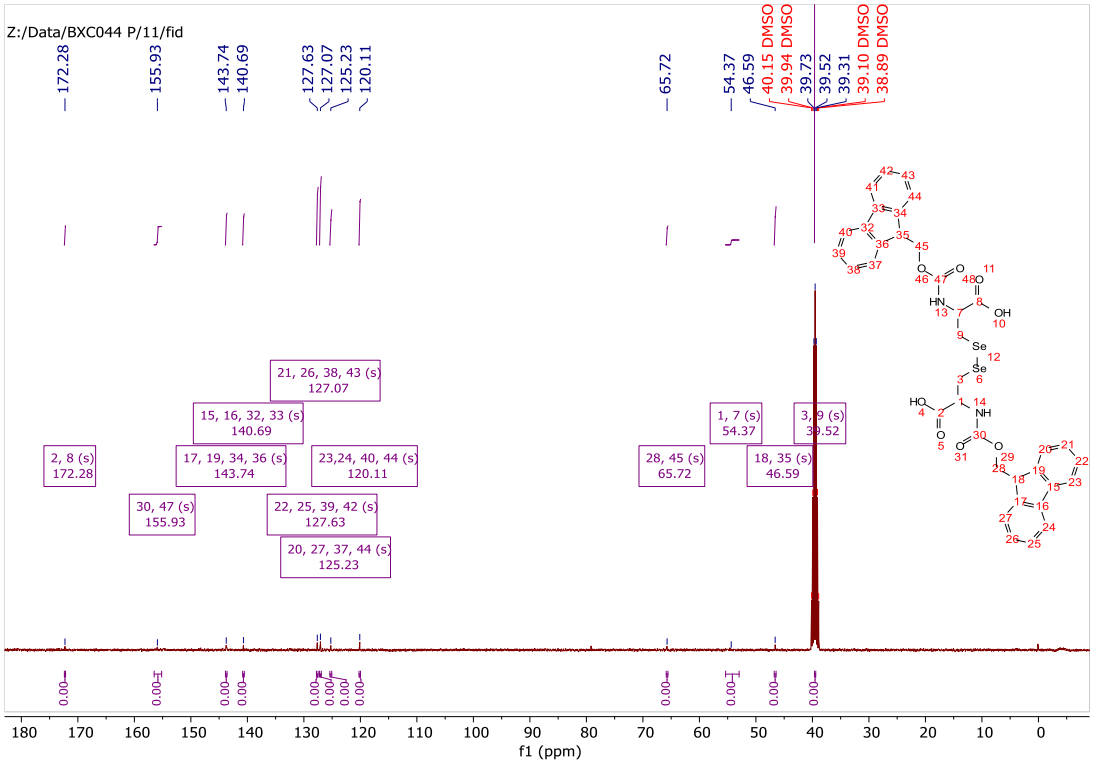
4.4.15 Single-step DNA-templated synthesis assay

See *Chapter 3 Experimental 3.4.12* for single-step DTS assay. Modified by changing DNA-NH_2 +1 position **2-2** with NVOC-Ala selenoester DNA +1 **4-28**. During the 20 minutes of incubation in 0.5 M $MgCl_2$ salt solution, the sample was exposed to a 365 nm handheld UV lamp.

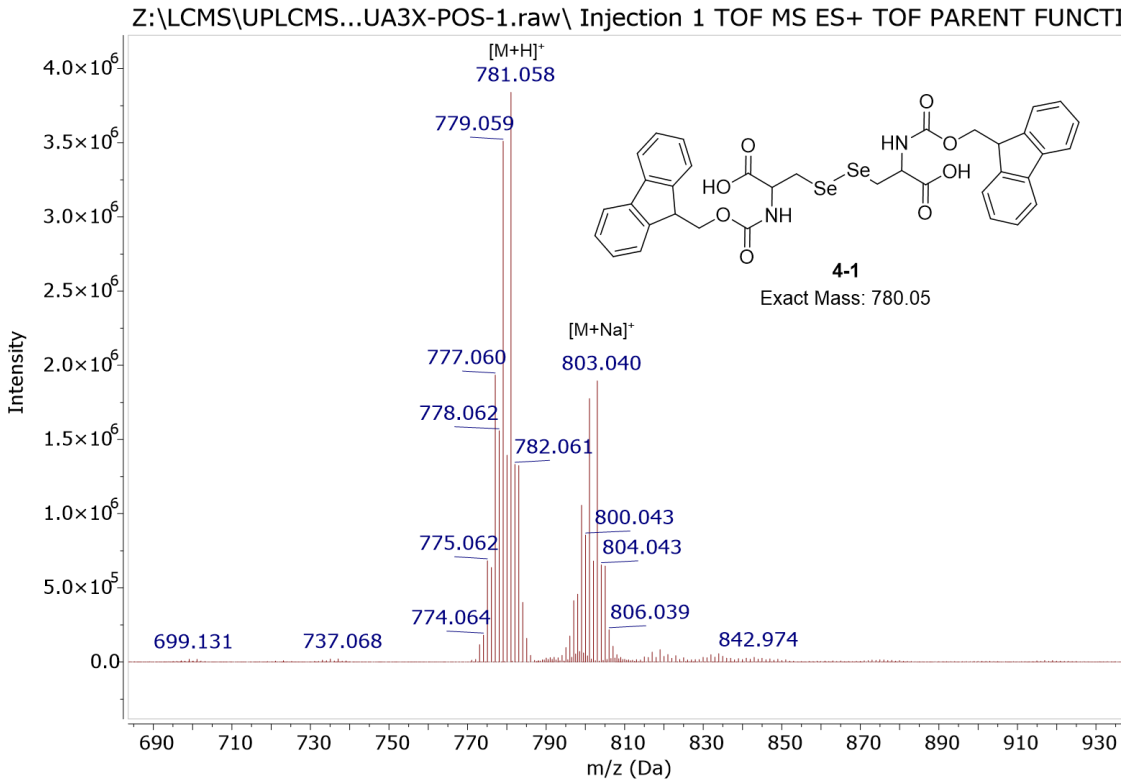
4.5 Supplementary data



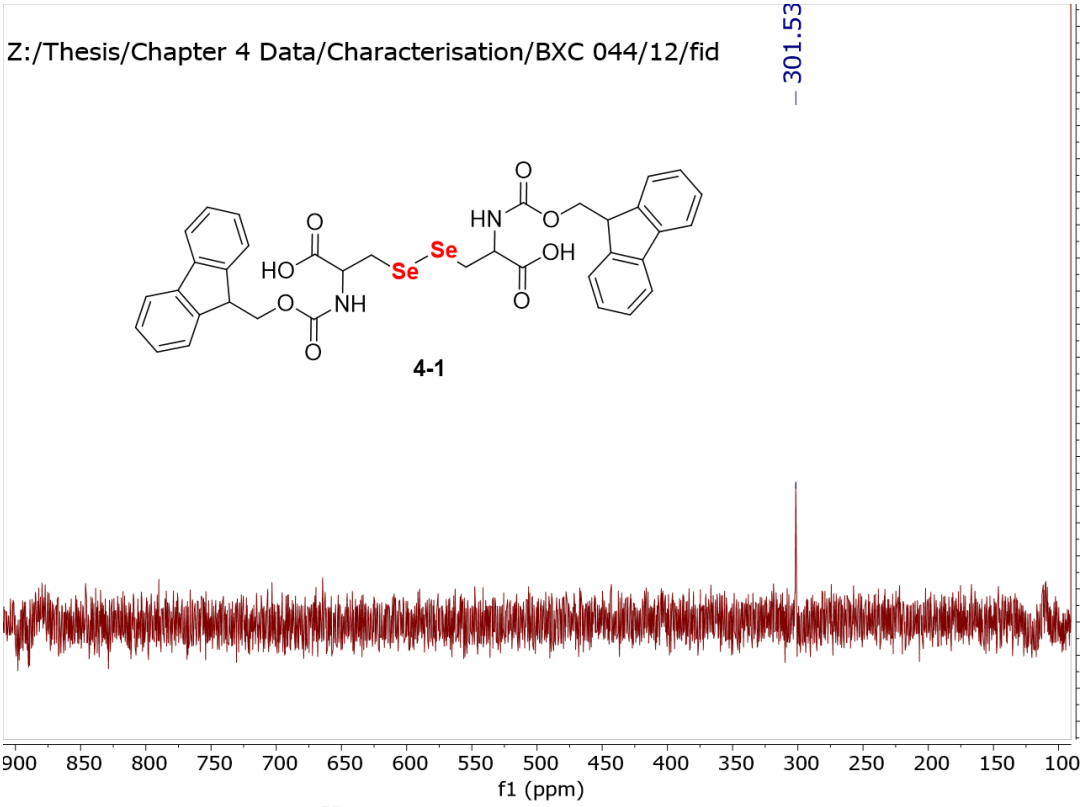
***Supplementary Figure 4-1**: ^{1}H NMR spectrum of (Fmoc-Sec-OH)$_2$ **4-1,** 400 MHz, DMSO-d$_6$.*



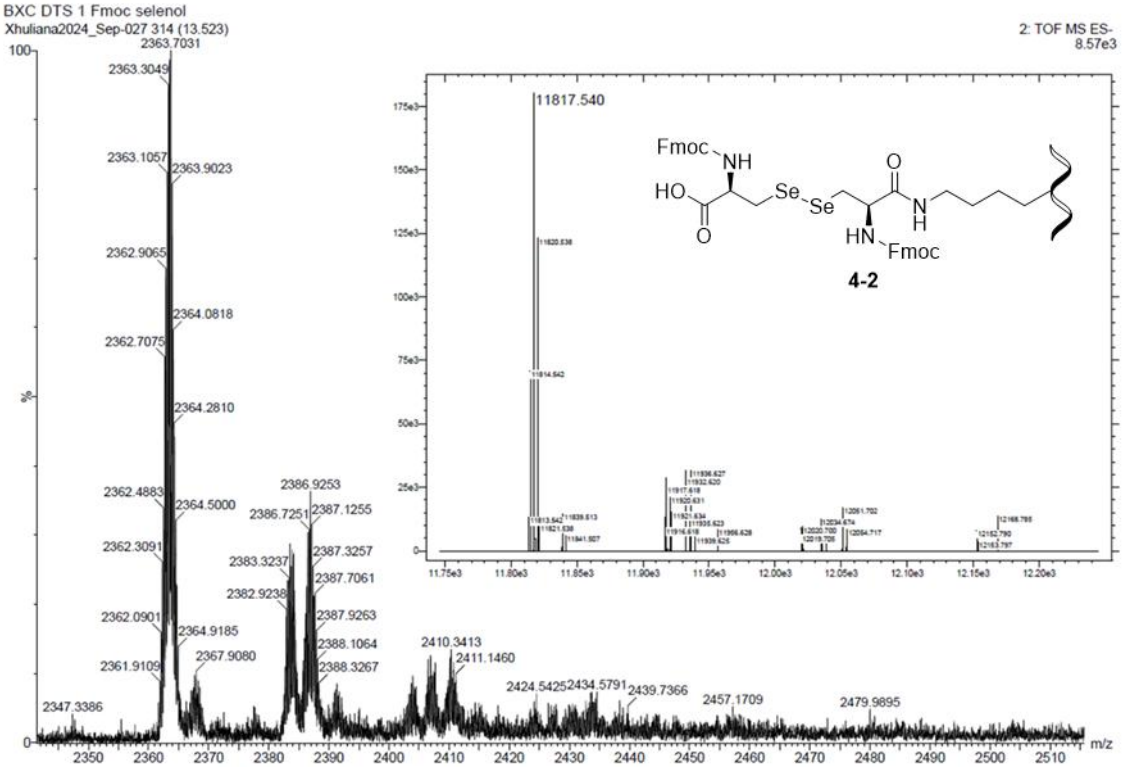
***Supplementary Figure 4-2:** ^{13}C NMR spectrum of (Fmoc-Sec-OH)$_2$ **4-1**,101 MHz, DMSO-d$_6$.*



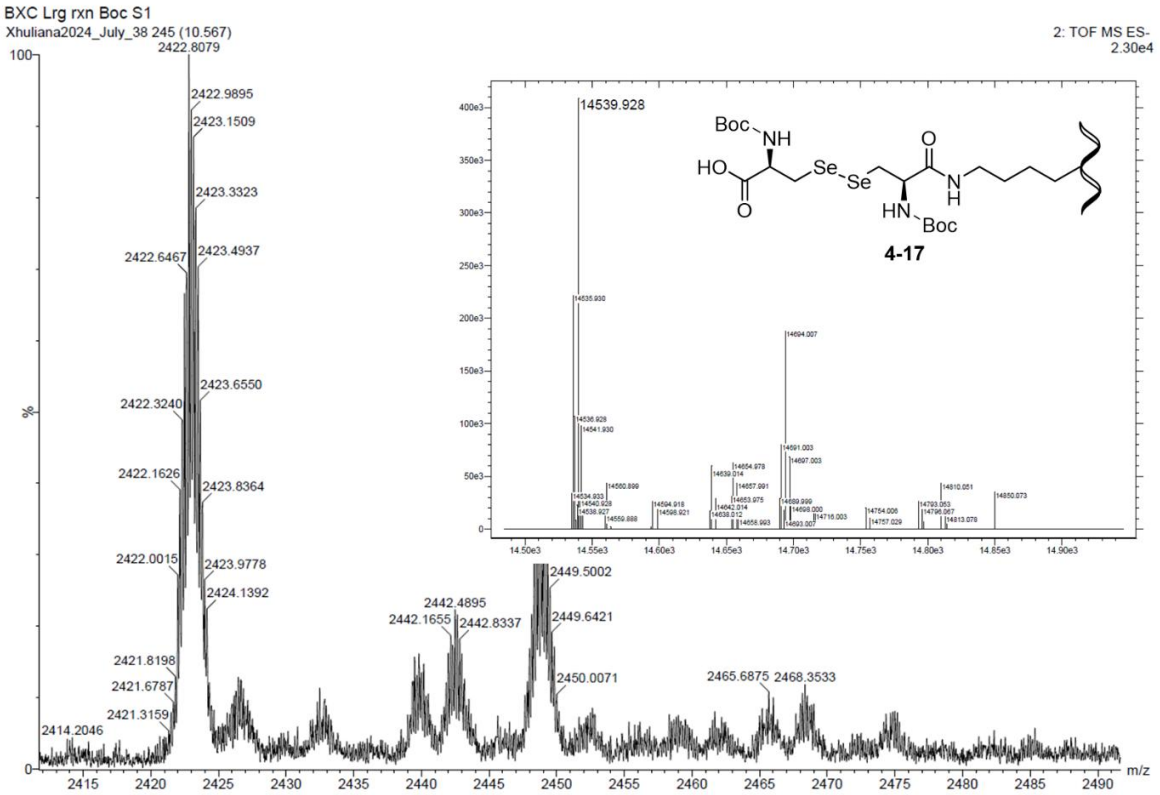
Supplementary Figure 4-3: *TOF MS ES+ of $(Fmoc\text{-}Sec\text{-}OH)_2$* ***4-1*** *$[M+H]^+$ expected: 781.05, found: 781.06 m/z.*



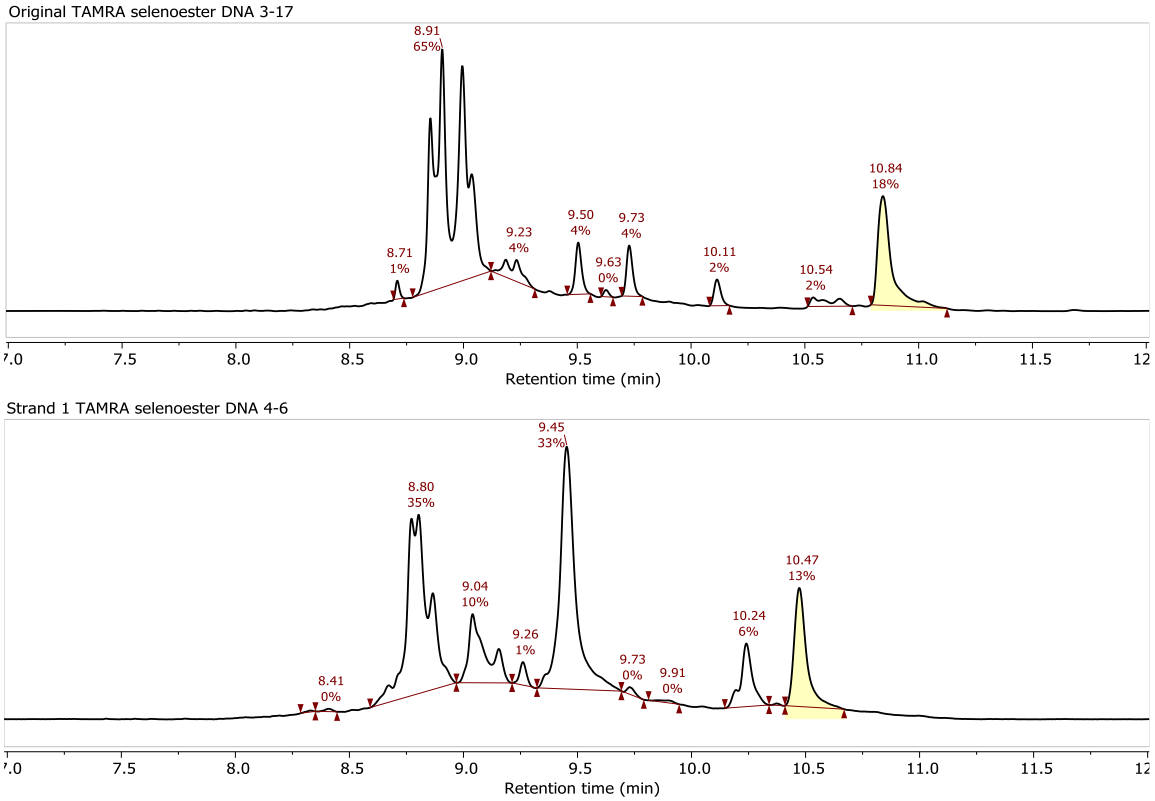
Supplementary Figure 4-4: *^{77}Se NMR spectrum of $(Fmoc\text{-}Sec\text{-}OH)_2$* ***4-1*** *($\delta_{Se/ppm}$ = 302), 76 MHz, DMSO-d_6.*



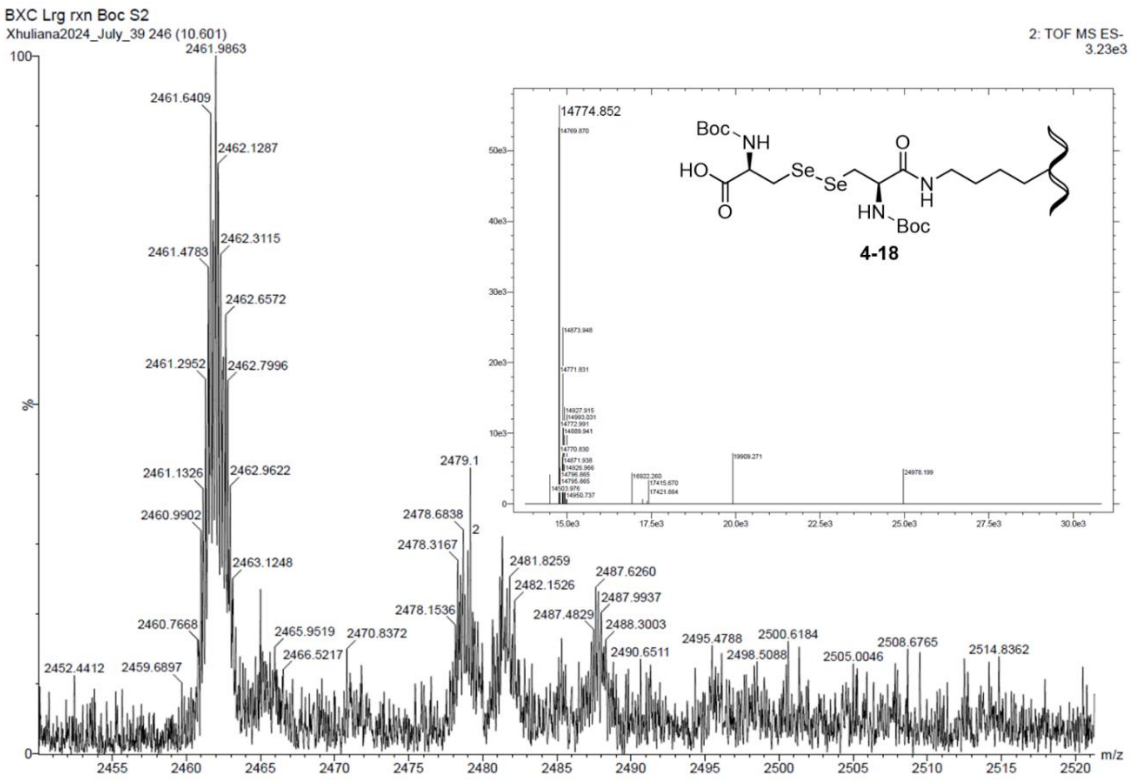
Supplementary Figure 4-5: *TOF MS ES− and deconvolution (ProMass HR) of Fmoc-Sec-DNA* ***4-2****. Expected Mass: 11818.94, found: 11817.54 m/z.*



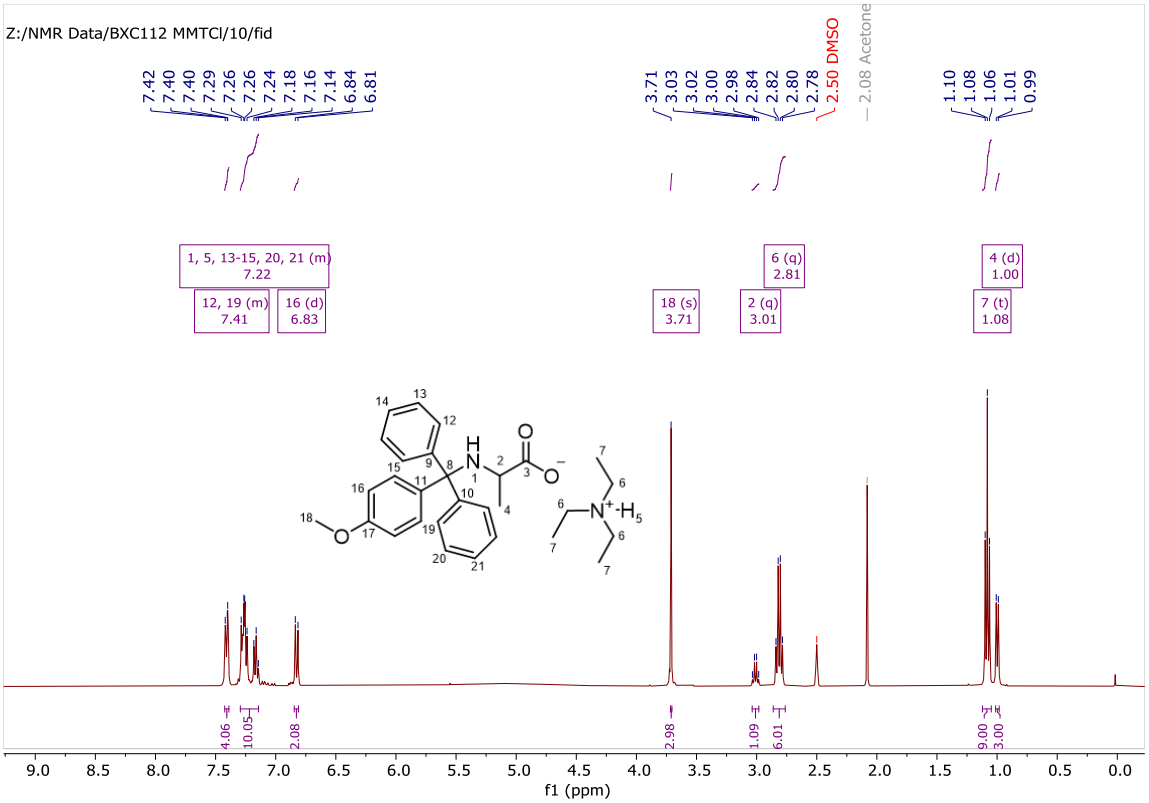
Supplementary Figure 4-6: *TOF MS ES− and deconvolution (ProMass HR) of* ***Strand 1*** *Boc-Sec-DNA* ***4-17****. Expected Mass: 14540.37, found: 14539.93 m/z.*



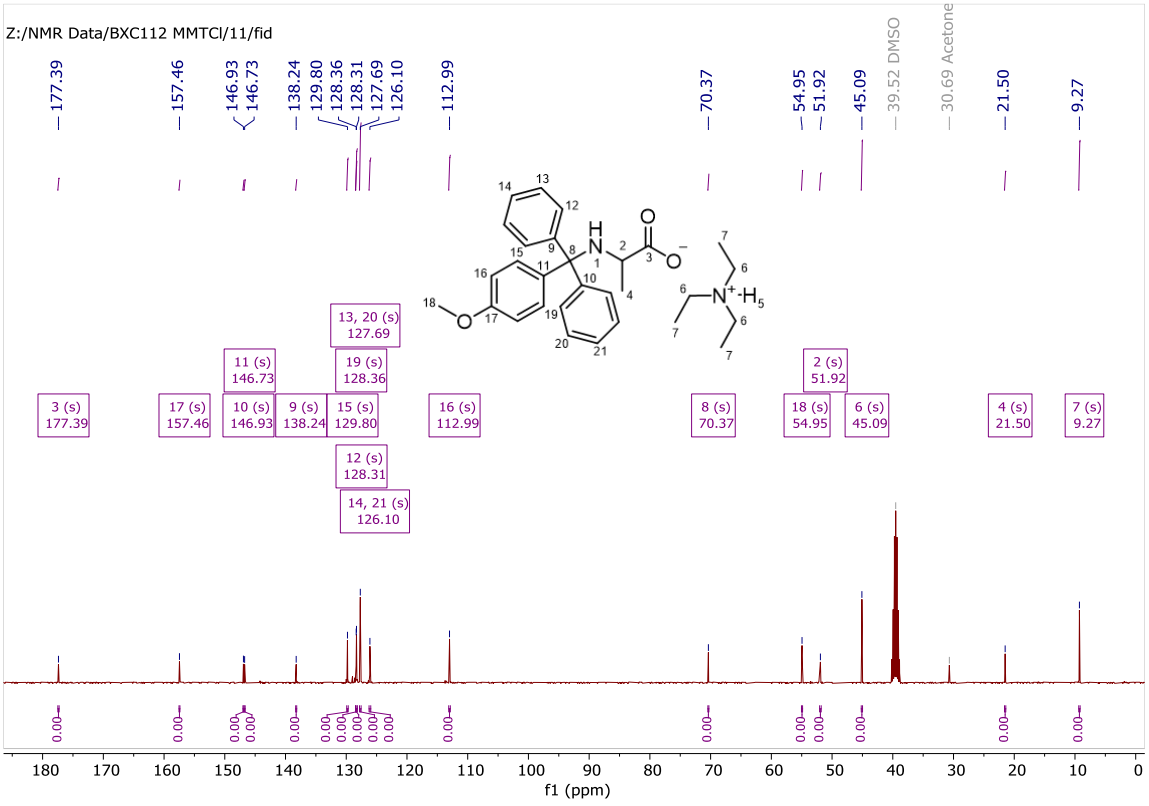
Supplementary Figure 4-7: *Calculation of original TAMRA selenoester DNA* ***3-17*** *(above) and Strand 1 TAMRA selenoester DNA* ***4-6*** *(below) yields using LC-MS PDA channel. Calculated using Mestrelab's Mestrenova software.*



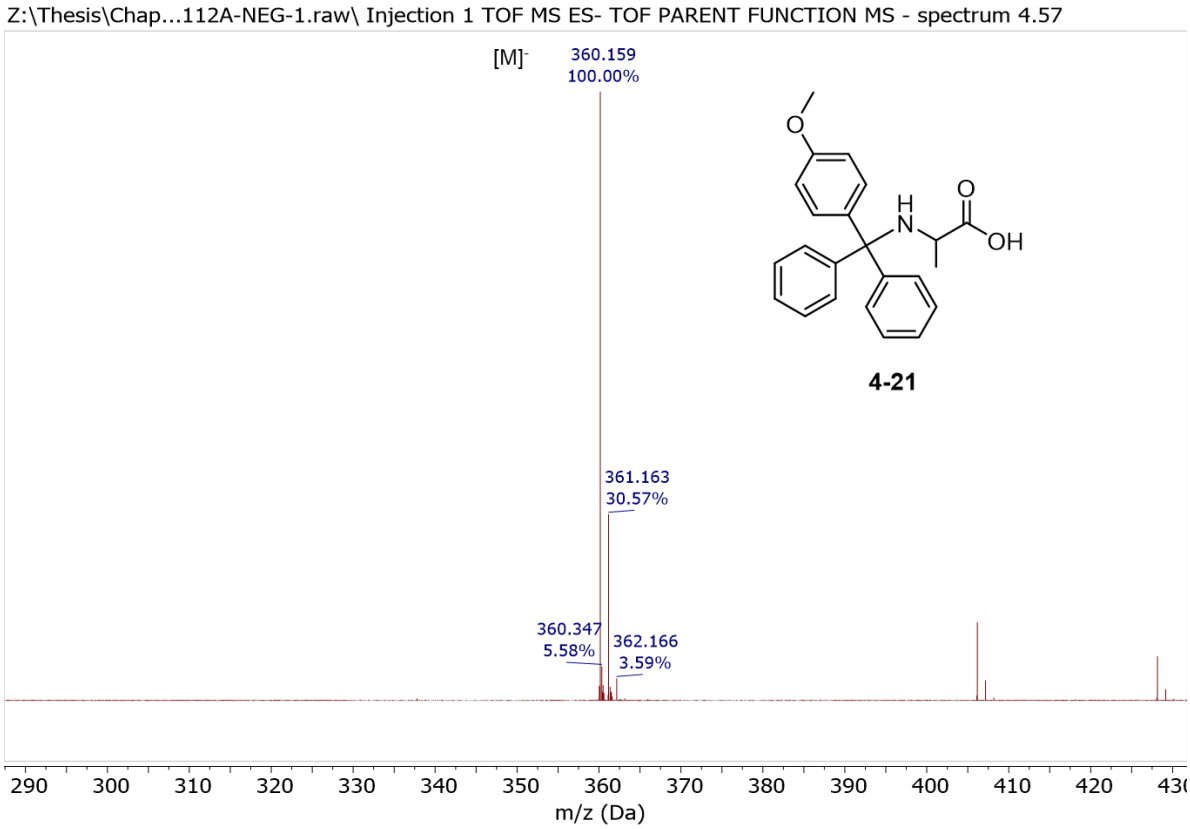
Supplementary Figure 4-8: *TOF MS ES− and deconvolution (ProMass HR) of* ***Strand 2*** *Boc-Sec-DNA* ***4-18****. Expected Mass: 14774.42, found: 14774.85 m/z.*



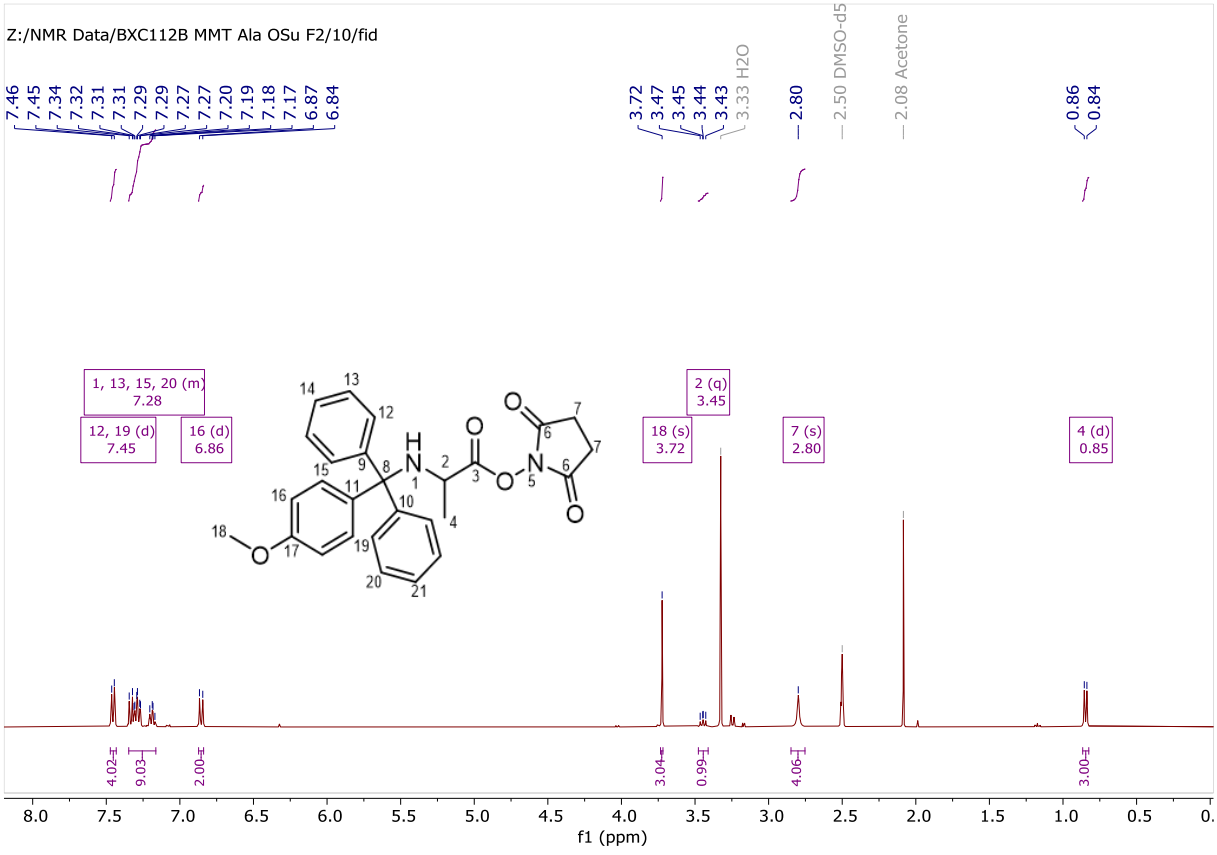
Supplementary Figure 4-9: *^{1}H NMR spectrum of MMT-Ala-OH* ***4-21,*** *400 MHz, DMSO-d_6.*



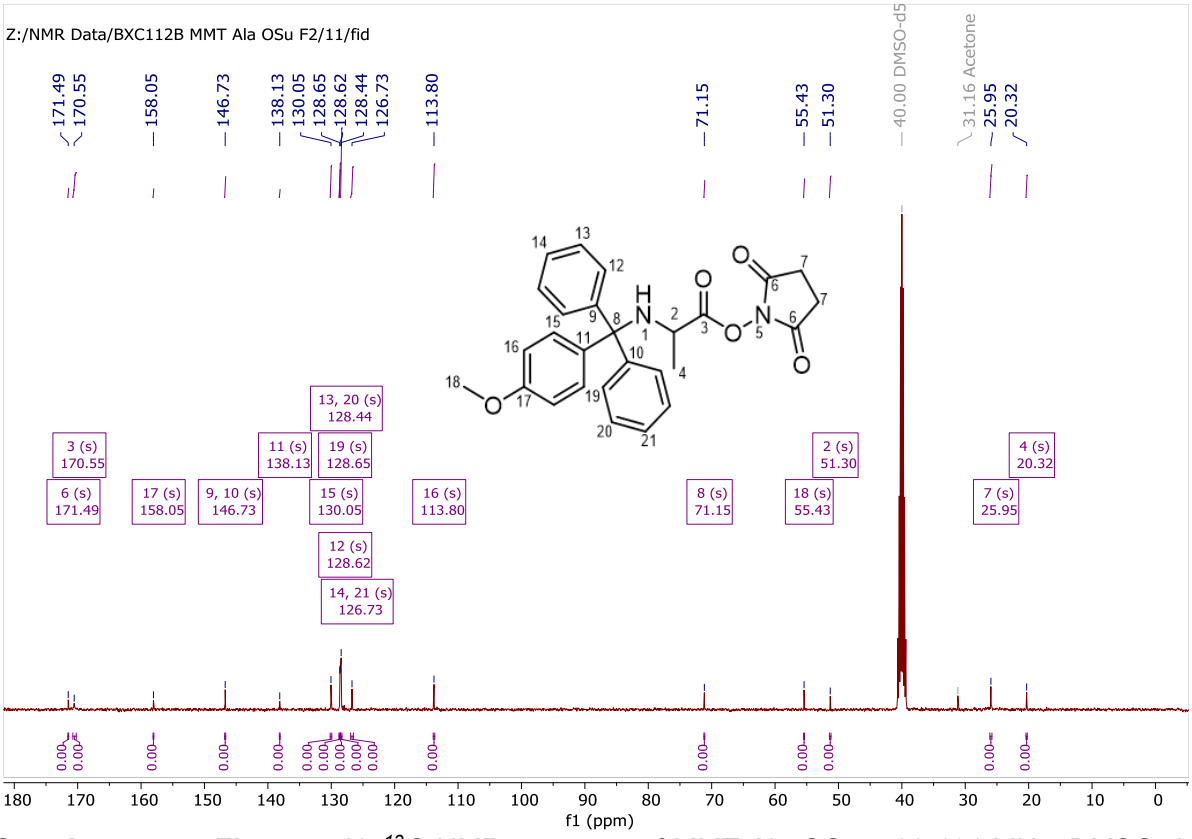
Supplementary Figure 4-10: *^{13}C NMR spectrum of MMT-Ala-OH* ***4-21****,101 MHz, DMSO-d_6.*



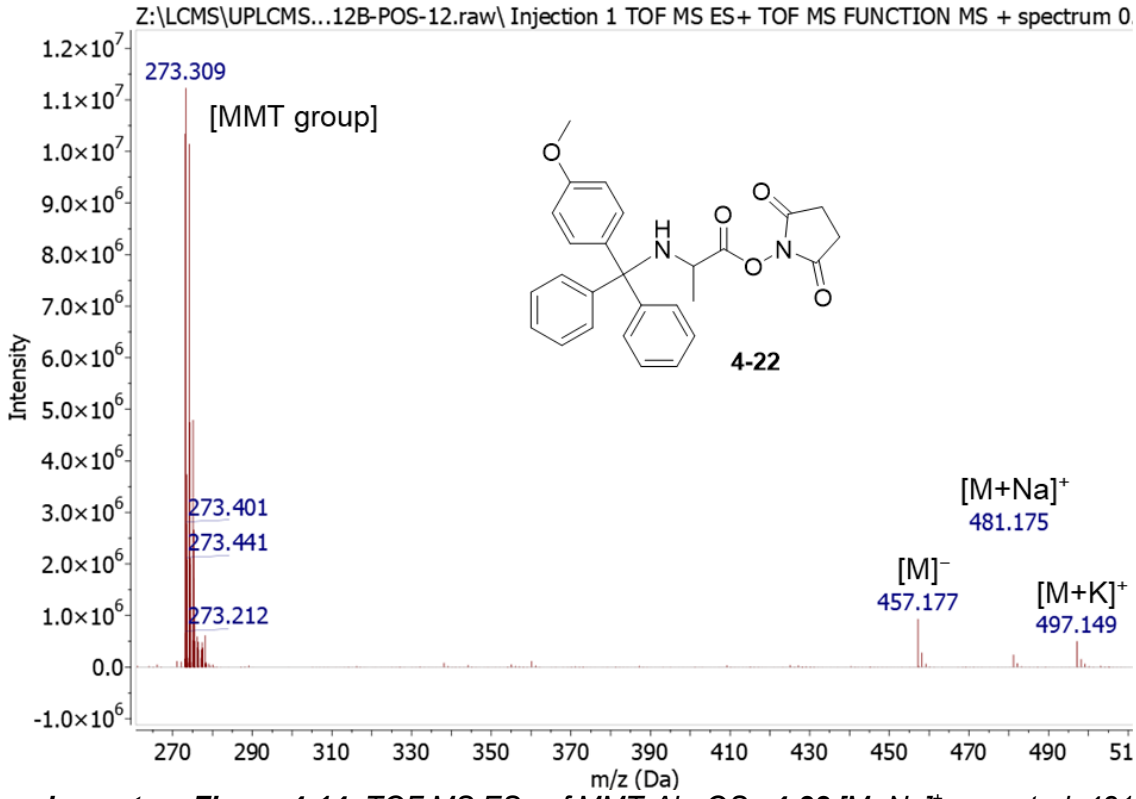
Supplementary Figure 4-11: *TOF MS ES− of MMT-Ala-OH* ***4-21*** *$[M]^-$ expected: 360.17, found: 360.16 m/z.*



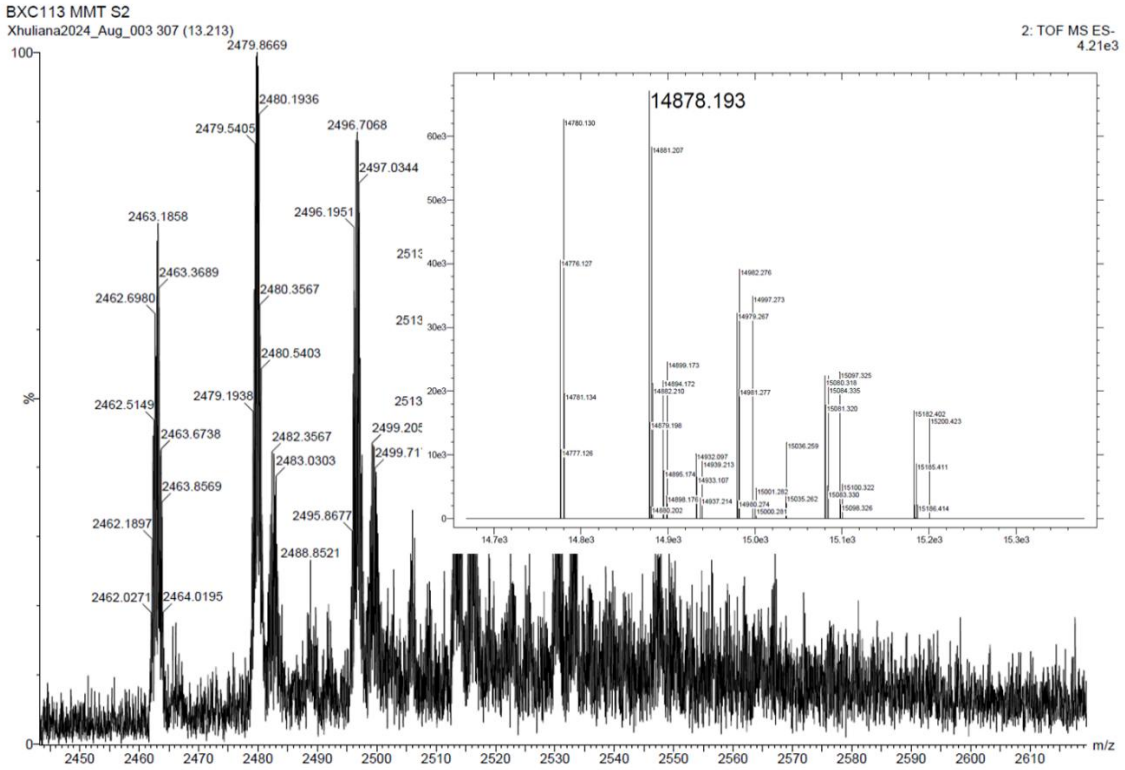
Supplementary Figure 4-12: *^{1}H NMR spectrum of MMT-Ala-OSu* ***4-22,*** *400 MHz, DMSO-d_6.*



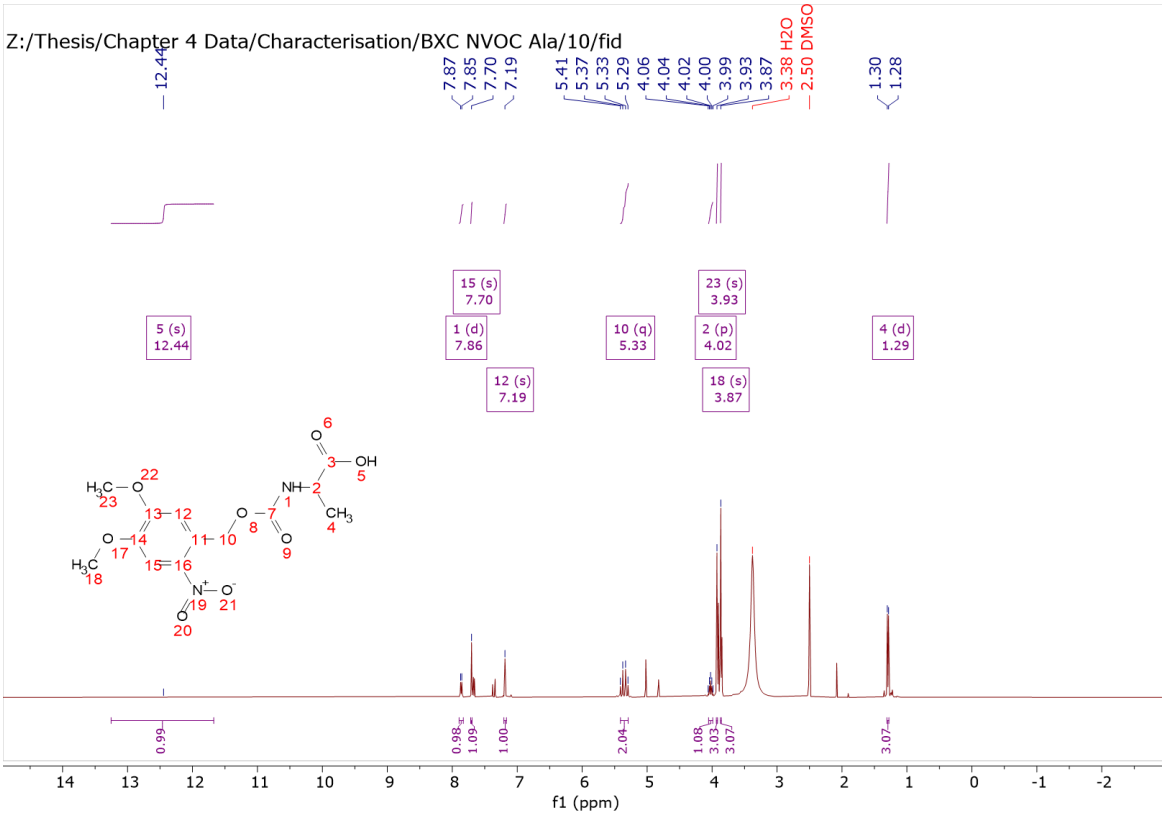
Supplementary Figure 4-13: *^{13}C NMR spectrum of MMT-Ala-OSu* ***4-22****,101 MHz, DMSO-d$_6$.*



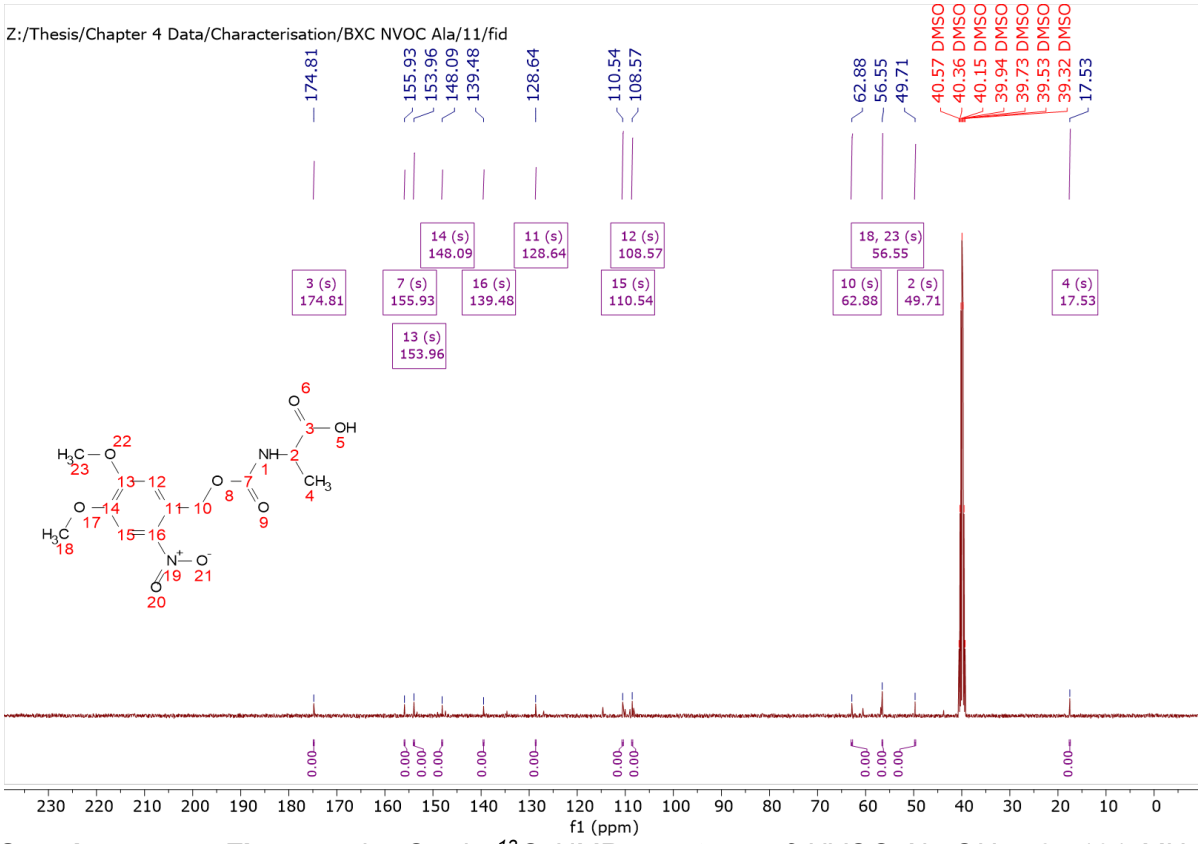
Supplementary Figure 4-14: *TOF MS ES+ of MMT-Ala-OSu* ***4-22*** *[M+Na]$^+$ expected: 481.18, found: 481.18 m/z.*



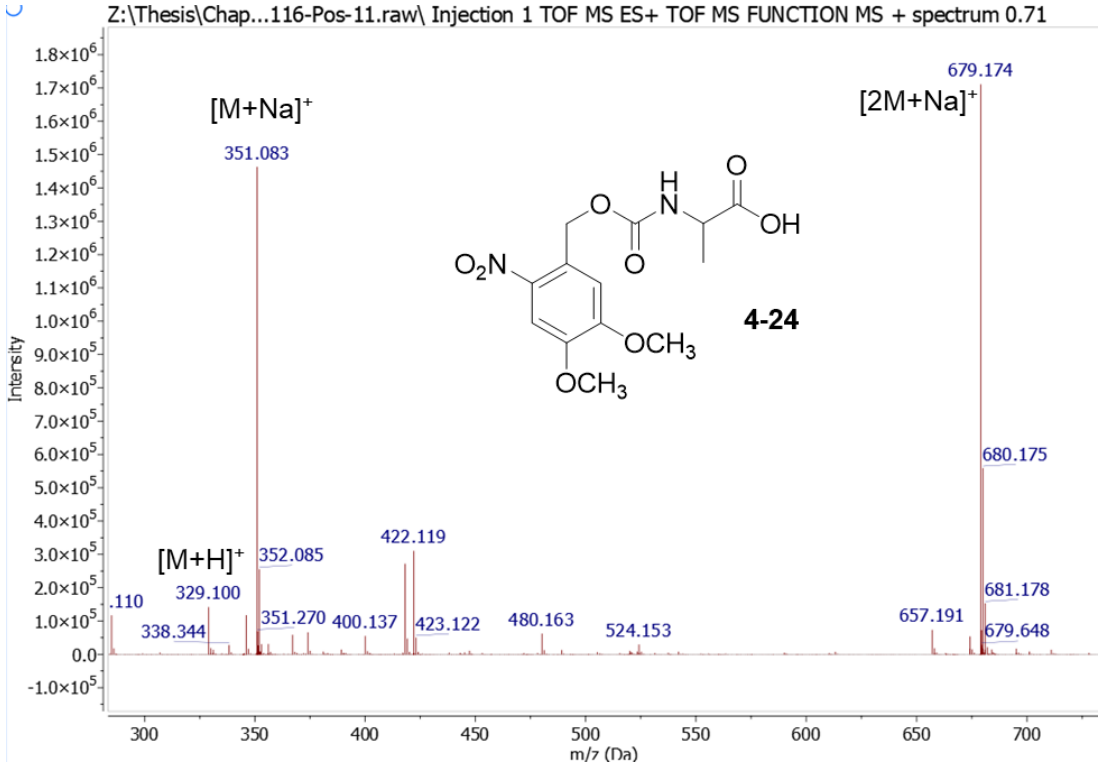
Supplementary Figure 4-15: *TOF MS ES− and deconvolution (ProMass HR) of MMT-Ala-OSu* ***4-22*** *with* ***Strand 2*** *Boc-Sec DNA* ***4-18****. Expected Mass: 14850.58 m/z, found: 14878.19 m/z.*



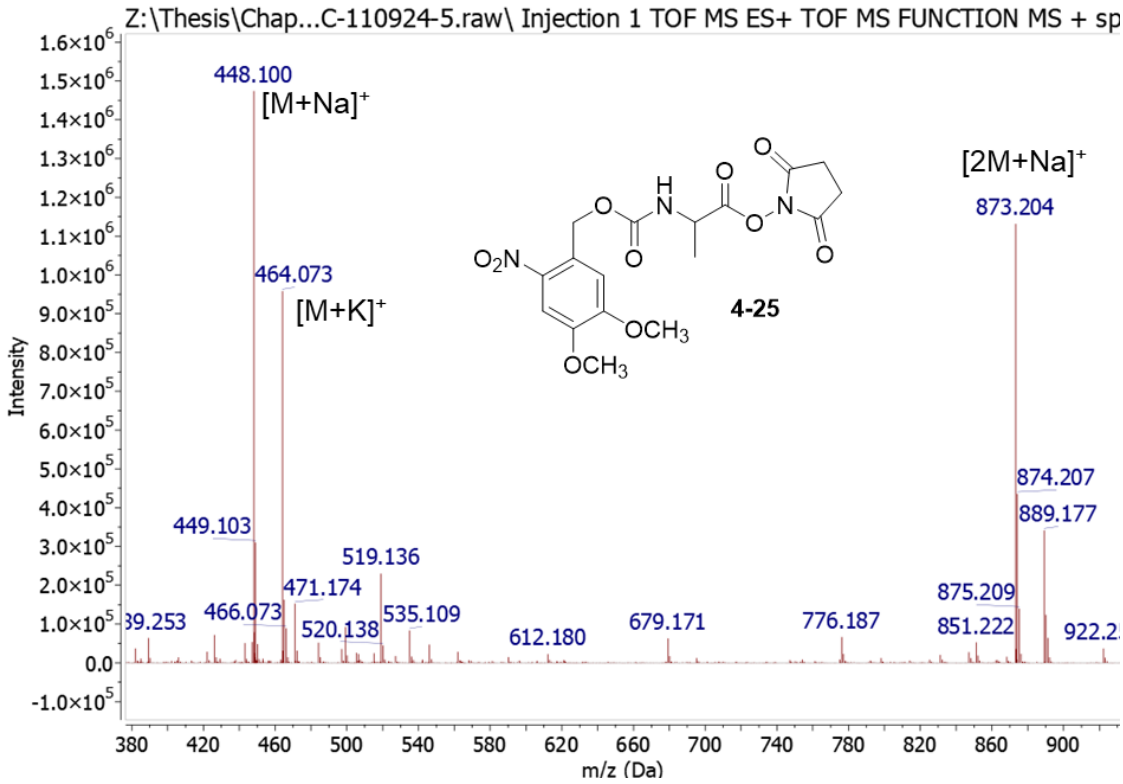
Supplementary Figure 4-16: *Crude* ^{1}*H NMR spectrum of NVOC-Ala-OH* ***4-24,*** *400 MHz, DMSO-*d_6*.*



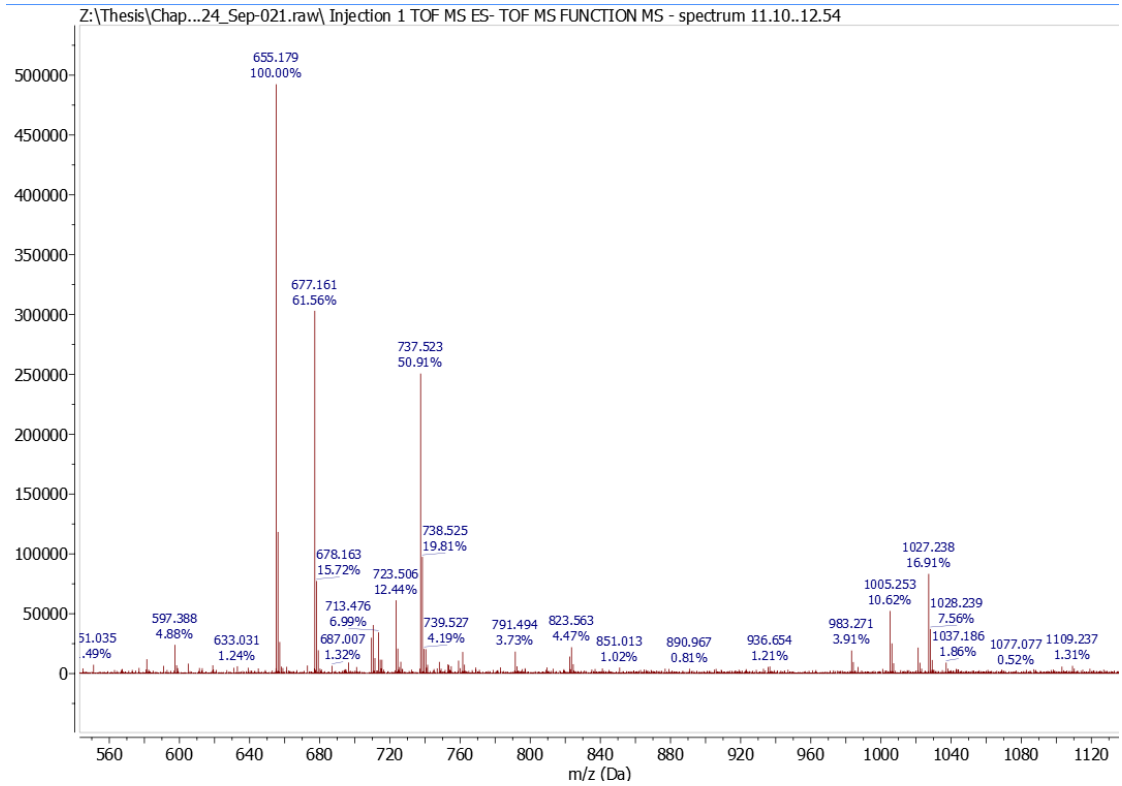
Supplementary Figure 4-17: *Crude ^{13}C NMR spectrum of NVOC-Ala-OH* ***4-24****,101 MHz, DMSO-d_6.*



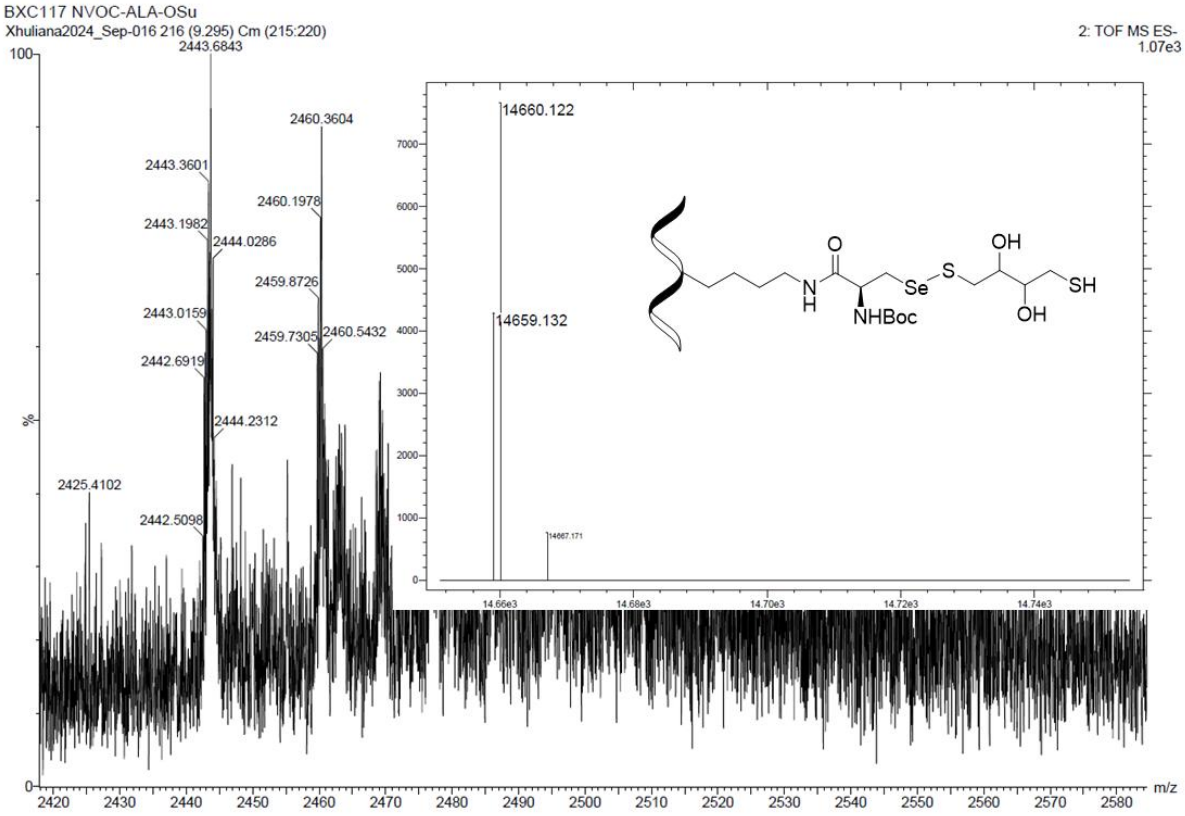
Supplementary Figure 4-18: *TOF MS ES+ of NVOC-Ala-OH* ***4-24*** *$[M+Na]^+$ expected: 351.09, found: 351.08 m/z.*



Supplementary Figure 4-19: *TOF MS ES+ of NVOC-Ala-OSu* ***4-25*** *$[M+Na]^+$ expected: 448.11, found: 488.10 m/z*



Supplementary Figure 4-20: *Combined TOF MS ES− of fragments in* ***Strand 2*** *NVOC-Alal selenoester DNA* ***4-26*** *synthesis assays.*



Supplementary Figure 4-21: *TOF MS ES− and deconvolution (ProMass HR) of Se-S DTT structure found in* ***Strand 2*** *NVOC-Alal selenoester DNA* ***4-26*** *synthesis assays. Expected Mass: 14659.42 m/z, found: 14660.12 m/z.*

4.6 References

1. Z. J. Gartner, M. W. Kanan and D. R. Liu, *J. Am. Chem. Soc.*, 2002, **124**, 10304-10306.
2. T. M. Snyder and D. R. Liu, *Angew. Chem. Int. Ed.*, 2005, **44**, 7379-7382.
3. Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder and D. R. Liu, *Science*, 2004, **305**, 1601-1605.
4. M. H. Hansen, P. Blakskjær, L. K. Petersen, T. H. Hansen, J. W. Højfeldt, K. V. Gothelf and N. J. V. Hansen, *J. Am. Chem. Soc.*, 2009, **131**, 1322-1327.
5. Y. He and D. R. Liu, *J. Am. Chem. Soc.*, 2011, **133**, 9972-9975.
6. Y. He and D. R. Liu, *Nat. Nanotechnol.*, 2010, **5**, 778-782.
7. W. Meng, R. A. Muscat, M. L. McKee, P. J. Milnes, A. H. El-Sagheer, J. Bath, B. G. Davis, T. Brown, R. K. O'Reilly and A. J. Turberfield, *Nat. Chem.*, 2016, **8**, 542-548.
8. J. Sayers, P. M. T. Karpati, N. J. Mitchell, A. M. Goldys, S. M. Kwong, N. Firth, B. Chan and R. J. Payne, *J. Am. Chem. Soc.*, 2018, **140**, 13327-13334.
9. Y. Wang, P. Liu, J. Chang, Y. Xu and J. Wang, *ChemBioChem*, 2021, **22**, 2918-2924.
10. J. G. Wetmur and N. Davidson, *J. Mol. Biol.*, 1968, **31**, 349-370.
11. S. S. Kulkarni, E. E. Watson, B. Premdjee, K. W. Conde-Frieboes and R. J. Payne, *Nature Proto.*, 2019, **14**, 2229-2257.
12. I. Sánchez-Campillo and J. B. Blanco-Canosa, *JACS Au*, 2024, **4**, 4374-4382.
13. M. J. Bedding, S. S. Kulkarni and R. J. Payne, in *Selenoprotein Structure and Function*, ed. E. Weerapana, Academic Press, Cambridge, U.S., 2022, vol. 662, ch. 15, pp. 363-399.
14. R. K. O'Reilly, A. J. Turberfield and T. R. Wilks, *Acc. Chem. Res.*, 2017, **50**, 2496-2509.
15. Invitrogen, *Life Technologies Corporation - Dynabeads™ Streptavidin Trial Kit*, 2018, **65801D**, MAN0015762.
16. S.-C. Hung, J. Ju, R. A. Mathies and A. N. Glazer, *Anal. Biochem.*, 1996, **243**, 15-27.
17. B. M. Brandsen, A. R. Hesser, M. A. Castner, M. Chandra and S. K. Silverman, *J. Am. Chem. Soc.*, 2013, **135**, 16014-16017.
18. J. C. Han and G. Y. Han, *Anal. Biochem.*, 1994, **220**, 5-10.
19. J. R. Peacock, R. R. Walvoord, A. Y. Chang, M. C. Kozlowski, H. Gamper and Y.-M. Hou, *RNA*, 2014, **20**, 758-764.

20. B. S. Vig, P. J. Lorenzi, S. Mittal, C. P. Landowski, H.-C. Shin, H. I. Mosberg, J. M. Hilfinger and G. L. Amidon, *Pharm. Res.*, 2003, **20**, 1381-1388.

21. W. Yang and D. G. Drueckhammer, *J. Am. Chem. Soc.*, 2001, **123**, 11004-11009.

22. J. Frommer, R. Oppenheimer, B. M. Allott, S. Núñez-Pertíñez, T. R. Wilks, L. R. Cox, J. Bath, R. K. O'Reilly and A. J. Turberfield, *Angew. Chem. Int. Ed.*, 2024, **163**, e202317482.

23. S. Flemer, *Protein Pept. Lett.*, 2014, **21**, 1257-1264.

24. D. R. Halpin, J. A. Lee, S. J. Wrenn and P. B. Harbury, *PLoS Biology*, 2004, **2**, e175.

25. S. A. Robertson, J. A. Ellman and P. G. Schultz, *J. Am. Chem. Soc.*, 1991, **113**, 2722-2729.

26. M. Kessler, R. Glatthar, B. Giese and C. G. Bochet, *Org. Lett.*, 2003, **5**, 1179-1181.

Chapter 5

Conclusions and future work

Multistep DTS for the autonomous, controlled synthesis of polymers has been limited due to the degradation of reactive tags, leading to product truncation and limited yields. Therefore, this thesis explored different methods to optimise the DTS transfer reaction, overcoming the hydrolysis of aminolysis electrophiles. Optimisation routes included investigating the cause of thioester protection in an across-the-helix DTS architecture (Chapter 2), determining the sequence dependency of DTS in the architecture (Chapter 2) and the incorporation of a selenoester electrophile into a single-step (Chapter 3) and multi-step (Chapter 4) DTS aminolysis reaction. This research demonstrated the complexities of small molecule-DNA conjugates and how an array of parameters affect the DTS transfer reaction.

Chapter 2 highlighted the impact a TAMRA fluorophore had in stabilising an aminolysis thioester conjugated to a DTS across-the-helix architecture. Changes in the TAMRA environment were observed through absorbance and fluorescence spectral shifts when the location of an abasic site was varied in a DNA sequence. It was hypothesised that the thioester stabilisation was a result of TAMRA's positioning along the minor groove of DNA, forcing the thioester into the hydrophobic core, away from incoming nucleophiles. However, future work would require 2D NMR analysis to conclude the exact positioning of the thioester with respect to the DNA helix. In addition, adjustments to the nucleobase sequence contributed to the degree of thioester stabilisation, DTS yield and altered the TAMRA environment further. Overall,

multiple interdependent interactions between TAMRA, the DNA sequence, the linker and the abasic sites operate within the across-the-helix architecture. Therefore, future work involving the across-the-helix architecture requires careful parameter consideration in order to obtain high DTS yields. Furthermore, to conclude whether the cause of the thioester stabilisation was a result of TAMRA-DNA interactions, rather than thioester-DNA interactions, the replacement of TAMRA with another fluorophore, *e.g*. fluorescein or Cy5, should be investigated in future work.

In Chapter 3, the selenoester successfully performed DTS within 3 hours, at 77% yield, as opposed to the original 68%, within 12 hours, for Chapter 2's thioester analogue. Additionally, the selenoester design performed successful DTS across both the major and minor grooves of the helix, improving the accessibility in the across-the-helix architecture. Overall, selenoesters demonstrated a successful candidate in improving the yields of multistep transfer DTS. Therefore, Chapter 4 implemented a selenoester into a multistep DTS strategy, utilising the across-the-helix architecture. Despite efforts to synthesise a TAMRA selenoester-amine conjugate, aminolysis was not observed in the original, shorter sequence. It was hypothesised that the incorporation of a small transferable functional group (Ala) led to the degradation of the selenoester, due to the increase in accessibility to water molecules. To tackle the instability of amine-selenoesters, future work should replace Ala with a larger transferable group, *e.g*. Phe, whilst utilising MMT/NVOC amine protection. Alternatively, the use of a different DNA architecture, *e.g*. DNA walker or HCR, with the selenoester electrophile, may provide a more suitable DTS multistep system.

Appendix: Experimental Methods

A.1 Materials

Reagents were purchased from Sigma-Aldrich, unless stated, and all solvents used were HPLC grade. Specialised chemicals Seleno-*l*-cystine, 6-TAMRA and 6-TAMRA-SE were purchased from Cambridge Bioscience. DNA samples were purchased from Integrated DNA Technologies (IDT). Dry solvents were purchased directly from Sigma-Aldrich. Samples were stored in 1.5 mL Eppendorf centrifuge vials, 200 µL PCR tubes (Axygen) or 10 mL glass vials at −20 °C. Where selenoesters were involved, samples were stored in the brown equivalent. Unless stated, all chemicals were used as received.

A.2 Reverse Phase High-Performance Liquid Chromatography (RP-HPLC)

A.2.1 Small molecule preparation and analysis:

Small molecule HPLC purification and analysis were performed on a Shimadzu Prominence Modular system using a Shim-pack GISS C18, 4.6 x 125 mm, 5 µm column at a flow rate of 0.5 mL min^{-1}. The mobile phase consisted of a gradient of 0.1% formic acid in H_2O (solvent A) and 0.1% formic acid in MeCN (solvent B). The buffer gradient is as follows: buffer B gradient 1% to 100% from 0 – 28 min, 1% from 28.01 – 30 min. Samples were prepared at 20 °C over 30 minutes.

A.2.2 DNA preparation and analysis:

DNA HPLC purification and analysis were performed on a Shimadzu Prominence Modular system, equipped with a CBM-20A system controller, LC-

20AD solvent deliver module, SIL-20AC HT autosampler, CTO-20AC column oven, SPD-M20A photodiode array UV-Vis detector, RF-20A spectrofluorometric detector and a FRC-10 fraction collector. Analysis was performed using a Waters XBridge™ BEH C18, 130 Å, 4.6 × 50 mm, 2.5 µm column, at a flow rate of 0.8 mL min^{-1}. The mobile phase consisted of 0.1 M TEAA (pH 7) in 95:5 H_2O:MeCN (solvent A) and 0.1 M TEAA (pH 7) in 30:70 H_2O:MeCN (solvent B). The buffer gradient is as follows: 1% buffer B for 3 minutes, 1% to 15% B over 17 minutes, 15% to 95% B over 5 min, 95% to 1% B over 2 min and finally 1% B for 4 min. Samples were performed at 60 °C over 29 minutes. Results were monitored *via* DNA absorption (262 nm) and fluorescence of the TAMRA moiety (λ_{ex} 550 nm, λ_{em} 580 nm). Preparative RP-HPLC samples were desalted using either 0.5 mL or 15 mL Amicon-Ultra 3 kDa centrifugal filters, depending on fraction size.

A.3 Liquid Chromatography-Mass spectrometry (LC-MS):

A.3.1 Small molecule Liquid Chromatography-Mass Spectrometry:

Small molecule LC-MS was performed using Waters Xevo G2-XS qToF coupled with a Waters Acquity UPLC H-Class Plus, PDA detector. Readings were measured using electrospray ionisation in both positive and negative modes (separate readings). A linear gradient of 0.1% formic acid in water to 0.1% formic acid in acetonitrile from 30% to 70%, over 15 minutes at a 0.2 mL^{-1} flow rate, in a Waters Acquity C18 BEH 2.1 x 50 mm, 1.7 µm particle size column was used.

A.3.2 Oligo Liquid Chromatography-Mass Spectrometry

Oligo LC-MS was performed using a Waters Xevo G2-XS qToF coupled with a Waters BioAcquity UPLC, PDA detector. Readings were measured in negative ion mode using an Oligo Waters ACQUITY Premier Oligonucleotide C18 Column, 130 Å, 2.1 x 50 mm, 1.7 µm, particle size column. A linear gradient of 75 mM triethylammonium acetate (TEAA, pH 7.0) in H_2O (Buffer A) and 75 mM triethylammonium acetate (pH 7.0) in MeCN (Buffer B), at 0.2 mL min^{-1} and 60 ºC. Data was deconvoluted using ProMass HR software. Mass spectra containing selenium display additional peaks due to the number of selenium isotopes: ^{74}Se, ^{76}Se, ^{77}Se, ^{78}Se, ^{80}Se (most abundant) and ^{82}Se.

A.4 Nuclear Magnetic Resonance spectrometry

^{1}H, ^{13}C (1D/2D) and ^{77}Se NMR data was obtained from a Bruker AVIII400, 400 MHz NMR spectrometer. Deuterated solvents were indicated and referenced at 298 K. Chemical shifts were measured on the δ scale in parts per million (ppm) and referenced against residual solvent peaks: $(CD_3)_2SO$ – ^{1}H: δ = 2.50 ppm, ^{13}C = 39.5 ppm and reported to 0.01 ppm. ^{1}H NMR multiplicity was indicated as s (singlet), d (doublet), t (triplet), q (quartet) or m (multiplet). Data was processed using Mestrelab's Mestrenova software.

A.5 Infrared spectroscopy

Infrared spectroscopy (IR) data was obtained from an Agilent Technologies Cary 360 FTIR spectrometer. Data was processed using OriginPro.

A.6 Fluorescence spectroscopy

Fluorescence spectroscopy data was obtained from an Edinburgh Instruments FS5 spectrofluorometer, in a Starna 16.45F-Q-3/Z15 cuvette (55 µL well), using a SC25 cuvette holder and a Xenon lamp. Data was collected and processed using Fluoracle.

A.7 UV-Vis spectroscopy

UV-Vis spectroscopy data was obtained from an Agilent Cary 3500 multicell, using a Xenon flash lamp light source in a Starna 16.45F-Q-3/Z15 cuvette (55 µL well). Data was collected and processed using the Cary UV Workstation.

A.8 Nanopure water

Using a Milli-Q Integral Water Purification System with a 0.22 µm filter, water was deionised/purified to 18.2 MΩ.

A.9 DNA chemistries

Ethanol precipitation and spin filtering were performed in an Eppendorf 5425 R 24-place lab standard centrifuge. Samples were mixed in an Eppendorf ThermoMixer C and concentrated in an Eppendorf 5305 Concentrator Plus. Incubated samples were stored in an Eppendorf Mastercycler Nexus GX2e. Amicon-Ultra 3 kDa 0.5 mL/15 mL centrifugal filters were purchased from Millipore. DNA concentrations were determined using a NanoDrop One/One UV-Vis Spectrophotometer at an absorbance of $\lambda = 260$ nm, with the corresponding extinction (ε) coefficient.