Peptide Stapling via Thiophosphoramidate Intermediates

ROBYN EMMA POULTON

Durham University

A thesis entitled

Peptide Stapling *via* Thiophosphoramidate Intermediates

Submitted by

Robyn Emma Poulton

Department of Chemistry

A Candidate for the Degree of Doctor of Philosophy

2024

ACKNOWLEDGMENTS

I would first like to express gratitude to my supervisor, Prof David R W Hodgson, for all of his help and guidance over the course of this PhD project and throughout the writing of this thesis. I would also like to thank the MoSMed CDT and EPSRC for funding the research undertaken during this project.

The research reported within this thesis would not have possible without the help of the departmental technical staff, in particular Mr Peter Stokes (MS) and Dr Eric Hughes (NMR). Thanks also go to Dr Carissa Lloyd and the Cobb lab for their aid with peptide synthesis, and to Miss Katherine Deck for her help with peptide purification by HPLC.

I would like to thank all members of the Hodgson group, past and present, but in particular Dr Carlotta Pagli and Dr Harriet Chidwick for welcoming me into the lab and helping make the early, COVID-19 impacted years enjoyable. Thanks also go to Mr Matthew Waterson for running some high-field NMR experiments for me during his booked NMR slot.

Thanks go to all of the friends I have been fortunate enough to make in my time at Durham for helping me navigate four particularly difficult years. Special thanks go to Izzy Cormack, Helen Sims and Abbey Butler for all of the support, coffee breaks, and venting sessions.

I must also thank my friends from my undergraduate degree at Heriot-Watt. You have all given me the best distractions (even if you thought I was crazy to do a PhD while you all got jobs). Again, special thanks must go to Taylor and Sarah George for absolutely everything – all of the trips up to yours, cuddles with Pepper and Echo, and far too many in-depth Marvel chats to count.

I'd like to thank my family – Mum, my sister Rachel, her husband Pete, their amazing kids Thomas and Jack, my brother Andrew, his wife Sophie, and their adorable triplets Poppy, Ziggy, and Bodhi. I can't put into words how much I have appreciated everything over the past four years. I don't know if I could have done this without you.

Most importantly, Dad. I don't even know what to say. Except I miss you more every day.

MEMORANDUM

The work that is presented in this thesis was carried out at Durham University between October 2020 and September 2024. The thesis is the work of the author, except where acknowledged by reference, and has not been submitted for any other degree. The copyright of this thesis lies solely with the author and no quotation from it should be published without written consent and information derived from it should be acknowledged.

This work has been presented at:

- 2nd Annual MoSMed Conference, December 2020, Durham University, Durham, UK – *oral presentation (online due to COVID-19 restrictions)*
- 3rd Annual MoSMed Conference, July 2022, Durham University, Durham, UK – *poster*
- 4th Annual MoSMed Conference, May 2023, Newcastle University, Newcastle, UK – *oral presentation*
- Durham Chemistry Annual Postgraduate Research Symposium, June 2023, Durham University, Durham, UK – *poster*
- RSC Chemical Biology Community Meeting, September 2023, Leeds University, Leeds, UK – *poster*
- 13th International Peptide Symposium, October 2023, Brisbane Convention & Exhibition Centre, Brisbane, Australia – *poster*
- Durham Chemistry Annual Postgraduate Research Symposium, June 2024, Durham University, Durham, UK – *oral presentation*

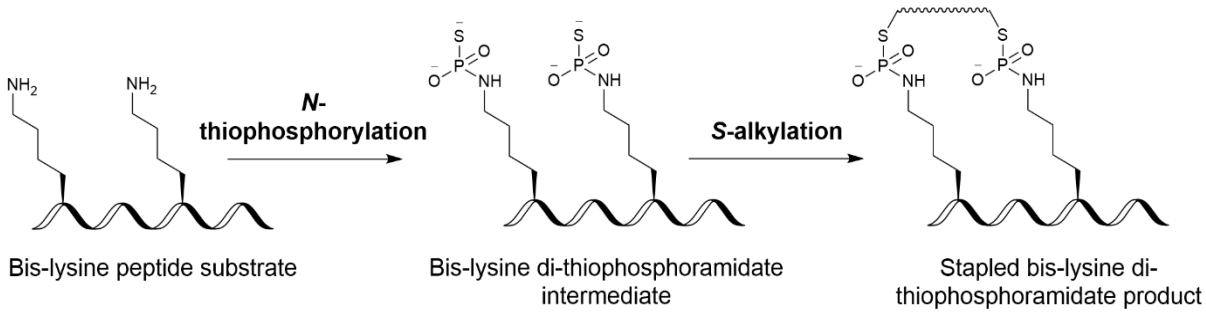
STATEMENT OF COPYRIGHT

FOR DAD

TABLE OF CONTENTS

ABSTRACT

Through the adaptation of 'click'-like aqueous *N*-thiophosphorylation chemistry established previously by the Hodgson group, we were able to staple a *bis*-lysine model peptide substrate in a novel one-pot, two-step stapling methodology *via* a di-thiophosphoramidate intermediate.



In order to reach the goal of obtaining a peptide stapled *via* a di-thiophosphoramidate intermediate, work first involved the optimisation of aqueous *N*-thiophosphorylation of a chromophore-labelled lysine residue model and a lysine-containing chromophore-labelled tetrapeptide model using thiophosphoryl chloride as the thiophosphorylation agent and $NaOH_{(aq)}$ as the base. Chromophore labelling allowed for the development of LC-MS DAD-based analytical assays to determine conversion from amine substrate to thiophosphoramidate product.

A study on the selectivity of the *N*-thiophosphorylation reaction was conducted on a series of unprotected tetrapeptides, designed with different *N*-terminal amino acids, following optimisation of aqueous *N*-thiophosphorylation. Through the implementation of a ^{31}P NMR spectroscopy-based assay, it was found that aqueous *N*-thiophosphorylation is selective for thiophosphorylation at an ε-NH_2 rather than an α-NH_2 group. It was also noted that the size of the *N*-terminal amino acid side chain had a greater impact on thiophosphorylation selectivity than the pK_{aH} of the tetrapeptide *N*-terminus.

Brief studies were also conducted on the use of an alternative thiophosphorylating agent (potassium thiophosphorodichloridate) and/ or an alternative base (triethylammonium bicarbonate solution).

Finally, a *bis*-lysine-containing peptide was designed and synthesised for stapling experiments. Through the use of corroborative LC-MS DAD and ^{1}H-^{31}P HMBC NMR spectroscopy analyses, peptide stapling was achieved using a one-pot, two-step procedure *via* a di-thiophosphoramidate intermediate using cheap and readily available reagents.

ABBREVIATIONS

A	5’-amino-5’-deoxyadenosine
Ac	Acetyl
Aib	2-aminoisobutyric acid
AFPS	Automated fast-flow peptide synthesis
AIBN	Azobisisobutyronitrile
Ala (A)	Alanine
Alloc	Allyloxycarbonyl
Arg (R)	Arginine
Asp (D)	Aspartic acid
BID BH3	BH3-interacting domain death agonist protein
Boc	tert-butyloxycarbonyl
C	5’-amino-5’-deoxycytidine
CAMP	Cationic antimicrobial peptide
CD	Circular dichroism
CLipPA	Cysteine lipidation on a peptide or amino acid
CuAAC	Copper catalysed azide-alkyne cycloaddition
Cys (C)	Cysteine
DAD	Diode array detection
DIC	*N,N*-diisopropylcarbodiimide
DIPEA	*N,N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformadime
DMPA	2,2-dimethoxy-2-phenylacetophenone
DNA	Deoxyribonucleic acid
EIC	Extracted ion count
eq	Equivalent
ESI	Electrospray ionisation
Et_3N	Triethylamine
Fmoc	9-fluorenylmethyloxycarbonyl
G	5’-amino-5’-deoxyguanosine
GLP-1	Glucagon-like peptide-1

Glu (E)	Glutamic acid
Gly (G)	Glycine
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High-performance liquid chromatography
KCl	Potassium chloride
$KPSOCl_2$	Potassium thiophosphorodichloridate
KTPA	Potassium thiophosphoramidate
LC	Liquid chromatography
LC-MS	Liquid chromatography-mass spectrometry
LC-MSMS	Liquid chromatography-tandem mass spectrometry
Lys (K)	Lysine
MeCN	Acetonitrile
MeI	Methyl iodide
MeOH	Methanol
Met (M)	Methionine
MS	Mass spectrometry
MSMS	Tandem mass spectrometry
NaOH	Sodium hydroxide
NCL	Native chemical ligation
Nle	Norleucine
NMR	Nuclear Magnetic Resonance
PDB	Protein Data Bank
PEG	Polyethylene glycol
Phe (F)	Phenylalanine
PPI	Protein-protein interaction
Pra	Propargylglycine
$POCl_3$	Phosphoryl chloride
$PSCl_3$	Thiophosphoryl chloride
$PSOCl_2^-$	Thiophosphorodichloridate ion
PyBOP	Benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
RCM	Ring-closing metathesis
S-Alk	*S*-alkylation

SAHB	Stabilised α-helix of BCL-2 domains
SMC	Suzuki-Miyaura cross-coupling
SPPS	Solid phase peptide synthesis
T	3’-amino-3’-deoxythymidine
TBAH	Tetrabutylammonium hydroxide
TCEP	Tris(2-carboxyethyl)phosphine
TEAB	Triethylammonium bicarbonate
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane
TP	Thiophosphorylation
TRAIL	Tumour necrosis factor-related apoptosis-inducing ligand
U	5’-amino-5’-deoxyuridine
UV	Ultraviolet
VLM	Visible-light-mediated

1. Introduction

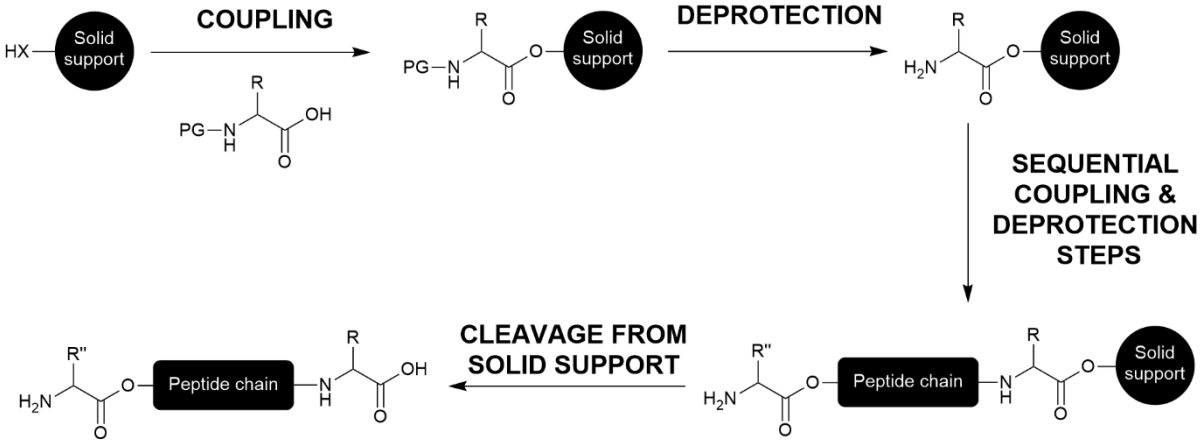
This chapter will provide a brief background on the use of proteins and peptides in medicine, before moving on to discuss limitations associated with peptide pharmaceuticals, and the various chemical modification methods used to improve the efficacy of peptide pharmaceuticals *in vivo*. Particular attention will be paid to peptide stapling as a method for improving the stability and biological activity of peptides. Finally, before outlining the aims and objectives of this project, thiophosphorylation of biologically relevant compounds will be discussed.

1.1. A brief history of peptides

Prior to discussing the chemistry associated with peptide therapeutics and the methodologies currently being researched to improve the scope of peptides as therapeutic agents, we will look at the history of peptide synthesis by chemical means. The current state of the global peptide therapeutic market will also be briefly explored.

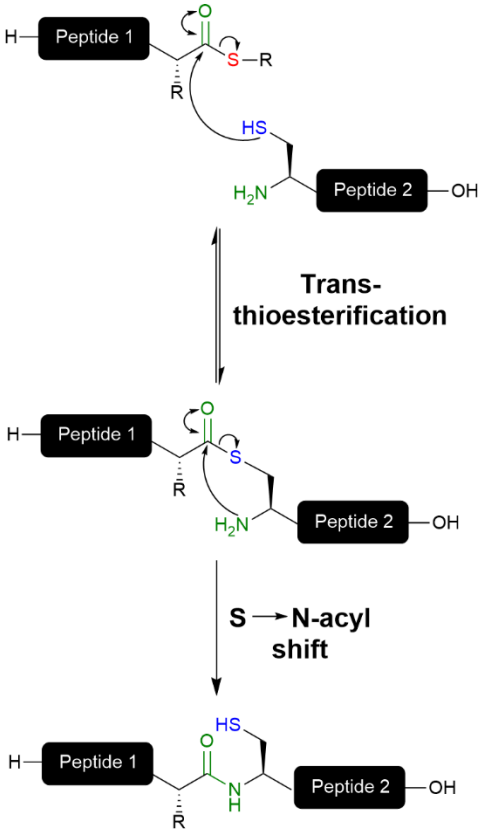
1.1.1. Development of chemical peptide synthesis

Although the first reported synthesis of an *N*-protected dipeptide was in 1881 (benzoylglycylglycine by Theodor Curtius)[1] and the first reported synthesis of an unprotected dipeptide was in 1901 (glycylglycine by Emil Fischer and Ernest Fourneau)[2], the field of chemical peptide synthesis remained relatively quiet until the 1960s. The advent of solid phase peptide synthesis (SPPS) in 1963 by Bruce Merrifield,[3] in which protected amino acids are individually coupled in a stepwise fashion to a growing peptide chain covalently bound to a solid support, resulted in a growth of research into peptide chemistry. While recombinant expression methods are in use for the synthesis of some peptide-based medications (e.g. insulin), the development of SPPS has allowed for the incorporation of non-proteinogenic amino acids in peptide-based medications (e.g. GLP-1 analogue semaglutide[4]) and the synthesis of peptides consisting of entirely D-amino acids (e.g. RE-E1P47 peptide investigated in pre-clinical testing as an HIV-1 inhibitor by Gomara *et al.*[5]). Scheme 1.1 depicts a schematic overview of how SPPS is used to synthesise peptides *via* chemical means.



Scheme 1.1 Schematic overview of solid phase peptide synthesis (SPPS). PG – protecting group; R – amino acid side chain.

While the development of SPPS represented a huge leap forward in the field of peptide chemistry, it is not without its limitations. Most notably, SPPS is often cited as being limited to the synthesis of peptides of a maximum of 40 – 50 amino acids in length.[6] In 1994, native chemical ligation (NCL) emerged as an efficient method for linking together unprotected peptide fragments, synthesised through SPPS, to generate peptide sequences with >50 amino acid residues.[6, 7] Since its development, NCL has been expanded upon,[8-10] and has been used extensively for the chemical synthesis of large proteins.[11-14] Scheme 1.2 illustrates the basic mechanism of NCL.



Scheme 1.2 Mechanism for native chemical ligation (NCL).

Using both SPPS and NCL, the total chemical synthesis of the 82-residue anti-thrombotic mosquito protein anophelin from the species *Anopheles gambiae* was achieved in 2018.[6, 15] A new method, referred to as automated fast-flow peptide synthesis (AFPS), for the rapid chemical synthesis of peptides up to 164 amino acid residues in length was reported by Pentelute *et al*. in 2020.[16]

The development of both SPPS and NCL have led to an explosion in research associated with the peptide-based pharmaceutical industry as the costs associated with peptide chemistry have fallen and the emphasis within pharmaceutical research circles has shifted away from drug activity towards drug specificity.[17] A brief look into some examples of peptide/ protein-based medications follows, before a more in depth look into some of the chemical advancements that have played parts in the rise of the peptide-based pharmaceutical industry.

1.1.2. Peptides as medications

While there has been a significant increase in interest in peptide pharmaceuticals over the last 30 years, the history of using proteins as medication dates back to the 1920s with the discovery of insulin by Frederick Banting in 1921 and the first administration to a human subject in January 1922.[18, 19] The insulin prepared for this first human trial was obtained from the pancreases of calf foetuses *via* alcohol extraction.[19] Following this, insulin was primarily obtained from bovine and porcine sources before

becoming the first licensed peptide-based medication produced through recombinant DNA technology by Genentech and Eli Lilly in 1982.[20, 21]

The first synthetic peptide drug made available in clinical practice was oxytocin in 1962 following its first reported chemical synthesis in 1953 by du Vigneaud *et al*. (*via* solution phase methods).[22-25] From 1962, a further 37 peptide-based medications where then approved in the 20th century.[26] Another 36 peptide drugs have been approved since 2000 (as of 2021),[27] with a further 90 peptide-based therapeutics undergoing clinical trials (as of 2024).[28] Figure 1.1 shows the current availability and primary therapeutic areas of the 74 peptide-based medications approved since 1962.

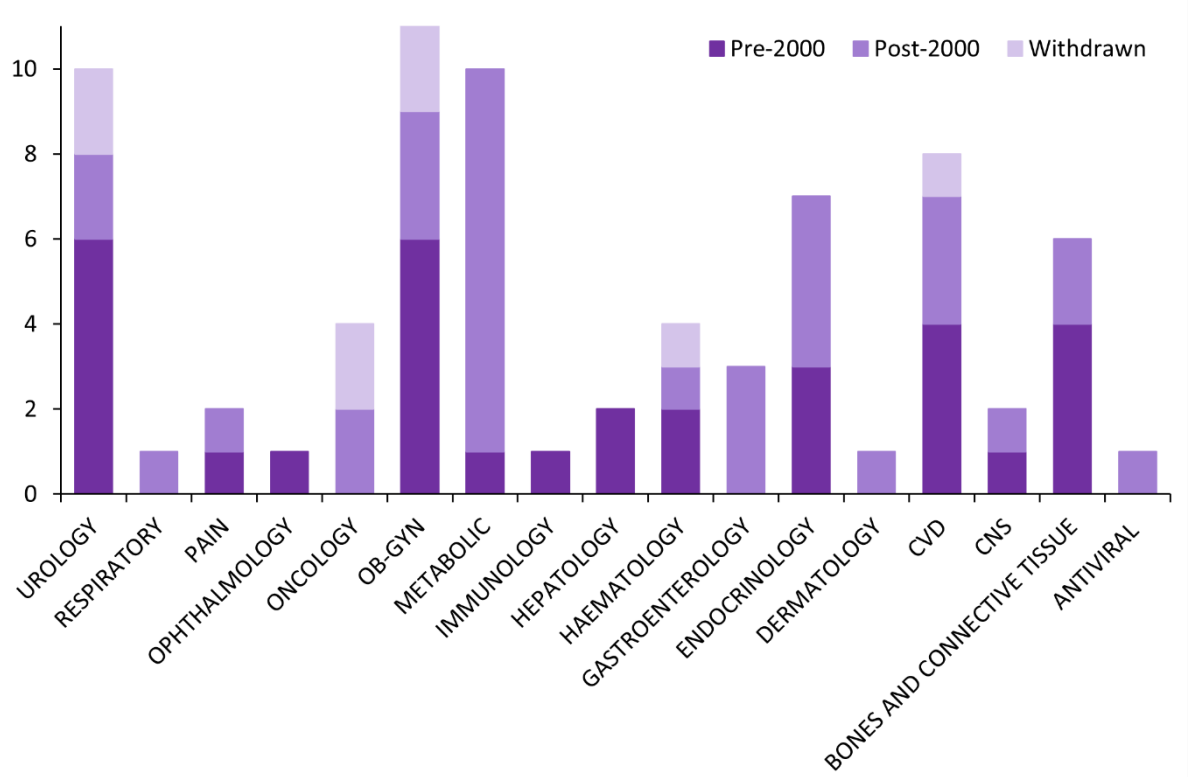


Figure 1.1 Breakdown of the primary-therapeutic areas of the 74 peptide-based drugs approved from 1962 onwards.[26, 27] Pre-2000 refers to the peptide drugs approved before 2000; post-2000 refers to the peptide drugs approved after 2000; withdrawn refers to the peptide drugs that were approved and have seen been withdrawn from the market.

Based on Figure 1.1, the largest areas in peptide drug development have been obstetrics and gynaecology, urology, and metabolic disorders. Research into metabolic disorders appears to have had the most significant growth since 2000, and this is reflected in the market value of the peptide pharmaceutical industry. In 2023, the peptide therapeutics market was valued at $45.67 billion (USD) with 22.9% of the market share accounted for by peptide therapeutics targeting metabolic disorders.[29] A significant contributor to the growth of the peptide therapeutic market with respect to metabolic disorders is semaglutide, a glucagon-like peptide-1 (GLP-1) analogue licensed in the EU for type 2

diabetes treatment in 2018, that has gained popularity as a weight loss medication.[30-33] Originally intended for use as a medication for type 2 diabetes (sold under the trade name Ozempic™), it has seen a 5000% rise in prescriptions in the US alone from 2018 to 2023 following its licensing as a weight loss medication (sold under the trade name Wegovy™).[34]

While there has been a notable uptick in research regarding peptide-based pharmaceuticals, and major advancements in the realm of peptide chemistry (Figure 1.2 summarises those briefly discussed here), there are still significant obstacles that need to be overcome before peptide-based pharmaceuticals become more commonplace. The following sections will outline some of the advantages of peptide-based drugs, some of their drawbacks, and the synthetic methods developed to lessen the impact of these drawbacks.

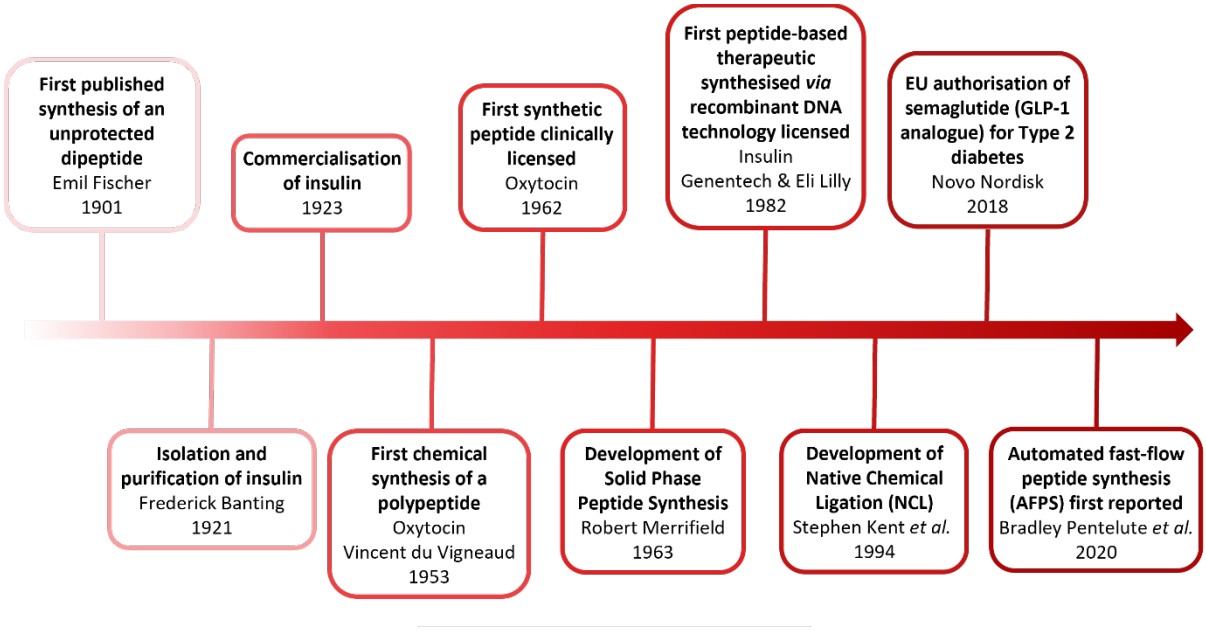


Figure 1.2 Summary timeline of various advancements in peptide chemistry and peptide-based pharmaceutics.

1.2. Peptide modifications to improve *in vivo* efficacy

Although the first protein-based medication, insulin, was commercialised in 1923,[20] there has been a lack of research with respect to peptide-based therapeutics when compared to their small drug molecule counterparts. The majority of drug development during the 20th century involved the synthesis of drug leads based on small molecule natural products (typically < 500 Da). While the low molecular weight of these small molecule drugs has allowed for oral administration, it also leads to reduced target specificity *in vivo*, resulting in potentially adverse side-effects.[17] Another significant disadvantage to small drug molecules is the potential for the evolution of toxic metabolites *in vivo*.[35]

Over the last couple of decades, however, there has been an increasing emphasis placed on drug specificity rather than activity,[17] with peptide-based drugs emerging as an area of particular interest within pharmaceutical research.

There are myriad advantages to peptide-based drugs over their small molecule counterparts, including their potential for targeting protein-protein interactions (PPIs) which have previously been considered 'undruggable' due to their large surface areas (PPI interfaces are ca. 1500-3000 $Å^2$ compared to the average protein-small drug molecule interface of ca. 300-1500 $Å^2$) and shallow interaction sites.[17, 36, 37] Due to their much larger size, peptides are able to bind their targets with much greater selectivity and specificity than small drug molecules, leading to a reduction in potentially adverse side-effects associated with off-target binding.[17, 38, 39] With respect to their metabolism *in vivo*, peptides degrade to generate non-toxic amino acids, and the short serum half-life of peptides minimises the risk of accumulation in tissues.[36, 38, 40, 41]

Despite these advantages, peptides were long considered to be poor drug candidates due to low oral bioavailability, poor resistance to proteases, rapid clearance by the kidneys and liver, and limited cell membrane permeability, as well as high manufacturing costs.[17, 26, 36, 38, 40-43] In an effort to combat some of the drawbacks associated with peptide-based therapeutics, an array of chemical modifications has been developed, some of which will be discussed now.

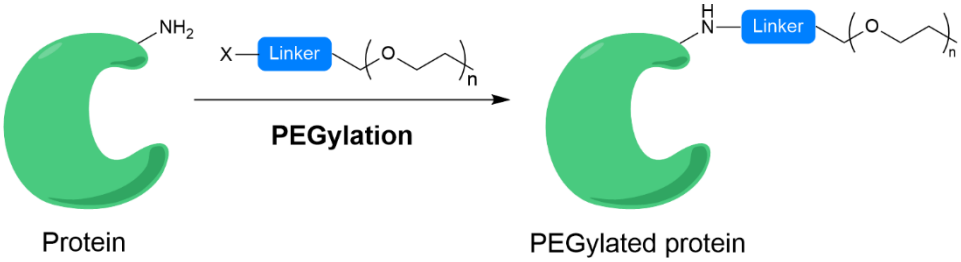
1.2.1. Chemical modifications used to improve peptide drug efficacy *in vivo*

Of the chemical modifications developed to improve the efficacy of peptide-based therapeutics *in vivo*, PEGylation, lipidation, and peptide cyclisation/ stapling have emerged as the key methods. PEGylation and lipidation will be briefly described in the following two sub-sections, while peptide stapling will be discussed in more detail in section 1.3.

1.2.1.1. PEGylation of peptides

The inclusion of a polyethylene glycol (PEG) group (known as PEGylation) is used to improve the *in vivo* stability and serum half-life of a peptide, as well as decrease the immunogenicity of the protein-based drug.[44, 45] PEG is a polymeric compound and is non-toxic, soluble in water, and non-immunogenic making it a suitable modifier for pharmaceuticals.[46] Since PEGylation was first used for commercial protein-based drugs in the 1990s, a total of 38 PEGylated pharmaceuticals have been approved by the FDA, with a further 82 active clinical trials running using PEGylated pharmaceuticals (as of May 2023).[47, 48] Of these 38 pharmaceuticals, three are PEGylated peptide-based medications. The first of these PEGylated peptide drugs to be approved was peginesatide (trade name – Omontys™), an erythropoiesis stimulator, in 2012.[48-51] The PEGylated peptide drug pegcetacoplan

(sold under trade names Empaveli™ and Syfovre™) was first approved in 2021 as an immunosuppressant for the treatment of paroxysmal nocturnal haemoglobinuria.[48, 52-54] Scheme 1.3 shows a schematic depiction of PEGylation on a peptide substrate.



Scheme 1.3 Schematic illustration of one-step protein PEGylation.

Various methods of peptide PEGylation have been explored, using both canonical and non-canonical amino acids. The canonical amino acids typically targeted for PEGylation are nucleophilic residues – lysine (at the ε-NH_2) and cysteine (at the thiol). However, PEGylation at these two nucleophilic residues is rarely site-specific as several of these residues are often present on a single protein.[47] A variety of linker groups can be used for the addition of a PEG group to a protein, depending on whether the target residue is a lysine or a cysteine. The structures of some commonly used linker groups are illustrated in Figure 1.3.

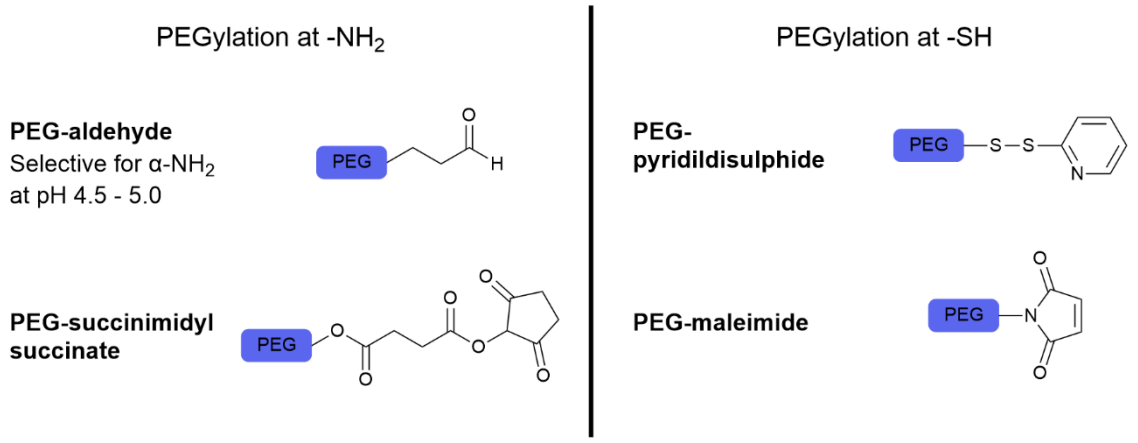


Figure 1.3 Structures of commonly used PEG linkers reactive to amino and thiol groups.[46]

The selectivity of protein PEGylation can, however, be enhanced through site-directed mutagenesis of a surface residue.[47] In 2013, a study was published in which tumour necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) underwent site-specific PEGylation at a mutated residue (N109C mutation) to improve the *in vivo* stability of TRAIL and increase the length of time TRAIL remained *in vivo* prior to clearance. When tested in mouse models, PEGylated N109C TRAIL was noted to

considerably inhibit tumour growth compared to the wildtype TRAIL protein, thus indicating that the pharmacokinetic characteristics of TRAIL had been improved following PEGylation.[45]

An alternative method for obtaining site-specific PEGylation of a protein/ peptide is the incorporation of noncanonical amino acids. A popular choice for these noncanonical amino acids are those containing an azide moiety, which can be specifically targeted *via* copper catalysed azide-alkyne cycloaddition (CuAAC) chemistry.[47] Using alkyne-functionalised PEG reagents, proteins containing a single (surface exposed) azide-functionalised amino acid can undergo site-specific PEGylation with up to 85% conversion.[55]

1.2.1.2. Lipidation of peptides

Lipidation of a protein/ peptide is typically done in an effort to extend the serum half-life of a protein *in vivo*, and to improve cell membrane permeability.[56] As of 2019, a total of seven lipidated peptide/ protein-based therapeutics had been approved for use, four of which are used in the treatment of metabolic disorders.[56-64] A particularly notable example of a lipidated peptide medication is semaglutide, previously mentioned in section 1.1.2. As previously stated, semaglutide is a GLP-1 analogue used in the treatment of type 2 diabetes and obesity.[30] Figure 1.4 compares the amino acid sequences of semaglutide and native human GLP-1.

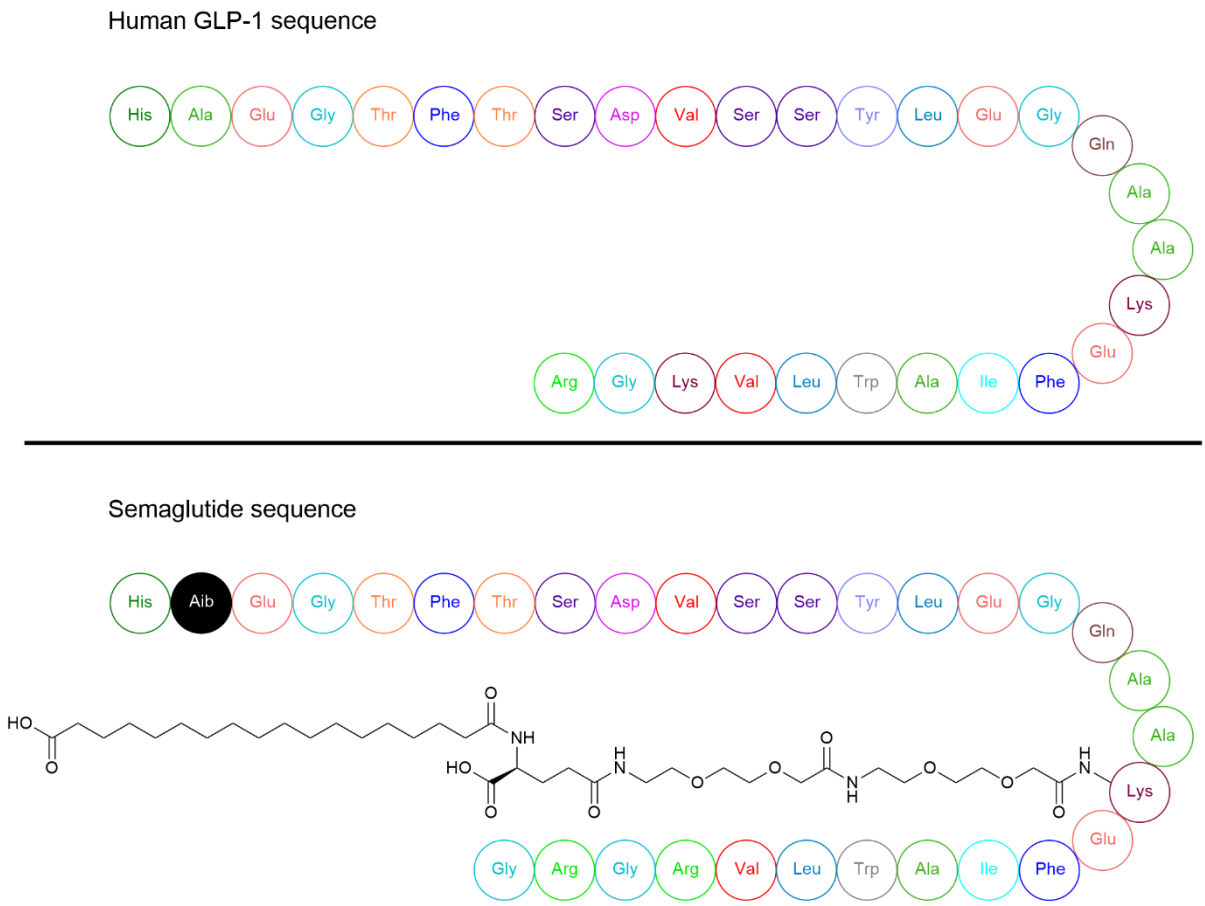


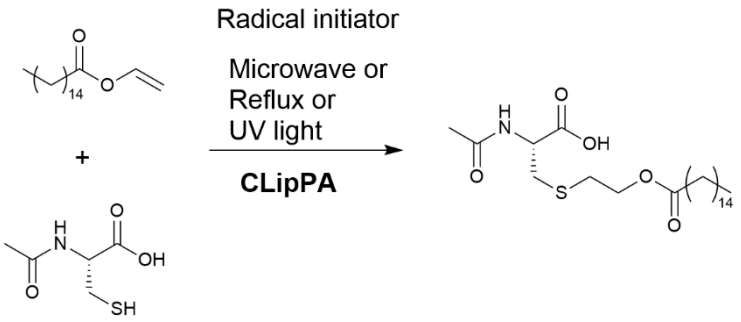
Figure 1.4 Comparison of the amino acid sequences of human GLP-1 and semaglutide. Aib – 2-aminoisobutyric acid.

Semaglutide is 94% homologous to human GLP-1, with two amino acid substitutions (Ala to Aib [increased enzymatic stabilisation *in vivo*] and Lys to Arg [to improve selectivity of lysine derivatisation]) and the addition of a hydrophilic spacer and fatty acid chain on the single remaining lysine residue.[4, 65] A comparison of semaglutide and native GLP-1 shows an over 100-fold increase in the *in vivo* half-life of the peptide (168 h vs. 1.5 h respectively).[56] The significant improvement in the *in vivo* half-life in humans is attributed to the enhanced albumin binding, achieved through the addition of the C_{18} fatty diacid chain on the lysine residue of semaglutide.[4]

While a wide range of methodologies exists for the lipidation of proteins and peptides including serine and lysine fatty acylation (as in the case of semaglutide),[66, 67] cysteine prenylation,[68-70] and *N*-terminal glycine myristoylation,[71-75] a synthetic method of cysteine palmitoylation will be described as an example.

In 2016 the Brimble group described the use of thiol-ene chemistry for lipidation of α-*N*-protected cysteine residues, in a method known as Cysteine Lipidation on a Peptide or Amino acid (CLipPA).[76]

Scheme 1.4 displays a schematic representation of CLipPA chemistry performed on an *N*-acetylated cysteine residue.



Scheme 1.4 Schematic representation of Cysteine Lipidation on a Peptide or Amino Acid (CLipPA).[76]

It was found that, using microwave irradiation and azobisisobutyronitrile (AIBN) as the radical initiator, quantitative conversion of Ac-Cys-OH to the *S*-palmitoylated derivative was achieved. The *S*-palmitoylated *N*-acetylated cysteine residue was then coupled to a resin-bound peptide model synthesised using standard Fmoc-SPPS methods, with the coupling reaction going to completion. Finally, it was demonstrated that CLipPA chemistry could be used for the direct conjugation of a palmitoyl group to a cysteine thiol in a semi-protected peptide.[76]

Having briefly considered PEGylation and lipidation as modifications used for the improvement of peptide and protein pharmacokinetic characteristics *in vivo*, we will now turn our attention to peptide stapling. The focus will be on the different methods of peptide stapling currently available, the use of canonical and noncanonical amino acids, and the reported conversions of each method discussed.

1.3. Peptide stapling

The first example of peptide stapling, an all-hydrocarbon staple, was reported in 2000 by the Verdine group.[77] Since then, myriad stapling methods have been developed using both unnatural and natural amino acids as the anchoring residues. As with PEGylation and lipidation, peptide stapling is a modification implemented in an effort to improve the *in vivo* pharmacokinetic characteristics of peptides believed to be of therapeutic benefit.

As mentioned in section 1.2, peptide therapeutics are of particular interest due to their ability to target PPI interfaces.[17] Based on entries in the Protein Data Bank (PDB), 62% of PPIs involve an α-helix at the interface.[78] Stapling aims to stabilise the α-helical structure of a peptide in the absence of the complete protein fold, and thus increase the affinity of the peptide to its target, resulting in a more potent therapeutic agent compared to the linear peptide.[79] Peptide stapling has also been reported as reducing proteolytic degradation of peptides by ca. 40-fold in *in vitro* trypsin digestion studies. As

proteases bind to peptide substrates in their extended conformation, the induction/ stabilisation of an α-helical conformation through stapling limits degradation by protease activity.[77, 80, 81] Finally, depending on the structure of the introduced linker, stapled peptides may also have increased cellular membrane permeability compared to their linear counterparts.[79, 82]

Before exploring the various different methods available to staple a peptide, the importance of staple design will be discussed.

1.3.1. Staple design

When it comes to designing a peptide staple, the staple position and length need to be considered. Depending on the desired extent of α-helix stabilisation, different anchoring residue intervals can be selected – shown in Figure 1.5. The longer the desired helix, the larger the interval between anchoring residues. In each case, the residues involved in stapling are found on the same side of the helix.[83]

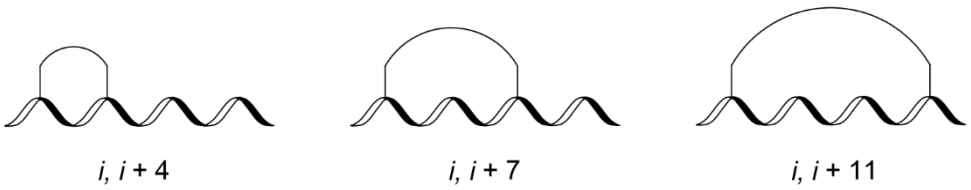


Figure 1.5 Representation of different stapling positions along a peptide.

One of the most important things to consider for the position of a peptide staple is to ensure that the anchor residues do not replace any residues critical for the interaction between the peptide therapeutic and its target. Should any critical binding residues be replaced, the binding affinity of the peptide would be reduced.[84-86] Staple position can also improve peptide binding affinity. Hydrocarbon-based staples positioned to flank the peptide binding interface have been reported to improve binding affinity through additional hydrophobic contacts between the hydrocarbon staple and hydrophobic protein regions adjacent to the target binding sites.[83, 87] The positioning of stapling amino acids can even impart protease resistance, by placing stapling residues adjacent to, or spanning over, a cleavage site.[80]

The length of the staple can have a significant impact on the stability of the α-helix. If the length of the intended staple is too short the staple will not form, resulting in minimal/ no helix stabilisation. If the staple is too long, however, it may be too flexible to impart enough stabilisation for an α-helix to form. Some work has been done with respect to staple length optimisation, most notably with all-hydrocarbon staples. When designing an all-hydrocarbon staple, an eight-carbon linker is optimal for an *i, i* + 4 staple and an 11-carbon linker is best for an *i, i* + 7 staple. The optimal linker length has not

been determined for an *i, i* + 11 all-hydrocarbon staple as these staples may be too flexible to stabilise an α-helix.[83]

Having considered the importance of staple design with respect to positioning and staple length, the focus can shift towards the various methods available for peptide stapling. Our discussion will separate these stapling methods by whether the staple is based around the use of noncanonical or canonical amino acids as anchor residues. The synthesis and physical impacts of each staple type will be explored, as well as the efficiency of each method.

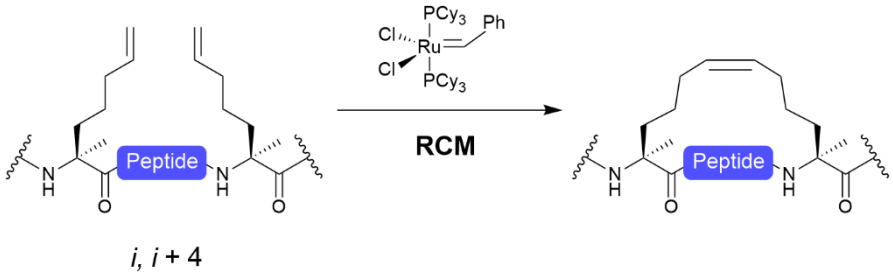
1.3.2. Noncanonical amino acid staples

The focus of this section will involve the discussion of a handful of different stapling procedures requiring the use of noncanonical amino acid(s).

The use of noncanonical residues can incur significant material costs to the research conducted, as well as the pharmaceutical created should it be approved for clinical usage. For example, 1 g of Fmoc-Lys-OH costs £32.70 while 1 g of Fmoc-azidolysine costs £691.00 (as of 2024).[88, 89] In some cases, the desired functionality may require modification of a canonical amino acid prior to incorporation into the peptide, thus resulting in a greater number of materials needed and waste generated overall. Many of these methods also require expensive catalysts to proceed.[77, 90, 91] That being said, stapling procedures based around the use of noncanonical amino acids are, often, site-specific and can proceed with high conversions to the stapled product.

1.3.2.1. All-hydrocarbon staples

As stated at the beginning of section 1.3, the first reported peptide staple was an all-hydrocarbon-based staple in 2000 by the Verdine group. Through the inclusion of two non-natural α,α-disubstituted amino acids bearing olefinic side chains of a variety of lengths, at either *i, i* + 4 or *i* +7 on the peptide chain, cyclisation was achieved through ring-closing metathesis (RCM), conducted on-resin immediately following SPPS, using a first-generation Grubbs catalyst, as shown in Scheme 1.5.[77, 92, 93]



Scheme 1.5 All-hydrocarbon staple formation through ring-closing metathesis (RCM) using a first-generation Grubbs catalyst.[77]

From the outset, peptide stapling with a hydrocarbon cross-link was found to be an extremely efficient synthesis method. Typically, >90% conversion to the stapled peptide was obtained after 2 h of metathesis, with the completion of metathesis after a further 2 h with fresh catalyst.[92] For cross-links formed between residues at *i, i* + 4, eight-carbon cross-links formed with >98% conversion from the linear peptide. For cross-links formed between residues at *i, i* + 7, both 11- and 12-carbon cross-links were formed with >98% conversion from the linear peptide. With respect to helix stabilisation, it was found that *i, i* + 4 staples, regardless of staple length, did not appear to improve the helical content of the stapled peptide compared to its unmodified control. As for the *i, i* + 7 staples, the 11-carbon cross-link resulted in the greatest improvement in helical content with a 26% increase compared to the unmodified control peptide.[77]

Finally, the Verdine group investigated whether the addition of a staple impacted the proteolytic stability of the peptide. It was found that the 11-carbon cross-link formed between residues positioned at *i, i* + 7 gave a 41-fold increase in the proteolytic stability of the peptide compared to the unmodified control. In fact, just the inclusion of the two α,α-disubstituted amino acids without metathesis lead to a 5-fold improvement in proteolytic stability compared to the control peptide.[77]

In 2004, Walensky *et al*. reported on the use of a hydrocarbon stapled peptide, based on an amphipathic α-helical BH3 segment (a necessary death domain), as a potential pro-apoptotic agent for leukaemia treatment. The α-helical BH3 segment mediates protein interactions between members of the BCL-2 protein family, which are critical in apoptosis control.[94, 95]

Analyses performed on a range of stapled BH3-based peptides, referred to as 'stabilised α-helix of BCL-2 domains' (SAHBs), indicated that the α-helical content, protease resistance, and cell membrane permeability were improved following stapling compared to the unmodified control peptide, with the lead peptide labelled as $SAHB_A$ (structure of $SAHB_A$ compared to unmodified peptide in Figure 1.6).[94]

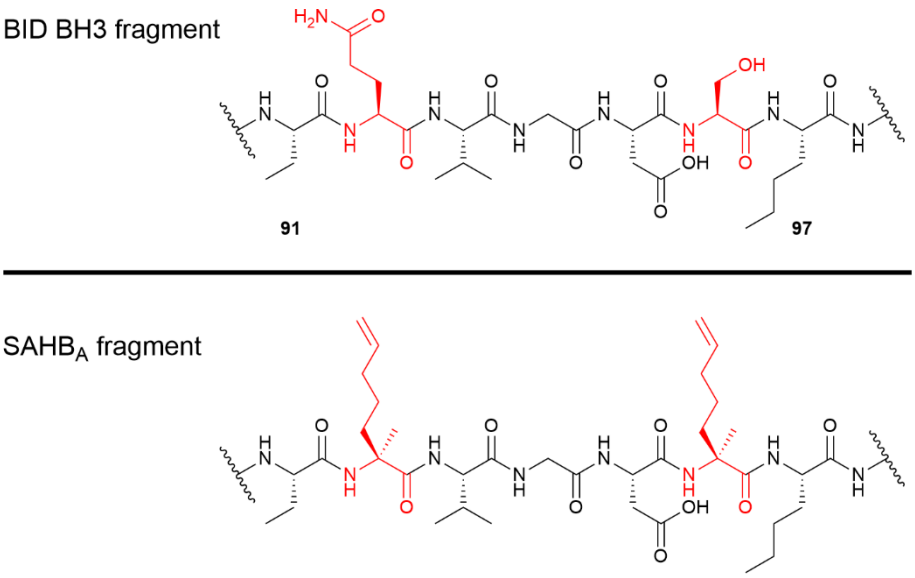
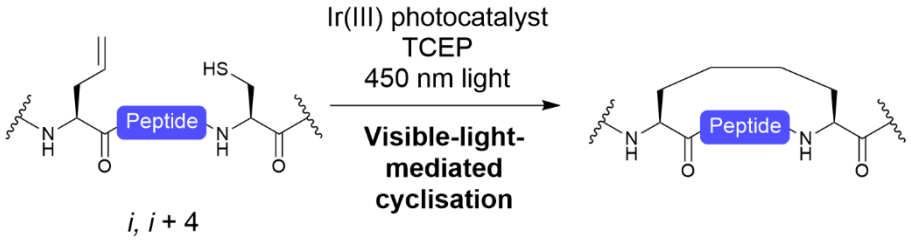


Figure 1.6 Fragment structures showing the modified residues of the unmodified control peptide (BID BH3 fragment) and SAHB$_A$ fragment. Red residues signify those replaced (BID BH3 fragment) with α,α-disubstituted olefinic residues (SAHB$_A$ fragment).[94]

A fluorescence polarisation assay also showed a 6-fold enhancement of peptide binding affinity for SAHB$_A$ compared to the unmodified peptide. Finally, an *in vivo* mouse trial was conducted to establish whether SAHB$_A$ repressed leukaemia growth. While control mice showed progressive leukaemic growth over the 5-day experimental period, the mouse cohort treated with SAHB$_A$ demonstrated suppressed tumour growth after 3 days and tumour reduction by day five.[94]

In all, hydrocarbon-based staples synthesised between two α,α-disubstituted amino acids not only constituted the first reported examples of peptide staples but also provided significant improvements in the pharmacokinetic properties of the modified peptides. The synthetic procedures gave >90% conversion after 2 h and were site-specific due to the inclusion of non-natural amino acids within the peptide chains.

More recently, in 2024, a new method for the chemoselective synthesis of a hydrocarbon staple was reported, based on intramolecular interception of visible-light-mediated (VLM) cysteine desulfurisation. The noncanonical (but commercially available) amino acid, Fmoc-allyl-Gly-OH (alGly), and a cysteine residue were incorporated into various model peptides at the *i, i* + 4 positions. The cyclisation reactions were performed on unprotected model peptides in aqueous solution, rather than on-resin, in the presence of an Ir(III) photocatalyst and tris(2-carboxyethyl)phosphine (TCEP) under irradiation with blue light (450 nm).[90]



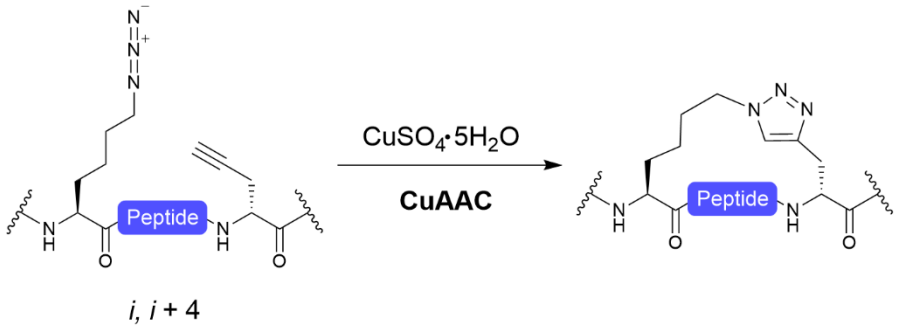
Scheme 1.6 Hydrocarbon staple formation through visible-light-mediated (VLM) cyclisation.[90]

The VLM reaction was reported as going to completion, through monitoring by analytical high-performance liquid chromatography (HPLC), and yields of up to 82% were obtained for *i, i* + 4 staples.[90]

1.3.2.2. Triazole-based stapling

The chemistry implemented in the synthesis of triazole-based staples was originally reported by Sharpless *et al*. in 2002, and involves the use of a Cu^{II}/ ascorbate aqueous system for 1,4-regioselective cycloaddition between an azide and terminal alkyne.[96]

The triazole staple was first described by Chorev *et al*. in 2008, utilising CuAAC or 'click' chemistry between an ε-azido lysine residue and propargylglycine (Pra) positioned at *i, i* + 4 relative to each other on the peptide chain. The introduction of the azido-functionalised residue could be done through two methods: (1) direct introduction during peptide synthesis; or (2) selective diazo transformation of a lysine residue on-resin.[97]



Scheme 1.7 Triazole staple formed between an ε-azido lysine residue and D-propargylglycine (D-Pra) through copper catalysed azide-alkyne cycloaddition (CuAAC).[98]

Unlike the hydrocarbon staple discussed in section 1.3.2.1, the CuAAC reaction is done following cleavage of the linear peptide from the resin. While the peptide was resin bound, Chorev *et al*. reported that all attempts to carry out CuAAC were unsuccessful. After the final side chain deprotection and cleavage from the resin, however, complete conversion was achieved in solution, with a final yield of 55% after purification by HPLC.[97]

Circular dichroism (CD) analysis performed on a variety of peptides stapled with *i, i* + 4 triazole bridges demonstrated that α-helicity is enhanced compared to both the wild-type control and the linear precursor (i.e. the modified peptide prior to undergoing the CuAAC reaction).[98, 99] The optimal triazole-based staple for α-helix stabilisation (with anchoring residues [ε-azido lysine and Pra] at *i, i* +4) was found to consist of an eight-atom linker – an example of this staple length is shown in Scheme 1.7. This particular triazole staple resulted in 90% peptide helicity while its linear parent was 45% helical under the same conditions, along with a ca. 38-fold improvement in binding affinity.[98] Various other examples of triazole staples also exist in the literature.[100-104]

The formation of double triazole stapled peptides has also been demonstrated, with a maximum reported peptide helical content of 99% (linear wild type reported as having 44% helical content). Proteolytic degradation of the double triazole stapled peptide was explored, and the stapled peptide was found to be markedly more resistant to proteolytic degradation compared to the wild type precursor.[98]

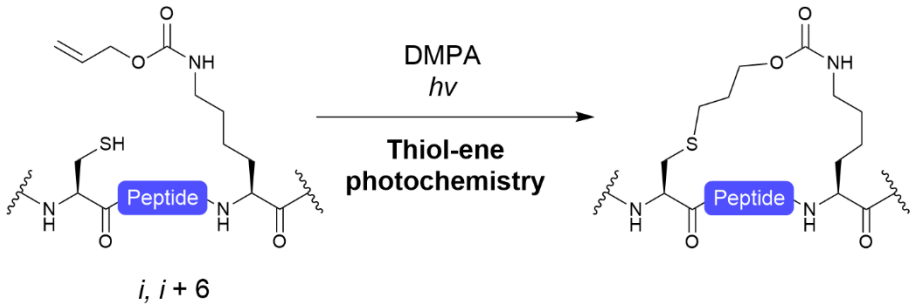
As with hydrocarbon staples, triazole-based staples have been shown to improve the α-helical stability (and content), binding affinity, and protease resistance of the modified peptides. The stapling procedure is reported to proceed to completion in solution, with product loss occurring during chromatographic purification.

1.3.2.3. Modified natural amino acids

A variety of alternative methods are available for peptide stapling based around the use of noncanonical amino acids synthesised through modification of natural amino acid residues. These alternative stapling methods include VLM thiol-ene chemistry[105] and Suzuki-Miyaura cross-coupling (SMC).[91] These two stapling methods will be briefly discussed in the following section.

1.3.2.3.1. Visible-light-mediated cyclisation

In 2010, a synthetic method for the on-resin stapling of a peptide using thiol-ene photochemistry was described. A model peptide based on Arg-Gly-Asp (RGD) was synthesised for proof-of-concept demonstration. An alkene functionality was introduced through the inclusion of a lysine residue modified with an allyloxycarbonyl (Alloc) protecting group (Lys(Alloc)). Through the inclusion of a cysteine residue and Lys(Alloc) residue at either end of the RGD motif, and introducing 2,2-dimethoxy-2-phenylacetophenone (DMPA) as the photoinitiator, peptide cyclisation was completed after 1 h of exposure to 365 nm wavelength light. Following resin cleavage and purification, 24% yield was obtained.[105]

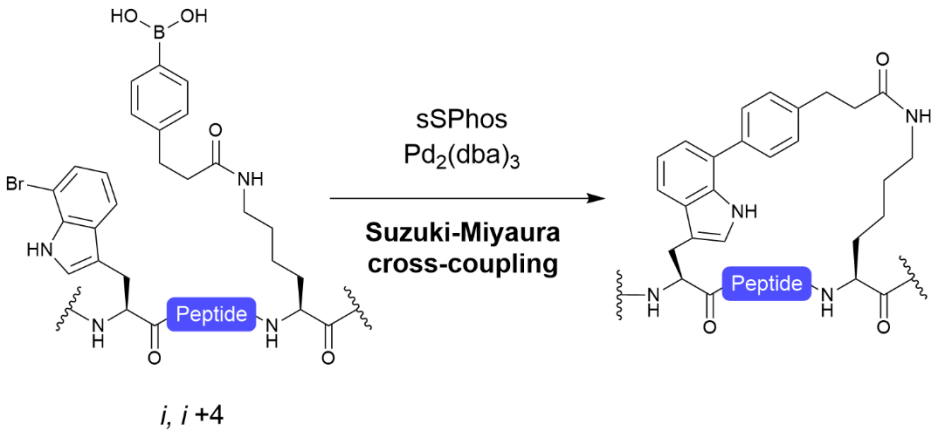


Scheme 1.8 Thiol-ene staple formed between a native cysteine residues and noncanonical Lys(Alloc) residue through thiol-ene photochemistry.[105]

A competitive binding assay was performed to ensure that this peptide staple structure would not negatively impact the ability of a peptide to bind with its protein target. The IC_{50} of the cyclised RGD-based peptide was measured, and found to be comparable with literature values, and the cyclised RGD-based peptide model was more potent than its linear parent.[105]

1.3.2.3.2. Suzuki-Miyaura cross-coupling-based cyclisation

A peptide stapling method based on Suzuki-Miyaura cross-coupling (SMC) was reported in 2022 by Sewald *et al*. This procedure involved the preparation of a brominated tryptophan residue (L-7-bromotryptophan shown in Scheme 1.9) prior to its incorporation into the peptide model SPPS. A lysine residue in the *i, i* + 4 position relative to L-7-bromotryptophan underwent selective on-resin modification with 4-(2-carboxyethyl)phenylboronic acid to generate the boronic acid motif required for SMC (lysine-based residue shown in Scheme 1.9). The cross-coupling reaction was carried out on-resin in the presence of sSPhos and $Pd_2(dba)_3$, followed by resin cleavage and purification by reverse phase HPLC. For the example shown in Scheme 1.9, the final yield was 14%.[91]



Scheme 1.9 Peptide stapling *via* Suzuki-Miyaura cross-coupling (SMC).[91]

Following stapling and isolation of the SMC stapled peptide, the α-helical content, binding affinity, and enzymatic stability of the stapled peptide were evaluated. The α-helical content of the linear precursor was 9%, while the helicity of the stapled peptide was found to be 21%. The binding affinity of the stapled peptide was shown to be five times greater than that of the linear precursor in a competitive fluorescence polarisation assay. As for the protease resistance, both the stapled peptide and its linear precursor were subjected to proteinase K digestion. While the stapled peptide was cleaved into 2 distinct fragments after 120 min, the linear precursor was cleaved into 4 fragments. The position of the peptide staple prevented enzymatic cleavage at two out of three cleavage sites as these sites were masked by the staple.[91]

1.3.3. Canonical amino acid staples

Having briefly described a variety of peptide stapling procedures requiring at least one noncanonical amino acid, our attention now turns to stapling methods reliant on canonical amino acids. Peptide stapling between two canonical amino acids allows for potentially cheaper synthesis, as there is no need to include expensive noncanonical residues. Stapling through natural amino acids also allows for the possibility of stapling peptides and small proteins synthesised through bacterial protein expression methods.

While these methods are often straightforward, site-specific modification can be an issue (as noted in section 1.2.1.1 with the PEGylation of peptides/ proteins), leading to loss of product through the formation of undesired side-products.

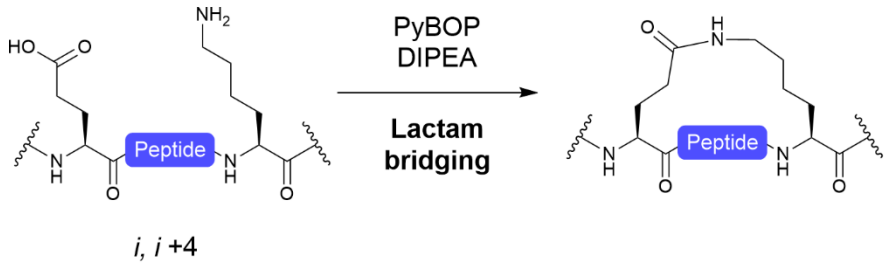
There are myriad different methods of peptide stapling relying solely on canonical amino acids,[106] including methods based around cysteine,[107-111] glutamic acid,[112-114] methionine,[115] histidine,[116-119] and tryptophan.[120, 121] Although a significant number of stapling methods require a single lysine residue as an anchor,[122-127] this section will focus on stapling methods dependent on two lysine residues as staple anchors.

1.3.3.1. Lactam-based staples

Two different methods for the synthesis of lactam-based staples will be described here. The first of these methods involves the intramolecular peptide cyclisation, *via* lactam formation, between a lysine and an acidic residue (either aspartic or glutamic acid).[5, 128-132] The second of these lactamisation methods involves stapling through a two-component process between a *bis*-lysine peptide and an activated diester.[133-136]

A review written by Taylor in 2002 cites various reports in which *i, i* + 4 lactam bridges have been used for the stabilisation of α-helices. As early as 1988 it was noted that the incorporation of an Asp to Lys (and/or Lys to Asp) lactam bridge led to an increase in helical content in growth hormone releasing factor (GRF (1-29)) analogues.[128, 137]

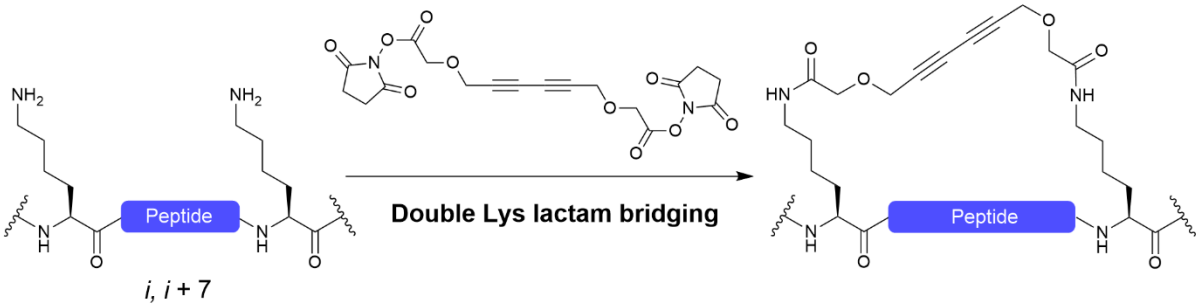
In 2020, Gomara *et al*. explored the use of two side chain lactam-bridged peptides (either stapled through Asp to Lys [*i, i* + 3] or Glu to Lys [*i, i* + 4]) as potential HIV-1 inhibitors. The cyclisation reactions were performed on-resin in the presence of PyBOP and DIPEA, although no conversion data or final yields of stapled peptides were reported. CD measurements showed that the Glu to Lys lactam bridged peptide (*i, i* + 4) appeared to have greater α-helix stabilisation compared to the linear precursor and the *i, i* + 3 lactam bridged peptide. The two lactam bridged peptides were also shown to have improved protease resistance compared to the linear control. As for the inhibitory properties of the stapled peptides, both the Asp to Lys (*i, i* + 3) and Glu to Lys (*i, i* + 4) lactam-bridged peptides exhibited enhanced HIV-1 inhibitory activity compared to the linear control (7-fold and 5-fold improvements respectively).[5]



Scheme 1.10 Peptide stapling *via* lactam bridging between glutamic acid and lysine.[5]

For the formation of an *i, i* + 7 spanning lactam-based staple, two-component synthesis strategies using either a *bis*-lysine-containing peptide and a diacid/ activated diester,[133] or a *bis*-acid-containing peptide and a diamine,[112] have been employed. As our focus is on the use of lysine residues as staple anchors, lactam stapling between a *bis*-lysine-containing peptide and an activated diester will be discussed.

In 2004, Fujimoto *et al*. described a two-component double lactam stapling procedure between a diacetylenic cross-linking agent and a *bis*-lysine-containing peptide (shown in Scheme 1.11). The stapling procedure was carried out in solution and yields of >70% stapled peptide were obtained. While the α-helical content of this stapled peptide was reported as ca. 65%, the helicity of the linear parent was not reported, so the impact of peptide stapling on helix stabilisation in this case cannot be fully quantified.[133]



Scheme 1.11 Peptide stapling *via* double lactam bridging between two lysine residues with a diacetylenic linker in a two-component stapling method.[133]

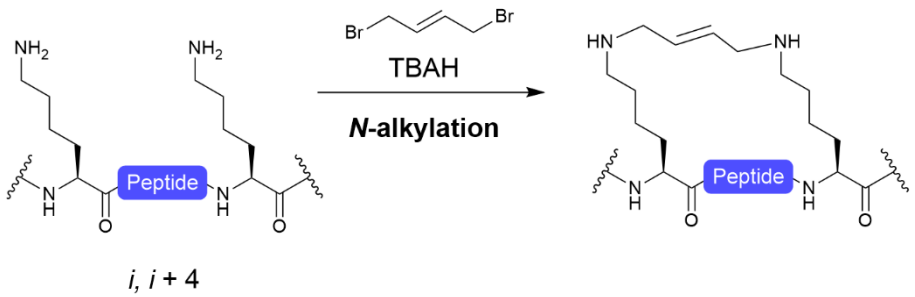
A further study by Fujimoto *et al*. in 2008 found that a double lactam staple spanning *i, i* + 10 significantly enhanced the α-helical content of a model peptide. The model, containing two lysine residues at the *i, i* + 10 positions, was noted as having ca. 20% α-helical content when in its linear form. Following solution phase stapling with naphthalene-2,6-dicarboxylic acid bis(2,5-dioxopyrrolidin-1-yl) ester, a yield of 35% stapled peptide was obtained with α-helical content of >95%.[138]

Peptides stapled through lactam bridging, either through side chain to side chain lactamisation or through two-component stapling between a *bis*-lysine-containing peptide and an activated diester, have been shown to exhibit enhanced α-helicity compared to their linear precursors over a range of staple lengths.[5, 128, 133, 138] The inclusion of a side chain to side chain lactam bridge has also been shown to improve peptide proteolytic stability and enhance biological activity.[5]

1.3.3.2. Lys-Lys direct *N*-alkylation

A study describing peptide stapling by direct *N*-alkylation was published by Zhang *et al*. in 2020, in which a series of stapled analogues of cationic antimicrobial peptides (CAMPs) was prepared and the antimicrobial properties of the cyclised peptides were investigated.[139]

Through on-resin cyclisation of a series of *bis*-lysine (*i, i* + 4) peptide models, using (*E*)-1,4-dibromobut-2-ene as the stapling agent and tetrabutylammonium hydroxide (TBAH) as the base (schematic representation in Scheme 1.12), a maximum yield of 19% was obtained. Following antimicrobial activity assays against both Gram-positive and Gram-negative bacteria, each of the stapled peptides showed similar activity to the linear control.[139]



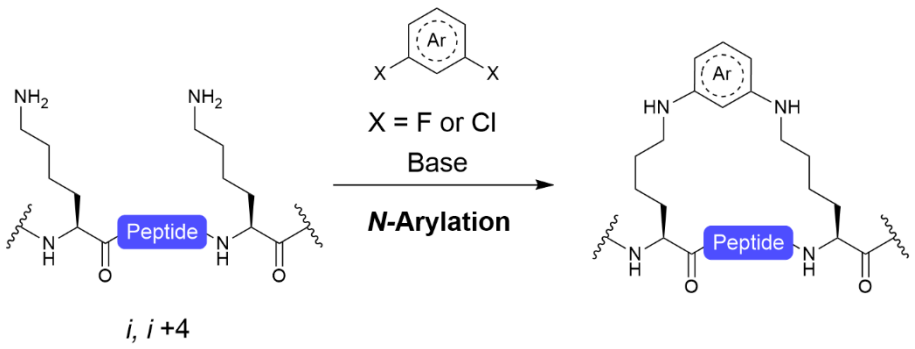
Scheme 1.12 Peptide stapling *via* direct *N*-alkylation of a *bis*-lysine peptide and (E)-1,4-dibromobut-2-ene.[140]

A second study carried out by the same group reported on the applicability of the *N*-alkylation stapling procedure to other natural and artificial CAMPs. Using the same stapling procedure as was reported in the 2020 study,[139] a maximum yield of 33% was obtained for the cyclisation of a Temporin-PTa-based peptide. An antimicrobial activity assay showed that, generally, the stapled peptides had either similar or slightly enhanced activities against Gram-positive bacteria compared to their linear counterparts.[140]

In general, direct *N*-alkylation between two lysine residues at *i, i* + 4 has been found to be a potentially useful stapling methodology for the cyclisation of CAMPs, despite low synthetic yields.[139, 140]

1.3.3.3. Lys-Lys arylation

An S_NAr-based stapling method was reported in 2016, in which two lysine residues undergo *N*-arylation to form a stapled peptide. A range of 7 aryl-based electrophiles were investigated as stapling agents against a model *bis*-lysine peptide with lysine residues positioned at *i, i* + 4. The stapling reactions were carried out in solution with either Tris base or DIPEA as the base and yields as high as 96% were obtained in 4 – 24 h at temperatures between room temperature and 37 °C.[141]



Scheme 1.13 Peptide stapling by *N*-arylation between two lysine residues and an aryl-based cross-linking agent.[141] The 1,3-dihalogenated aryl ring illustrated is representative of the reagents used during this stapling procedure.

In order to investigate the utility of the *N*-arylation stapling method in biological settings, a stapled peptide based on a known p53 peptide inhibitor of MDM2 protein (pDI) was synthesised, with the lysine residues positioned at *i, i* + 7, with a yield of 71%. The proteolytic stability of the stapled pDI-based peptide was found to be greater than that of the linear pDI-based peptide. The stapled peptide was also found to have significantly enhanced cell membrane permeability compared to the linear parent peptide. Finally, the binding affinity of the *N*-aryl-based stapled peptide was compared to that of the linear parent, with the stapled peptide showing an approximate two-fold improvement.[141]

Various other *N*-arylation-based stapling methods between two lysine residues have also been described.[142, 143]

Over the course of this section, a variety of different peptide stapling methodologies have been presented. Tables 1.1 and 1.2 briefly summarise each staple type discussed, showing an abbreviated representation of the staple structure, and highlighting the reported conversions and yields associated with each staple type.

Table 1.1 Summary of the stapling methods requiring noncanonical amino acids discussed throughout section 1.3.

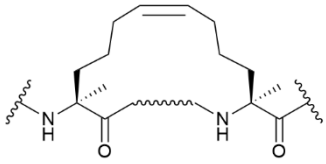
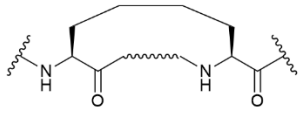
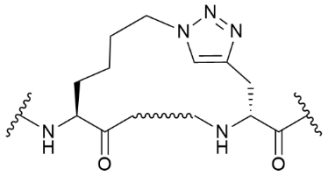
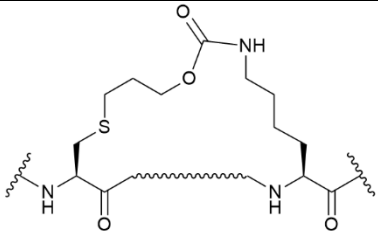
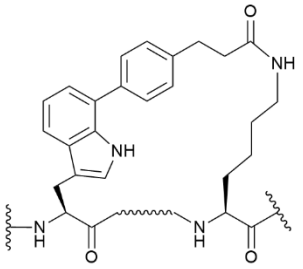
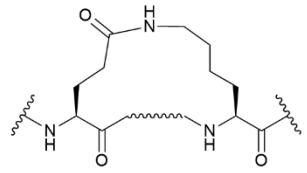
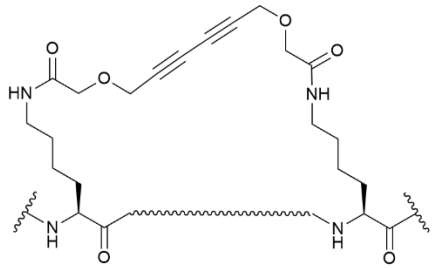
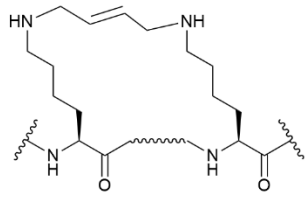
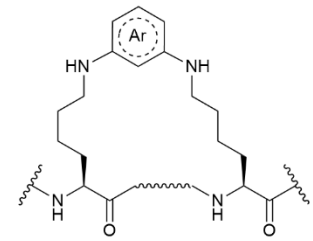
Staple structure	Stapling efficiency	Final yield
 Hydrocarbon staple	>98%[77] To completion[92]	Not reported[77, 92]
 VLM Hydrocarbon staple	To completion[90]	Up to 82%[90]
 Triazole staple	To completion[97]	55%[97]
 Thiol-ene staple	To completion[105]	24%[105]
 SMC staple	Not reported[91]	14%[91]

Table 1.2 Summary of the stapling methods requiring canonical amino acids discussed throughout section 1.3.

Staple structure	Stapling efficiency	Final yield
 Lactam staple (one Lys)	Not reported[5]	
 Lactam staple (two Lys)	Not reported	>70%[133]
 Alkyl-based staple	Not reported[140]	Up to 33%[140]
 Aryl-based staple	Not reported[141]	Up to 96%[141]

Having focussed predominantly on established peptide/ protein-based chemistry, the following section will move onto the development of chemical thiophosphorylation as we begin to look into the chemistry we aim to exploit within this project for the development of a novel peptide stapling method.

1.4. Thiophosphorylation

The exploitation of thiophosphorylation of proteins began with its use as a tool for the study of kinase activity[144] and mapping protein phosphorylation sites.[145] For both applications, thiophosphorylation was performed enzymatically using adenosine 5'-*O*-(γ-thiotriphosphate) (ATPγS) as the thiophosphorylating agent.

In more recent years, chemical methods of thiophosphorylation, typically using thiophosphoryl chloride ($PSCl_3$) or thiophosphoramidate ions, have been employed in the synthesis of aminonucleoside *N*-thiophosphoramidates,[146, 147] *N,S*-bridging dinucleotide analogues,[148] and thiophosphorylated peptides.[149]

Over the course of section 1.4, the development of chemical thiophosphorylation methods using a variety of thiophosphorylating agents will be explored.

1.4.1. Chemical methods of thiophosphorylation

The first reported chemical thiophosphorylation was in 1999 by Lasker *et al*. who conducted a study into the stability of thiophosphohistidine compared to phosphohistidine. A short peptide sequence (Ac-HGGGGAAAL-NH_2) was thiophosphorylated at the histidine residue using $PSCl_3$ in an overnight reaction. The major product was found to be the more thermodynamically stable 3-thiophosphohistidine isomer over the 1-thiophosphohistidine isomer.[150]

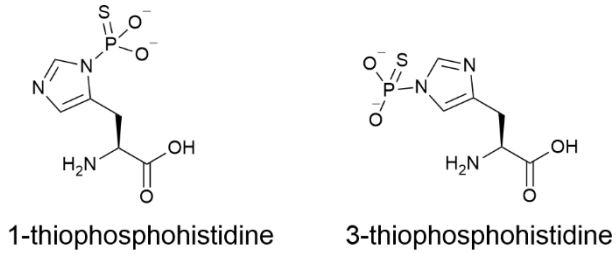


Figure 1.7 Two possible thiophosphorylated isomers of histidine.

To determine the stability of the thiophosphopeptide (Ac-(tpH)GGGGAAAL-NH_2) compared to the phosphopeptide (Ac-(pH)GGGGAAAL-NH_2) , both modified peptides were treated at acidic pH (0-7) for 3 h at room temperature and at pH 1 for 3 h at room temperature to determine time- and pH-dependence of hydrolyses. In both cases, the thiophosphopeptide was found to be more resistant to hydrolysis than the phosphopeptide. The thiophosphopeptide was more stable at pH 1 compared to pH 3 for the phosphopeptide in the pH dependence screening. In the time dependence screening held at pH 1, phosphopeptide showed significant hydrolysis after 1 h, while the thiophosphopeptide showed minimal degradation over 3 h.[150]

Further work on the thiophosphorylation of a single histidine residue was carried out in 2000. This work, instead of using $PSCl_3$, used potassium thiophosphoramidate (KTPA) as the thiophosphorylating agent. While this study also found that 3-thiophosphohistidine was the major product isomer over 1-thiophosphohistidine, the researchers found that thiophosphorylation of histidine was competing with the hydrolysis of KTPA, with hydrolysis of KTPA proving to be the dominant process. After 24 h,

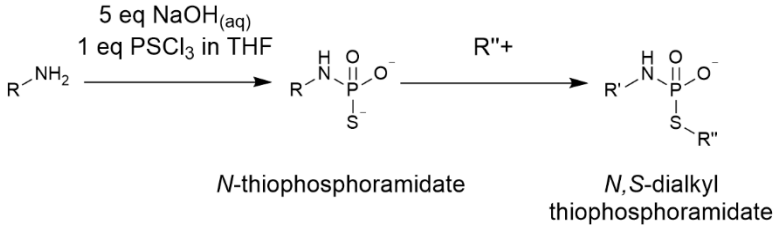
ca. 5% of the phosphorus present within the thiophosphorylation mixture was consumed in the formation of thiophosphohistidine, while ca. 70% of the phosphorus present was inorganic thiophosphate.[151]

Despite the low conversion to thiophosphohistidine demonstrated in the 2000 study, a study conducted in 2010 exclusively used thiophosphoramidate ions for the thiophosphorylation of free amino acids and thymidylate synthase. Diammonium thiophosphoramidate was used in a 3-fold excess to successfully thiophosphorylate histidine, lysine, cysteine, and serine.[149]

Thymidylate synthase was thiophosphorylated in two separate experiments using diammonium thiophosphoramidate at pH 7.5 (0.2 M Tris-HCl buffer) and pH 7.4 (0.1 M ammonium carbonate buffer). Both thiophosphorylations carried out on thymidylate synthase resulted in modification at available serine, lysine, and histidine residues, as determined by ^{31}P NMR analysis. While ratios of thiophosphorylated residues are not reported, the impact of the buffer used is noted. Compared to thiophosphorylation using Tris-HCl buffer (pH 7.5), thiophosphorylation using ammonium carbonate buffer (pH 7.4) resulted in reduced amounts of thiophosphorylated histidines, and much smaller corresponding thiophospholysine and thiophosphoserine signals. An activity assay carried out using thiophosphorylated thymidylate synthase found that enzyme activity was reduced by ca. 50% over a 24 h period compared to a control.[149]

Each of the examples discussed above employs aqueous thiophosphorylation methods on either amino acid or peptide substrates to generate thiophosphoramidates. The reaction times for these thiophosphorylations ranged from overnight,[150] 24 – 72 h at 4 °C,[149] and 72 h at room temperature.[151]

An aqueous thiophosphorylation method, reported in 2011 by Trmčić and Hodgson, gave conversions to *N*-thiophosphoramidates from primary amines of >90% in 1 h, using $PSCl_3$ as the thiophosphorylating agent. They also demonstrated subsequent *S*-alkylation of the anionic sulfur introduced via thiophosphorylation. Conversions to the *N*,*S*-dialkyl thiophosphoramidate from *N*-thiophosphoramidate were largely found to be > 80%.[152]

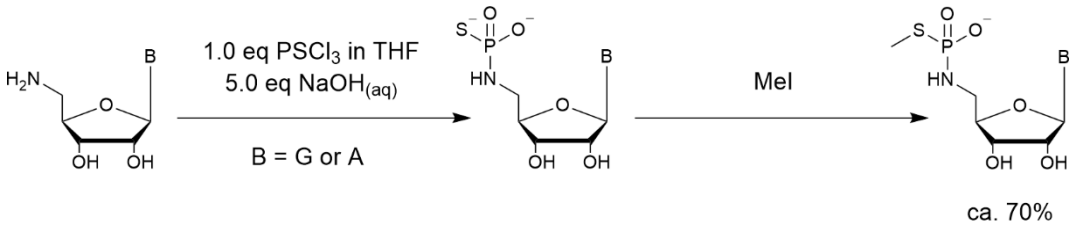


Scheme 1.14 General scheme for the one-pot *N*-thiophosphorylation and *S*-alkylation on a variety of amines using a range of alkylating agents.[152]

This work demonstrated that rapid, 'click-like' chemistry was possible for the aqueous synthesis of *N*-thiophosphoramidates. From here the Hodgson group began to explore potential uses for aqueous thiophosphorylation on compounds with biological relevance, largely focussing on the synthesis of aminonucleoside thiophosphoramidates. The following section will focus significantly on the aqueous *N*-thiophosphorylation of aminonucleosides, with a brief look into aqueous *N*-thiophosphorylation of phenylalanine.

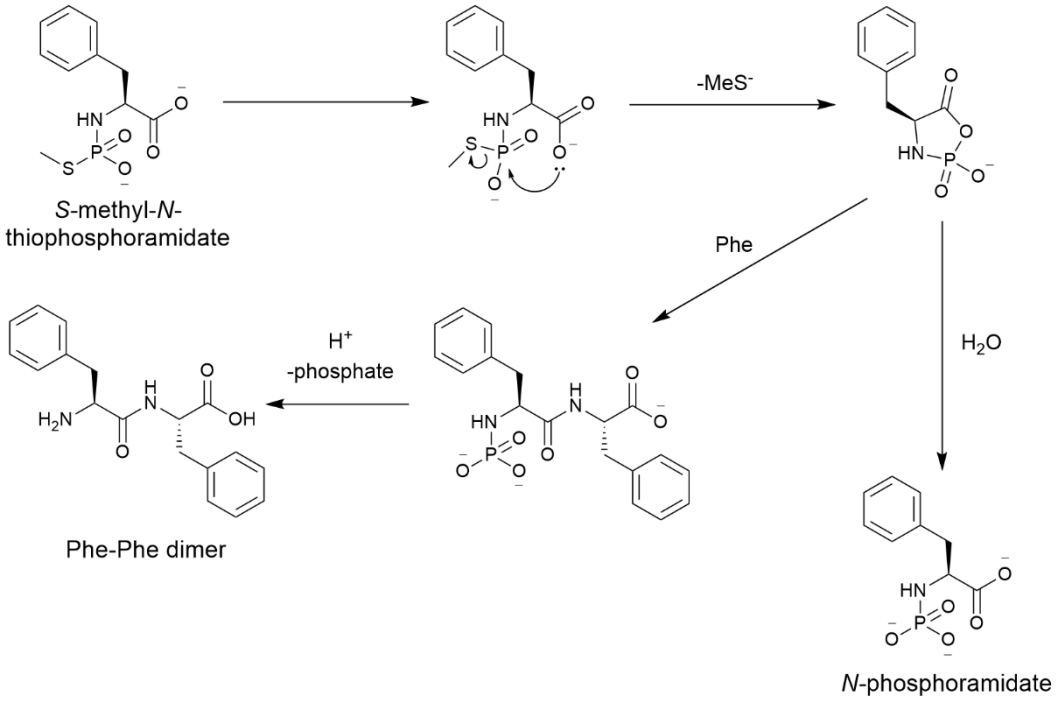
1.4.2. 'Click-like' aqueous *N*-thiophosphorylation of aminonucleosides and phenylalanine

The aqueous *N*-thiophosphorylation method optimised by Trmčić and Hodgson on primary amines mentioned in section 1.4.1 was adjusted for *N*-thiophosphorylation of two aminonucleosides (5'-amino-5'-deoxyguanosine and 5'-amino-5'-deoxyadenosine) and an amino acid (phenylalanine). Both aminonucleosides were taken straight through to *S*-alkylation with methyl iodide (MeI) using a continuous tripartite method, and both had conversions of ca. 70% from the aminonucleoside substrate to *S*-alkylated aminonucleoside-*N*-thiophosphoramidates.[146]



Scheme 1.15 *N*-thiophosphorylation and *S*-alkylation of 5'-aminonucleosides.[146]

As for phenylalanine, 85% was converted to the thiophosphoramidate derivative. The subsequent *S*-alkylation using MeI, however, gave a poor conversion to the *S*-methyl-*N*-thiophosphoramidate (24% according to ^{31}P NMR analysis). The major by-products appeared to be desulfurised *N*-phosphorylated phenylalanine, inorganic phosphate, and phenylalanine dimer. The significant presence of these by-products supported the idea of *N*,*S*-dialkyl thiophosphoramidate decomposition via intramolecular attack by the carboxylate group. A representative mechanism for this decomposition is illustrated in Scheme 1.16.



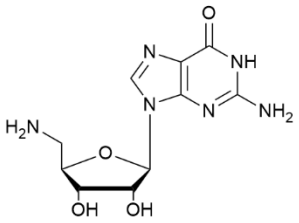
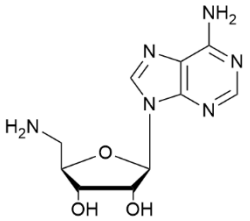
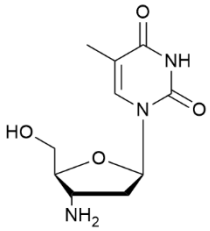
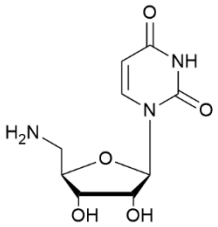
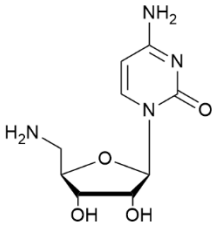
Scheme 1.16 Proposed decomposition routes for *S*-methyl *N*-thiophosphorylated phenylalanine.[146]

N-thiophosphorylation of amino acids was set aside at this stage, and the research focus remained on aminonucleosides.

5'-amino-5'-deoxyguanosine was used to find the optimal pH for aqueous *N*-thiophosphorylation of aminonucleosides using $PSCl_3$. A series of *N*-thiophosphorylations was conducted at pH 11.0, 11.5, 12.0, and 12.5 under pH control. Optimal conversion to *N*-thiophosphoramidate was obtained at pH 12, with 93% conversion during thiophosphorylation.[147]

Following on from pH optimisation, a series of five aminonucleosides underwent pH-controlled *N*-thiophosphorylation using $PSCl_3$ (1.0 eq) at pH 12. Table 1.3 shows the results of this experimental series, with the structures of each aminonucleoside illustrated.

Table 1.3 Conversion results following *N*-thiophosphorylation at pH 12 using $PSCl_3$ on aminonucleosides **G**, **A**, **T**, **U**, and **C**.[147]

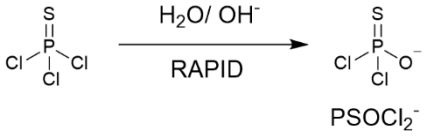
Aminonucleoside	Conversion using $PSCl_3$/ %
5'-amino-5'-deoxyguanosine **G** 	92
5'-amino-5'-deoxyadenosine **A** 	99
3'-amino-3'-deoxythymidine **T** 	65
5'-amino-5'-deoxyuridine **U** 	91
5'-amino-5'-deoxycytidine **C** 	92

N-thiophosphorylation was able to proceed with >90% conversions to the corresponding thiophosphoramidate for each of the 5'-aminonucleosides (**G**, **A**, **U** and **C**). Thiophosphorylation of the 3'-aminonucleoside (**T**), however, resulted in a significantly reduced conversion to thiophosphoramidate (65%). This reduction in thiophosphorylation is most likely due to the more sterically hindered environment surrounding 3'-NH_2 compared to 5'-NH_2.[147]

While aqueous thiophosphorylation with $PSCl_3$ had been shown to proceed with high conversions to thiophosphoramidate and minimal side-product formation, the use of an alternative thiophosphorylating agent under aqueous conditions was explored. A study on the hydrolysis of thiophosphorodichloridate ions published in 2012 suggested that they could be a useful alternative to $PSCl_3$.[153] With this in mind, potassium thiophosphorodichloridate ($KPSOCl_2$) was synthesised and its potential as a thiophosphorylating agent in aqueous conditions was explored.

1.4.2.1. Aqueous *N*-thiophosphorylation of aminonucleosides using $KPSOCl_2$

As $KPSOCl_2$ is water soluble it was hypothesised that it would be a more selective alternative to $PSCl_3$ under aqueous conditions.[153] $PSCl_3$ is known to be reactive with water, rather than soluble, and the first hydrolysis product is thiophosphorodichloridic acid (the conjugate base form of which is thiophosphorodichloridate [$PSOCl_2^-$]).



Scheme 1.17 First hydrolysis of $PSCl_3$ in aqueous alkaline conditions to generate thiophosphorodichloridate ions ($PSOCl_2^-$).

The half-life of $PSOCl_2^-$ is largely pH-independent, at ca. 3.2 min across the pH range ~2 to ~13, which would allow for thorough mixing in aqueous solutions within this pH range. As with $PSCl_3$, $PSOCl_2^-$ shows good selectivity for aminolysis over hydrolysis in the presence of amines. The large pH-independent region for hydrolysis of $PSOCl_2^-$ would even allow thiophosphorylation of the most basic alkyl amines (e.g. ε-NH_2 of lysine [pK_{aH} = 10.53][154]).[153]

A series of aqueous *N*-thiophosphorylations was set up using $KPSOCl_2$ (1.0 eq) as the thiophosphorylating agent and the aminonucleosides shown in Table 1.3 as the substrates. The results of this series are shown in Table 1.4.

Table 1.4 Conversion results following *N*-thiophosphorylation at pH 12 using $KOPSCl_2$ on aminonucleosides **G**, **A**, **T**, **U**, and **C**.[147]

Aminonucleoside	Conversion using $KPSOCl_2$/ %
5'-amino-5'-deoxyguanosine **G**	97
5'-amino-5'-deoxyadenosine **A**	90
3'-amino-3'-deoxythymidine **T**	48
5'-amino-5'-deoxyuridine **U**	82
5'-amino-5'-deoxycytidine **C**	80

For each thiophosphorylation using $KPSOCl_2$, lower conversions to thiophosphoramidate were recorded, with the exception of aminonucleoside **G**. The reduced conversions to thiophosphoramidate using $KPSOCl_2$ compared to $PSCl_3$ could be explained by enhanced selectivity of the thiophosphorylating agent used under aqueous conditions. $PSCl_3$ is insoluble in water, so thiophosphorylations using $PSCl_3$ may only be occurring at the water-solvent interfaces. In contrast to this, $KPSOCl_2$ is water soluble so there would be a greater chance of $KPSOCl_2$ undergoing hydrolysis rather than aminolysis in the homogeneous aqueous thiophosphorylation solution.[147]

As with the *N*-thiophosphorylations undertaken using $PSCl_3$, the lowest conversion to corresponding thiophosphoramidate was seen with the 3'-aminonucleoside substrate **T**. Again, this reduced conversion to thiophosphoramidate can be explained due to the sterically hindered environment surrounding the 3'-NH_2 group.

1.5. Aims and objectives

This project aims to build on work done previously by the Hodgson group through the implementation of aqueous *N*-thiophosphorylation on peptide substrates to develop a novel method for peptide stapling that is efficient, chemoselective, and more affordable than alternative methods that have already been established and discussed in section 1.3.

It was hypothesised that selective *N*-thiophosphorylation of the ε-NH_2 on a lysine side chain was possible due to the inherent nucleophilicity of the ε-NH_2 group. It was also thought that a two-component stapling methodology could be developed using a di-thiophosphoramidate intermediate, formed through thiophosphorylation of a model *bis*-lysine-containing peptide substrate, and a *bis*-alkyl halide linker, as shown in Figure 1.8.

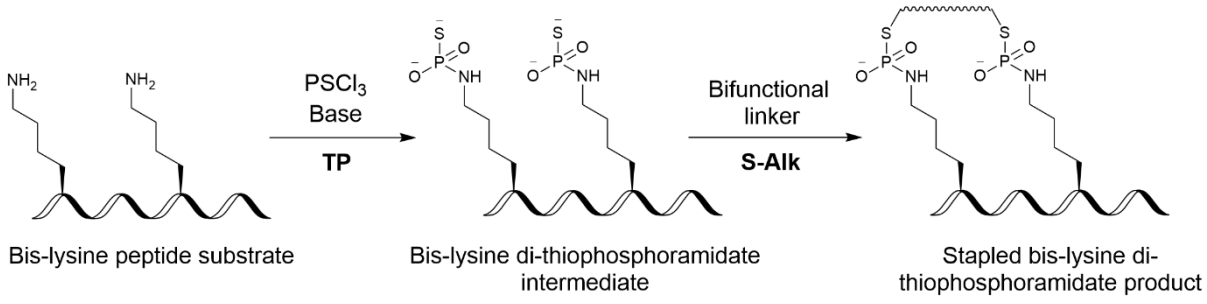


Figure 1.8 Proposed one-pot, two-step method for stapling a *bis*-lysine peptide substrate *via* a di-thiophosphoramidate intermediate. TP – thiophosphorylation; S-Alk – *S*-alkylation.

Through the use of inexpensive, readily available reagents (thiophosphoryl chloride and $NaOH_{(aq)}$) for the modification of unprotected peptides devoid of non-natural amino acids, we hope to deliver an

accessible and affordable novel peptide stapling methodology that is not particularly labour intensive and results in a high yield of stapled peptide.

In order to realise this ambition, aqueous *N*-thiophosphorylation must first be optimised. *S*-Alkylation of the resultant thiophosphoramidates must also be performed to ensure no unexpected phenomena occur during the formation of *N,S*-alkyl thiophosphoramidates (as with the *S*-alkylation of the phenylalanine-based thiophosphoramidate discussed in section 1.4.2). Additionally, the synthesis of *N,S*-alkyl thiophosphoramidates using model systems will provide some insight into the stability of such species during isolation by chromatographic means.

While it is hypothesised that selective aqueous *N*-thiophosphorylation at a lysine side chain is possible, a selectivity study must be performed using an array of lysine-containing unprotected model peptides to determine thiophosphorylation selectivity between the ε-NH_2 on a lysine side chain and other nucleophilic groups commonly found in peptides.

Finally, a model peptide substrate must be designed containing two lysine residues positioned at appropriate intervals. This model peptide substrate will undergo preliminary stapling experiments *via* our proposed one-pot, two-step stapling methodology illustrated in Figure 1.8.

With these aims in mind, the specific objectives of this thesis were:

1. To optimise the aqueous *N*-thiophosphorylation reaction on a model chromophore-labelled tetrapeptide.
2. To determine the selectivity of the *N*-thiophosphorylation reaction between the lysine ε-NH_2 and the *N*-terminal α-NH_2, as well as other nucleophilic side chain groups.
3. To design and staple a *bis*-lysine-containing peptide model.

These objectives are explored in Chapter 2 (objective 1), Chapter 3 (objective 2), and Chapters 5 and 6 (objective 3).

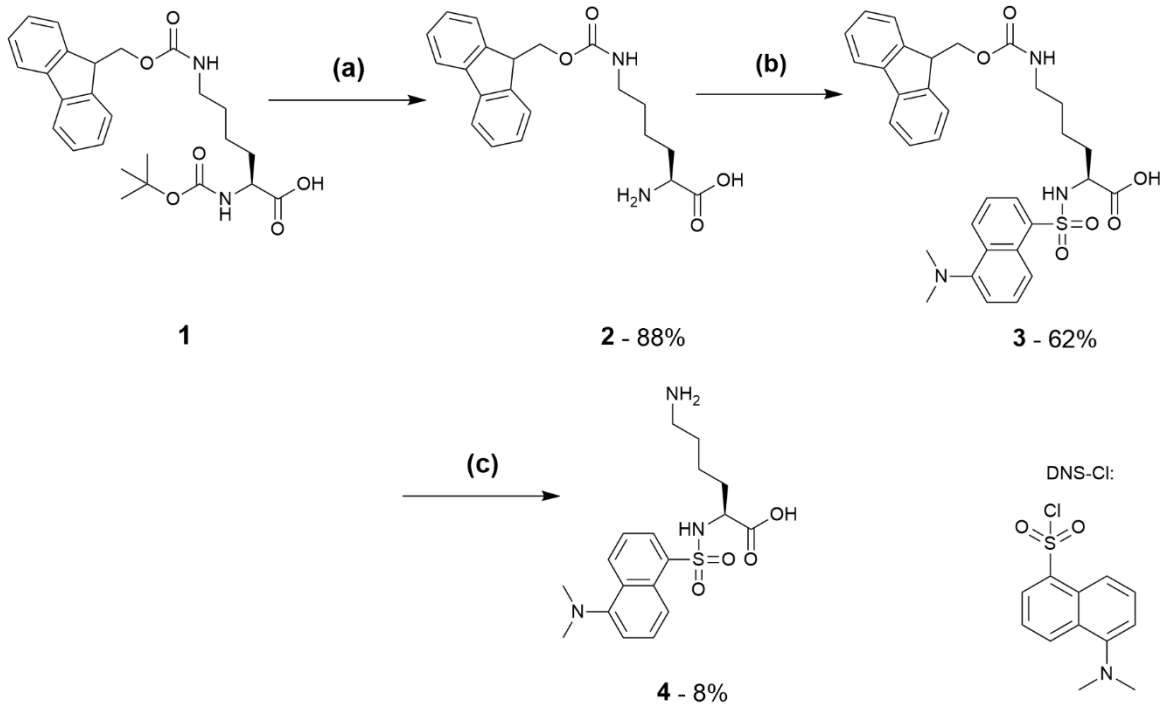
2. Optimisation of aqueous *N*-thiophosphorylation

This first results chapter will cover the synthesis and use of a model α-*N*-dansyl-labelled lysine system for the optimisation of aqueous *N*-thiophosphorylation, before moving on to a larger dansyl-labelled tetrapeptide model (section 2.3). Evaluation of the advantages and disadvantages of liquid chromatography-mass spectrometry (LC-MS) and ^{31}P NMR spectroscopy will be outlined to explain why LC-MS was initially chosen as the primary analysis method for assessing *N*-thiophosphorylation.

2.1. Synthesis of α-*N*-dansyl-labelled lysine model 4

In order to explore conditions for *N*-thiophosphorylation of amino-acid based systems using $PSCl_3$, a model system based on a single lysine residue was designed and prepared. A model system rather than a small peptide fragment was chosen for three reasons: (i) restrictions in free movement between laboratories, and limited time within laboratories brought about by the COVID-19 pandemic; (ii) simpler analyses *via* LC-MS and NMR spectroscopy; and (iii) availability of significant quantities of test material to allow optimisation of the chemical and analytical methods. Once the aqueous thiophosphorylation method had been optimised and analytical techniques developed, we planned to address larger peptides.

The lysine-based model was synthesised from orthogonally protected lysine **1** (α-*N*-Boc-ε-*N*-Fmoc-L-lysine) following the route illustrated in Scheme 2.1.



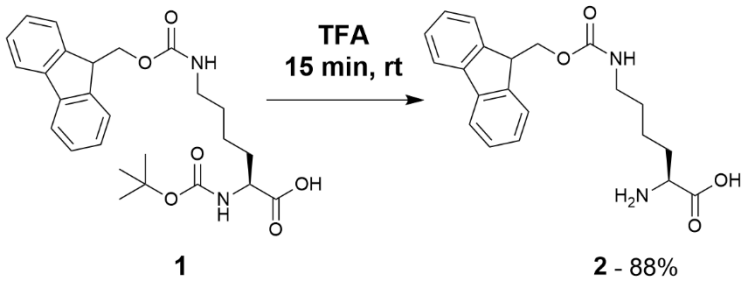
Scheme 2.1 Synthetic route for the preparation of α-*N*-dansyl-labelled lysine **4. (a)** TFA, 15 min (88%); **(b)** DNS-Cl, NaOH, MeCN, 1 h 20 min, rt (62%); **(c)** pyrrolidine, MeCN, 24 h, rt (8%).

As stated previously in the introduction (section 1.5), lysine was chosen as a target due to the inherent nucleophilicity of its side chain amine group (ε-NH_2) which should favour selective thiophosphorylation at the ε-NH_2 rather than completing hydrolysis of the thiophosphorylating agent.

This section will detail the synthesis, optimisation, and purification of α-*N*-dansyl-labelled lysine model **4**, starting with the removal of the α-*N*-Boc group and labelling with the dansyl group in the α-NH_2 position. The removal of the ε-*N*-Fmoc group and purification of the α-*N*-dansyl-lysine product **4** will conclude this sub-section before moving on to the optimisation of *N*-thiophosphorylation of **4** with $PSCl_3$.

2.1.1. Removal of α-*N*-Boc group from α-*N*-Boc-ε-*N*-Fmoc-L-lysine **1**

The removal of the α-*N*-Boc protecting group from orthogonally protected lysine **1** to generate ε-*N*-Fmoc-L-lysine **2** was performed using standard methodology based on trifluoroacetic acid (TFA)[155] (see Scheme 2.2 and experimental details in section 7.1).



Scheme 2.2 The removal of α-*N*-Boc group from α-, ε-orthogonally-protected lysine **1** to give ε-protected lysine **2** (88%).

LC-MS analysis with diode array detection (DAD) was used to confirm removal of the α-*N*-Boc protecting group by comparing retention times of the starting material **1** and ε-*N*-Fmoc-L-lysine **2** (Figure 2.1).

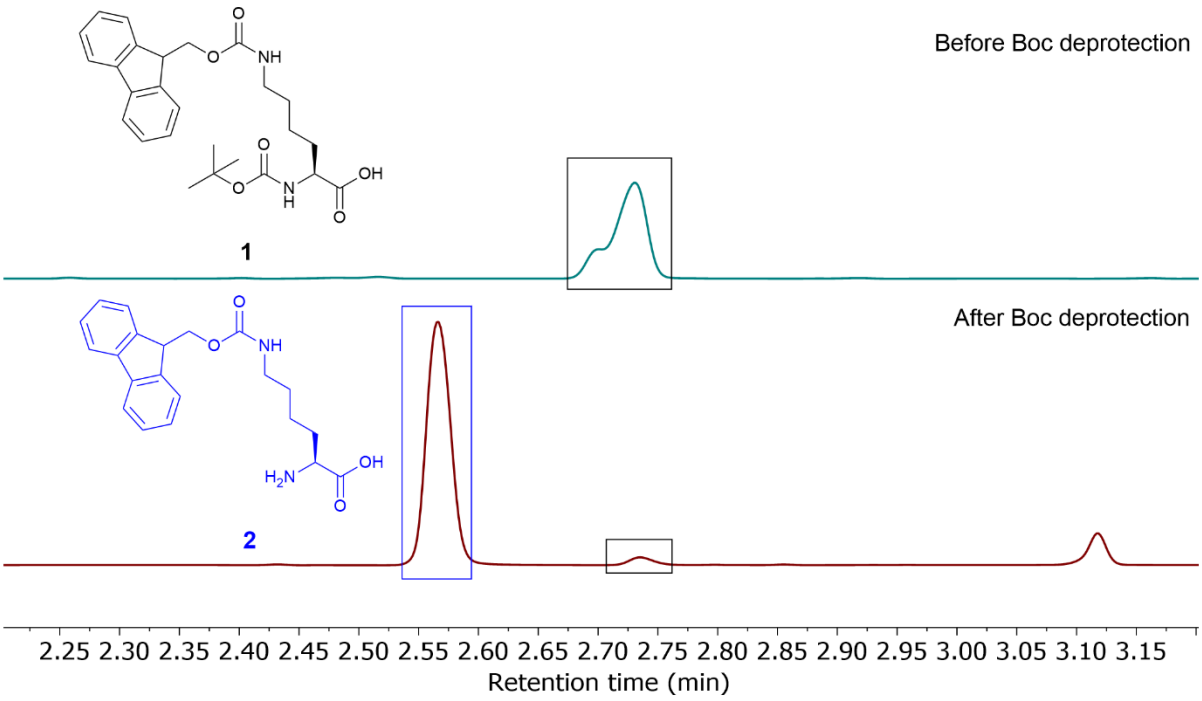


Figure 2.1 Total absorbance chromatographs from LC-MS analysis of α-, ε-orthogonally-protected lysine **1** (*m/z* 469 Da) before Boc deprotection (top trace) and crude ε-*N*-Fmoc-L-lysine **2** (*m/z* 369 Da) following Boc deprotection (bottom trace).

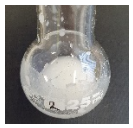
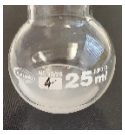
The Boc deprotection resulted in 88% conversion to ε-*N*-Fmoc-L-lysine **2** after 15 min, with a retention time shift from 2.73 min for **1** down to 2.57 min for **2**. The LC-MS experimental method is reverse phase (details in Chapter 7) resulting in the shortened retention time of **2** compared to **1** following the removal of the hydrophobic Boc group.

Having confirmed that a standard Boc deprotection method was adequate in this case, a suitable base and solvent system needed to be found for the α-*N*-dansylation step.

2.1.1.1. Solubility study on ε-*N*-Fmoc-L-lysine **2**

While all reagents used for the α-*N*-dansylation step were known to be soluble in dimethylformamide (DMF), the high boiling point of the solvent (153 °C) made it difficult to remove expediently, so it was an unattractive option. To find a more convenient solvent for the dansylation step, a solubility trial was carried out. Three solvents were selected (water, acetonitrile [MeCN] and DMF) as well as two bases (triethylamine [Et_3N] and 1 M volumetric standard NaOH solution). The bases were added to keep the pH of the solutions above the isoelectric point (pI) of the lysine derivative, thus ensuring that the α-NH_2 position remains unprotonated throughout the acid-generating dansylation process. The performance of each solvent system was determined through visual assessment – if significant amounts of precipitate were still visible following addition of both base and solvent, it was decided that the solvent system used was unsuitable for the dansylation step. Details of the trial are presented in Table 2.1. It should be noted that the quantities of ε-*N*-Fmoc-L-lysine **2** differ considerably between each experiment as the TFA salt generated through the removal of the α-*N*-Boc group was extremely sticky and difficult to handle quantitatively.

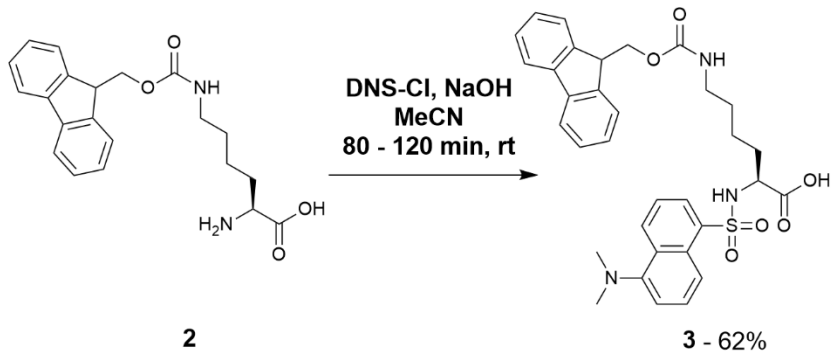
Table 2.1 Details and visual results of the solubility trial carried out on ε-*N*-Fmoc-L-lysine **2**.

Run	Lysine 2/ g	Solvent	Vol. solvent/ mL	Base	Vol. base/ mL	Amount base/ Eq.	Visual result
1	0.0732	H_2O	1.0	Et_3N	0.15	5.4	 Insoluble
2	0.1435	H_2O	2.0	NaOH	0.30	0.8	 Insoluble
3	0.0530	MeCN	1.0	Et_3N	0.10	5.2	 Insoluble
4	0.1411	MeCN	2.0	NaOH	0.30	0.8	 Partially soluble
5	0.0844	DMF	1.0	Et_3N	1.00	31.3	 Insoluble
6	0.1149	DMF	2.0	NaOH	1.00	3.2	 Soluble

All experiments with Et_3N as the base resulted in minimal, if any, visible dissolution of **2**. The mixtures of NaOH and DMF completely dissolved **2**, however, as stated previously, DMF was to be avoided due to difficulties in its removal. The MeCN and NaOH system was, therefore, chosen as the most appropriate for initial α-dansylation experiments.

2.1.2. Addition of dansyl label at the α-NH_2 position of ε-*N*-Fmoc-L-lysine **2**

Following the decision to work with MeCN and $NaOH_{(aq)}$ as the solvent and base, experimentation moved on to the addition of an α-dansyl group at the newly deprotected α-NH_2 position.



Scheme 2.3 Addition of the dansyl label on to ε-*N*-Fmoc-lysine **2** to give α-*N*-dansyl-labelled ε-*N*-Fmoc-lysine **3** (62%). Average concentration of solution of **2** used for dansylation – 57.4 mM; average concentration of solution of **2** in total reaction volume – 47.2 mM.

Dansyl chloride was chosen as the chromophore label due to its large extinction coefficient (approximately 3.4×10^3 $M^{-1}cm^{-1}$ at λ = 339 nm when bound to bovine serum albumin[156]). This large extinction coefficient means that very small quantities of the material can be used for UV detection and quantification, thus making LC-MS DAD analysis the preferred choice rather than ^1H or ^{31}P NMR spectroscopies. The significantly reduced material cost of LC-MS analysis compared to NMR (μg vs mg) already made it an appealing option, but the addition of a strong chromophore allows for more simple quantitative analysis through the integration of peaks within absorbance chromatographs (e.g. Figure 2.2) rather than (potentially complex) NMR spectra.

The use of a dansyl-labelled model also allowed for the development of a robust workflow and assay using LC-MS DAD analysis. The early development of this assay would become extremely beneficial, as it was subsequently planned to be employed in the analyses of larger peptide-based models (section 2.3 and Chapters 3, 4, and 5).

After finding an appropriate solvent and base system for the dansylation step, work commenced on the optimisation of both the number of mole equivalents (eq) of NaOH used, and the concentration of aqueous NaOH solution used.

2.1.2.1. Investigation into the impact of varying the number of equivalents of $NaOH_{(aq)}$ on dansylation of ε-*N*-Fmoc-L-lysine **2**

At least 3 eq of $NaOH_{(aq)}$ were used to dissolve every equivalent of ε-*N*-Fmoc-L-lysine **2** in initial dansylations as this gave full dissolution of **2** in mixtures with MeCN. These conditions raised the pH of the lysine derivative solution high enough (ca. 11-12 determined by pH paper) so that the addition of approximately 1 eq acid-forming dansyl chloride would not cause the pH to drop below 9.0. It was believed to be important to keep the pH of the dansylation reaction above 9.0 to ensure the lysine α-amine group was unprotonated and available as a nucleophile (the pK_{aH} of lysine α-NH_2 is 8.95[154]). However, the rate of hydrolysis of dansyl chloride increases rapidly above pH 9.5.[157] Due to this, a small working window between pH 9.0 and 9.5 was sought after for optimal dansylation conditions.

A short series of reactions was carried out varying the number of equivalents of $NaOH_{(aq)}$ used. Figure 2.2 displays the total absorbance chromatographs generated for each reaction. Details of the experimental procedures can be found in section 7.1.1.

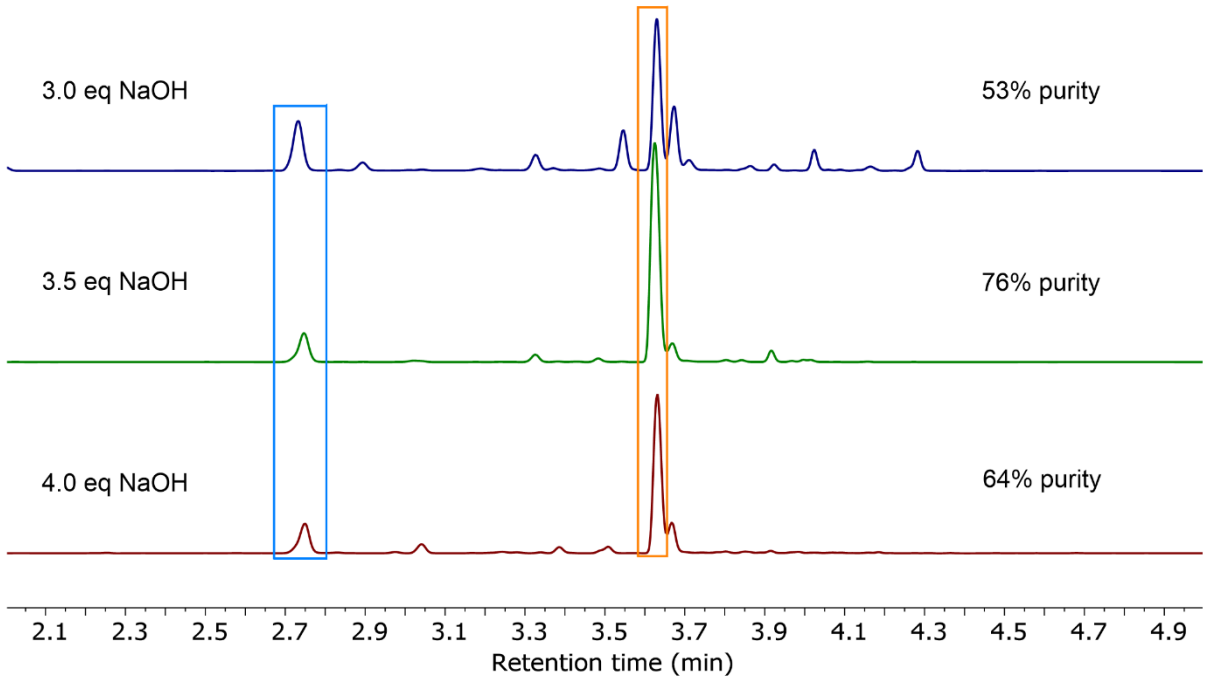


Figure 2.2 Total absorbance chromatographs from LC-MS analysis of $NaOH_{(aq)}$ equivalents series. The light blue box surrounding peaks at 2.7 min shows unreacted ε-*N*-protected lysine **2**, while the orange box around peak at 3.6 min shows α-*N*-dansylated ε-*N*-protected lysine **3** in the crude product.

This short series showed that 3.5 eq $NaOH_{(aq)}$ resulted in the greatest conversion from **2** (peak at 2.7 min) to α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** (peak at 3.6 min), with **3** making up 76% of the crude dansylation mixture after work-up, based on total absorbance. 3.0 eq $NaOH_{(aq)}$ was insufficient to keep

the pH above 9, thus ensuring the α-amine remains largely unprotonated, whereas 4.0 eq $NaOH_{(aq)}$ appeared to result in increased hydrolysis of dansyl chloride.

2.1.2.2. Investigation into the impact of varying the concentration of $NaOH_{(aq)}$ on dansylation of ε-*N*-Fmoc-L-lysine **2**

Having found that 3.5 eq $NaOH_{(aq)}$ for every 1.0 eq lysine derivative **2** gave the greatest conversion to α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**, work moved on to finding the optimal concentration of aqueous NaOH solution. Again, a series of experiments was devised where the concentration of NaOH used during the dansylation step was varied from ca. 0.5 M to 2.0 M, while keeping the number of equivalents of $NaOH_{(aq)}$ at 3.5 eq relative to ε-*N*-Fmoc-L-lysine **2**. For the $NaOH_{(aq)}$ concentrations below 1.0 M, water was added to 1 M volumetric standard solution to lower the concentration to the appropriate level. For $NaOH_{(aq)}$ concentrations between 1.0 and 2.0 M, water was added to 2 M volumetric standard solution to lower the concentration to the appropriate level. 1.0 and 2.0 M experiments were carried out using volumetric standard NaOH solutions.

Table 2.2 and Figure 2.3 show the results from this reaction series and the general experimental procedure can be found in section 7.1. A positive linear relationship can be seen up to the use of 1.1 M NaOH solution for dansylation, with the increase in the concentration of $NaOH_{(aq)}$ used corresponding to an increase in the percentage conversion from ε-*N*-Fmoc-L-lysine **2** to α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**. The increase in conversion level appears to be optimal at 1.1 M $NaOH_{(aq)}$ with 91% conversion to **3** (Table 2.2 and Figure 2.3). At concentrations of $NaOH_{(aq)}$ above 1.1 M, the conversions to α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** begin to reduce rapidly, illustrated by a conversion of 73% to **3** from **2** when using 1.2 M NaOH solution for the dansylation step.

Table 2.2 Results for the investigation in the impact of the concentration of NaOH solution used during the dansylation step. These data are graphical illustrated in Figure 2.3. The data indicating the greatest conversion of **2** to **3** highlighted in red. Percentage conversions were determined through LC-MS DAD analysis.

NaOH solution used/ M	Conversion from 2 to 3/ %	Amount of NaOH used/ Eq
0.54	29	3.51
0.75	58	3.51
1.00	81	3.49
1.11	91	3.50
1.21	73	3.49
1.31	76	3.49
2.00	43	3.51

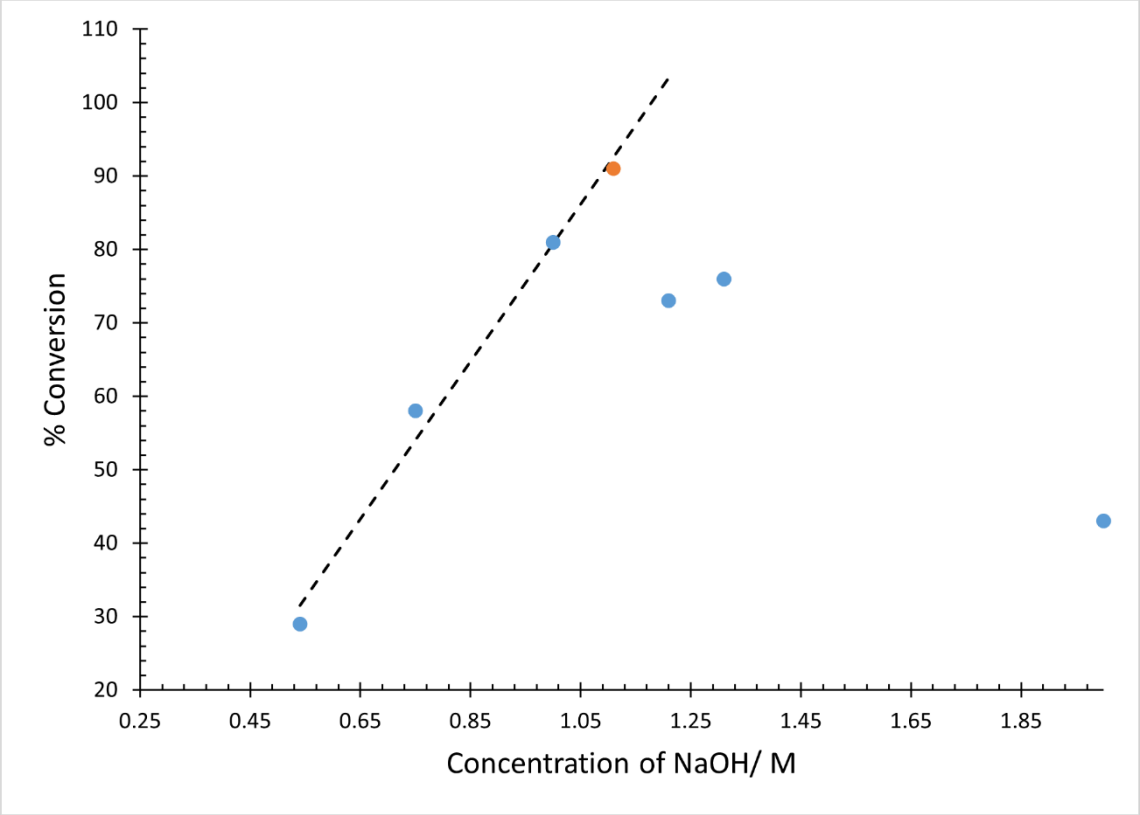


Figure 2.3 Percentage conversion to α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** from ε-*N*-Fmoc-L-lysine **2** as a function of the concentration of NaOH solution used for the dansylation reaction (graphic representation of data in Table 2.2). Data point indicating the greatest conversion from **2** to **3** is highlighted in orange.

2.1.2.3. Purification of α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**

Before moving on to the ε-Fmoc deprotection to give α-*N*-dansyl-L-lysine **4**, purification of the crude α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** was attempted. In order to facilitate optimisation of the purification of α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** by chromatography, a scale up was carried out from approximately 200 mg to ca. 1 g of α-*N*-Boc-ε-*N*-Fmoc-L-lysine **1** per reaction. Fortunately, the Boc deprotection and the dansylation step were simple linear scale ups and similar yields were obtained compared to the smaller scale reactions. On average over 11 reaction attempts, the conversion to α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** from ε-*N*-Fmoc-L-lysine **2** was found to be 56% in scaled up reactions based on LC-MS analysis.

Purification was carried out using a Teledyne CombiFlash Next Gen 100 system available in the laboratory by normal phase column chromatography with a hexane-ethyl acetate gradient solvent system. The chromatograms in Figure 2.4 show LC-MS analyses of samples of α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** before and after purification. Prior to purification, the sample was 71% α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**. Following purification, sample homogeneity improved to 86%.

The average purity of all α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** samples was 85% following purification using the method described above. Rather than spending an extended period of time trying to further optimise chromatographic resolution, it was decided to use the part purified material for trails of the Fmoc removal step because it was subsequently found that the remaining impurities were spectators and could be removed readily after the Fmoc deprotection.

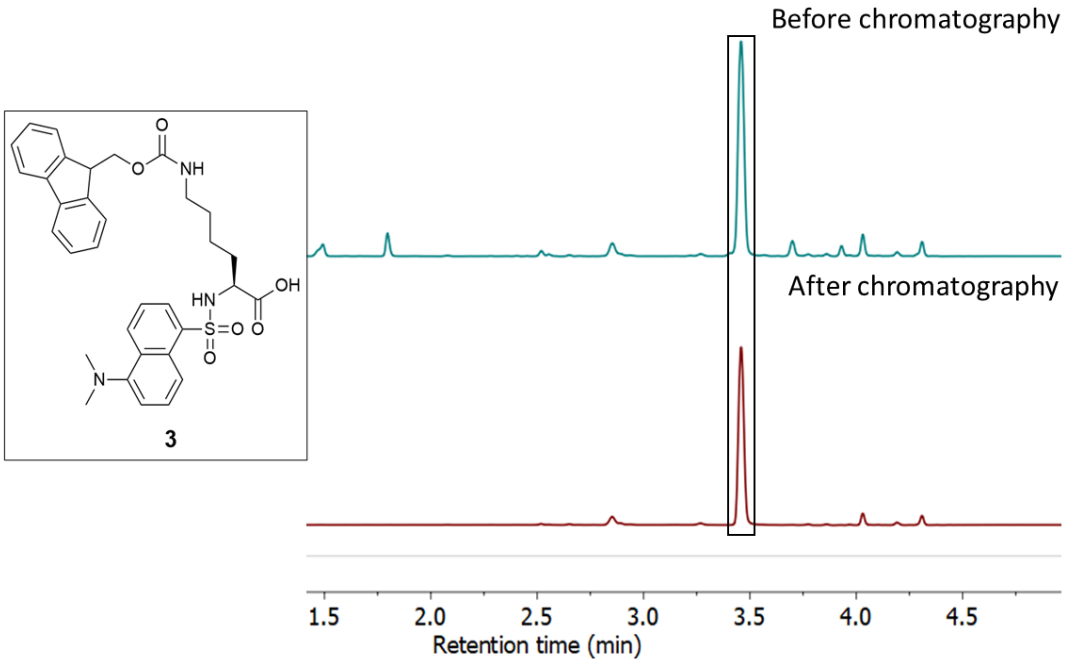
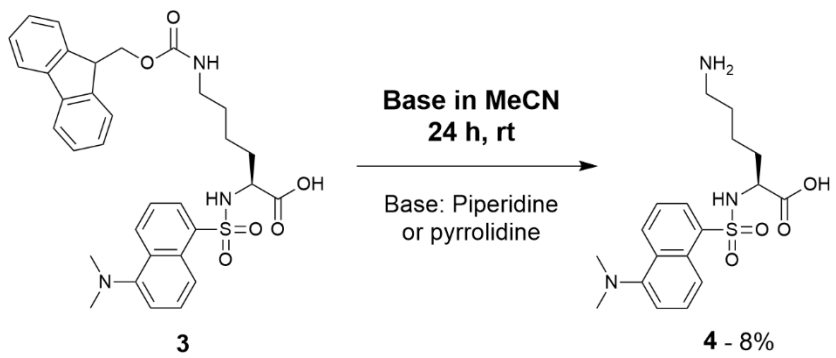


Figure 2.4 Total absorbance chromatographs from LC-MS analysis before and after chromatography. The top chromatograph shows crude α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine derivative **3**; the bottom chromatograph shows lysine derivative **3** after preparative chromatography. The black box around the peaks at 3.45 min highlights the presence of lysine derivative **3**.

2.1.3. Fmoc deprotection and purification of α-*N*-dansyl-L-lysine **4**

With partially purified α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** in hand, ε-*N*-Fmoc deprotection and purification of the final α-*N*-dansyl-L-lysine product **4** began.



Scheme 2.4 Removal of the ε-*N*-Fmoc group on α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** to give α-*N*-dansyl-L-lysine **4** (8%). Average concentration of solution of **3** in total reaction volume – 337.7 mM.

This section will be split into two parts, the first will focus on the Fmoc deprotection. The second will detail the purification of the α-dansyl-labelled product.

2.1.3.1. Removal of ε-*N*-Fmoc protecting group from α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**

Two amine bases were screened for the removal of Fmoc from α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**: piperidine and pyrrolidine. Both of these methods, and their results, will be discussed in the following two sub-sections.

2.1.3.1.1. ε-Fmoc deprotection of α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** using piperidine

The 'standard' method for removal of an Fmoc protecting group is 20% piperidine in DMF[155], however, as discussed in section 2.1.1.1, the use of MeCN was favoured to facilitate solvent removal following reaction termination.

The Fmoc deprotection was run for 24 h, and during this time, aliquots were periodically withdrawn for monitoring by LC-MS analysis. The reaction was ended after 24 h (Figure 2.5) by removal of solvents under vacuum.

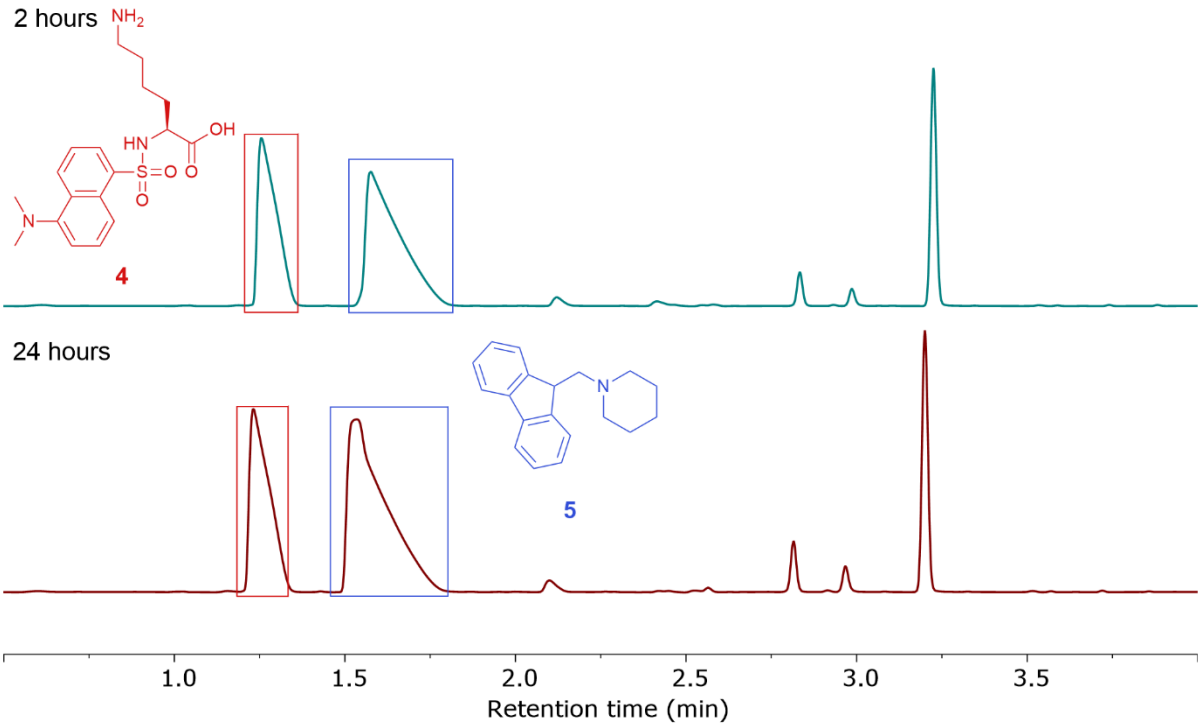


Figure 2.5 LC-MS total absorbance chromatographs taken to monitor the Fmoc deprotection of α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**. The top chromatograph was taken after 2 h, the bottom chromatograph was taken after 24 h. Each peak has been labelled using coloured boxes with the corresponding compound identity. Red boxes indicate α-*N*-dansyl-L-lysine **4** (*m/z* 380 Da); blue boxes indicate piperidine-fluorenyl adduct by-product **5** (*m/z* 264 Da).

Figure 2.5 suggests that the Fmoc deprotection was complete within 2 h, but the reaction was left running overnight to ensure total deprotection. Over the course of the reaction, a major side-product formed – piperidine-fluorenyl adduct **5**, confirmed by *m/z* 264 Da. Therefore, as stated in section 2.1.2.3, purification of the crude α-*N*-dansyl-L-lysine **4** was required before using the lysine-based model for thiophosphorylations and the purification details can be found in section 2.1.3.2 below.

However, as will be detailed in section 2.1.3.2, the removal of piperidine-fluorenyl adduct **5** proved particularly difficult so the use of pyrrolidine as base was investigated.

2.1.3.1.2. ε-*N*-Fmoc deprotection of α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3** using pyrrolidine

As with the ε-Fmoc deprotection reactions carried out using piperidine, the reactions using pyrrolidine were run for 24 h, and reaction progress was monitored periodically by LC-MS. The reaction was ended after 24 h (Figure 2.6) by removal of solvents under vacuum.

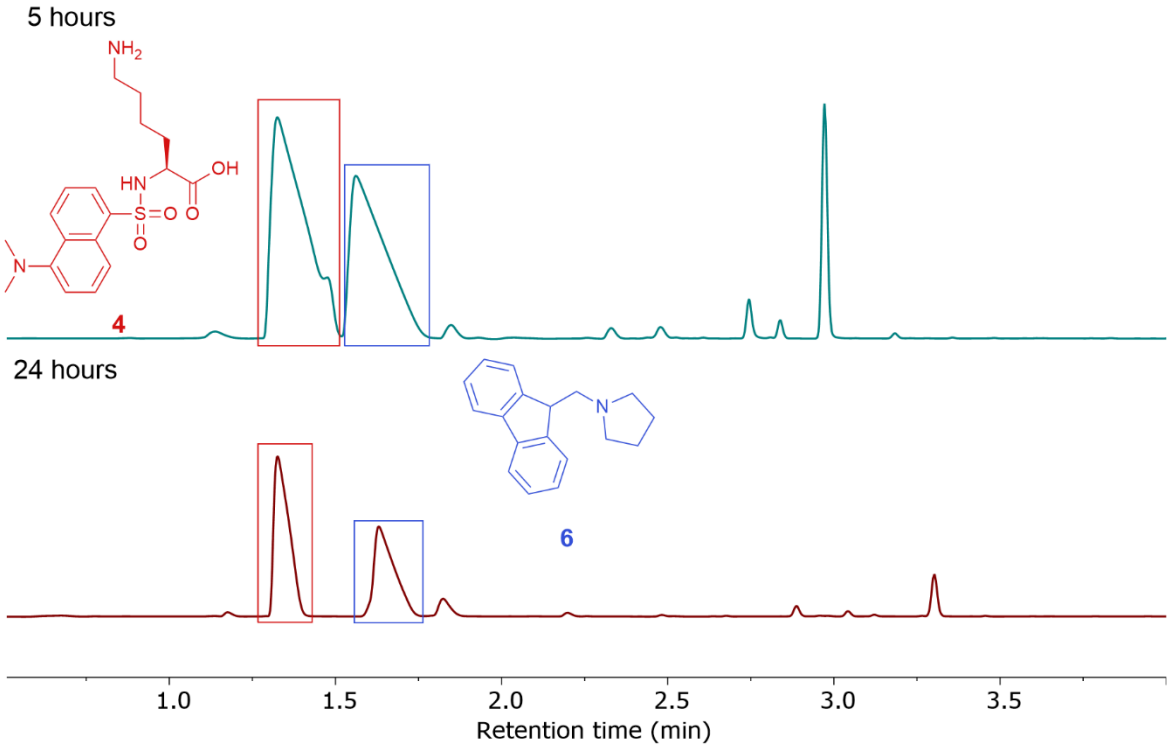


Figure 2.6 LC-MS total absorbance chromatographs taken to monitor the Fmoc deprotection of α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**. The top chromatograph was taken after 5 h, the bottom chromatograph was taken after 24 h. Each peak has been labelled using coloured boxes with the corresponding compound identity. Red boxes indicate α-*N*-dansyl-L-lysine **4** (*m/z* 380 Da); blue boxes indicate fluorenyl-pyrrolidine adduct by-product **6** (*m/z* 250 Da).

As with the Fmoc deprotection using piperidine (section 2.1.3.1.1), the reaction appeared to have run to completion well before the 24 h end point, but the reaction was continued to ensure total

deprotection. The major side-product generated over the course of the reaction is the pyrrolidine-fluorenyl adduct **6**. Removal of this adduct and purification of crude α-N-dansyl-L-lysine **4** is covered in section 2.1.3.2. Fortunately, pyrrolidine-fluorenyl adduct **6** proved to be easier to remove than its piperidine counterpart allowing work to move on to thiophosphorylations using the lysine-based model (discussed in section 2.2).

2.1.3.2. Purification of α-*N*-dansyl-L-lysine **4**

Two different purification methods were explored concurrently – selective precipitation (section 2.1.3.2.1) and reverse phase chromatography (section 2.1.3.2.2).

2.1.3.2.1. Precipitations

Precipitation was tested as a purification method due to its simplicity and potential rapidity. As α-*N*-dansyl-L-lysine **4** is more hydrophilic than either piperidine adduct **5** or pyrrolidine adduct **6**, precipitation appeared to be an appealing option for purification.

Using samples of crude α-*N*-dansyl-L-lysine **4** from Fmoc deprotection reactions using piperidine, several precipitation experiments were run. Crude samples were dissolved in HCl in MeOH (1.25 M), and diethyl ether was added dropwise to try to selectively precipitate **4** (further method details found in section 7.1.1). Multiple precipitations were carried out in an attempt to purify crude α-*N*-dansyl-L-lysine **4**, each following the same procedure. Samples were submitted for LC-MS analysis following each precipitation. Figure 2.7 shows the results from a purification attempt using multiple precipitations.

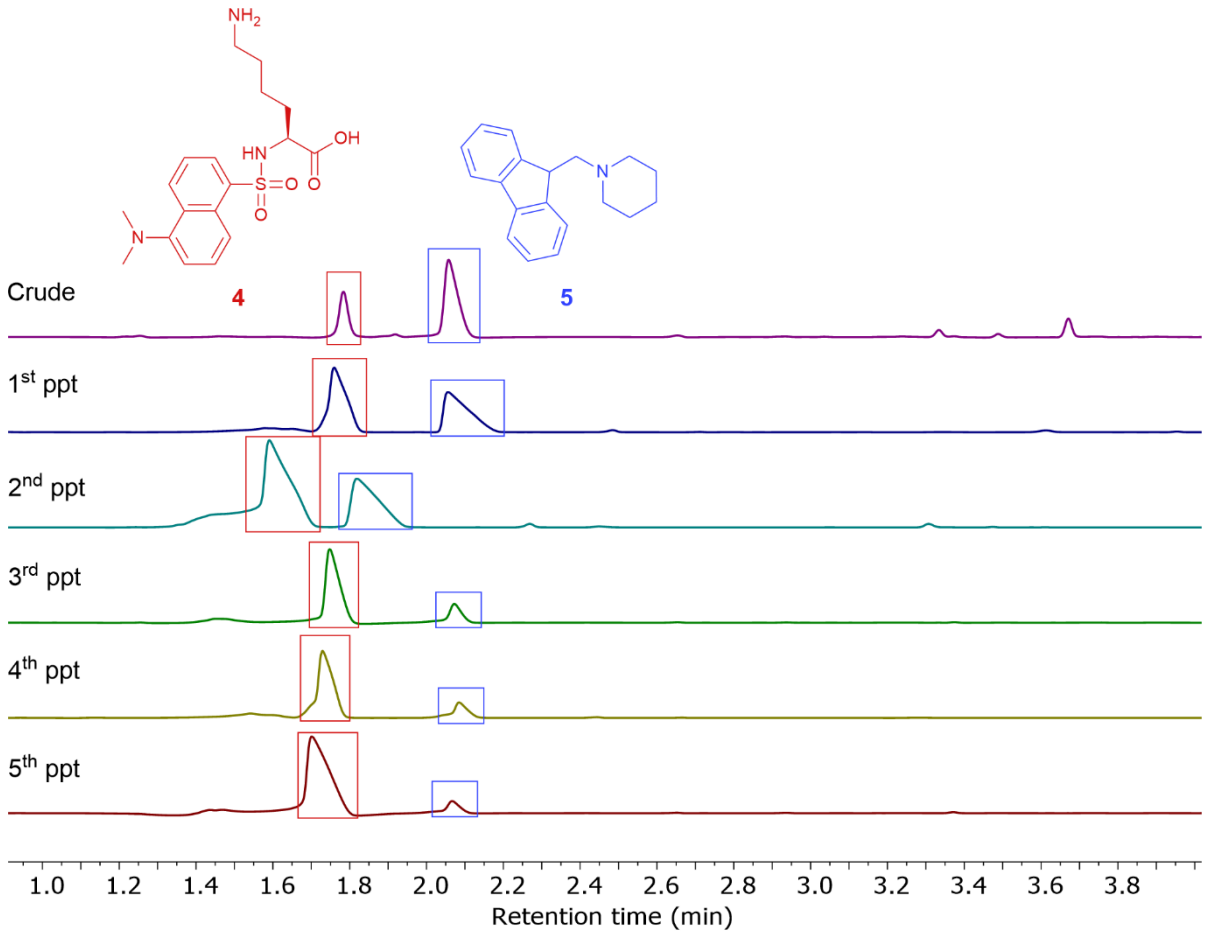


Figure 2.7 LC-MS total absorbance chromatographs of a series of precipitations done to try to purify α-*N*-dansyl-L-lysine **4** (*m/z* 380 Da). The top chromatograph shows crude **4** following Fmoc deprotection. The other five traces show the results from each precipitation. Red boxes correspond to **4**; blue boxes correspond to piperidine-fluorenyl adduct **5** (*m/z* 264 Da).

Table 2.3 Data from the precipitation series showing the amounts of α-*N*-dansyl-L-lysine **4** and piperidine-fluorenyl adduct **5**. Percentages were calculated solely from the area under the peaks corresponding to **4** and **5** highlighted in Figure 2.7.

Precipitation	Proportion of 4/ %	Proportion of 5/ %
Crude	27	73
1st	52	48
2nd	56	44
3rd	85	15
4th	87	13
5th	94	6

While both Figure 2.7 and Table 2.3 show that precipitation was improving the purity of **4**, going from 27% to 94% **4** over 5 precipitations, the improvements were diminishing and potentially may never have realised sample homogeneity. Thus, precipitation strategies were paused, and focus turned fully to chromatography.

2.1.3.2.2. Reverse phase chromatography

With precipitation having proven inadequate for purification of crude α-*N*-dansyl-L-lysine **4**, reverse phase chromatography was taken forward as the sole purification strategy. Reverse phase chromatography was chosen over normal phase due to the polar nature of α-*N*-dansyl-L-lysine **4**. Reverse phase chromatography was carried out on crude samples of α-*N*-dansyl-labelled lysine model **4** using the Teledyne flash column chromatography system, with 15.5 g C_{18} Aq Gold columns, available in the Hodgson lab (details in section 7.1.1).

After finding piperidine-fluorenyl adduct **5** difficult to remove by precipitation, pyrrolidine was used as the organic base during the Fmoc deprotection reaction (section 2.1.3.1.2).

The methods and results for the purification of crude materials from both the piperidine-based and pyrrolidine-based reactions were largely the same, so only the purification of **4** and removal of pyrrolidine-fluorenyl adduct **6** from a pyrrolidine-based Fmoc deprotection will be discussed. The key aim for the purification was to obtain a sample that was homogeneous by LC-MS DAD analysis so that we could use the areas under UV-vis absorbance peaks to monitor conversion of model **4** to ε-*N*-thiophosphoramidate **7**.

Crude α-*N*-dansyl-L-lysine **4** solution was injected on to a C_{18} column and the analyte was eluted using a triethylammonium acetate (pH 4.7) buffer and MeCN solvent system (further details of the equipment used and chromatography set-up can be found in Chapter 7).

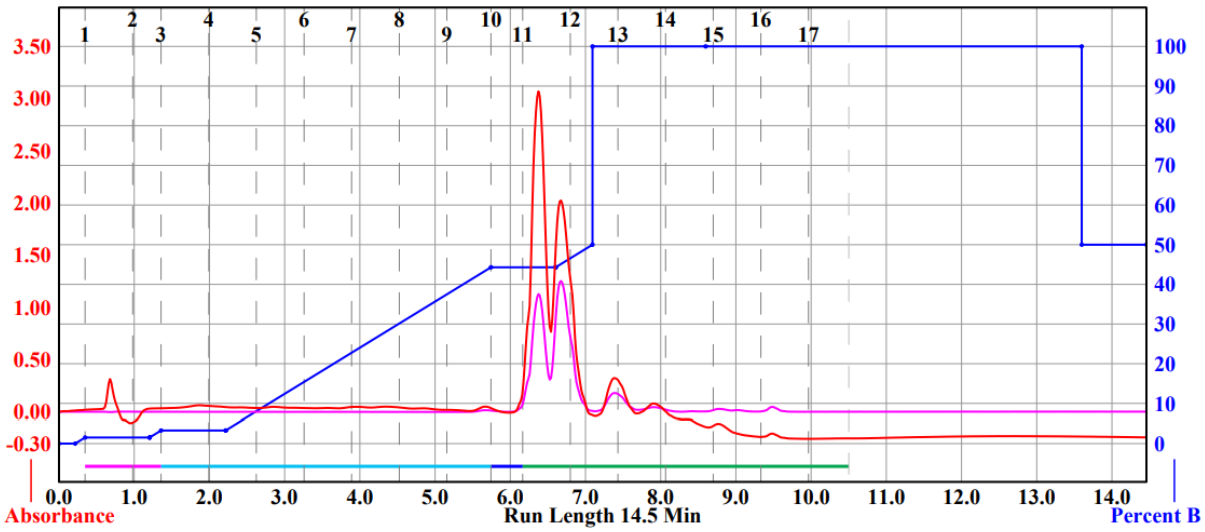


Figure 2.8 Chromatograph from the first reverse phase run to purify α-*N*-dansyl-L-lysine **4**. Each of the numbers running along the top of the chromatograph (1-17) correspond to 18 mL fractions taken. Individual peaks are indicated by the different coloured lines running below the chromatography trace, above the Run Length axis.

All peaks were collected in 18 mL fractions and samples from the major peaks between minutes 6 and 7 of the run (fractions 11 and 12) were submitted for LC-MS analysis. The results from both of these fractions indicated that this first reverse phase chromatography run had not fully resolved α-*N*-dansyl-L-lysine **4** (Figure 2.9).

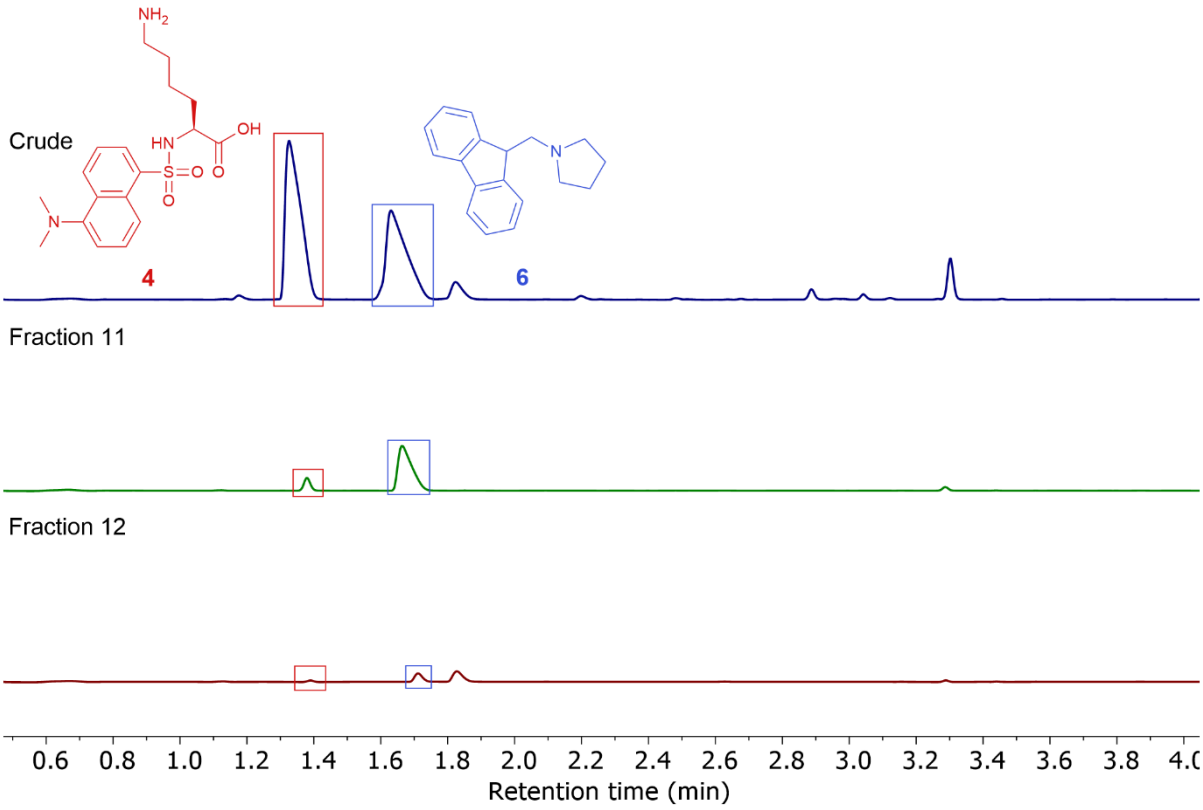


Figure 2.9 LC-MS chromatographs of crude α-*N*-dansyl-L-lysine **4** (top trace); fraction 11 (middle trace); and fraction 12 (bottom trace). The peaks relating to the species of interest are surrounded by boxes of the correlating colour – red boxes for α-*N*-dansyl-L-lysine **4** (*m/z* 380 Da) and blue boxes for pyrrolidine-fluorenyl adduct **6** (*m/z* 250 Da).

Both fractions 11 and 12 contain **4** and adduct **6** so they were collected, freeze dried, and a second reverse phase chromatography run was done on the resulting material. The second chromatographic purification run was carried out using the same mobile and stationary phases, however, the peaks were collected in smaller 5 mL fractions rather than 18 mL.

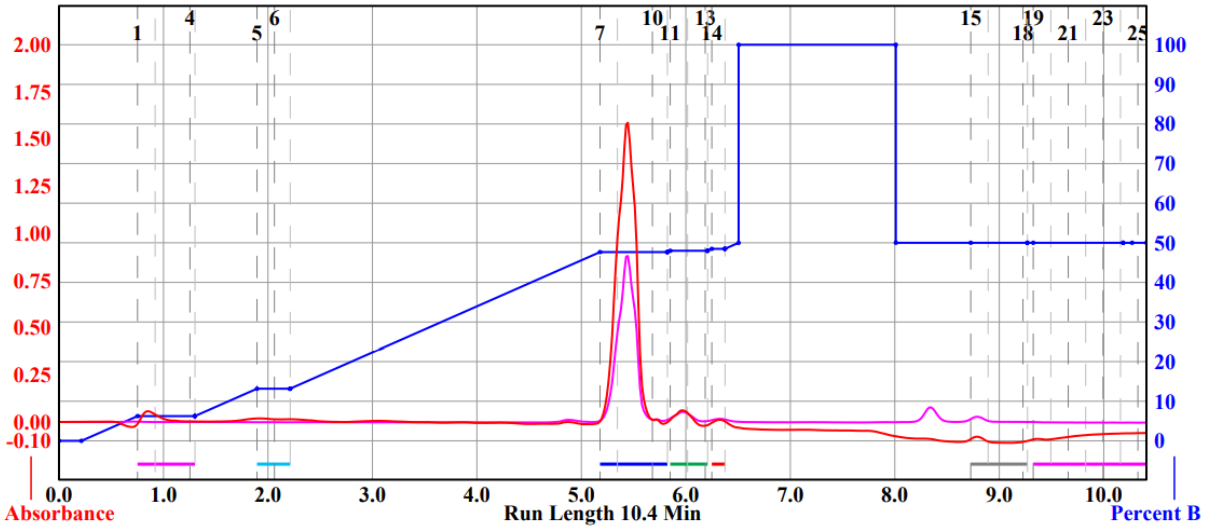


Figure 2.10 Chromatograph from the second reverse phase run to purify α-*N*-dansyl-L-lysine **4**. Each of the numbers running along the top of the chromatograph correspond to 5 mL fractions taken. Individual peaks are indicated by the different coloured lines running just below the chromatography trace, above the Run Length axis.

As with the first attempt at purification, LC-MS samples were submitted from the fractions eluted in the major peak between minutes 5 and 6 on the chromatograph shown in Figure 2.10 (fractions 7-10). The results from these LC-MS experiments are shown in Figure 2.11.

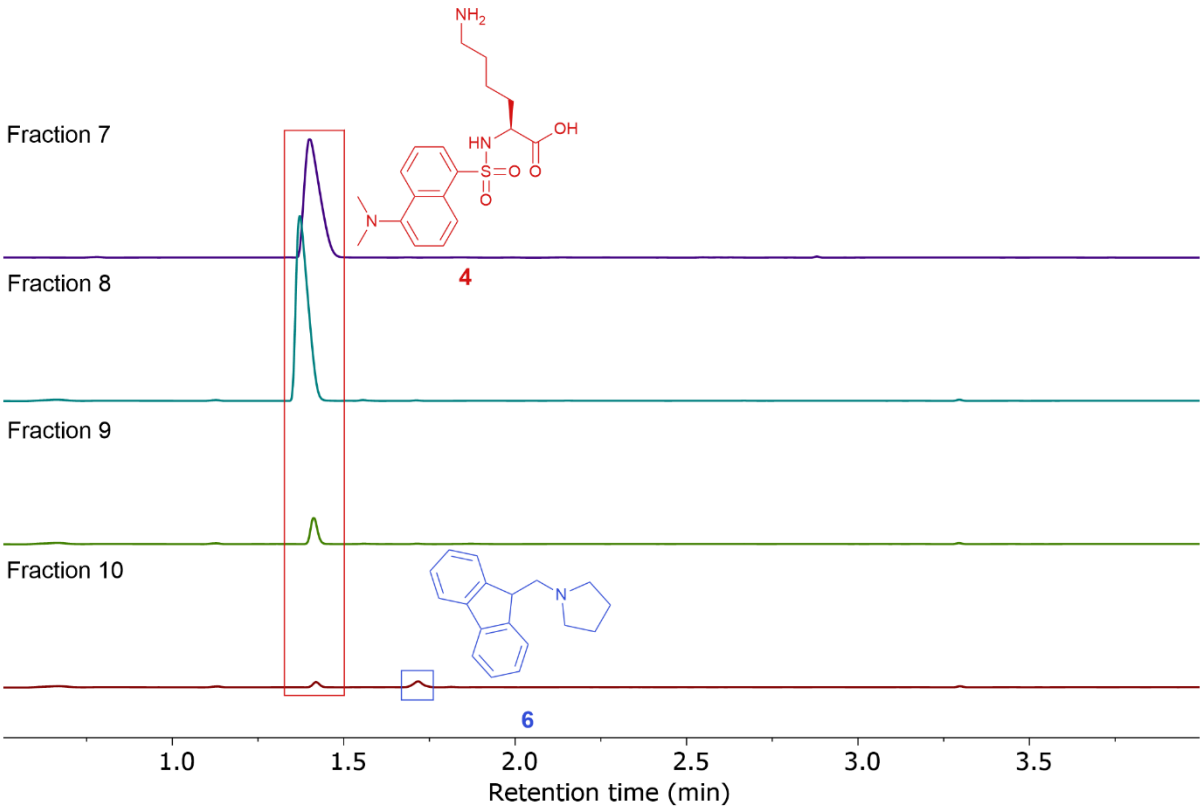


Figure 2.11 LC-MS chromatographs of fractions 7 – 10 in descending order. The peaks relating to the species of interest are surrounded by boxes of the correlating colour – red boxes for α-*N*-dansyl-L-lysine **4** (*m/z* 264 Da) and blue box for pyrrolidine-fluorenyl adduct **6** (*m/z* 250 Da).

As shown in Figure 2.11, fractions 7 – 9 contain solely α-*N*-dansyl-L-lysine **4** while fraction 10 contains both α-*N*-dansyl-L-lysine **4** and pyrrolidine-fluorenyl adduct **6**. Fractions 7 – 9 were collected together, organic solvent was removed on the rotary evaporator and the remaining aqueous buffer was removed through lyophilisation. The resulting solid material was then submitted for LC-MS analysis to confirm LC-UV-vis purity.

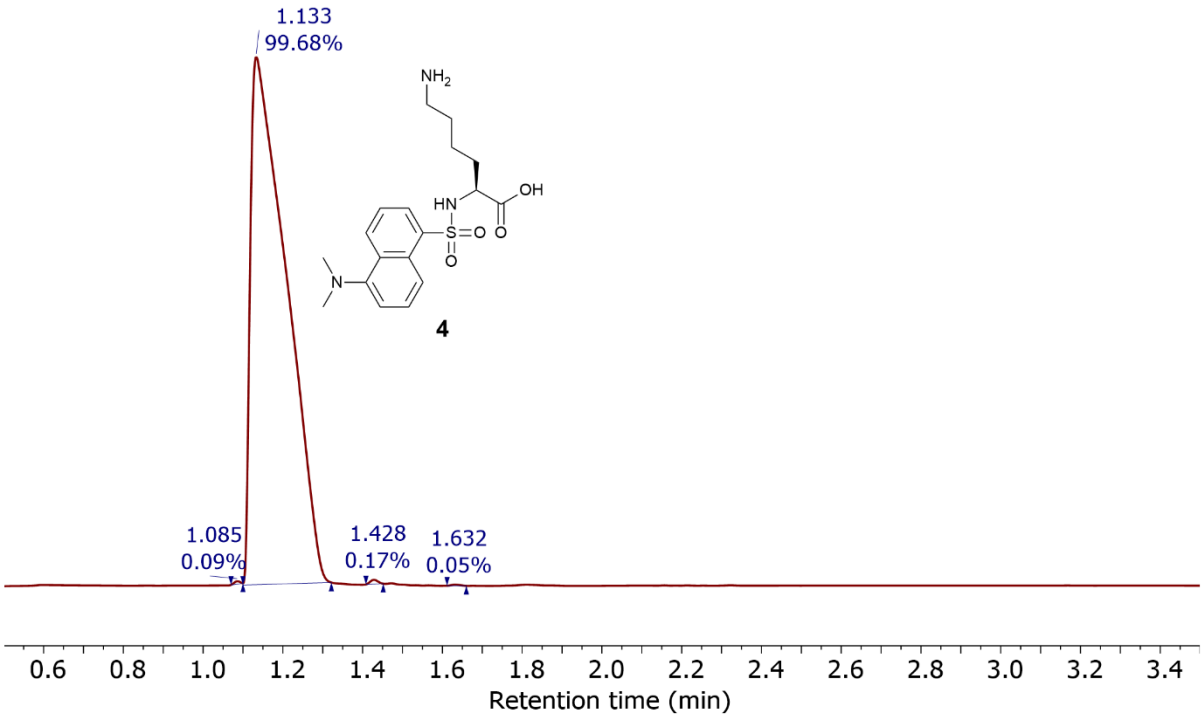


Figure 2.12 LC-MS chromatograph of purified α-*N*-dansyl-L-lysine **4** (*m/z* 380 Da) following two reverse phase chromatography runs.

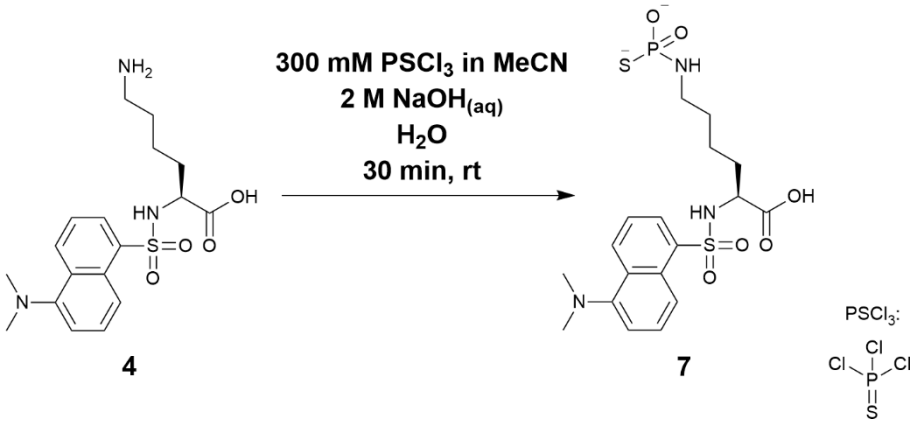
Following a double purification method involving two reverse phase chromatography runs, it was possible to obtain samples of **4** at 99.7% purity (by LC-MS DAD analysis). ^{1}H and ^{13}C NMR spectra of model **4** contained triethylammonium acetate signals remaining from the buffer used during chromatography, but the spectra otherwise indicated pure **4**.

Due to the relative ease and success of chromatography for the purification of crude samples of α-*N*-dansyl-L-lysine **4**, it was decided that this would be the method used in each case rather than precipitation. As stated at the beginning of this section, crude samples of **4** synthesised using the piperidine-based Fmoc deprotection method were also able to be purified by the C_{18} reverse phase method.

Now that a robust purification method had been developed to give 99.7% pure **4**, the lysine-based model could be used for optimising *N*-thiophosphorylation at the ε-NH_2 position under aqueous conditions. It should be noted that concentration corrections were carried out for solutions of α-*N*-dansyl-L-lysine **4** made up from samples purified by reverse phase chromatography to ensure the correct quantities of base and thiophosphorylating agent were being used for each optimisation reaction. The concentration corrections were done through comparing integrations of ^{1}H NMR signals for α-*N*-dansyl-L-lysine **4**, Et_3N, and acetic acid.

2.2. Optimisation of aqueous *N*-thiophosphorylation on α-*N*-dansyl-L-lysine 4

Having successfully synthesised dansyl-labelled lysine model **4** and purified it to give 99.7% LC-MS DAD purity, work turned to the aqueous *N*-thiophosphorylation of the model system. The general scheme of the *N*-thiophosphorylation is shown in Scheme 2.5. The *N*-thiophosphorylation procedure was adapted from previous work done within the Hodgson group.[152]



Scheme 2.5 Aqueous *N*-thiophosphorylation of α-*N*-dansyl-L-lysine **4** to give ε-*N*-thiophosphoramidate **7**.

The optimisation of *N*-thiophosphorylation was carried out through varying the number of equivalents of $PSCl_3$ used (either 1.0 eq or 1.2 eq). For both different numbers of equivalents of $PSCl_3$ used, the numbers of equivalents of $NaOH_{(aq)}$ were also varied.

The first *N*-thiophosphorylation series was carried out using 1.0 eq $PSCl_3$ while varying the number of equivalents of 2 M $NaOH_{(aq)}$ used in each reaction (details in section 7.1.2). The second *N*-thiophosphorylation series used the same procedures as the first but using 1.2 eq $PSCl_3$. The results from both series are displayed in Table 2.4 and Figure 2.13.

An LC-MS based assay, rather than a ^{31}P NMR based assay, was used to examine the degree of conversion from α-*N*-dansyl-L-lysine **4** to ε-*N*-thiophosphoramidate **7**. As stated previously in section 2.1.2, LC-MS was chosen as the primary analysis technique for the *N*-thiophosphorylation step due to the significantly lower material cost. All *N*-thiophosphorylation reactions carried out during this optimisation screening were done on a 25 µL scale (2.5 µmol of α-*N*-dansyl-L-lysine **4** used for each reaction), allowing large amounts of information to be gleaned from ca. 20 mg of α-*N*-dansyl-L-lysine **4**. For comparison, at least 5 mg of **4** would be needed for every *N*-thiophosphorylation reaction in order to obtain readable ^{31}P NMR spectra.

Table 2.4 Data from *N*-thiophosphorylation series of α-*N*-dansyl-L-lysine **4** to ε-*N*-thiophosphoramidate **7**. Conversions calculated from area under relevant peaks seen in LC-MS total absorbance chromatographs. Concentration of solution of **4** – 200 mM; concentration of solution of **4** in total reaction volume – 100 mM.

	1.0 eq 300 mM $PSCl_3$	**1.2 eq 300 mM $PSCl_3$**
2 M NaOH/ eq	**Conversion to 7/ %**	**Conversion to 7/ %**
6.0	82	70
6.5	84	86
7.0	87	94
7.5	92	94
8.0	88	97

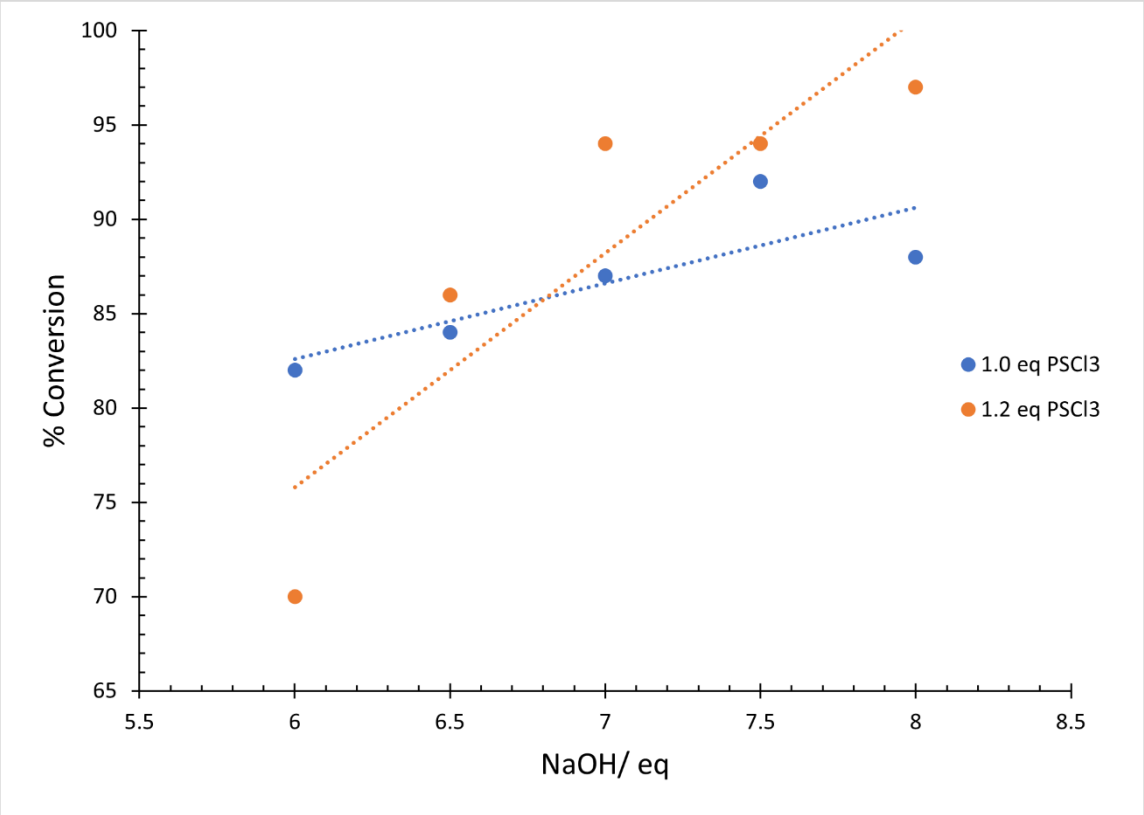
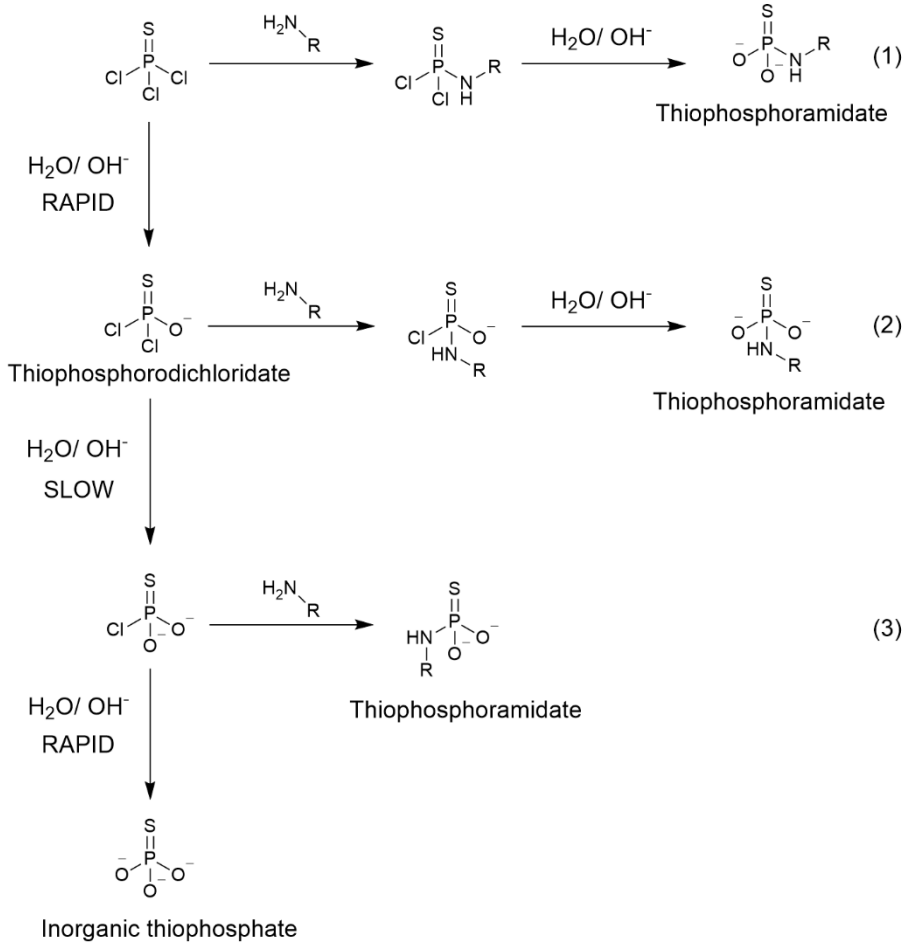


Figure 2.13 Percentage conversion of α-*N*-dansyl-L-lysine **4** to ε-*N*-thiophosphoramidate **7** during *N*-thiophosphorylation conditions screening (graphic representation of data presented in Table 2.4).

The general, approximately linear, trends seen for both the 1.0 eq $PSCl_3$ and 1.2 eq $PSCl_3$ series show that the greater the number of equivalents of 2 M $NaOH_{(aq)}$ used, the greater the conversion to ε-*N*-thiophosphoramidate **7**, with the greatest conversion seen using 1.2 eq $PSCl_3$ and 8.0 eq NaOH (97%). Although the scale of these reactions (ca. 25 µL) prevented pH monitoring before and after *N*-thiophosphorylation, it can be assumed for both series that with a greater number of equivalents of 2 M $NaOH_{(aq)}$ used, a greater proportion of the ε-amine group on **4** was in its unprotonated form thus leading to a greater conversion to thiophosphoramidate **7**.

While the mechanism of aqueous *N*-thiophosphorylation with $PSCl_3$ is not completely understood, there are various potential pathways (Scheme 2.6).



Scheme 2.6 Potential pathways for the thiophosphorylation of a primary amine under basic aqueous conditions.

Of the pathways indicated in Scheme 2.6, the most likely candidate to explain the trends seen in Figure 2.13, is route (2), based on the initial hydrolysis kinetics of the $PSCl_3$ oxy- analogue phosphoryl chloride ($POCl_3$). The initial hydrolysis of $POCl_3$ to phosphorodichloridic acid is essentially instantaneous,[158] with the same likely to be true for $PSCl_3$ hydrolysing to thiophosphorodichloridic acid (or the thiophosphorodichloridate anion as is the case under basic conditions) following diffusion from MeCN to water during *N*-thiophosphorylation. The subsequent hydrolysis of thiophosphorodichloridate to inorganic thiophosphate is largely pH independent and the relatively long half-life of thiophosphorodichloridate allows for thorough mixing in aqueous solutions (half-life of 3.2 min over pH range spanning from ~2 to ~13 at 25 °C).[153] This extended half-life of the thiophosphorodichloridate anion at pH > 11 allows for reaction with α-*N*-dansyl-L-lysine **4** with high conversions to ε-*N*-thiophosphoramidate **7**.

The slight reduction in conversion seen in Figure 2.13 for 1.0 eq $PSCl_3$ and 8.0 eq 2 M $NaOH_{(aq)}$ can, therefore, be explained assuming the pH of this reaction was greater than 13. While no significant data have been collected for the hydrolysis of thiophosphorodichloridate at pH > 13, Delley *et al*. reported a potential slight increase in the rate of hydrolysis of thiophosphorodichloridate around pH 13.2.[153] Through extrapolation of the data collected by Delley *et al*. for the hydrolysis of thiophosphorodichloridate, and assuming the same k_{obs}-pH rate profile trend as for hydrolysis of phosphorodichloridate,[159] it can be assumed that the rate of hydrolysis of thiophosphorodichloridate increases rapidly at pH > ~13. As the hydrolysis of thiophosphorodichloridate is the rate determining step in the total hydrolysis of $PSCl_3$ to inorganic thiophosphate, should the rate of this intermediate step increase, the chance of aminolysis decreases, resulting in a reduced conversion from α-*N*-dansyl-L-lysine **4** to ε-*N*-thiophosphoramidate **7**.

The reduced conversion seen in Figure 2.13 for 1.2 eq $PSCl_3$ and 6.0 eq 2 M $NaOH_{(aq)}$ can also be explained by considering the potential pH of the reaction. These conditions are suggested to have had the lowest pH of all those run during the optimisation screening as the equivalents of NaOH are the lowest compared to the equivalents of acid forming $PSCl_3$. If the pH of this reaction, therefore, was closer to the pK_{aH} of ε-NH_2 group of a lysine side chain (10.53)[154] than any of the other reactions in this screening, there would be a smaller proportion of ε-NH_2 in its unprotonated, nucleophilic form. This in turn would result in a reduced conversion of α-*N*-dansyl-L-lysine **4** to ε-*N*-thiophosphoramidate **7**.

Having found optimal *N*-thiophosphorylation conditions for α-*N*-dansyl-L-lysine **4** (1.2 eq $PSCl_3$ and 7.0-8.0 eq 2 M $NaOH_{(aq)}$), several tetrapeptide models were designed in order to test the optimised conditions on larger systems and to learn about the selectivity of aqueous *N*-thiophosphorylation between competing *N*-sites i.e. ε-NH_2 of a lysine side chain vs α-NH_2 of *N*-terminal amino acids.

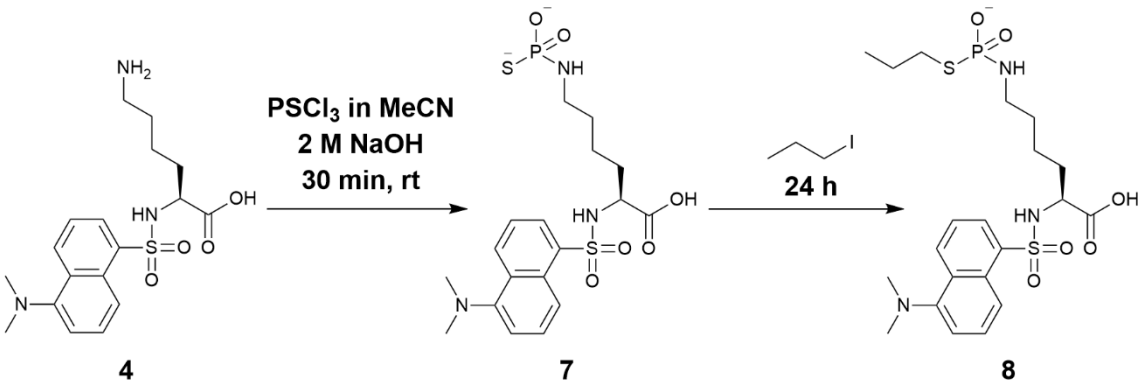
2.2.1. *S*-alkylation of ε-*N*-thiophosphoramidate **7**

Before advancing with *N*-thiophosphorylation of larger models, however, *S*-alkylation with 1-iodopropane on ε-*N*-thiophosphoramidate **7** will be briefly discussed. As will be seen in section 3.3.1, *S*-alkylation of thiophosphorylated tetrapeptide models will be used in an investigation into the selectivity of *N*-thiophosphorylation. The *S*-alkylation methodology is introduced here to illustrate that there was potential scope to use *S*-alkylation during the selectivity investigation for simple thiophosphoramidate 'trapping', as well as the additional benefit of stabilising the acid-labile N-P bond of a thiophosphoramidate.

While the N-P bond in (thio)phosphoramidates is known to be labile under acidic conditions, alkylation at the anionic sulfur introduced has been shown to significantly increase thiophosphoramidate half-

life under both physiological and acidic conditions.[146] Due to this reported increase, *S*-alkylation was performed on ε-*N*-thiophosphoramidate **7** to 'trap' it as an *S*-alkyl ε-*N*-thiophosphoramidate and facilitate its isolation by reverse phase chromatography and allow detailed spectroscopic analysis.

The *S*-alkylation procedure shown in Scheme 2.7 was adapted from work previously done by the Hodgson group,[146] using 1-iodopropane as the alkylating agent.



Scheme 2.7 *N*-thiophosphorylation and subsequent *S*-alkylation (with 1-iodopropane) of α-*N*-dansyl-L-lysine **4** to give *S*-propyl ε-*N*-thiophosphoramidate **8** *via* ε-*N*-thiophosphoramidate **7**.

The aqueous *N*-thiophosphorylation reaction developed and optimised in section 2.2 was scaled up linearly from 25 µL to 2.3 mL scale (from 2.5 µmol to 275 µmol of **4**), using 1.0 eq $PSCl_3$ and 7.0 eq $NaOH_{(aq)}$, to give 80% conversions to ε-*N*-thiophosphoramidate **7** determined by LC-MS DAD analysis. *S*-alkylation was carried out immediately, with 7.0 eq 1-iodopropane (300 mM solution in MeCN) added to the crude thiophosphoramidate mixture.

Figure 2.14 shows the reaction progress following *N*-thiophosphorylation and *S*-alkylation. While *N*-thiophosphorylation proceeded with 80% conversion to ε-*N*-thiophosphoramidate **7**, as previously stated, the *S*-alkylation step appears to have resulted in 95% conversion of **7** to *S*-propyl ε-*N*-thiophosphoramidate **8**, with all ε-*N*-thiophosphoramidate **7** consumed over the course of the reaction. The main by-product seen was the hydrolysis product of ε-*N*-thiophosphoramidate **7** into ε-*N*-phosphoramidate **7b** (5%), and there was minimal evidence of direct ε-*N*-alkylation of α-*N*-dansyl-L-lysine **4** with 1-iodopropane (2% of all dansyl-labelled lysine species present – not highlighted in Figure 2.14) over the 24 h reaction period.

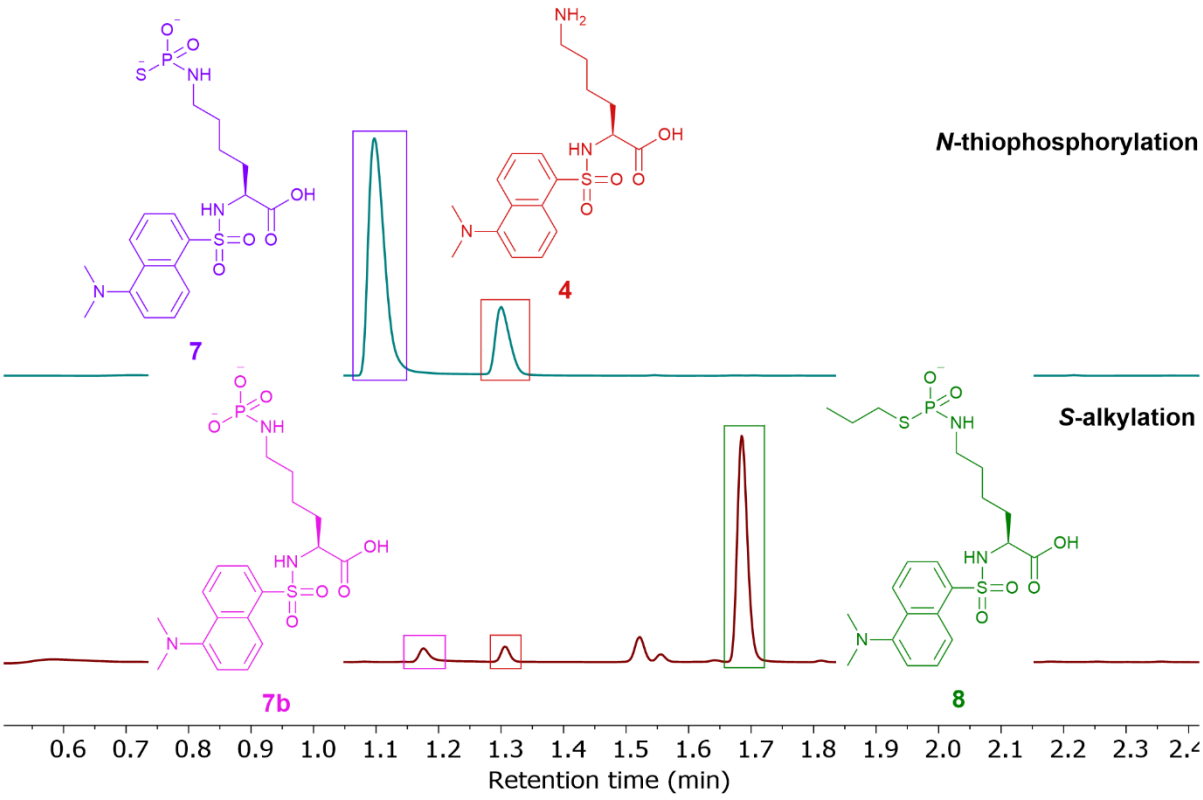
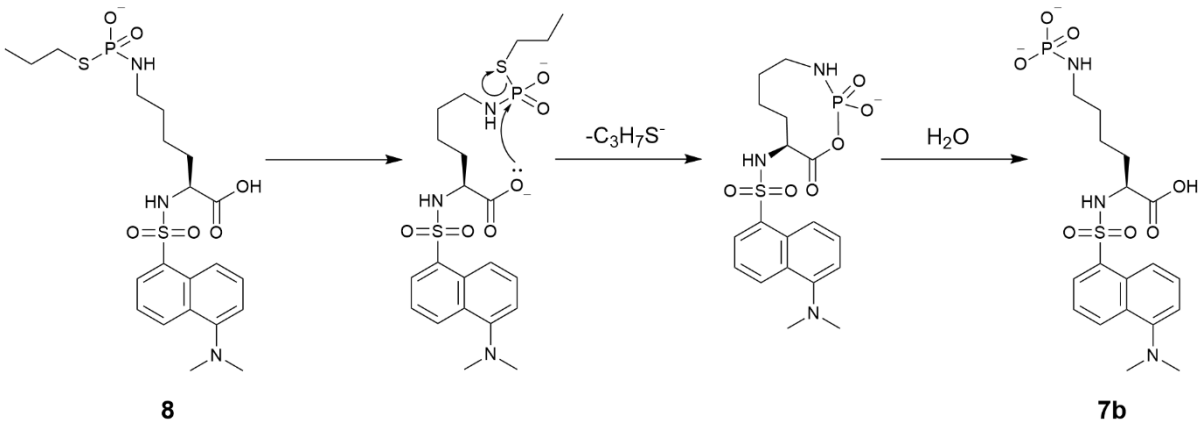


Figure 2.14 Total absorbance chromatographs from LC-MS analysis after *N*-thiophosphorylation of α-*N*-dansyl-L-lysine **4** (*m/z* 380 Da) to give ε-*N*-thiophosphoramidate **7** (*m/z* 476 Da), and *S*-alkylation of **7** with 1-iodopropane to give *S*-propyl ε-*N*-thiophosphoramidate **8** (*m/z* 518 Da). Small quantity of ε-*N*-phosphoramidate **7b** (*m/z* 460 Da) seen as the main by-product during the *S*-alkylation step.

Fortunately, the conversion to *S*-propyl ε-*N*-thiophosphoramidate **8** was much greater than that seen for the *S*-alkylation of thiophosphorylated phenylalanine done previously by Trmčić (discussed in section 1.4.2). The poor conversion to *S*-methyl α-*N*-thiophosphorylated phenylalanine (24%) was explained through neighbouring group-assisted decomposition of *S*-methyl α-*N*-thiophosphorylated phenylalanine with assistance coming from the carboxylate group (Scheme 1.16).[146] As the *N*-thiophosphorylation occurred at the α-NH_2 group of the amino acid, the resultant thiophosphoramidate was held in a convenient, nearby position for the proposed intramolecular decomposition to occur via a 5-membered ring intermediate.

However, *S*-propyl ε-*N*-thiophosphoramidate **8** has a far greater distance between the thiophosphoramidate and carboxylate groups than *S*-methyl α-*N*-thiophosphorylated phenylalanine. This distance largely explains the relative lack of decomposition of **8** to **7b** through reduced likelihood of intramolecular cyclisation via a medium-sized ring transition state. Another factor to consider for the relative lack of decomposition from **8** to **7b** is the nature of the proposed heterocyclic intermediates. The 5-membered ring intermediate for the decomposition of *S*-methyl α-*N*-thiophosphorylated phenylalanine is more energetically favourable than its 9-membered counterpart

(shown in Scheme 2.8). If the same energies are assumed for heteroatomic rings as are seen for cycloalkane rings, the strain energy of a 5-membered ring is 6.2 kcal mol^{-1}, but is 12.6 kcal mol^{-1} for a 9-membered ring.[160] Along with this, the entropy lost cyclising pentane to cyclopentane is -13.3 cal mol^{-1} K^{-1} and the entropy difference becomes less favourable as rings become bigger, with 4.5 cal mol^{-1} K^{-1} taken as the representative entropy lost for each bond rotation frozen.[160, 161]



Scheme 2.8 Proposed mechanism for decomposition of *S*-propyl ε-*N*-thiophosphoramidate **8**, via a 9-membered intermediate ring, to ε-*N*-phosphoramidate **7b** based on the decomposition route for *S*-alkylated *N*-thiophosphorylated phenylalanine suggested by Trmčić.[146]

2.2.1.1. Purification and isolation of *S*-propyl ε-*N*-thiophosphoramidate **8**

The crude *S*-alkylation mixture obtained in section 2.2.1 was purified by reverse phase flash chromatography, and the structure of isolated species **8** was confirmed by LC-MS (Figure 2.15) and ^{1}H-coupled ^{31}P NMR analyses (Figure 2.16).

A minor impurity was detected by LC-MS analysis with a retention time of 2.13 min with an *m/z* 291 Da, making up <1% of the sample (by LC-UV-vis detection).

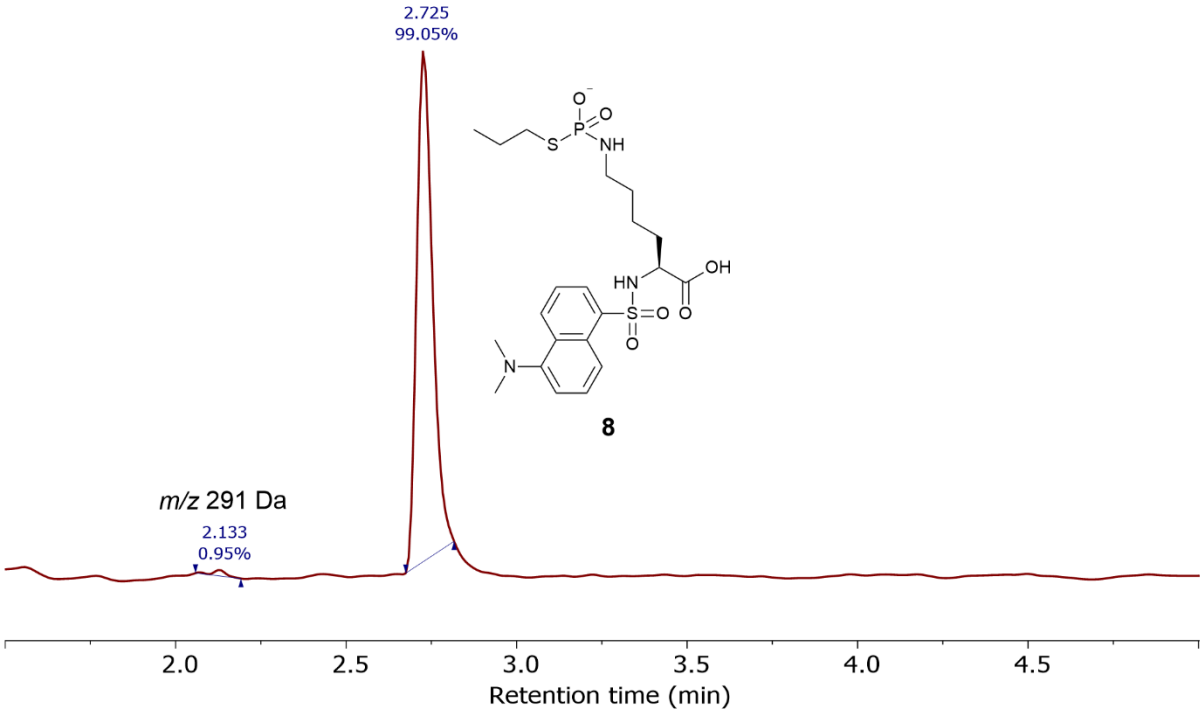


Figure 2.15 Total absorbance chromatograph from LC-MS analysis of purified *S*-propyl ε-*N*-thiophosphoramidate **8**.

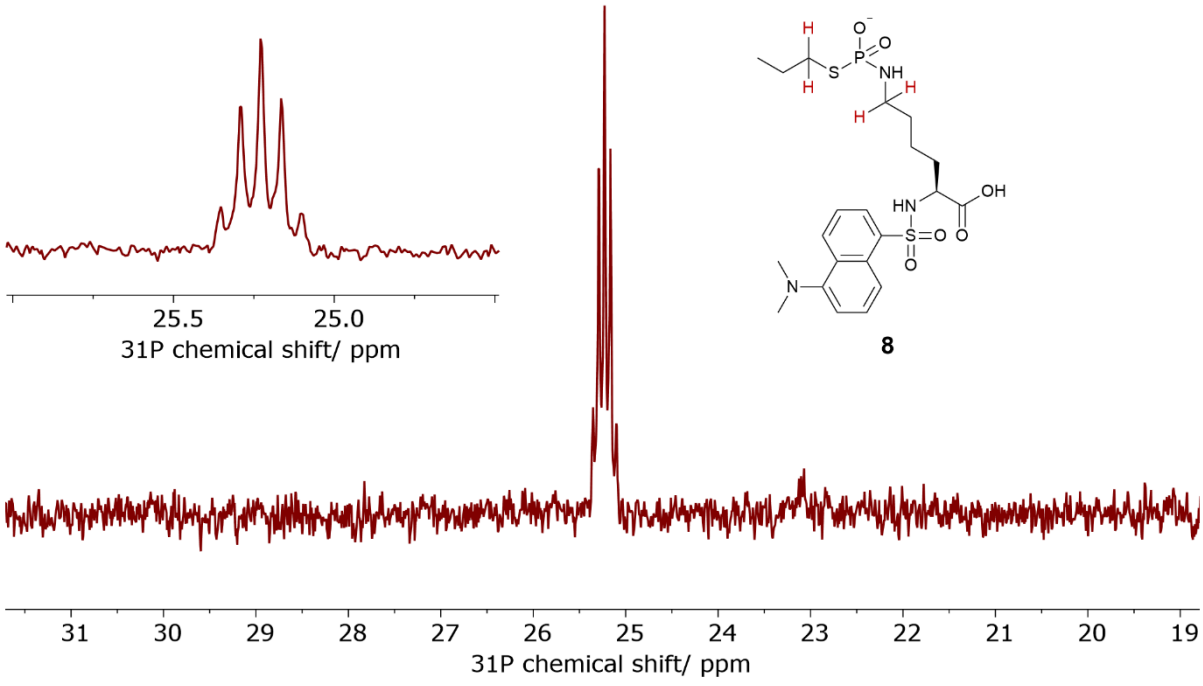


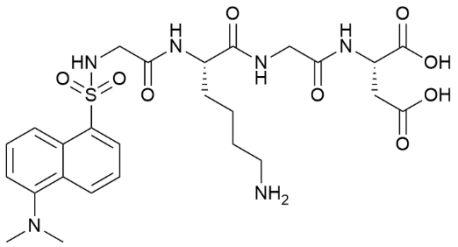
Figure 2.16 ^{1}H-coupled ^{31}P NMR spectrum following the purification of *S*-propyl ε-*N*-thiophosphoramidate **8**. Zoomed in section between 24.5 and 26.0 ppm shows the apparent quintet at 25.23 ppm in greater detail. Structure of **8** indicates protons coupling to phosphorus atom to generate the apparent quintet signal shown. Sample dissolved in D_2O.

The ^{31}P NMR spectrum shows an apparent quintet signal at 25.23 ppm, indicative of *S*-propyl ε-*N*-thiophosphoramidate **8**. The ^{31}P signal was expected to give a signal resembling a quintet as the single phosphorus atom is coupling to the 2 ε-methylene protons on the lysine side chain, as well as the 2 methylene protons on the propyl chain bound to the sulfur atom of the thiophosphoramidate group (all relevant protons are highlighted in red on the structure shown in Figure 2.16).

Having shown that ε-*N*-thiophosphoramidate **7** could be successfully 'trapped' with an alkyl group and isolated as an *S*-alkyl ε-*N*-thiophosphoramidate with >99% purity (by LC-UV-vis analysis), work moved on to *N*-thiophosphorylation of a larger peptide model.

2.3. Optimisation of aqueous *N*-thiophosphorylation on α-*N*-dansyl-labelled tetrapeptide model 9

As stated in section 2.1, restrictions in movement between laboratories due to the COVID-19 pandemic prevented the synthesis of peptides in the initial stages of the programme, as the only peptide synthesiser available was in a different research laboratory. Therefore, the α-*N*-dansyl-labelled tetrapeptide model **9** was designed (Figure 2.17) and ordered from Thermo Fisher Scientific.



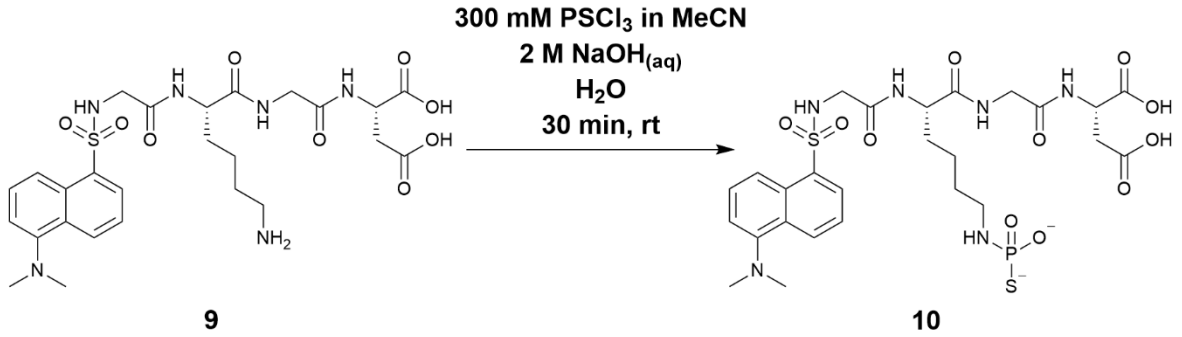
9

Figure 2.17 α-*N*-dansyl-labelled tetrapeptide model **9**.

The dansyl label was incorporated to aid with the use of LC-MS analysis with DAD. As mentioned in sections 2.1.2 and 2.2, LC-MS analysis incurs a significantly lower material cost compared to NMR spectroscopic analysis, which is why LC-MS was chosen as the primary analysis method.

The tetrapeptide was designed to allow the ε-NH_2 of the lysine side chain to be as unhindered as possible, hence the two glycine residues either side of lysine. The aspartic acid residue was included at the *C*-terminus to aid in peptide solubility because the carboxylic acid groups were expected to exist in their conjugate base forms under the basic conditions of our conjugation processes.

The general reaction protocol was the same as for the *N*-thiophosphorylations carried out on α-*N*-dansyl-L-lysine **4**.



Scheme 2.9 *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** to give ε-*N*-thiophosphopeptide **10**.

Since a starting point for optimised aqueous *N*-thiophosphorylation conditions had already been established on α-*N*-dansyl-L-lysine **4**, optimisation on tetrapeptide **9** was limited to a series of experiments varying the number of equivalents of $PSCl_3$ used (adjusting the number of equivalents of $NaOH_{(aq)}$ to accommodate the excess of acid forming $PSCl_3$).

When the optimal conditions found for *N*-thiophosphorylation of α-*N*-dansyl-L-lysine **4** were tested on tetrapeptide **9**, the conversion to ε-*N*-thiophosphopeptide **10** was lower than expected (82% conversion to **10** compared to 92% conversion to ε-*N*-thiophosphoramidate **7**). The details and conversions of this experiment and the rest of the optimisation series carried out on tetrapeptide **9** are displayed in Table 2.5 and Figure 2.18.

Table 2.5 Data from *N*-thiophosphorylation series of α-*N*-dansyl tetrapeptide **9** to ε-*N*-thiophosphopeptide **10**. Conversions calculated from area under relevant peaks seen in LC-MS total absorbance chromatographs.

Equivalents of $PSCl_3$	Equivalents of $NaOH_{(aq)}$	Conversion from 9 to 10/ %
1.0	7.5	82
2.0	14.7	89
7.5	35.7	95

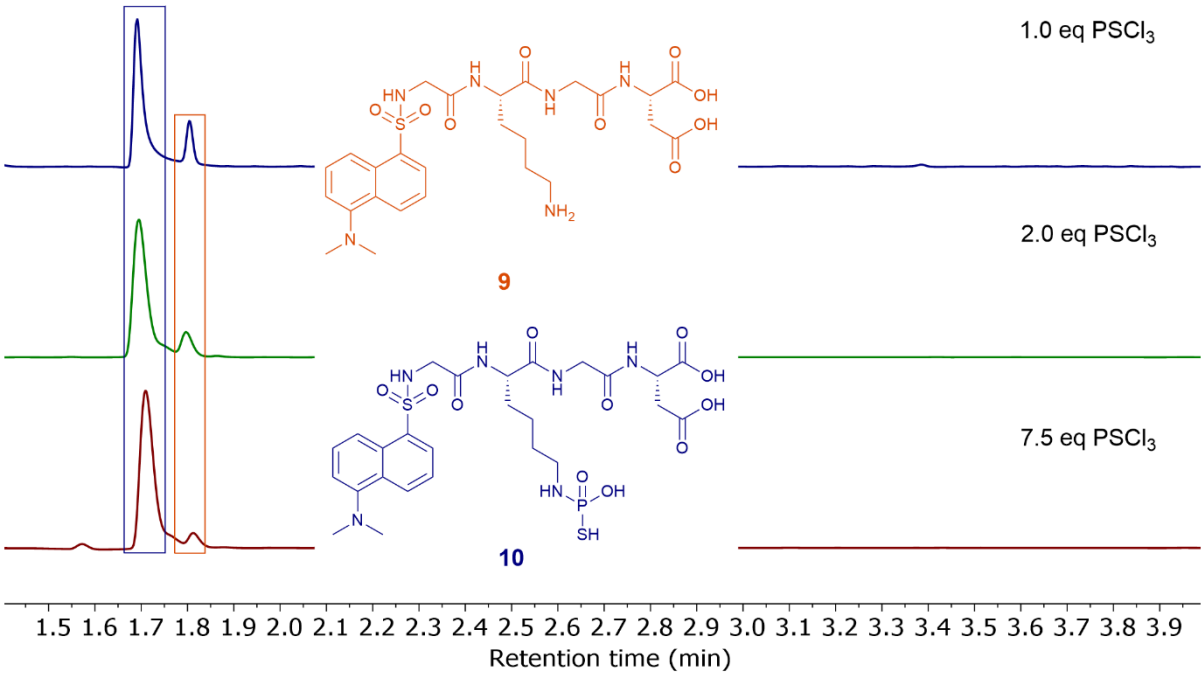


Figure 2.18 LC-MS total absorbance chromatographs following *N*-thiophosphorylation of α-*N*-dansyl tetrapeptide **9** (*m/z* 609 Da) to give ε-*N*-thiophosphopeptide **10** (*m/z* 705 Da).

Optimisation using tetrapeptide **9** (9.51 mM solution in water) resulted in conversion to ε-*N*-thiophosphopeptide **10** of up to 95% using 7.5 eq 300 mM $PSCl_3$ solution and 35.7 eq 2 M NaOH solution. As with the optimisation series run on α-*N*-dansyl-L-lysine **4** in section 2.2, the experimental scale was too small (< 25 µL) to allow for pH checks before and after *N*-thiophosphorylation, but the mechanism of *N*-thiophosphorylation on **9** is believed to be the same as for *N*-thiophosphorylation of **4** (i.e. route (2) in Scheme 2.6).

Having been optimised on a model tetrapeptide, the scope of aqueous *N*-thiophosphorylation has been successful expanded towards the synthesis of peptide thiophosphoramidates.

2.4. Summary

This project began with the synthesis and use of a single amino acid model, α-*N*-dansyl-L-lysine **4**, for the optimisation of aqueous *N*-thiophosphorylation. The synthesis of model lysine **4** was achieved through the modification of an orthogonally protected lysine residue (**1**) by first replacing the Boc protecting group with a dansyl group at the α-NH_2 position. The Fmoc protecting group at the ε-NH_2 position was then removed to give α-*N*-dansyl-L-lysine **4** as the desired model. The dansyl group was introduced to give the lysine model a strong chromophore to facilitate the use of LC-MS analysis with DAD. Purification of α-*N*-dansyl-L-lysine **4** proved a challenge, but two reverse phase chromatography

runs (with triethylammonium acetate solution [pH 4.7]: acetonitrile gradient as the mobile phase) resulted in the isolation of 99.7% homogeneous product (by LC-UV-vis analysis).

Optimisation of aqueous *N*-thiophosphorylation began on α-*N*-dansyl-L-lysine **4**, based on a method previously established within the group.[152] Two short experimental series were devised in which the number of equivalents of $PSCl_3$ and $NaOH_{(aq)}$ were varied in order to find the optimal conditions for *N*-thiophosphorylation on **4**. Several of the optimisation experiments resulted in conversions from α-*N*-dansyl-L-lysine **4** to ε-*N*-thiophosphoramidate **7** of > 90%, with the greatest conversion seen with 1.2 eq $PSCl_3$ and 8.0 eq $NaOH_{(aq)}$ (97% by LC-UV-vis analysis). A scaled-up thiophosphorylation reaction (2.5 µmol to 275 µmol **4**) also led to high conversions on a 30 min time scale at room temperature, with 80% conversion to **7** using 1.0 eq $PSCl_3$ and 7.0 eq $NaOH_{(aq)}$.

S-alkylation of ε-*N*-thiophosphoramidate **7** with 1-iodopropane presented a conversion of up to 95% from ε-*N*-thiophosphoramidate **7** to *S*-propyl ε-*N*-thiophosphoramidate **8**, with minimal decomposition to ε-*N*-phosphoramidate **7b** over the course of the reaction. While not pursued within this project due to time constraints, further work may be done on alternative functionalisations of mono-thiophosphorylated peptides at the introduced *S*-atom, such as PEGylation and lipidation.

Finally, aqueous *N*-thiophosphorylation of an α-*N*-dansyl-labelled tetrapeptide model (**9**) was optimised using the results from the optimisation on α-*N*-dansyl-L-lysine **4** as a starting point. Again, a short series of experiments was devised varying the number of equivalents of $PSCl_3$ used, but this time the amount of $NaOH_{(aq)}$ used was adjusted to accommodate the large excess of $PSCl_3$, rather than testing several different numbers of $NaOH_{(aq)}$ equivalents. 7.5 eq $PSCl_3$ and 35.7 eq $NaOH_{(aq)}$ were found to give a conversion of 95% from tetrapeptide **9** to ε-*N*-thiophosphopeptide **10**.

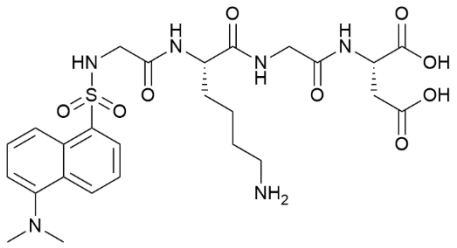
While aqueous *N*-thiophosphorylation was optimised on a tetrapeptide model to consistently give > 90% conversions to thiophosphoramidate, no information on the selectivity of *N*-thiophosphorylation had been gathered. Based on pK_{aH} values of amino acid α-amine groups, *N*-thiophosphorylation is still expected to preferentially occur at the lysine side chain. Data collected on the selectivity of *N*-thiophosphorylation on unprotected model peptides are presented in Chapter 3.

3. Selectivity Investigation

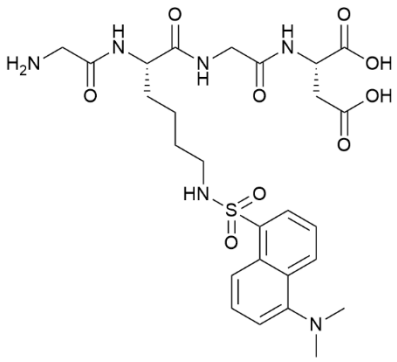
Having optimised aqueous *N*-thiophosphorylation on an α-*N*-dansyl-labelled tetrapeptide model (section 2.3), work moved towards investigating the selectivity of the *N*-thiophosphorylation reaction. In the previous chapter, all thiophosphorylations were conducted on substrates with only one available thiophosphorylation site: the ε-NH_2 of a lysine derivative. While conversions over 95% were obtained on both lysine-based model **4** and tetrapeptide model **9**, these studies gave no insight into the chemoselectivity of the thiophosphorylation reaction between the lysine ε-NH_2 group and other nucleophilic functionalities found on unprotected peptide substrates.

3.1. Aqueous *N*-thiophosphorylation on ε-*N*-dansyl-labelled tetrapeptide model 11

In order to obtain quantitative data on how the optimised aqueous thiophosphorylation would proceed with unprotected lysine-containing peptides, a comparison experiment was devised using both the ε-*N*-dansyl-labelled and α-*N*-dansyl-labelled isomers of the tetrapeptide model (**11** and **9** respectively).



9



11

Figure 3.1 Structure comparison of α-*N*-dansyl-labelled tetrapeptide **9** and ε-*N*-dansyl-labelled tetrapeptide **11**.

Figure 3.1 shows that the same peptide sequence (GKGD) was used for ε-*N*-dansyl labelled tetrapeptide **11** as for α-*N*-dansyl-labelled tetrapeptide **9** to keep as many variables as possible consistent for the thiophosphorylation comparison.

Parallel *N*-thiophosphorylations of **11** and **9** were performed over 30 min at room temperature as indicated in Scheme 3.1 for **11** and Scheme 2.9 for **9**.

Scheme 3.1 *N*-thiophosphorylation of ε-*N*-dansyl-labelled tetrapeptide **11** to give ε-*N*-dansyl-labelled α-*N*-thiophosphopeptide **12**.

In each of the experiments, 300 mM $PSCl_3$ solution (10.0 eq) and 2 M NaOH solution (70.0 eq) were used. LC-MS DAD was the preferred analysis method rather than ^{31}P NMR spectroscopy due to the significantly lower material cost incurred. The *N*-thiophosphorylation results are displayed in Figure 3.2.

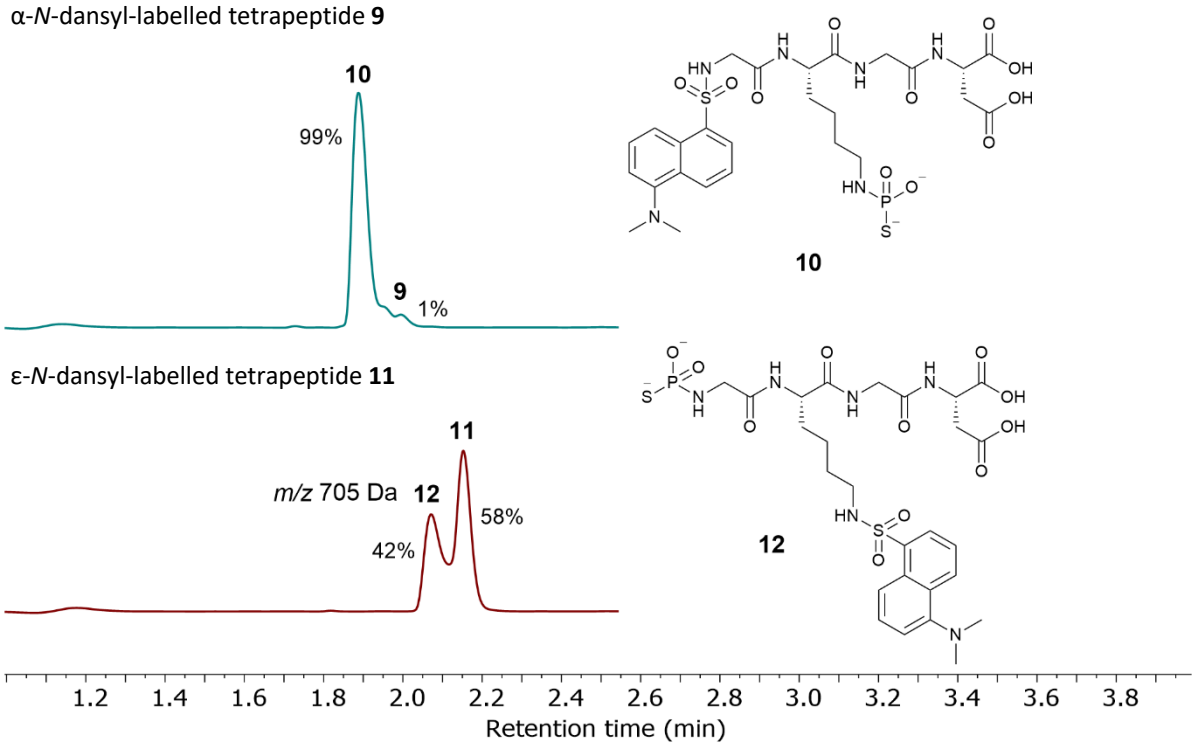


Figure 3.2 LC-MS total absorbance chromatographs following aqueous *N*-thiophosphorylations of α-*N*-dansyl tetrapeptide **9** (*m/z* 609 Da) to generate ε-*N*-thiophosphopeptide **10** (*m/z* 705 Da), and ε-*N*-dansyl tetrapeptide **11** (*m/z* 609 Da) to generate α-*N*-thiophosphopeptide **12** (*m/z* 705 Da). Concentration of solution of **9** – 9.5 mM; concentration of **9** in total reaction volume – 5.75 mM; concentration of solution of **11** – 4 mM; concentration of **11** in total reaction volume – 3.14 mM.

ε-*N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** essentially ran to completion over the 30 min reaction time, with a 99% conversion to thiophosphopeptide **10** (by LC-UV-vis analysis). α-*N*-

thiophosphorylation of ε-*N*-dansyl-labelled tetrapeptide **11**, however, resulted in 42% conversion to thiophosphopeptide **12** over the same time period.

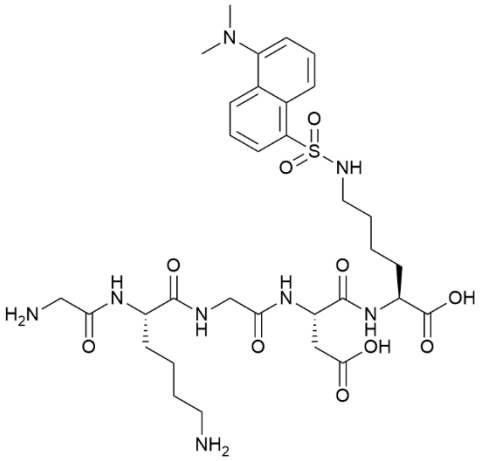
The limited thiophosphorylation of **11** compared to that of **9** may be explained by two factors. First, the pK_{aH} of the target amino groups differ (α-NH_2 of glycine is 9.60; ε-NH_2 of lysine is 10.53)[154]. In general, the nucleophilicity of a species increases with its basicity (i.e. the species pK_{aH}). In this case the ε-NH_2 group on the lysine residue would be more nucleophilic than the *N*-terminal glycine α-NH_2 group, thus leading to the reduction in thiophosphorylation at α-NH_2 compared to ε-NH_2. The second factor to consider is the steric environment surrounding the target amino group. The *N*-terminal α-NH_2 of **11** is in a more sterically hindered environment than the lysine ε-NH_2 of **9**, so there is less chance of thiophosphorylation at the *N*-terminus than at the lysine side chain.

This reduced conversion to thiophosphopeptide **12** compared to thiophosphopeptide **10** suggested that aqueous *N*-thiophosphorylation is selective for the formation of ε-*N*-thiophosphoramidates over α-*N*-thiophosphoramidates. However, as this was not a direct competition reaction (i.e. both peptide substrates present in equal amounts in the same reaction vessel), the results shown in Figure 3.2 cannot be used to accurately determine thiophosphorylation selectivity. α-*N*-dansyl-labelled tetrapeptide **9** and ε-*N*-dansyl-labelled tetrapeptide **11** were deemed to be unsuitable substrates for direct competition experiments as they are isomers, thus could not be distinguished readily by LC-MS analyses, which would be an essential part of the assay, due of the similar retention times of the substrates and products.

In an effort to work around this issue, additional mass spectrometry (MS) methods were explored for investigating the selectivity of *N*-thiophosphorylation by finding the preferred thiophosphorylation position between ε-NH_2 and α-NH_2.

3.2. Using tandem mass spectrometry to investigate thiophosphorylation position on dansyl-labelled pentapeptide model 13

In order to find the preferred thiophosphorylation target, a peptide substrate was designed with both an unprotected *N*-terminal amino group and an unprotected lysine side chain. In a continued effort to keep as many variables as possible the same, a dansyl-labelled pentapeptide was designed and purchased from Thermo Fisher Scientific (Figure 3.3).



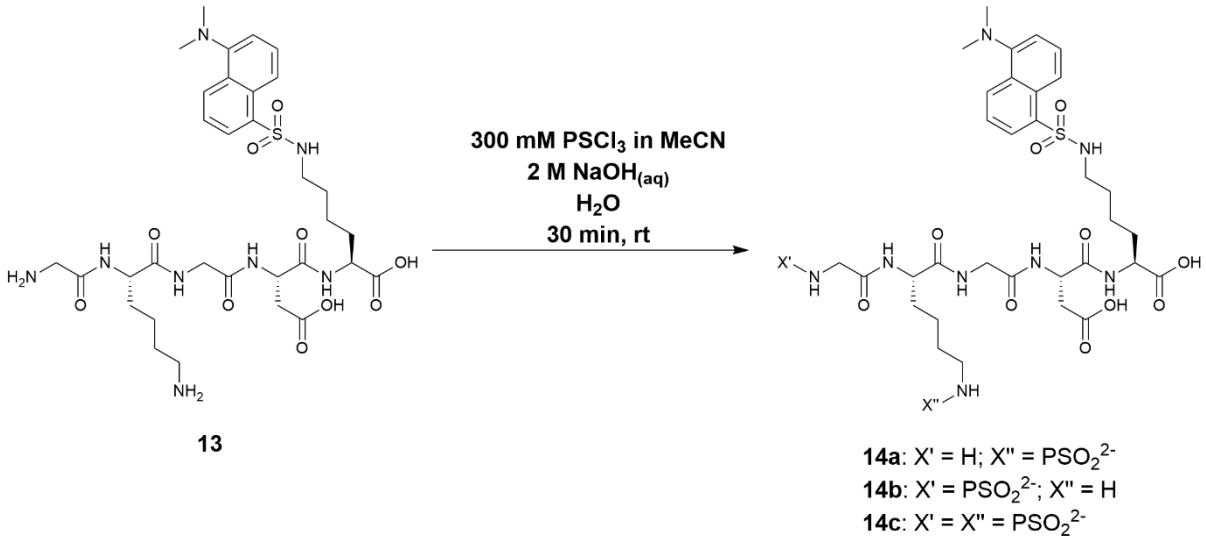
13

Figure 3.3 Structure of dansyl-labelled pentapeptide model **13**.

The sequence of pentapeptide **13** kept the same motif as dansyl-labelled tetrapeptide models **9** and **11**, but with a second lysine residue at the *C*-terminus. This second lysine residue was labelled with a dansyl group at the ε-NH_2 to facilitate the use of LC-MS DAD analysis.

3.2.1. *N*-thiophosphorylation of dansyl-labelled pentapeptide **13**

Dansyl-labelled pentapeptide **13** underwent thiophosphorylation under the same conditions as used for the thiophosphorylation of **11** in section 3.1 (10.0 eq $PSCl_3$ solution and 70.0 eq 2 M NaOH solution). As peptide **13** contained two unprotected amino groups, 3 thiophosphorylated peptide products were expected. The anticipated products were the mono-thiophosphoramidate species **14a** and **14b**, and di-thiophosphoramidate species **14c** (see Scheme 3.2).



Scheme 3.2 *N*-thiophosphorylation of dansyl-labelled pentapeptide **13** to give ε-*N*-thiophosphoramidate **14a**, α-*N*-thiophosphoramidate **14b**, and α,ε-di-thiophosphoramidate **14c**. Concentration of solution of **13** – 4 mM; concentration of **13** in total reaction volume – 3.14 mM.

LC-MS analysis of the crude thiophosphorylation mixture after the 30 min reaction time suggested that peptide **13** had been entirely consumed. Figure 3.4 shows the total absorbance chromatograph obtained from LC-MS analysis. The single major peak at 2.042 min corresponds to species with *m/z* 833 Da – the molecular mass of mono-thiophosphoramidates **14a** and **14b**. There was no evidence, by either LC-UV-vis or mass extracted ion count (EIC) analyses, of the formation of di-thiophosphoramidate **14c**.

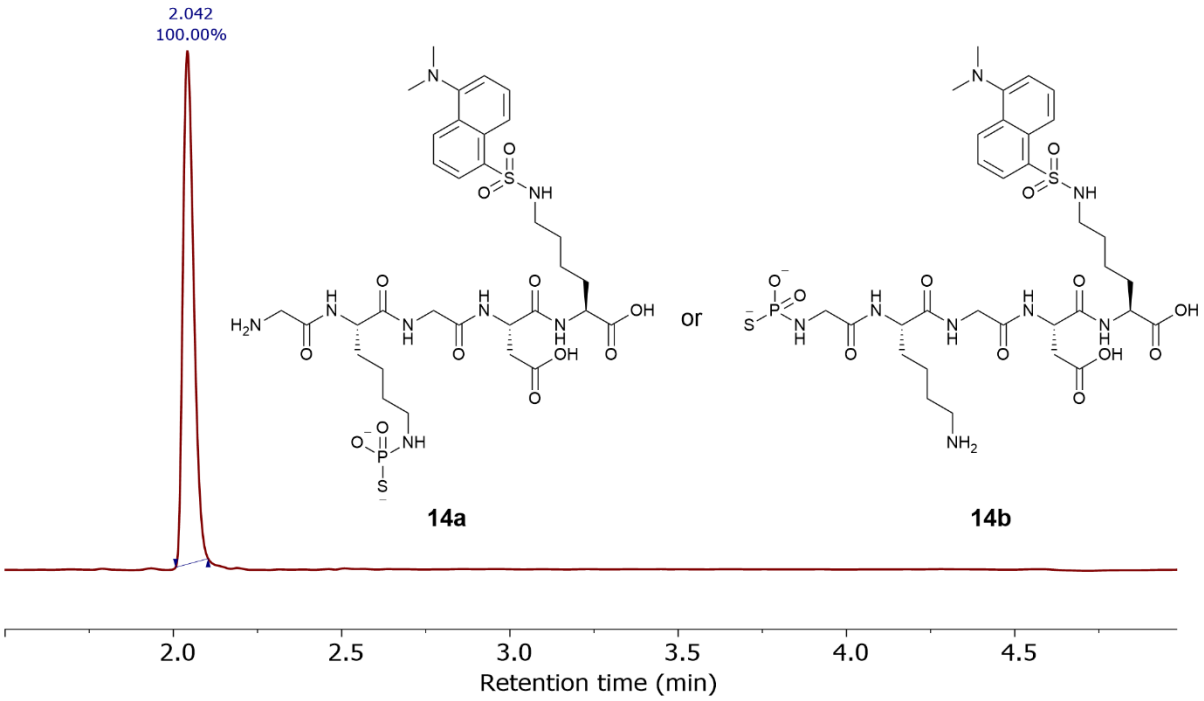


Figure 3.4 LC-MS total absorbance chromatograph following *N*-thiophosphorylation of dansyl-labelled pentapeptide **13** (*m/z* 737 Da) to give thiophosphopeptide **14a** or **14b** (*m/z* 833 Da).

LC-MS showed that thiophosphorylation of peptide **13** had resulted in complete conversion to thiophosphoramidate **14a** and/or **14b**, however, the resulting single peak in the product chromatogram give no indication of the site(s) of thiophosphorylation or their relative proportions. Thus, it was hypothesised that tandem mass spectrometry (MSMS) analysis may be able to determine the dominant thiophosphorylation position based on the fragmentation patterns generated. For example, if the major fragment (daughter ion) indicated the loss of *N*-terminal glycine from the thiophosphoramidate it could then be assumed that thiophosphorylation occurred at predominantly at ε-NH_2 of the lysine side chain.

3.2.1.1. Tandem MS analysis of crude suspected thiophosphoramidates **14a** and/or **14b** under acidic conditions

The MSMS spectrum showing the daughter ions of thiophosphoramidate **14a** and/or **14b** ions (*m/z* 833 Da using positive ion ESI) is shown in Figure 3.5.

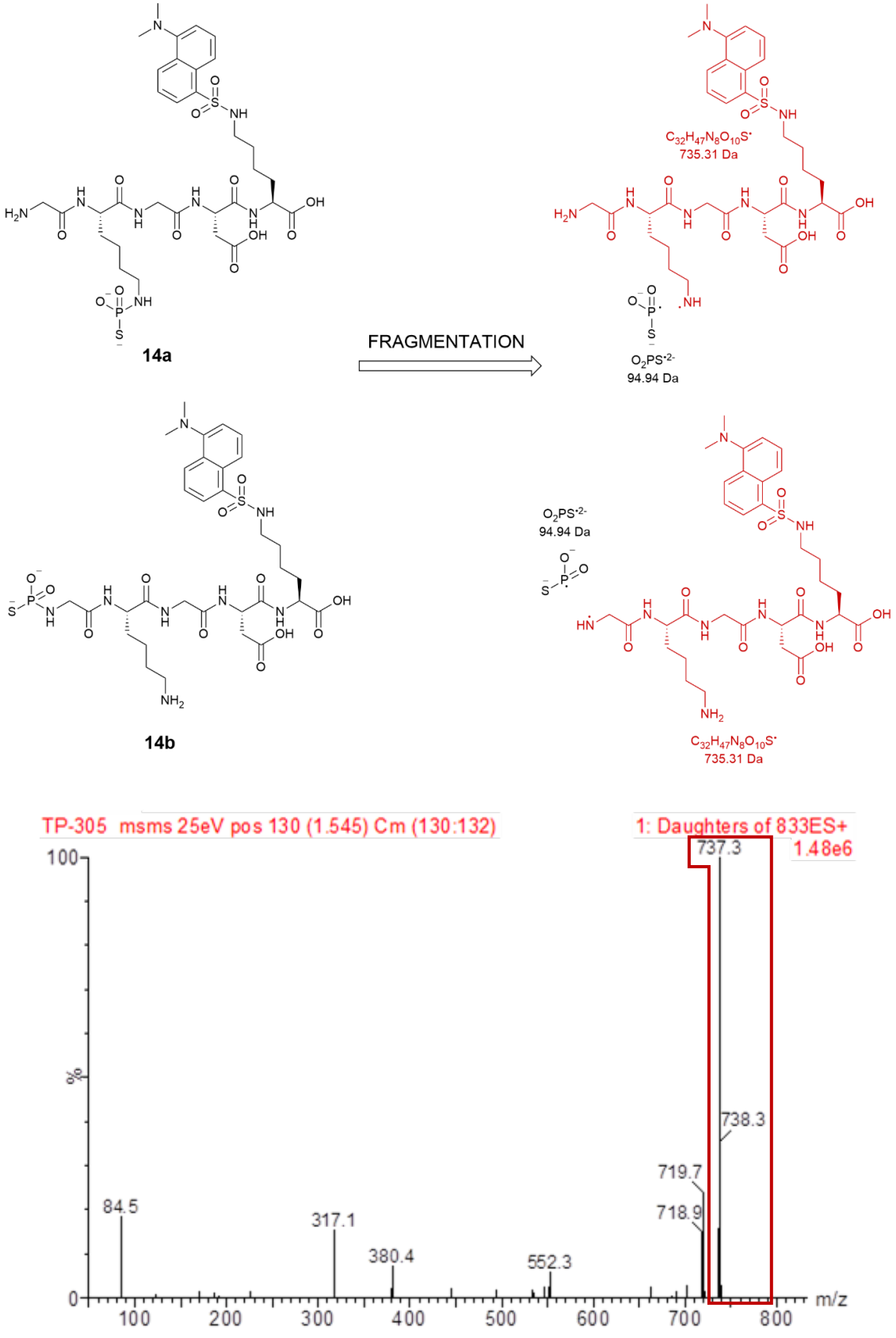


Figure 3.5 MSMS spectrum of daughter ions from thiophosphoramidates **14a** and/or **14b** (*m/z* 833 Da [ES+]). Example fragmentation to pentapeptide **13** from **14a** and/or **14b** at the P-N bond is illustrated above the MSMS spectrum. The red fragment structure corresponds to the MS signal outlined in red.

Unfortunately, the major daughter ion generated through MSMS analysis of suspected thiophosphoramidates **14a** and **14b** was dansyl-labelled pentapeptide **13** (*m/z* 737 Da), indicating that the thiophosphate group was the first to fragment off the parent ion(s).

The observed fragmentation pattern can be explained by the lability of the P-N bond of a thiophosphoramidate in acidic conditions. As 0.1% formic acid solution is a commonly used solvent for LC-based analytical methods, the significant loss of thiophosphate may have been impacted by the conditions used during the LC-MSMS experiment.

To minimise the potential impact of the LC solvent system used, a basic buffer solution (pH 8) was prepared and used for a follow-up LC-MSMS experiment.

3.2.1.2. Tandem MS analysis of crude suspected thiophosphoramidates **14a** and/or **14b** under basic conditions

An ammonium bicarbonate buffer (pH 8) was employed as a mobile phase solvent for the second LC-MSMS experiment. The basic LC-MSMS experiment was run using the same thiophosphoramidate mixture synthesised in section 3.2.1 (LC-MS of thiophosphoramidate **14a** and/or **14b** shown in Figure 3.4).

Figure 3.6 shows the resultant MSMS spectrum. It should be noted that the thiophosphoramidate molecular ion mass is 831 Da under basic conditions as it is formed through negative ion ESI.

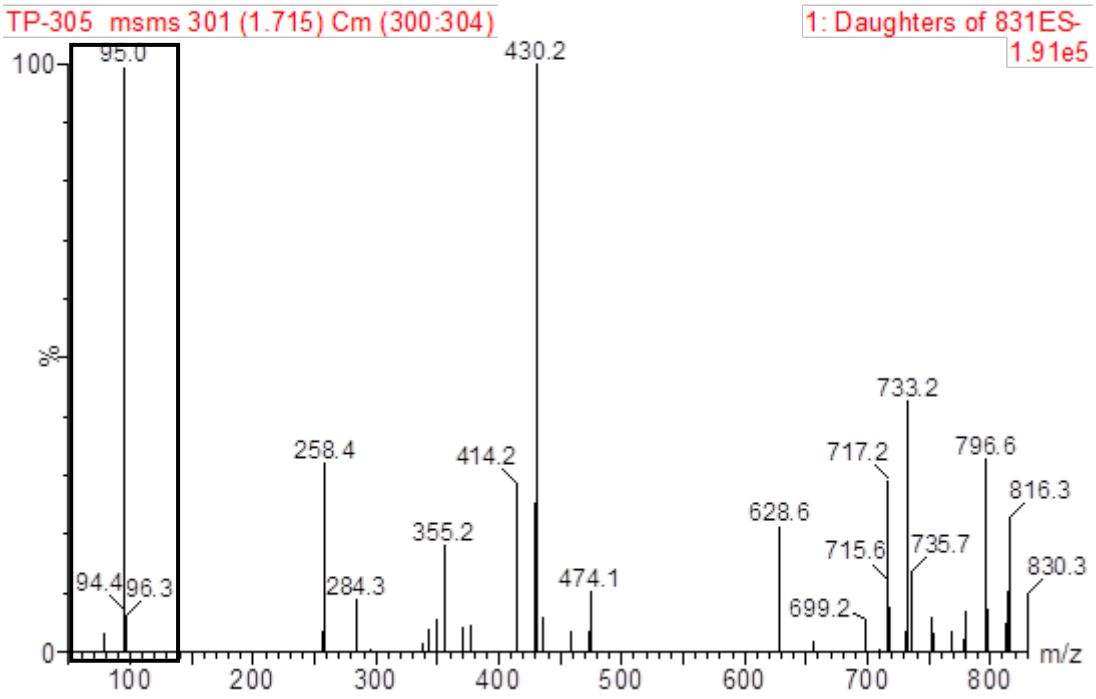


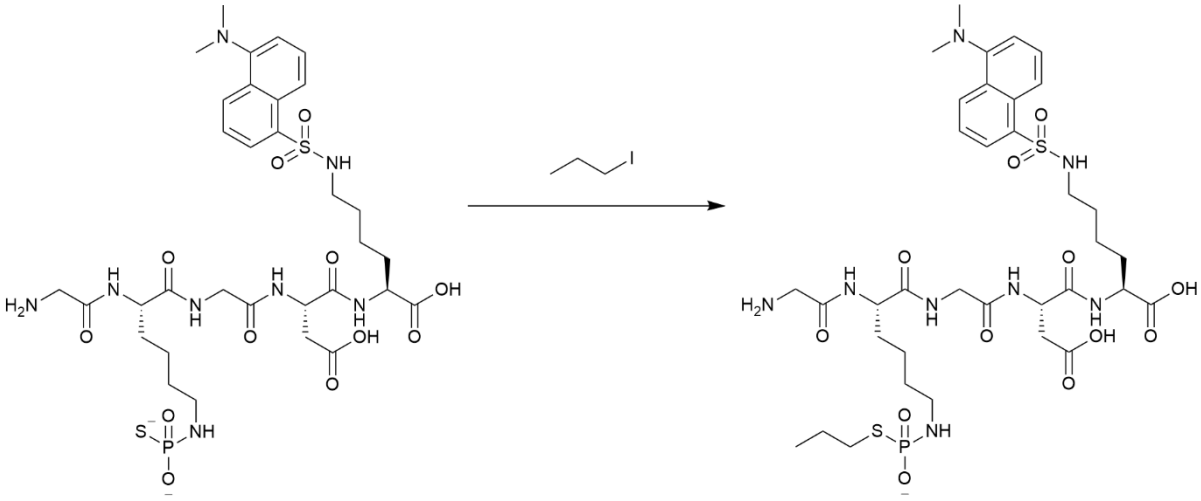
Figure 3.6 MSMS spectrum of daughter ions from thiophosphoramidate **14a** or **14b** (*m/z* 831 Da [ES-]). For an example of how parent ion(s) **14a** or **14b** fragments to give inorganic thiophosphate, see Figure 3.5. The black fragment seen in Figure 3.5 corresponds to the MS signal outlined in black.

As seen in section 3.2.1.1, the daughter ions generated suggest that the thiophosphate group (*m/z* 95 Da [ES-]) was a major fragment released from the parent ion(s) (*m/z* 831 Da).

Despite using basic LC conditions, the P-N bond of the thiophosphoramidate was still proving to be labile under LC-MSMS conditions. Since alkylation at the anionic sulfur of a thiophosphoramidate has been shown to extend thiophosphoramidate half-life in acidic conditions,[146] one-pot *N*-thiophosphorylation and *S*-alkylation was performed in an attempt to stabilise the N-P thiophosphoramidate bond.

3.2.2. *S*-alkylation of thiophosphopeptides **14a-c**

The one-pot, two-step *N*-thiophosphorylation and *S*-alkylation method introduced in section 2.2.1 was adjusted for use on pentapeptide **13** (10.0 eq $PSCl_3$ solution and 70.0 eq 2 M NaOH solution for *N*-thiophosphorylation, as in section 3.2.1; 23.3 eq 1-iodopropane solution for *S*-alkylation). The thiophosphorylation scheme can been seen in Scheme 3.2. Scheme 3.3 shows one of the possible *S*-alkylation products, starting from ε-*N*-thiophosphoramidate **14a** to generate *S*-propyl ε-*N*-thiophosphoramidate **15a**.



Scheme 3.3 *S*-alkylation of ε-*N*-thiophosphoramidate **14a** to *S*-propyl ε-*N*-thiophosphoramidate **15a**.

S-alkylation of α-*N*-thiophosphoramidate **14b** gave *S*-propyl α-*N*-thiophosphoramidate **15b**, and *S*-alkylation of di-thiophosphoramidate **14c** gave *bis*-*S*-propyl di-thiophosphoramidate **15c**.

The one-pot *N*-thiophosphorylation and *S*-alkylation of pentapeptide **13** resulted in significant conversion (47%) to *S*-propyl thiophosphoramidate **15a** and/or **15b**, and 29% conversion to *bis*-*S*-propyl di-thiophosphoramidate **15c**. The LC-MS total absorbance chromatograph obtained from the crude *S*-alkylation mixture is shown in Figure 3.7, with the structures of *S*-propyl thiophosphopeptides **15a-c** placed alongside their corresponding peaks.

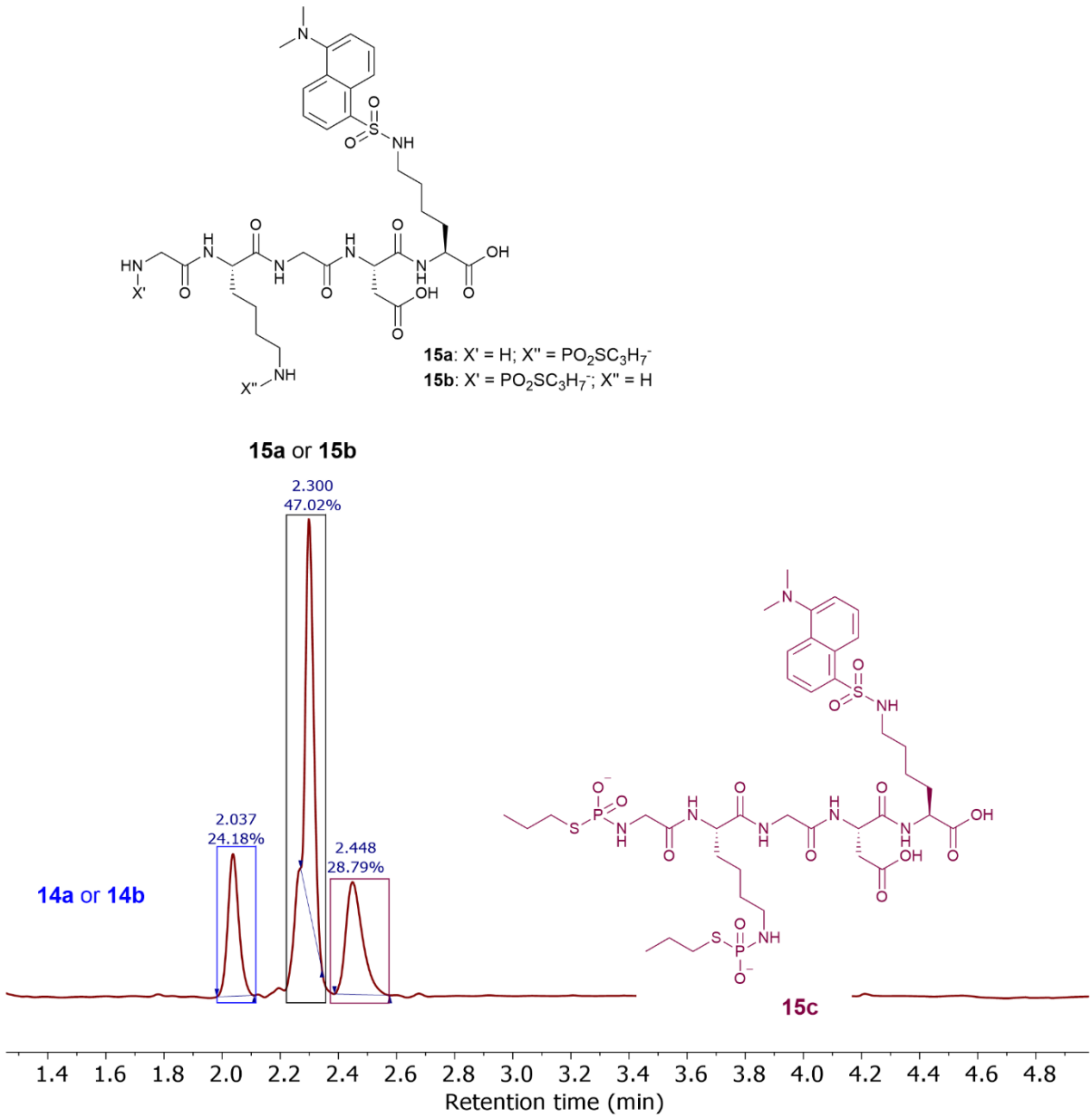


Figure 3.7 LC-MS total absorbance chromatograph following one-pot *N*-thiophosphorylation and *S*-alkylation of peptide **13** to generate *S*-alkylated thiophosphopeptides **15a-c** (**15a-b**: *m/z* 875 Da; **15c**: *m/z* 1013 Da) via thiophosphopeptides **14a-c** (**14a-b**: *m/z* 833 Da; **14c**: *m/z* 929 Da).

Since tandem MS experiments involve the separation of ions by mass before fragmentation, there was no need to isolate *S*-propyl thiophosphoramidate **15a** and/or **15b** prior to MSMS analysis.

3.2.2.1. Tandem MS analysis of crude *S*-alkylated thiophosphopeptides **15a-c**

Having set the tandem MS experiment to isolate ions of *m/z* 875 Da (molecular mass of both *S*-propyl thiophosphoramidates **15a** and **15b**) before fragmentation, Figure 3.8 shows the MSMS spectrum generated for 875 Da (ES+) daughter ions.

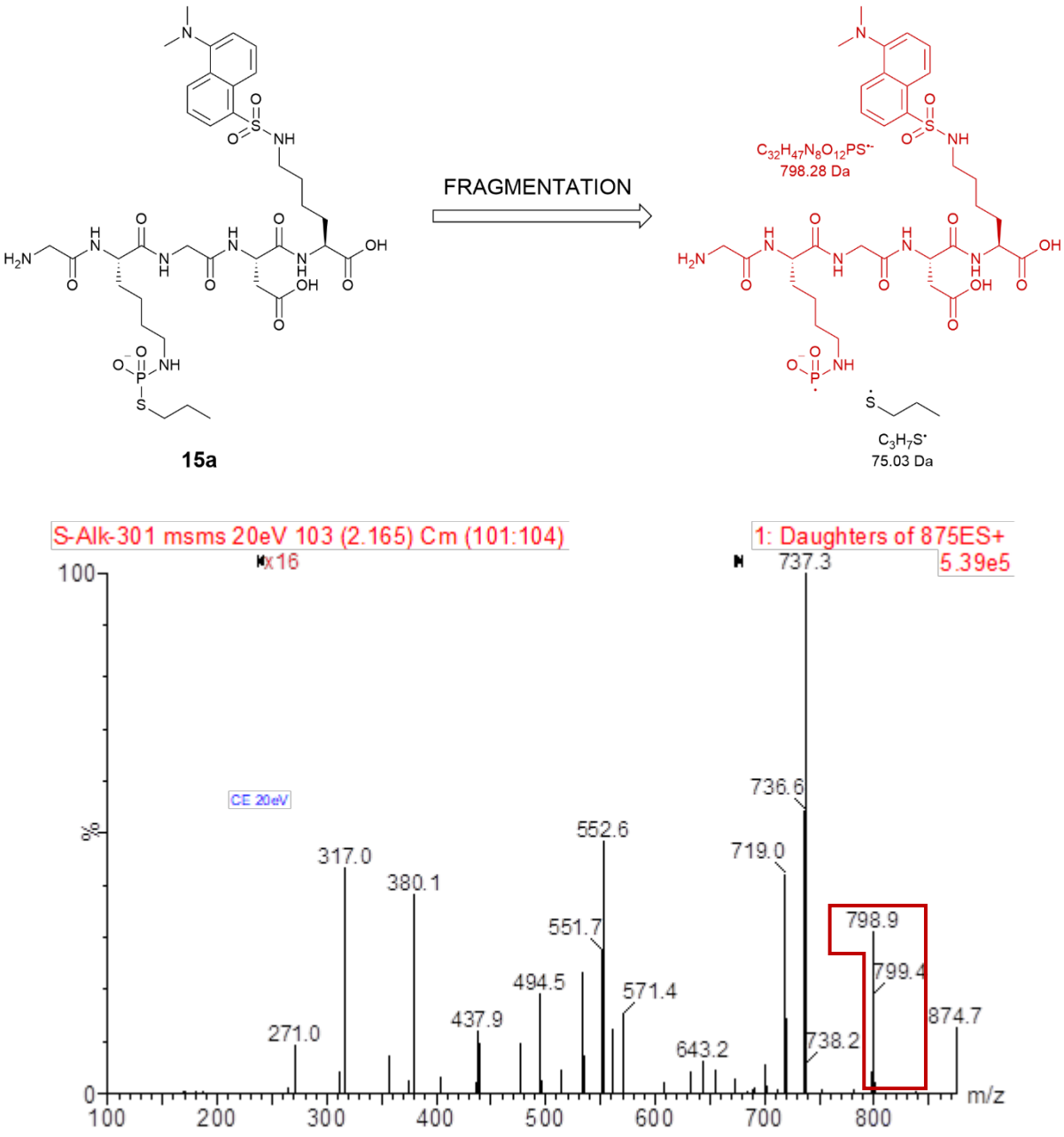


Figure 3.8 MSMS spectrum of daughter ions from *S*-propyl *N*-thiophosphoramidates **15a** and/or **15b** (*m/z* 875 Da). Example fragmentation of **15a** at the P-S bond is illustrated above the MSMS spectrum (P-S bond breakage would also be seen for **15b**). The red fragment structure corresponds to the MS signal outlined in red.

Once again, as seen in section 3.2.1.1, a major daughter ion seen is that of dansyl-labelled pentapeptide **13** (*m/z* 737 Da – fragmentation at P-N bond illustrated in Figure 3.5), indicating the loss of the *S*-alkylated thiophosphate group.

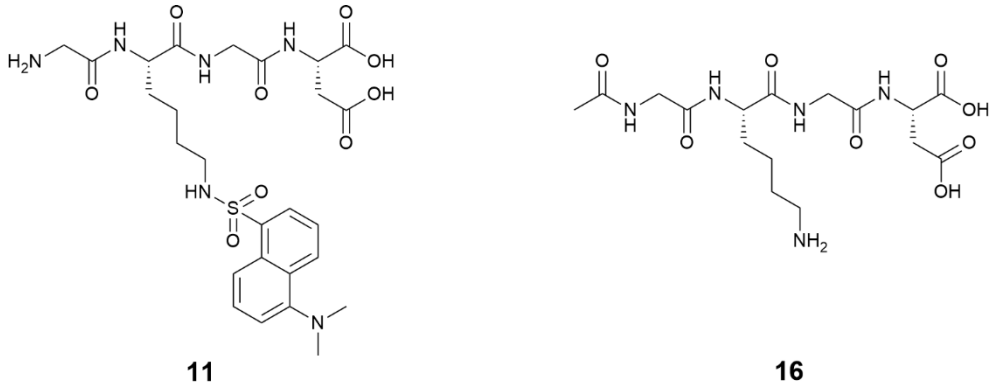
In addition, as illustrated in Figure 3.8, the first loss seen in the MSMS spectrum is 1-propanethiol (875 Da down to 799 Da – Δm 76 Da). The loss of 1-propanethiol suggests that not only is the labile P-N thiophosphoramidate bond susceptible to fragmentation, as seen in Figures 3.5 and 3.6, but the P-S bond is susceptible to fragmentation during tandem MS analysis.

Based on the MSMS results gathered both before and after thiophosphoramidate *S*-alkylation, it was decided that the N-P bond was too labile, even under soft ionisation conditions and after a move to less acidic LC conditions, to locate the position of thiophosphorylation using mass spectrometry techniques. Instead, direct competition reactions between an α-*N*-acetyl protected tetrapeptide and an ε-*N*-dansyl-labelled tetrapeptide were explored as a means of calculating the selectivity of *N*-thiophosphorylation between ε- and α-amine groups by LC-MS DAD analysis.

3.3. Competition reactions between α-*N*-acetyl-protected tetrapeptide 16 and ε-*N*-dansyl-labelled tetrapeptide 11 to study *N*-thiophosphorylation selectivity

To ascertain whether *N*-thiophosphorylation was selective for the ε-NH_2 or α-NH_2 position through a competition reaction, equal quantities of two different, selectively protected, peptides were required. As LC-MS DAD analysis was still preferred over NMR spectroscopy due to the lower material costs incurred, the two peptide models also had to have different molecular masses to facilitate their identification in LC-MS spectra.

The requirement for peptides of different masses led to the decision to use ε-*N*-dansyl-labelled tetrapeptide **11** (introduced in section 3.1) and a new model peptide – α-*N*-acetyl protected tetrapeptide **16**.



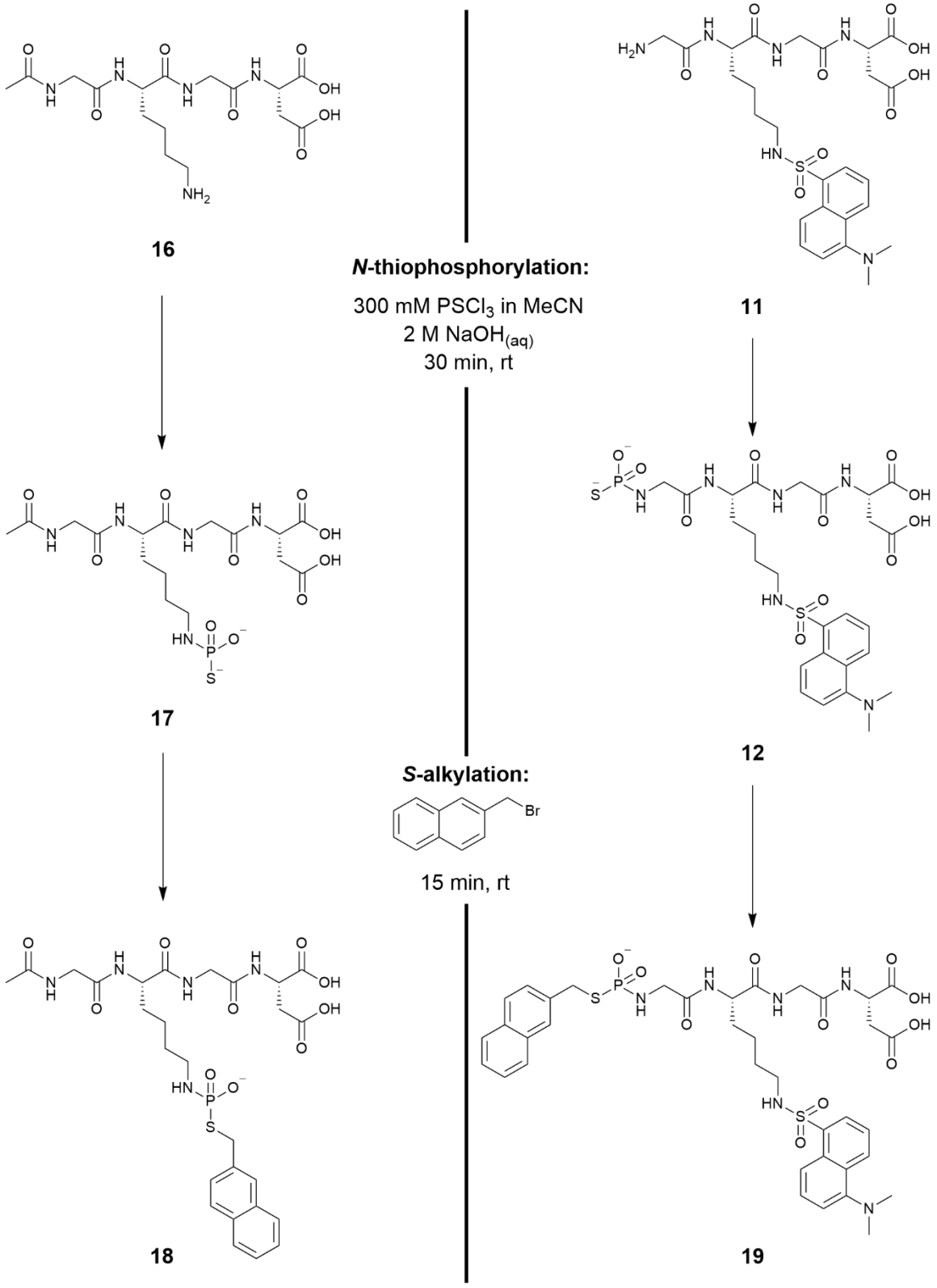
11 **16**

Figure 3.9 Structures of ε-*N*-dansyl-labelled tetrapeptide **11** and α-*N*-acetyl protected tetrapeptide **16**.

However, as can be seen in the peptide structures shown in Figure 3.9, only tetrapeptide **11** has a UV-active moiety present. To resolve this issue, the *N*-thiophosphoramidates generated during thiophosphorylation were 'trapped' through *S*-alkylation using a UV-active alkylating agent: 2-(bromomethyl)naphthalene.

For the *N*-thiophosphorylation step 1.0 eq of each of the tetrapeptides **11** and **16**, 20.0 eq 2 M NaOH solution and 1.0 eq $PSCl_3$ solution were allowed to react for 30 min at room temperature. The subsequent *S*-alkylation step involved the addition of 2.0 eq 2-(bromomethyl)naphthalene before allowing that mixture to react for 15 min at room temperature. The general scheme of the competition reactions is in illustrated in Scheme 3.4 and further experimental details are provided in section 7.2.4.

The method used for attempted determination of *N*-thiophosphorylation selectivity using LC-MS total absorbance chromatographs will be discussed in section 3.3.1.



Scheme 3.4 One-pot, two-step *N*-thiophosphorylation and *S*-alkylation competition reaction between α-*N*-acetyl protected tetrapeptide **16** and ε-*N*-dansyl-labelled tetrapeptide **11** to generate α-*N*-acetyl *S*-naphthyl ε-*N*-thiophosphoramidate **18** and ε-*N*-dansyl *S*-naphthyl α-*N*-thiophosphoramidate **19** via α-*N*-acetyl ε-*N*-thiophosphoramidate **17** and ε-*N*-dansyl α-*N*-thiophosphoramidate **12** respectively.

3.3.1. Use of LC-MS analysis for determining selectivity of *N*-thiophosphorylation

Following the synthetic procedure outlined in section 3.3 and Scheme 3.4, competition reactions for the evaluation of *N*-thiophosphorylation selectivity were conducted. An example of the LC-MS chromatographs obtained following *N*-thiophosphorylation and *S*-alkylation is shown in Figure 3.10.

Since only ε-*N*-dansyl-labelled tetrapeptide **11** contains a chromophore, the reaction progress following *N*-thiophosphorylation can only be noted for the conversion of tetrapeptide **11** to ε-*N*-dansyl-labelled α-*N*-thiophosphoramidate **12** (3%). When compared to the conversion from tetrapeptide **11** to α-*N*-thiophosphoramidate **12** seen in section 3.1 (42%), this result of 3% conversion was surprisingly low. However, it must be considered that the thiophosphorylation described in section 3.1 used 10.0 eq $PSCl_3$ solution to every 1.0 eq tetrapeptide **11**. In the competition reaction represented in Scheme 3.4 and Figure 3.10, 1.0 eq $PSCl_3$ solution was used for every 1.0 eq of tetrapeptide **11**. Alongside this, a second tetrapeptide (**16**) is present in the same proportions as tetrapeptide **11**, and tetrapeptide **16** has an available ε-amine group which is believed to be more nucleophilic than the available α-amino group on tetrapeptide **11** (discussed in section 3.1).

The subsequent *S*-alkylation of the crude thiophosphoramidate mixture by a chromophore allows for estimates of conversions to *S*-naphthyl α-*N*-thiophosphoramidate **19** to be calculated. By only accounting for the peaks corresponding to ε-*N*-dansyl-labelled tetrapeptide **11** and its derivatives in the *S*-alkylation chromatograph shown in Figure 3.10, the conversion to *S*-naphthyl α-*N*-thiophosphoramidate **19** from tetrapeptide **11** was found to be 11%. This conversion, however, is not an accurate depiction as the UV absorbance contribution of the naphthalene moiety added during *S*-alkylation has not been considered.

It is also not possible to obtain a clear idea of the conversion from α-*N*-acetyl protected tetrapeptide **16** to α-*N*-acetyl protected *S*-naphthyl ε-*N*-thiophosphoramidate **18** as only naphthalated derivative **18** is visible on the LC-MS total absorbance chromatographs seen in Figure 3.10.

That being said, it was possible to calculate an approximate reaction selectivity between α-*N*-thiophosphorylation and ε-*N*-thiophosphorylation using the areas under the relevant peaks on the total absorbance chromatographs. An example of these selectivity calculations using the chromatograph peaks displayed follows Figure 3.10 in section 3.3.1.1.

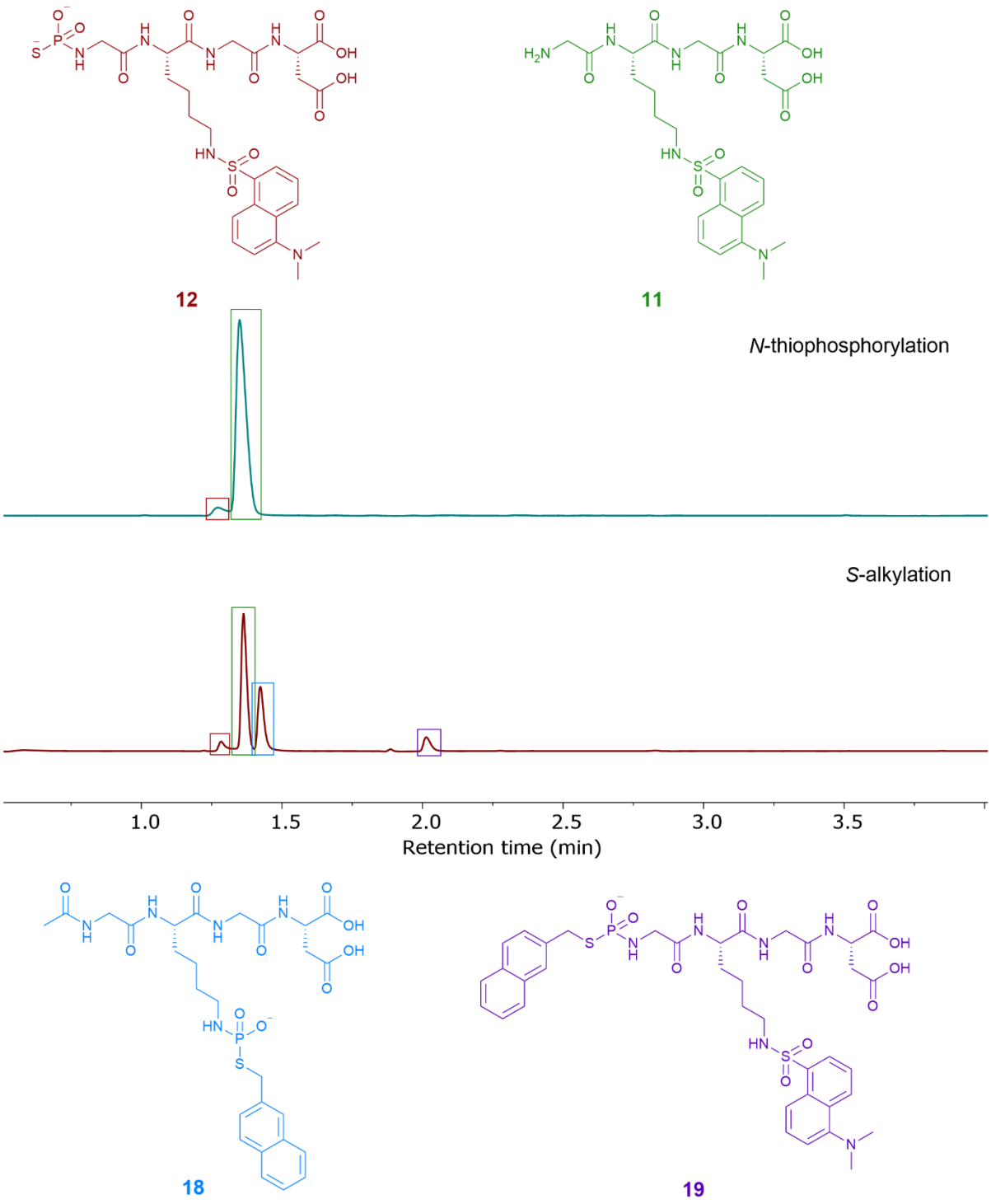


Figure 3.10 LC-MS total absorbance chromatographs following *N*-thiophosphorylation competition experiment, and subsequent *S*-alkylation 'trapping' reaction, between ε-*N*-dansyl-labelled tetrapeptide **11** (indicated by the green boxes – *m/z* 609 Da) and α-*N*-acetyl protected tetrapeptide **16** (not UV active so not seen). ε-*N*-dansyl-labelled α-*N*-thiophosphoramidate **12** is indicated by the small red boxes – *m/z* 705 Da; ε-*N*-dansyl-labelled *S*-naphthyl α-*N*-thiophosphoramidate **19** is indicated by the purple box – *m/z* 845 Da; α-*N*-acetyl protected *S*-naphthyl ε-*N*-thiophosphoramidate **18** is indicated by the blue box – *m/z* 654 Da.

3.3.1.1. Selectivity calculations based on LC-MS total absorbance chromatographs

First, an adjustment factor for the concentration changes of the peptide substrates between *N*-thiophosphorylation and *S*-alkylation needs to be calculated. This adjustment factor accounts for the aliquot removed from the *N*-thiophosphorylation mixture for LC-MS analysis, and the dilution following the addition of alkylating agent to the reaction stock (for *S*-alkylation). While the concentration of both peptide substrates **11** and **16** are impacted, the adjustment factor would only be applied to ε-dansyl-labelled tetrapeptide **11** and its derivatives, as only these peptides are UV active without the addition of the naphthalene group. Therefore, the reduction in the size of these peaks following sample dilution (with the removal of an aliquot for LC-MS analysis after *N*-thiophosphorylation and the addition of the alkylating agent solution) would impact the apparent UV contribution from the naphthalene moiety on **19**, resulting in an inaccurate estimation of *N*-thiophosphorylation selectivity.

The concentration adjustment factor is calculated from the ratio of the area under the peaks corresponding to tetrapeptide **11** in both the *N*-thiophosphorylation and *S*-alkylation total absorbance chromatographs. The area under the peak corresponding to tetrapeptide **11** was chosen as it was assumed that no thiophosphorylation occurred during the *S*-alkylation stage, thus the relative amount of tetrapeptide **11** following both *N*-thiophosphorylation and *S*-alkylation should remain the same.

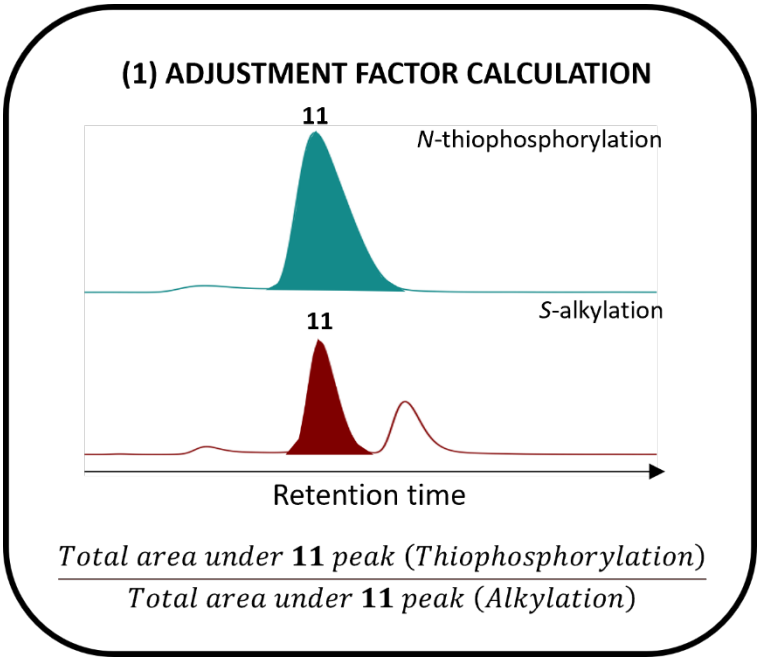


Figure 3.11 Illustration of the calculation of the concentration adjustment factor. The equation shown gives the final concentration adjustment factor. The total absorbance chromatographs are those seen in Figure 3.10.

For this competition reaction, the concentration adjustment factor was found to be 2.5.

To find the UV absorbance contribution of the naphthalene moiety in ε-*N*-dansyl-labelled *S*-naphthyl α-*N*-thiophosphoramidate **19**, the areas under peaks corresponding to ε-*N*-dansyl-labelled tetrapeptide **11** and its derivatives are summated for both the *N*-thiophosphorylation and *S*-alkylation chromatographs. The total area under the peaks corresponding to **11**, **12** and **19** following *S*-alkylation was multiplied by the concentration adjustment factor. The total area under the peaks corresponding to **11** and **12** in the *N*-thiophosphorylation chromatogram was subtracted from the adjusted total area under the peaks corresponding to **11**, **12** and **19** in the *S*-alkylation chromatogram to give the UV contribution of the naphthalene moiety attached to **19**.

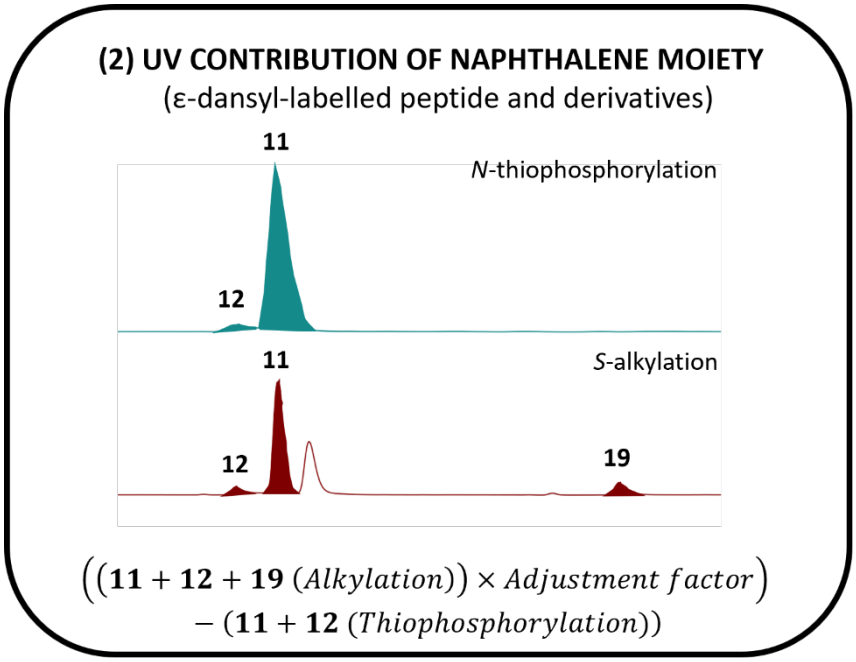


Figure 3.12 Illustration of the calculation of the UV absorbance contribution of the naphthalene moiety on ε-*N*-dansyl-labelled *S*-naphthyl α-*N*-thiophosphoramidate **19**. The equation shown gives the UV absorbance contribution of the naphthalene moiety on **19**. The total absorbance chromatographs are those seen in Figure 3.10.

The total UV absorbance contribution of the naphthalene moiety on α-*N*-acetyl protected *S*-naphthyl ε-*N*-thiophosphoramidate **18** is simply the area under its corresponding peak in the *S*-alkylation chromatograph.

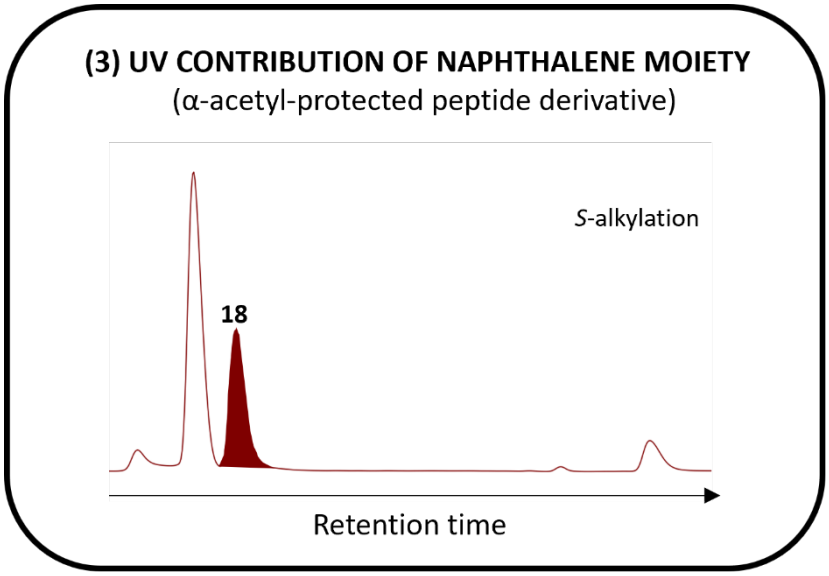


Figure 3.13 Illustration of finding the UV absorbance contribution of the naphthalene moiety on α-*N*-acetyl-protected *S*-naphthyl ε-*N*-thiophosphoramidate **18**. The total absorbance chromatographs are those seen in Figure 3.10.

Finally, to determine an approximate selectivity for *N*-thiophosphorylation between ε- and α-*N*-thiophosphorylation, the area under the peak corresponding to **18** and the absorbance contribution of the naphthalene moiety on **19** are divided against each other.

(4) SELECTIVITY DETERMINED THROUGH LC-MS ANALYSIS

$$\frac{Naphthalene\ UV\ contribution\ (\mathbf{18})}{Naphthalene\ UV\ contribution\ (\mathbf{19})} = \frac{\varepsilon - thiophosphoramidate}{\alpha - thiophosphoramidate}$$

Figure 3.14 Equation to determine the selectivity of *N*-thiophosphorylation between ε-*N*-thiophosphorylation and α-*N*-thiophosphorylation through LC-MS total absorbance chromatograms.

The *N*-thiophosphorylation selectivity calculated for this competition reaction comes to 1.2, suggesting that there is 1.2 × more thiophosphorylation occurring at the available ε-NH_2 of α-*N*-acetyl protected tetrapeptide **16** compared to the α-NH_2 of ε-*N*-dansyl-labelled tetrapeptide **11**.

The competition reaction was carried out twice more to ensure the selectivity values obtained were reproducible. Table 3.1 details the results of the three competition reactions. The selectivity values were calculated using the method described above.

Table 3.1 Details and results obtained from the three competition reactions carried out at room temperature, with 30 min *N*-thiophosphorylation and 15 min *S*-alkylation.

Competition reaction number	Equivalents of $PSCl_3$	Equivalents of NaOH	Equivalents of 2-(bromomethyl)naphthalene	Selectivity (ε-NH_2/ α-NH_2)
1	1.0	20.0	2.0	1.2
2	1.0	20.0	2.0	2.9
3	1.0	20.0	2.0	6.7

While each of the competition reactions outlined in Table 3.1 were carried out using the same number of equivalents for each reagent, the selectivities calculated in each case differ significantly. While each selectivity suggests that *N*-thiophosphorylation is selective for ε-NH_2 over α-NH_2, an accurate idea of the reaction selectivity was not obtained from this reaction series.

Since each of these competition reactions were conducted at ambient room temperature, it was thought that the variation in reaction selectivities may have been due to the varying reaction temperatures. To test this theory, two competition reactions were performed simultaneously in a temperature-controlled cold room, set at 4 °C.

3.3.2. Cold room competition reactions to determine selectivity

For the competition reactions conducted at 4 °C, all of the reagents were kept the same and were used with the same number of equivalents as are detailed in Table 3.1. Aside from the temperature at which the reactions were run, the sole experimental difference was the reaction time. Instead of 30 min *N*-thiophosphorylation and 15 min *S*-alkylation, these reaction times were extended to 24 h and 3 h respectively.

The total absorbance chromatographs obtained following the *N*-thiophosphorylation and *S*-alkylation steps are displayed in Figure 3.15. Both chromatographs resemble those shown in Figure 3.10, but through the application of the selectivity calculation method described in section 3.3.1.1, the selectivity of the *N*-thiophosphorylation reaction was found to be 4.0 (i.e. 4.0 × more thiophosphorylation at ε-NH_2 than α-NH_2). As this was the first competition reaction carried out under temperature control, a second reaction was performed. Unfortunately, the selectivity for this second temperature-controlled competition reaction was found to be 5.6.

This continued inconsistency in the calculated reaction selectivity suggested that there may be an alternative cause unrelated to temperature control, so the LC-MS data were further questioned.

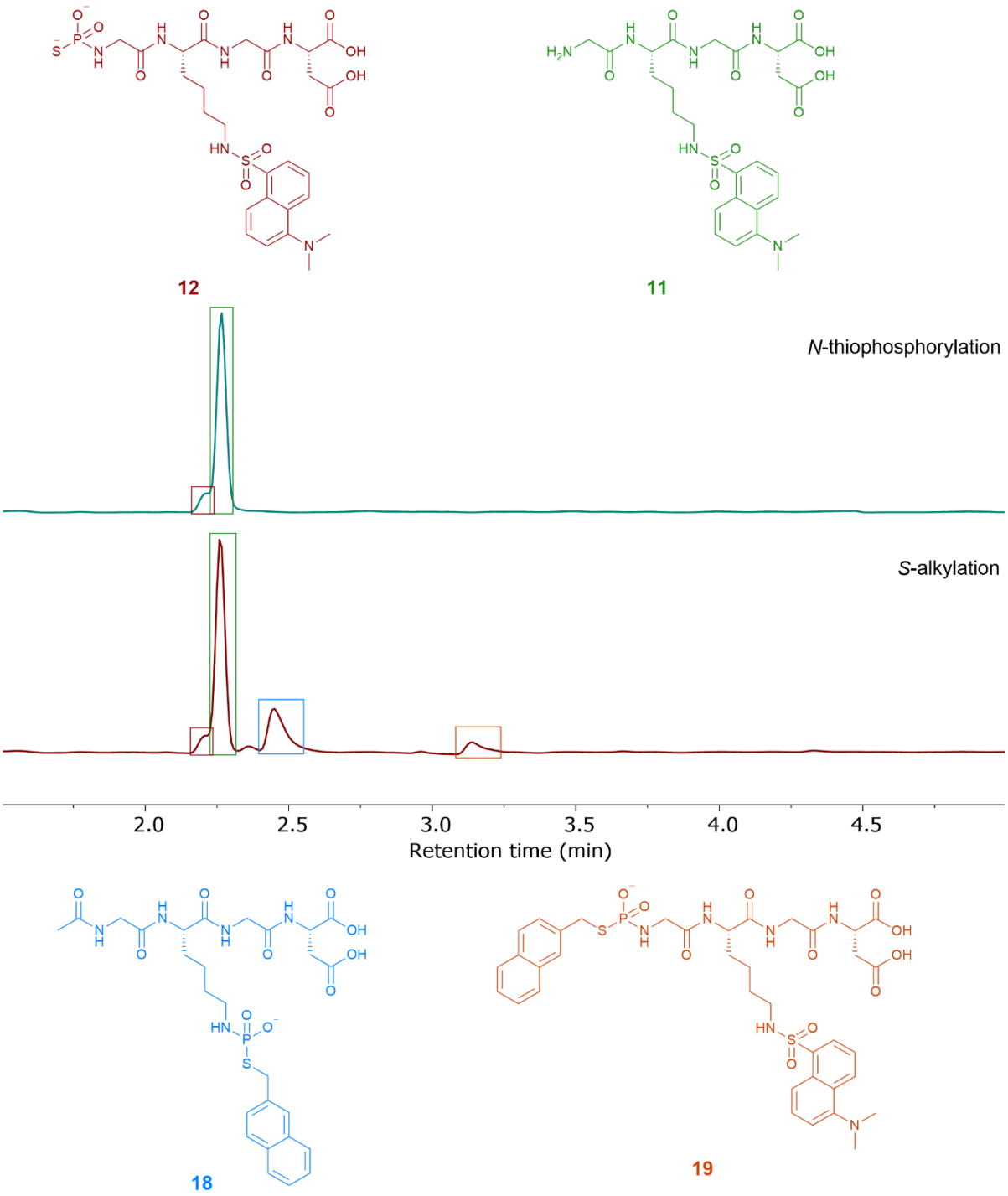


Figure 3.15 LC-MS total absorbance chromatographs following *N*-thiophosphorylation competition experiment and subsequent *S*-alkylation 'trapping' reaction set in a temperature-controlled cold room (4 °C), between ε-*N*-dansyl-labelled tetrapeptide **11** (indicated by the green boxes – *m/z* 609 Da) and α-*N*-acetyl protected tetrapeptide **16** (not UV active so not seen). ε-*N*-dansyl-labelled α-*N*-thiophosphoramidate **12** is indicated by the small red boxes – *m/z* 705 Da; ε-*N*-dansyl-labelled *S*-naphthyl α-*N*-thiophosphoramidate **19** is indicated by the orange box – *m/z* 845 Da; α-*N*-acetyl protected *S*-naphthyl ε-*N*-thiophosphoramidate **18** is indicated by the blue box – *m/z* 654 Da.

3.3.2.1. Further analysis of competition reaction *S*-alkylation chromatographs

As the total absorbance chromatographs collected following *N*-thiophosphorylation only show ε-*N*-dansyl-labelled tetrapeptide **11** and ε-*N*-dansyl-labelled α-*N*-thiophosphoramidate **12**, these were disregarded, and the heightened focus was placed on the *S*-alkylation chromatographs.

While exploring the mass spectra for each individual product peak obtained following *S*-alkylation, a species with *m/z* 255 Da was noted as being eluted alongside α-*N*-acetyl protected *S*-naphthyl ε-*N*-thiophosphoramidate **18**. Based on the mass EICs, it was difficult to determine whether the 255 Da species was simply a fragment from **18** or a discrete species as the retention time for both ion peaks was 1.45 min (Figure 3.16).

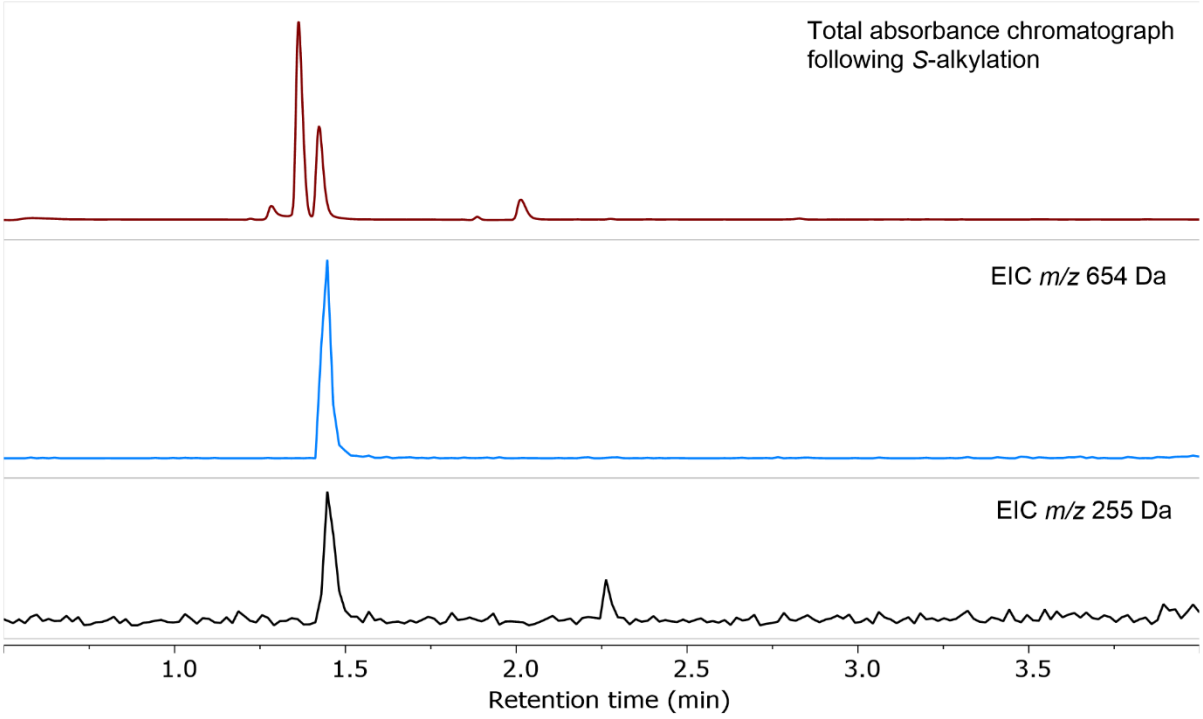


Figure 3.16 LC-MS total absorbance chromatograph following competition reaction *S*-alkylation at room temperature (red trace – same *S*-alkylation chromatograph as in Figure 3.10) with mass EICs for α-*N*-acetyl protected *S*-naphthyl ε-*N*-thiophosphoramidate **18** (blue trace – *m/z* 654 Da) and suspected by-product (black trace – *m/z* 255 Da).

LC-MS experiments conducted on temperature-controlled competition samples used a different LC procedure. Rather than the 5 min LC gradient used for the analysis of non-temperature-controlled samples (shown in Figures 3.10 and 3.16), a 7 min LC gradient was employed for the temperature-controlled samples (shown in Figure 3.15 and 3.17) (LC gradient details can be found in Chapter 7). This was done in an effort to improve the separation between the peaks eluted between 1.2 and 1.5 min in the chromatogram shown in Figure 3.16. The thought was that improved separation between

these peaks would lead to greater consistency in the selectivities calculated from LC-UV-vis analysis, as any overlap in species elution, and therefore overlap in LC-UV-vis peaks, would be minimised.

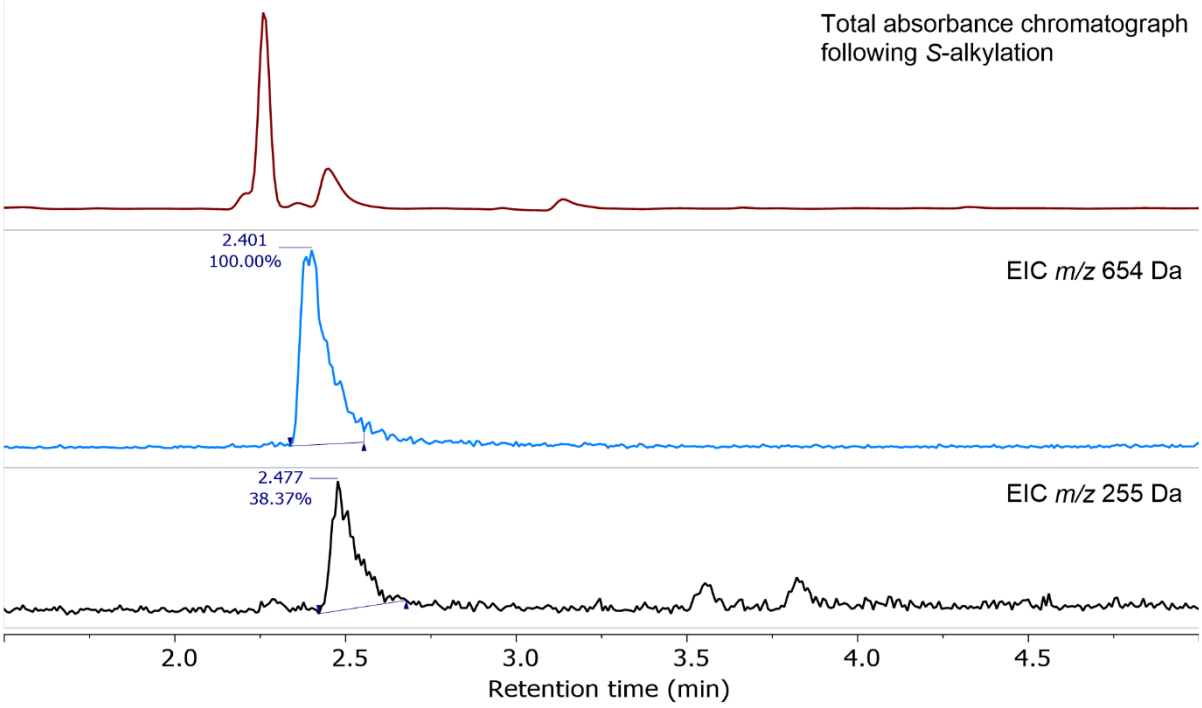


Figure 3.17 LC-MS total absorbance chromatograph following temperature-controlled competition reaction *S*-alkylation (red trace – same *S*-alkylation chromatograph as in Figure 3.15) with mass EICs for α-*N*-acetyl protected *S*-naphthyl ε-*N*-thiophosphoramidate **18** (blue trace – *m/z* 654 Da) and suspected by-product (black trace – *m/z* 255 Da).

This extended chromatographic run resulted in slightly different retention times between α-*N*-acetyl-protected *S*-naphthyl ε-*N*-thiophosphoramidate **18** and *m/z* 255 Da species, however the retention times may not be distinct enough to definitively conclude that the two species are unrelated – i.e. that the *m/z* 255 Da species is not a hydrolysis product formed from **18** over the course of the chromatography run (retention times for thiophosphoramidate **18** was 2.40 min, and *m/z* 255 Da species was 2.48 min).

A control experiment was carried out to determine whether *m/z* 255 Da species would still form in the absence of a peptide substrate. $PSCl_3$, $NaOH_{(aq)}$ and 2-(bromomethyl)naphthalene (in the same number of equivalents as outlined in Table 3.1) were added to a reaction vessel and left stirring for an hour at room temperature. LC-MS analysis of the reaction mixture after the reaction indicated the formation of a species with *m/z* 255 Da.

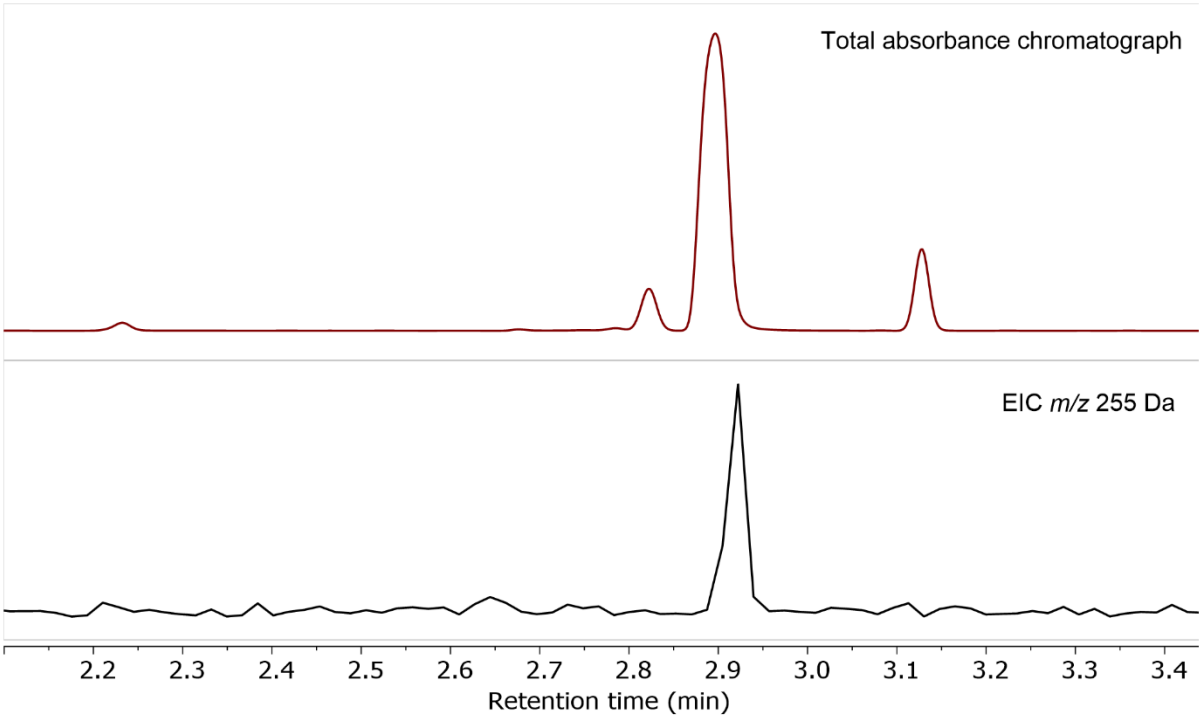
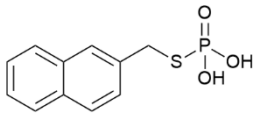


Figure 3.18 LC-MS analysis following the control thiophosphorylation of 2-(bromomethyl)naphthalene. The red trace is the total absorbance chromatograph, and the black trace is the mass EIC for *m/z* 255 Da.

Based on the formation of *m/z* 255 Da species in the absence of a peptide substrate, it can be assumed that this species must consist of a combination of thiophosphate and naphthalene-based substituents. With this in mind, Figure 3.19 illustrates a possible structure for the 255 Da species.



20

Figure 3.19 Proposed structure for *m/z* 255 Da (**20**) based on the LC-MS results following the control thiophosphorylation of 2-(bromomethyl)naphthalene.

Assuming that this proposed structure is accurate for **20**, it may form one of two ways over the course of the competition reactions discussed in sections 3.3.1 and 3.3.2. Either it is a direct *S*-alkylation of any inorganic thiophosphate present in the crude reaction mixture, or it is a hydrolysis product formed during the *S*-alkylation stage.

Regardless of how it is formed, the co-elution of **20** with α-*N*-acetyl-protected *S*-naphthyl ε-*N*-thiophosphoramidate **18** in the crude competition reaction mixtures leads to an inaccurate representation of the apparent UV absorbance for **18** in LC-MS chromatographs, which in turn leads to inaccurate selectivities calculated from the total absorbance chromatographs. The selectivities calculated using the method outlined in section 3.3.1.1 lean in favour of ε-NH_2 thiophosphorylation,

as there is no equivalent co-elution of **20** with ε-*N*-dansyl-labelled *S*-naphthyl α-*N*-thiophosphoramidate **19** at 2.01 min (Figure 3.16).

As both tandem MS and LC-MS analyses were unable to give a clear idea for the selectivity between ε-NH_2 thiophosphorylation and α-NH_2 thiophosphorylation, NMR analysis was reassessed as a potential tool for investigating the selectivity of the *N*-thiophosphorylation reaction.

3.4. ^{31}P NMR assay for investigating the selectivity of *N*-thiophosphorylation on a series of unprotected tetrapeptides

Despite previously choosing MS-based methods over NMR analysis due to the much higher material cost associated with NMR, it was thought that a ^{31}P NMR-based assay could give a more accurate estimation of *N*-thiophosphorylation selectivity on an unprotected peptide.

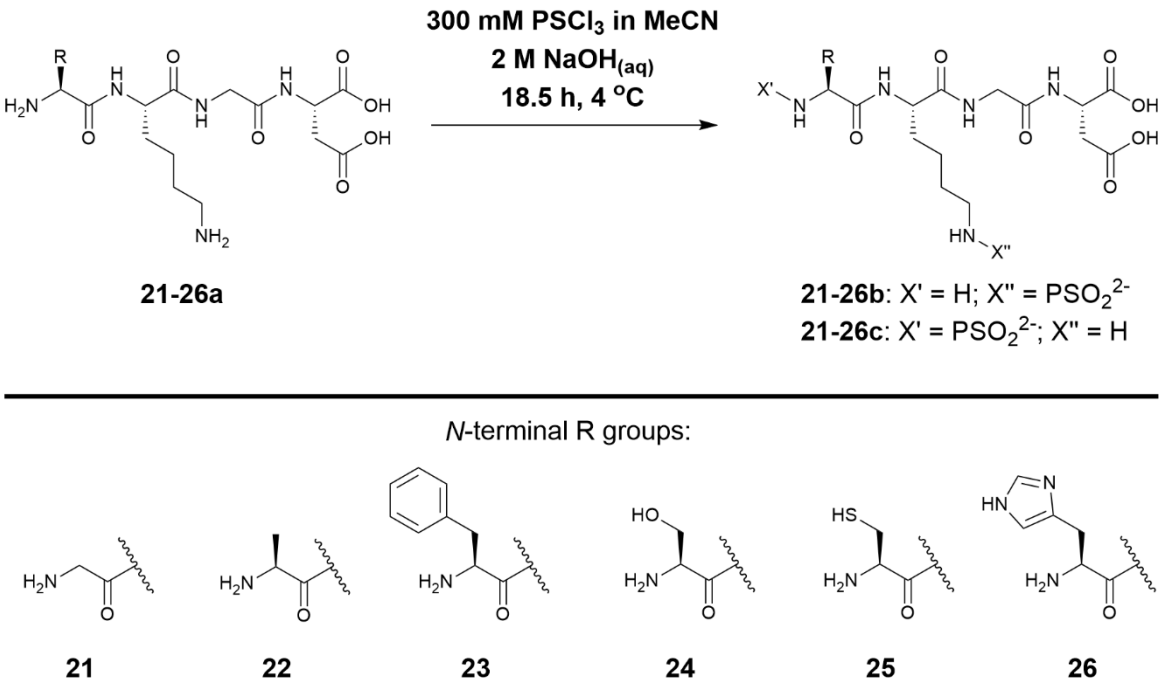
Since ^{31}P NMR chemical shifts are determined by the chemical environment surrounding each phosphorus atom, there is little chance of the expected signals overlapping. The signal splitting patterns seen on the ^{31}P NMR spectra will also give information on the thiophosphorylation site (ε-NH_2 or α-NH_2), which MS-based techniques failed to clarify.

Using ^{31}P NMR as the primary analysis method for this reaction series also negates the need for a chromophore label on the peptide, so *S*-alkylation is no longer necessary to trap the thiophosphoramidate as it was for LC-MS DAD analysis.

3.4.1. Aqueous *N*-thiophosphorylation of unprotected tetrapeptides **21a – 26a**

A series of six unprotected tetrapeptides was designed, each with the same three amino acid residues from the *C*-terminus and a different *N*-terminal amino acid (general sequence: XKGD). For the *N*-terminal amino acids, a selection of hydrophobic and hydrophilic side chains was chosen. Three different sizes of hydrophobic side chains were chosen in order to see how steric effects may impact *N*-thiophosphorylation selectivity. As for the hydrophilic side chains, different side chains were selected to see how the presence of other nucleophilic residues may impact the *N*-thiophosphorylation selectivity.

Each of the six unprotected tetrapeptides underwent aqueous *N*-thiophosphorylation individually, using 1.0 eq $PSCl_3$ in MeCN solution and 7.0 eq 2 M $NaOH_{(aq)}$ solution for 18.5 h at 4 °C (Scheme 3.5).



Scheme 3.5 Aqueous *N*-thiophosphorylation of unprotected tetrapeptides **21-26a** to give either ε-*N*-thiophosphoramidates **21-26b**, or α-*N*-thiophosphoramidates **21-26c**.

The resulting crude thiophosphorylation mixtures were submitted for ^{31}P NMR analyses.

3.4.2. ^{31}P NMR assay results

In all six thiophosphorylation reactions, > 90% of the $PSCl_3$ was consumed over the 18.5 h reaction period (indicated through the reduction in $PSCl_3$ phosphorus signal around 31.6 ppm). The ^{31}P spectra for each of the six thiophosphorylations are displayed in Figure 3.20.

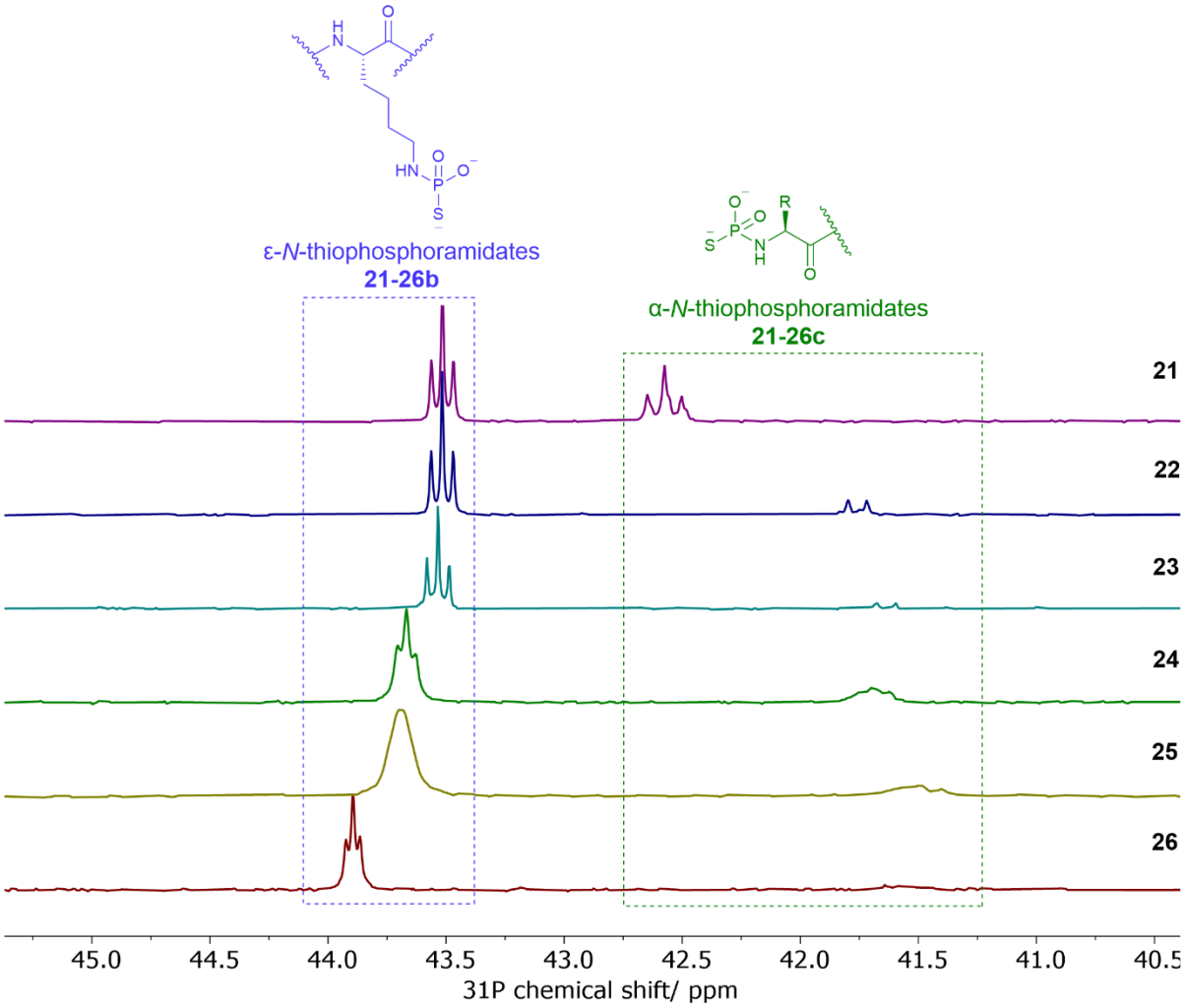


Figure 3.20 ^{31}P NMR spectra following *N*-thiophosphorylation series performed on unprotected tetrapeptides **21-26a**. The dotted blue box indicates ε-*N*-thiophosphoramidates **21-26b** signals between 43.5 – 44.0 ppm. The dotted green box indicates α-*N*-thiophosphoramidates **21-26c** signals between 41.5 – 42.7 ppm.

To determine which signal corresponded to which thiophosphoramidate, the number of hydrogens coupling to the thiophosphoryl phosphorus needed to be considered. For ε-*N*-thiophosphoramidates **21-26b**, the phosphorus atom is coupling to the two methylene hydrogens on the ε-C of the lysine residue resulting in an apparent triplet signal. For α-*N*-thiophosphoramidates **22-26c**, the phosphorus atom is coupling to the single hydrogen present at the α-C centre of the *N*-terminal amino acids resulting in a doublet signal.

For *N*-terminal glycine α-*N*-thiophosphoramidate **21c**, both signals appear as triplets since the phosphorus atoms in both potential thiophosphoramidates are coupled to methylene hydrogens (ε-CH_2 and α-CH_2). The chemical shifts of the two triplets in the ^{31}P NMR spectrum following **21a** thiophosphorylation were therefore compared to the chemical shifts of the doublet signals known to represent α-*N*-thiophosphoramidates **22-26c**. The triplet signal on the **21b-c** peptide spectrum with a

chemical shift closer to those of thiophosphoramidates **22-26c** was assumed to be the triplet signal representative of *N*-terminal glycine α-*N*-thiophosphoramidate **21c**.

3.4.2.1. *N*-thiophosphorylation selectivity calculation based on ^{31}P NMR signal integrations

Having determined the ^{31}P NMR signals corresponding to the α- and ε-*N*-thiophosphoramidates, these signals were used to calculate the *N*-thiophosphorylation selectivity.

Using the ^{31}P NMR spectrum of thiophosphorylated *N*-terminal alanine peptide **22a** as an example, the integrations of ε-*N*-thiophosphoramidate **22b** (apparent triplet signal at 43.52 ppm with an integration of 1.00) and α-*N*-thiophosphoramidate **22c** (doublet signal at 41.76 ppm with an integration of 0.18) were taken. A ratio of these two integrations was calculated and this gave the thiophosphorylation selectivity.

$$\frac{\varepsilon - N - thiophosphoramidate\ \mathbf{22b}}{\alpha - N - thiophosphoramidate\ \mathbf{22c}} = \frac{1.00}{0.18} = 5.6$$

Therefore, according to the signal integrations, 5.6 × more ε-*N*-thiophosphoramidate **22b** was formed than α-*N*-thiophosphoramidate **22c**.

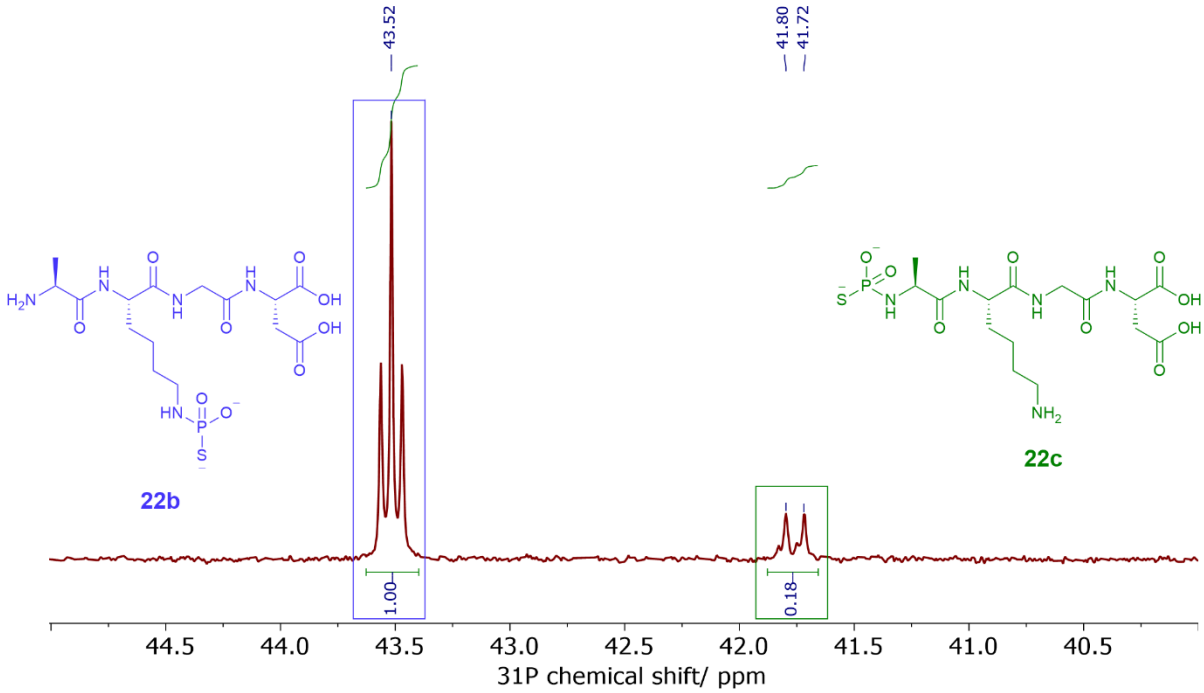
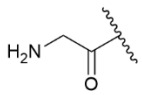
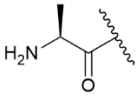
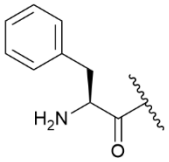
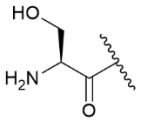
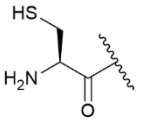
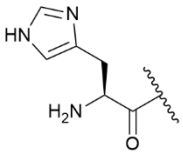


Figure 3.21 ^{31}P NMR spectrum of thiophosphorylated *N*-terminal alanine tetrapeptide **22a**. The blue box surrounding the apparent triplet signal at 43.52 ppm indicates ε-*N*-thiophosphoramidate **22b**, and the green box surrounding the doublet signal at 41.76 ppm indicates α-*N*-thiophosphoramidate **22c**.

The selectivities for the remaining five tetrapeptides that were exposed to thiophosphorylation conditions were calculated using ^{31}P NMR signal integrations and the selectivity results are displayed in Table 3.2.

Table 3.2 *N*-thiophosphorylation selectivity results from the thiophosphorylation reaction series performed on unprotected tetrapeptides **21-26a**. The selectivity value for each peptide is based on the ^{31}P NMR signal integrations for ε-*N*-thiophosphoramidates **21-26b** and α-*N*-thiophosphoramidates **21-26c**. Concentration of each tetrapeptide solution – 100 mM; concentration of tetrapeptide in total reaction solution – 60 mM.

Peptide	pK_{aH}[154]	NaOH/ eq	$PSCl_3$/ eq	ε-NH_2/ ^{31}P signal integration	α-NH_2/ ^{31}P signal integration	Selectivity (ε-NH_2/ α-NH_2)
 **21a**	9.60	7.0	1.0	1.0	0.69	**1.4**
 **22a**	9.69	7.0	1.0	1.0	0.18	**5.6**
 **23a**	9.13	7.0	1.0	1.0	0.07	**14.3**
 **24a**	9.15	7.0	1.0	1.0	0.29	**3.4**
 **25a**	10.28	7.0	1.0	1.0	0.19	**5.3**
 **26a**	9.17	7.0	1.0	1.0	0.15	**6.7**

Based on the selectivities found through comparing ^{31}P NMR signal integrations, sterics appear to be the dominant factor in determining selectivity. The lowest selectivity seen was for *N*-terminal glycine tetrapeptide **21a**, with the ε-NH_2 to α-NH_2 ratio being 1.4. In contrast, the highest selectivity seen was for *N*-terminal phenylalanine tetrapeptide **23a**, an ε-NH_2 to α-NH_2 ratio of 14.3. Since glycine is the smallest amino acid (van der Waals volume of 71.8 Å^3),[162] its α-NH_2 is the least sterically hindered of all natural amino acids. Phenylalanine, on the other hand, is one of the largest naturally occurring

amino acids (van der Waals volume of 186.4 Å^3)[162] with a phenyl ring on the side chain, so the *N*-terminus of **23a** is significantly more sterically hindered than the *N*-terminus of **21a**.

The pK_{aH} of the *N*-terminus does not appear to have a significant impact on the reaction selectivity between α-NH_2 and ε-NH_2. As stated in section 3.1, the nucleophilicity of a compound generally increases with its pK_{aH}. Based on this general trend, *N*-terminal cysteine peptide **25a** should have the lowest selectivity as it has the most nucleophilic *N*-terminus.

Figure 3.22 shows how the expected selectivity trends based on α-NH_2 pK_{aH} and van der Waals volume (Å^3) compared to the observed selectivities, with the *N*-terminal amino acid expected to give the lowest selectivity on the left-hand side, and the *N*-terminal amino acid expected to give the greatest selectivity on the right-hand side. The top trend shown is based on the pK_{aH} of each *N*-terminal α-NH_2 and the middle trend shown is based on the van der Waals volume (Å^3) of each *N*-terminal residue. The experimental selectivities trend (the bottom trend shown in Figure 3.22) shows the observed selectivities following thiophosphorylation of unprotected tetrapeptides **21-26a**.

As shown in Figure 3.22, the experimental selectivity trend is far more closely corelated to the volume of the *N*-terminal residue. The main exception is that alanine at the *N*-terminus results in greater selectivity for ε-*N*-thiophosphorylation than expected. Despite this, it can be concluded that *N*-thiophosphorylation selectivity between a lysine side chain ε-NH_2 and *N*-terminal α-NH_2 is largely dependent on the size of the side chain on the N-terminal residue.

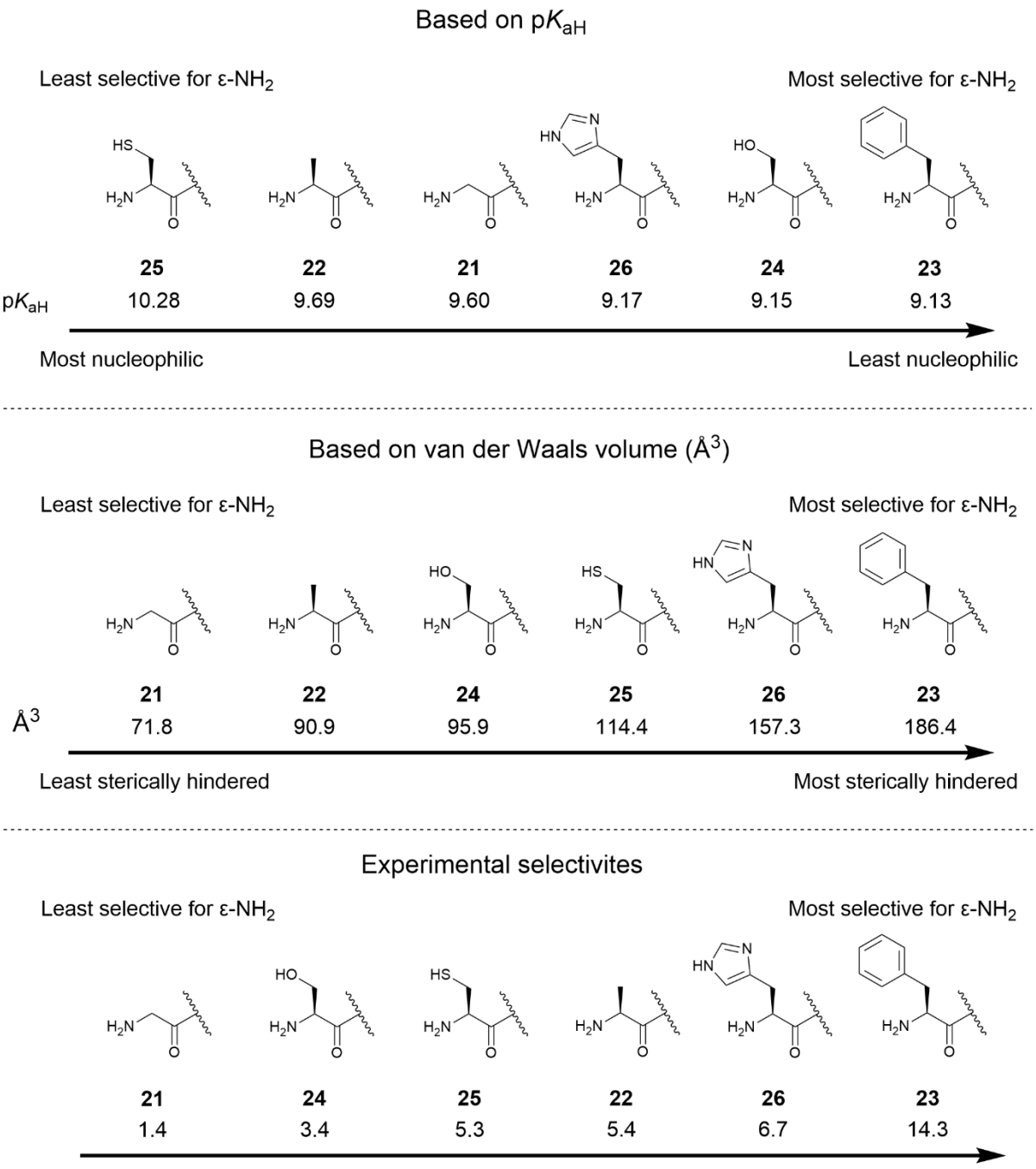


Figure 3.22 Comparison of expected thiophosphorylation selectivity trends based on α-NH_2 pK_{aH} values (top trend) and amino acid van der Waals volumes (Å^3) (middle trend) with the observed thiophosphorylation selectivities (bottom trend). For each trend, the (expected) lowest selectivity for ε-NH_2 is on the left and the (expected) highest selectivity for ε-NH_2 is on the right. The pK_{aH} values are from Lide.[154] The van der Waals volume values for F, H, S, A, and G are from Rellick;[162] van der Waals volume for C is from Pontius.[163]

While aqueous *N*-thiophosphorylation has been found to be selective for ε-*N*-thiophosphorylation over α-*N*-thiophosphorylation, the impact of other unprotected nucleophilic residues must also be considered. Within this selectivity series, two nucleophilic residues (cysteine and histidine) were included as *N*-terminal amino acids in peptides **25a** and **26a** respectively. The thiophosphorylations of these two peptides will be discussed here.

3.4.2.2. *N*-terminal cysteine tetrapeptide **25a** thiophosphorylation

The presence of a highly nucleophilic cysteine residue was seen to lead to the production of a variety of by-products over the course of the thiophosphorylation reaction. To reiterate the comments in section 3.4.1, all thiophosphorylations carried out on unprotected tetrapeptide substrates used the same number of equivalents of $PSCl_3$ solution and $NaOH_{(aq)}$ solution, and each thiophosphorylation was run for the same length of time (18.5 h) and at the same temperature (4 °C). The ^{31}P NMR spectrum of *N*-terminal cysteine thiophosphoramidate **25b** and **25c** mixture is shown in Figure 3.23.

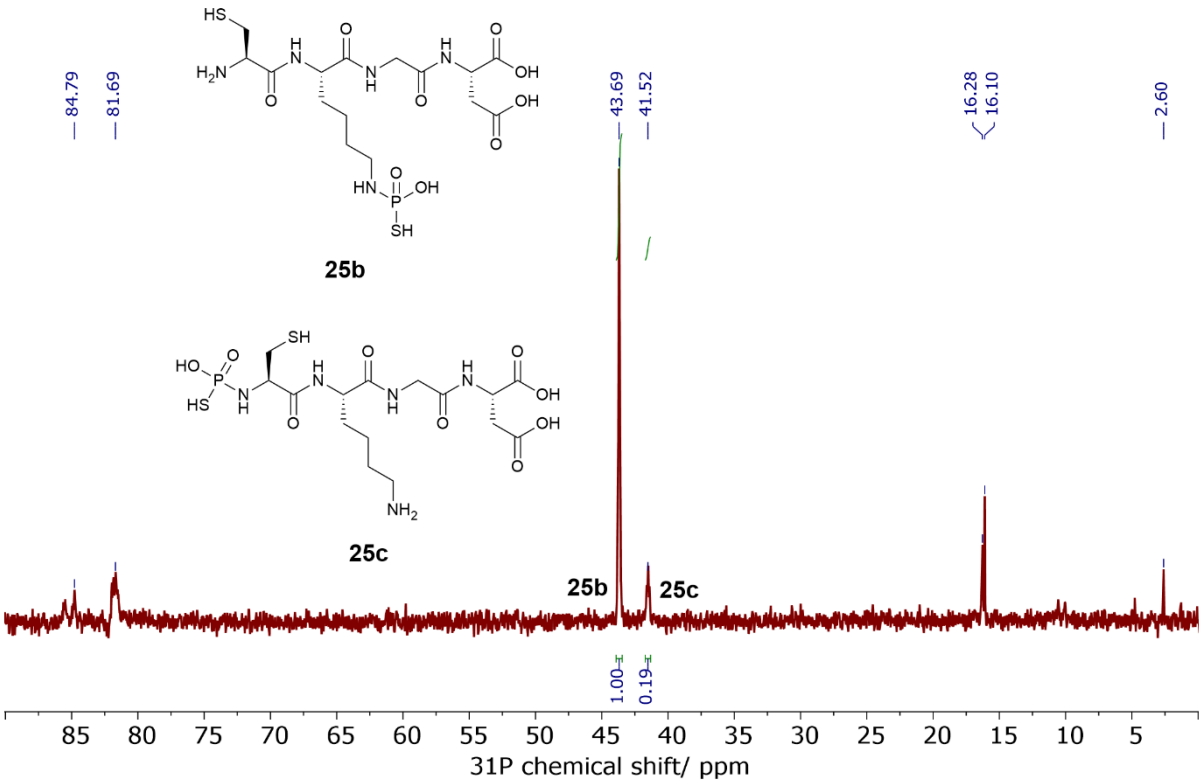
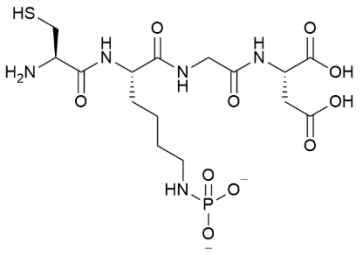


Figure 3.23 ^{31}P NMR spectrum following the *N*-thiophosphorylation of *N*-terminal cysteine tetrapeptide **25a**. The structures of thiophosphoramidates **25b** and **25c** are shown for reference.

The signals at 43.69 and 41.52 ppm were identified in Figure 3.20 as ε-*N*-thiophosphoramidate **25b** and α-*N*-thiophosphoramidate **25c** respectively. The minor signal at 2.60 ppm is inorganic phosphate. The remaining signals highlighted at 84.79, 81.69 and 16.28 ppm indicate the various by-products formed during the thiophosphorylation reaction.

The signal at 16.28 ppm is likely to be desulfurised thiophosphoramidate (the ^{31}P signal does not indicate which phosphoramidate isomer is formed, so the structure of the ε-*N*-phosphoramidate isomer **25d** is shown in Figure 3.24 as an example). As none of the other 5 tetrapeptides used for the ^{31}P NMR assay in section 3.4.2 indicated the formation of a phosphoramidate by-product (see Appendix 1, Figures 8.25 to 8.30 for full spectra), its presence here suggests that a cysteine residue

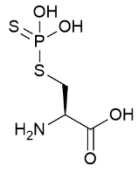
adjacent to a thiophosphoramidate group may cause the thiophosphoramidate to be more susceptible to the loss of sulfur. In order to maximise the formation of *N*-thiophosphoramidate and minimise the potential for desulfurisation, thiophosphorylations of unprotected, cysteine-containing peptides and proteins should perhaps be run for a reduced reaction time to limit the potential for subsequent desulfurisation.



25d

Figure 3.24 Structure of *N*-terminal cysteine ε-*N*-phosphoramidate **25d**.

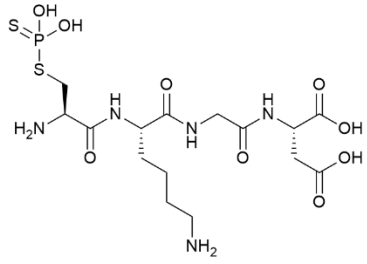
The signal at 84.79 ppm corresponds closely to a calculated value for *S*-thiophosphorylated cysteine reported in 2010 (87.3 ppm). In this case, the *S*-thiophosphorylated cysteine was assumed to be the thione tautomer (Figure 3.25).[149]



'Thione' tautomer

Figure 3.25 The structure of the 'thione' tautomer of *S*-thiophosphorylated cysteine.

Based on the reported value for *S*-thiophosphorylated cysteine as the thione tautomer, it can be assumed that the signal at 84.79 ppm observed in Figure 3.23 may be *S*-thiophosphorylated tetrapeptide in the thione tautomeric form (structure **25e** in Figure 3.26).

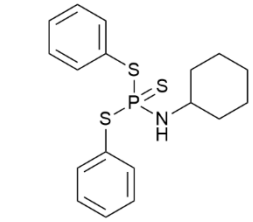


25e

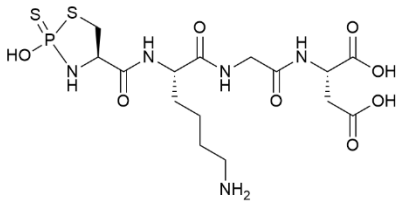
Figure 3.26 Proposed potential structure for the thiophosphorylation by-product indicated by ^{31}P NMR signals at 84.79 ppm in Figure 3.23.

As for the signal seen at 81.69 ppm in Figure 3.23, the potential identity of this structure is not clear. As the chemical shift is close to that of the reported *S*-thiophosphorylated cysteine value,[149] it can be assumed that the species represented by the signal at 81.69 ppm is also thiophosphorylated at the cysteine thiol group.

A phosphoramidotrithioate species reported in 2022 by Zhang *et al.* (structure in Figure 3.27) was shown to have a ^{31}P chemical shift of 80.58 ppm.[164] It was, therefore, thought that the species corresponding to the ^{31}P chemical shift at 81.69 ppm in Figure 3.23 may bare a structural resemblance to the phosphoramidotrithioate species. A potential structure for the *N*-terminal cysteine thiophosphorylated by-product corresponding to the chemical shift at 81.69 ppm (**25f**) is also shown in Figure 3.27. Table 3.3 summarises the identities of each of the ^{31}P signals noted in Figure 3.23.



Phosphoramidotrithioate species



25f

Figure 3.27 Structure of phosphoramidotrithioate species reported by Zhang *et al*. in 2022 (left hand structure),[164] and potential structure for the thiophosphorylation by-product indicated by ^{31}P NMR signals at 81.69 ppm in Figure 3.23 (right hand structure).

Table 3.3 Summary of the identities and apparent proportions of each ^{31}P signal seen in Figure 3.23. The proportion of each phosphorus-containing species present was determined from the ^{31}P signal integrations in Figure 3.23.

^{31}P chemical shift/ ppm	**Identity**	**Proportion/ %**
2.60	Inorganic phosphate	2
16.28	**25d**	10
41.52	**25c**	10
43.69	**25b**	55
81.69	**25f**	18
84.79	**25e**	5

While the selectivity for ε-*N*-thiophosphorylation over α-*N*-thiophosphorylation for *N*-terminal cysteine tetrapeptide **25a** was found to be 5.3, Figure 3.23 shows that there are many more phosphorus-containing products than the expected thiophosphoramidates **25b** and **25c**. Based on signal integrations, only 55% of all phosphorus-containing species generated over the thiophosphorylation of **25a** was found to correspond to ε-*N*-thiophosphoramidate **25b**. Based on the results summarised in Table 3.3, thiophosphorylation at an unprotected cysteine residue may compete with thiophosphorylation at an ε-amine group. To minimise the production of unwanted

dithiophosphate by-products during thiophosphorylation, peptides/ proteins containing cysteine residues should be avoided.

That being said, the ^{31}P NMR spectrum shown in Figure 3.23 suggests that cysteine residues may be a potential target for thiophosphorylation. Due to time constraints, however, the thiophosphorylation of cysteine residues was not explored further during this project.

3.4.2.3. *N*-terminal histidine tetrapeptide **26a** thiophosphorylation

As with the thiophosphorylation of *N*-terminal cysteine peptide **25a**, the thiophosphorylation of *N*-terminal histidine peptide **26a** resulted in more than just the expected thiophosphoramidate products (**26b** and **26c**). But, unlike in the thiophosphorylation of **25a**, the thiophosphorylation of **26a** resulted in only one significant by-product. The ^{31}P NMR spectrum of *N*-terminal histidine thiophosphoramidate **26b** and **26c** mixture is shown in Figure 3.28.

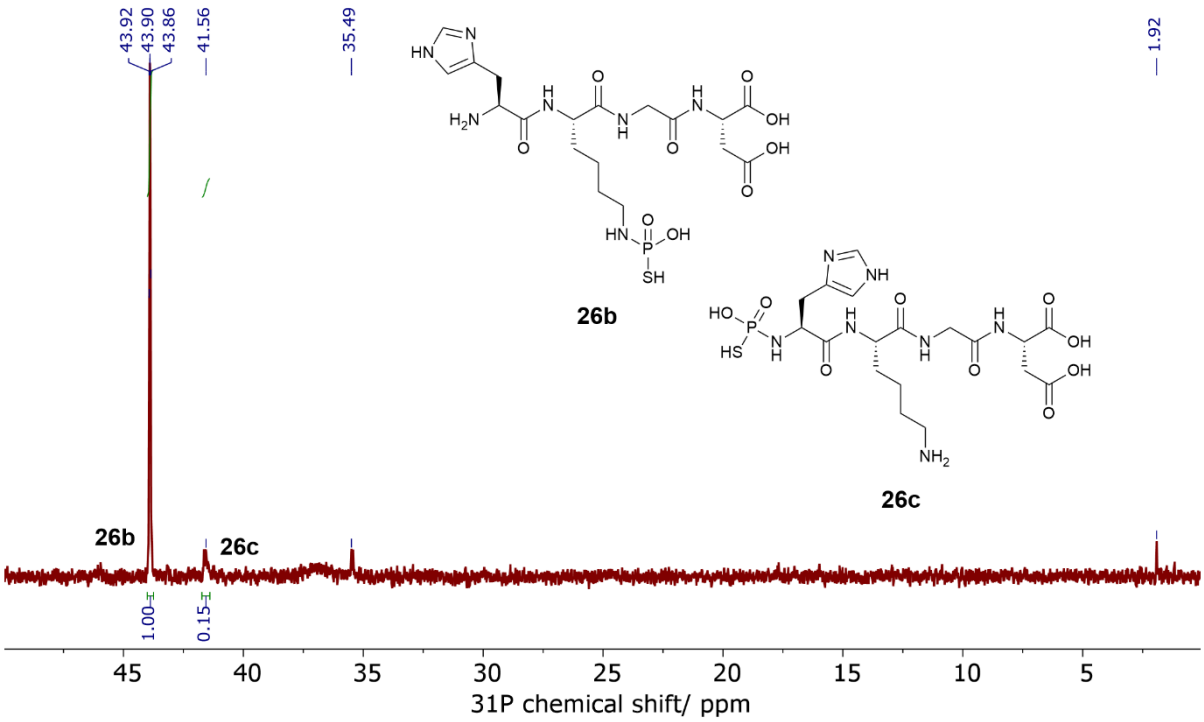


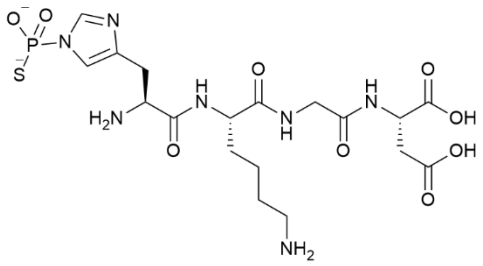
Figure 3.28 ^{31}P NMR spectrum following the *N*-thiophosphorylation of *N*-terminal histidine tetrapeptide **26a**. The structures of thiophosphoramidates **26b** and **26c** are shown for reference.

The signals at 43.90 and 41.56 ppm were identified in Figure 3.20 as ε-*N*-thiophosphoramidate **26b** and α-*N*-thiophosphoramidate **26c** respectively. The minor singlet signal at 1.92 ppm is inorganic phosphate.

The remaining signal at 35.49 ppm is representative of either 1-thiophosphohistidine or 3-thiophosphohistidine. While it is difficult to conclusively determine which of the two isomers the signal at 35.49 ppm represents, an informed estimation can be made.

The more thermodynamically stable isomer of thiophosphohistidine is 3-thiophosphohistidine.[149] The 1-*N* position of the histidine imidazole is sterically unfavoured due to the adjacent alkyl group so, despite the 1-*N* position being more nucleophilic than 3-*N*, thiophosphorylation occurs predominantly at the 3-*N* position of the histidine imidazole ring.[151] The observed ^{31}P chemical shift for the thiophosphohistidine product is 35.49 ppm. This chemical shift value aligns well with the ^{31}P chemical shift for 3-thiophosphohistidine reported in the literature (35.2 ppm).[151]

Based on the fact that 3-thiophosphohistidine is the more stable isomer of thiophosphorylated histidine, and the observed thiophosphohistidine signal in Figure 3.28 is extremely similar to that of 3-thiophosphohistidine reported in the literature, it was hypothesised that the thiophosphohistidine isomer formed during the thiophosphorylation of **26a** was 3-thiophosphohistidine. The structure of 3-thiophosphohistidine tetrapeptide **26d** is displayed in Figure 3.29. Table 3.4 summarises the identities of each ^{31}P signal noted in Figure 3.28.



3-thiophosphohistidine

26d

Figure 3.29 Proposed structure of 3-thiophosphohistidine tetrapeptide **26d** by-product.

Table 3.4 Summary of the identities and apparent proportions of each ^{31}P signal seen in Figure 3.28. The proportion of each phosphorus-containing species present was determined from the ^{31}P signal integrations in Figure 3.28.

^{31}P chemical shift/ ppm	**Identity**	**Proportion/ %**
1.92	Inorganic phosphate	3
35.49	**26d**	7
41.56	**26c**	12
43.90	**26b**	78

Through consideration of the nucleophilicity of cysteine and histidine side chains (8.18 and 6.00 respectively)[154], it was expected that thiophosphorylation of *N*-terminal cysteine tetrapeptide **25a** would result in a greater proportion of by-products compared to the thiophosphorylation of *N*-

terminal histidine tetrapeptide **26a**. This expectation was realised since, as can be seen in Tables 3.3 and 3.4, a far greater proportion of $PSCl_3$ was consumed in the formation of *N*-terminal histidine ε-*N*-thiophosphoramidate **26b** than *N*-terminal cysteine ε-*N*-thiophosphoramidate **25b** (78% compared to 55%). Also, when comparing the direct thiophosphorylation of cysteine and histidine residues, 2.5 × more $PSCl_3$ was consumed in the synthesis of cysteine dithiophosphoramidate **25f** compared to the synthesis of 3-thiophosphohistidine tetrapeptide **26d** (18% and 7% respectively [Tables 3.3 and 3.4]).

Based on the results summarised in Table 3.4, the presence of an unprotected histidine residue during peptide/ protein thiophosphorylation does not significantly compete with aqueous *N*-thiophosphorylation at an ε-amine group under strongly basic conditions. It is, therefore, not necessary to avoid the inclusion of histidine residues in peptides bound for chemoselective ε-NH_2 thiophosphorylation.

3.4.3. *S*-alkylation and isolation of *S*-naphthyl ε-*N*-thiophosphoramidate **23d**

Having determined that aqueous *N*-thiophosphorylation of unprotected peptides is selective for ε-NH_2 groups over α-NH_2 in section 3.4.2.1, a one-pot, two-step *N*-thiophosphorylation and *S*-alkylation reaction was performed on *N*-terminal phenylalanine tetrapeptide **23a** for the synthesis and subsequent isolation of *S*-naphthyl ε-*N*-thiophosphoramidate **23d**.

The synthesis and isolation of *S*-naphthyl ε-*N*-thiophosphoramidate **23d** was carried out in order to ensure that the alkylated thiophosphopeptide was stable under preparative reverse phase chromatography conditions, as **23d** is more closely related to the target stapled peptide system that will be discussed in Chapter 5 than *S*-propyl ε-*N*-thiophosphoramidate **8** which was previously isolated (section 2.2.1).

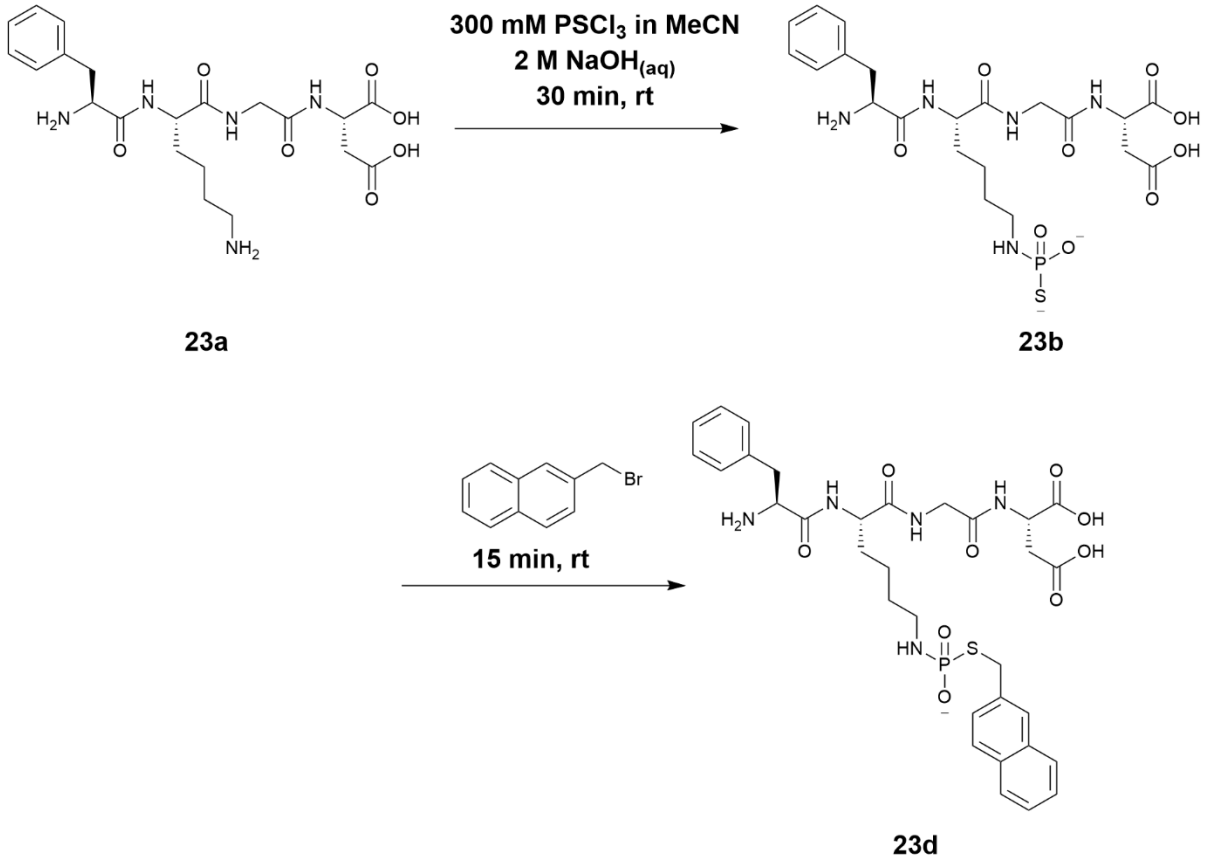
3.4.3.1. Synthesis of *S*-naphthyl ε-*N*-thiophosphoramidate **23d**

N-terminal phenylalanine tetrapeptide **23a** was chosen for this isolation experiment as it gave the greatest selectivity for ε-*N*-thiophosphoramidate formation (discussed in section 3.4.2.1). It was therefore hypothesised to give rise to the greatest overall conversion to the target *S*-naphthyl ε-*N*-thiophosphoramidate out of unprotected tetrapeptides **21-26a**.

Aqueous *N*-thiophosphorylation of **23a** was carried out using 7.0 eq $NaOH_{(aq)}$ and 0.75 eq $PSCl_3$ solution. The number of equivalents of $PSCl_3$ used was reduced to minimise the formation of α-*N*-thiophosphoramidate **23c**, thus attempting to maximise the proportion of the crude *S*-alkylation mixture consisting of *S*-naphthyl ε-*N*-thiophosphoramidate **23d**. Following *N*-thiophosphorylation, 1.0 eq 2-(bromomethyl)naphthalene solution was added for the *S*-alkylation step. 2-

(bromomethyl)naphthalene was chosen as the alkylating agent to facilitate the use of LC-MS DAD analysis alongside ^{31}P NMR spectroscopy to obtain a more accurate picture of the complex product mixture formed following the one-pot *N*-thiophosphorylation and *S*-alkylation reactions.

The reaction scheme for the synthesis of *S*-naphthyl ε-*N*-thiophosphoramidate **23d** is shown in Scheme 3.6 and experimental details can be found in section 7.2.5.1.



Scheme 3.6 One-pot, two-step *N*-thiophosphorylation and *S*-alkylation of *N*-terminal phenylalaninyl peptide **23a** to give *S*-naphthyl ε-*N*-thiophosphoramidate **23d** via ε-*N*-thiophosphoramidate **23b**.

^{31}P NMR spectra were recorded for the crude reaction mixtures following both the *N*-thiophosphorylation and *S*-alkylation steps, and these spectra can be seen in Figure 3.30.

The teal ^{31}P spectrum in Figure 3.30 shows the results following *N*-thiophosphorylation of **23a**. According to signal integrations, the ε-NH_2 (**23b**) to α-NH_2 (**23c**) thiophosphoramidate ratio is 20, suggesting that the reduction in the number of equivalents of $PSCl_3$ used during the thiophosphorylation stage did result in enhanced thiophosphorylation selectivity.

The red ^{31}P spectrum in Figure 3.30 shows the results following the *S*-alkylation step. Although the *S*-alkylation reaction did not result in the total consumption of ε-*N*-thiophosphoramidate **23b** (apparent

triplet signal highlighted in blue at 43.78 ppm on red ^{31}P spectrum in Figure 3.30), an approximate 81% conversion to *S*-naphthyl ε-*N*-thiophosphoramidate **23d** (apparent quintet signal highlighted in red around 23.12 ppm on red ^{31}P spectrum in Figure 3.30) was seen.

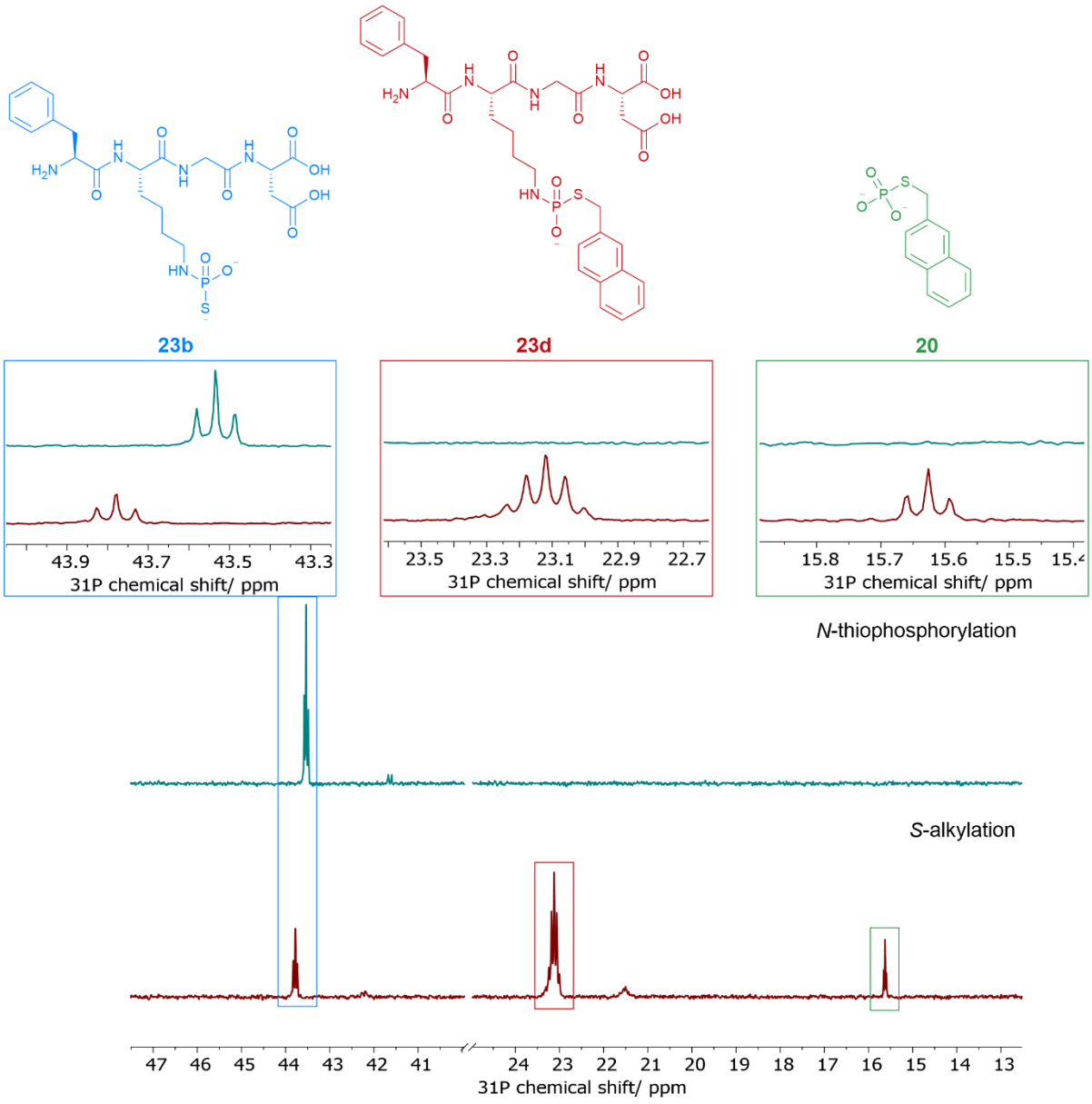


Figure 3.30 ^{31}P NMR spectra following *N*-thiophosphorylation and *S*-alkylation of *N*-terminal phenylalaninyl peptide **23a**. The expanded regions represent: (blue) the apparent triplet signal ca. 43.5-43.9 ppm corresponding to ε-*N*-thiophosphoramidate **23b**; (red) the apparent quintet signal at 23.1 ppm corresponding to *S*-naphthyl ε-*N*-thiophosphoramidate **23d**; (green) the triplet signal at 15.6 ppm corresponding to *S*-naphthyl thiophosphate **20**.

A triplet signal at 15.63 ppm can also be seen on the red ^{31}P spectrum in Figure 3.30, which corresponds to *S*-naphthyl thiophosphate **20** (structure shown in green in Figure 3.30). The one-pot *N*-thiophosphorylation and *S*-alkylation reactions carried out for the synthesis of *S*-naphthyl ε-*N*-thiophosphoramidate **23d** were a slightly adjusted method to that used in the competition reactions

outlined in section 3.3. Therefore, the presence of a triplet signal corresponding to **20** in Figure 3.30 supports the hypothesis presented in section 3.3.2.1 that *S*-naphthyl thiophosphate **20** was formed during the *S*-alkylation step of the competition reactions, thus interfering with the *N*-thiophosphorylation selectivity calculations based on LC-MS results. However, it remains unclear to what extent *S*-naphthyl thiophosphate **20** is formed through hydrolysis of the thiophosphoramidate N-P bond or through the direct *S*-alkylation of thiophosphate.

While ^{31}P NMR spectroscopy has been a valuable tool for the determination of thiophosphorylation selectivity between ε-NH_2 and α-NH_2, it is only able to show the presence of species containing a phosphorus atom. Consequently, the UV active alkylating agent 2-(bromomethyl)naphthalene was used to facilitate the use of LC-MS DAD analysis alongside ^{31}P NMR to further understand the product mixture following *S*-alkylation. Figure 3.31 shows the LC-MS total absorbance chromatogram generated following *S*-alkylation of the *N*-terminal phenylalanine thiophosphoramidate (**23b** and **23c**) mixture. No LC-MS data were collected prior to the *S*-alkylation stage as neither tetrapeptide **23a** nor thiophosphoramidates **23b**/ **23c** contain any significantly UV active components and so are not visible on LC-UV-vis chromatographs.

As stated previously, the *N*-thiophosphorylation stage of this synthesis proceeded with an excess of *N*-terminal phenylalanine tetrapeptide **23a**. The excess of **23a** explains the presence of different *N*-naphthyl tetrapeptide species noted in Figure 3.31 (**23e** [only the ε-*N*-naphthalene isomer shown for simplicity] and **23h**).

As expected, LC-MS indicated a significant proportion of *S*-naphthyl ε-*N*-thiophosphoramidate **23d**. While the ^{31}P NMR spectrum in Figure 3.30 indicates the formation of the α-*N* isomer of **23d** (small signal at 21.5 ppm on red ^{31}P NMR spectrum – not highlighted), only the structure of the ε-*N* isomer, **23d**, is shown in Figure 3.31 for simplicity.

Figure 3.31 also suggests that some α,ε-di-thiophosphoramidate was formed during thiophosphorylation, and that this di-thiophosphoramidate has been *bis*-alkylated to give *bis*-*S*-naphthyl α,ε-*N*-di-thiophosphoramidate **23g** (*m/z* 938 Da).

Due to the complexity of the mixture obtained following *S*-alkylation, it was not possible to determine the relative proportions of each product highlighted in Figure 3.31, especially as neither ^{31}P NMR spectroscopy nor LC-MS DAD analysis were able to show how much *N*-terminal phenylalanine tetrapeptide **23a** was consumed over the course of the reactions.

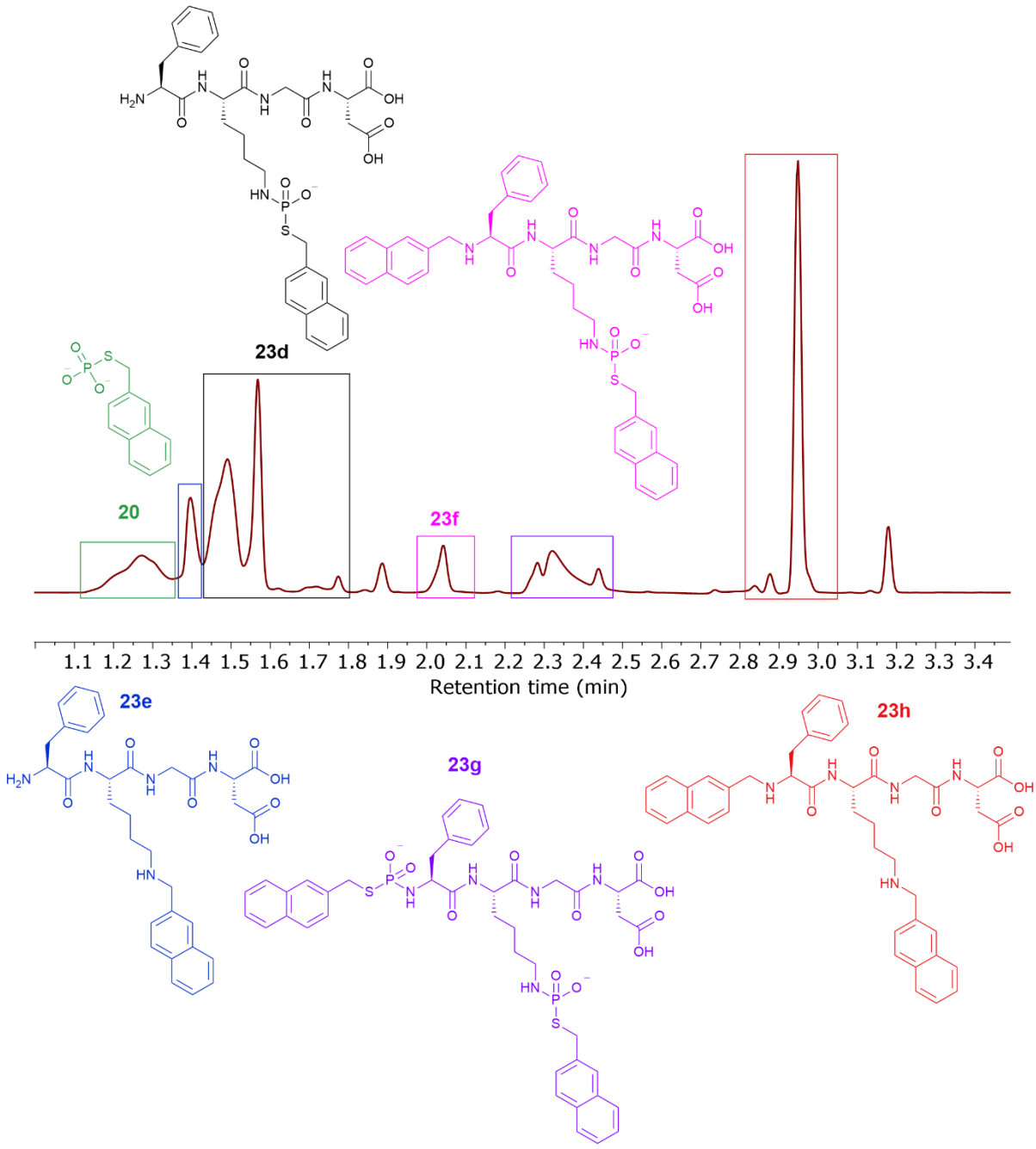


Figure 3.31 Crude LC-MS total absorbance chromatograph following one-pot *N*-thiophosphorylation and *S*-alkylation of *N*-terminal phenylalanine **23a** to synthesise *S*-naphthyl ε-*N*-thiophosphoramidate **23d** (indicated by the black box – *m/z* 702 Da). Various by-products are highlighted, and their structures shown. From left to right: *S*-naphthalated thiophosphate **20** (in green – *m/z* 255 Da); ε-*N*-naphthyl tetrapeptide **23e** (in blue – *m/z* 606 Da); α-*N*-naphthyl *S*-naphthyl ε-*N*-thiophosphoramidate **23f** (in pink – *m/z* 842 Da); *bis*-*S*-naphthyl α,ε-*N*-di-thiophosphoramidate **23g** (in purple – *m/z* 938 Da); α,ε-*N*-dinaphthyl tetrapeptide **23h** (in red – *m/z* 746 Da).

Having successfully synthesised crude *S*-naphthyl ε-*N*-thiophosphoramidate **23d**, the isolation of **23d** could now begin.

3.4.3.2. Isolation and characterisation of *S*-naphthyl ε-*N*-thiophosphoramidate **23d**

Due to its polar nature, *S*-naphthyl ε-*N*-thiophosphoramidate **23d** was isolated through preparative reverse phase chromatography, on a Teledyne ISCO CombiFlash system using a 15.5 g HP C_{18} Gold column and water/ MeCN gradient mobile phase (further details in section 7.2.5.1).

Figures 3.32 and 3.33 show the ^{31}P NMR and LC-UV-vis data collected following isolation of **23d** respectively. For both figures, the isolated spectra are compared with the spectra obtained from the crude material prior to purification.

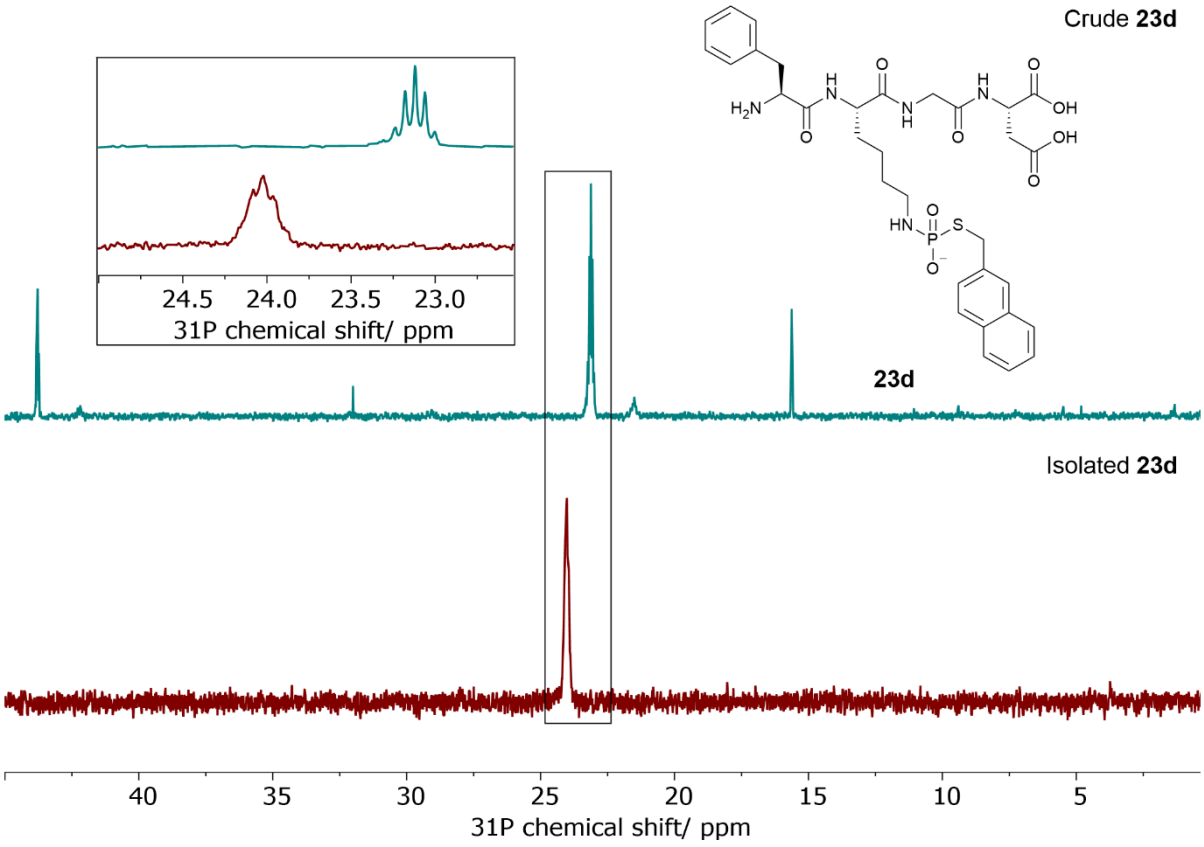


Figure 3.32 ^{31}P NMR spectra of crude **23d** (top spectrum) and isolated **23d** (bottom spectrum). The zoomed in region between 22.5 and 25.0 ppm represents the signal corresponding to *S*-naphthyl ε-*N*-thiophosphoramidate **23d**.

Figure 3.32 shows that **23d** was isolated to 100% purity, based on the ^{31}P NMR spectrum taken after purification (the red spectrum in Figure 3.32) only showing a signal at 24.02 ppm – indicative of *S*-naphthyl ε-*N*-thiophosphoramidate **23d**. Compound **23d** was also found to be stable at 4 °C over three weeks by ^{31}P NMR analysis (spectrum not shown).

However, as stated in section 3.4.3.1, ^{31}P NMR spectroscopic analysis can only give information on the presence of phosphorus-containing compounds. LC-MS DAD analysis was, therefore, employed to

examine whether *S*-naphthyl ε-*N*-thiophosphoramidate **23d** had been isolated from all other naphthalated species noted in Figure 3.31.

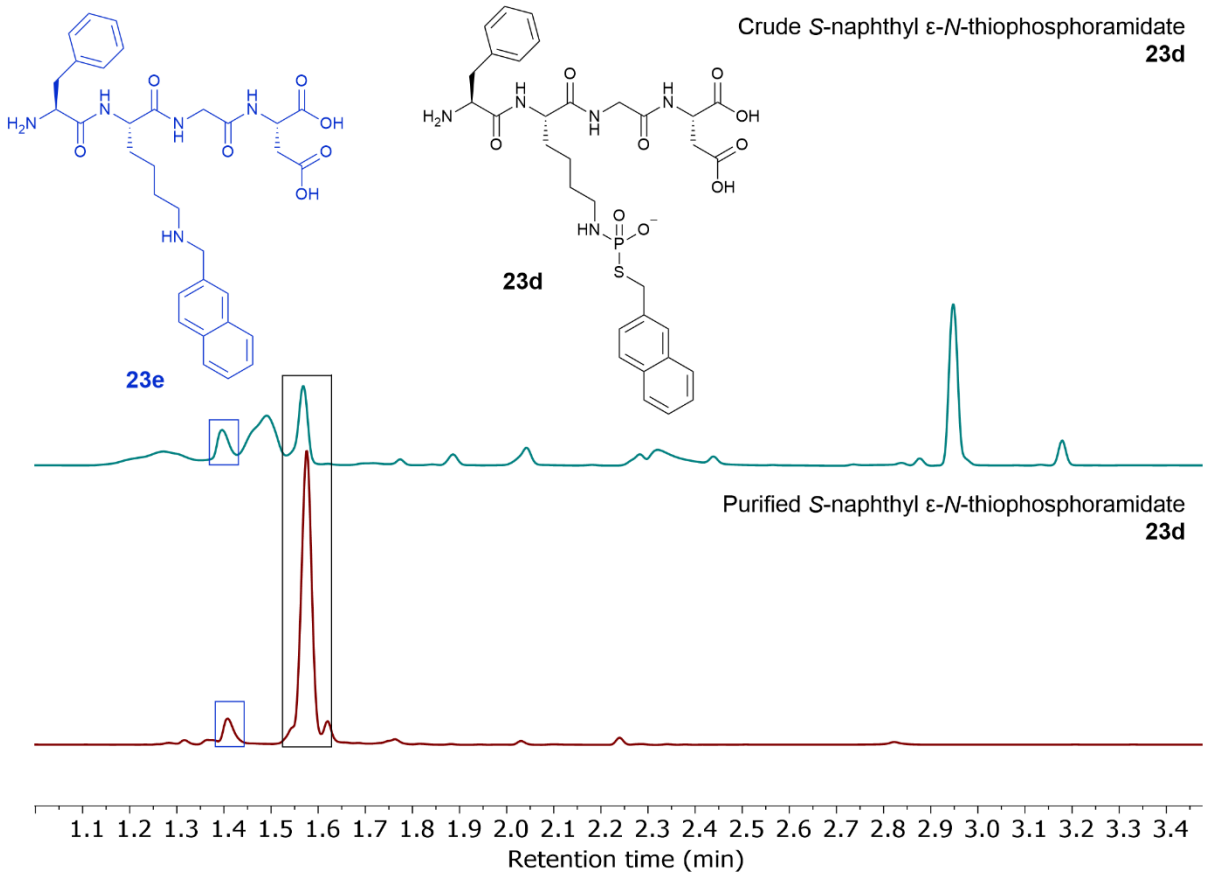


Figure 3.33 LC-MS total absorbance chromatograms before (top trace) and after (bottom trace) purification and attempted isolation of *S*-naphthyl ε-*N*-thiophosphoramidate **23d** (shown in black – *m/z* 702 Da) by preparative reverse phase chromatography. The major remaining impurity, ε-*N*-naphthyl tetrapeptide **23e** (in blue – *m/z* 606 Da), is also highlighted.

According to LC-UV-vis analysis, the isolated product was not 100% homogeneous. *S*-naphthyl ε-*N*-thiophosphoramidate **23d** was purified to 84% homogeneity, with *N*-naphthalated tetrapeptide **23e** (8%) seen as the largest impurity (cannot determine whether **23e** is the α-*N* or ε-*N* isomer or a mixture of the two – the ε-*N* isomer is shown as an example).

As compound **23d** was not being used for any further investigations or analysis, it was decided that an 84% homogenous sample was adequate to allow confirmation of the desired thiophosphoramidate species. Through the purification and characterisation of this *S*-alkylated thiophosphoramidate, we were able to demonstrate that it was stable under neutral chromatographic conditions and could be stored at 4 °C for three weeks with minimal degradation. These results suggest that a peptide stapled via *S*-alkylation of a thiophosphoramidate intermediate (see Chapter 5) will be stable during chromatography and, therefore, be able to be isolated.

3.5. Summary

Following the optimisation of *N*-thiophosphorylation on α-*N*-dansyl-labelled tetrapeptide model **9**, an investigation into the selectivity of *N*-thiophosphorylation between α- and ε-amine groups began. Initial comparisons of *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide model **9** and ε-*N*-dansyl-labelled tetrapeptide **11** using the technique described in section 3.1, suggested that thiophosphorylation may occur preferentially at an ε-NH_2 group rather than at an *N*-terminal α-NH_2 group.

N-thiophosphorylations were then carried out on a pentapeptide model (**13**) with both an unprotected *N*-terminus and lysine side chain, with the aim of using tandem MS analysis to locate the preferred *N*-thiophosphorylation site based on fragmentation patterns. Unfortunately, fragmentation of both mono-thiophosphoramidate (**14a** and **14b**) and *S*-alkylated mono-thiophosphoramidate (**15a** and **15b**) derivatives of pentapeptide **13** indicated that the first losses seen were either that of the thiophosphate group (from **14a/14b**) or the alkyl thiol (from **15a/15b**). For that reason, it was determined that tandem MS analysis was not appropriate for investigating the selectivity of the aqueous *N*-thiophosphorylation reaction.

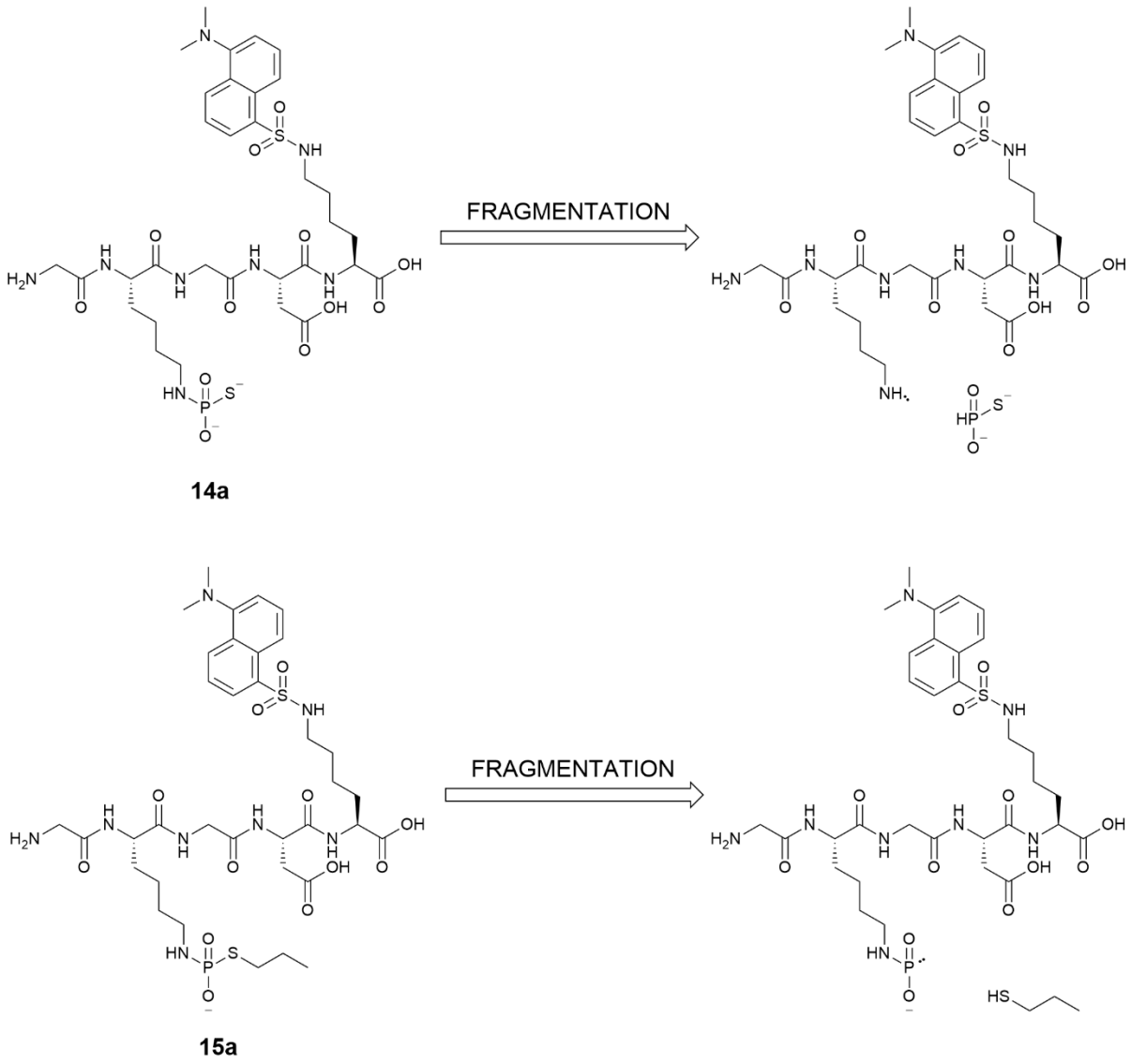


Figure 3.34 First loss fragmentations seen by tandem MS analysis for *N*-thiophosphoramidates **14a** and **14b** and *S*-propyl *N*-thiophosphoramidates **15a** and **15b**. The ε-*N*-thiophosphoramidate derivatives (**14a** and **15a**) of pentapeptide **13** are depicted here for simplicity.

While still probing the scope of MS-based analysis methods for investigating the selectivity of *N*-thiophosphorylation, one-pot, two-step *N*-thiophosphorylation and *S*-alkylation competition reactions were performed and analysed by LC-MS DAD analysis. These competition reactions employed α-*N*-acetyl protected tetrapeptide **16** (for thiophosphorylation at the lysine side chain ε-NH_2) and ε-*N*-dansyl-labelled tetrapeptide **11** (for thiophosphorylation at the *N*-terminal α-NH_2). Since tetrapeptide **16** contained no UV active components, the thiophosphoramidates formed (ε-*N*-thiophosphoramidate **17** and α-*N*-thiophosphoramidate **12**) were 'trapped' via *S*-alkylation with 2-(bromomethyl)naphthalene to generate α-*N*-acetyl *S*-naphthyl ε-*N*-thiophosphoramidate **18** and ε-*N*-dansyl α-*N*-thiophosphoramidate **19** (Scheme 3.4).

Total absorbance chromatographs were taken following the *N*-thiophosphorylation and *S*-alkylation steps, and *N*-thiophosphorylation reaction selectivity was calculated through the method outlined in section 3.3.1.1. When done in triplicate, it was found that the *N*-thiophosphorylation selectivities calculated by this method were irreproducible. Upon further analysis of the *S*-alkylation chromatographs, it was found that a UV active by-product (*S*-naphthyl thiophosphate **20**) had a very similar retention time as α-*N*-acetyl *S*-naphthyl ε-*N*-thiophosphoramidate **18** (Figures 3.16 – 3.18). The co-elution of **20** with **18** led to the over-representation of **18** by LC-UV-vis analysis, and therefore gave inaccurate *N*-thiophosphorylation selectivities calculated by LC-MS total absorbance chromatographs through overstating the proportion of ε-*N*-thiophosphoramidate compared to α-*N*-thiophosphoramidate.

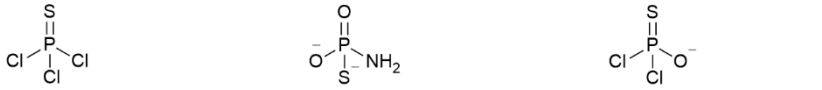
As neither MS-based analysis method was able to give reproducible selectivity data for *N*-thiophosphorylations between an α- and ε-amino groups, an ^{31}P NMR spectroscopy-based assay was developed. Six α,ε-unprotected tetrapeptides with different *N*-terminal amino acids (**21-26a**) were thiophosphorylated individually and analysed by ^{31}P NMR spectroscopy. As the ε-*N*-thiophosphoramidates and α-*N*-thiophosphoramidates gave different ^{31}P chemical shifts due to the differing chemical environments, calculation of the *N*-thiophosphorylation selectivity simply involved comparison of signal integrations (section 3.4.2.1). From this ^{31}P NMR-based assay, it was found that *N*-terminal glycine tetrapeptide **21a** gave the lowest selectivity with the ε-NH_2 to α-NH_2 thiophosphorylation ratio being 1.4. *N*-terminal phenylalanine tetrapeptide **23a** gave the highest selectivity with the ε-NH_2 to α-NH_2 thiophosphorylation ratio being 14.3. It was also concluded that the size of the *N*-terminal residue side chain (i.e. steric effects) is the key factor in determining how selective *N*-thiophosphorylation will be for ε-NH_2 compared to α-NH_2 when performed on an unprotected peptide.

With respect to the impact of other nucleophilic residues (i.e. cysteine and histidine) on the selectivity of thiophosphorylation, it was found that thiophosphorylation of the cysteine thiol competes with lysine ε-NH_2 thiophosphorylation (18% vs 55% respectively). On the other hand, histidine thiophosphorylation was minimal (7%) compared to ε-*N*-thiophosphorylation (78%) under aqueous hydroxide conditions.

Finally, *S*-naphthyl ε-*N*-thiophosphoramidate **23d** was synthesised and part-purified to ensure that *S*-alkylated thiophosphopeptides are stable under reverse phase chromatography conditions. According to ^{31}P NMR spectroscopy, **23d** was able to be isolated as a 100% homogeneous product, while LC-UV-vis analysis indicated the continued presence of (non-phosphorus-containing) by-product *N*-naphthyl tetrapeptide **23e** – giving **23d** a purity of 84% by LC-UV-vis analysis.

4. Investigation of alternative conditions for aqueous *N*-thiophosphorylation of peptide substrates

During the introduction of this thesis (specifically section 1.4.1), several chemical thiophosphorylation methods were discussed, using one of three thiophosphorylating agents. These thiophosphorylating agents were $PSCl_3$ (already used throughout experimental work discussed in Chapters 2 and 3), thiophosphoramidate ions, and thiophosphorodichloridate ions ($PSOCl_2^-$).



Thiophosphoryl chloride $PSCl_3$ | Thiophosphoramidate | Thiophosphorodichloridate ion $PSOCl_2^-$

Figure 4.1 Structures of three different thiophosphorylating agents. From left to right: thiophosphoryl chloride ($PSCl_3$); thiophosphoramidate ion; and thiophosphorodichloridate ion ($PSOCl_2^-$).

In 2000, Pirrung *et al.* noted poor conversions from histidine to 3-thiophosphohistidine in the presence of excess potassium thiophosphoramidate, with the hydrolysis of potassium thiophosphoramidate being the dominant process. Following a 3-day reaction, approximately 5% of the phosphorus-containing species were reported as being 3-thiophosphohistidine, while 80% was inorganic thiophosphate.[151] Thus, it was decided that thiophosphoramidate ions would not be explored for peptide thiophosphorylation during this project.

Thiophosphorodichloridate ions were used as a thiophosphorylating agent by the Hodgson group in 2014, for the *N*-thiophosphorylation on 5'-aminonucleosides.[147] Researchers found that, compared with *N*-thiophosphorylation of 5'-aminonucleosides using $PSCl_3$, $PSOCl_2^-$ typically gave reduced conversions to thiophosphorylated 5'-aminonucleosides (on average, reduced by 9 to 11%). Despite this, these conversions to *N*-thiophospho-5'-aminonucloside using $PSOCl_2^-$ were generally greater than 80%.[147]

While the results from the thiophosphorylation of 5'-aminonucleosides with $PSOCl_2^-$ suggested that peptide thiophosphorylation using $PSOCl_2^-$ may not generate the high conversions seen with $PSCl_3$, the potential of $PSOCl_2^-$ as a thiophosphorylating agent for peptide substrates was investigated. This decision was based on the fact that the work previously carried out with $PSOCl_2^-$ by the Hodgson group was limited to the 5'-aminonucleoside thiophosphorylations.[147] As there had not been any research into selective aqueous *N*-thiophosphorylation of peptide substrates with $PSOCl_2^-$, the scope of $PSOCl_2^-$ as a thiophosphorylating agent was tested. Section 4.1 of this chapter will begin by discussing the results obtained from aqueous *N*-thiophosphorylations carried out using $PSOCl_2^-$ as the

thiophosphorylating agent and comparing these results to those obtained from thiophosphorylations using $PSCl_3$. In section 4.2, alternative aqueous base conditions (the use of triethylammonium bicarbonate [TEAB]) for peptide thiophosphorylation will be discussed. Since harsh basic conditions can lead to epimerisation at α-C centres, the use of a milder source of base, such as TEAB, could minimise this risk. A secondary benefit for trialling TEAB solution as the base during peptide thiophosphorylation is the inherent pH control when using buffered solutions, which is particularly advantageous for laboratories without a pH stat set up.

The *N*-thiophosphorylation of 5'-aminonucleosides using potassium thiophosphorodichloridate ($KPSOCl_2$) previously performed by Conway *et al* in 2014 implemented an autotitrator to add 1 M potassium hydroxide solution in order to hold the reaction pH at a predetermined value over the reaction. Through the use of this pH stat set up, it was found that pH 12 led to the greatest conversion to thiophosphoramidate from 5'-amino-5'-deoxyguanosine when using 1 M $KOH_{(aq)}$ and 1.0 eq $KPSOCl_2$.[147] By using a buffer solution (i.e. TEAB solution) as the base source during aqueous *N*-thiophosphorylation, it was hoped that the pH control afforded by a pH stat could be imitated without requiring the use of one.

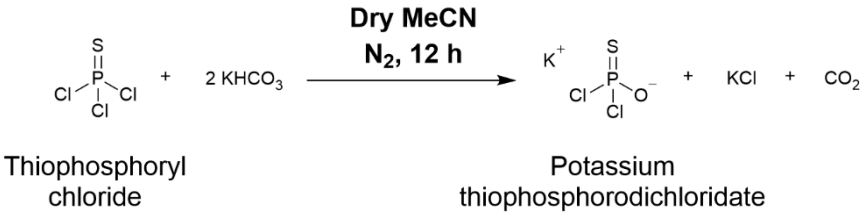
4.1. *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide 9 using potassium thiophosphorodichloridate and $NaOH_{(aq)}$

It was hypothesised that potassium thiophosphorodichloridate ($KPSOCl_2$) could be an effective alternative for aqueous *N*-thiophosphorylation to $PSCl_3$ due to its solubility in water. A large pH-independent hydrolysis range (from ~2 to ~13) and a half-life in water of ca. 3.2 min also added to the idea that $KPSOCl_2$ could support homogeneous reactions under aqueous conditions.[153]

Unlike $PSCl_3$, however, $KPSOCl_2$ requires synthesis prior to use. This section will briefly detail the synthesis of $KPSOCl_2$ before discussing the optimisation of aqueous *N*-thiophosphorylation on α-*N*-dansyl-labelled tetrapeptide **9**.

4.1.1. Synthesis of potassium thiophosphorodichloridate ($KPSOCl_2$)

The procedure followed for the synthesis of $KPSOCl_2$ was reported by Delley *et al*. in 2012,[153] based on earlier work published in 2004.[165] Scheme 4.1 shows the general procedure followed.



Scheme 4.1 Synthetic scheme for the synthesis of potassium thiophosphorodichloridate ($KPSOCl_2$). Procedure followed from Delley *et al*. (2012).[153]

A solution of $KPSOCl_2$ in MeCN was separated from KCl, evolved during the reaction, by filtration, and an assessment of P-containing species within the sample was made using ^{31}P NMR spectroscopy. Figure 4.2 shows the resultant ^{31}P NMR spectrum.

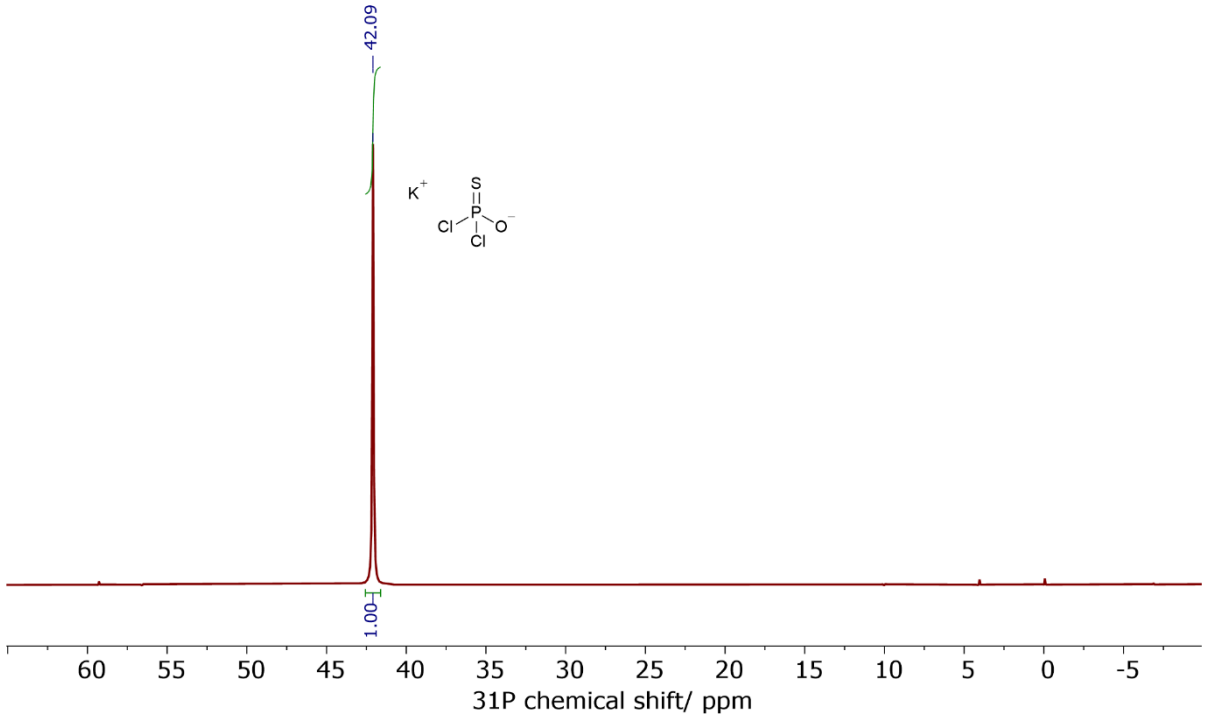


Figure 4.2 ^{31}P NMR spectrum of potassium thiophosphorodichloridate ($KPSOCl_2$).

The singlet signal at 42.09 ppm was believed to correspond to $KPSOCl_2$ due to it close proximity to the ^{31}P chemical shift value reported by Delley (39.8 ppm).[153] The ^{31}P NMR spectrum of $KPSOCl_2$ in Figure 4.2 shows 99% homogeneity, with minor signals at -0.05, 4.04, and 59.28 ppm totalling ~1% (integrals not shown).

4.1.2. Optimisation of aqueous *N*-thiophosphorylation using $KPSOCl_2$ as the thiophosphorylating agent

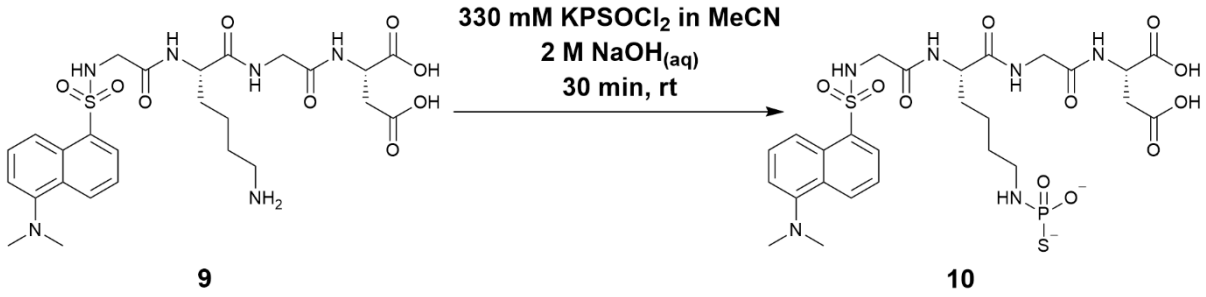
With a $KPSOCl_2$ solution in MeCN prepared to 99% homogeneity with thiophosphorodichloridate ion concentration assumed to be 330 mM based on the concentrations of reagents used in the

preparation, work began on the optimisation of aqueous *N*-thiophosphorylation using $KPSOCl_2$ as the thiophosphorylating agent. Optimisation reactions using $KPSOCl_2$ were carried out using α-*N*-dansyl-labelled tetrapeptide **9**.

Tetrapeptide model **9** was used for this optimisation series using $KPSOCl_2$ and $NaOH_{(aq)}$ as there were LC-MS data available from a previous *N*-thiophosphorylation optimisation series using $PSCl_3$ and $NaOH_{(aq)}$ (see section 2.3). These two sets of data could therefore be compared to see which thiophosphorylating agent resulted in a greater conversion to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** from α-*N*-dansyl-labelled tetrapeptide **9**.

Using α-*N*-dansyl-labelled tetrapeptide **9** also facilitates the use of LC-MS DAD analysis, allowing for smaller scale reactions and thus reducing the material cost of each reaction when compared to reactions carried out for ^{31}P NMR spectroscopic analysis.

The general reaction procedure for the aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** was adapted from the procedure discussed in section 2.3, and the procedure is summarised in Scheme 4.2.



Scheme 4.2 Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** to generate α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** using $KPSOCl_2$ as the thiophosphorylating agent.

A preliminary *N*-thiophosphorylation reaction was carried out using 1.2 eq $KPSOCl_2$ in MeCN solution and 7.0 eq $NaOH_{(aq)}$. A small excess of $KPSOCl_2$ was used rather than a 1:1 ($KPSOCl_2$: tetrapeptide **9**) ratio because, based on the work conducted by Conway *et al* (2014),[147] $KPSOCl_2$ was assumed to result in lower conversions to thiophosphoramidate than $PSCl_3$. Figure 4.3 shows the LC-MS total absorbance chromatograph obtained following the aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** using 1.2 eq $KPSOCl_2$ and 7.0 eq $NaOH_{(aq)}$.

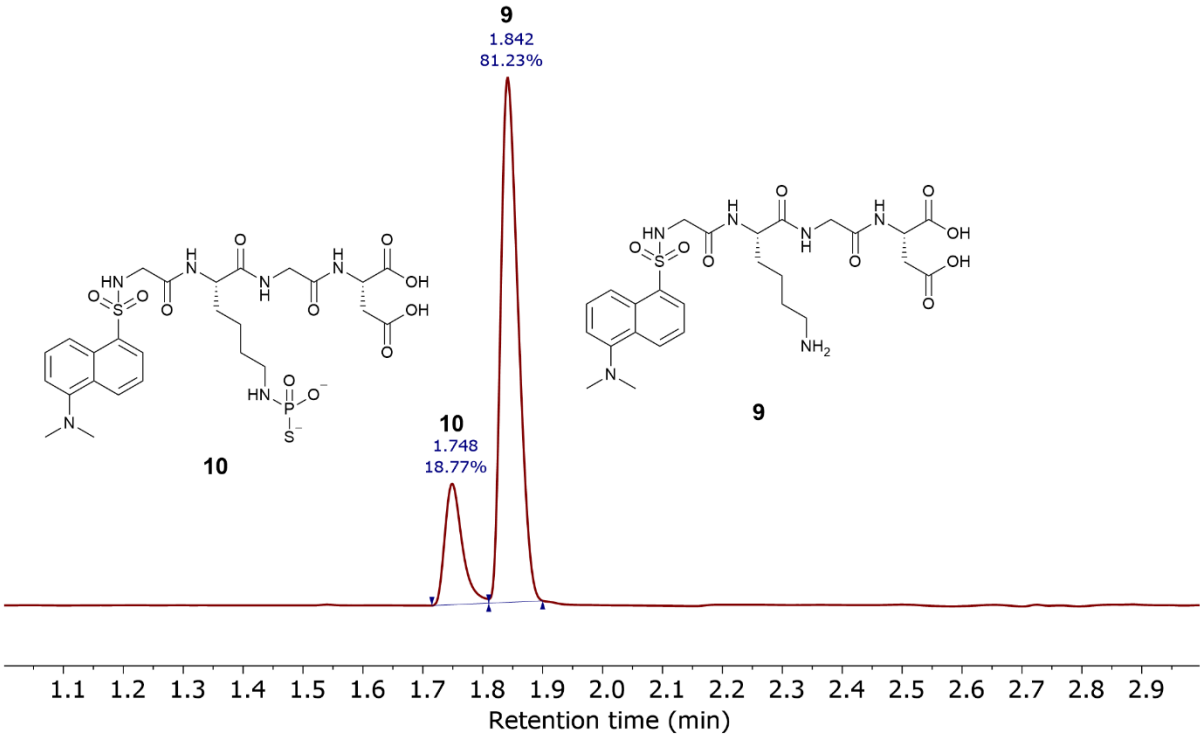


Figure 4.3 LC-MS total absorbance chromatograph following aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** to give α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** with 1.2 eq $KPSOCl_2$ and 7.0 eq $NaOH_{(aq)}$. The concentrations of each reagent in the total reaction volume were: α-*N*-dansyl-labelled tetrapeptide **9** = 3.9 mM; $KPSOCl_2$ = 4.8 mM; $NaOH_{(aq)}$ = 29.1 mM.

The total absorbance chromatograph in Figure 4.3 reveals that a conversion of 19% from α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** was obtained with 1.2 eq $KPSOCl_2$ and 7.0 eq $NaOH_{(aq)}$.

Based on this poor conversion using 1.2 eq $KPSOCl_2$ and 7.0 eq $NaOH_{(aq)}$, a series on experiments increasing the number of equivalents of $KPSOCl_2$ was devised. The number of equivalents of $NaOH_{(aq)}$ used was also adjusted for each thiophosphorylation to account for the acid that forms as a result of consumption of $KPSOCl_2$. The results collected from this *N*-thiophosphorylation series are shown in both Table 4.1 and Figure 4.4, with the experimental details in section 7.3.2.1. The conversion results for **9** to **10** when using $PSCl_3$ are included in Table 4.1 to allow comparisons between the two thiophosphorylating agents.

Table 4.1 Aqueous *N*-thiophosphorylation series of α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** using $KPSOCl_2$ as the thiophosphorylating agent (in red). Data for conversions from **9** to **10** during aqueous *N*-thiophosphorylation using $PSCl_3$ shown for comparison (in black). Conversions calculated from area under relevant peaks seen in LC-MS total absorbance chromatographs.
[a] Number of equivalents of $PSCl_3$ used.

Thiophosphorylating agent/ eq	NaOH/ eq		Conversion from 9 to 10/ %	
	$KPSOCl_2$	$PSCl_3$	$KPSOCl_2$	$PSCl_3$
1.2 (1.0)[a]	7.0	7.5	19	82
2.5 (2.0)[a]	14.6	14.7	51	89
5.0	29.2	-	81	-
7.5	43.7	35.7	95	95
10.0	58.3	70.0	97	99

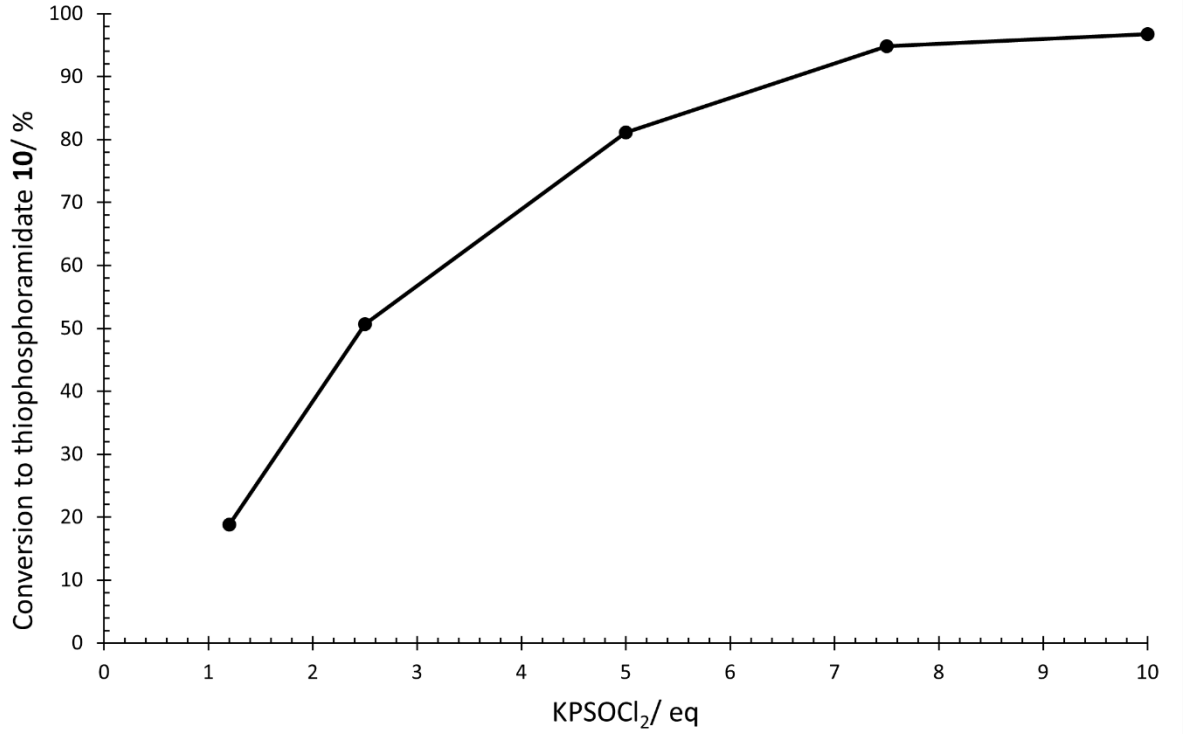


Figure 4.4 Percentage conversion of α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** during aqueous *N*-thiophosphorylations using $KPSOCl_2$ as the thiophosphorylating agent (graphic representation of the $KPSOCl_2$ data detailed in Table 4.1).

As with the optimisation of aqueous *N*-thiophosphorylation using $PSCl_3$ (section 2.3), increasing the number of equivalents of $KPSOCl_2$ resulted in an increase in the conversion of α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10**, with a plateau being reached around 7.5 to 10.0 eq $KPSOCl_2$. Based on the results shown in Table 4.1 and Figure 4.4, at least 5.0 eq $KPSOCl_2$ should be used for aqueous *N*-thiophosphorylation to obtain similar conversions to those seen using 1.0 eq $PSCl_3$.

While the conversions seen to **10** from **9** with 7.5 to 10.0 eq $KPSOCl_2$ closely match with those seen with $PSCl_3$, when using 1.0 to 2.5 eq $KPSOCl_2$ there is significantly less conversion to

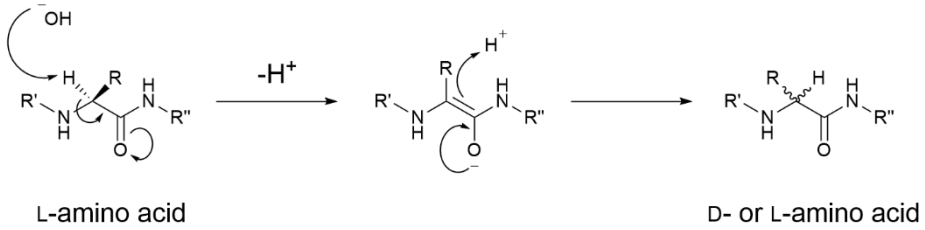
thiophosphoramidate when comparing to similar proportions of $PSCl_3$. This reduced conversion may be due to the differences in water solubility of the two thiophosphorylating agents. As previously explained in Chapter 2, thiophosphorodichloridate ions are soluble in water and have a half-life of ca. 3.2 min in the pH range ~2 to ~13,[153] thus resulting in a homogeneous *N*-thiophosphorylation reaction mixture. Aminolysis of $KPSOCl_2$ under aqueous conditions is therefore likely to be competing directly with $KPSOCl_2$ hydrolysis, with hydrolysis appearing to be the dominant process when using 1.0 to 2.5 eq $KPSOCl_2$, with reagent concentrations of amine and thiophosphorylating agent being 4 mM and 330 mM, respectively.

On the other hand, $PSCl_3$ is sparingly soluble in water and could remain in micro-droplets (possibly with MeCN cosolvent) or other phase-separated bodies following addition to the basic aqueous α-*N*-dansyl-labelled tetrapeptide **9** solution. This phase separation between $PSCl_3$ and water may give rise to greater reaction selectivity for $PSCl_3$, as was previously hypothesised by Conway *et al.* (2014), since $PSCl_3$ and tetrapeptide **9** (which shows limited solubility in MeCN) would largely react at $PSCl_3$-MeCN/water interfaces rather than in homogeneous solution.[147] Aminolysis at solvent interfaces rather than in a homogeneous mixture would, in turn, result in reduced loss of $PSCl_3$ through hydrolysis.

Having optimised the aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** using $KPSOCl_2$ as the thiophosphorylating agent and demonstrating its use as an alternative thiophosphorylating agent to $PSCl_3$, we began investigating whether milder basic conditions could be used as an alternative to the harsh conditions that have been employed up to this point.

4.2. Investigating the use of triethylammonium bicarbonate (TEAB) for aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide 9

Although using $NaOH_{(aq)}$ for aqueous *N*-thiophosphorylation on peptide substrates has resulted in > 90% conversions to thiophosphoramidate (as reported in Chapter 2), harsh basic conditions, such as the use of hydroxide solutions, have been reported to cause epimerisation at α-carbon centres through α-H abstraction.[166]



Scheme 4.3 Mechanism of epimerisation *via* base-catalysed proton abstraction.

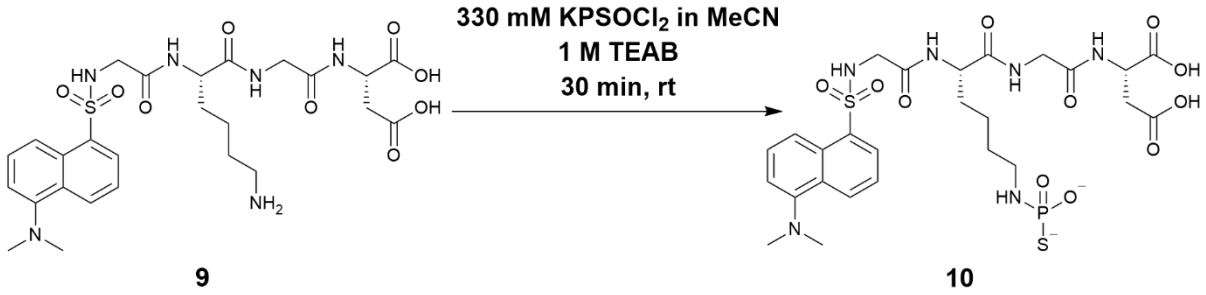
In order to minimise any risk of epimerisation, milder basic conditions were explored as an alternative to using sodium hydroxide. Triethylammonium bicarbonate (TEAB) solution was chosen as the alternative base due to its high volatility, allowing any excess to be easily removed following reaction.

The pH control afforded by buffer solutions was also a potential advantage of the use of TEAB solutions for aqueous *N*-thiophosphorylation. Not only does TEAB solution allow for pH-controlled reactions without the use of a pH stat but, as the TEAB buffering range in use for these reactions is between pH 10 and 11, the pH of the thiophosphorylation reaction will remain well within the pH reactivity plateau of $KPSOCl_2$ reported by Delley *et al*.[153] In addition, pH control in this reaction offers the potential to control the ionisation states of competing amine nucleophiles, thus potentially allowing for improvements in selectivity.

Aqueous *N*-thiophosphorylations were carried out on α-*N*-dansyl-labelled tetrapeptide **9** with TEAB using both $KPSOCl_2$ and $PSCl_3$ to appreciate how TEAB compares to $NaOH_{(aq)}$ for both thiophosphorylating agents.

4.2.1. Using $KPSOCl_2$ and TEAB as reagents for aqueous *N*-thiophosphorylation of peptide substrates

The general reaction procedure for aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** with TEAB and $KPSOCl_2$ was adapted from the procedure discussed in section 4.1.2 and is shown in Scheme 4.4.



Scheme 4.4 Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** to generate α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** using $KPSOCl_2$ as the thiophosphorylating agent and triethylammonium bicarbonate (TEAB) as the base.

An initial *N*-thiophosphorylation attempt using the conditions shown in Scheme 4.4 with 7.5 eq $KPSOCl_2$ and ca. 10 eq TEAB (pH 10) was unsuccessful, with no discernible conversion of α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** noted by LC-MS DAD analysis.

Based on this poor conversion result, it was decided that a series of *N*-thiophosphorylations would be carried out varying the number of equivalents of TEAB used, while consistently using 7.5 eq of $KPSOCl_2$. We chose to carry out these *N*-thiophosphorylations with 7.5 eq $KPSOCl_2$ rather than 10.0 eq as there had not been a significant increase in ε-*N*-thiophosphoramidate **10** formation, as noted in Table 4.1 with 10.0 eq $KPSOCl_2$ compared to 7.5 eq. The pH of the TEAB solution used was also raised from 10 to 10.5 in subsequent experiments, in order to have the pH of the *N*-thiophosphorylation mixture closer to the pK_{aH} of the lysine ε-NH_2.

4.2.1.1. Optimisation of the number of equivalents of 1 M TEAB for aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** using $KPSOCl_2$

A series of aqueous *N*-thiophosphorylations of α-*N*-dansyl-labelled tetrapeptide **9** using $KPSOCl_2$ was designed in which the numbers of equivalents of 1 M TEAB (pH 10.5) were varied against the numbers of equivalents of **9** used.

Starting with ca. 50 eq of TEAB, the experimental series was constructed by increasing the number of equivalents of TEAB by 25 eq increments up to 150.0 eq. Thereafter, from 150.0 eq to 350.0 eq TEAB, the number of equivalents of TEAB were increased in 50.0 eq increments. The same stock of 1 M TEAB (pH 10.5) was used for the initial series of *N*-thiophosphorylations. Using the same TEAB stock solution throughout the *N*-thiophosphorylation series is particularly important as getting consistent pH through the addition of CO_2 is difficult. As well as this, the loss of CO_2 from a TEAB solution will lead to pH changes. Both of these problems can be avoided by using the same stock solution of TEAB. The

same stock of 330 mM $KPSOCl_2$ in MeCN solution was also used for each of the 9 reactions included in the series.

The percentage conversions from α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** for each *N*-thiophosphorylation within the series are displayed in Figure 4.5 and Table 4.2. The results from this *N*-thiophosphorylation series were obtained from LC-MS DAD analysis, with the conversions calculated from the areas under peaks corresponding to tetrapeptide **9** and thiophosphoramidate **10** in the total absorbance chromatographs generated from LC-UV-vis detection.

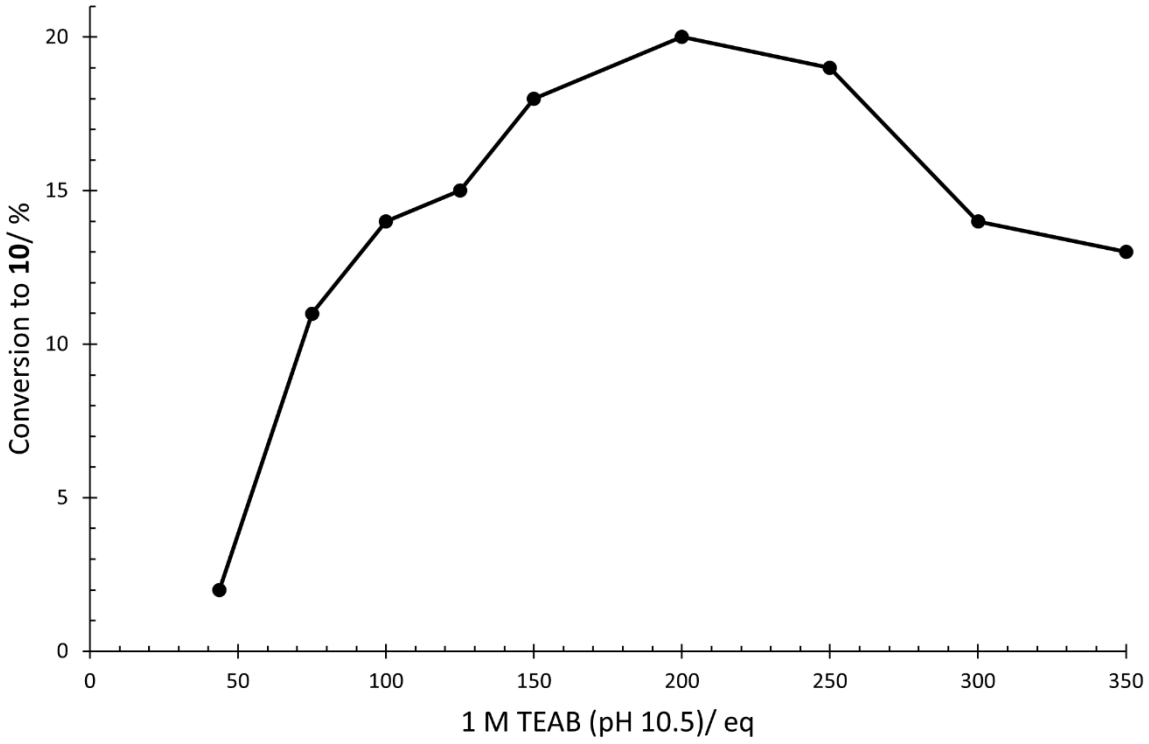


Figure 4.5 Percentage conversion of α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** during aqueous *N*-thiophosphorylation using 7.5 eq $KPSOCl_2$ as the thiophosphorylating agent and TEAB as the source of base (graphic representation of the $KPSOCl_2$ data detailed in Table 4.2). Concentration of solution of **10** used for each thiophosphorylation reaction – 4 mM.

Table 4.2 Data from the aqueous *N*-thiophosphorylation series of α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** using 7.5 eq $KPSOCl_2$ as the thiophosphorylating agent and TEAB as the base. Conversions calculated from areas under relevant peaks seen in LC-MS total absorbance chromatographs.

TEAB (pH 10.5)/ eq	Conversion from 9 to 10/ %
43.7	2
75.0	11
100.0	14
125.0	15
150.0	18
200.0	20
250.0	19
300.0	14
350.0	13

It is immediately apparent from these data that *N*-thiophosphorylations carried out with 1 M TEAB do not achieve conversion levels as high as those recorded when using $NaOH_{(aq)}$. The greatest conversion from tetrapeptide **9** to ε-*N*-thiophosphoramidate **10** seen in Figure 4.5 was using 200.0 eq TEAB and 7.5 eq $KPSOCl_2$ (20%). Below 200.0 eq TEAB solution, there is a positive trend seen – each increase in the number of equivalents of TEAB results in an improved conversion from tetrapeptide **9** to ε-*N*-thiophosphoramidate **10**. The low conversions from **9** to **10** seen below 200.0 eq TEAB are probably caused by the reaction pH being below the pK_{aH} of the lysine ε-NH_2 moiety (10.53)[154], in analogy to what was seen when insufficient $NaOH_{(aq)}$ was used during the optimisation of aqueous *N*-thiophosphorylation on α-*N*-dansyl-L-lysine model **4**, which was discussed in section 2.2. Above 200.0 eq TEAB solution, there is a negative trend seen – each increase in the number of equivalents of TEAB leads to a reduced conversion from **9** to **10**. While there was no clear explanation for the observed trend above 200.0 eq TEAB solution, it may be due to increased competition between hydrolysis and aminolysis of $KPSOCl_2$.

When compared with the conversion from tetrapeptide **9** to ε-*N*-thiophosphoramidate **10** seen with $NaOH_{(aq)}$ in section 4.1.2, the greatest conversion to **10** with TEAB solution reported in Figure 4.5 is extremely poor (95% **9** to **10** with 43.7 eq 2 M $NaOH_{(aq)}$ compared to 20% **9** to **10** with 200.0 eq 1 M TEAB respectively). Both *N*-thiophosphorylations involved the use of 7.5 eq 330 mM $KPSOCl_2$ in MeCN (from the same stock solution). The same 4 mM tetrapeptide **9** stock solution was used for both the TEAB and $NaOH_{(aq)}$ experiments. The significant reductions in conversions to thiophosphoramidate can therefore be attributed to the nature of the species within TEAB and/or the pH used during the reactions.

For the *N*-thiophosphorylation carried out on tetrapeptide **9** using 7.5 eq $KPSOCl_2$ and 43.7 eq $NaOH_{(aq)}$ the pH was ca. 11 after the reaction had ended (tested with pH indicator paper). For the *N*-

thiophosphorylation carried out with tetrapeptide **9** using 7.5 eq $KPSOCl_2$ and 200.0 eq TEAB, however, the pH of the TEAB solution appeared to be lower at ca. 10.5 at room temperature. As pH 10.5 is close to the pK_{aH} of the lysine ε-NH_2 (10.53)[154], it is likely that approximately half of the ε-amine group of α-*N*-dansyl-labelled tetrapeptide **9** was in its protonated (i.e. ε-NH_3^+) form during the reaction, and was unable to undergo thiophosphorylation.

Based on this, an experiment was carried out with a 1 M TEAB solution made up to ca. pH 10.8 rather than 10.5 in order to see whether this minor pH change may result in a greater proportion of α-*N*-dansyl-labelled tetrapeptide **9** undergoing *N*-thiophosphorylation. The procedure used for this experiment was the same as that illustrated in Scheme 4.4.

The total absorbance chromatograph obtained through LC-MS DAD analysis following aqueous *N*-thiophosphorylation of tetrapeptide **9** with 7.5 eq $KPSOCl_2$ and 200.0 eq 1 M TEAB solution (pH 10.8) is shown in Figure 4.6. This chromatograph is compared against that obtained from the *N*-thiophosphorylation of tetrapeptide **9** using 7.5 eq $KPSOCl_2$ and 200.0 eq 1 M TEAB solution (pH 10.5) to illustrate the difference in conversion made through raising the base pH by 0.3.

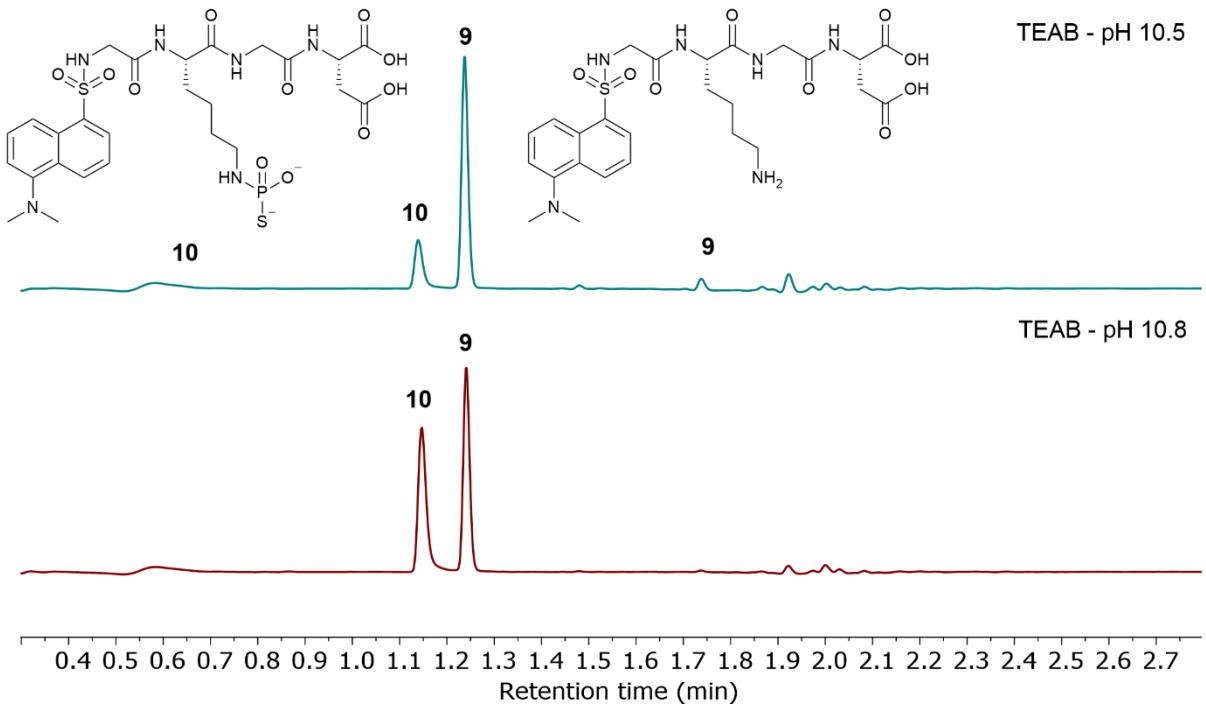
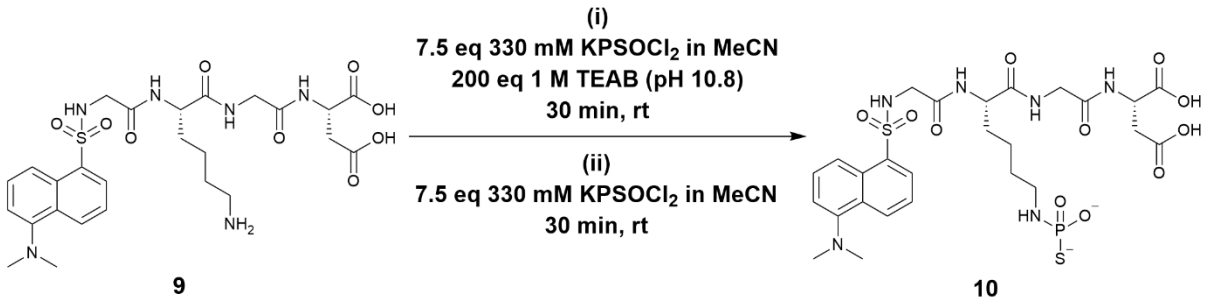


Figure 4.6 LC-MS total absorbance chromatographs illustrating the improved conversion from α-*N*-dansyl-labelled tetrapeptide **9** (*m/z* 609 Da) to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** (*m/z* 705 Da) seen when using TEAB (pH 10.8) (bottom chromatograph) compared to TEAB (pH 10.5) (top chromatograph). The concentrations of each reagent in the total reaction volume were: α-*N*-dansyl-labelled tetrapeptide **9** = 2.1 mM; $KPSOCl_2$ = 15.9 mM; TEAB solution = 423.1 mM.

Based on these two absorbance chromatographs, increasing the pH of TEAB solution resulted in a significant improvement in the proportion of tetrapeptide **9** undergoing *N*-thiophosphorylation, with conversion going up from 20% (1 M TEAB [pH 10.5]) to 46% (1 M TEAB [pH 10.8]). This pH increase of 0.3 likely resulted in a greater proportion of α-*N*-dansyl-labelled tetrapeptide **9** in the unprotonated form, thus allowing a greater conversion to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** over the same reaction time (30 min). Although increasing the pH of 1 M TEAB solution resulted in a 2.3-fold increase in the ε-*N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9**, this conversion was still < 50%.

4.2.1.2. Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** with multiple additions of $KPSOCl_2$

In order to try to further increase levels of *N*-thiophosphorylation, an experiment with two sequential additions of $KPSOCl_2$ (with an equal number of equivalents in each [7.5 eq]) was conducted. The first part of the $KPSOCl_2$ double addition experiment followed the procedure described in section 4.2.1. After 30 min of stirring, a second 7.5 eq dosage of $KPSOCl_2$ was added to the thiophosphorylating mixture, as is outlined in Scheme 4.5 (a more detailed procedure can be found in section 7.3.3.3).



Scheme 4.5 Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** to generate α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** using multiple additions of 7.5 eq $KPSOCl_2$.

A 5.0 µL aliquot was submitted for LC-MS analysis 30 min after each of the additions of $KPSOCl_2$, and the apparent conversions from α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** were calculated from the areas under the corresponding peaks on the total absorbance chromatographs (Table 4.3).

Table 4.3 Data from aqueous *N*-thiophosphorylation reaction using multiple additions of 7.5 eq $KPSOCl_2$. Conversions calculated from areas under relevant peaks seen in LC-MS total absorbance chromatographs.

Dose number	$KPSOCl_2$/ eq	Conversion from 9 to 10/ %
1	7.5	48
2	7.5	74

As expected, the final conversion following two additions of $KPSOCl_2$ from tetrapeptide **9** to thiophosphoramidate **10** was significantly improved compared to the single addition (48% increased

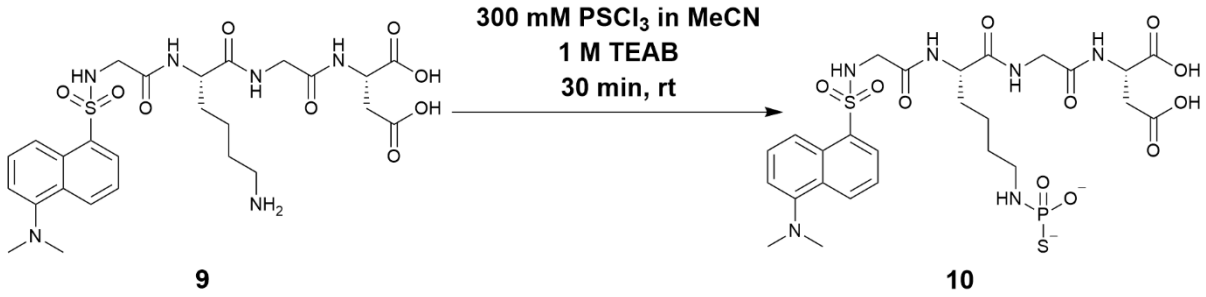
to 74%). That being said, this final conversion (74%) was still not on par with that seen following *N*-thiophosphorylation of tetrapeptide **9** with 7.5 eq $KPSOCl_2$ and 43.7 eq $NaOH_{(aq)}$ reported in section 4.1.2 (95%).

Based on all of the results obtained from aqueous *N*-thiophosphorylations carried out using $KPSOCl_2$ and TEAB solution, it was decided that these reaction conditions would not be any further. The large excess of $KPSOCl_2$ necessary to obtain conversions to ε-*N*-thiophosphoramidate greater than 70% could lead to significant formation of several undesired thiophosphorylated by-products should this methodology be employed on an unprotected peptide.

While aqueous *N*-thiophosphorylation using $KPSOCl_2$ and TEAB solution had proven less effective for formation of α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10**, the use of $PSCl_3$ and TEAB for the formation of ε-*N*-thiophosphoramidate **10** was also explored.

4.2.2. Using $PSCl_3$ and TEAB for aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9**

A 30 min aqueous *N*-thiophosphorylation reaction was performed using 7.5 eq $PSCl_3$ and 200.0 eq TEAB solution (pH 10.8) at room temperature (Scheme 4.6).



9 **10**

Scheme 4.6 Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** to generate α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** using $PSCl_3$ as the thiophosphorylating agent and TEAB as the base.

These conditions were employed in order to mirror those used in the 'optimised' single addition $KPSOCl_2$ and TEAB-based *N*-thiophosphorylation discussed in section 4.2.1.1, which resulted in 46% conversion from α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10**.

The use of 7.5 eq $PSCl_3$ and 200.0 eq TEAB solution (pH 10.8) for *N*-thiophosphorylation resulted in an observed conversion to ε-*N*-thiophosphoramidate **10** (based on LC-UV-vis analysis) of 70% (Figure 4.7). This conversion of 70% aligns with a trend noted by Conway *et al*. in 2014 and in section 4.1.2 of this chapter, which is that the use of $PSCl_3$ for aqueous *N*-thiophosphorylation typically results in greater

conversions to thiophosphoramidate compared to the same number of equivalents of $KPSOCl_2$.[147] This difference in thiophosphorylating efficiency has been largely attributed to $PSCl_3$ being relatively insoluble in water, while $KPSOCl_2$ is soluble. Greater reaction selectivity may be seen with $PSCl_3$ (between hydrolysis and aminolysis) as reactions occur solely at MeCN-water interfaces, while reactions with $KPSOCl_2$ are homogeneous resulting in hydrolysis of $KPSOCl_2$ competing with *N*-thiophosphorylation to a greater extent.[147]

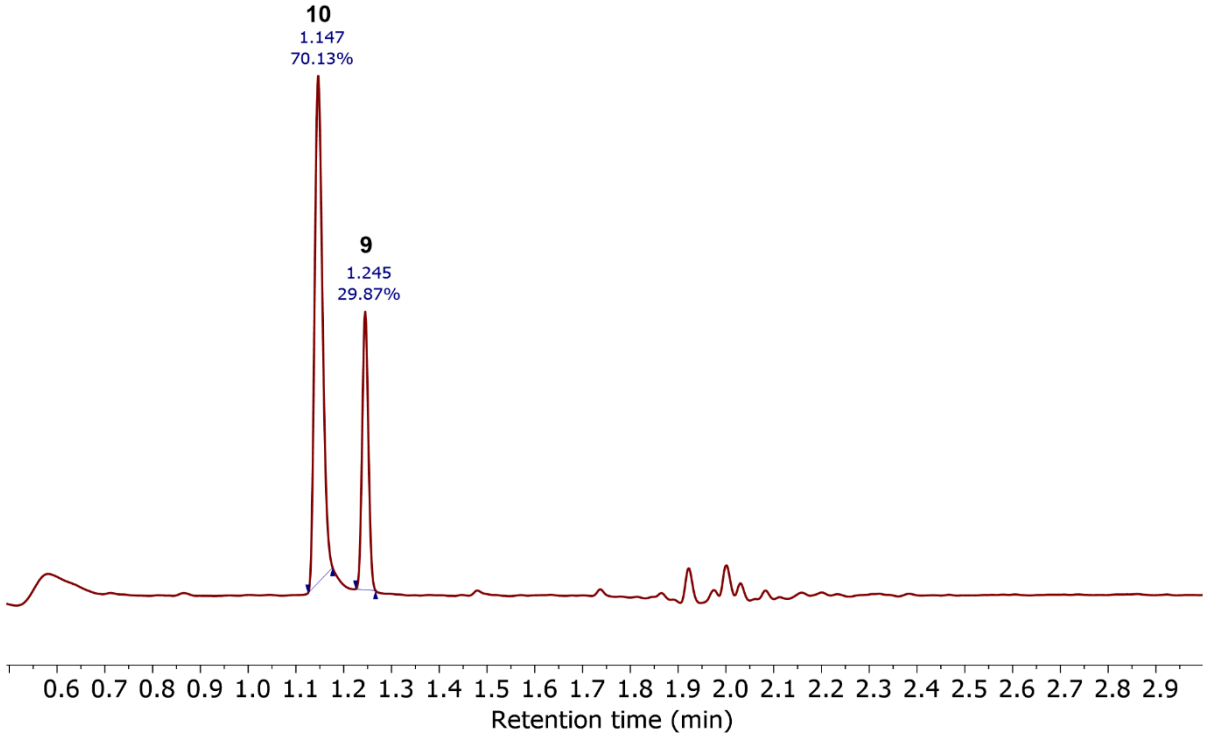
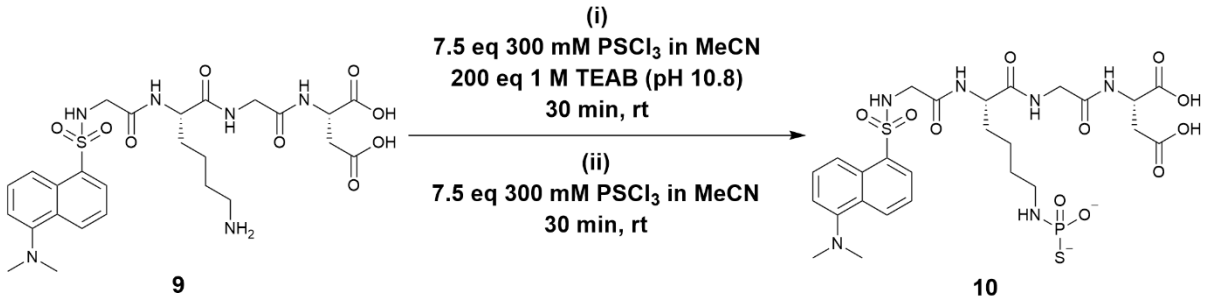


Figure 4.7 LC-MS total absorbance chromatograph following aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** to give α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** using 7.5 eq $PSCl_3$ and 200.0 eq TEAB (pH 10.8). The concentrations of each reagent in the total reaction volume were: α-*N*-dansyl-labelled tetrapeptide **9** = 2.1 mM; $PSCl_3$ = 15.8 mM; TEAB solution = 421.0 mM.

Despite this clear improvement in conversion to thiophosphoramidate **10** using 7.5 eq $PSCl_3$ (70%) rather than 7.5 eq $KPSOCl_2$ (46%) with 200.0 eq 1 M TEAB solution (pH 10.8), it did not surpass the conversions seen with $PSCl_3$ and $NaOH_{(aq)}$ in Chapter 2. Therefore, as with $KPSOCl_2$ and TEAB-based *N*-thiophosphorylations in section 4.2.1.2, an experiment involving two sequential additions of 7.5 eq $PSCl_3$ was carried out to determine whether conversions of > 90% are possible with $PSCl_3$ and TEAB.

4.2.2.1. Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** with multiple additions of $PSCl_3$

The experimental procedure used was the same as that used in section 4.2.1.2 except for the substitution of 330 mM $KPSOCl_2$ with 300 mM $PSCl_3$ (Scheme 4.7).



Scheme 4.7 Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** to generate α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** using multiple additions of 7.5 eq $PSCl_3$.

Samples were taken 30 min after the first and second addition of $PSCl_3$ and submitted for LC-MS DAD analysis. The conversions from α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** were calculated from the areas under the corresponding peaks on the total absorbance chromatographs (Table 4.4).

Table 4.4 Data from aqueous *N*-thiophosphorylation reaction with multiple additions of 7.5 eq $PSCl_3$. Conversions calculated from area under relevant peaks seen in LC-MS total absorbance chromatographs.

Dose number	$PSCl_3$/ eq	Conversion from 9 to 10/ %
1	7.5	59
2	7.5	88

As was seen in section 4.2.1.2, the final conversion following the second addition of $PSCl_3$ was significantly greater than the conversion following the first addition (88% compared to 59% respectively). The 88% conversion from tetrapeptide **9** to ε-*N*-thiophosphoramidate **10** is also much closer to the conversions seen with $PSCl_3$ and $NaOH_{(aq)}$ included in Table 4.1 in section 4.1.2.

The results from this aqueous *N*-thiophosphorylation using 200.0 eq 1 M TEAB (pH 10.8) and two additions of 7.5 eq $PSCl_3$ suggest that this methodology could be used for the ε-*N*-thiophosphorylation of peptide substrates that may be particularly susceptible to epimerisation at α-C centres, with up to 90% conversion to thiophosphoramidate possible.

For the remainder of the work presented within this thesis, however, all *N*-thiophosphorylations will be carried out using the $PSCl_3$ and $NaOH_{(aq)}$ conditions optimised on peptide substrates in Chapter 2 and used throughout Chapter 3. This decision was made largely due to the $NaOH_{(aq)}$-based thiophosphorylations only requiring a single addition of $PSCl_3$ in order to generate > 90% conversions to thiophosphoramidate. A single addition of thiophosphorylating agent, whether for a 30 min room temperature reaction or an 18 h overnight reaction (as shown in Chapter 3), keeps *N*-thiophosphorylation of a peptide substrate simple. In addition, the ability to use 'off the shelf' reagents without any further processing (e.g. $PSCl_3$ and 2 M volumetric standard NaOH solution) was a major driving force and advantage of this approach. Thus, it was deemed unnecessary to complicate

the simple aqueous thiophosphorylation method through the addition of extra steps to prepare thiophosphorylating agent ($KPSOCl_2$) and/or organic base solution (TEAB) during the remaining work.

4.3. Summary

Following the development of a robust and reliable aqueous *N*-thiophosphorylation method for the modification of peptide substrates using $PSCl_3$ and 2 M $NaOH_{(aq)}$ in Chapter 2, and an exploration of the selectivity of aqueous *N*-thiophosphorylation on unprotected peptide models in Chapter 3, we chose to investigate whether an alternative thiophosphorylating agent and a milder base could generate similar levels of conversion to thiophosphoramidate. It was also hoped that through the control of pH, additional selectivity between competing *N*-nucleophiles could be achieved.

Potassium thiophosphorodichloridate ($KPSOCl_2$) was chosen as the alternative thiophosphorylating agent based on work previously published by the Hodgson group.[147, 153] Using 7.5 eq 330 mM $KPSOCl_2$ in MeCN solution and 43.7 eq 2 M NaOH solution, a conversion of 95% to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** was obtained in a 30 min room temperature thiophosphorylation.

Section 4.2 focussed on the use of basic conditions, milder than hydroxide, to minimise the risk of epimerisation at α-C centres and potentially improve selectivity among competing *N*-nucleophiles. 1 M TEAB solutions with pHs ca. 10.5 to 11, were initially tested for *N*-thiophosphorylations with $KPSOCl_2$ as the thiophosphorylating agent. Through a series of aqueous *N*-thiophosphorylations carried out varying the number of equivalents of 1 M TEAB solution, and with 7.5 eq $KPSOCl_2$, it was found that 200.0 eq 1 M TEAB (pH 10.5) resulted in the greatest conversion to ε-*N*-thiophosphoramidate **10** at 20%. In an effort to improve on this rather poor conversion, the pH of the 1 M TEAB solution was raised from 10.5 to 10.8, which (using 200.0 eq 1 M TEAB and 7.5 eq $KPSOCl_2$) resulted in an improved conversion to ε-*N*-thiophosphoramidate **10** of 46%. For a final experiment to determine whether aqueous *N*-thiophosphorylation using TEAB solution and $KPSOCl_2$ was able to form thiophosphoramidate in conversions of > 50%, an experiment involving two sequential additions of 7.5 eq $KPSOCl_2$ was carried out. The total conversion from α-*N*-dansyl-labelled tetrapeptide **9** to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10** following this experiment was 74%. Following this, research into the use of $KPSOCl_2$ for selective aqueous *N*-thiophosphorylation was paused because of the relative inefficiency and inconvenience of $KPSOCl_2$-based reactions compared with $PSCl_3$.

As a final attempt using 1 M TEAB solution (pH 10.8) for aqueous *N*-thiophosphorylation, a thiophosphorylation reaction was carried out on α-*N*-dansyl-labelled tetrapeptide **9** using 200.0 eq 1 M TEAB solution and 7.5 eq 300 mM $PSCl_3$ in MeCN solution. This experiment resulted in 70% conversion to α-*N*-dansyl-labelled ε-*N*-thiophosphoramidate **10**. A further *N*-thiophosphorylation

reaction was carried out involving two sequential additions of 7.5 eq $PSCl_3$ in order to investigate whether a double addition of $PSCl_3$ under TEAB conditions could lead to conversions similar to those obtained with $PSCl_3$ and 2 M $NaOH_{(aq)}$. The total conversion from tetrapeptide **9** to ε-*N*-thiophosphoramidate **10** following two additions of 7.5 eq $PSCl_3$ solution was 88%. While this was a promising level of conversion to thiophosphoramidate, we chose not to continue using TEAB as the base because the large excess of $PSCl_3$ required to obtain thiophosphoramidate conversions > 85% could result in the formation of greater levels of side-products and consequently increased problems with product isolation.

Since aqueous *N*-thiophosphorylation with $PSCl_3$ and $NaOH_{(aq)}$ generates consistently high conversions to thiophosphoramidate from our tetrapeptide substrate with a single addition of $PSCl_3$, we decided to continue with the methodology developed throughout Chapters 2 and 3 for the peptide stapling investigation which will be discussed in Chapter 5.

5. Synthesis and stapling of *bis*-lysine heptamer

Having concluded that $PSCl_3$ and $NaOH_{(aq)}$ were the most effective combination for aqueous *N*-thiophosphorylations, we began designing and synthesising a short, *bis*-lysine peptide (sections 5.1 and 5.2) for use during proof-of-concept peptide stapling experiments (sections 5.3 to 5.5).

5.1. Design of an α-*N*-dansyl-labelled heptamer

During the early stages of peptide design we reviewed the available literature on stapled peptides, looking specifically at stapled sequences based on naturally occurring proteins. As one of the aims of this work was to develop a selective peptide stapling method that can be used on peptides without the need for protecting groups, designing a peptide model based on a naturally occurring sequence was thought to be the most appropriate course of action.

Based on a literature search, it was decided that the peptide model used for this work would be based on a fragment of the human BH3-interacting domain death agonist protein (BID BH3), previous used by Walensky *et al.* in 2004 for the activation of apoptotic pathways in leukaemia cells. In their study, as outlined in Chapter 1, an all-hydrocarbon staple was employed over a known α-helical region to impart enhanced α-helicity when compared to the wild-type peptide, as well as improving biological activity both *in vitro* and *in vivo* (mouse studies).[94]

While the fragment used by Walensky *et al*. contained 23 amino acid residues (spanning residues 80 to 102 of human BID BH3),[94] it was decided that the peptide fragment would be shortened further for this project in order to simplify peptide synthesis and minimise the number of competing nucleophilic amino acid residues. The eventual peptide sequence chosen for the *bis*-lysine peptide model was based on the 92 to 98 residue region of human BID BH3. Figure 5.1 shows the structures of the fragment designed for use in this project, wild-type fragment, and the fragment used by Walensky *et al*.[94]

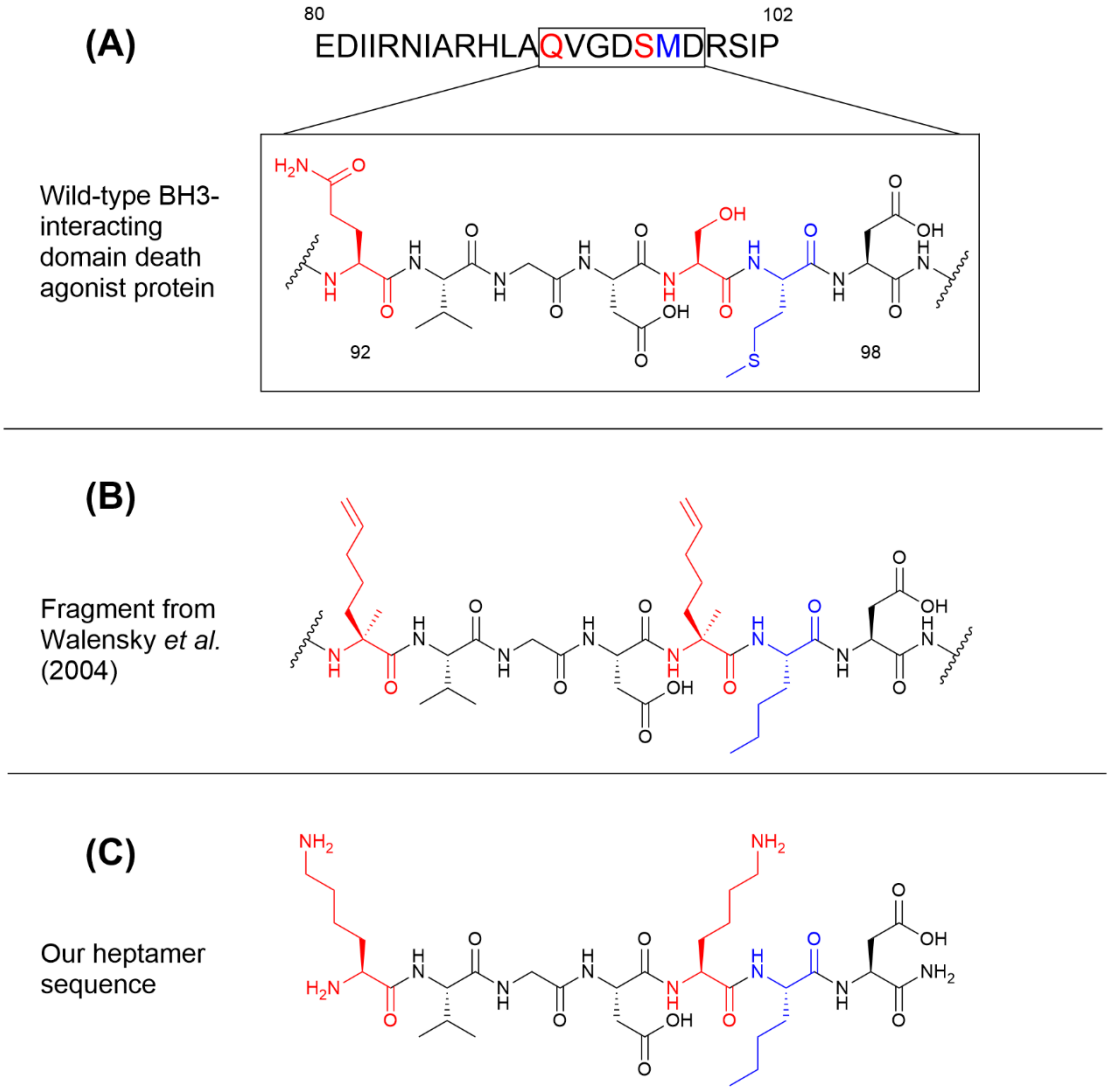


Figure 5.1 Peptide fragment structures for (A) wild-type BID BH3 protein (residues 92 to 98, with single letter code sequence showing residues 80 to 102); (B) seven-residue fragment from peptide used by Walensky *et al*. (2004)[94]; (C) seven-residue sequence used in our *bis*-lysine heptamer model. Red residues indicate the position of amino acid residues involved in stapling, and the blue residue indicates the change from Met to Nle.

As shown in Figure 5.1, the lysine residues for our heptamer model were positioned at the same points as the unnatural olefinic amino acids in the Walensky model (i.e. the *i, i* + 4 positions). The lysine residues were positioned at these sites because Walensky *et al*. showed that the amino acids in these positions could be substituted with amino acids containing peptide stapling precursor groups without inhibiting either the formation of secondary structural motifs (i.e. α-helices) or the biological activity of the stapled peptide.[94]

The methionine (Met [M]) residue of the wild-type fragment was replaced with the unnatural amino acid norleucine (Nle) to prevent any potential *S*-thiophosphorylation during thiophosphorylation reactions. Nle is isosteric with Met and has been shown to replace Met in bacterial protein

synthesis,[167, 168] and so the inclusion of Nle was believed to potentially have minimal impact on the activity or structure of the peptide.

With the final amino acid sequence of the heptamer model set, a dansyl group was added to the *N*-terminus to act as a chromophore. As with all dansyl-labelled models previously used throughout this project (most notably α-*N*-dansyl-labelled tetrapeptide **9**), the addition of the dansyl group facilitated the use of LC-MS DAD analysis. The dansyl label was positioned on the *N*-terminus to prevent any α-*N*-thiophosphoramidate side-products from forming during stapling attempts. Figure 5.2 shows the final structure of α-*N*-dansyl-labelled heptamer **27**.

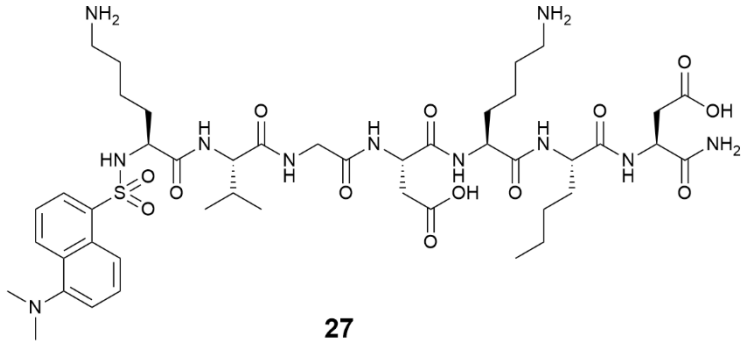


Figure 5.2 Structure of α-*N*-dansyl-labelled heptamer **27**.

5.2. Synthesis of α-*N*-dansyl-labelled heptamer 27

With a *bis*-lysine peptide model structure in mind, work began on the synthesis of α-*N*-dansyl-labelled heptamer **27**. Unlike all of the previously used tetra- and pentapeptide models, which were ordered from Thermo Fisher Scientific, heptamer **27** was synthesised in the Chemistry Department at Durham University. Restrictions on movement between laboratories implemented during the COVID-19 pandemic had been removed, thus allowing mg-scale microwave-assisted solid-phase peptide synthesis (SPPS) to be conducted on site.

5.2.1. Synthesis of α-*N*-dansyl-labelled heptamer **27**

α-*N*-dansyl-labelled heptamer **27** was synthesised by SPPS using a CEM Liberty Blue Automated Microwave Peptide Synthesiser, with the aid of Dr Carissa Lloyd.

Rink amide MBHA resin was used (not preloaded), and couplings were carried out with 0.2 M Fmoc-protected amino acid solutions in DMF (5 eq), 1 M DIC activator solution in DMF (10 eq) and 1 M oxyma solution in DMF (5 eq). Couplings were carried out *via* '0.10-Single Coupling (HS)' cycles. Deprotection was carried out after each coupling with 20% piperidine in DMF solution followed by DMF washes. Heptamer **27** was cleaved from the resin following a final Fmoc deprotection using a TFA:TIPS:H_2O solution (95:2.5:2.5 v/v). A more detailed description of the SPPS method can be found

in section 7.4.1. Figure 5.3 shows an LC-UV-vis chromatograph of crude heptamer **27** following cleavage from the resin.

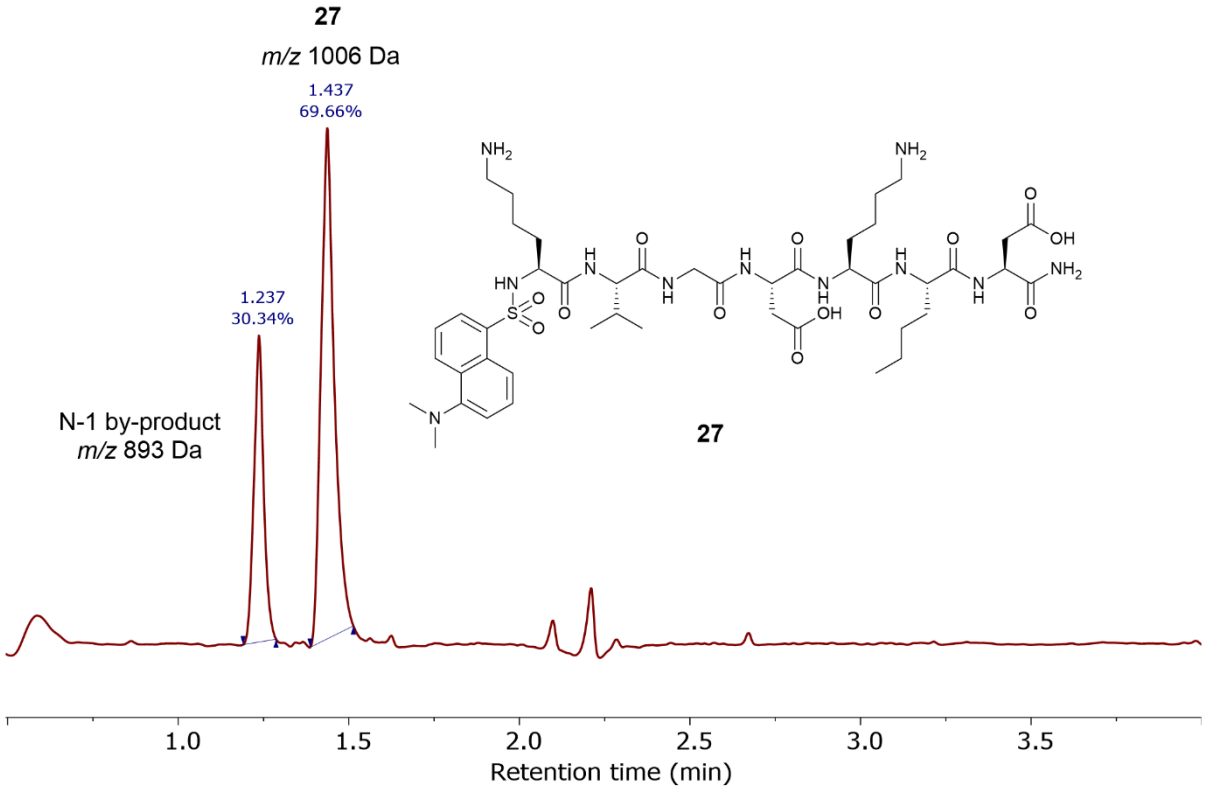


Figure 5.3 LC-MS total absorbance chromatograph of crude α-*N*-dansyl-labelled heptamer **27** (*m/z* 1006 Da) following cleavage from Rink amide MBHA resin.

Based on LC-UV-vis analysis, the crude SPPS product contained 70% α-*N*-dansyl-labelled heptamer **27** (*m/z* 1006 Da) and 30% N-1 by-product (*m/z* 893 Da). This N-1 by-product was determined to be a hexamer of the same sequence as heptamer **27** but missing the norleucine residue.

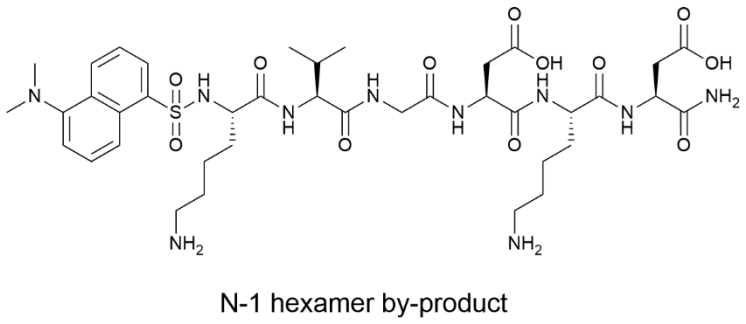


Figure 5.4 Structure of N-1 hexamer by-product.

As Figure 5.4 shows, the N-1 hexamer by-product was a *bis*-lysine peptide and so would not behave as a spectator during stapling reactions and would undergo *N*-thiophosphorylation alongside heptamer **27**. For this reason, while other batches of α-*N*-dansyl-labelled heptamer **27** were later

synthesised with greater homogeneity (ca. 75% heptamer **27**) and without an N-1 hexamer by-product, we decided to purify the crude 70% heptamer **27** sample.

5.2.2. Purification of α-*N*-dansyl-labelled heptamer **27**

Initial attempts at purification of α-*N*-dansyl-labelled heptamer **27** were carried out using the Teledyne CombiFlash NextGen 100 system available in the Hodgson laboratory at Durham University. However, when this chromatography system proved to be ineffective, heptamer **27** was purified *via* high-performance liquid chromatography (HPLC) using an Agilent 1260 Infinity II system available in the Cobb laboratory at Durham University.

5.2.2.1. Purification of α-*N*-dansyl-labelled heptamer **27** with Teledyne CombiFlash NextGen 100 system

Preliminary purification attempts on crude α-*N*-dansyl-labelled heptamer **27** were carried out using reverse phase flash chromatography. A Teledyne 5.5 g HP C_{18} Gold column was loaded with ca. 15 mg crude heptamer **27** dissolved in water/MeCN solution (50:50, v/v, 1 mL) and a solvent gradient of water (0.1% TFA) and MeCN (0.1% TFA) was then applied to the loaded column. Further details of the methodology used can be found in section 8.4.1. All eluted materials were collected, with UV-absorbing species being collected in 2 mL fractions, and non-absorbing species being collected in 10 mL fractions. Figure 5.5 shows the UV chromatograph obtained from this reverse phase chromatography run.

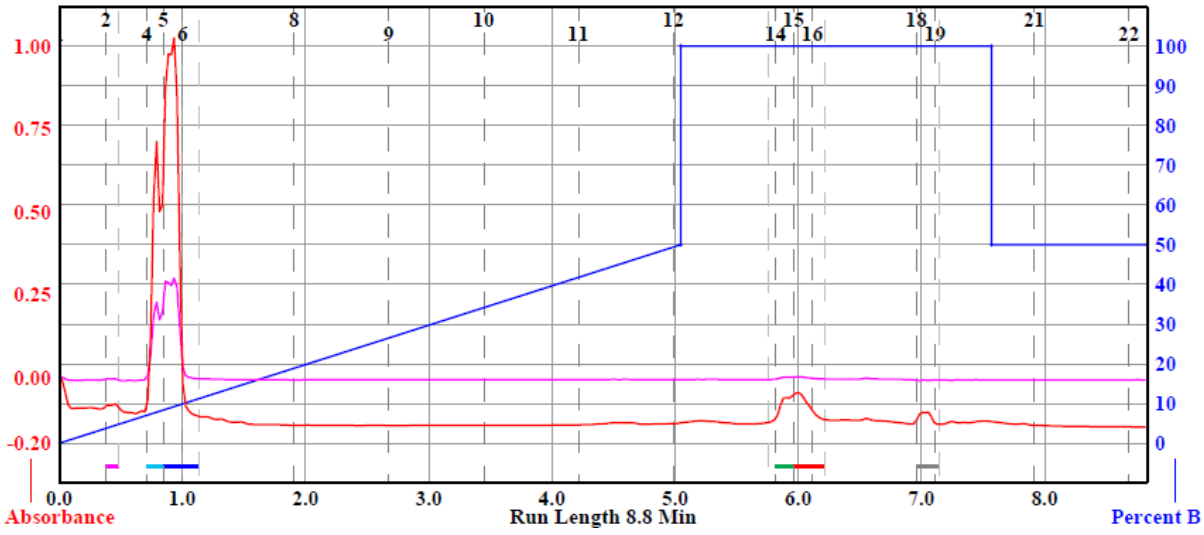


Figure 5.5 Chromatograph following preliminary reverse phase chromatography run to purify α-*N*-dansyl-labelled heptamer **27**.

From the red chromatograph displayed in Figure 5.5, two significant UV peaks can be seen. The first was at retention time ca. 0.7 – 1.1 min, and the eluted materials were collected in fractions 4 to 6. The

second significant peaks were seen around 5.8 – 6.2 min and all eluted materials were collected in fractions 14 to 16. LC-MS DAD analysis of fractions 4 to 6 and 14 to 16 indicated the presence of α-*N*-dansyl-labelled heptamer **27**, however, only fractions 14 to 16 were homogeneous. Fractions 4 to 6, while containing α-*N*-dansyl-labelled heptamer **27**, also contained the N-1 hexamer by-product.

Fractions 4 to 6 were, therefore, collected, the solvents were removed, and a second reverse phase chromatography run was carried out on the resulting material. The column and solvent system used remained the same for this second chromatography run, but the solvent gradient was made shallower in an attempt to improve separation of the desired heptamer and the N-1 hexamer by-product (gradient details can be found in section 7.4.1).

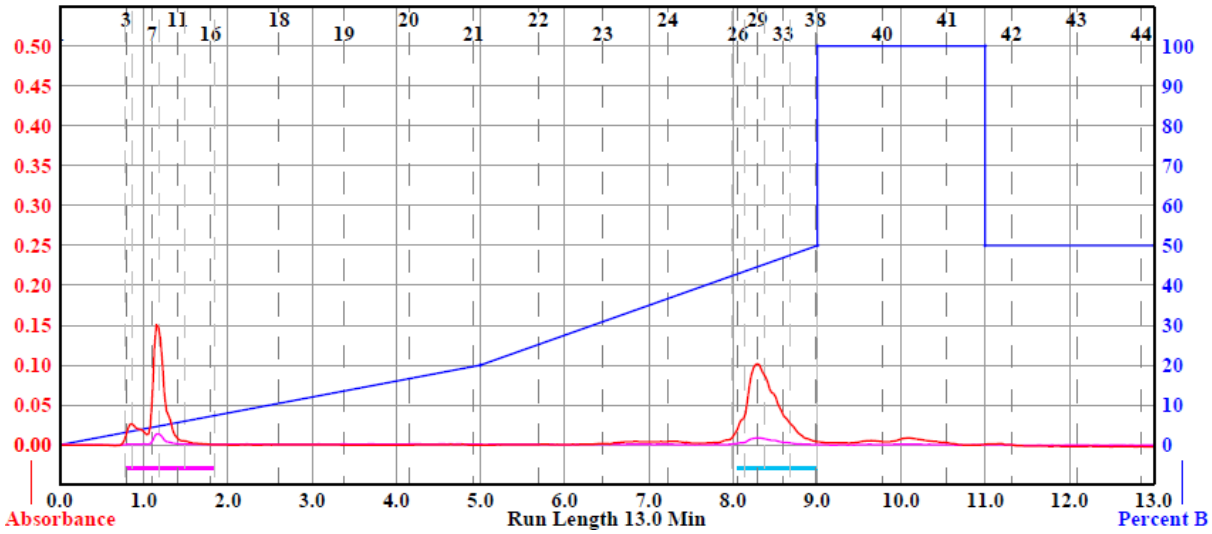


Figure 5.6 Chromatograph from second reverse phase chromatography run performed to purify α-*N*-dansyl-labelled heptamer **27**. Sample loaded on to C_{18} column was collected from fractions 4 – 6 of the first reverse phase chromatography run.

Figure 5.6 shows the UV chromatograph obtained from the second reverse phase chromatography run carried out on fractions 4 to 6 from the initial chromatographic separation. As with the first purification attempt, there are two significant peaks visible on the chromatograph. The first of these is from 0.7 – 1.8 min. From LC-MS DAD data, it was determined that this first peak contained the N-1 hexamer by-product. The second peak, seen between 8 – 9 min, was found to contain the α-*N*-dansyl-labelled heptamer **27** (data obtained from LC-MS DAD analysis).

Following these two purifications, fractions 14 to 16 from the first chromatography run and fractions 26 to 36 from the second chromatography run (i.e. the fractions found to be > 97 % **27**) were collected together. The solvents were removed, the solid residue was dissolved in water, and the resultant solution was analysed by LC-MS DAD, and the total absorbance chromatograph obtained is displayed in Figure 5.7.

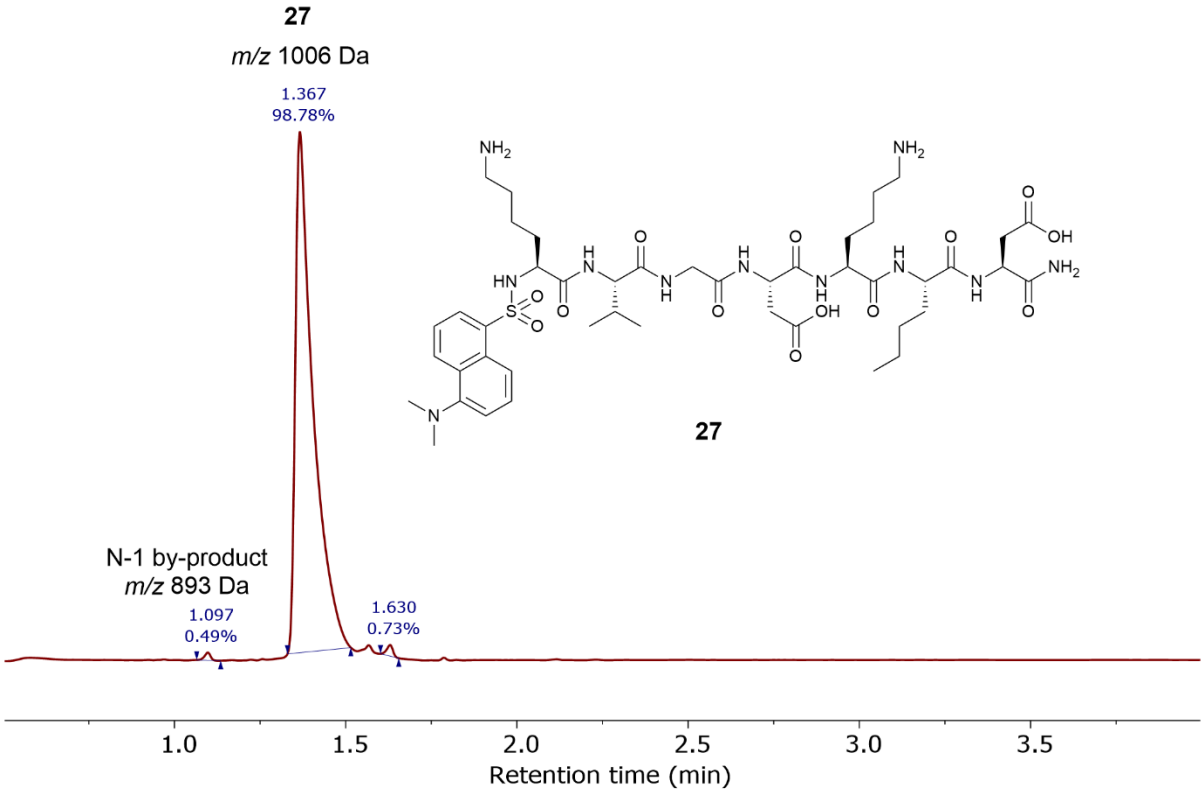


Figure 5.7 LC-MS total absorbance chromatograph of α-*N*-dansyl-labelled heptamer **27** (*m/z* 1006 Da) following purification *via* two reverse phase chromatograph columns using a Teledyne CombiFlash NextGen 100 system.

Based on LC-UV-vis analysis, the two reverse phase flash chromatography purifications resulted in the isolation of α-*N*-dansyl-labelled heptamer **27** to ca. 98.8% homogeneity. Trace impurities totalling ca. 1.2%, including the N-1 hexamer by-product (*m/z* 893 Da), remained.

While we were able to purify crude α-*N*-dansyl-labelled heptamer **27** to ca. 98.8% homogeneity *via* two reverse phase flash chromatography runs using the Teledyne system available in the Hodgson lab, it ultimately proved to be extremely time consuming. A second batch of α-*N*-dansyl-labelled heptamer **27** was synthesised by automated SPPS concurrently with the purification discussed above. The second synthesis of heptamer **27** resulted in a product of greater purity in its crude form (74% heptamer **27**) with no N-1 hexamer by-product being present.

As a supply of 'purer' α-*N*-dansyl-labelled heptamer **27** had been obtained following the second SPPS run, we chose to go ahead with exploratory stapling experiments using the crude peptide from the second heptamer synthesis (74% heptamer **27**). This decision was made largely in the interest of time remaining for this project. The preliminary stapling experiments carried out of the second batch of α-*N*-dansyl-labelled heptamer **27** will be discussed in section 5.3.

Before detailing these preliminary experiments, however, the purification of the hexamer-contaminated batch of heptamer **27**, using an HPLC method, will be described.

5.2.2.2. Purification of α-*N*-dansyl-labelled heptamer **27** with Agilent 120 Infinity II HPLC system

For this new approach to chromatographic purification, a preparative HPLC system was employed rather than the Teledyne flash system used previously. The HPLC system was employed in an effort to minimise the number of chromatography runs necessary to obtain heptamer **27** at high levels of purity.

The HPLC system was used, with the aid of Miss Katherine Deck, to successfully isolate α-*N*-dansyl-labelled heptamer **27** following a single 30 min run. The column used was an Agilent 5 Prep C_{18} column while the mobile phase was the same as that used for the Teledyne purifications discussed in section 5.2.2.1 (0.1% TFA in water solution for the aqueous solvent and 0.1% TFA in MeCN solution for the organic solvent). Mobile phase gradient details can be found in section 7.4.1. Figure 5.8 shows the chromatograph obtained following HPLC purification of crude α-*N*-dansyl-labelled heptamer **27**.

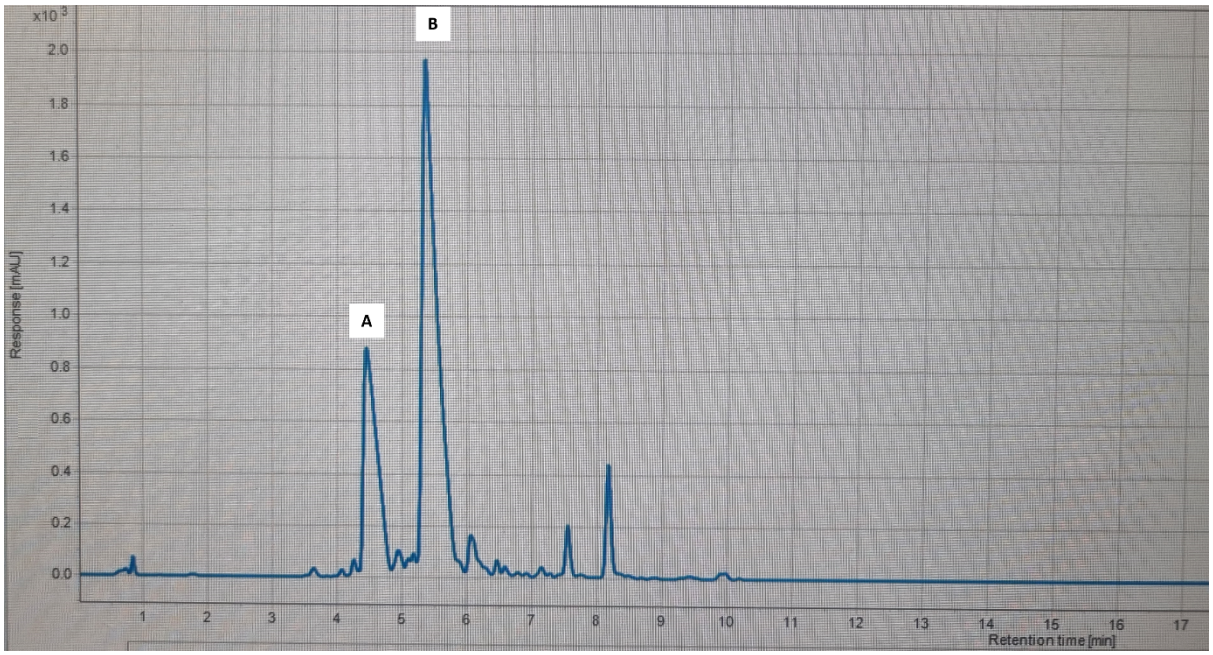


Figure 5.8 220 nm chromatograph following HPLC purification of crude α-*N*-dansyl-labelled heptamer **27**.

Following LC-MS analysis of peaks A and B, it was found that peak A contained the N-1 hexamer by-product while peak B contained isolated α-*N*-dansyl-labelled heptamer **27**. Figure 5.9 shows the total absorbance chromatograph obtained from LC-MS analysis of peak B following purification of hexamer contaminated α-*N*-dansyl-labelled heptamer **27**. Based on LC-UV-vis analysis, α-*N*-dansyl-labelled heptamer **27** was isolated to > 99% homogeneity.

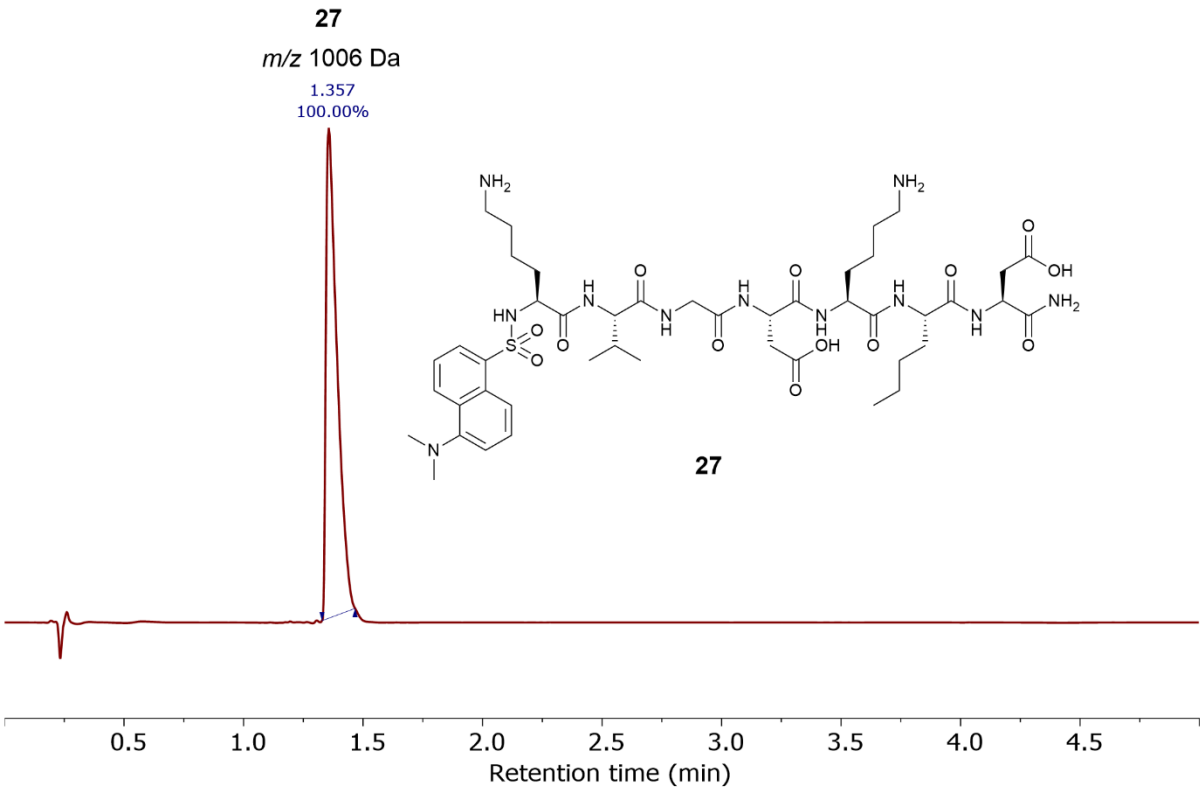
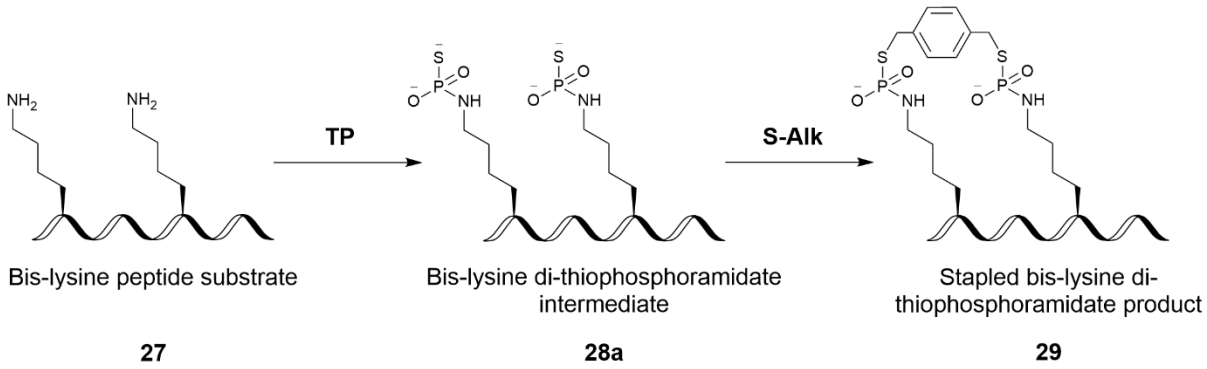


Figure 5.9 LC-MS total absorbance chromatograph of α-*N*-dansyl-labelled heptamer **27** (*m/z* 1006 Da) following purification *via* reverse phase HPLC using an Agilent 1260 Infinity II system.

Following isolation, heptamer **27** was employed in scaled up stapling experiments in an effort to observe quantitative conversion levels for both steps of the stapling process and to isolate sufficient quantities of stapled heptamer for characterisation by NMR methods in addition to LC-MS (sections 5.4 and 5.5). As mentioned, however, in section 5.2.2.1, the results of the preliminary stapling experiments carried out on crude samples of heptamer **27** must first be discussed.

5.3. Preliminary stapling experiments on crude α-*N*-dansyl-labelled heptamer 27

The method used for the stapling reactions carried out on samples of crude heptamer **27** was based on the one-pot, two-step *N*-thiophosphorylation and *S*-alkylation experiments described in section 2.2.1 (with α-*N*-dansyl-L-lysine **4** as the substrate) and section 3.4.3 (with *N*-terminal phenylalanine tetrapeptide **23a** as the substrate). In short, α-*N*-dansyl-labelled heptamer **27** was thiophosphorylated under aqueous conditions before stapling (*S*-alkylating) the resulting di-thiophosphoramidate using a di-alkyl halide linker. A schematic overview of *N*-thiophosphorylation and stapling of heptamer **27** to generate α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**, is illustrated in Scheme 5.1.



Scheme 5.1 Sketched representation of stapling on a *bis*-lysine peptide substrate (i.e. α-*N*-dansyl-labelled heptamer **27**) to give a stapled *bis*-lysine di-thiophosphoramidate product (i.e. α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**) *via* a *bis*-lysine di-thiophosphoramidate intermediate (i.e. α-*N*-dansyl-labelled di-thiophosphoramidate **28a**). TP = *N*-thiophosphorylation; S-Alk = *S*-alkylation.

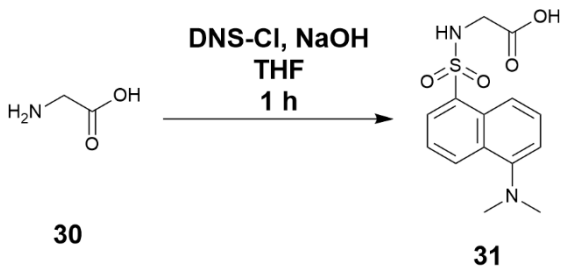
Through the use of crude α-*N*-dansyl-labelled heptamer **27** robust workflows and analysis methods were established and were applied to our complex stapling mixtures before implementing these procedures on homogeneous samples of heptamer **27**.

While there were many advantages to using crude samples of heptamer **27** for initial investigations into peptide stapling *via* a di-thiophosphoramidate intermediate, the use of impure starting material meant that the apparent concentration of aqueous heptamer **27** solution (prepared based on the mass of crude peptide) was not a realistic representation of the true concentration of heptamer **27** present in the peptide solution prior to thiophosphorylation. In order to obtain a more accurate picture of the concentration of α-*N*-dansyl-labelled heptamer **27** present in each substrate solution, a calibration plot was created based on varying concentrations of a dansyl-labelled amino acid standard.

5.3.1. Calibration plot for determining concentration of α-*N*-dansyl-labelled heptamer **27** in substrate stock solutions

To create the calibration plot for determining the concentration of α-*N*-dansyl-labelled heptamer **27**, a dansyl-labelled amino acid model was required. Dansyl glycine **31** was chosen for the standard curve model as, while it has been noted that the extinction coefficient of a protein-bound dansyl group can differ significantly based on the protein,[156] it was thought that the UV absorbance of a dansyl group bound to the α-NH_2 of a glycine residue would not differ significantly when compared to a dansyl group bound to the *N*-terminus of heptamer **27**.

Dansyl glycine **31** was prepared based on a procedure found in the literature,[169] and purified to ca. 98% homogeneity (by LC-UV-vis analysis).



Scheme 5.2 Dansylation of glycine **30** to give dansyl glycine **31**.

A 10 mM stock solution of dansyl glycine **31** was prepared, a series of dilutions was carried out and each diluted sample was analysed by LC-MS DAD. A plot of the UV absorbance at 325 nm as a function of dansyl glycine **31** concentration (Figure 5.10) was then constructed. The absorbance at 325 nm was selected for use as it represents λ_{max} for dansyl, and should not be impacted by any absorbance contributions for amide bonds (absorbance around 190 to 220 nm).[170, 171] The area under each LC-UV peak corresponding to dansyl glycine **31** (identity confirmed by MS) was plotted against the LC-MS sample concentration based on dilution of the 10 mM stock of dansyl glycine **31**.

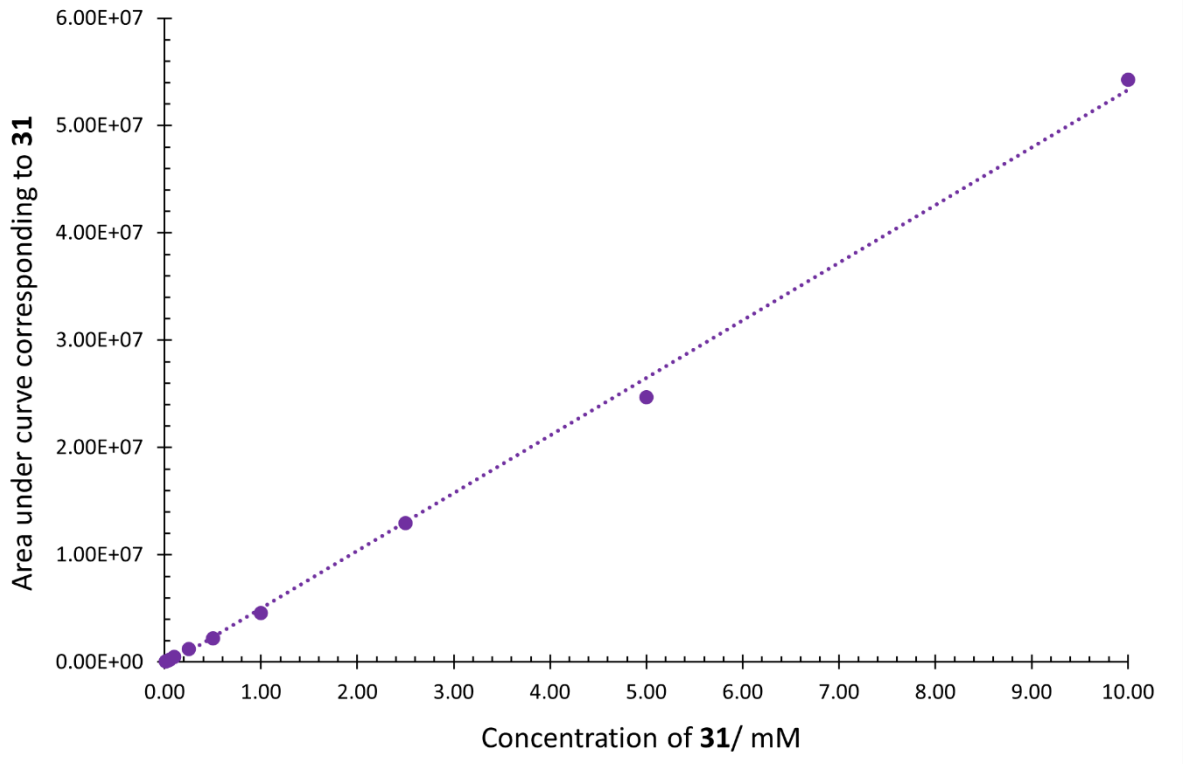


Figure 5.10 Plot of the area under each peak corresponding to dansyl glycine **31** taken from absorbance measurements at 325 nm of LC-MS chromatographs at various concentrations of dansyl glycine **31**.

From the calibration plot shown in Figure 5.10, the concentration of α-*N*-dansyl-labelled heptamer **27** can be estimated based on the area under the peak corresponding to heptamer **27** on a 325 nm

chromatograph of the crude substrate solution using the rearranged equation for determining an unknown y value from a linear plot:

Basic equation for finding a y value from a linear plot: $y = mx + c$

Where *y* is a value on the y axis, *m* is the gradient of the slope, *x* is a value on the x axis, and *c* is the point at which the slope intercepts the x axis.

Rearranged to find an unknown value on the x axis: $x = \frac{y-c}{m}$

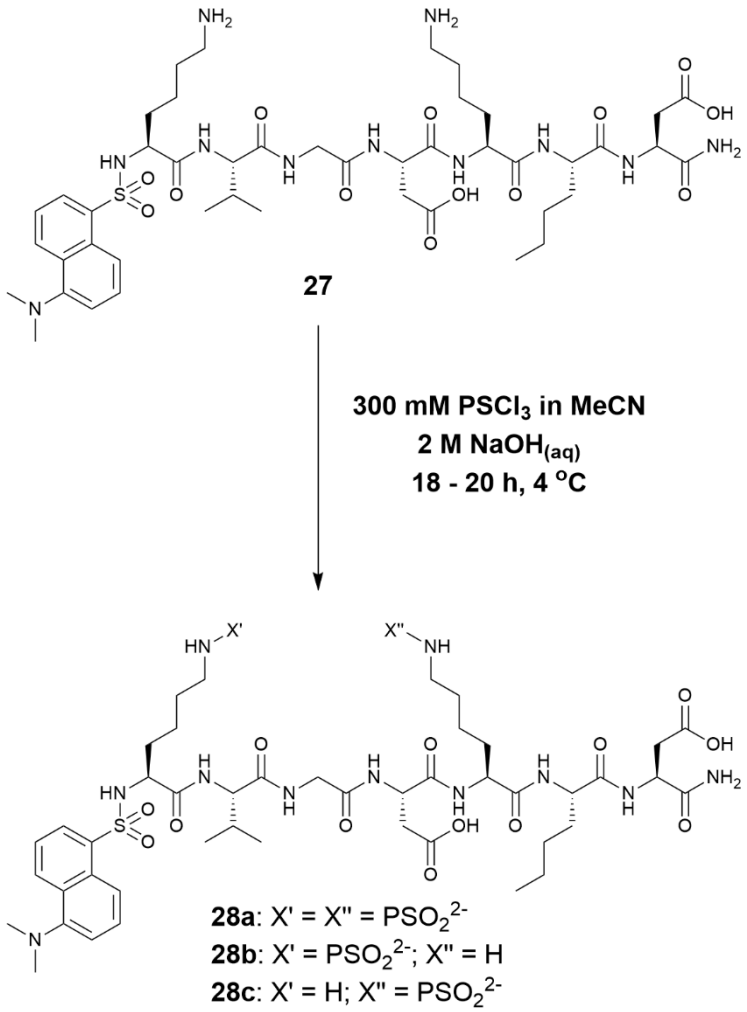
The constants *c* and *m* can be substituted for - 341903 and 5 × 10^6 respectively to give the final equation:

$$\frac{Area\ under\ the\ peak\ corresponding\ to\ \mathbf{27} + 341903}{5 \times 10^6} = Concentration\ of\ \mathbf{27}\ (mM)$$

With a calibration plot established to estimate concentrations of α-*N*-dansyl-labelled heptamer **27** in crude substrate samples, exploratory stapling experiments could begin.

5.3.2. Aqueous *N*-thiophosphorylation and *S*-alkylation of α-*N*-dansyl-labelled heptamer **27**

The proposed stapling method was based on the one-pot, two-step *N*-thiophosphorylation and *S*-alkylation methodology first reported by Trmčić and Hodgson,[152] so the initial stapling investigation began with aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled heptamer **27**. As heptamer **27** contains two lysine residues requiring thiophosphorylation, 2.0 eq $PSCl_3$ and 11.0 eq $NaOH_{(aq)}$ were used during the initial *N*-thiophosphorylation conducted on α-*N*-dansyl-labelled heptamer **27**. The reaction scheme is illustrated in Scheme 5.3.



Scheme 5.3 *N*-thiophosphorylation of α-*N*-dansyl-labelled heptamer **27** to form α-*N*-dansyl-labelled di-thiophosphoramidate **28a**.

Scheme 5.3 highlights the three expected thiophosphoramidate products formed during aqueous *N*-thiophosphorylation of heptamer **27** – the desired α-*N*-dansyl-labelled di-thiophosphoramidate **28a** and the two α-*N*-dansyl-labelled mono-thiophosphoramidate isomers **28b** and **28c**.

Figure 5.11 shows the 325 nm chromatographs obtained from LC-MS DAD analysis of the crude heptamer **27** substrate prior to *N*-thiophosphorylation, and the product mixture following *N*-thiophosphorylation.

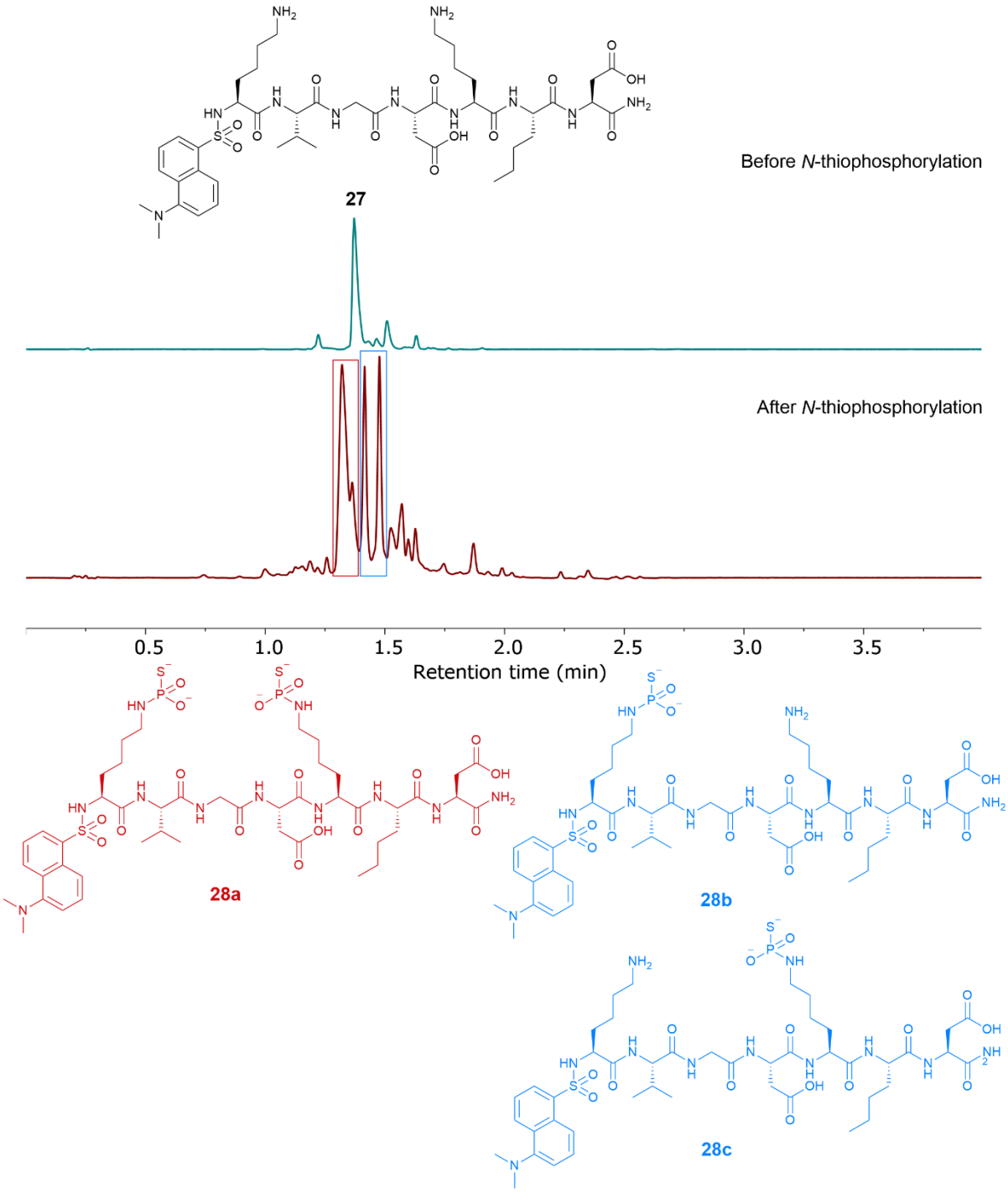


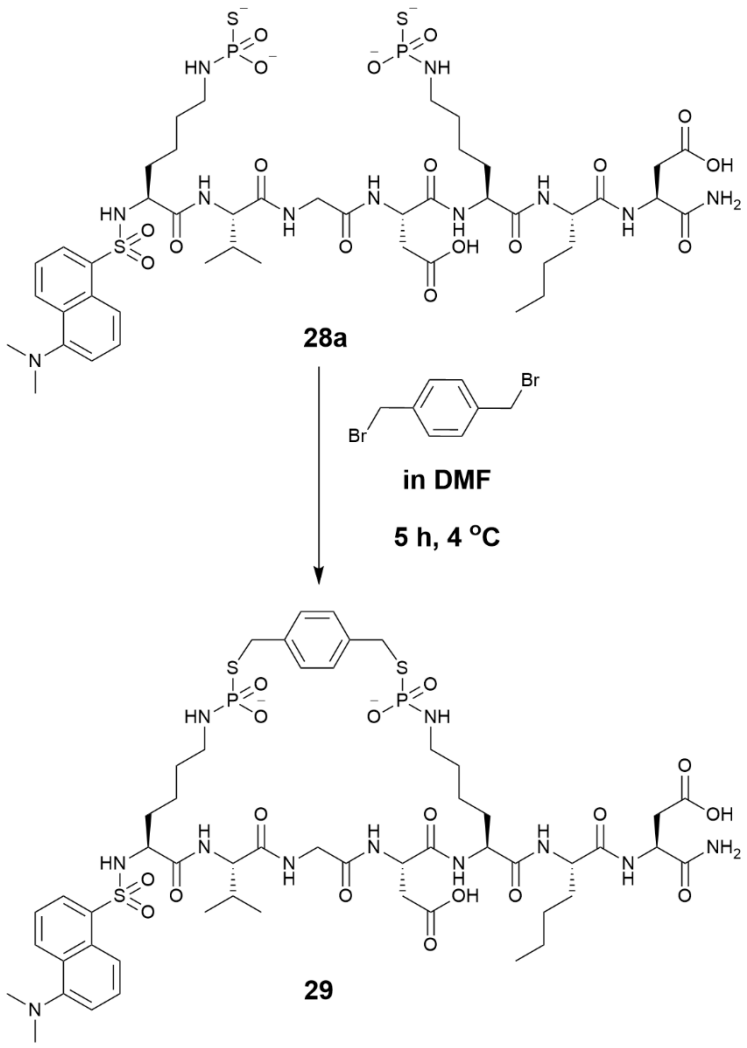
Figure 5.11 325 nm chromatographs obtained from LC-MS DAD analysis following aqueous *N*-thiophosphorylation of crude α-*N*-dansyl-labelled heptamer **27** (*m/z* 1006 Da) (top trace) to give α-*N*-dansyl-labelled di-thiophosphoramidate **28a** (*m/z* 1198 Da – indicated in red) (bottom trace). The structures of the two major thiophosphorylation side-products (α-*N*-dansyl-labelled mono-thiophosphoramidates **28b** and **28c** (*m/z* 1102 Da) are highlighted in blue. The concentrations of each reagent in the total reaction volume were: heptamer **27** = 21.4 mM; $PSCl_3$ = 42.9 mM; NaOH solution = 235.8 mM.

Following aqueous *N*-thiophosphorylation in the presence of 2.0 eq $PSCl_3$, LC-MS DAD analysis indicated that α-*N*-dansyl-labelled heptamer **27** was entirely consumed and formed a mixture of di-thiophosphoramidate **28a** and the two mono-thiophosphoramidate isomers **28b** and **28c**. By only

considering the peaks corresponding to the three thiophosphoramidate products, ca. 57% of heptamer **27** was converted to α-*N*-dansyl-labelled di-thiophosphoramidate **28a**, with the remaining 43% of heptamer **27** transformed to one of the two α-*N*-dansyl-labelled mono-thiophosphoramidates **28b** and **28c** (cannot determine which mono-thiophosphoramidate peak on the 325 nm chromatograph in Figure 5.11 corresponds to which thiophosphoramidate isomer by MS).

Following the *N*-thiophosphorylation step, 1.0 eq α,α'-dibromo-*p*-xylene (10.2 mM solution in DMF) was added to the thiophosphoramidate mixture and the stapling step was run for 5 h at 4 °C. A rigid, xylene-based linker was chosen for these stapling experiments over a flexible, alkyl linker as staple rigidity impacts the stabilisation of an α-helix.[138, 172] As lysine side chains are already long and flexible, a rigid linker (such as a xylene-based linker) will theoretically lead to greater stabilisation of an α-helix than an alkyl linker. That being said, due to the flexibility of the two lysine side chains, the α-helical secondary structure of α-*N*-dansyl-labelled heptamer **27** is not expected to be stabilised dramatically through stapling.

The reaction scheme for the *S*-alkylation-based stapling step is illustrated in Scheme 5.4. While various by-products were noted to form through *S*-alkylation of α-*N*-dansyl-labelled mono-thiophosphoramidates **28b** and **28c**, these are not illustrated here for the sake of simplicity.



Scheme 5.4 *S*-alkylation of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to give α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**.

Figure 5.12 shows the 325 nm chromatographs obtained from LC-MS DAD analysis of the crude thiophosphoramidate mixture before stapling (top trace) and the stapling mixture after 5 h reaction time at 4 °C (bottom trace).

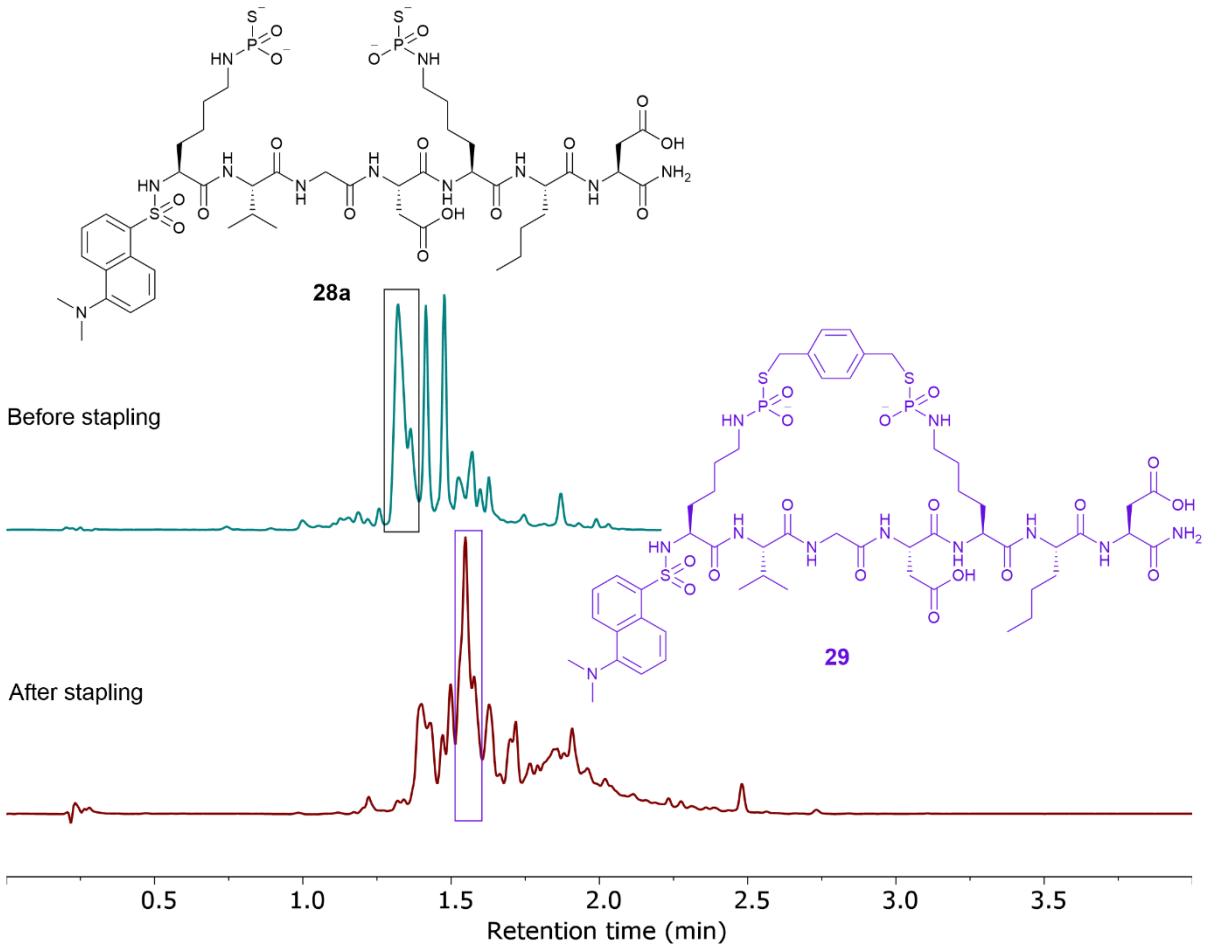
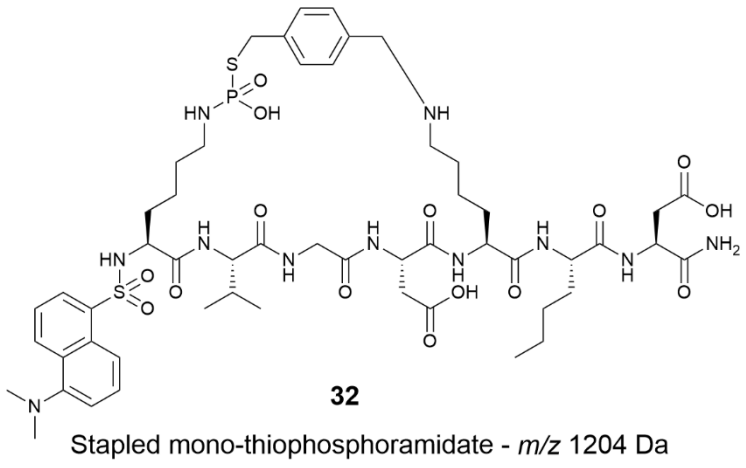


Figure 5.12 325 nm chromatographs obtained from LC-MS DAD analysis following *S*-alkylation of crude α-*N*-dansyl-labelled di-thiophosphoramidate **28a** (*m/z* 1198 Da) (top trace) to give α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (*m/z* 1300 Da) (bottom trace). The concentrations of each reagent in the total reaction volume were: α-*N*-dansyl-labelled di-thiophosphoramidate **28a** = 6.91 mM; α,α'-dibromo-*p*-xylene = 6.91 mM.

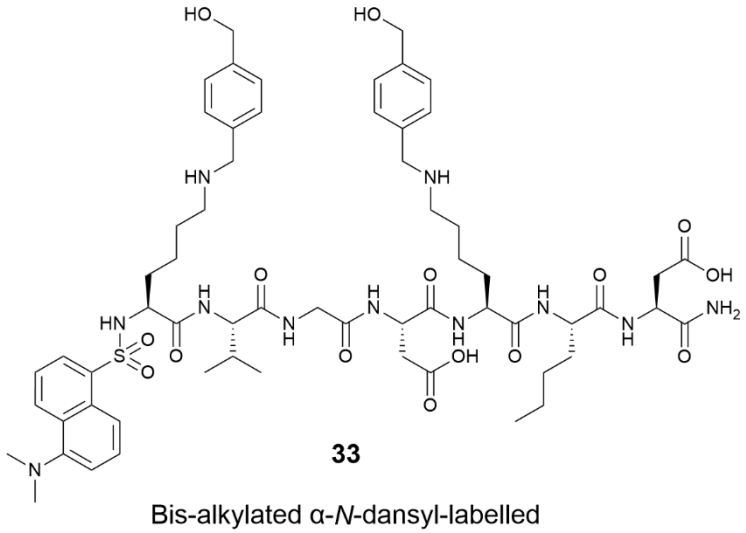
Following 5 h *S*-alkylation of the crude *N*-thiophosphorylation mixture with 1.0 eq α,α'-dibromo-*p*-xylene, LC-MS DAD analysis gave a complex chromatograph at 325 nm. While approximate conversions cannot be calculated due to the lack of baseline separation between species eluted from 1.25 – 2.00 min, the peak at 1.548 min did show the formation of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (*m/z* 1300 Da).

The extracted ion count for α-*N*-dansyl-labelled di-thiophosphoramidate **28a** (*m/z* 1198 Da) indicated that it was not entirely converted to stapled di-thiophosphoramidate **29**. This can largely be explained through the generation of various side-products formed over the course of the *S*-alkylation step. Two of the most significant side-products had *m/z* values of 1204 Da (**32**) and 1246 Da (**33**). Proposed structures for these two side-products are shown in Figure 5.13.



32

Stapled mono-thiophosphoramidate - *m/z* 1204 Da



33

Bis-alkylated α-*N*-dansyl-labelled heptamer - *m/z* 1246

Figure 5.13 Proposed structures for the side-products with *m/z* 1204 Da (stapled mono-thiophosphoramidate **32**) and *m/z* 1246 Da (*bis*-alkylated α-*N*-dansyl-labelled heptamer **33**).

The formation of various *S*- and *N*-alkylated side-products (some of which have undergone *bis*-alkylation) alongside the synthesis of the desired stapled heptamer product therefore led to the overconsumption of α,α'-dibromo-*p*-xylene, preventing the complete conversion of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**.

It should be noted that the formation of *bis*-alkylated α-*N*-dansyl-labelled heptamer **33** suggests that aqueous *N*-thiophosphorylation had not led to the complete conversion of α-*N*-dansyl-labelled heptamer **27** to one of the three expected thiophosphoramidate derivatives as previously thought, based on the LC-MS data shown in Figure 5.11, or that loss of thiophosphoryl groups could have occurred, perhaps due to the use of insufficient NaOH during thiophosphorylation, resulting in hydrolysis of acid-sensitive thiophosphoramidate groups.

In order to ensure that *N*-thiophosphorylation was carried out to near completion prior to the addition of the bifunctional linker, an experimental series was conducted on heptamer **27** in which the number of equivalents of $PSCl_3$ was varied.

5.3.2.1. Varying the number of equivalents of $PSCl_3$ used for aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled heptamer **27**

Three aqueous *N*-thiophosphorylations using either 2.0, 2.5, or 3.0 eq 300 mM $PSCl_3$ in MeCN solution were carried out on crude α-*N*-dansyl-labelled heptamer **27**. In each experiment, the number of equivalents of 2 M $NaOH_{(aq)}$ was adjusted to accommodate the increasing number of equivalents of acid forming $PSCl_3$ (method details can be found in section 7.5.1.2.1). The reaction scheme is illustrated in Scheme 5.3 (on page 157).

The LC-MS DAD data collected following each of the three *N*-thiophosphorylation reactions can be seen in Figure 5.14 and Table 5.1. The chromatographs in Figure 5.14 were taken at 325 nm, with only the thiophosphoramidate derivatives of heptamer **27** (α-*N*-dansyl-labelled di-thiophosphoramidate **28a** and α-*N*-dansyl-labelled mono-thiophosphoramidates **28b** and **28c**) highlighted.

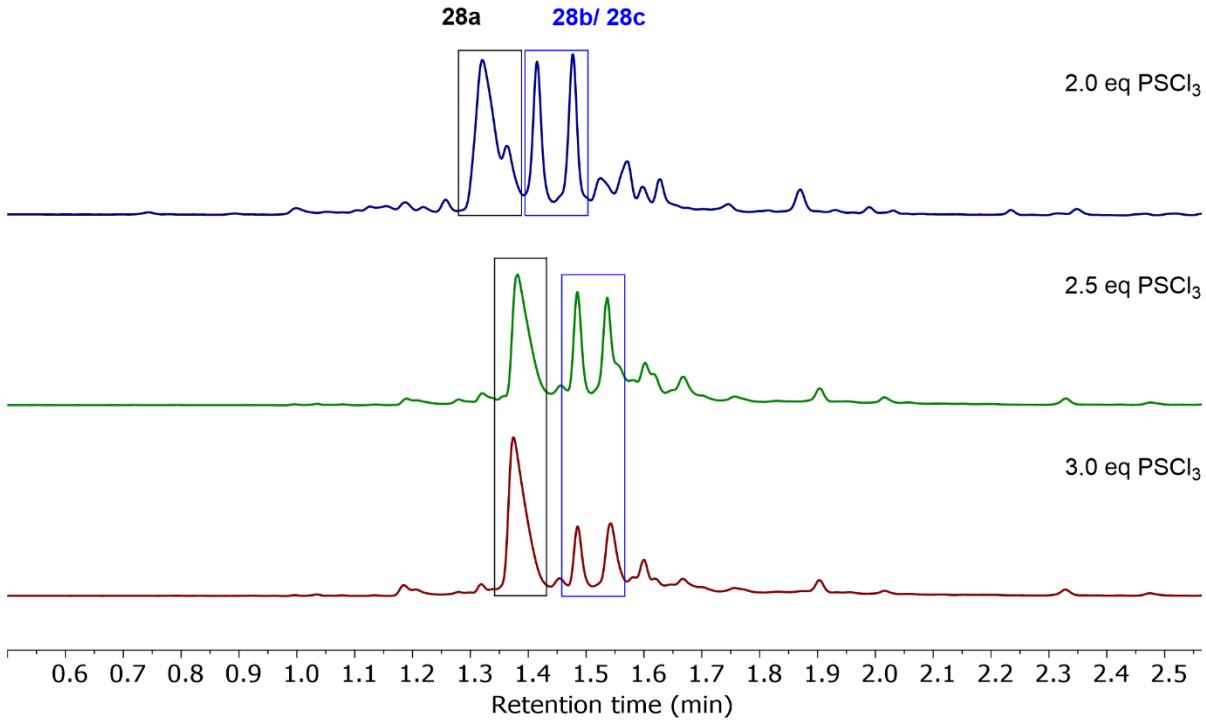


Figure 5.14 325 nm chromatographs obtained from LC-MS DAD analysis following an aqueous *N*-thiophosphorylation series performed on α-*N*-dansyl-labelled heptamer **27** in which the number of equivalents of $PSCl_3$ was varied. Peaks corresponding to α-*N*-dansyl-labelled di-thiophosphoramidate **28a** (*m/z* 1198 Da) are highlighted in black, and peaks corresponding to the two α-*N*-dansyl-labelled mono-thiophosphoramidates **28b** and **28c** (*m/z* 1102 Da) are highlighted in blue.

Table 5.1 shows the proportions of each of the thiophosphoramidate derivatives following the overnight *N*-thiophosphorylation reactions. The proportions of the two mono-thiophosphoramidate products (**28b** and **28c**) were grouped together as it was not possible to determine which mono-thiophosphoramidate peak in the LC-MS chromatographs corresponded to which mono-thiophosphoramidate isomer.

Table 5.1 Approximate proportions of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** and the two α-*N*-dansyl-labelled mono-thiophosphoramidates **28b** and **28c** formed during the aqueous *N*-thiophosphorylation series varying the number of equivalents of $PSCl_3$. Proportions were determined through consideration of the thiophosphoramidate peaks only (from 325 nm LC-MS DAD chromatographs shown in Figure 5.14).

$PSCl_3$/ eq	**Proportion of 28a/ %**	**Combined proportion of 28b and 28c/ %**
2.0	57	43
2.5	59	41
3.0	68	32

Based on the results obtained for LC-MS DAD analysis, it was found that using 3.0 eq 300 mM $PSCl_3$ in MeCN solution with 14.0 eq 2 M $NaOH_{(aq)}$ resulted in the greatest conversion from α-*N*-dansyl-labelled heptamer **27** to α-*N*-dansyl-labelled di-thiophosphoramidate **28a** (ca. 68%). While there remained a significant proportion of α-*N*-dansyl-labelled mono-thiophosphoramidates **28b** and **28c**, it was decided that 3.0 eq $PSCl_3$ would be the upper limit for the number of equivalents of $PSCl_3$ used. This upper limit of 3.0 eq $PSCl_3$ was decided on because, in theory, this number of equivalents of thiophosphorylating agent should result in greater conversion to a di-ε-thiophosphoramidate intermediate while limiting the formation of potential thiophosphorylated side-products should this one-pot, two-step stapling methodology be implemented on an unprotected peptide substrate containing other nucleophilic residues in the future.

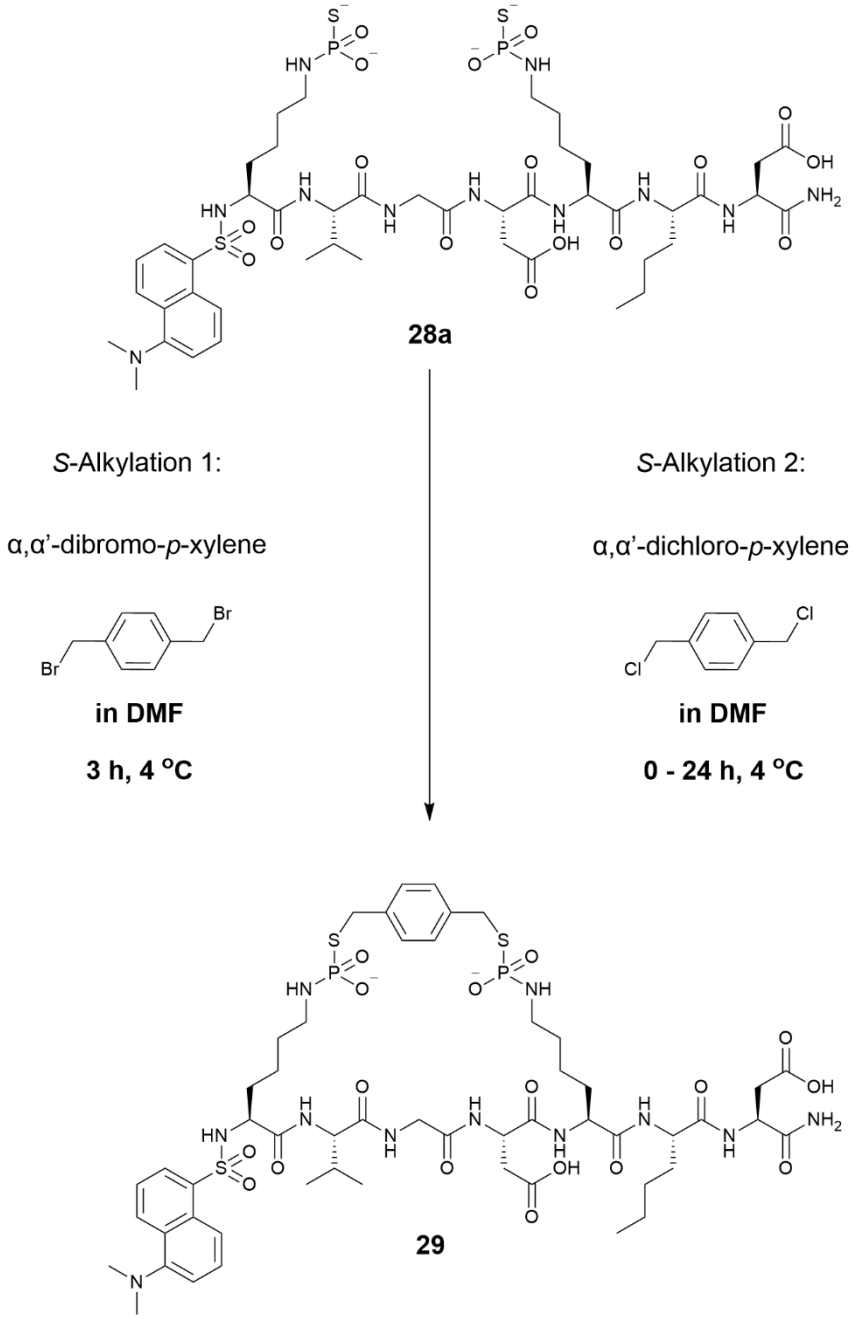
Having determined that 3.0 eq $PSCl_3$ resulted in the greatest conversion to α-*N*-dansyl-labelled di-thiophosphoramidate **28a** from α-*N*-dansyl-labelled heptamer **27**, we chose to look into how the choice of dihalogenated xylene-based linker may impact the *S*-alkylation step of our one-pot, two-step stapling procedure.

5.3.2.2. *S*-alkylation of di-thiophosphoramidate **28a** with a dibromo- linker versus a dichloro- linker

For the initial stapling experiment discussed in section 5.3.2, a dibromo-xylene linker was used. Following the *S*-alkylation step, a vast array of both *S*- and *N*-alkylated side-products appeared to have been formed (based on *m/z* values seen following LC-MS DAD analysis). In an effort to limit the formation of alkylated side-products, a dichloro-xylene linker was investigated.

As bromide is known to be a better leaving group than chloride in S_N2 reactions, meaning that organobromide reagents are more reactive and often less selective than organochloride reagents, we hypothesised that substituting α,α'-dibromo-*p*-xylene with α,α'-dichloro-*p*-xylene would result in the generation of fewer unwanted side-products while still resulting in the formation of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** from α-*N*-dansyl-labelled di-thiophosphoramidate **28a**.

To compare the dibromo- and dichloro-based stapling agents, two stapling experiments were run concurrently. Both reactions involved the use of 1.0 eq linker solutions (made up to 10 mM in DMF) following overnight *N*-thiophosphorylation. The procedure is shown in Scheme 5.5.



Scheme 5.5 Comparative *S*-alkylations of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to give α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** using either 1.0 eq α,α'-dibromo-*p*-xylene (*S*-Alkylation 1) or 1.0 eq α,α'-dichloro-*p*-xylene (*S*-Alkylation 2).

While the *S*-alkylation using α,α'-dibromo-*p*-xylene was run for 3 h, the *S*-alkylation using α,α'-dichloro-*p*-xylene was allowed to run for a total of 24 h, with aliquots taken for LC-MS DAD analysis at various time points (3, 6, and 24 h). After 3 h, there was still a significant presence of di-thiophosphoramidate **28a**, so the reaction was allowed to continue. The proportions of stapled di-thiophosphoramidate **29** at 6, 24, and 48 h did not change significantly following 6 h, so it was determined that the reaction had gone to completion (i.e. the α,α'-dichloro-*p*-xylene linker had been consumed) after 6 h.

In an effort to obtain a clearer picture regarding the conversion of α-*N*-dansyl-labelled heptamer **27** to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**, a longer LC gradient was employed in LC-MS DAD analysis. LC-MS data previously presented throughout this chapter consisted of a 5 min LC gradient which was insufficient for baseline separation between species. From this point on, however, all LC-MS chromatographs were obtained using a 10 min LC gradient. This longer chromatography gradient was implemented to try to improve the separation of heptamer **27**-based species, and therefore allow for more accurate estimations of the final conversion to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**. Details of the 10 min LC gradient LC-MS DAD analysis can be found in section 7.5.1.2.2, but it should be noted that the same column, solvent system, and mass spectrometer were used for all LC-MS data presented in Chapter 5.

The LC-MS data obtained following the implementation of the 10 min LC gradient analysis method for both *S*-alkylation reactions are shown in Figure 5.15. The peak highlighted in blue and labelled as *m/z* 1249 Da is a major side product formed during the stapling step, the structure of which is shown in Figure 5.17.

.

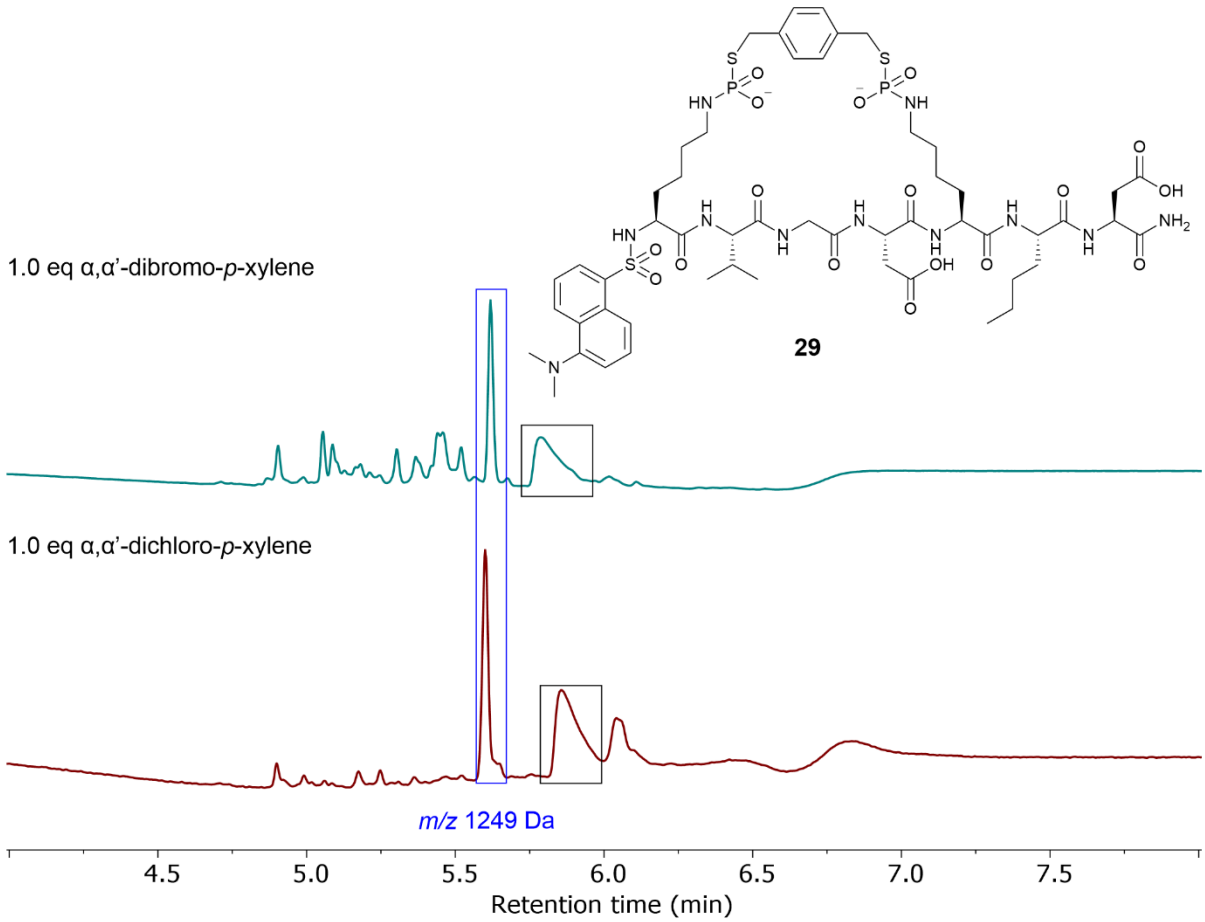


Figure 5.15 325 nm chromatographs obtained from LC-MS DAD analyses following *S*-alkylations performed on crude α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to generate α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (*m/z* 1300 Da) using either 1.0 eq α,α'-dibromo-*p*-xylene (top trace) or 1.0 eq α,α'-dichloro-*p*-xylene (bottom trace). The peaks corresponding to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** are highlighted in black. The concentrations of each reagent in the total reaction volumes were: α-*N*-dansyl-labelled di-thiophosphoramidate **28a** = 4.59 mM (for both reactions); α,α'-dibromo-*p*-xylene = 4.59 mM; α,α'-dichloro-*p*-xylene = 4.58 mM.

By comparing the LC-UV-vis chromatographs in Figure 5.15, the *S*-alkylation using α,α'-dichloro-*p*-xylene appears to have resulted in a significantly 'cleaner' reaction, meaning there appears to be far fewer side-product peaks and those that are present are generally smaller following the dichloro-stapling compared to the dibromo- stapling. Approximately 32% of the total UV absorbance at 325 nm is accounted for by α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (5.8 min; *m/z* 1300 Da) following stapling with the α,α'-dichloro-*p*-xylene linker (bottom trace). In contrast, ca. 25% of the total UV absorbance at 325 nm is accounted for by α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (5.8 min; *m/z* 1300 Da) following stapling with the α,α'-dibromo-*p*-xylene linker (top trace).

Based on the marked increase in UV absorbance associated with stapled di-thiophosphoramidate **29**, all subsequent stapling experiments were carried out using α,α'-dichloro-*p*-xylene as the stapling agent.

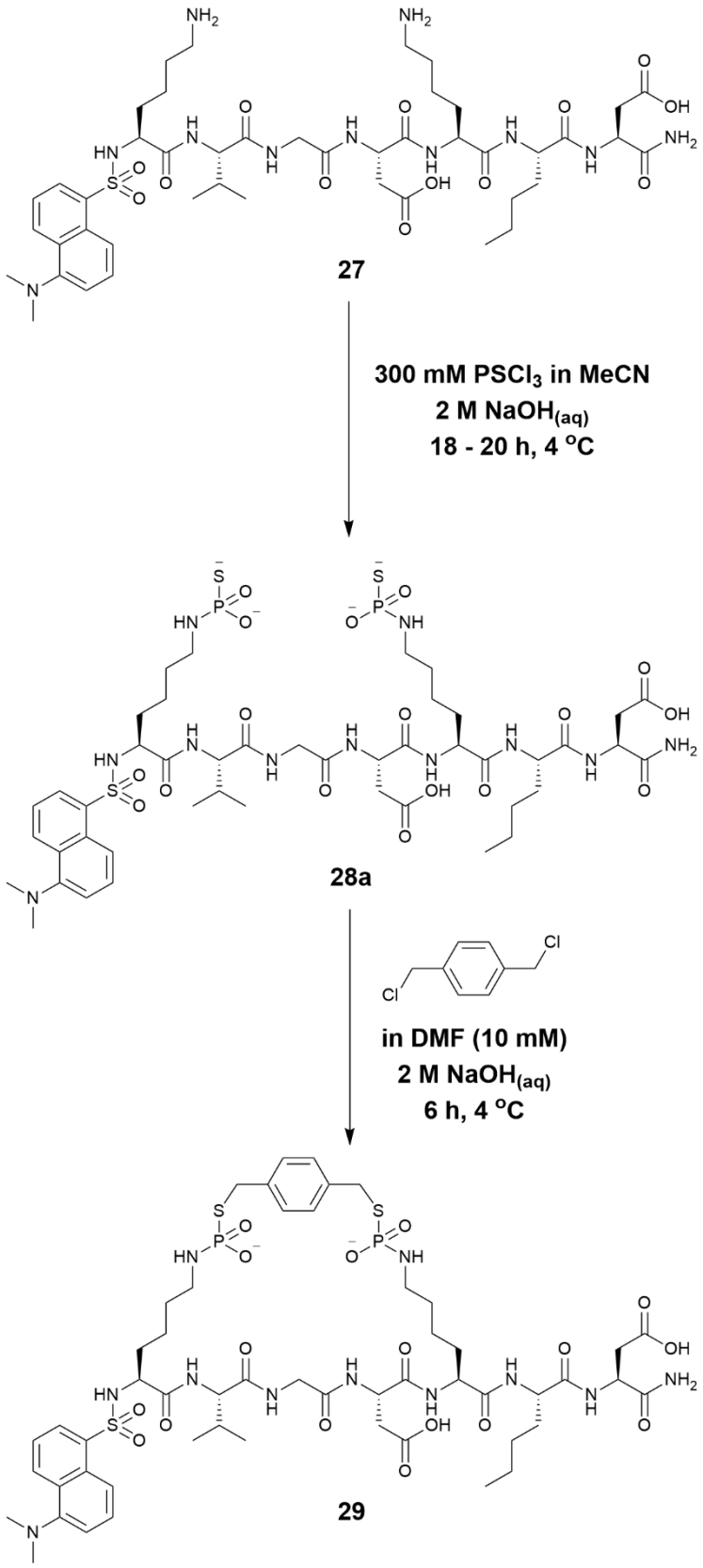
5.4. Stapling experiments on isolated α-*N*-dansyl-labelled heptamer 27

Having obtained promising results from initial one-pot, two-step stapling experiments carried out on crude samples of α-*N*-dansyl-labelled heptamer **27**, isolated samples of heptamer **27** (discussed in section 5.2.2.2) underwent experiments to further optimise the stapling procedure.

5.4.1. Complications from using DMF as a co-solvent during *S*-alkylation

A small-scale stapling experiment was performed on a sample of isolated α-*N*-dansyl-labelled heptamer **27**, implementing the *N*-thiophosphorylation procedure optimised in section 5.3.2.1 (3.0 eq $PSCl_3$ and 14.0 eq $NaOH_{(aq)}$) and using 1.0 eq α,α'-dichloro-*p*-xylene as the bifunctional linker during the stapling step. A further 2.0 eq of 2 M $NaOH_{(aq)}$ were added alongside the dichloro- linker to minimise any potential reduction in pH following the stapling step.

An outline of the stapling procedure is illustrated in Scheme 5.6.



Scheme 5.6 Overview of one-pot, two-step stapling procedure to transform α-*N*-dansyl-labelled heptamer **27** to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** *via* an α-*N*-dansyl-labelled di-thiophosphoramidate **28a** intermediate.

While the overnight *N*-thiophosphorylation step proceeded with a 90% conversion to α-*N*-dansyl-labelled di-thiophosphoramidate **28a**, the stapling step appeared to result in ca. 24% conversion to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** following a 6 h reaction. Further analysis of the LC-MS data obtained from the crude stapling mixture revealed the significant presence of two by-products with *m/z* values of 1249 Da and 1493 Da. Figure 5.16 shows the LC-UV absorbance chromatograph at 325 nm obtained following the 6 h stapling step along with the extracted ion count (EIC) traces of the 1249 Da and 1493 Da side-products, and the EIC trace corresponding to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (*m/z* 1300 Da).

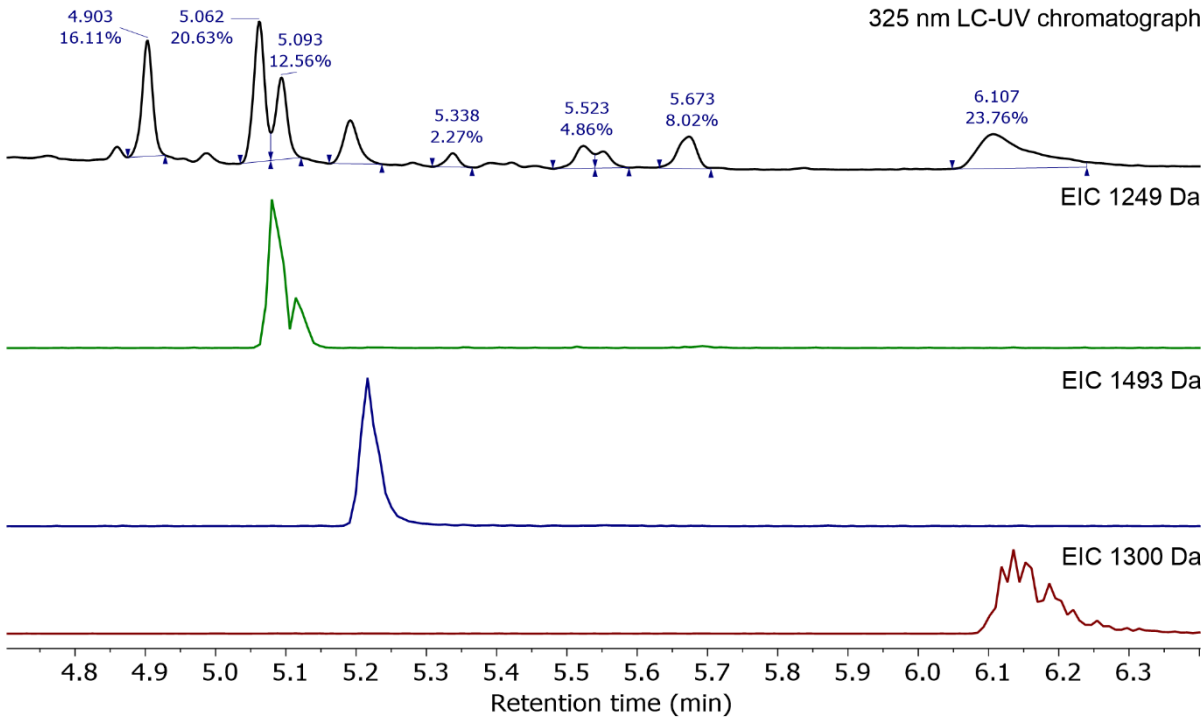
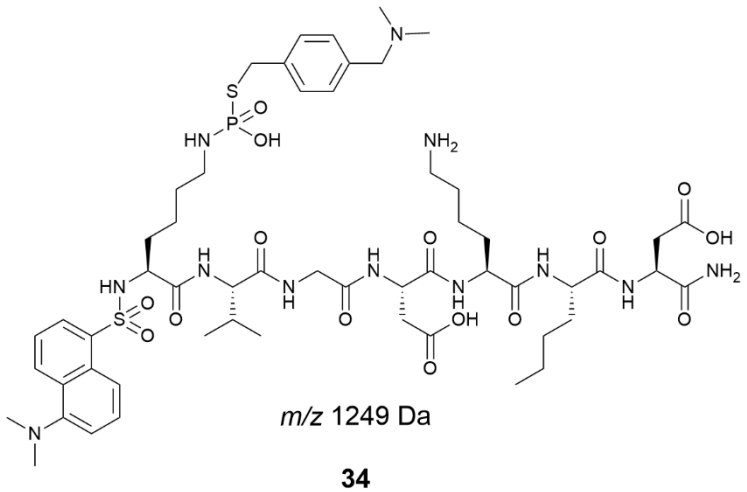


Figure 5.16 Data collected following 6 h *S*-alkylation of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** with 1.0 eq α,α'-dichloro-*p*-xylene in DMF solution. In descending order, the traces show: 325 nm chromatograph following 6 h *S*-alkylation of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to generate α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (*m/z* 1300 Da); extracted ion count (EIC) for by-product with *m/z* 1249 Da; EIC for by-product with *m/z* 1493 Da; EIC for α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (*m/z* 1300 Da).

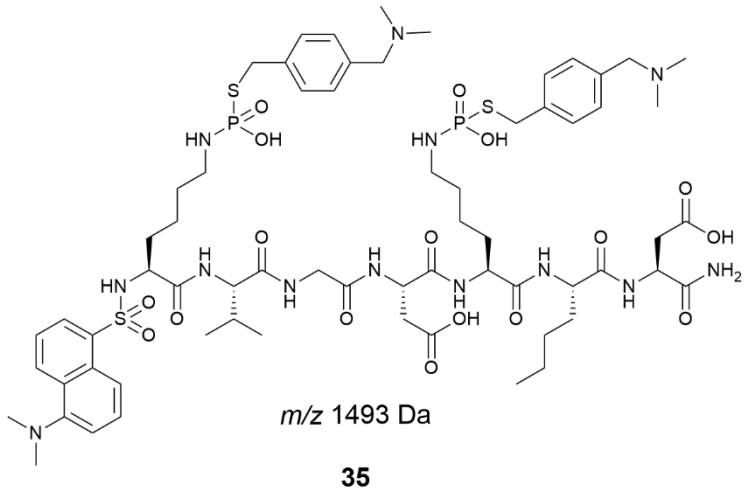
Based on the LC-UV chromatograph at 325 nm, the 1249 Da species appears to be present in more than one isomer as the LC-UV peaks corresponding to the 1249 Da species have a similar pattern to the two possible mono-thiophosphoramidate isomers **28b** and **28c** formed during *N*-thiophosphorylation of α-*N*-dansyl-labelled heptamer **27** (as seen in Figure 5.11). Under this assumption, it was believed that the 1249 Da by-product was formed through reaction at one of the two lysine side chains present on heptamer **27**. The 1493 Da species, on the other hand, is eluted as one peak suggesting only one isomer exists within this reaction mixture. This assumption led us to

believe that the 1493 Da by-product was formed through reaction at both lysine side chains present on heptamer **27**.

Following a consultation of the literature, it was found that DMF is known to react with NaOH to produce dimethylamine and formate.[173] With this in mind, potential structures of both the 1249 Da and 1493 Da by-products were determined and are shown in Figure 5.17.



34



35

Figure 5.17 Structures for the two major by-products formed through reaction with dimethylamine.

For both by-products **34** and **35**, the second chlorine on the xylene-based linker has been substituted by dimethylamine, thus preventing any potential for peptide cyclisation.

5.4.2. Using MeCN as the *S*-alkylation step co-solvent

Having found that DMF was not an inert solvent under the reaction conditions used, the solvent used for dissolving the xylene-based linker was changed to MeCN and another small-scale stapling experiment was run. The one-pot, two-step stapling procedure used was the same as outlined in Scheme 5.6, but with the 10 mM α,α'-dichloro-*p*-xylene solution made with MeCN instead of DMF.

The overnight aqueous *N*-thiophosphorylation step proceeded with 90% conversion to α-*N*-dansyl-labelled di-thiophosphoramidate **28a** according to LC-MS DAD analysis (325 nm LC-UV absorbance chromatograph).

The *S*-alkylation mixture was analysed by LC-MS DAD after 6 h, and the resulting 325 nm LC-UV absorbance chromatograph is shown in Figure 5.18 with each peak labelled with the corresponding compound number. Despite the *S*-alkylation having not gone to completion within 6 h, the 6 h *S*-alkylation chromatograph was chosen for Figure 5.18 (MeCN co-solvent) to compare with the 325 nm LC-UV chromatograph shown in Figure 5.16 (DMF co-solvent) as both chromatographs were taken at the same time point following the addition of the stapling agent.

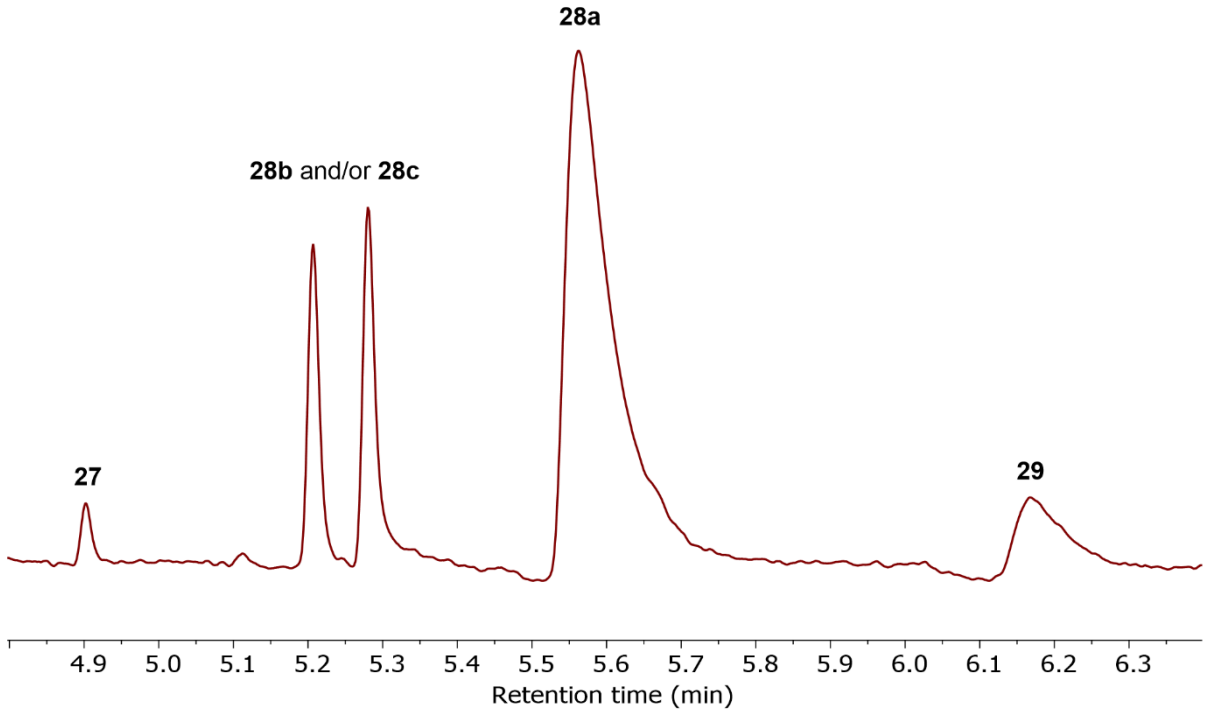


Figure 5.18 325 nm chromatograph obtained from LC-MS DAD analysis following aqueous *N*-thiophosphorylation and 6 h *S*-alkylation using α,α'-dichloro-*p*-xylene in MeCN solution performed on α-*N*-dansyl-labelled heptamer **27**. Each of the peaks have been labelled with the corresponding species number. The concentrations of each reagent in the total reaction volumes were: α-*N*-dansyl-labelled di-thiophosphoramidate **28a** = 4.12 mM; α,α'-dichloro-*p*-xylene = 4.12 mM.

When compared to the 325 nm LC-UV chromatograph obtained following *S*-alkylation of crude α-*N*-dansyl-labelled di-thiophosphoramidate **28a** with α,α'-dichloro-*p*-xylene solution in DMF presented in Figure 5.16, *S*-alkylation carried out using α,α'-dichloro-*p*-xylene solution in MeCN appears to be a more selective reaction. As Figure 5.18 shows, the only significant product following *S*-alkylation was α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (6.2 min; 1300 Da).

While there are relatively few heptamer-based side-products formed over the course of the 6 h *S*-alkylation experiment when using MeCN as the linker solvent, the *S*-alkylation reaction appears to be slower than when DMF was used as the linker solvent. In addition, the 325 nm chromatograph in Figure 5.16 (DMF co-solvent) does not appear to show any significant peak indicative of α-*N*-dansyl-labelled di-thiophosphoramidate **28a**, however Figure 5.18 (MeCN co-solvent) shows a large peak at 5.6 min corresponding to the di-thiophosphoramidate intermediate. The relative percentage proportions of the peaks labelled in Figure 5.18 are reported in Table 5.2.

Table 5.2 Percentage proportions of each species indicated in Figure 5.18. Proportions were calculated from the areas under the labelled peaks (from a 325 nm LC-MS DAD chromatograph). The percentage proportions calculated from the peaks corresponding to α-*N*-dansyl-labelled mono-thiophosphoramidates **28b** and **28c** were combined as the peak identities cannot be definitively assigned.

Species number	Retention time/ min	Mass/ Da	Proportion/ %
27	4.90	1006	2
28b and **28c**	5.21/5.28	1102	21
28a	5.56	1198	66
29	6.17	1300	11

As Table 5.2 shows, α-*N*-dansyl-labelled di-thiophosphoramidate **28a** makes up 66% of all species detected at 325 nm in the crude *S*-alkylation mixture, while α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** accounts for 11%. A potential explanation for this low conversion to stapled di-thiophosphoramidate **29** is the 'salting-out' effect observed between the aqueous and organic solvents used in the *S*-alkylation reaction.

NaOH is an inorganic salt made of up two kosmotropic (i.e. water ordering) ions (Na^+ and OH^-). When the concentration of kosmotropic ions in a miscible water-organic solvent surpasses a certain threshold concentration, a bi-phasic mixture forms made up of a 'water-rich' phase and an 'organic-rich' phase.[174] So, while water and MeCN are miscible solvents, the addition of a relatively large quantity of NaOH can result in the formation of a biphasic mixture due to this salting-out effect.

The bi-phasic mixture seen during *S*-alkylation may therefore lead to greater reaction selectivity between the stapling agent and di-thiophosphoramidate **28a**. Both α-*N*-dansyl-labelled heptamer **27** and α-*N*-dansyl-labelled di-thiophosphoramidate **28a** are insoluble in MeCN, while α,α'-dichloro-*p*-xylene is insoluble in water. *S*-alkylation reactions between di-thiophosphoramidate **28a** and the

xylene-based stapling agent are therefore limited to occurring at water-MeCN solvent interfaces rather than in a homogeneous solution, resulting in a more selective reaction and fewer, if any, side-products formed.

Having found stapling conditions that result in clean conversion of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**, the synthesis of stapled di-thiophosphoramidate **29** was scaled up with a view to isolation and characterisation.

5.5. Scaled up stapling and isolation of α-*N*-dansyl-labelled heptamer 27

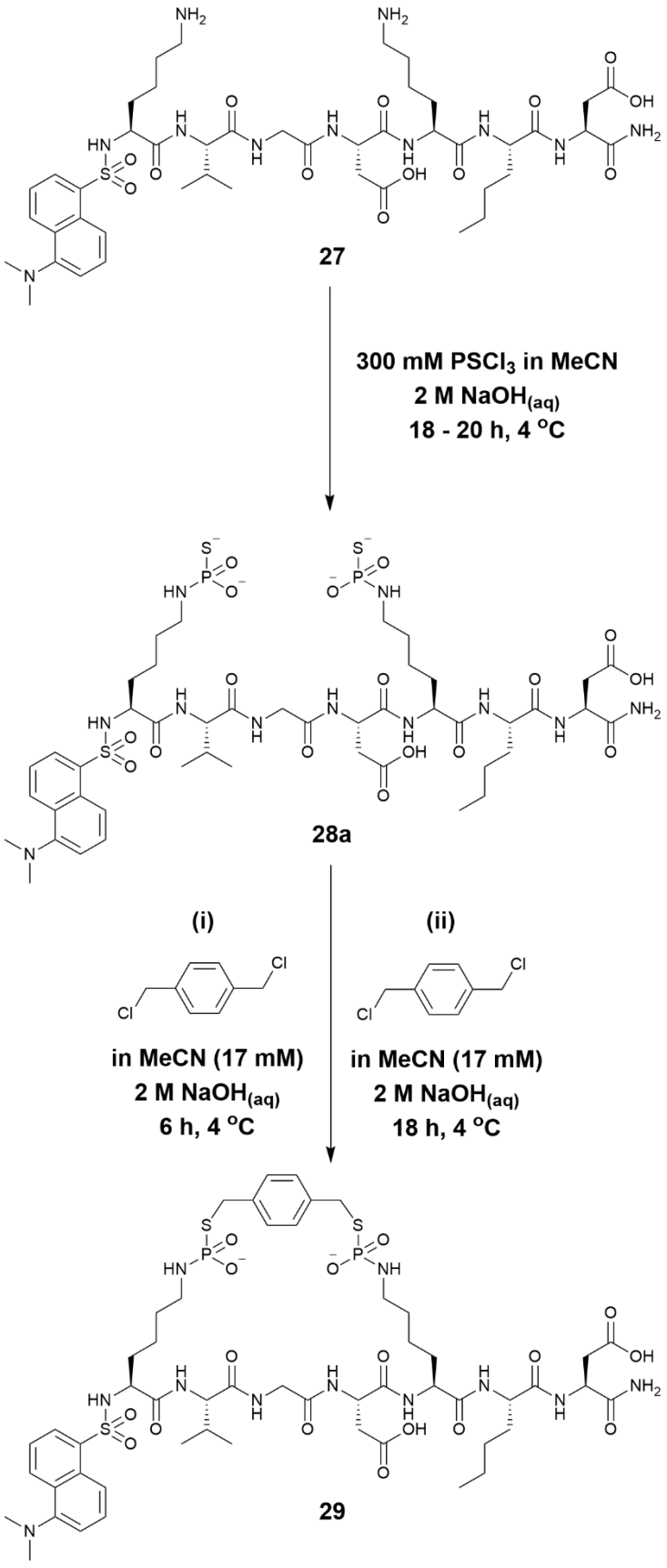
Despite having developed a one-pot, two-step procedure for the stapling of *bis*-lysine-containing peptide substrates that results in the generation of a stapled di-thiophosphoramidate peptide with minimal side-product evolution, a high conversion to the desired stapled di-thiophosphoramidate peptide product had not been obtained. It was decided, therefore, to conduct a scaled-up experiment involving two additions of the stapling agent, α,α'-dichloro-*p*-xylene, at 6 h intervals during the *S*-alkylation step. The second addition of α,α'-dichloro-*p*-xylene was implemented as the *S*-alkylation was found to have plateaued 6 h after the first addition of α,α'-dichloro-*p*-xylene.

5.5.1. Scaled up stapling of α-*N*-dansyl-labelled heptamer **27**

In order to obtain a greater conversion from α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** an experiment was carried out involving two separate additions of 1.0 eq α,α'-dichloro-*p*-xylene during the *S*-alkylation step, 6 hours apart. The reaction scheme is illustrated in Scheme 5.7.

A fresh stock solution of α,α'-dichloro-*p*-xylene in MeCN was made up at a slightly higher concentration (17 mM rather than the 10 mM solution used previously) in an effort to improve the conversion to the desired stapled peptide. By increasing the concentration of the linker solution, the relative concentration of the linker in the total *S*-alkylation mixture increases.

As mentioned in section 5.4.1, 2.0 eq 2 M $NaOH_{(aq)}$ were added with each of the two 1.0 eq α,α'-dichloro-*p*-xylene additions to account for any potential reductions in pH.



Scheme 5.7 Overview of the one-pot, two-step stapling procedure to transform α-*N*-dansyl-labelled heptamer **27** to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** *via* the α-*N*-dansyl-labelled di-thiophosphoramidate **28a** intermediate.

For the scaled-up stapling reaction, the same 11.7 mM stock solution of heptamer **27** was used as for the smaller scale reactions previously described in section 5.4. The scale-up of the one-pot, two-step stapling procedure was done in a linear fashion, with the same numbers of equivalents of each reagent being used for both steps as had been done previously. Table 5.3 compares the volumes of each reagent used in both the small-scale stapling reaction described in section 5.4.2 and the scaled-up reaction described in the following section. Only the details of the first linker addition during the scaled-up *S*-alkylation reaction are included for simplicity.

Table 5.3 Reagent details of the *N*-thiophosphorylation and first addition of the *S*-alkylation for both the small-scale and scaled-up stapling reactions performed on α-*N*-dansyl-labelled heptamer **27**. The moles of reagents listed during the *S*-alkylation take into account the sample removed between *N*-thiophosphorylation and *S*-alkylation for LC-MS DAD analysis. The concentration of the linker solution (α,α'-dichloro-*p*-xylene) differed between the small and large-scale reactions so the appropriate concentration in displayed in brackets alongside the volume used.

Reaction step	**Reagent**	**Number of Equivalents**	**Small-scale**		**Scale-up**	
			Vol/ µL	**Mol/ µmol**	**Vol/ µL**	**Mol/ µmol**
N-thiophosphorylation	11.7 mM stock of isolated **27**	1.0	42.5	0.5	500.0	5.8
	2 M $NaOH_{(aq)}$	14.0	3.5	7.0	40.9	81.9
	300 mM $PSCl_3$ in MeCN	3.0	5.0	1.5	58.5	17.5
S-alkylation	α,α'-dichloro-*p*-xylene	1.0	36.0 *(10 mM)*	0.4	338.2 *(17 mM)*	5.7
	2 M $NaOH_{(aq)}$	2.0	0.4	0.8	5.7	11.4

LC-MS DAD analysis was carried out following the overnight *N*-thiophosphorylation. The 325 nm chromatograph obtained following the overnight *N*-thiophosphorylation revealed that α-*N*-dansyl-labelled heptamer **27** had been entirely consumed, with 86% converted to α-*N*-dansyl-labelled di-thiophosphoramidate **28a**.

LC-MS DAD analysis was also conducted during the *S*-alkylation step, with aliquots taken 6 h after the first addition of 1.0 eq of α,α'-dichloro-*p*-xylene and 18 h after the second addition of 1.0 eq of α,α'-dichloro-*p*-xylene (total *S*-alkylation reaction time was 24 h). The 325 nm chromatographs obtained following *S*-alkylation for 6 and 24 h are displayed in Figure 5.19.

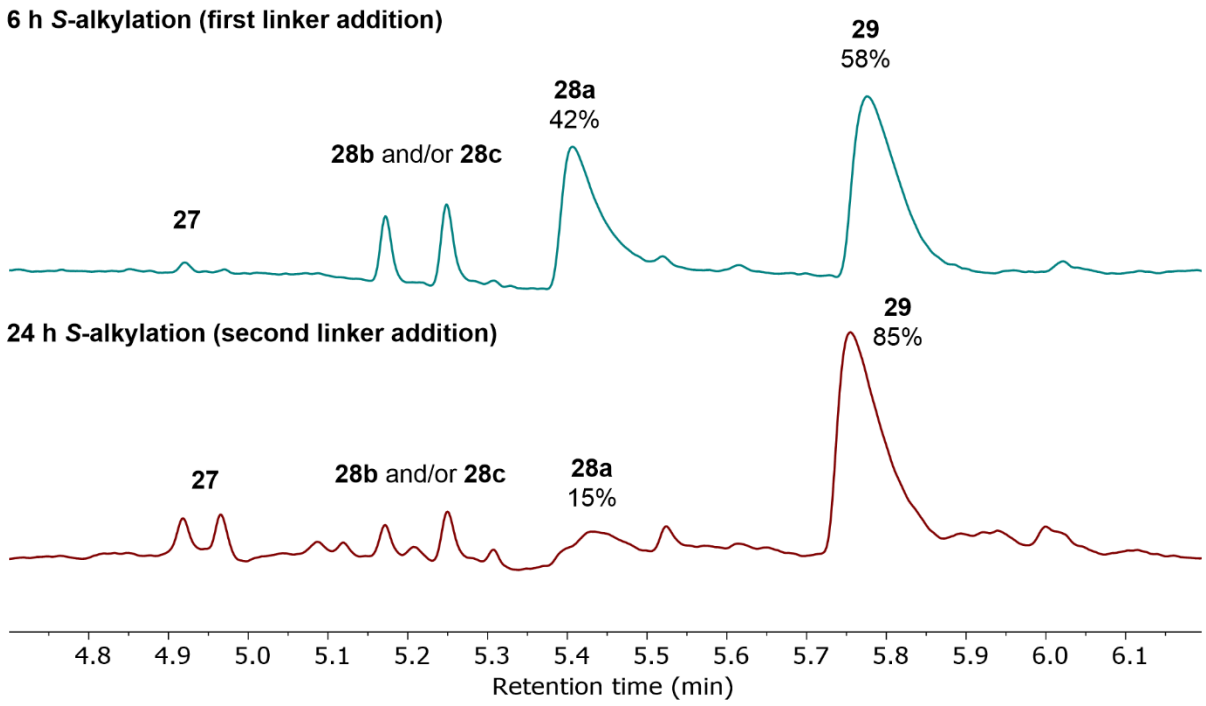


Figure 5.19 325 nm chromatograph obtained from LC-MS DAD analysis following a double addition *S*-alkylation of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to form α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**. The top trace shows the 6 h *S*-alkylation progress after the first addition of α,α'-dichloro-*p*-xylene. The bottom trace shows the 24 h *S*-alkylation progress after the second addition of α,α'-dichloro-*p*-xylene. The conversion to **29** from **28a** at 6 and 24 h were calculated from the areas under the peaks corresponding to **28a** and **29**. The concentrations of each reagent in the total reaction volumes were: α-*N*-dansyl-labelled di-thiophosphoramidate **28a** = 6.09 mM (after the first linker addition) and 4.63 mM (after the second linker addition); α,α'-dichloro-*p*-xylene = 6.09 mM (after the first linker addition) and 9.05 mM (after the second linker addition).

Following 6 h *S*-alkylation of the crude *N*-thiophosphorylation mixture with 1.0 eq α,α'-dichloro-*p*-xylene, ca. 58% α-*N*-dansyl-labelled di-thiophosphoramidate **28a** was converted to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**. While scaling up the thiophosphorylation step appears to have improved the conversion to stapled heptamer **29** over 6 h compared to the small scale thiophosphorylations previously discussed (section 5.4.2), there was still a significant presence of di-thiophosphoramidate intermediate **28a**.

After 24 h total *S*-alkylation reaction time and a second addition of 1.0 eq α,α'-dichloro-*p*-xylene after 6 h, the apparent conversion from di-thiophosphoramidate **28a** to stapled di-thiophosphoramidate **29** was 85%. It should be noted that while the conversion from di-thiophosphoramidate **28a** to stapled di-thiophosphoramidate **29** was 85%, stapled di-thiophosphoramidate **29** only accounted for ca. 67% of the total crude *S*-alkylation mixture after 24 h (based on LC-UV vis analysis at 325 nm).

Having obtained > 80% conversion of α-*N*-dansyl-labelled di-thiophosphoramidate **28a** to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**, we chose to attempt to isolate stapled di-thiophosphoramidate **29** for more in depth characterisation by LC-MS DAD and NMR spectroscopy.

5.5.2. Isolation of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**

α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** was isolated by HPLC using an Agilent 1260 Infinity II system and an Agilent 5 Preparative C_{18} column, with the help of Miss Katherine Deck. The mobile phase used was a neutral water/ MeCN gradient, as had been used for the isolation of other *N,S*-alkylated thiophosphoramidates (discussed in Chapters 2 and 3).

A crude mixture of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** was injected on to the HPLC column and a 20 min 10-95% MeCN gradient was run. The resulting 220 nm chromatograph is shown in Figure 5.20. The wavelength was set at 220 nm to ensure any species containing a peptide bond (UV absorbance at 190 – 220 nm)[170, 171] would be shown, not just the species containing a dansyl group.

In an effort to minimise any hydrolysis of stapled di-thiophosphoramidate **29**, 50 µL 2 M NaOH solution was added to the bottom of each test tube used for fraction collection prior to the HPLC run to keep the collected fractions at a high pH.

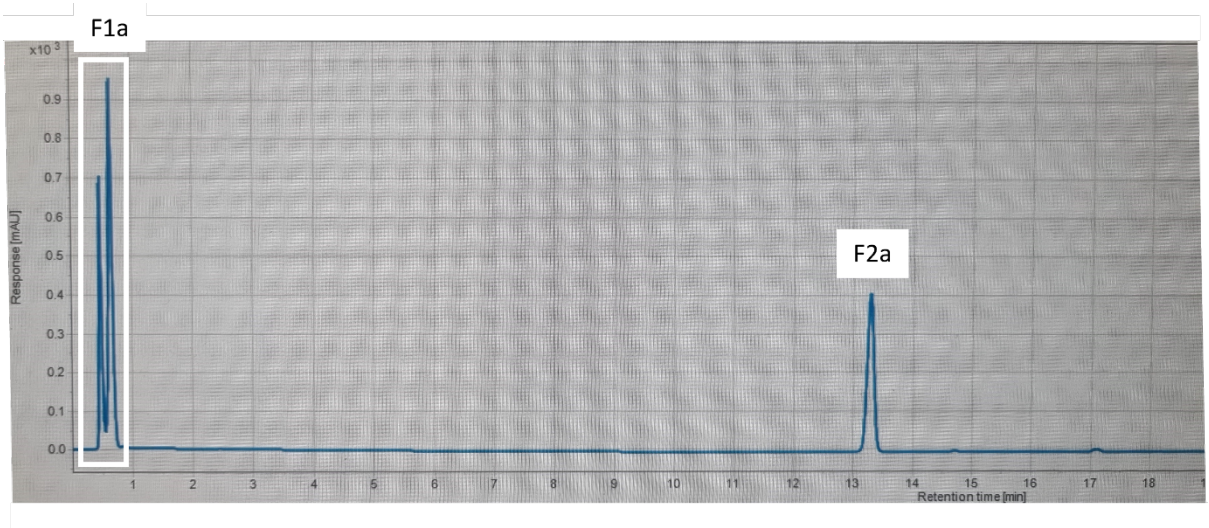


Figure 5.20 220 nm chromatograph from the first HPLC run to isolate α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**. The peaks highlighted by the white box (F1a) were found to contain α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**.

Although it is not entirely clear why stapled di-thiophosphoramidate **29** eluted so close to the solvent front (at 0.415 – 0.915 min), it may be due to the ionisation state of **29** at column injection. Following the end of the *S*-alkylation step, the pH of the crude mixture was ca. 12. At this high pH, stapled di-

thiophosphoramidate **29** has a net negative charge as both thiophosphoramidate groups are deprotonated as well as both aspartic acid residues. As the pH of the crude *S*-alkylation mixture was not adjusted prior to HPLC, the highly ionised stapled peptide may have simply run through the C_{18} column with the 90% water solvent front.

LC-MS DAD analysis was carried out on fractions F1a and F2a and stapled di-thiophosphoramidate **29** was present in F1a, highlighted in Figure 5.20 by the white box. It is not clear which of these two peaks contained stapled di-thiophosphoramidate **29** as both were collected as a single fraction due to a lack of baseline separation. Figure 5.21 displays the 325 nm chromatograph obtained from LC-MS DAD analysis of F1a, along with the extracted ion counts of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (m/z 1300 Da) and a potential hydrolysis product (m/z 1318 Da).

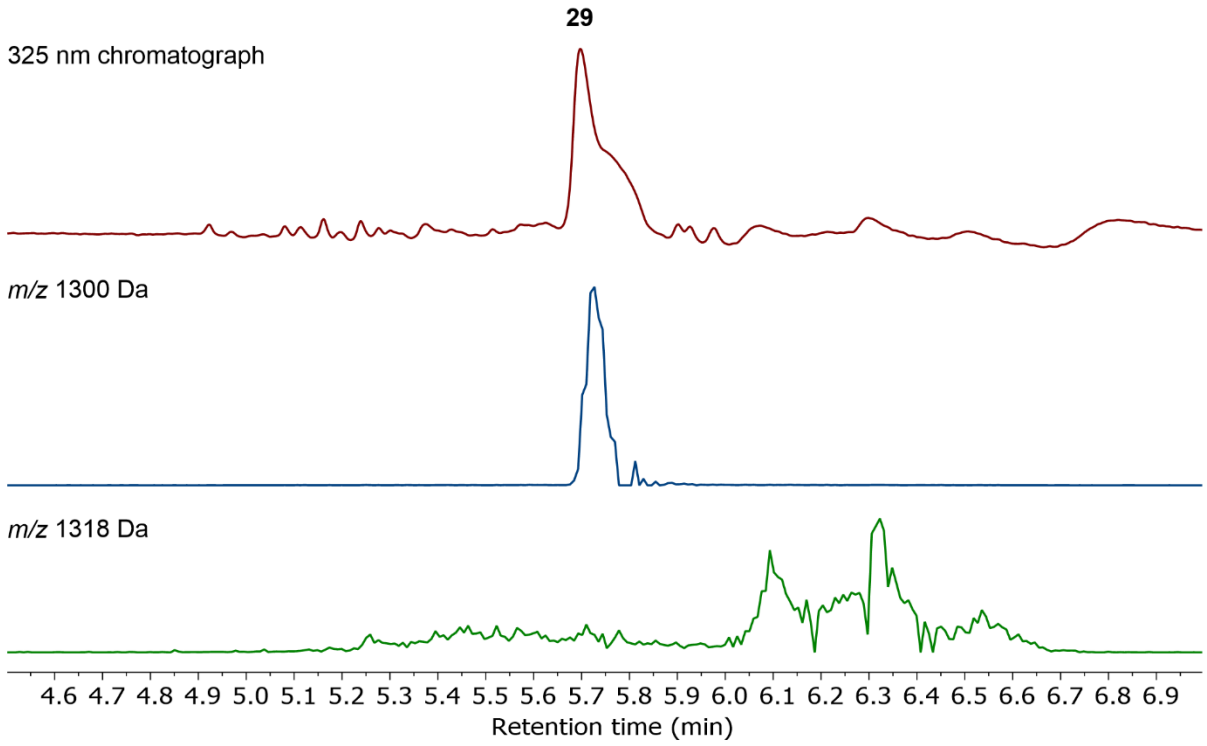


Figure 5.21 LC-MS after first HPLC run. First trace is the 325 nm chromatograph for F1a; the second trace is the mass extracted ion count of **29** (*m/z* 1300 Da); the third trace is the mass extracted ion count of potential hydrolysis product (*m/z* 1318 Da).

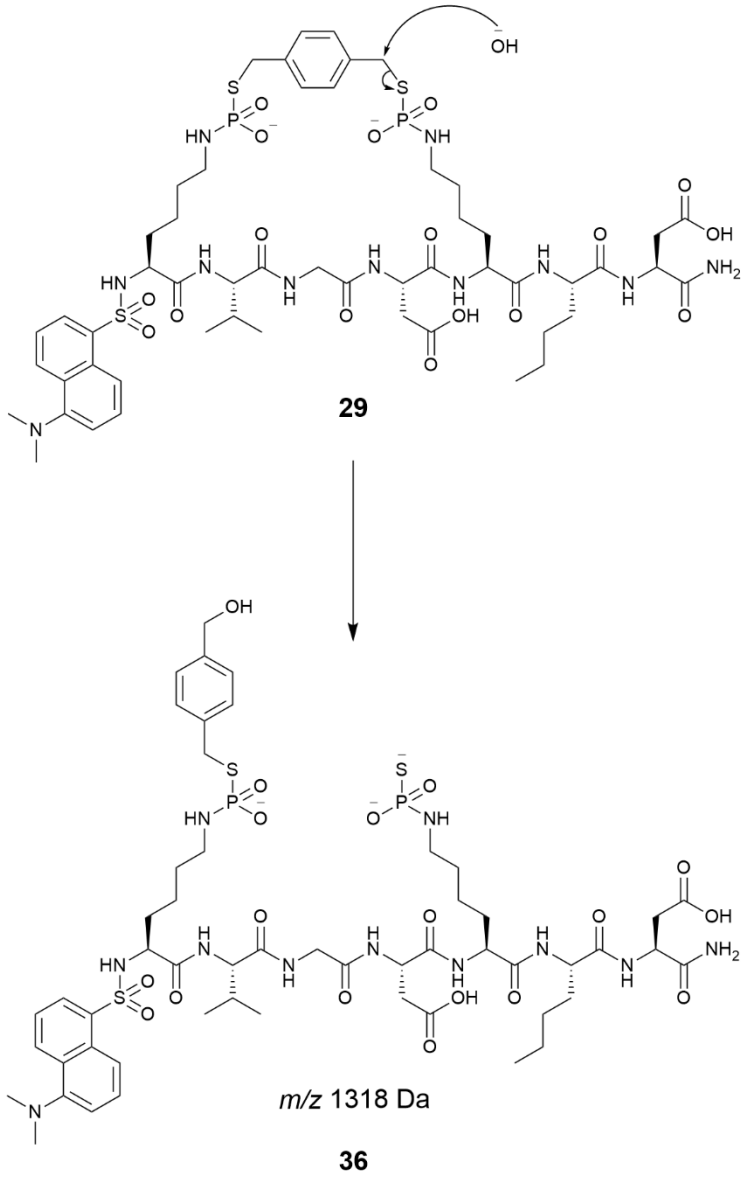
Following this initial HPLC run there was no improvement in the purity of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** according to LC-UV vis analysis, most likely due to the due to the lack of resolution between the two HPLC peaks eluted at 0.415 – 0.915 min.

The peak at 5.72 min seen on the 325 nm chromatograph correlated to stapled di-thiophosphoramidate **29** through inspection of mass extracted ion counts – **29** has a mass of 1300 Da, and the major peak in the EIC at *m/z* 1300 Da aligns with the major peak seen on the 325 nm

chromatograph. Unfortunately, the 325 nm chromatograph obtained following HPLC did not suggest any significant improvement to the purity of **29** compared to the 325 nm chromatograph obtained from the crude *S*-alkylation mixture (24 h 325 nm chromatograph shown in Figure 5.19). Although improper isolation of stapled di-thiophosphoramidate **29** through the collection of two peaks as a single fraction during HPLC impacted the final purity of the isolated stapled heptamer, the LC-MS DAD data displayed in Figure 5.21 suggested that base-catalysed hydrolysis of stapled di-thiophosphoramidate **29** may have also impacted compound purity.

As previously mentioned, 2 M $NaOH_{(aq)}$ (50.0 μL) was added to each of the collection tubes prior to fraction collection to minimise hydrolysis of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** at neutral pH. It seems, however, the large amount of hydroxide used (0.1 mmol compared to a maximum of 5.8 μmol heptamer substrate) may have resulted in the base-catalysed hydrolysis of a small amount of stapled di-thiophosphoramidate **29**, leading to the formation of the *m/z* 1318 Da species noted in Figure 5.21.

Scheme 5.8 illustrates the potential base-catalysed hydrolysis of stapled di-thiophosphoramidate **29** that resulted in the formation of by-product **36** (*m/z* 1318 Da).



Scheme 5.8 Potential mechanism for the hydroxide catalysed hydrolysis of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** to form hydrolysis by-product **36**.

A second preparative HPLC run was carried out on the same sample of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** to improve the sample purity for more in depth analyses. For this run, however, the pH of the injected sample was adjusted from ca. 13 (from the addition of 2 M $NaOH_{(aq)}$ to the collector tubes prior to the first HPLC run) down to ca. 8 with 0.01 M acetic acid to force **29** to elute at a longer retention time, and to improve the separation between the two peaks eluted close to the solvent front. Due to the base-catalysed hydrolysis seen after the first HPLC run, no base was added to the collector tubes prior to the second HPLC run.

Figure 5.22 shows the 220 nm chromatograph obtained from the second HPLC run. The column, solvent system, and gradient length were the same as for the first HPLC run shown in Figure 5.20.

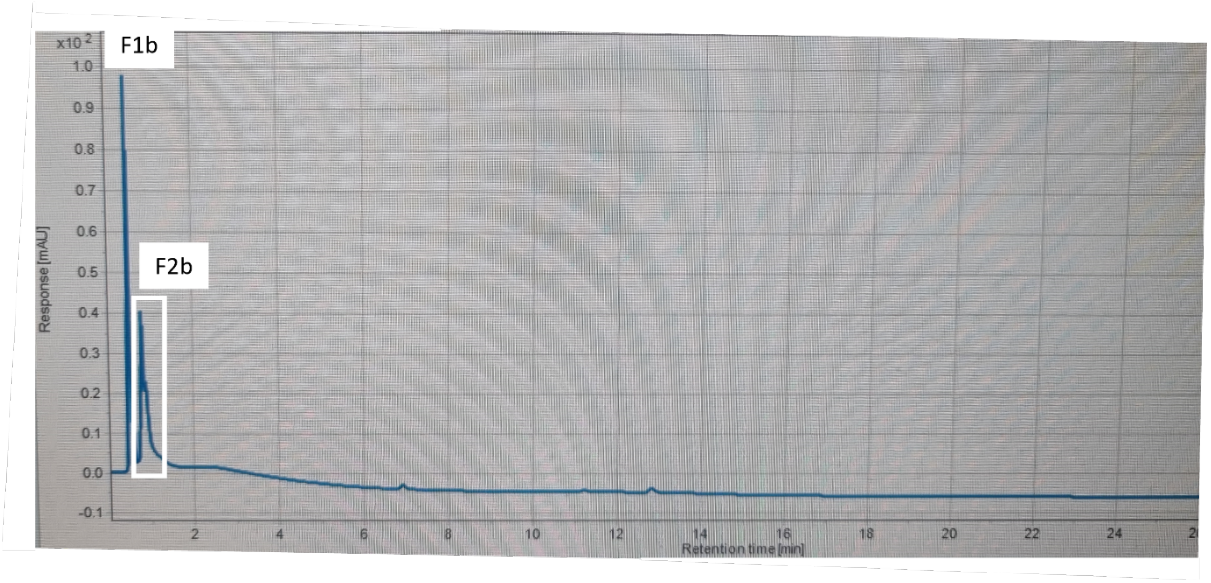


Figure 5.22 220 nm chromatograph from the second HPLC run to isolate α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**. The peak highlighted by the white box (F2b) was found to contain α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**.

As with the first HPLC run, α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** eluted early in the HPLC run, at 0.7 – 1.2 min. In this case, however, there was baseline separation between the two peaks eluted near the solvent front. These two peaks were collected as individual fractions (F1b and F2b), the solvents were removed by lyophilisation, the recovered solids were dissolved in water and submitted for LC-MS DAD analysis.

F2b was found to contain stapled di-thiophosphoramidate **29** (highlighted in white in Figure 5.22), and the 325 nm chromatograph obtained from LC-UV vis analysis is shown in Figure 5.23.

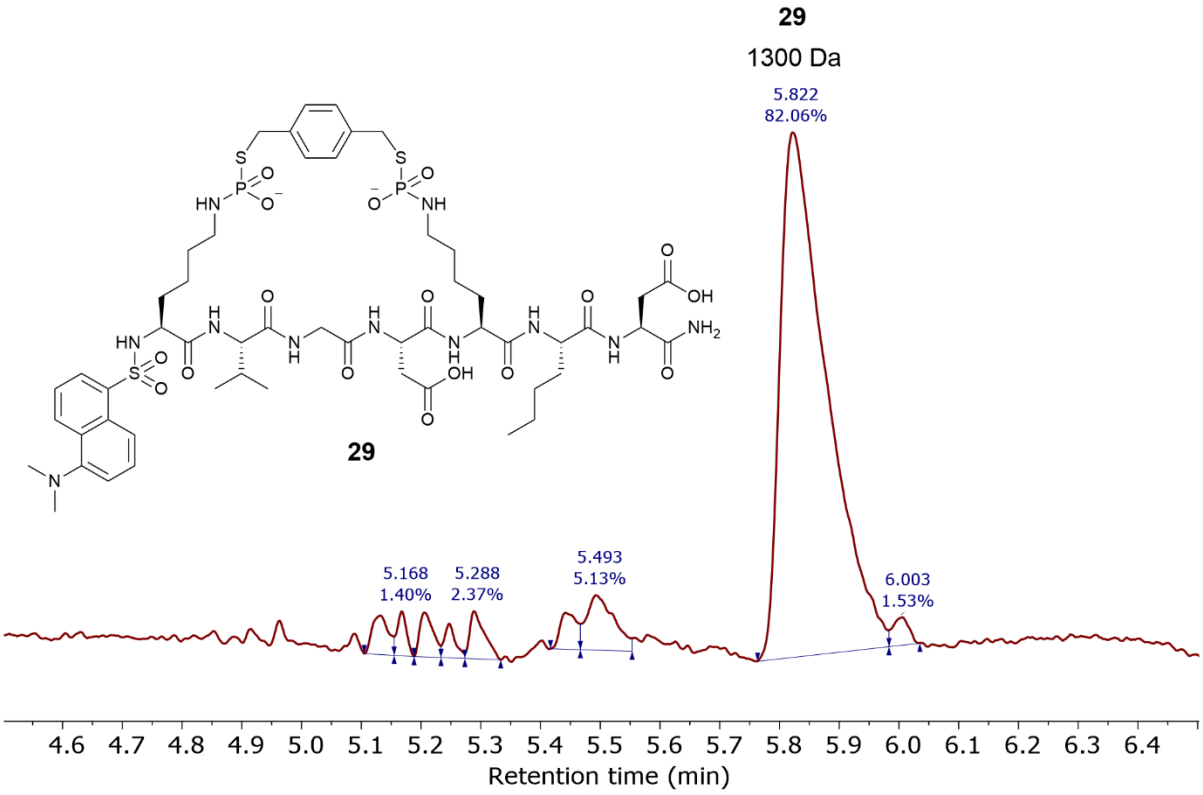


Figure 5.23 325 nm chromatograph obtained from LC-MS DAD analysis of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (*m/z* 1300 Da) following the second HPLC run.

Based on LC-UV vis analysis at 325 nm, the two HPLC purifications had resulted in a < 1 mg sample of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** at 82% purity.

A 2D ^{1}H-^{31}P HMBC NMR experiment was run in order to confirm the staple position in stapled di-thiophosphoramidate **29** through the coupling of phosphorus atoms present in both thiophosphoramidate groups to nearby hydrogen atoms. Prior to obtaining the 2D NMR spectrum, the chemical shifts of the phosphorus and hydrogen atoms in question could be predicted through consultation of work previously described in this thesis and carried out in the Hodgson group. The predictions of the phosphorus chemical shifts were based on those seen for *S*-propyl ε-*N*-thiophosphoramidate **8** (Chapter 2, Figure 2.16 – δ_P 25.23 ppm) and *S*-naphthyl ε-*N*-thiophosphoramidate **23d** (Chapter 3, Figure 3.32– δ_P 24.02 ppm). The chemical shifts predicted for the ε-CH_2 protons on the lysine side chains were also based on those seen for *S*-propyl ε-*N*-thiophosphoramidate **8** (Appendix 1, Figure 8.17 – δ_H 2.42 ppm). In respect of the chemical shifts of the -SCH_2 protons, the predicted δ_H value was based on -SCH_2 proton chemical shift values reported by Trmčić for a series of related *N,S*-alkylated thiophosphoramidate species (δ_H ca. 4.5).[175]

Figure 5.24 outlines the predicted chemical shifts for each pair of methylene protons and both phosphorus centres. The structures and chemical shifts of the related species are shown for reference.

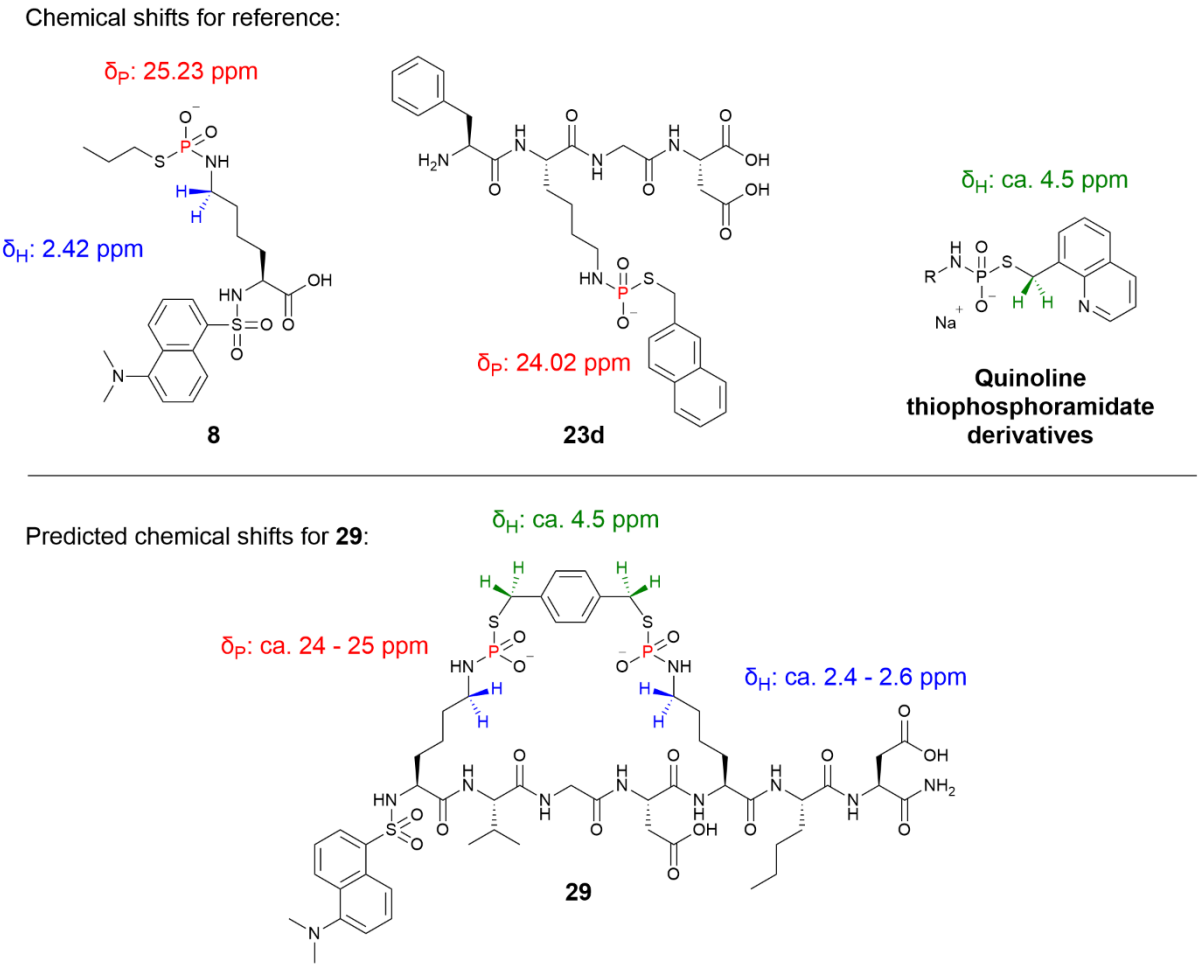


Figure 5.24 Structures and observed chemical shifts of relevant atoms for *S*-propyl ε-*N*-thiophosphoramidate **8** (Chapter 2), *S*-naphthyl ε-*N*-thiophosphoramidate **23d** (Chapter 3), and the general structure of quinoline thiophosphoramidate derivatives reported by Trmčić.[175] The structure of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** and the predicted chemical shifts of the phosphorus and hydrogen atoms of interest are also shown. The phosphorus atoms are highlighted in red; the ε-CH_2 protons are highlighted in blue; the -SCH_2 protons are highlighted in green.

The ^{1}H-^{31}P HMBC NMR spectrum for α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** is shown in Figure 5.25. Each of the relevant ^{1}H and ^{31}P signals (along the x- and y- axes respectively) are highlighted with coloured boxes corresponding to the coloured hydrogen and phosphorus atoms in the diagram of the staple structure of **29**.

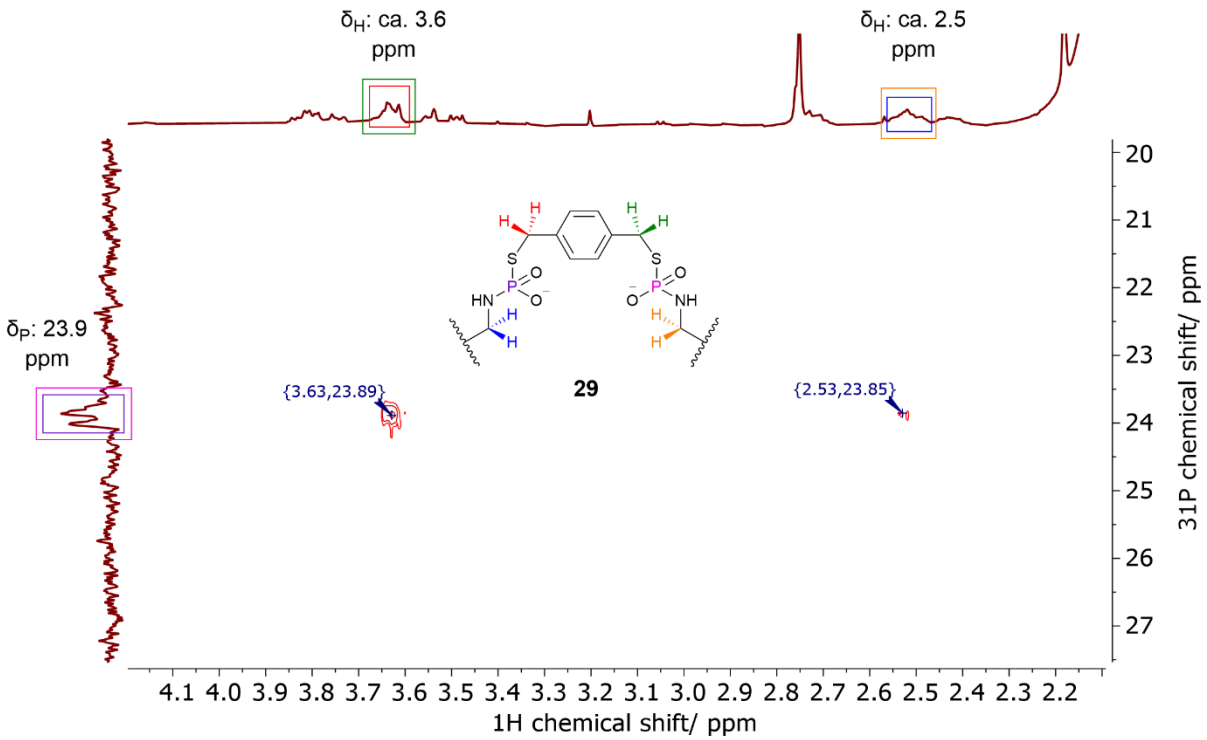


Figure 5.25 ^{1}H-^{31}P HMBC NMR spectrum of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** following the attempted isolation. The structure of the di-thiophosphoramidate-based staple in **29** is illustrated with the phosphorus atoms and protons of interest coloured to correspond to the coloured boxes surrounding the ^{31}P and ^{1}H signals highlighted in the 1D spectra along the axes.

As Figure 5.25 shows, the predicted ^{1}H and ^{31}P chemical shifts, based on those seen for *S*-propyl ε-*N*-thiophosphoramidate **8** (ε-CH_2 protons) and *S*-naphthyl ε-*N*-thiophosphoramidate **23d** (phosphorus atoms) respectively, were similar. The observed ^{1}H chemical shift for the -SCH_2 (benzylic) protons on stapled di-thiophosphoramidate **29** is shifted upfield compared to the predicted chemical shift value (predicted value – 4.5 ppm; observed value – 3.6 ppm). The observed upfield shift compared to the predicted value may be due to the nature of the aromatic group adjacent to the benzylic protons (benzene and quinoline respectively).

The ^{1}H-^{31}P HMBC NMR spectrum shown in Figure 5.25 provides corroborating evidence for the formation of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** when viewed alongside the data obtained through LC-MS DAD analysis. While LC-MS DAD analysis confirmed the molecular mass of **29** (*m/z* 1300 Da) formed through the one-pot, two-step stapling of α-*N*-dansyl-labelled heptamer **27** *via* the α-*N*-dansyl-labelled di-thiophosphoramidate intermediate **28a**, the ^{1}H-^{31}P HMBC spectrum shows the position of the thiophosphoramidate groups and the xylene-based staple through multi-bond correlations.

As the two ^{31}P NMR signals at ca. 23.9 ppm could be assigned to the thiophosphoramidate centres, the two HMBC signals in line with the ^{31}P signals at 23.9 ppm were thought to correspond to ^{1}H signals

for the ε-CH_2 protons and -SCH_2 benzylic protons. The HMBC signal at (2.53, 23.85) could be assigned to the ε-CH_2 protons on the lysine side chains correlating to the thiophosphoramidate group P-centres due to the ^{1}H signal (2.53 ppm) being in line with ^{1}H signals observed for the ε-CH_2 protons in *S*-propyl ε-*N*-thiophosphoramidate **8**. Following this assignment, it could be deduced that the HMBC signal at (3.63, 23.89) indicated a correlation between the same thiophosphoramidate group (^{31}P signal at 23.89 ppm) and -SCH_2 benzylic protons adjacent to the benzene ring in the peptide staple (^{1}H signal at 3.63 ppm). Thus, through the assignment of these two HMBC signals, we could confirm the presence of *N,S*-bridging thiophosphoramidate groups positioned on the ε-NH_2 of the two lysine side chains.

5.6. Summary

Following the design and synthesis of α-*N*-dansyl-labelled heptamer **27**, initial stapling experiments were performed on crude samples of **27** in order to develop a robust workflow for the analysis of crude stapled mixtures while simultaneously optimising the purification of **27** by chromatographic methods.

These preliminary stapling experiments were based on the one-pot, two-step *N*-thiophosphorylation and *S*-alkylation reactions previously carried out on α-*N*-dansyl-L-lysine **4** in Chapter 2 and *N*-terminal phenylalanine tetrapeptide **23a** in Chapter 3. The first step consisted of the aqueous *N*-thiophosphorylation of heptamer **27** to obtain α-*N*-dansyl-labelled di-thiophosphoramidate intermediate **28a**, before continuing with the *S*-alkylation step involving stapling of the heptamer through the anionic sulfur atoms on the two thiophosphoramidate groups introduced in the thiophosphorylation step to give α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (Scheme 5.1). Through these stapling experiments on crude heptamer **27**, it was determined that 3.0 eq $PSCl_3$ solution was optimal for the synthesis of di-thiophosphoramidate intermediate **28a** while limiting the generation of unwanted thiophosphorylated side-products in future reactions on unprotected peptides (Figure 5.14). The crude stapling reactions also allowed us to find that using α,α'-dichloro-*p*-xylene as the stapling agent rather than α,α'-dibromo-*p*-xylene results in fewer alkylated side-products, therefore giving neater conversions to stapled di-thiophosphoramidate **29** from di-thiophosphoramidate intermediate **28a** (Figure 5.15).

Following the isolation of α-*N*-dansyl-labelled heptamer **27** discussed in section 5.2.2, the stapling procedure was further optimised through the use of MeCN as a co-solvent during *S*-alkylation rather than DMF as it was found that DMF was reacting with $NaOH_{(aq)}$ present in the reaction mixture, resulting in the generation of side-products **34** and **35** (Figure 5.17).

With optimised conditions and reagents found for this novel stapling methodology, the reaction was scaled up linearly from ca. 100 µL to ca. 1 mL. To obtain greater conversions from di-thiophosphoramidate intermediate **28a** to stapled heptamer **29**, a second addition of the stapling reagent was implemented 6 h after the first addition. Through this second addition of the stapling agent, the conversion from **28a** to **29** was improved from 58% after a single addition of stapling agent to 85% after the second addition (Figure 5.19).

Finally, an attempt to isolate α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** by reverse phase HPLC was carried out. Using a 20 min 10-95% MeCN gradient with water, stapled heptamer **29** was isolated to 82% purity through two runs of chromatography. Figure 5.26 shows LC chromatographs taken at 325 nm of crude **29**, **29** after the first HPLC run, and **29** after the second HPLC run. The purity of each sample of stapled di-thiophosphoramidate **29**, as determined by LC-UV-vis analysis, is shown on the left of each trace.

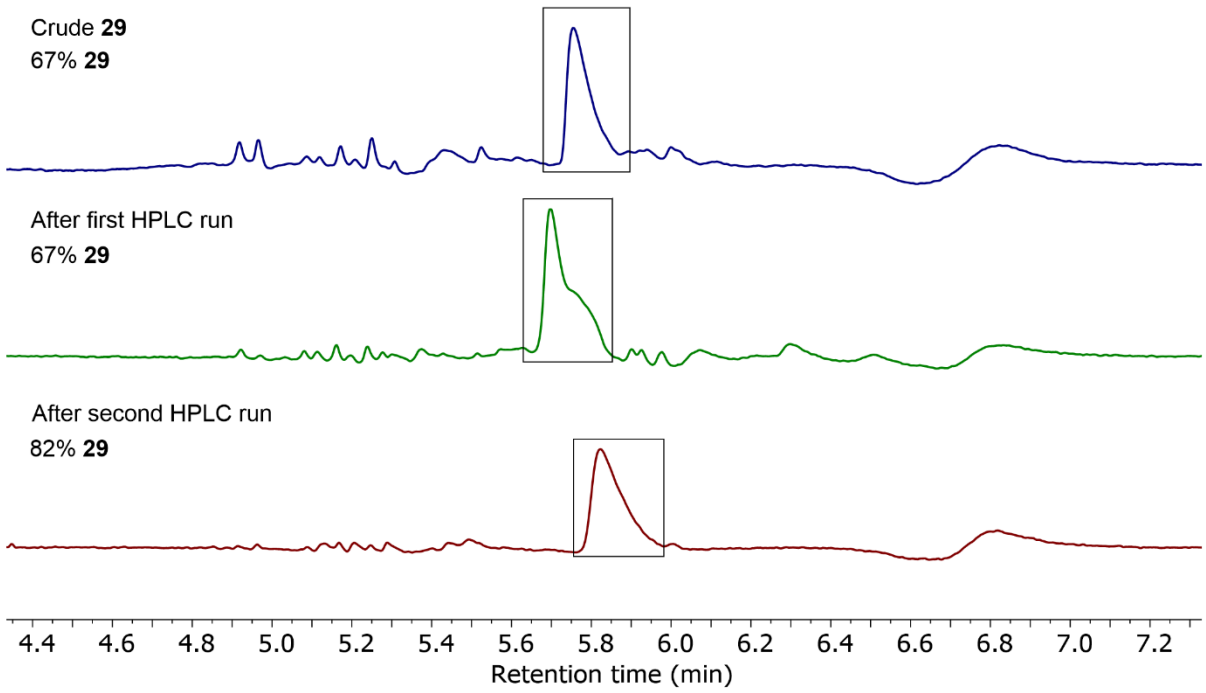


Figure 5.26 Stacked 325 nm chromatographs of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** obtained from LC-MS DAD analysis at the different stages throughout sample purification.

Following the second HPLC run, a < 1 mg yield of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** was obtained at 82% homogeneity. This sample was submitted for ^{1}H-^{31}P HMBC NMR spectroscopy analysis which provided corroborating evidence for the synthesis of **29** through correlations between the two phosphorus centres of the ε-*N*-thiophosphoramidate groups, the ε-CH_2 protons on the lysine

side chains, and the -SCH_2 benzylic protons on the xylene-based staple adjacent to the ε-*N*-thiophosphoramidate groups.

Due to time constraints, the purification method used to isolate **29** could not be improved. Based on how quickly **29** was eluted in both HPLC runs (Figures 5.20 and 5.22), it may be necessary to use ion exchange or ion pairing chromatography to increase the retention time of **29** and prevent elution close to the solvent front and other fast-running impurities.

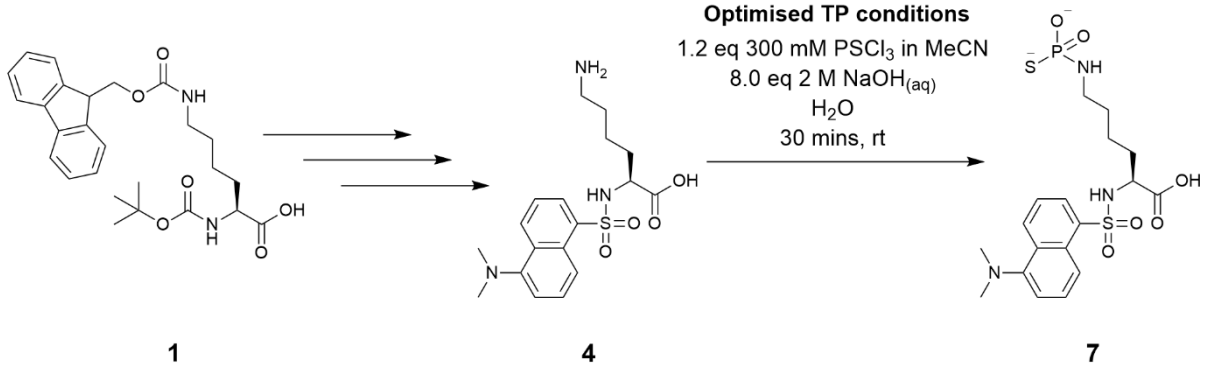
6. Conclusions and future work

Peptide stapling offers improved pharmacokinetic characteristics to peptide drugs, aiding in protease resistance and stabilising the α-helical conformation of many peptides.[79] While there are myriad peptide stapling methods already reported in the literature, many of these require the use of unnatural amino acids or orthogonal protection for selective modification. We have presented a simple one-pot, two-step stapling methodology that is selective for the ε-NH_2 group found on lysine residues, and that uses inexpensive, readily available reagents.

To round off the work presented in this thesis, a summary of the findings is presented and recommendations are outlined for the continuation of the work described herein. As conclusions were included at the end of each chapter, this summary will be brief, with the aim of pulling together the overarching picture of the work described in the thesis.

6.1. Summary of optimisation of aqueous *N*-thiophosphorylation

In order to develop the one-pot, two-step stapling procedure, a series of aims and objectives were laid out (section 1.5). The first of these was the design and synthesis of a lysine-based single residue model for the optimisation of aqueous *N*-thiophosphorylation at an ε-NH_2 group. Starting from an orthogonally protected lysine residue, α-*N*-Boc-ε-*N*-Fmoc-L-lysine **1**, α-*N*-dansyl-L-lysine **4** was synthesised using a three-step procedure. Following purification of α-*N*-dansyl-L-lysine **4** by reverse phase flash chromatography, a series of experiments was carried out for the optimisation of aqueous *N*-thiophosphorylation (section 2.2). The optimal conditions for aqueous *N*-thiophosphorylation on a lysine ε-NH_2 group are displayed in Scheme 6.1.



Scheme 6.1 Simplified synthetic route for the synthesis of α-*N*-dansyl thiophosphoramidate **7** from α-*N*-Boc-ε-*N*-Fmoc-L-lysine **1**. Refer back to Schemes 2.1 – 2.4 for details on the synthesis of α-*N*-dansyl-L-lysine **4**. TP – thiophosphorylation.

After optimising aqueous *N*-thiophosphorylation on a lysine-based small molecule model, a lysine-containing tetrapeptide model (α-*N*-dansyl tetrapeptide **9**) was designed for optimisation of *N*-thiophosphorylation on a larger system (section 2.3). After a short experimental series, it was found that using 7.5 eq 300 mM $PSCl_3$ in MeCN and 35.7 eq 2 M $NaOH_{(aq)}$ resulted in up to 95% conversion of tetrapeptide **9** to ε-*N*-thiophosphorylated tetrapeptide **10** over a 30 min reaction at room temperature.

6.2. Summary of selectivity investigation

With optimised conditions for aqueous *N*-thiophosphorylation in hand, a study was conducted in which the selectivity of aqueous *N*-thiophosphorylation between ε-NH_2 and α-NH_2 positions was explored. Following a few attempts at determining thiophosphorylation selectivity through MS- and LC-UV vis-based methods (sections 3.2 and 3.3), a ^{31}P NMR spectroscopy-based assay was developed.

A series of unprotected tetrapeptide models was designed, where each peptide had the same general sequence (XKGD) but with different *N*-terminal amino acids. These *N*-terminal amino acids covered a range of different sized side chains, as well as different nucleophilicities, to see how changes at the *N*-terminal position may impact selectivity of the thiophosphorylation reaction. Figure 6.1 summarises the reaction selectivities observed between the lysine ε-NH_2 and the *N*-terminal α-NH_2, where a selectivity greater than 1.0 indicates that a greater proportion of ε-*N*-thiophosphoramidate was formed than α-*N*-thiophosphoramidate.

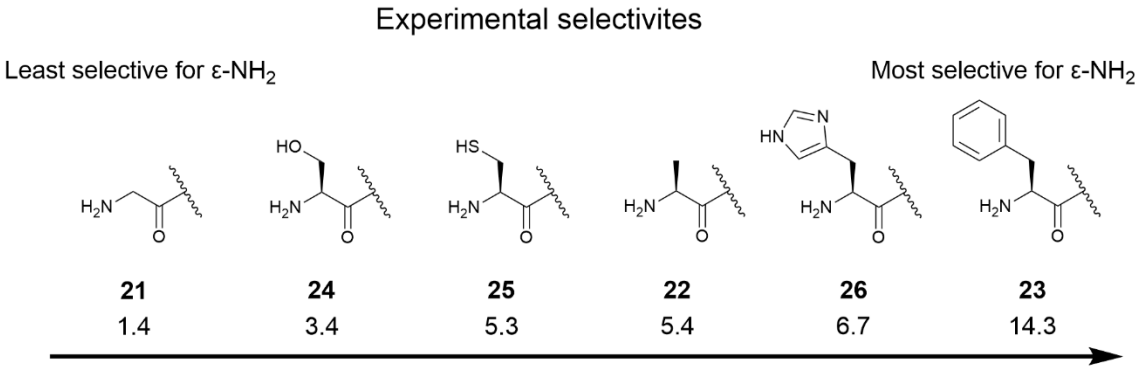


Figure 6.1 Observed selectivities for aqueous *N*-thiophosphorylation between lysine ε-NH_2 and α-NH_2 of various different *N*-terminal amino acids.

In every case, the selectivity was found to be greater than 1.0, meaning aqueous *N*-thiophosphorylation of unprotected peptides is selective for reaction at ε-NH_2 over α-NH_2. Although several amino acids were chosen for the *N*-termini, further study may be conducted using either arginine or tryptophan as *N*-terminal amino acids in order to obtain a clearer picture of how the size/

nucleophilicity of the *N*-terminal amino acid impacts thiophosphorylation selectivity. Further study into whether the nature of the residues adjacent to the target lysine may impact thiophosphorylation should also be conducted.

S-alkylation of the anionic sulfur moiety introduced *via* thiophosphorylation of *N*-terminal phenylalanine tetrapeptide **23a** was conducted using 2-bromomethyl(naphthalene) to facilitate isolation of *N*-terminal phenylalanine ε-*N*-thiophosphoramidate **23b** as *S*-naphthyl ε-*N*-thiophosphoramidate **23d** (Scheme 3.6 in section 3.4.3.1). While, in the interest of time *S*-alkylation was not explored any further at this stage, additional experimental work may be conducted to explore the scope of site-specific functionalisation, such as PEGylation and lipidation, at the anionic sulfur of the thiophosphoramidate group in competition with an unprotected cysteine residue.

6.3. Summary of investigation of alternative conditions for aqueous *N*-thiophosphorylation of peptide substrates

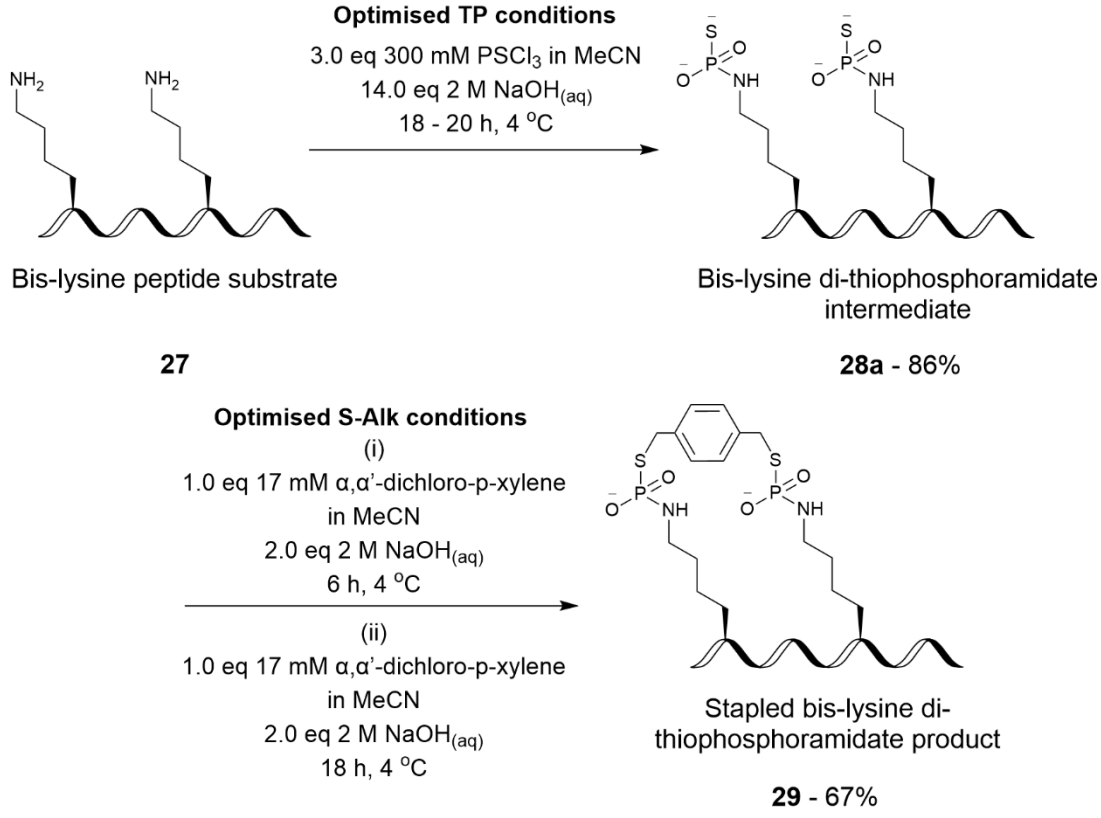
After establishing the selectivity of the aqueous *N*-thiophosphorylation reaction with $PSCl_3$ and $NaOH_{(aq)}$, we chose to investigate whether alternative, milder thiophosphorylation conditions could be used to obtain thiophosphoramidates more efficiently.

Thiophosphorylations were trialed on α-*N*-dansyl tetrapeptide **9** using each of the following three combinations of thiophosphorylating agent and base individually: potassium thiophosphorodichloridate ($KPSOCl_2$) and $NaOH_{(aq)}$ (section 4.1.2); $KPSOCl_2$ and triethylammonium bicarbonate (TEAB) solution (section 4.2.1); $PSCl_3$ and TEAB solution (section 4.2.2). It was found over the course of these experiments that none of the combinations tried were able to reproduce the conversions seen when using $PSCl_3$ and $NaOH_{(aq)}$, so investigations into the use of alternative thiophosphorylation methods were concluded.

Further work can, however, be proposed for the investigation into different conditions for aqueous *N*-thiophosphorylation of peptide/ protein substrates. Despite TEAB being chosen as the source of aqueous base for these trials because of its volatility, it was shown to be an inappropriate choice due to the large excess necessary to obtain 'optimal' results (section 4.2). Future experimentation in this area should consist of trials involving a variety of inexpensive, off-the-shelf aqueous buffers known to have a buffering range between pH 10.5 to 12.

6.4. Summary of di-thiophosphorylation and stapling of α-*N*-dansyl-labelled heptamer 27 *via bis*-*S*-alkylation using a dihalogenated linker

The final part of this project was to implement the aqueous *N*-thiophosphorylation methodology as the first step in a one-pot, two-step peptide stapling procedure. Through some method optimisation (sections 6.2 and 6.3), we were able to obtain ca. 67% conversion from α-*N*-dansyl-labelled heptamer **27** to α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** *via* α-*N*-dansyl-labelled di-thiophosphoramidate intermediate **28a**. Following two reverse phase chromatography purifications (section 5.5.2), we were able to purify stapled peptide **29** to 82% homogeneity (based on LC-UV vis analysis at 325 nm) and obtain a corroborative 2D NMR spectrum to confirm the synthesis of a di-thiophosphoramidate-based staple between two lysine residues (Figure 5.25). Scheme 6.2 provides an overview of the optimised stapling procedure.



Scheme 6.2 Optimised conditions for the one-pot, two-step stapling of α-*N*-dansyl-labelled heptamer **27** to give α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** *via* α-*N*-dansyl-labelled di-thiophosphoramidate **28a**.

Further optimisation for the purification of the stapled peptide is required, perhaps through the use of ion exchange or ion pairing chromatography rather than reverse phase HPLC due to the highly polar nature of the stapled peptide.

Further work for this particular stapled peptide should involve circular dichroism (CD) to see whether the α-helicity of the peptide has changed following stapling. Purification optimisation is required to obtain samples of stapled peptide at appropriate purities for *in vitro* cell culture testing. *In vitro* cell culture tests should be carried out in order to learn whether this staple design aids in cell penetration of peptides.

To investigate the limits of the novel stapling methodology, a variety of *bis*-lysine peptides, containing amino acids covering a wide range of nucleophilicities and reactivities, should be designed and synthesised for stapling. This variety of *bis*-lysine-containing peptides should also have the lysine residues positioned at differing intervals (e.g. some at *i, i+4*, and some at *i, i+7*) to aid in staple length optimisation for α-helix stabilisation. A range of bifunctional linkers should also be used during staple optimisation trials, as linker identity can impact the α-helix stabilisation afforded by a staple.[138]

7. General Experimental Procedures

All air or moisture sensitive reactions were carried out under a nitrogen atmosphere.

Liquid chromatography-mass spectrometry (LC-MS) experiments were carried out using either a single quadrupole detector (SQD) mass spectrometer and an Acquity ultra performance liquid chromatography (UPLC) system (Waters Ltd, UK), or a quadrupole time-of-flight (QToF) mass spectrometer and an Acquity UPLC system (Waters Ltd, UK) unless otherwise specified.

For SQD LC-MS experiments, the column was an Acquity UPLC BEH C_{18} column (2.1 mm × 50 mm with 1.7 µm particle size). The mobile phase was 0.1% formic acid solution (v/v): acetonitrile gradient with a flow rate of 0.6 mL min^{-1}. Gradient details were as follows: 5% acetonitrile from 0 – 0.2 min, increase up to 95% acetonitrile from 0.2 – 4 min, held at 95% acetonitrile from 4 – 4.5 min, and reduced back to 5% acetonitrile from 4.5 – 5 min.

For QToF LC-MS experiments, the column was an Acquity UPLC BEH C_{18} column (2.1 mm × 100 mm with 1.7 µm particle size). The mobile phase was 0.1% formic acid solution (v/v): acetonitrile gradient with a flow rate of 0.6 mL min^{-1}. Gradient details were as follows: 0-99% acetonitrile from 0 – 5 min, held at 99% acetonitrile from 5 – 6 min, reduced back to 0% acetonitrile from 6 – 6.1 min, and held at 0% acetonitrile from 6.1 – 7 min.

Standard 1D ^{1}H, ^{13}C and ^{31}P Nuclear Magnetic Resonance (NMR) experiments were run on a Bruker Neo-400 spectrometer with operating frequencies of 400.20 MHz (^{1}H), 100.63 MHz (^{13}C) and 162.00 MHz (^{31}P) at 25 °C.

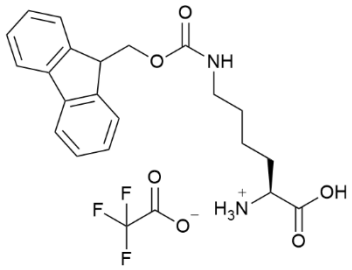
The 2D ^{1}H-^{31}P HMBC NMR experiment conducted in Chapter 5 were run on a Varian DD2-500 spectrometer with operating frequencies of 499.53 MHz (^{1}H) and 202.21 MHz (^{31}P).

For the purification of compounds, preparative chromatography was used. Flash column chromatography (normal and reverse phase) was performed on a Teledyne CombiFlash NextGen 100 system. High-performance liquid chromatography (HPLC) was performed on an Agilent 1260 Infinity II system. Column and solvent details are stated alongside synthetic procedures where relevant.

7.1. Procedures from optimisation of *N*-thiophosphorylation

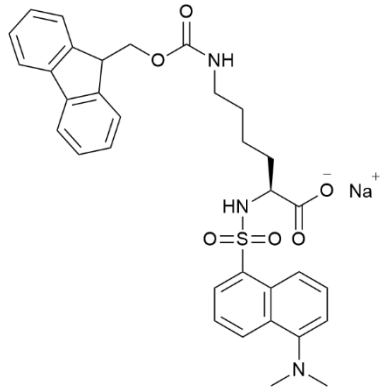
7.1.1. Synthesis of α-*N*-dansyl-labelled lysine model **4**

ε-*N*-fluorenylmethyloxycarbonyl-L-lysine 2



α-*N*-*tert*-butyloxycarbonyl-ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **1** (2.00 g, 4.26 mmol, 1.0 eq) and trifluoroacetic acid (1.31 mL; 17.07 mmol, 4.0 eq) were added to a round-bottom flask and stirred for 15 min. Reaction mixture was co-evaporated with MeCN (3 × 8 mL) and diethyl ether (10 mL) under reduced pressure. The sticky white residue was used for further experimentation without purification. δ_H (400 MHz, DMSO-d_6) 1.37 (4H, m, γ-CH_2 and δ-CH_2), 1.70 (2H, m, β-CH_2), 2.96 (2H, q, *J* 5.7, ε-CH_2), 3.42 (1H, t, *J* 5.7, α-C*H*), 4.20 (1H, t, *J* 7.1, C*H*-CH_2), 4.29 (2H, d, *J* 7.1, CH-CH_2), 7.33 (2H, t, *J* 7.4, C4'-*H* and C9'-*H*), 7.41 (2H, t, *J* 8.5, C3'-*H* and C10'-*H*), 7.68 (2H, d, *J* 7.6, C2'-*H* and C11'-*H*), 7.89 (2H, d, *J* 7.6, C5'-*H* and C8'-*H*). δ_c (100 MHz, DMSO-d_6) 22.58 (γ-*C*H_2), 29.57 (β-*C*H_2), 30.82 (δ-*C*H_2), 40.51 (ε-*C*H_2), 47.20 (*C*H-CH_2), 53.84 (α-*C*H), 65.68 (CH-*C*H_2), 120.58 (*C*5'-H and *C*8'-H), 125.62 (*C*2'-H and *C*11'-H), 127.52 (*C*3'-H and *C*10'-H), 128.07 (*C*4'-H and *C*9'-H), 141.19 (*C*6' and *C*7'), 144.38 (*C*1' and *C*12'), 156.58 (O-*C*(=O)-NH), 171.75 (-*C*(=O)-OH). *m/z* 369.335 ($[M+H]^+$, 2.57 min).

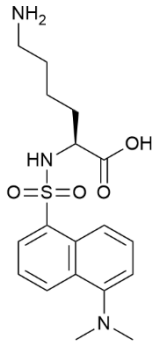
α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-fluorenylmethyloxycarbonyl-L-lysine 3



Optimised procedure:

MeCN (60 mL) was added to the TFA salt of ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **2** (1.96 g, 4.19 mmol, 1.0 eq). $NaOH_{(aq)}$ (1.1 M, 13.4 mL, 14.7 mmol, 3.5 eq) was added to dissolve **2**. 5-(dimethylamino)naphthalene-1-sulfonyl chloride (0.99 g, 3.66 mmol, 0.87 eq) was dissolved in MeCN (7 mL) and added to the solution of **2**. If precipitation of materials was observed, further $NaOH_{(aq)}$ (1.1 M) was added in 100.0 µL increments until the precipitate re-dissolved. The reaction mixture was stirred for 80 – 120 min at room temperature and the solvents were removed under reduced pressure. The bright yellow residue was dissolved in ethyl acetate (200 mL) and washed with 0.1 N citric acid solution (3 × 150 mL). The organic layer was dried over $MgSO_4$, filtered, and evaporated under reduced pressure. The crude product (**3**) was purified by flash chromatography using a Teledyne CombiFlash NextGen 100 system (hexane/ ethyl acetate gradient) using a 40 g HP Silica Gold column to give dansylated product **3** (1.35 g, 62%). δ_H (400 MHz, MeOD) 0.99 (2H, m, γ-CH_2), 1.12 (2H, m, β-CH_2), 1.55 (2H, quin, *J* 6.4, δ-CH_2), 2.70 (2H, t, *J* 6.7, ε-CH_2), 2.83 (6H, s, C6''-N-(CH_3)$_2$), 3.70 (1H, m, α-C*H*), 4.18 (1H, t, *J* 6.7, C*H*-CH_2), 4.34 (2H, d, *J* 6.7, CH-CH_2), 7.21 (1H, d, *J* 7.9, C5''-*H*), 7.31 (2H, t, *J* 7.3, C4'-*H* and C8'-*H*), 7.39 (2H, t, *J* 7.9, C3'-*H* and C10'-*H*), 7.54 (2H, app. q, *J* 6.2, C4''-*H* and C9''-*H*), 7.64 (2H, d, *J* 8.1, C2'-*H* and C11'-*H*), 7.80 (2H, d, *J* 7.9, C'-*H* and C8'-*H*), 8.22 (1H, d, *J* 7.3, C10''-*H*), 8.41 (1H, d, *J* 8.5, C3''-*H*), 8.52 (1H, d, *J* 8.5, C8''-*H*). δ_c (100 MHz, MeOD) 23.15 (γ-*C*H_2), 29.40 (β-*C*H_2), 32.79 (δ-*C*H_2), 40.96 (ε-*C*H_2), 45.52 (2 × N-*C*H_3), 48.21 (*C*H-CH_2), 56.53 (α-*C*H), 67.15 (CH-*C*H_2), 116.15 (*C*5''-H), 120.64 (*C*5'-H and *C*8'-H), 123.92 (*C*3''-H), 125.84 (*C*2'-H and *C*11'-H), 127.85 (*C*3'-H and *C*10'-H), 128.47 (*C*4'-H and *C*9'-H), 128.74 (*C*4''-H and *C*9''-H), 130.05 (*C*8''-H), 130.74 (*C*7''), 130.84 (*C*2''), 131.03 (*C*10''-H), 136.84 (*C*6''), 142.32 (*C*6' and *C*7'), 145.05 (*C*1' and *C*12'), 152.73 (*C*1''-SO_2-), 158.37 (O-*C*(=O)-NH), 174.92 (-*C*(=O)-OH). *m/z* 602.406 ($[M+H]^+$, 3.15 min).

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-L-lysine 4



Optimised pyrrolidine-based deprotection and purification method:

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **3** (0.39 g, 0.65 mmol, 1.0 eq) was dissolved in MeCN (3 mL). Pyrrolidine (0.30 mL, 3.65 mmol, 5.6 eq) was added and the reaction mixture was stirred overnight. The mixture was then evaporated under reduced pressure. Crude **4** was dissolved in triethylammonium acetate buffer (pH 4.7, ca. 2 mL) and purified through two reverse phase flash chromatography runs (triethylammonium acetate buffer/ MeCN – gradient) using a 15.5 g C_{18} Aq Gold column on a Teledyne CombiFlash NextGen 100 system. The purified product was lyophilised and re-dissolved twice in water (2 mL), followed by a final lyophilisation to remove residual triethylammonium acetate to give deprotected product **4** (15.7 mg, 6%). δ_H (400 MHz, D_2O) 0.51 (2H, m, γ-C*H*$_2$), 0.74 (2H, m, β-C*H*$_2$), 1.27 (2H, m, δ-C*H*$_2$), 1.77 (2H, m, ε-C*H*$_2$), 2.73 (6H, s, C6''-N-(C*H*$_3$)$_2$), 3.21 (1H, m, α-C*H*), 7.24 (1H, d, *J* 7.6, C5''-*H*), 7.52 (2H, t, *J* 6.9, C4''-*H* and C9''-*H*), 8.07 (1H, d, *J* 7.6, C10''-*H*), 8.25 (1H, d, *J* 8.6, C3''-*H*), 8.43 (1H, d, *J* 8.6, C8''-*H*). δ_C (100 MHz, D_2O) 22.31 (γ-*C*H$_2$), 30.86 (β-*C*H$_2$), 33.44 (δ-*C*H$_2$), 40.14 (ε-*C*H$_2$), 44.98 (2 × N-*C*6''-H$_3$), 59.59 (α-*C*H), 115.50 (*C*5''-H), 121.15 (*C*3''-H), 123.96 (*C*4''-H), 127.49 (*C*9''-H), 128.04 (*C*8''-H), 128.58 (*C*10''-H), 128.83 (*C*7''), 129.59 (*C*2''), 137.44 (*C*6''), 150.05 (*C*1''), 182.87 (-*C*(=O)-OH). *m/z* 380.326 ([M+H]$^+$, 1.34 min).

Final piperidine-based deprotection and purification method:

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **3** (0.23 g, 0.37 mmol, 1.0 eq) was dissolved in MeCN (1 mL). Piperidine (0.22 mL, 2.25 mmol, 6.07 eq) was added and the reaction mixture was stirred overnight. The mixture was evaporated under reduced pressure. Crude deprotected product **4** was dissolved in triethylammonium acetate buffer (pH 4.7, ca. 2 mL), filtered to remove remaining solid, and purified through three reverse phase flash chromatography runs (triethylammonium acetate buffer/ MeCN – gradient) using a 15.5 g C_{18} Aq Gold column on a Teledyne CombiFlash NextGen 100 system. The purified product was lyophilised and re-dissolved

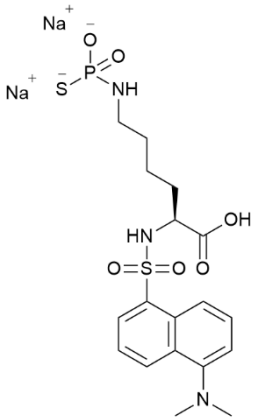
twice in water (2 mL), followed by a final lyophilisation to remove residual triethylammonium acetate to give deprotected product **4** (16.1 mg, 11%).

Precipitation purification method:

Crude α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-L-lysine **4** (1.0 eq; 0.116 mmol; 44.0 mg) was dissolved in HCl solution in MeOH (1.25 M; 26.9 eq; 2.5 mL; 3.12 mmol). Diethyl ether (9 mL) was added to the HCl in MeOH solution, and a pale-yellow precipitate immediately formed. The mixture was centrifuged (8 min, 10,000 rpm). Liquids were carefully removed, and the precipitate was washed twice with diethyl ether (7 mL). The precipitate was dissolved in MeOH (0.5 mL), and the precipitation steps were repeated 4 times. An aliquot from the MeOH solution was submitted for LC-MS analysis after each precipitation.

7.1.2. Optimisation of aqueous *N*-thiophosphorylation on α-*N*-dansyl-L-lysine **4**

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophospho-L-lysine 7



General procedure adapted from Trmčić *et al.* (2011).[152]

An aqueous solution of α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-L-lysine **4** (12.5 µL; 2.5 µmol; 1.0 eq) was added to a 2 mL glass HPLC vial with NaOH solution (2 M volumetric standard; varied volumes) and water (varied volumes to make total aqueous volume ca. 25 µL). $PSCl_3$ in MeCN solution (300 mM; varied volumes) was added and the mixture was stirred for 30 min at room temperature. δ_H (400 MHz, D_2O) 0.86 (2H, m, γ-CH_2), 1.08 (2H, m, δ-CH_2), 1.37 (2H, q, *J* 5.4, β-CH_2), 2.42 (2H, t, *J* 7.3, ε-CH_2), 2.81 (6H, s, C6''-N-(CH_3)$_2$), 3.29 (1H, t, *J* 6.4, α-C*H*), 7.32 (1H, d, *J* 8.3, C5''-*H*), 7.57 (2H, app. q, *J* 8.3, C4''-*H* and C9''-*H*), 8.08 (1H, d, *J* 7.8, C10''-*H*), 8.29 (1H, d, *J* 8.2, C3''-*H*), 8.48 (1H, d, *J* 9.1, C8''-*H*). δ_P (162 MHz, D_2O) 43.58 (app. t, $^3J_{P\text{-}H}$ 8.1). *m/z* 476.318 Da ($[M+H]^+$, 1.16 min).

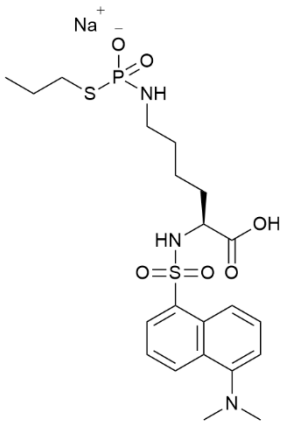
Table 7.1 Experimental details for *N*-thiophosphorylation series on lysine model **4** using 1.0 eq $PSCl_3$. The conversions to thiophosphoramidate **7** were determined from the relevant peaks seen in LC-MS total absorbance chromatographs.

300 mM $PSCl_3$/ eq	300 mM $PSCl_3$/ µL	2 M NaOH/ eq	2 M NaOH/ µL	Water/ µL	Conversion to 7/ %
1.0	8.3	6.0	7.5	5.0	82
1.0	8.3	6.5	8.1	4.4	84
1.0	8.3	7.0	8.7	3.8	87
1.0	8.3	7.5	9.4	3.1	92
1.0	8.3	8.0	10.0	2.5	88

Table 7.2 Experimental details for *N*-thiophosphorylation series on lysine model **4** using 1.2 eq $PSCl_3$. The conversions to thiophosphoramidate **7** were determined from the relevant peaks seen in LC-MS total absorbance chromatographs.

300 mM $PSCl_3$/ eq	300 mM $PSCl_3$/ µL	2 M NaOH/ eq	2 M NaOH/ µL	Water/ µL	Conversion to 7/ %
1.2	10.0	6.0	7.5	5.0	70
1.2	10.0	6.5	8.1	4.4	86
1.2	10.0	7.0	8.7	3.8	94
1.2	10.0	7.5	9.4	3.1	94
1.2	10.0	8.0	10.0	2.5	97

_S_-propyl α-_N_-5(dimethylamino)naphthalene-1-sulfonyl-ε-_N_-thiophospho-L-lysine 8



Scaled up N-thiophosphorylation:

An aqueous solution of α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-L-lysine **4** (2.1 mL; 0.275 mmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 960.0 µL; 1.925 mmol; 7.0 eq) were added to a round-bottomed flask. $PSCl_3$ in MeCN solution (300 mM; 920.0 µL; 0.276 mmol; 1.0 eq) was added and the mixture was stirred for 30 min at room temperature.

S-Alkylation:

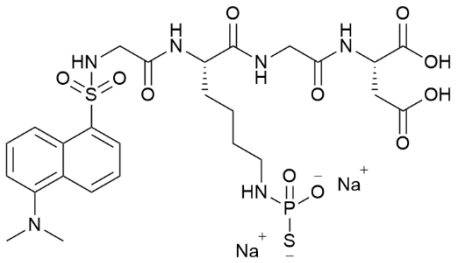
A solution of 1-iodopropane in MeCN solution (300 mM; 6.42 mL; 1.925 mmol; 7.0 eq) was added to the crude *N*-thiophosphorylation mixture and the solution mixture was stirred for 24 h at room

temperature. Crude product **8** was purified by reverse phase flash chromatography using a Teledyne CombiFlash NextGen 100 system (mobile phase: water/ MeOH gradient; column: 15.5 g HP C_{18}Aq Gold column). δ_H (400 MHz, D_2O) 0.54 (2H, m, γ-C*H*$_2$), 0.79-0.91 (5H, m, C3'-*H*$_2$ and δ-C*H*$_2$), 1.32 (2H, m, β-C*H*$_2$), 1.45 (2H, sext, *J* 7.18, C2'-*H*$_2$), 2.42 (2H, m, ε-C*H*$_2$), 2.80 (6H, s, C6''-N-(C*H*$_3$)$_2$), 3.30 (1H, m, α-C*H*), 3.37 (2H, m, C1'-*H*$_2$), 7.30 (1H, m, C5''-*H*), 7.62 (2H, m, C4''-*H* and C9''-*H*), 8.17 (1H, d, *J* 7.19, C10''-*H*), 8.27 (1H, d, *J* 9.02, C3''-*H*), 8.39 (1H, d, *J* 8.25, C8''-*H*). δ_P (162 MHz, D_2O) 25.23 (app. quint, $^3J_{P\text{-}H}$ 9.57). *m/z* 518.156 ([M+H]$^+$, 1.68 min).

Note: C' denotes a carbon in the propyl moiety, C'' denotes a carbon in the dansyl moiety. Carbon numbering on Fmoc and dansyl moieties can be found in the SI.

7.1.3. Optimisation of aqueous *N*-thiophosphorylation on α-*N*-dansyl-labelled tetrapeptide model **9**

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophosphopeptide 10



General procedure:

An aqueous solution of α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-labelled tetrapeptide **9** (10.0 µL; 0.095 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; varying volumes) were added to a 2 mL glass HPLC vial. $PSCl_3$ in MeCN solution (300 mM; varying volumes) was added and the mixture was stirred for 30 min at room temperature. *m/z* 705.350 Da ([M+H]$^+$, 1.71 min).

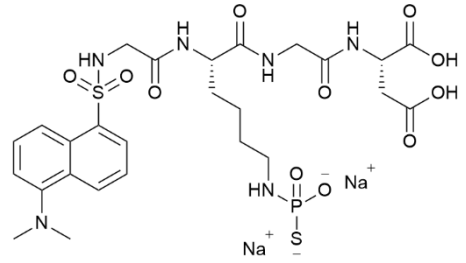
Table 7.3 Experimental details for *N*-thiophosphorylation series on model tetrapeptide **9**. The conversions to thiophosphoramidate **7** were determined from the relevant peaks seen in LC-MS total absorbance chromatographs.

300 mM $PSCl_3$/ eq	300 mM $PSCl_3$/ µL	2 M NaOH/ eq	2 M NaOH/ µL	Conversion to 10/ %
1.0	0.30	7.5	0.35	82
2.0	0.65	14.7	0.70	89
7.5	2.40	35.7	1.70	95

7.2. Procedures from selectivity investigation

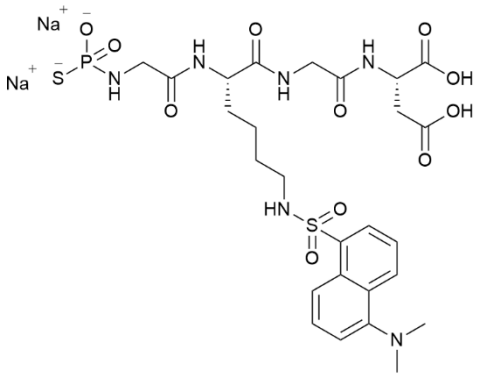
7.2.1. Aqueous *N*-thiophosphorylations of α-*N*-dansyl-labelled tetrapeptide **9** and ε-*N*-dansyl-labelled tetrapeptide **11**

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophosphopeptide 10



An aqueous solution of α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-labelled tetrapeptide **9** (10.0 µL; 0.095 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 3.34 µL; 6.66 µmol; 70.0 eq) were added to a 2 mL glass HPLC vial. $PSCl_3$ in MeCN solution (300 mM; 3.17 µL; 0.95 µmol; 10.0 eq) was added and the mixture was stirred for 30 min at room temperature. Conversion of 99% to **10** from **9** was seen by LC-UV-vis analysis. *m/z* 705.424 Da ($[M+H]^+$, 1.89 min).

ε-*N*-5(dimethylamino)naphthalene-1-sulfonyl-α-*N*-thiophosphopeptide 12



An aqueous solution of ε-*N*-5(dimethylamino)naphthalene-1-sulfonyl-labelled tetrapeptide **11** (10.0 µL; 0.04 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 1.40 µL; 2.80 µmol; 70.0 eq) were added to a 2 mL glass HPLC vial. $PSCl_3$ in MeCN solution (300 mM; 1.33 µL; 0.40 µmol; 10.0 eq) was added and the mixture was stirred for 30 min at room temperature. Conversion of 42% to **12** from **11** was seen by LC-UV-vis analysis. *m/z* 705.424 Da ($[M+H]^+$, 2.07 min).

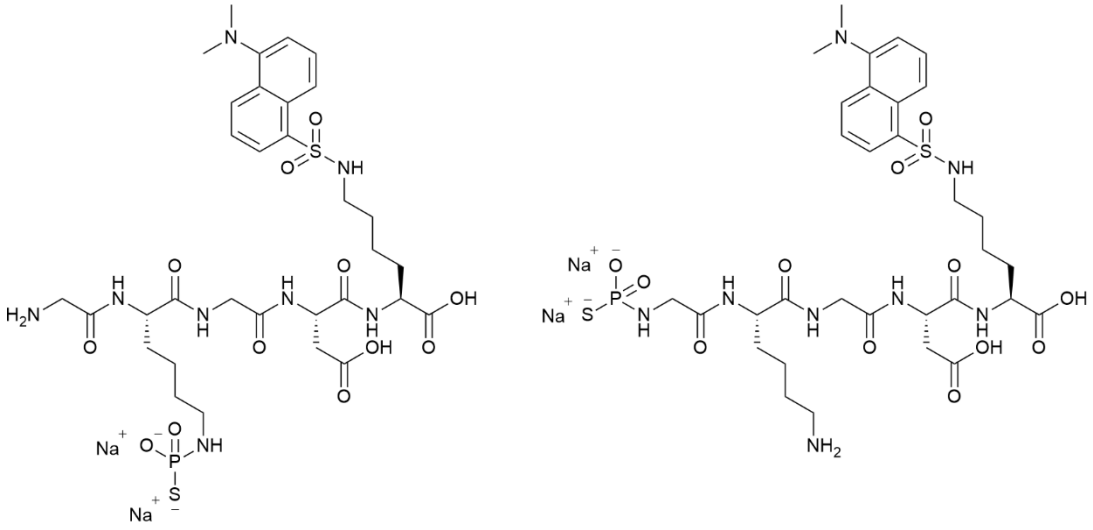
7.2.2. *N*-thiophosphorylation of dansyl-labelled pentapeptide model **13**

Tandem MS experimental details:

The tandem MS experiments to determine thiophosphoramidate positioning were carried out using a triple quadrupole detector (TQD) mass spectrometer and an Acquity UPLC system (Waters Ltd, UK). The column was an Acquity UPLC BEH C_{18} column (2.1 mm × 50 mm with 1.7 µm particle size). The mobile phase was 0.1% formic acid solution (v/v): MeCN gradient with a flow rate of 0.6 mL min^{-1}. The gradient details were as follows: 5% MeCN from 0 – 0.2 min, increase up to 95% MeCN from 0.2 – 4 min, held at 95% MeCN from 4 – 4.5 min, and reduced back to 5% MeCN from 4.5 – 5 min.

For the tandem MS experiments run under basic conditions, the 0.1% formic acid solution (v/v) was replaced with a 10 mM solution of ammonium bicarbonate adjusted to pH 8 with acetic acid.

ε-*N'*-5(dimethylamino)naphthalene-1-sulfonyl ε-*N*-thiophosphopentapeptide 14a and ε-*N'*-5(dimethylamino)naphthalene-1-sulfonyl α-*N*-thiophosphopentapeptide 14b

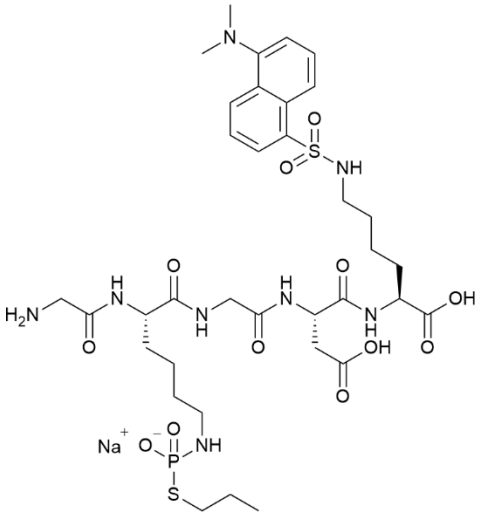


14a **14b**

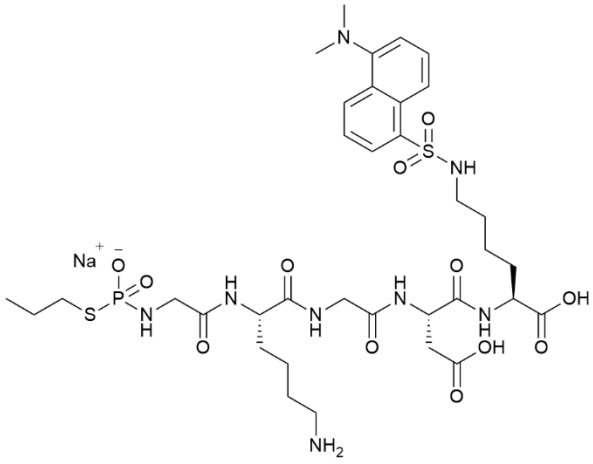
An aqueous solution of ε-*N'*-5(dimethylamino)naphthalene-1-sulfonyl-labelled pentapeptide **13** (10.0 µL; 0.04 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 1.40 µL; 2.80 µmol; 70.0 eq) were added to a 2 mL glass HPLC vial. $PSCl_3$ in MeCN solution (300 mM; 1.33 µL; 0.40 µmol; 10.0 eq) was added and the mixture was stirred for 30 min at room temperature. Conversion to **14a** and/or **14b** from **13** noted as 100% by LC-UV-vis analysis. *m/z* 833.551 Da ($[M+H]^+$, 2.04 min).

7.2.3. *S*-alkylation of dansyl-labelled thiophosphopentapeptides **14a** and **14b**

ε-*N'*-5(dimethylamino)naphthalene-1-sulfonyl *S*-propyl ε-*N*-thiophosphopentapeptide 15a and ε-*N'*-5(dimethylamino)naphthalene-1-sulfonyl *S*-propyl α-*N*-thiophosphopentapeptide 15b



15a



15b

N-thiophosphorylation procedure:

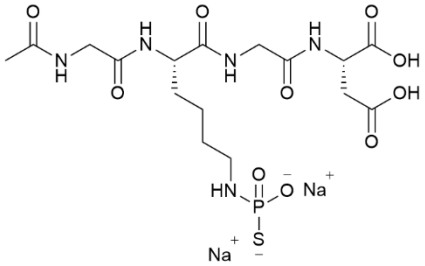
An aqueous solution of ε-*N'*-5(dimethylamino)naphthalene-1-sulfonyl-labelled pentapeptide **13** (22.5 µL; 0.09 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 3.15 µL; 6.30 µmol; 70.0 eq) were added to a 2 mL glass HPLC vial. $PSCl_3$ in MeCN solution (300 mM; 3.0 µL; 0.90 µmol; 10.0 eq) was added and the mixture was stirred for 30 min at room temperature.

S-alkylation procedure:

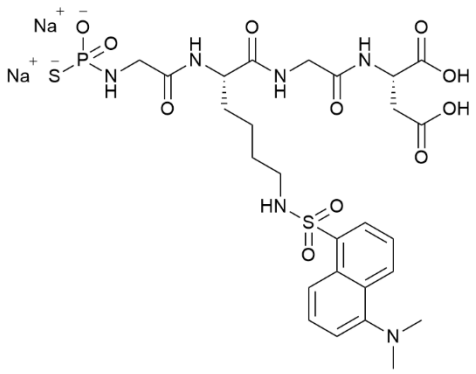
A solution of 1-iodopropane in MeCN (500 mM; 4.20 µL; 2.1 µmol; 23.3 eq) was added to the *N*-thiophosphorylation mixture and the mixture was stirred for 24 h at room temperature. Conversion to **15a** and/or **15b** from **13** noted as 47% by LC-UV-vis analysis. *m/z* 875.423 Da ($[M+H]^+$, 2.30 min).

7.2.4. Competition reactions between α-*N*-acetyl-protected tetrapeptide **16** and ε-*N*-dansyl-labelled tetrapeptide **11**

ε-*N*-5(dimethylamino)naphthalene-1-sulfonyl-labelled *S*-naphthyl-α-*N*-thiophosphoramidate 19 and α-*N*-acetyl-protected *S*-naphthyl-ε-*N*-thiophosphoramidate 18



17



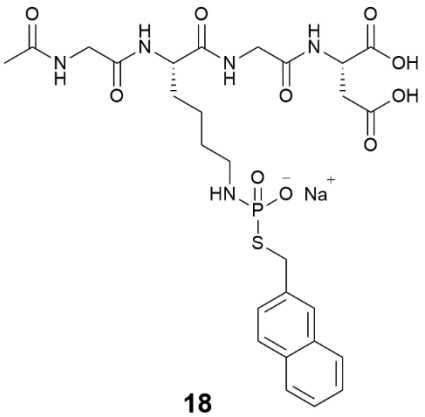
12

<u>*Room temperature procedure – LC-MS data obtained using SQD mass spectrometer:*</u>

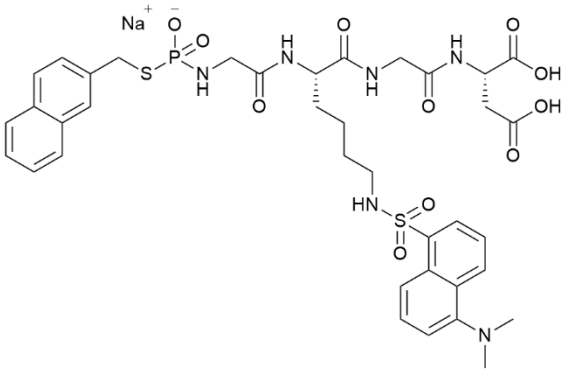
N-thiophosphorylation:

Aqueous solutions of ε-*N*-5(dimethylamino)naphthalene-1-sulfonyl-labelled tetrapeptide **11** (20.0 µL; 0.08 µmol; 1.0 eq) and α-*N*-acetyl protected tetrapeptide **16** (20.0 µL; 0.08 µmol; 1.0 eq) were added to a 2 mL glass HPLC vial with NaOH solution (2 M volumetric standard; 0.80 µL; 1.60 µmol; 20.0 eq). $PSCl_3$ in MeCN solution (300 mM; 0.27 µL; 0.08 µmol; 1.0 eq) was added and the mixture was stirred for 30 min at room temperature. Conversion to **12** from **11** was noted as 3% by LC-UV-vis analysis. *m/z* 705.317 Da ($[M+H]^+$, 1.95 min).

S-alkylation:



18



19

2-(bromomethyl)naphthalene in MeCN solution (300 mM; 0.4 µL; 0.12 µmol; 2.0 eq [after aliquot removal from *N*-thiophosphorylation mixture for LC-MS analysis]) was added to the *N*-

thiophosphorylation mixture and the resulting *S*-alkylation mixture was stirred for 15 min at room temperature. **18** *m/z* 654.350 Da ($[M+H]^+$, 1.42 min). **19** *m/z* 845.409 Da ($[M+H]^+$, 2.01 min).

Cold room (4 °C) procedure – LC-MS data obtained using QToF mass spectrometer:

N-thiophosphorylation:

Aqueous solutions of ε-*N*-5(dimethylamino)naphthalene-1-sulfonyl-labelled tetrapeptide **11** (10.0 µL; 0.04 µmol; 1.0 eq) and α-*N*-acetyl protected tetrapeptide **16** (10.0 µL; 0.04 µmol; 1.0 eq) were added to a 2 mL glass HPLC vial with NaOH solution (2 M volumetric standard; 0.40 µL; 0.80 µmol; 20.0 eq). $PSCl_3$ in MeCN solution (300 mM; 0.13 µL; 0.04 µmol; 1.0 eq) was added and the mixture was stirred for 24 h at 4 °C. Conversion to **12** from **11** was noted as 2% by LC-UV-vis analysis. *m/z* 705.186 Da ($[M+H]^+$, 2.20 min).

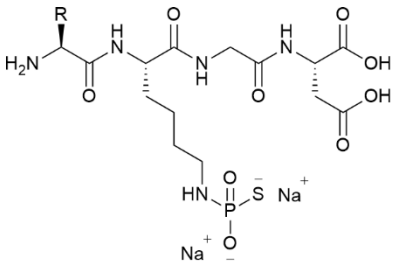
S-alkylation:

A solution of 2-(bromomethyl)naphthalene in MeCN (300 mM; 0.25 µL; 0.073 µmol; 2.0 eq [after aliquot removal from *N*-thiophosphorylation mixture for LC-MS analysis]) was added to the *N*-thiophosphorylation mixture and the resulting *S*-alkylation mixture was stirred for 3 h at 4 °C. **18** *m/z* 654.200 Da ($[M+H]^+$, 2.45 min). **19** *m/z* 845.243 Da ($[M+H]^+$, 3.13 min).

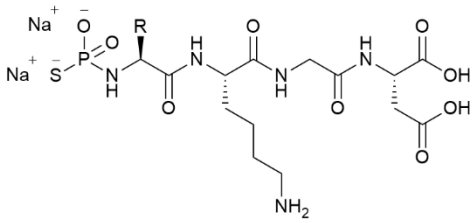
7.2.5. Aqueous *N*-thiophosphorylations of unprotected peptides **21-26a** for a ^{31}P NMR-based assay

ε-*N*-thiophosphoramidates 21-26b and α-*N*-thiophosphoramidates 21-26c

The general *N*-thiophosphorylation procedure is outlined, with the experimental details displayed in Table 8.4. The ^{31}P NMR spectroscopy results are displayed in Table 8.5.



21-26b
ε-*N*-thiophosphoramidate



21-26c
α-*N*-thiophosphoramidate

General procedure:

An aqueous solution of unprotected tetrapeptide **21-26a** (1.0 eq), NaOH solution (2 M volumetric standard; 7.0 eq), and $PSCl_3$ in MeCN solution (300 mM; 1.0 eq) were added to a 2 mL glass HPLC vial and stirred for 18.5 h at 4 °C.

Table 7.4 Experimental details of the *N*-thiophosphorylation series performed on unprotected tetrapeptides **21-26**.

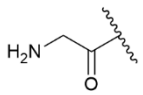
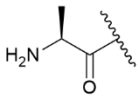
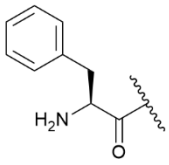
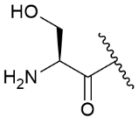
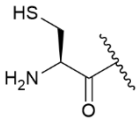
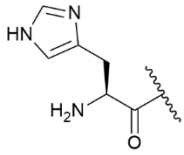
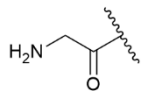
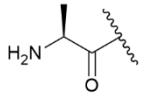
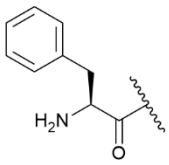
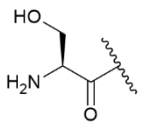
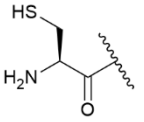
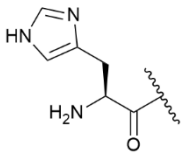
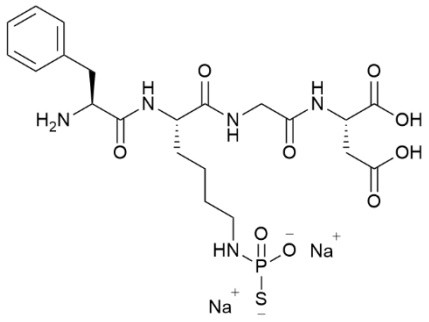
***N*-terminal amino acid**	**Tetrapeptide**		**$NaOH_{(aq)}$**		**$PSCl_3$**	
	µL	**µmol**	**µL**	**µmol**	**µL**	**µmol**
 **21**	133.3	13.3	46.5	93.1	44.3	13.3
 **22**	128.5	12.8	44.8	89.6	42.7	12.8
 **23**	107.5	10.7	37.5	74.9	35.7	10.7
 **24**	123.4	12.3	43.1	86.1	41.0	12.3
 **25**	118.7	11.8	41.3	82.6	39.3	11.8
 **26**	109.8	11.0	38.5	77.0	36.7	11.0

Table 7.5 ^{31}P NMR chemical shift details and reaction selectivity of the *N*-thiophosphorylation series performed on unprotected tetrapeptides **21-26**.

***N*-terminal amino acid**	**δ_P/ ppm**		***N*-thiophosphorylation selectivity (ε-NH_2/ α-NH_2)**
	Thiophosphorylation at ε-NH_2 (integration; multiplicity; $^3J_{P-H}$)	**Thiophosphorylation at α-NH_2 (integration; multiplicity; $^3J_{P-H}$)**	
 **21**	43.52 (1.0; t; 7.66 Hz)	42.57 (0.69; t; 11.97 Hz)	1.4
 **22**	43.52 (1.0; t; 7.48 Hz)	41.76 (0.18; d; 12.89 Hz)	5.6
 **23**	43.54 (1.0; t; 7.62 Hz)	41.64 (0.07; d; 12.67 Hz)	14.3
 **24**	43.67 (1.0; t; 6.28 Hz)	41.70 (0.29; d; 12.11 Hz)	3.4
 **25**	43.69 (1.0; br s; n/a)	41.52 (0.19; d; 15.33 Hz)	5.3
 **26**	43.90 (1.0; t; 4.95 Hz)	41.56 (0.15; m; n/a)	6.7

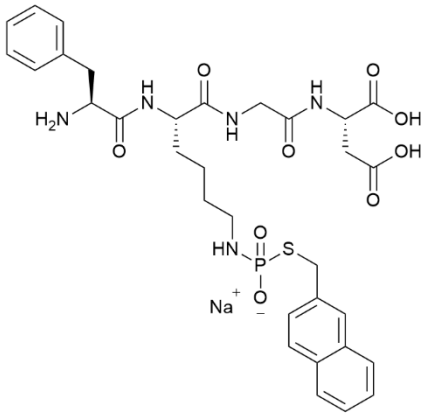
7.2.5.1. The synthesis and isolation of *S*-naphthyl ε-*N*-thiophosphoramidate **23d**

***N*-terminal phenylalanine ε-*N*-thiophosphoramidate 23b**



An aqueous solution of *N*-terminal phenylalanine tetrapeptide **23a** (160.0 µL; 16.0 µmol; 1.0 eq), NaOH solution (2 M volumetric standard; 56.0 µL; 112.0 µmol; 7.0 eq), and $PSCl_3$ in MeCN solution (300 mM; 40.0 µL; 12.0 µmol; 0.75 eq) were added to a 2 mL glass HPLC vial and stirred for 30 min at room temperature. δ_P (162 MHz, $CDCl_3$) 43.39 (br. s).

***N*-terminal phenylalanine *S*-naphthyl ε-*N*-thiophosphoramidate 23d**



A solution of 2-(bromomethyl)naphthalene in MeCN (300 mM; 53.3 µL; 16.0 µmol; 1.0 eq) was added to the *N*-thiophosphorylation mixture. The resulting *S*-alkylation mixture was stirred for 15 min at room temperature. The crude *S*-naphthyl ε-*N*-thiophosphoramidate **23d** mixture was purified through reverse phase chromatography following *S*-alkylation. δ_P (162 MHz, D_2O) 24.02 (app. quint, $^3J_{P\text{-}H}$ 9.67). *m/z* 702.465 ($[M+H]^+$, 1.57 min).

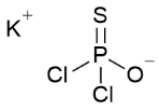
Reverse phase chromatography details:

The column used was a 15.5 g HP C_{18} Gold column. The solvent system was a water/ MeCN gradient. The gradient details were as follows: 0% MeCN from 0 – 0.3 min, increased up to 50% MeCN from 0.3 – 7.0 min, immediately increased to 100% MeCN at 7.0 min, held at 100% MeCN from 7.0 – 8.6 min, immediately reduced back down to 50% MeCN at 8.6 min, and held at 50% MeCN from 8.6 – 10.0 min.

Final collected product was shown to be 84% homogenous **23d** by LC-UV-vis analysis, but 100% homogenous **23d** by ^{31}P NMR spectroscopy.

7.3. Procedures from investigation of alternative conditions for *N*-thiophosphorylation of peptide substrates

7.3.1. Synthesis of potassium thiophosphorodichloridate ($KPSOCl_2$)



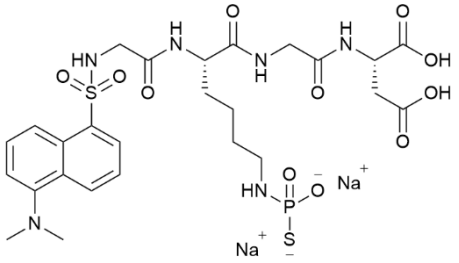
Procedure followed from Delley *et al.* (2012).[153]

A solution of $PSCl_3$ (0.85 mL; 8.33 mmol; 1.0 eq) in dry MeCN (25 mL) was added to a round-bottomed flask containing oven dried potassium bicarbonate (1.67 g; 16.7 mmol; 2.0 eq), and the mixture was stirred vigorously under nitrogen for 12 h. The KCl precipitate formed was filtered off under vacuum. The supernatant remaining was the $KPSOCl_2$ in MeCN solution (330 mM) and was stored in a freezer in a sealable vial within a polythene bag containing anhydrous magnesium sulfate. δ_P (162 MHz, CD_3CN) 42.09 (s).

7.3.2. Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** using $KPSOCl_2$ and $NaOH_{(aq)}$

7.3.2.1. Optimisation of aqueous *N*-thiophosphorylation with $KPSOCl_2$ and $NaOH_{(aq)}$

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophosphopeptide 10



10

General procedure:

An aqueous solution of α-*N*-dansyl-labelled tetrapeptide **9** (4 mM; 40.0 µL; 0.16 µmol; 1.0 eq), NaOH solution (2 M volumetric standard; varying volumes), and $KPSOCl_2$ in MeCN solution (330 mM; varying volumes) were added to a 2 mL glass HPLC vial and stirred for 30 min at room temperature. *m/z* 705.377 ($[M+H]^+$, 1.13 min).

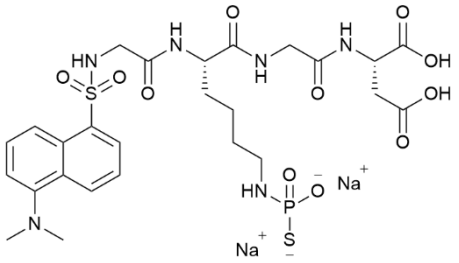
Table 7.6 Experimental details of the optimisation of aqueous *N*-thiophosphorylation performed on α-*N*-dansyl-labelled tetrapeptide **9** using $KPSOCl_2$ and $NaOH_{(aq)}$. The conversions to thiophosphoramidate **10** were determined from the relevant peaks seen in LC-MS total absorbance chromatographs.

$KPSOCl_2$		**$NaOH_{(aq)}$**		**Conversion to 10/ %**
Equivalents	**Volume/ µL**	**Equivalents**	**Volume/ µL**	
1.2	0.60	7.0	0.60	19
2.5	1.21	14.6	1.17	51
5.0	2.42	29.2	2.33	81
7.5	3.67	43.7	3.49	95
10.0	4.84	58.3	4.67	97

7.3.3. Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** using triethylammonium bicarbonate (TEAB) solution and $KPSOCl_2$

7.3.3.1. Optimisation of the number of equivalents of TEAB solution (pH 10.5) for aqueous *N*-thiophosphorylation with $KPSOCl_2$

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophosphopeptide 10



10

General procedure:

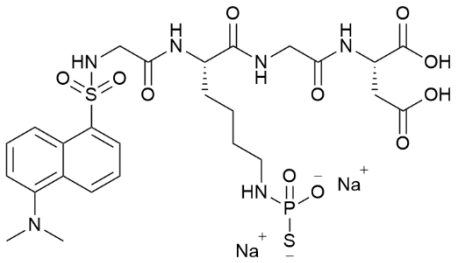
An aqueous solution of α-*N*-dansyl-labelled tetrapeptide **9** (4 mM; 20.0 µL; 0.08 µmol; 1.0 eq), TEAB solution (1 M; pH 10.5; varying volumes), and $KPSOCl_2$ in MeCN solution (330 mM; 1.82 µL; 0.6 µmol; 7.5 eq) were added to a 2 mL glass HPLC vial and stirred for 30 min at room temperature. *m/z* 705.352 ($[M+H]^+$, 1.14 min).

Table 7.7 Experimental details of the optimisation of aqueous *N*-thiophosphorylation performed on α-*N*-dansyl-labelled tetrapeptide **9** using 7.5 eq $KPSOCl_2$ and 1 M TEAB solution (pH 10.5). The conversions to thiophosphoramidate **10** were determined from the relevant peaks seen in LC-MS total absorbance chromatographs.

1 M TEAB solution (pH 10.5)		**Conversion to 10/ %**
Equivalents	**Volume/ µL**	
43.7	3.5	2
75.0	6.0	11
100.0	8.0	14
125.0	10.0	15
150.0	12.0	18
200.0	16.0	20
250.0	20.0	19
300.0	24.0	14
350.0	28.0	13

7.3.3.2. Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** with $KPSOCl_2$ and TEAB solution (pH 10.8)

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophosphopeptide 10



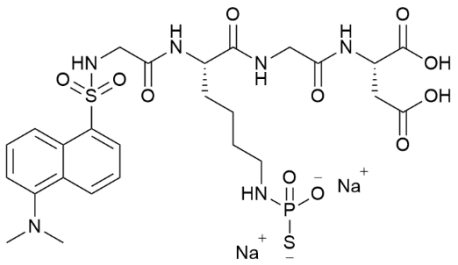
10

An aqueous solution of α-*N*-dansyl-labelled tetrapeptide **9** (4 mM; 20.0 µL; 0.08 µmol; 1.0 eq), TEAB solution (1 M; pH 10.8; 16.0 µL; 16.0 µmol; 200.0 eq), and $KPSOCl_2$ in MeCN solution (330 mM; 1.82 µL; 0.6 µmol; 7.5 eq) were added to a 2 mL glass HPLC vial and stirred for 30 min at room temperature. Conversion determined from LC-UV-vis analysis was 46%. *m/z* 705.355 ($[M+H]^+$, 1.14 min).

7.3.3.3. Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** with TEAB solution (pH 10.8) and multiple additions of $KPSOCl_2$

LC-MS DAD analysis was carried out using the QToF mass spectrometer in order to attempt to improve the baseline separation of the ε-*N*-thiophosphoramidate **10** and α-*N*-dansyl-labelled tetrapeptide **9** peaks.

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophosphopeptide 10



10

An aqueous solution of α-*N*-dansyl-labelled tetrapeptide **9** (4 mM; 50.0 µL; 0.2 µmol; 1.0 eq), TEAB solution (1 M; pH 10.8; 40.0 µL; 40.0 µmol; 200.0 eq), and $KPSOCl_2$ in MeCN solution (330 mM; 4.5 µL; 1.5 µmol; 7.5 eq) were added to a 2 mL glass HPLC vial and stirred for 30 min at room temperature. A second addition of $KPSOCl_2$ in MeCN solution (330 mM; 4.5 µL; 1.5 µmol; 7.5 eq) was added to the

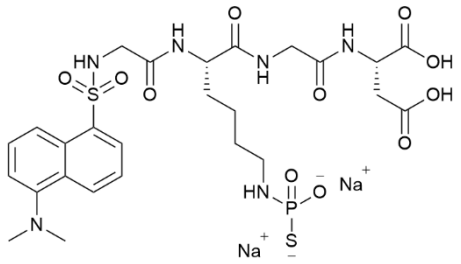
aqueous *N*-thiophosphorylation mixture and was stirred for a further 30 min at room temperature. *m/z* 705.352 ($[M+H]^+$, 1.14 min).

Table 7.8 Conversion results from α-*N*-dansyl-labelled tetrapeptide **9** to ε-*N*-thiophosphoramidate **10** after the first and second addition of 7.5 eq $KPSOCl_2$ as determined by LC-UV-vis analysis.

Addition of 7.5 eq $KPSOCl_2$	**Conversion to 10/ %**
1	48
2	74

7.3.4. Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** using TEAB solution (pH 10.8) and $PSCl_3$

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophosphopeptide 10



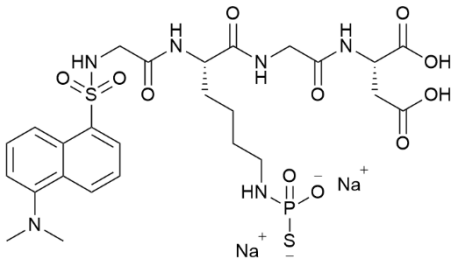
10

An aqueous solution of α-*N*-dansyl-labelled tetrapeptide **9** (4 mM; 20.0 µL; 0.08 µmol; 1.0 eq), TEAB solution (1 M; pH 10.8; 16.0 µL; 16.0 µmol; 200.0 eq), and $PSCl_3$ in MeCN solution (300 mM; 2.0 µL; 0.6 µmol; 7.5 eq) were added to a 2 mL glass HPLC vial and stirred for 30 min at room temperature. Conversion determined from LC-UV-vis analysis was 70%. *m/z* 705.355 ($[M+H]^+$, 1.14 min).

7.3.4.1. Aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled tetrapeptide **9** with TEAB solution (pH 10.8) and multiple additions of $PSCl_3$

LC-MS DAD analysis was carried out using the QToF mass spectrometer in order to attempt to improve the baseline separation of the ε-*N*-thiophosphoramidate **10** and α-*N*-dansyl-labelled tetrapeptide **9** peaks.

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl-ε-*N*-thiophosphopeptide 10



10

An aqueous solution of α-*N*-dansyl-labelled tetrapeptide **9** (4 mM; 50.0 µL; 0.2 µmol; 1.0 eq), TEAB solution (1 M; pH 10.8; 40.0 µL; 40.0 µmol; 200.0 eq), and $PSCl_3$ in MeCN solution (300 mM; 5.0 µL; 1.5 µmol; 7.5 eq) were added to a 2 mL glass HPLC vial and stirred for 30 min at room temperature. A second addition of $PSCl_3$ in MeCN solution (300 mM; 5.0 µL; 1.5 µmol; 7.5 eq) was added to the aqueous *N*-thiophosphorylation mixture and was stirred for a further 30 min at room temperature. *m/z* 705.181 ($[M+H]^+$, 2.07 min).

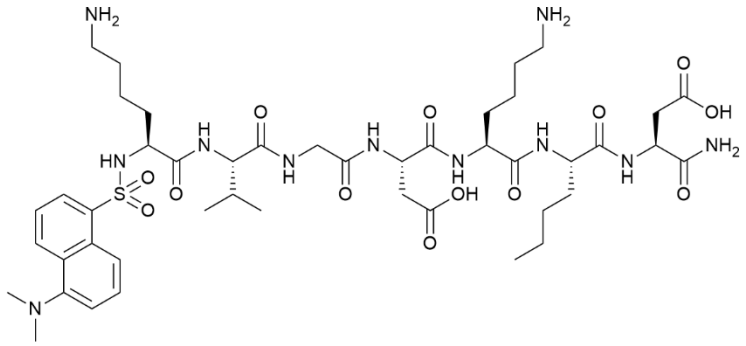
Table 7.9 Conversion results from α-*N*-dansyl-labelled tetrapeptide **9** to ε-*N*-thiophosphoramidate **10** after the first and second addition of 7.5 eq $PSCl_3$ as determined by LC-UV-vis analysis.

Addition of 7.5 eq $PSCl_3$	**Conversion to 10/ %**
1	59
2	88

7.4. Procedures from design and synthesis of *bis*-lysine heptamer for peptide stapling

7.4.1. Synthesis of α-*N*-dansyl-labelled heptamer **27**

Synthesis of α-*N*-dansyl-labelled heptamer **27** was carried out with a CEM Liberty Blue Automated Microwave Peptide Synthesiser using a standard Fmoc solid-phase peptide synthesis (SPPS) methodology.



Peptide synthesis procedure:

α-*N*-dansyl-labelled heptamer **27** was synthesised on a 0.10 mmol scale with Rink amide MBHA resin (0.284 g) in a PTFE reactor. Each amino acid was coupled using 2.5 mL Fmoc-amino acid in DMF solution (0.2 M; 5 eq), 1 mL activator solution (DIC in DMF; 1 M; 10 eq) and 0.5 mL base solution (oxyma in DMF; 1 M; 5 eq) *via* '0.10-Single Coupling (HS)' cycles. The amino acid derivatives used were Fmoc-Asp(OtBu)-OH, Fmoc-Nle-OH, Fmoc-Lys(Boc)-OH, Fmoc-Gly-OH, Fmoc-Val-OH, and α-*N*-dansyl-labelled ε-*N*-Fmoc-L-lysine **3**. Fmoc deprotections were carried out following each coupling stage with 4.5 mL 20% piperidine in DMF solution (v/v).

The resin-bound heptamer **27** was washed into a fritted syringe from the PTFE reactor with DMF and washed with DCM (3 × 4 mL). Heptamer **27** was cleaved from the resin with a TFA/TIPS/H_2O solution (95:2.5:2.5 v/v, 3 mL) by shaking for 2 h in an orbital shaker platform at 400 rpm. The cleavage solution was filtered into a falcon tube containing diethyl ether (25 mL). The suspension was centrifuged (5 min; 4,000 rpm), and solvent carefully removed. The crude peptide was washed with diethyl ether (25 mL), centrifuged (5 min; 4,000 rpm), and solvent was carefully removed. Heptamer **27** was dissolved in water and lyophilised. 202 mg of crude heptamer **27** (70% homogeneity) was obtained. *m/z* 1006.593 ($[M+H]^+$, 1.375 min).

*Purification of α-N-dansyl-labelled heptamer **27** with a Teledyne CombiFlash NextGen 100 system:*

Crude α-*N*-dansyl-labelled heptamer **27** was dissolved in water and injected on to a pre-equilibrated 5.5 g HP C_{18} Gold column. The solvent system used was a water (0.1% TFA)/ MeCN (0.1% TFA) gradient. The gradient details were as follows: 0 – 50% MeCN from 0 – 5.0 min, immediately increased to 100% MeCN at 5.0 min and held at 100% MeCN from 5.0 – 7.5 min. Immediately reduced back to 50% MeCN at 7.5 min and held at 50% MeCN until 9.0 min.

A second reverse phase chromatograph run was carried out on the material that was not totally isolated during the first run. The same column and solvent system were used, but the solvent gradient was changed. The gradient details were as follows: 0 – 10% MeCN from 0 – 5.0 min, 10-50% MeCN from 5.0 – 9.0 min, 100% MeCN from 9.0 – 11.0 min, 50% MeCN from 11.0 – 13.0 min.

All fractions from both chromatography runs consisting of homogeneous solutions of heptamer **27** were collected together. The solvents were removed, and the final collected product was shown to be > 99% **27** by LC-UV-vis analysis.

*Purification of α-N-dansyl-labelled heptamer **27** with an Agilent 1260 Infinity II system:*

Crude α-*N*-dansyl-labelled heptamer **27** was dissolved in water and injected on to an Agilent 5 Prep C_{18} column. The solvent system used was a water (0.1% TFA)/ MeCN (0.1% TFA) gradient. The gradient details were as follows: 10% MeCN from 0 – 2.0 min, 10 – 95% MeCN from 2.0 – 22.0 min, 95% MeCN from 22.0 – 27.0 min, 95 – 10% MeCN from 27.0 – 29.5 min, 10% MeCN from 29.5 – 30.0 min.

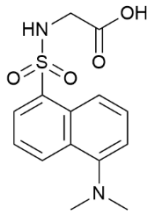
The fraction found to consist of a homogeneous solution of heptamer **27** had the solvents removed. The final collected product was shown to be > 99% **27** by LC-UV-vis analysis.

7.5. Procedures from stapling of α-*N*-dansyl-labelled heptamer 27 *via* a di-thiophosphoramidate intermediate

7.5.1. Preliminary stapling experiment on α-*N*-dansyl-labelled heptamer **27**

7.5.1.1. Synthesis of dansyl-glycine for the creation of a calibration plot

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl glycine 31



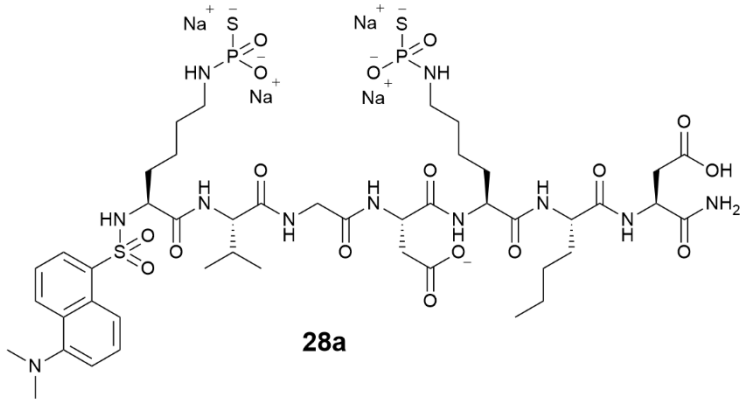
31

Procedure adapted from Cai *et al*. (2022)[169]:

Glycine (0.149 g; 1.98 mmol; 1.0 eq) and dansyl chloride (0.342 g; 1.27 mmol; 0.64 eq) were dissolved in THF (12 mL). NaOH solution (1 M volumetric standard; 2.25 mL; 2.25 mmol; 1.14 eq) was added to the THF solution. The reaction mixture was stirred for 1 h at room temperature. Solvents were then removed under vacuum. The crude dansylated product **31** was purified by flash chromatography (ethyl acetate/ methanol gradient) using a 4 g Silica column to give purified **31** (0.234 g, 60%). δ_H (400 MHz, D_2O) 3.72 (2H, s, α-C*H*$_2$), 7.78 (2H, app. q, *J* 7.58, C4''-*H* and C9''-*H*), 8.00 (1H, d, *J* 7.75, C5''-*H*), 8.21 (1H, d, *J* 7.58, C10''-*H*), 8.35 (1H, d, *J* 7.75, C3''-*H*), 8.63 (1H, d, *J* 9.09, C8''-*H*). δ_C (100 MHz, D_2O) 43.51 (α-*C*H$_2$), 119.42 (*C*5''-H), 125.49 (*C*3''-H), 125.75 (*C*4''-H), 126.57 (*C*9''-H), 126.84 (*C*8''-H), 127.97 (*C*10''-H), 128.70 (*C*7''), 130.33 (*C*2''), 135.01 (*C*6''), 138.35 (*C*1''), 172.62 (-*C*OOH). *m/z* 309.133 ($[M+H]^+$, 1.62 min).

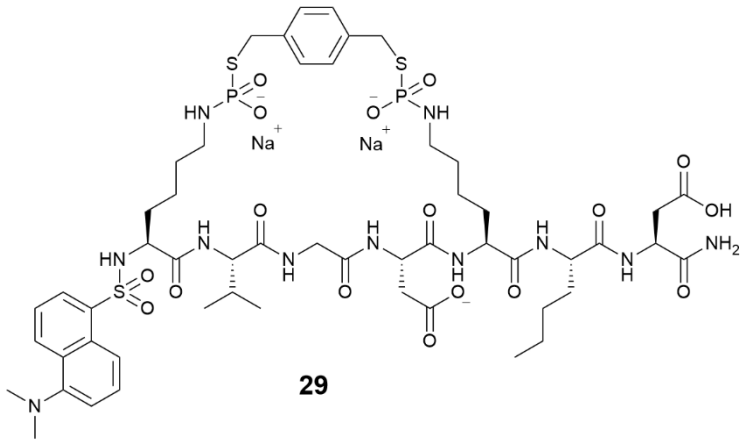
7.5.1.2. Aqueous *N*-thiophosphorylation and *S*-alkylation of α-*N*-dansyl-labelled heptamer **27**

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl di-ε-*N*-thiophosphoramidate 28a



An aqueous solution of crude α-*N*-dansyl-labelled heptamer **27** (29 mM; 83.4 µL; 2.42 µmol; 1.0 eq), NaOH solution (2 M volumetric standard; 13.3 µL; 26.60 µmol; 11.0 eq) and $PSCl_3$ in MeCN solution (300 mM; 16.1 µL; 4.84 µmol; 2.0 eq) were added to a 2 mL glass HPLC vial and stirred overnight at 4 °C. *m/z* 1198.837 Da ($[M+H]^+$, 1.32 min).

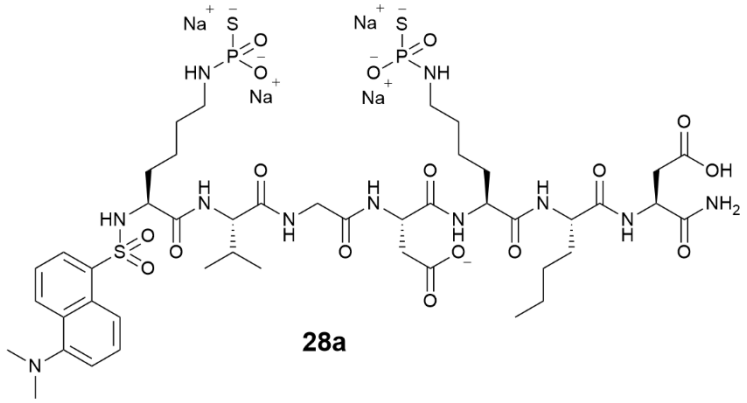
α-*N*-5(dimethylamino)naphthalene-1-sulfonyl stapled di-ε-*N*-thiophosphoramidate 29



A solution of α,α'-dibromo-*p*-xylene in DMF (10.2 mM; 237.3 µL; 2.42 µmol; 1.0 eq) was added directly to the aqueous *N*-thiophosphorylation mixture. The resultant *S*-alkylation mixture was stirred for 6 h at 4 °C. *m/z* 1300.750 Da ($[M+H]^+$, 1.548 min).

7.5.1.2.1. Varying the number of equivalents of $PSCl_3$ used for aqueous *N*-thiophosphorylation of α-*N*-dansyl-labelled heptamer **27**

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl di-ε-*N*-thiophosphoramidate 28a



An aqueous solution of crude α-*N*-dansyl-labelled heptamer **27** (14.5 mM; 50.0 µL; 0.73 µmol; 1.0 eq), NaOH solution (2 M volumetric standard; varying volumes) and $PSCl_3$ in MeCN solution (300 mM; varying volumes) were added to a 2 mL glass HPLC vial and stirred overnight at 4 °C.

Table 7.10 Experimental details of the optimisation of aqueous *N*-thiophosphorylation performed on α-*N*-dansyl-labelled heptamer **27**. The conversion to α-*N*-dansyl-labelled di-thiophosphoramidate **28a** was calculated from the LC-MS chromatographs (325 nm) obtained following *N*-thiophosphorylation.

$PSCl_3$			**$NaOH_{(aq)}$**			**Conversion to**
Eq	**Vol/ µL**	**µmol**	**Eq**	**Vol/ µL**	**µmol**	**28a/ %**
2.0	4.83	1.46	11.0	3.99	8.03	57
2.5	6.04	1.82	12.5	4.53	9.12	59
3.0	7.25	2.19	14.0	5.08	10.22	68

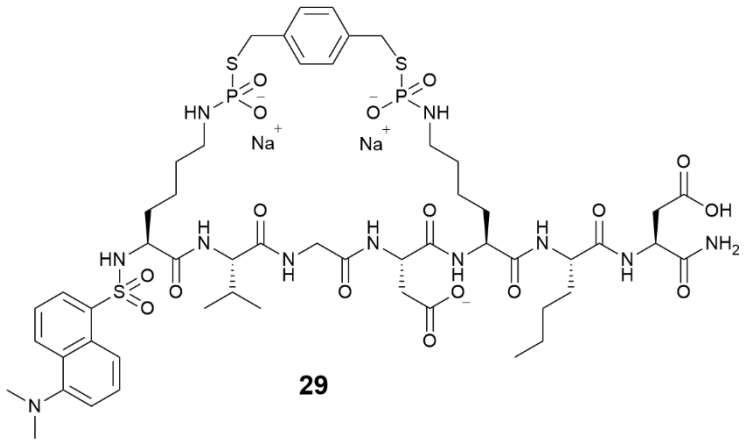
7.5.1.2.2. *S*-alkylation of di-thiophosphoramidate **28a** with a dibromo- linker versus a dichloro- linker

All LC-MS experiments conducted on samples containing α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** from this point onwards were carried out using a 10 min LC gradient.

The equipment specifications were as follows: The column was an Acquity UPLC BEH C_{18} column (2.1 mm × 100 mm with 1.7 µm particle size). The mobile phase was 0.1% formic acid solution (v/v): acetonitrile gradient with a flow rate of 0.450 mL min^{-1}.

Gradient details were as follows: 1% acetonitrile from 0.01 – 3.0 min, increase from 1 to 40% acetonitrile from 3.0 – 5.0 min (linear gradient), 40% acetonitrile from 5.0 – 6.0 min, decrease from 40 to 1% acetonitrile from 6.0 – 6.1 min, and 1% acetonitrile from 6.1 – 9.9 min.

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl stapled di-ε-*N*-thiophosphoramidate 29



29

N-thiophosphorylation – same procedure used for both stapling experiments:

An aqueous solution of crude α-*N*-dansyl-labelled heptamer **27** (12.7 mM; 39.4 µL; 0.5 µmol; 1.0 eq), NaOH solution (2 M volumetric standard; 3.5 µL; 7.0 µmol; 14.0 eq) and $PSCl_3$ in MeCN solution (300 mM; 5.0 µL; 1.5 µmol; 3.0 eq) were added to a 2 mL glass HPLC vial and stirred overnight at 4 °C. 5 min LC gradient: *m/z* 1198.533 Da ($[M+H]^+$, 1.38 min).

S-Alkylation using α,α'-dibromo-p-xylene:

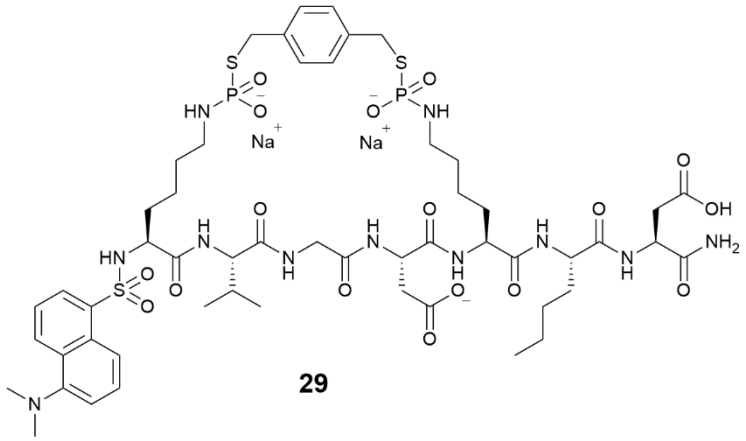
A solution of α,α'-dibromo-*p*-xylene in DMF (10.2 mM; 38.8 µL; 0.40 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 0.4 µL; 0.80 µmol; 2.0 eq) were added directly to the aqueous *N*-thiophosphorylation mixture. The resultant *S*-alkylation mixture was stirred for 6 h at 4 °C. 10 min LC gradient: *m/z* 1300.759 Da ($[M+H]^+$, 5.79 min).

S-Alkylation using α,α'-dichloro-p-xylene:

A solution of α,α'-dichloro-*p*-xylene in DMF (10.0 mM; 39.0 µL; 0.4 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 0.4 µL; 0.8 µmol; 2.0 eq) were added directly to the aqueous *N*-thiophosphorylation mixture. The resultant *S*-alkylation mixture was stirred for 6 h at 4 °C. 10 min LC gradient: *m/z* 1300.712 Da ($[M+H]^+$, 5.85 min).

7.5.2. Stapling experiments on isolated α-*N*-dansyl-labelled heptamer **27**

α-*N*-5(dimethylamino)naphthalene-1-sulfonyl stapled di-ε-*N*-thiophosphoramidate 29



29

7.5.2.1. Complication from using DMF as a co-solvent during *S*-alkylation

N-thiophosphorylation:

An aqueous solution of isolated α-*N*-dansyl-labelled heptamer **27** (11.7 mM; 42.47 µL; 0.50 µmol; 1.0 eq), NaOH solution (2 M volumetric standard; 3.48 µL; 6.96 µmol; 14.0 eq) and $PSCl_3$ in MeCN solution (300 mM; 4.97 µL; 1.49 µmol; 3.0 eq) were added to a 2 mL glass HPLC vial and stirred overnight at 4 °C. 5 min LC gradient: *m/z* 1198.623 Da ($[M+H]^+$, 1.40 min).

S-Alkylation:

A solution of α,α'-dichloro-*p*-xylene in DMF (10.0 mM; 44.8 µL; 0.45 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 0.45 µL; 0.90 µmol; 2.0 eq) were added directly to the aqueous *N*-thiophosphorylation mixture. The resultant *S*-alkylation mixture was stirred for 6 h at 4 °C. 10 min LC gradient: *m/z* 1300.725 Da ($[M+H]^+$, 6.107 min).

7.5.2.2. Using MeCN as the co-solvent during *S*-alkylation

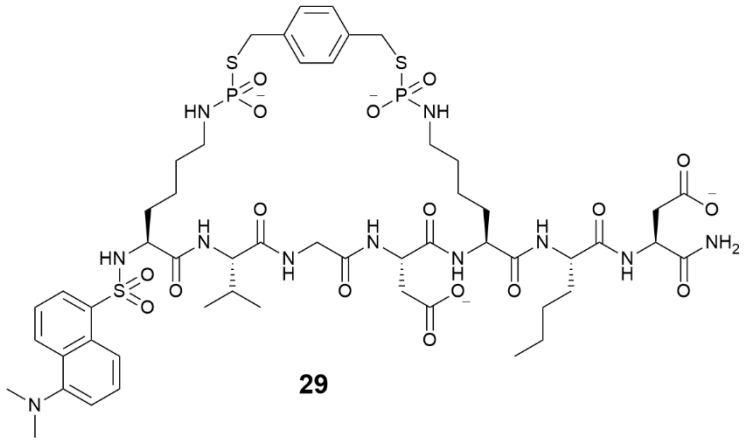
N-thiophosphorylation:

An aqueous solution of isolated α-*N*-dansyl-labelled heptamer **27** (11.7 mM; 42.47 µL; 0.50 µmol; 1.0 eq), NaOH solution (2 M volumetric standard; 3.48 µL; 6.96 µmol; 14.0 eq) and $PSCl_3$ in MeCN solution (300 mM; 4.97 µL; 1.49 µmol; 3.0 eq) were added to a 2 mL glass HPLC vial and stirred overnight at 4 °C. 5 min LC gradient: *m/z* 1198.649 Da ($[M+H]^+$, 1.335 min).

S-Alkylation:

A solution of α,α'-dichloro-*p*-xylene in MeCN (10.0 mM; 36.0 µL; 0.36 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 0.36 µL; 0.72 µmol; 2.0 eq) were added directly to the aqueous *N*-thiophosphorylation mixture. The resultant *S*-alkylation mixture was stirred for 6 h at 4 °C. 10 min LC gradient: *m/z* 1300.793 Da ($[M+H]^+$, 5.877 min).

7.5.3. Scaled up stapling and isolation of α-*N*-dansyl-labelled heptamer **27**



29

N-thiophosphorylation:

An aqueous solution of isolated α-*N*-dansyl-labelled heptamer **27** (11.7 mM; 500.0 µL; 5.85 µmol; 1.0 eq), NaOH solution (2 M volumetric standard; 40.9 µL; 81.9 µmol; 14.0 eq) and $PSCl_3$ in MeCN solution (300 mM; 58.5 µL; 17.55 µmol; 3.0 eq) were added to a 2 mL glass HPLC vial and stirred overnight at 4 °C. 5 min LC gradient: *m/z* 1198.643 Da ($[M+H]^+$, 1.400 min).

S-Alkylation:

A solution of α,α'-dichloro-*p*-xylene in MeCN (17.0 mM; 338.2 µL; 5.75 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 5.78 µL; 11.55 µmol; 2.0 eq) were added directly to the aqueous *N*-thiophosphorylation mixture. The resultant *S*-alkylation mixture was stirred for 6 h at 4 °C. A second addition of both the α,α'-dichloro-*p*-xylene in MeCN solution (17.0 mM; 334.7 µL; 5.69 µmol; 1.0 eq) and NaOH solution (2 M volumetric standard; 5.69 µL; 11.38 µmol; 2.0 eq) were added to the *S*-alkylation mixture, and the resultant mixture was stirred for a further 18 h at 4 °C. δ_P (202 MHz, D_2O, decoupled spectrum) 24.01, 23.86. ^1H-^{31}P HMBC δ_H-δ_P (D_2O) 2.53,23.85; 3.63,23.89. 10 min LC gradient: *m/z* 1300.824 Da ($[M+H]^+$, 5.755 min).

*Purification of α-N-dansyl-labelled stapled di-thiophosphoramidate **29** using an Agilent 1260 Infinity II system:*

Crude α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29** (1 mL) was injected on to an Agilent 5 Prep C_{18} column. The solvent system used was a water/MeCN gradient. The gradient details were as follows: 10% MeCN from 0 – 2.0 min, 10 – 95% MeCN from 2.0 – 22.0 min, 95% MeCN from 22.0 – 27.0 min, 95 – 10% MeCN from 27.0 – 29.5 min, 10% MeCN from 29.5 – 30.0 min. Each fraction was collected into a test tube containing 50.0 µL 2 M $NaOH_{(aq)}$ in an effort to keep **29** at a high pH to minimise hydrolysis.

As the peak containing **29** was collected as a single fraction with another peak, a second HPLC run was carried out. Prior to injection, the sample of **29** was made up to 1 mL with water and pH adjusted to ca. 8 with 0.1 M acetic acid. The HPLC run was carried out using the same column, solvent system, and gradient as the first HPLC run. The fraction containing **29** was lyophilised to remove solvents.

8. Appendix 1 – supplementary spectra

8.1. Chapter 2 compounds

8.1.1. ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **2**

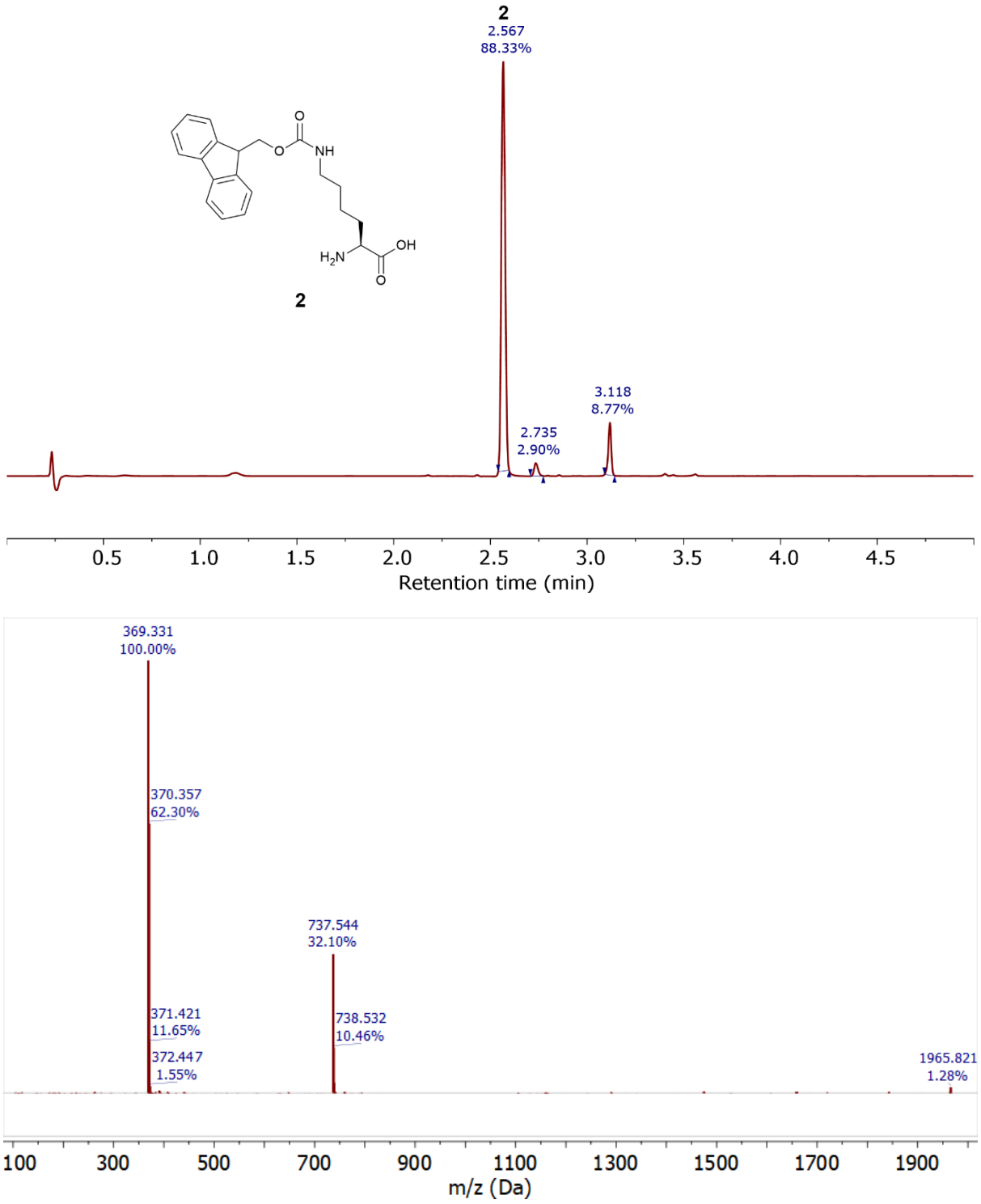


Figure 8.1 LC-MS total absorbance chromatograph and mass spectrum for peak at 2.567 min.

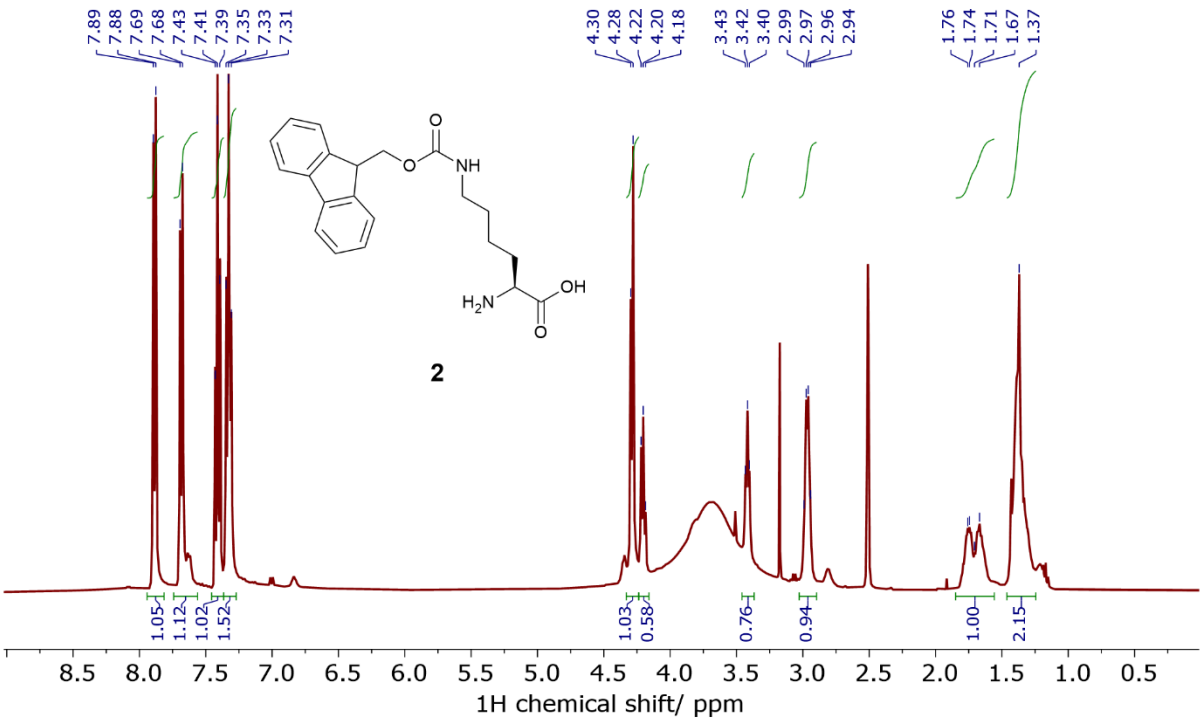


Figure 8.2 ^{1}H NMR spectrum of ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **2**.

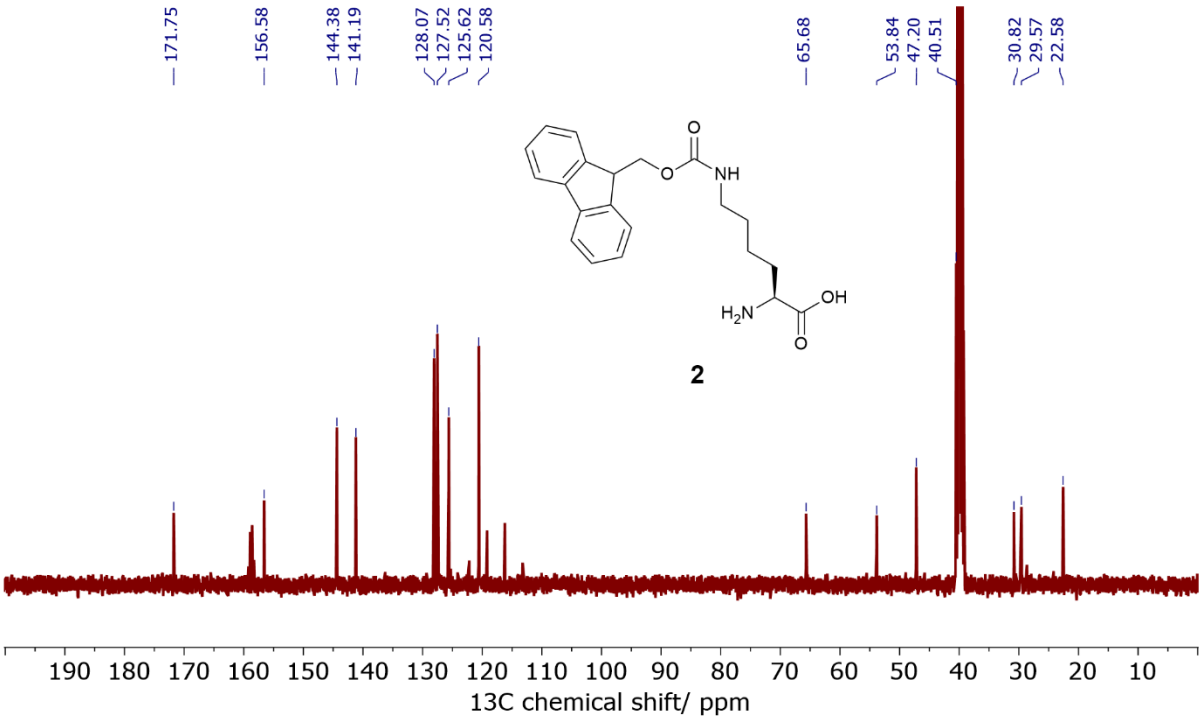


Figure 8.3 ^{13}C NMR spectrum of ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **2**.

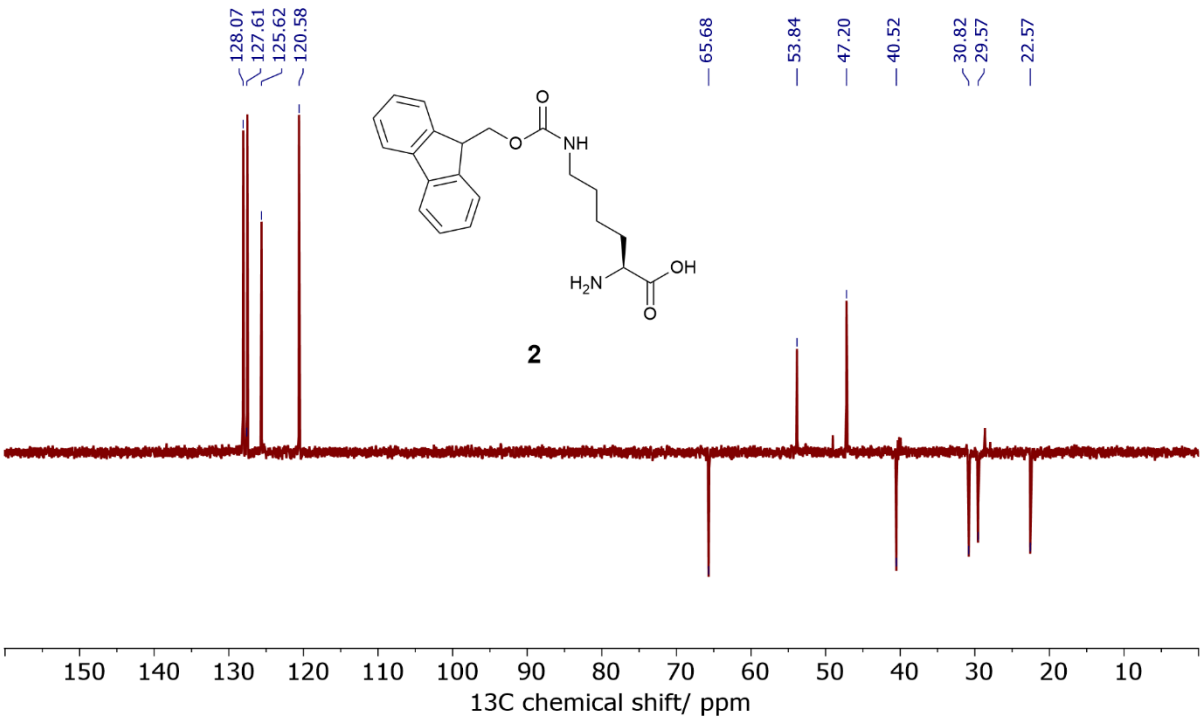


Figure 8.4 DEPT-135 NMR spectrum of ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **2**.

8.1.2. α-*N*-dansyl ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **3**

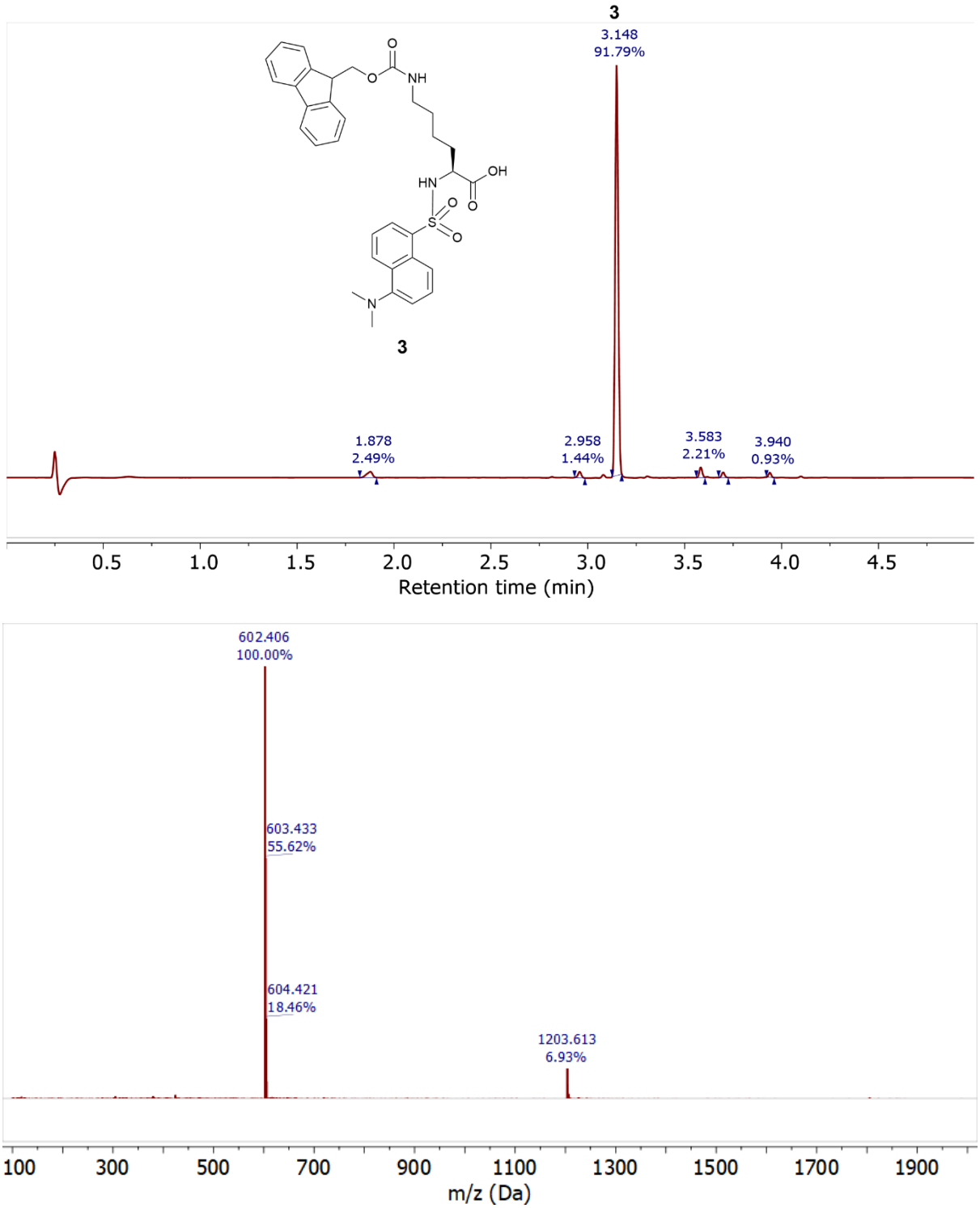


Figure 8.5 LC-MS total absorbance chromatograph and mass spectrum for peak at 3.148 min.

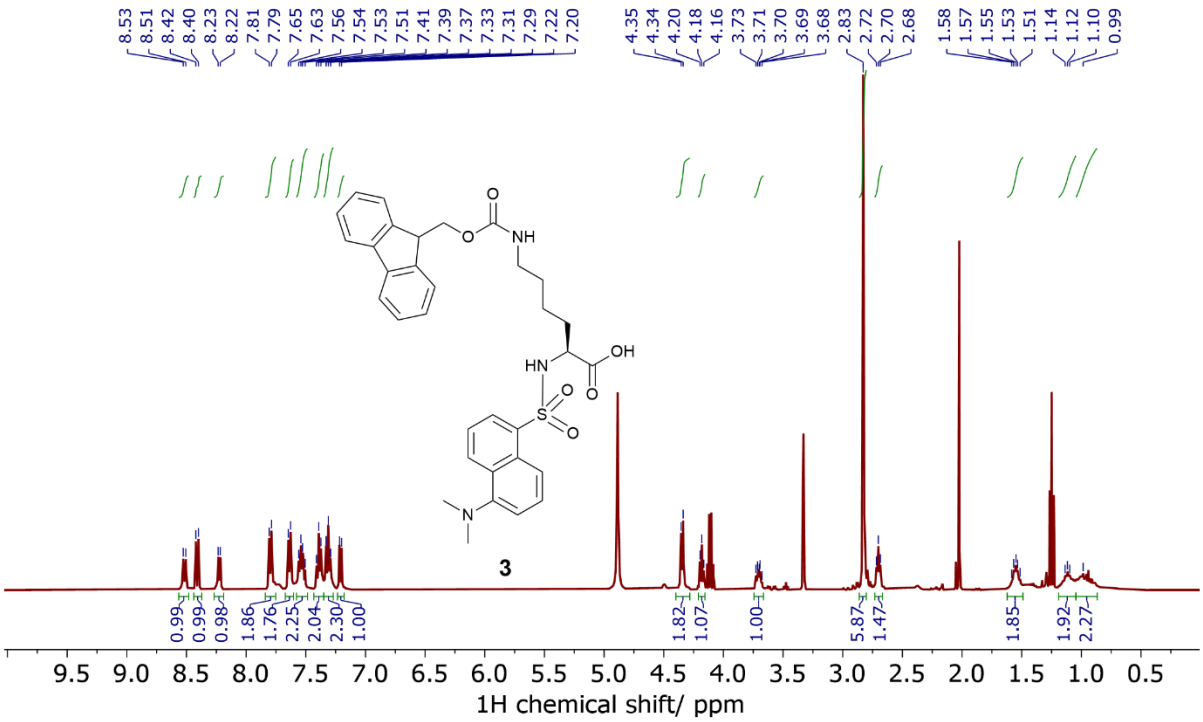


Figure 8.6 ^{1}H NMR spectrum of α-*N*-dansyl ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **3**.

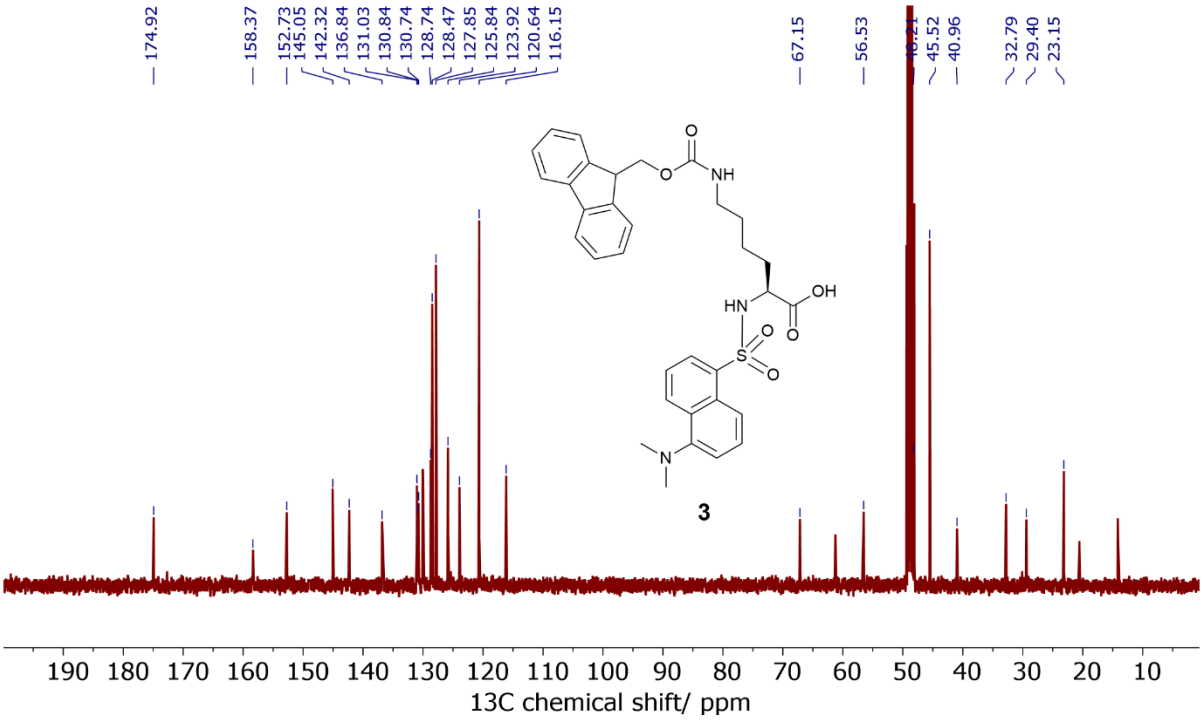


Figure 8.7 ^{13}C NMR spectrum of α-*N*-dansyl ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **3**.

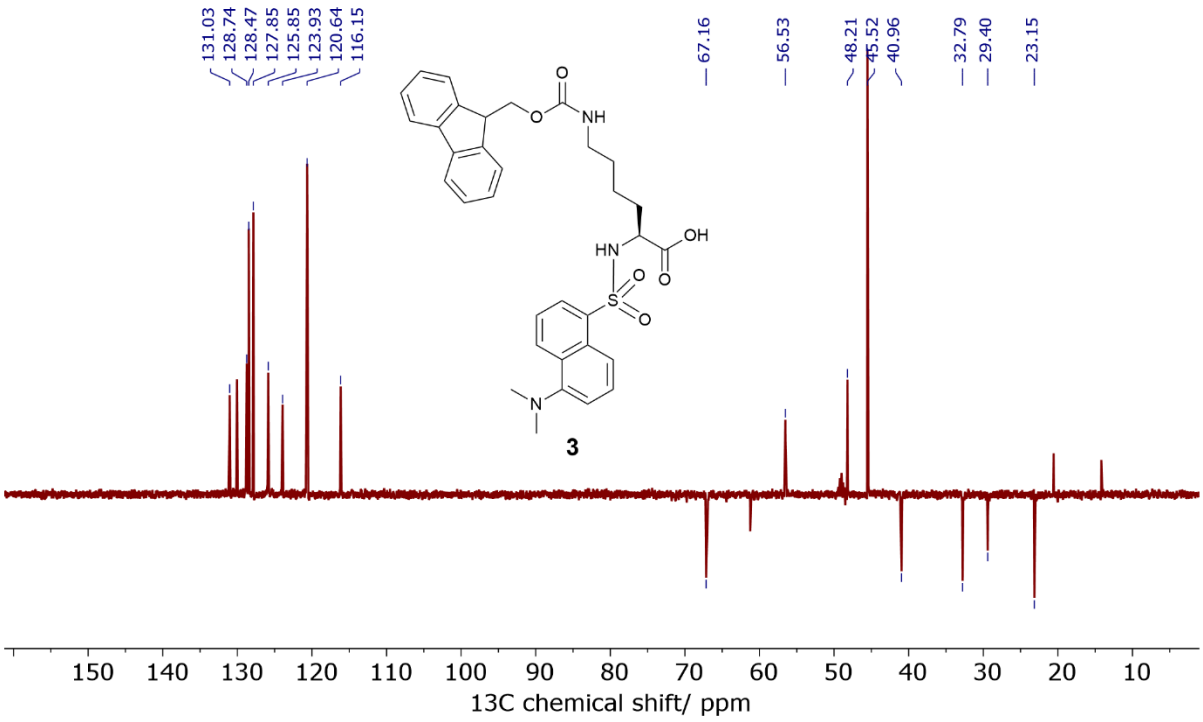


Figure 8.8 DEPT-135 NMR spectrum of α-*N*-dansyl ε-*N*-fluorenylmethyloxycarbonyl-L-lysine **3**.

8.1.3. α-*N*-dansyl-L-lysine **4**

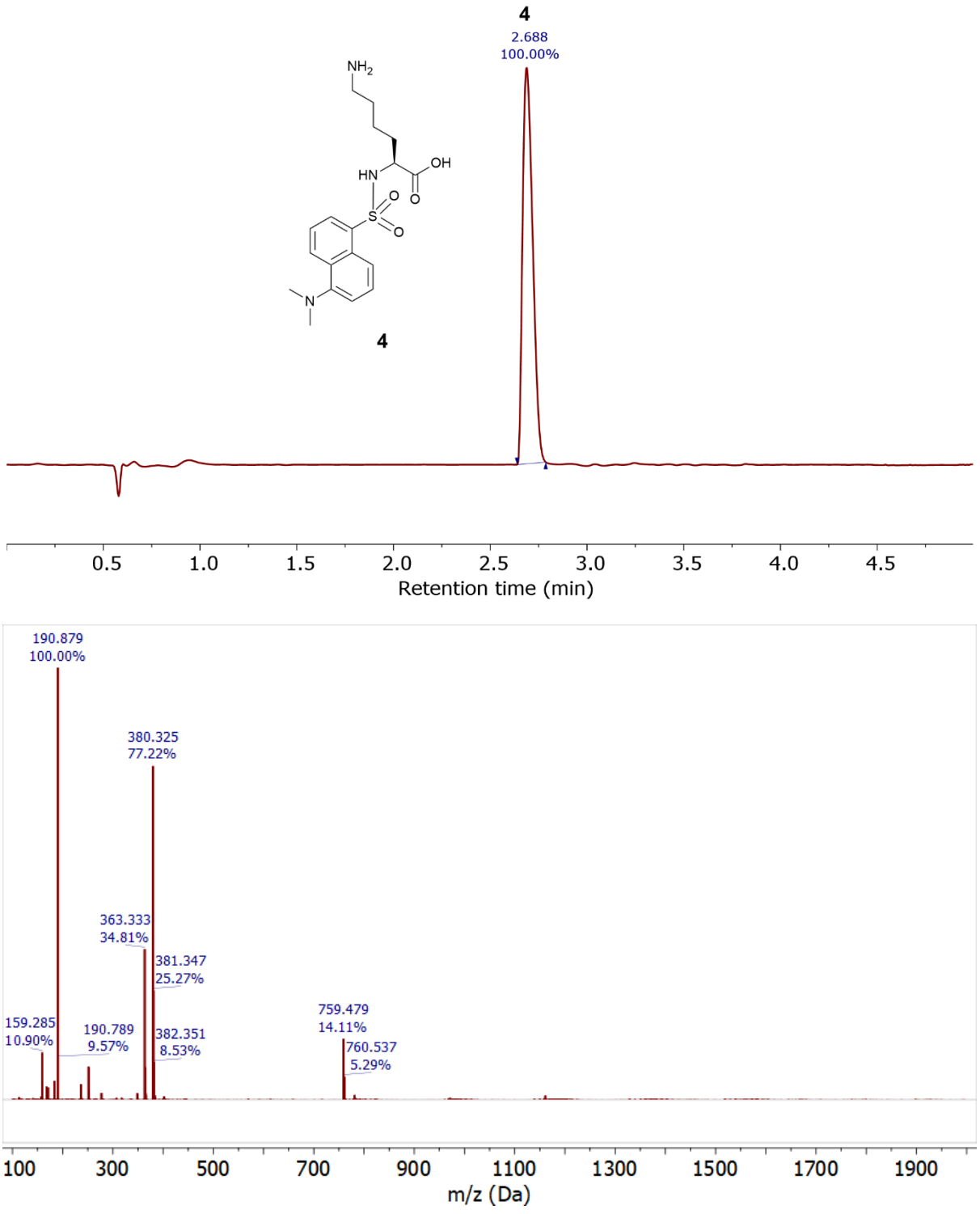


Figure 8.9 LC-MS total absorbance chromatograph and mass spectrum for peak at 2.688 min.

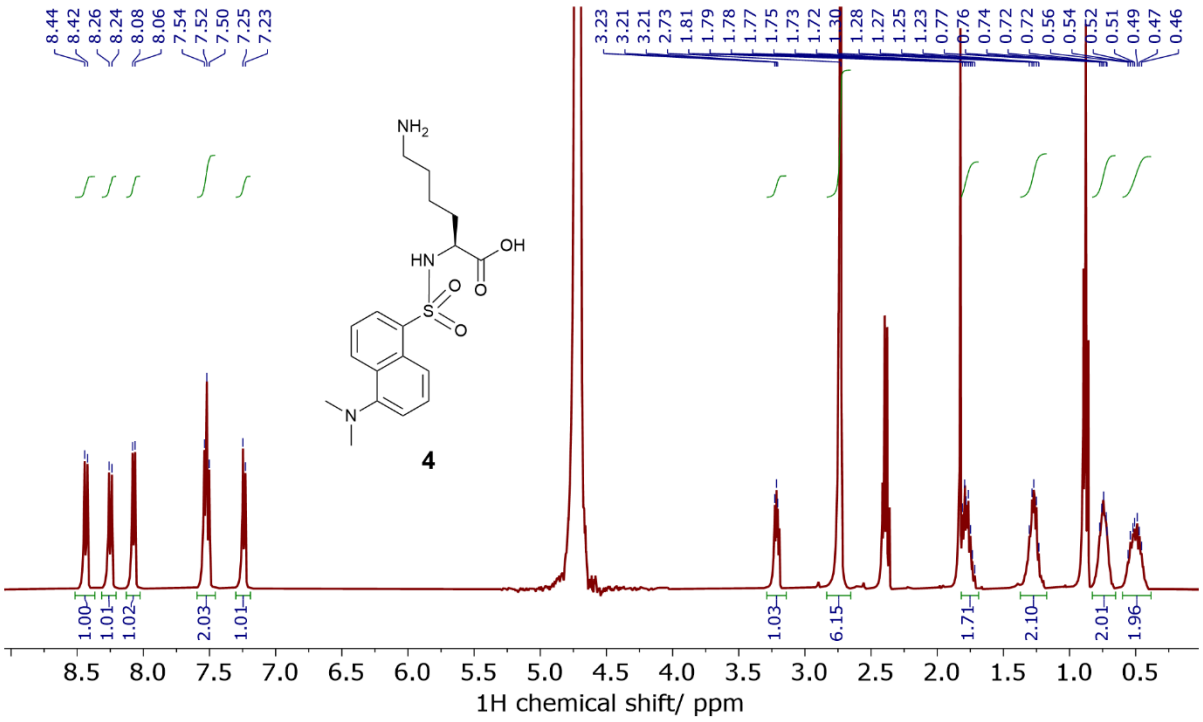


Figure 8.10 ^{1}H NMR spectrum of α-*N*-dansyl-L-lysine **4**.

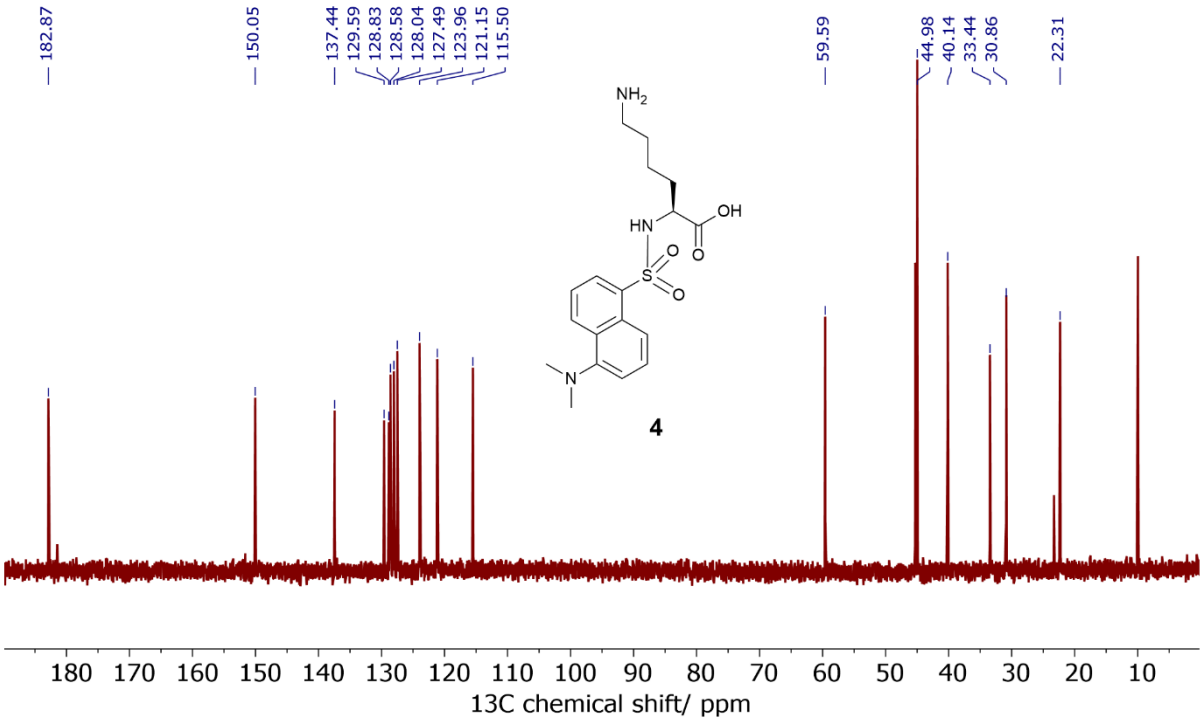


Figure 8.11 ^{13}C NMR spectrum of α-*N*-dansyl-L-lysine **4**.

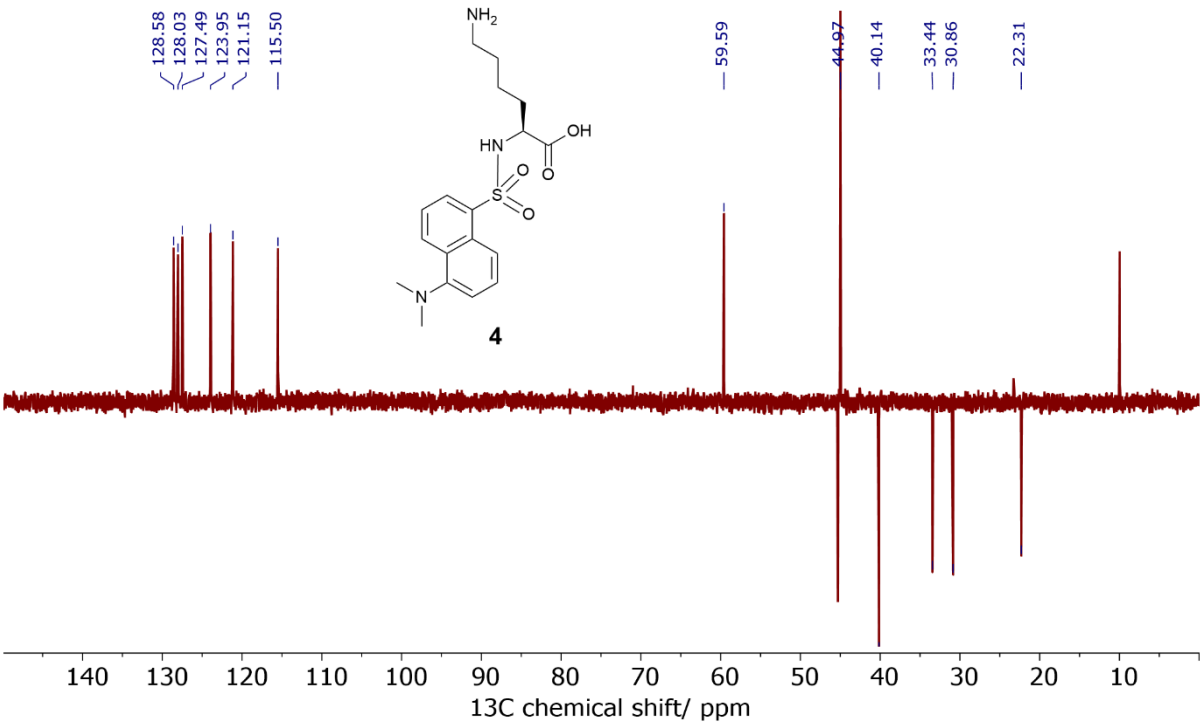


Figure 8.12 DEPT-135 NMR spectrum of α-*N*-dansyl-L-lysine **4**.

8.1.4. α-*N*-dansyl ε-*N*-thiophospho-L-lysine **7**

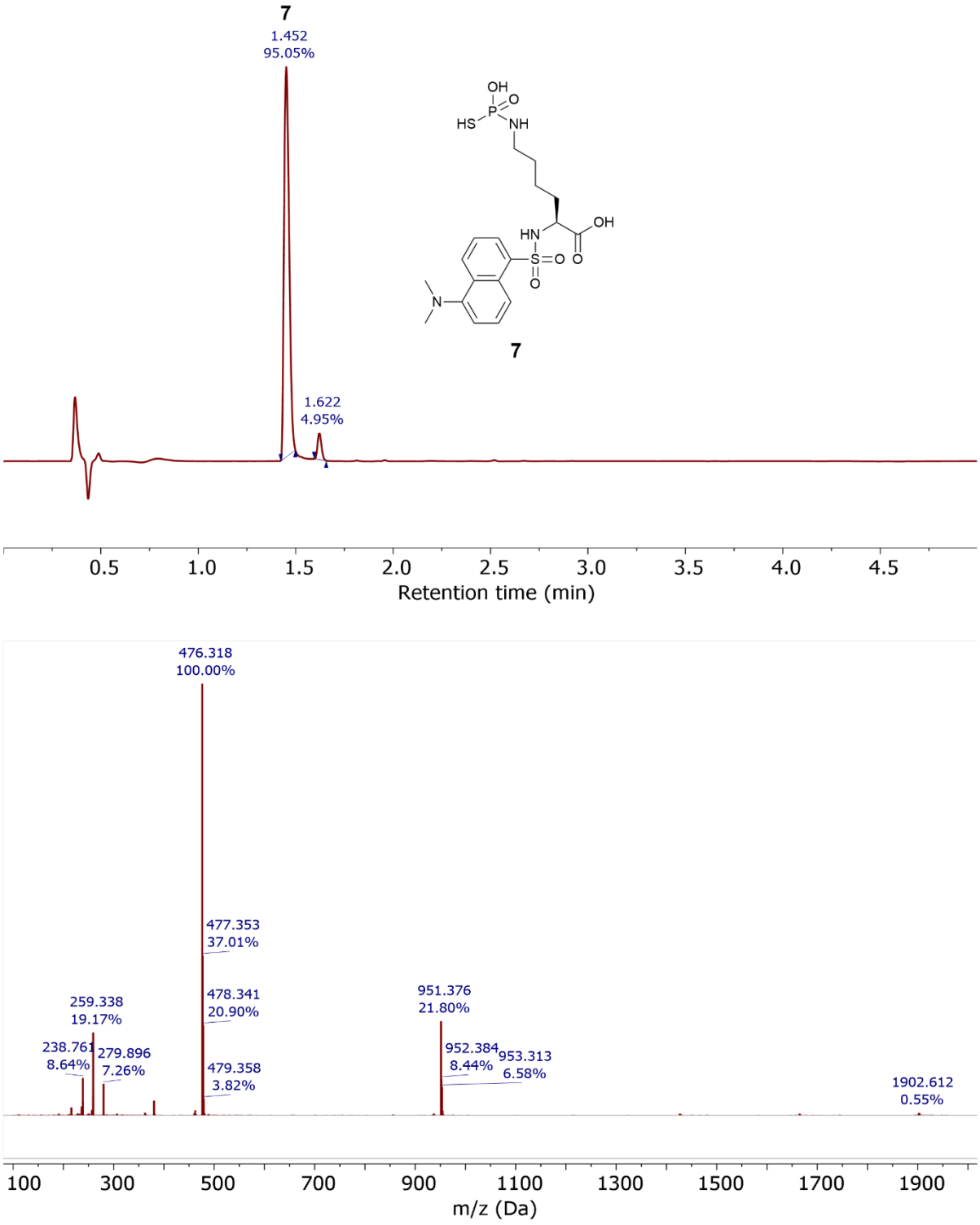


Figure 8.13 LC-MS total absorbance chromatograph and mass spectrum for peak at 1.452 min.

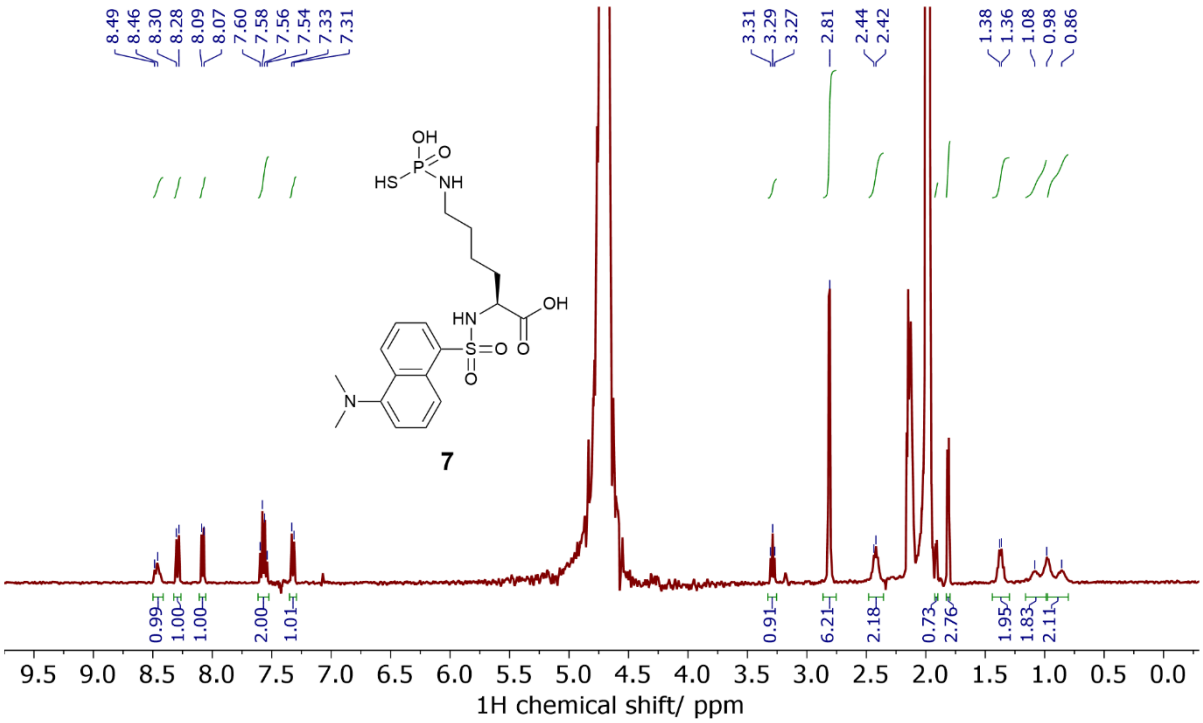


Figure 8.14 ^{1}H NMR spectrum of α-*N*-dansyl ε-*N*-thiophospho-L-lysine **7**.

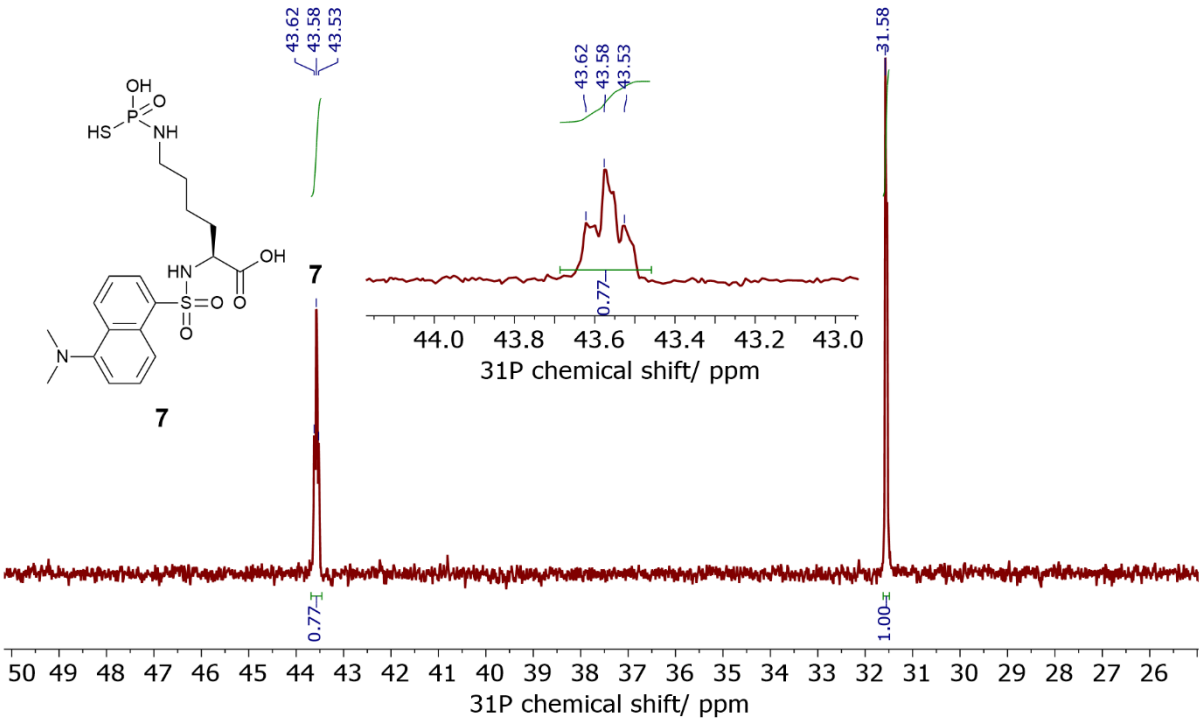


Figure 8.15 ^{31}P NMR spectrum of α-*N*-dansyl ε-*N*-thiophospho-L-lysine **7**.

8.1.5. *S*-propyl α-*N*-dansyl ε-*N*-thiophospho-L-lysine **8**

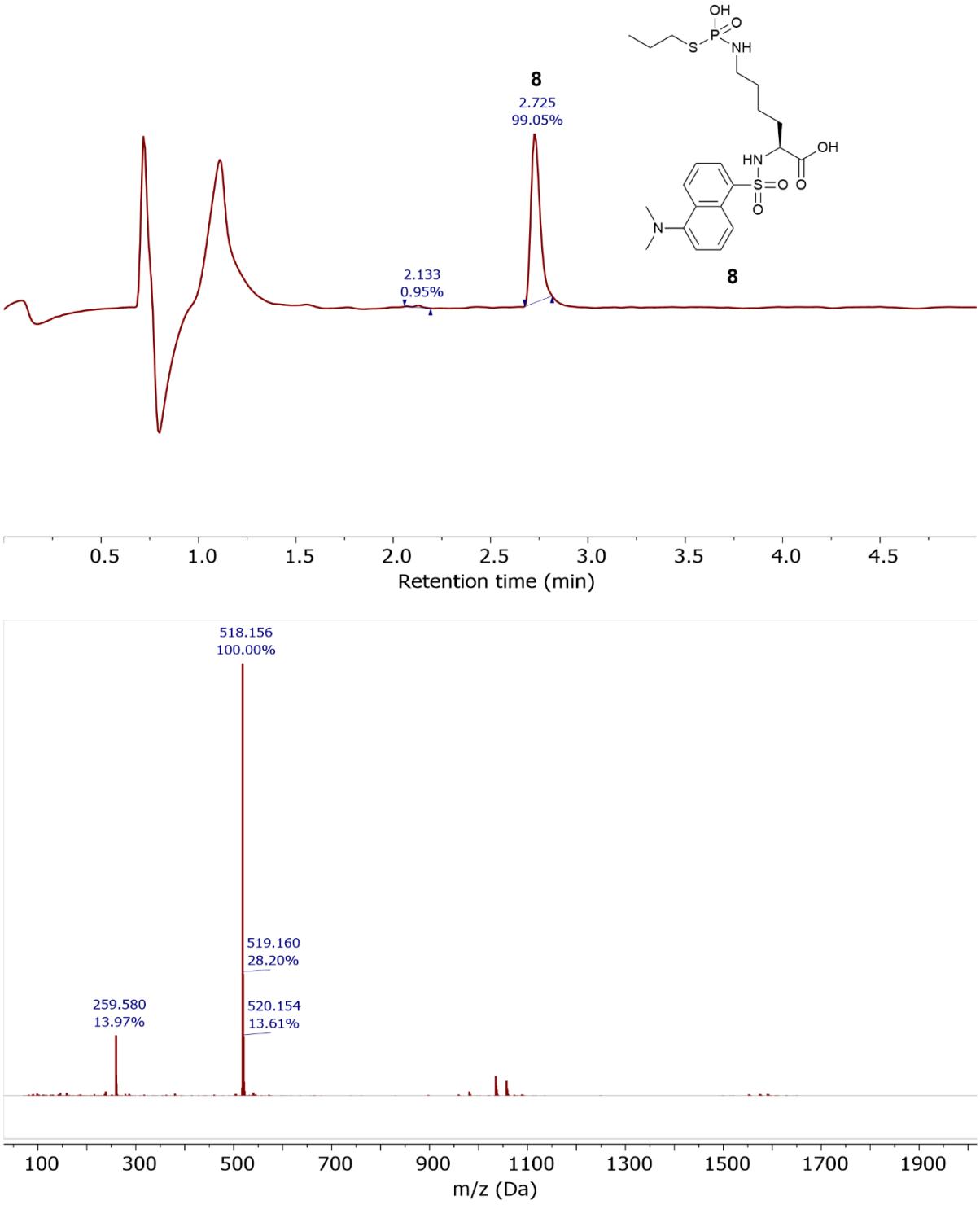


Figure 8.16 LC-MS total absorbance chromatograph and mass spectrum for peak at 1.452 min.

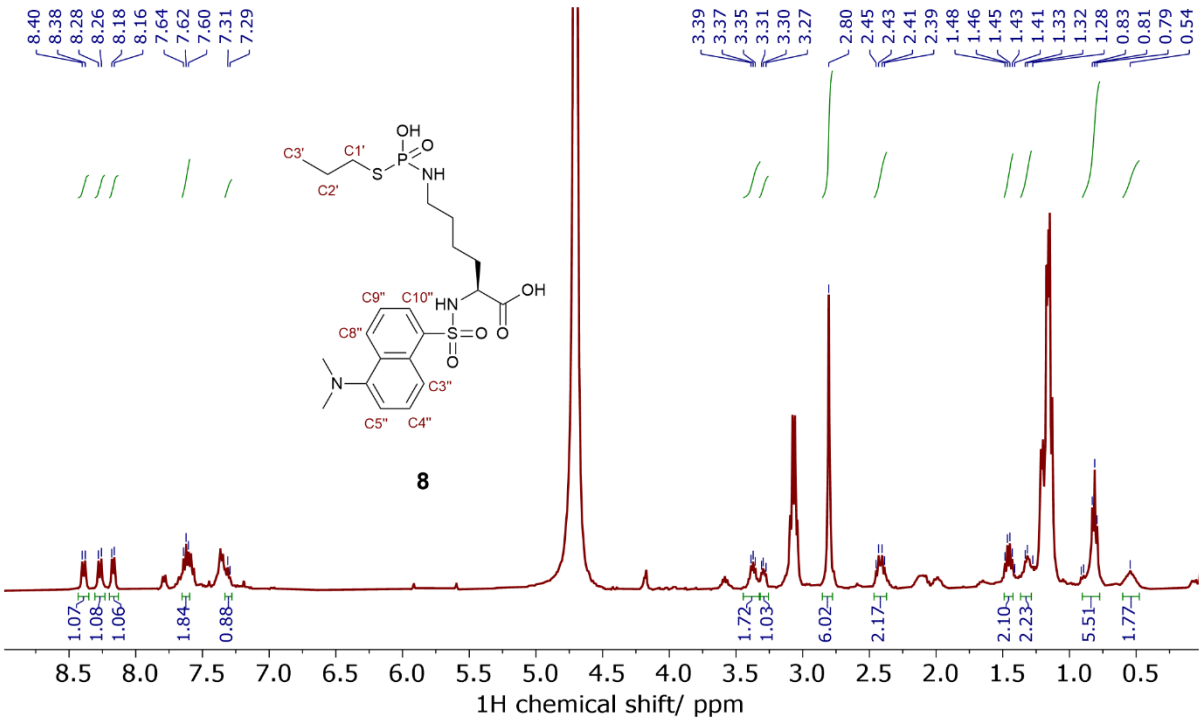


Figure 8.17 ^{1}H NMR spectrum of *S*-propyl α-*N*-dansyl ε-*N*-thiophospho-L-lysine **8**.

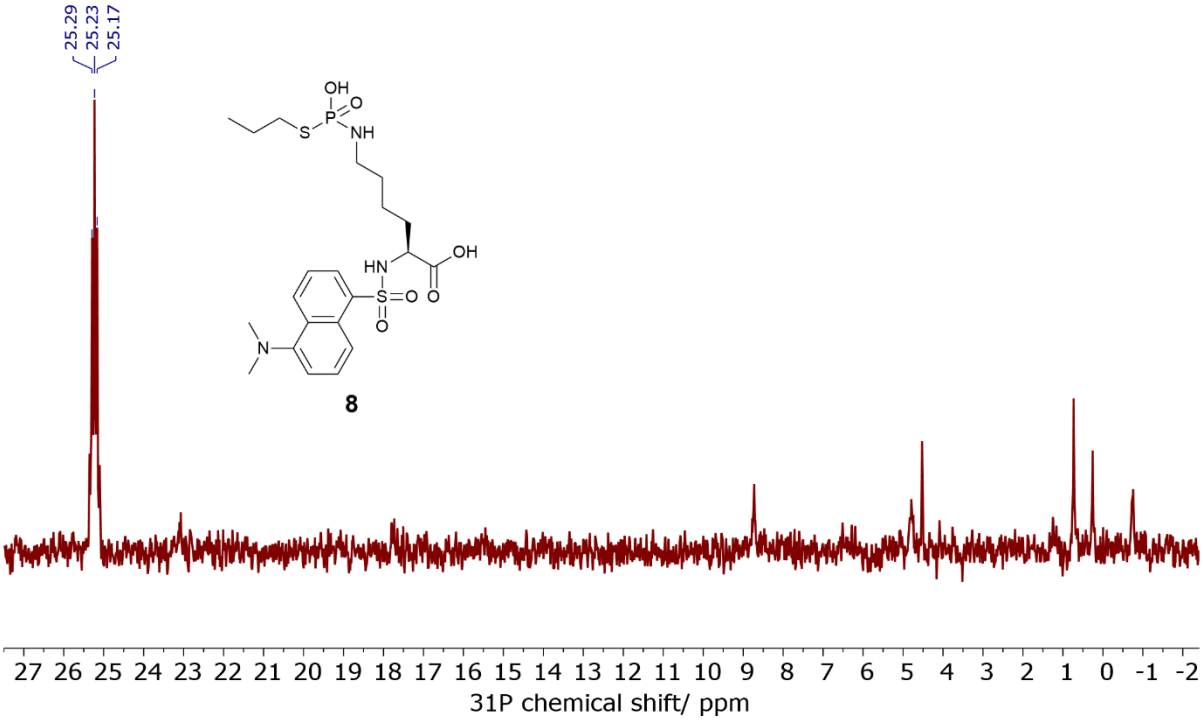
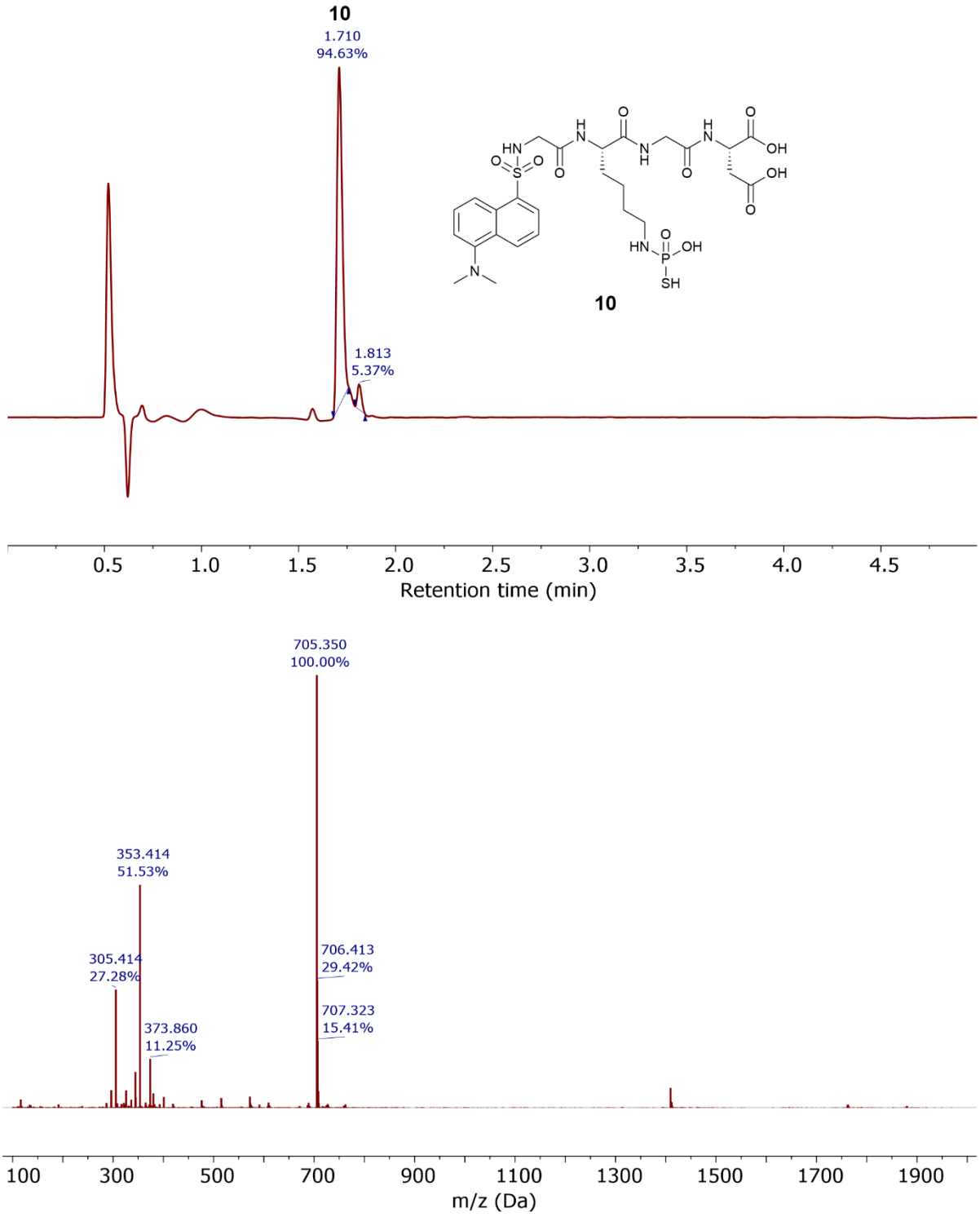


Figure 8.18 ^{31}P NMR spectrum of *S*-propyl α-*N*-dansyl ε-*N*-thiophospho-L-lysine **8**.

8.1.6. α-*N*-dansyl ε-*N*-thiophosphopeptide **10**



10

Figure 8.19 LC-MS total absorbance chromatograph and mass spectrum for peak at 1.710 min.

8.2. Chapter 3 compounds

8.2.1. ε-*N*-dansyl α-*N*-thiophosphopeptide **12**

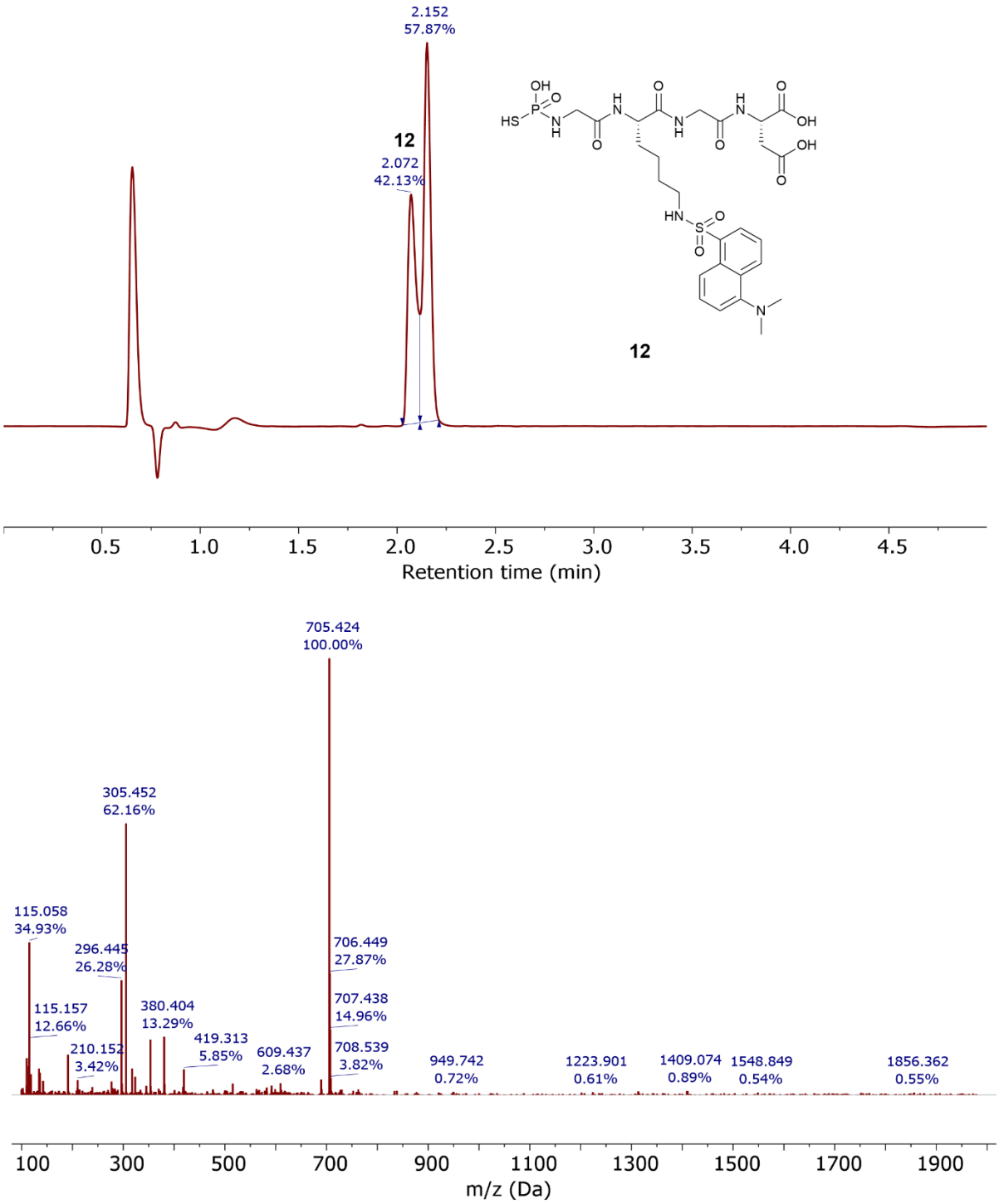


Figure 8.20 LC-MS total absorbance chromatograph and mass spectrum for peak at 2.072 min.

8.2.2. ε-*N*'-dansyl ε-*N*-thiophosphopentapeptide **14a** and/or ε-*N*'-dansyl α-*N*-thiophosphopentapeptide **14b**

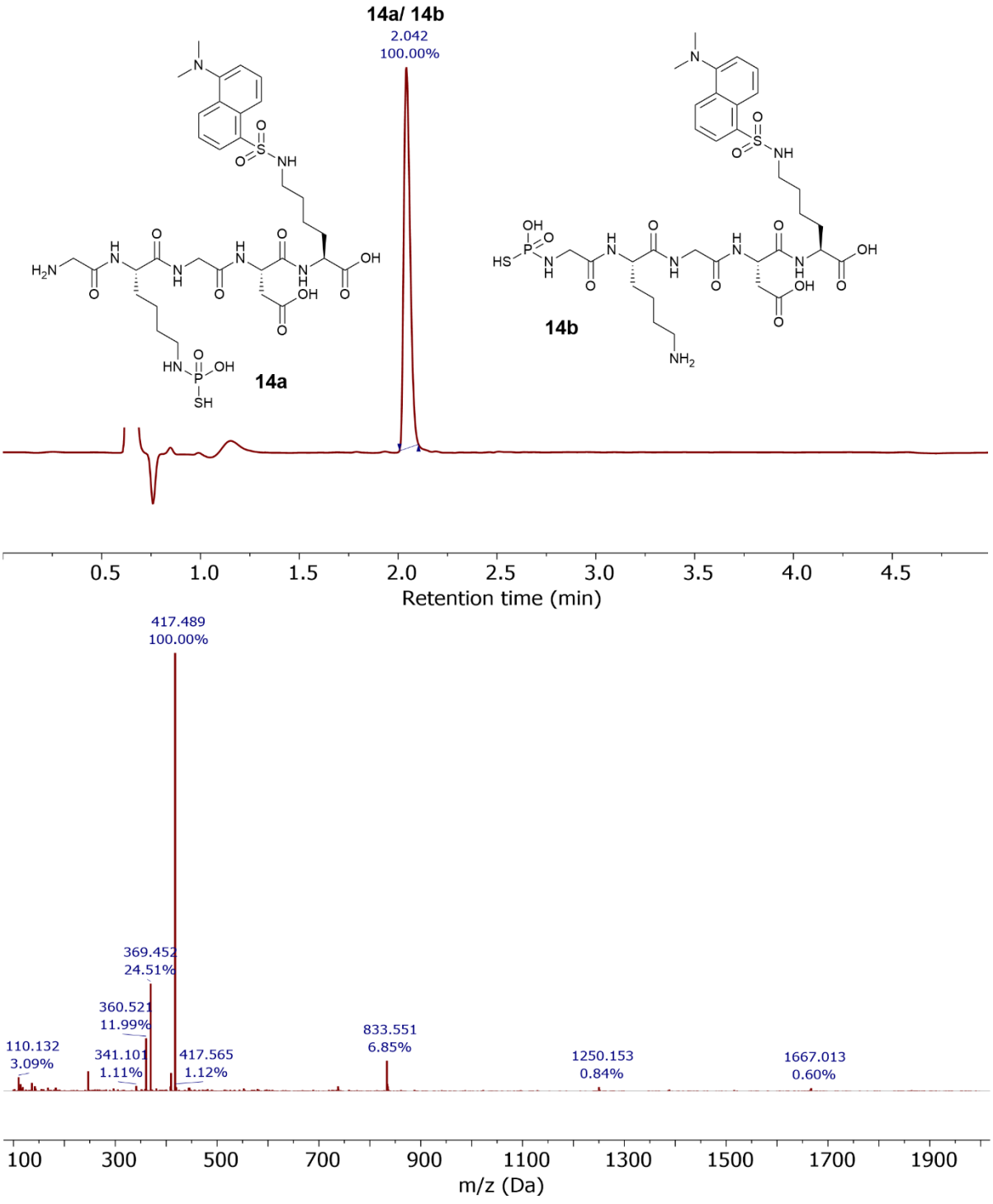


Figure 8.21 LC-MS total absorbance chromatograph and mass spectrum for peak at 2.042 min.

8.2.3. ε-*N*'-dansyl *S*-propyl ε-*N*-thiophosphopentapeptide **15a** and/or ε-*N*'-dansyl *S*-propyl α-*N*-thiophosphopentapeptide **15b**

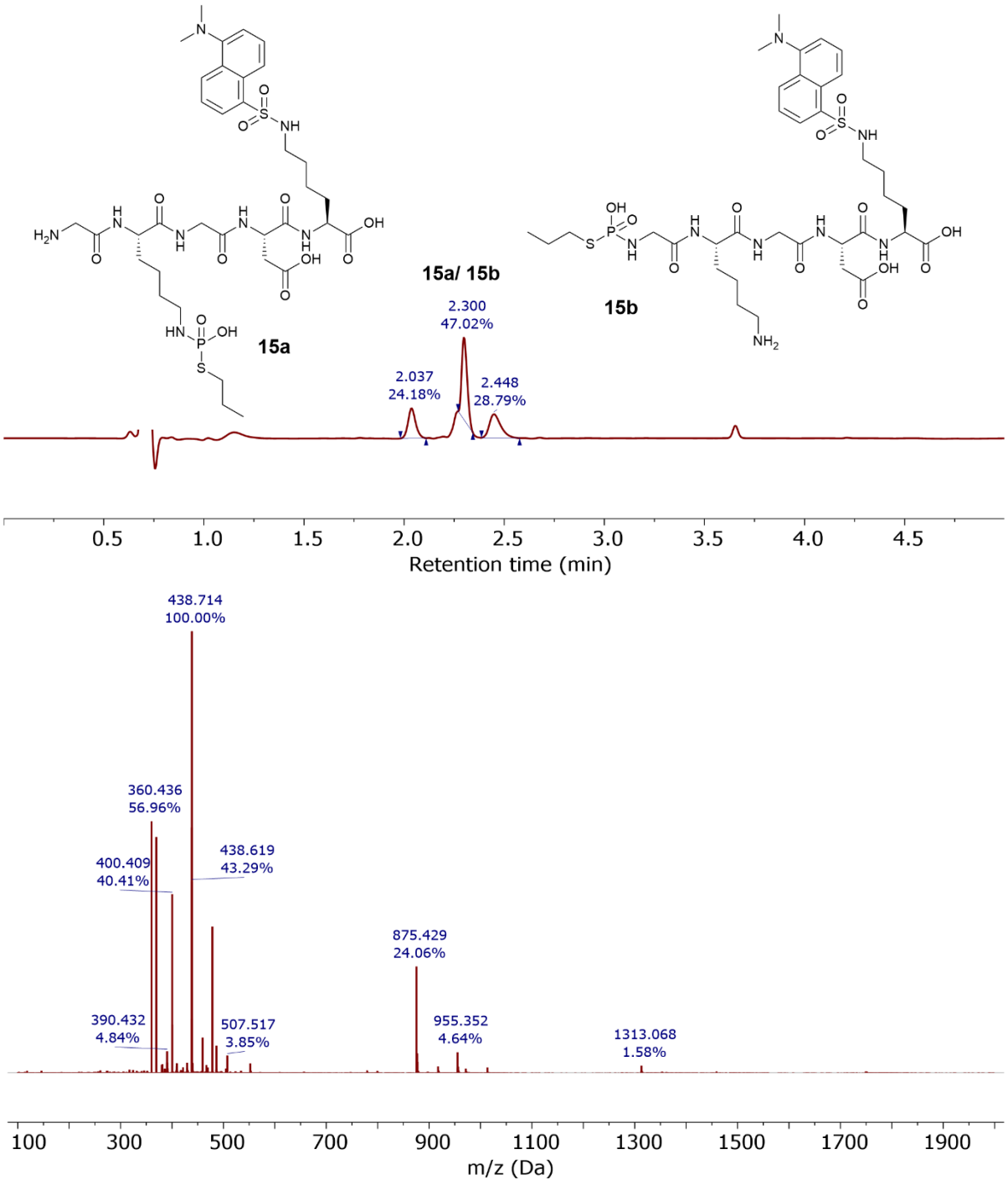


Figure 8.22 LC-MS total absorbance chromatograph and mass spectrum for peak at 2.300 min.

8.2.4. α-*N*-acetyl *S*-naphthyl ε-*N*-thiophosphoramidate **18**

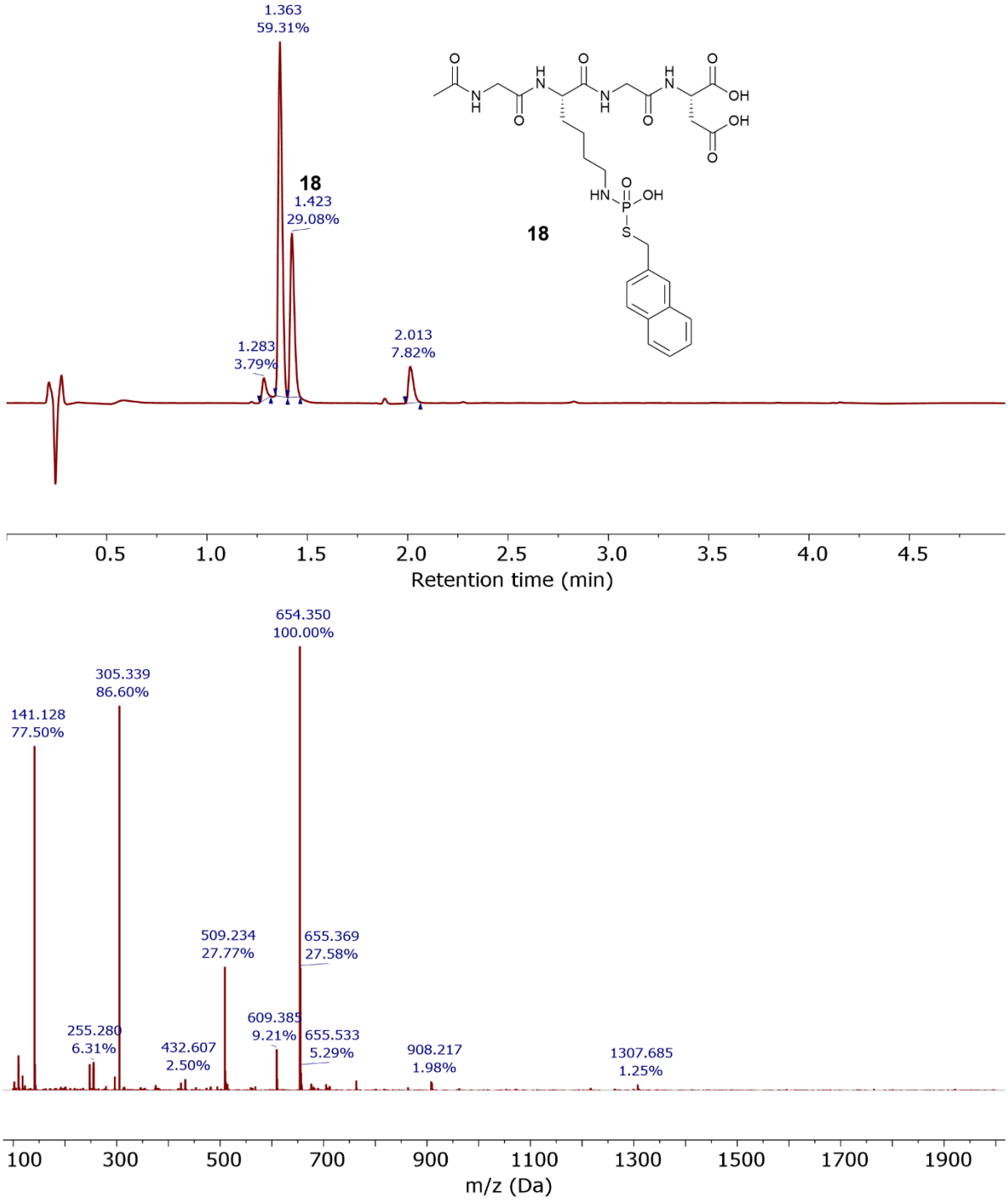


Figure 8.23 LC-MS total absorbance chromatograph and mass spectrum for peak at 1.423 min.

8.2.5. ε-*N*-dansyl *S*-naphthyl α-*N*-thiophosphopeptide **19**

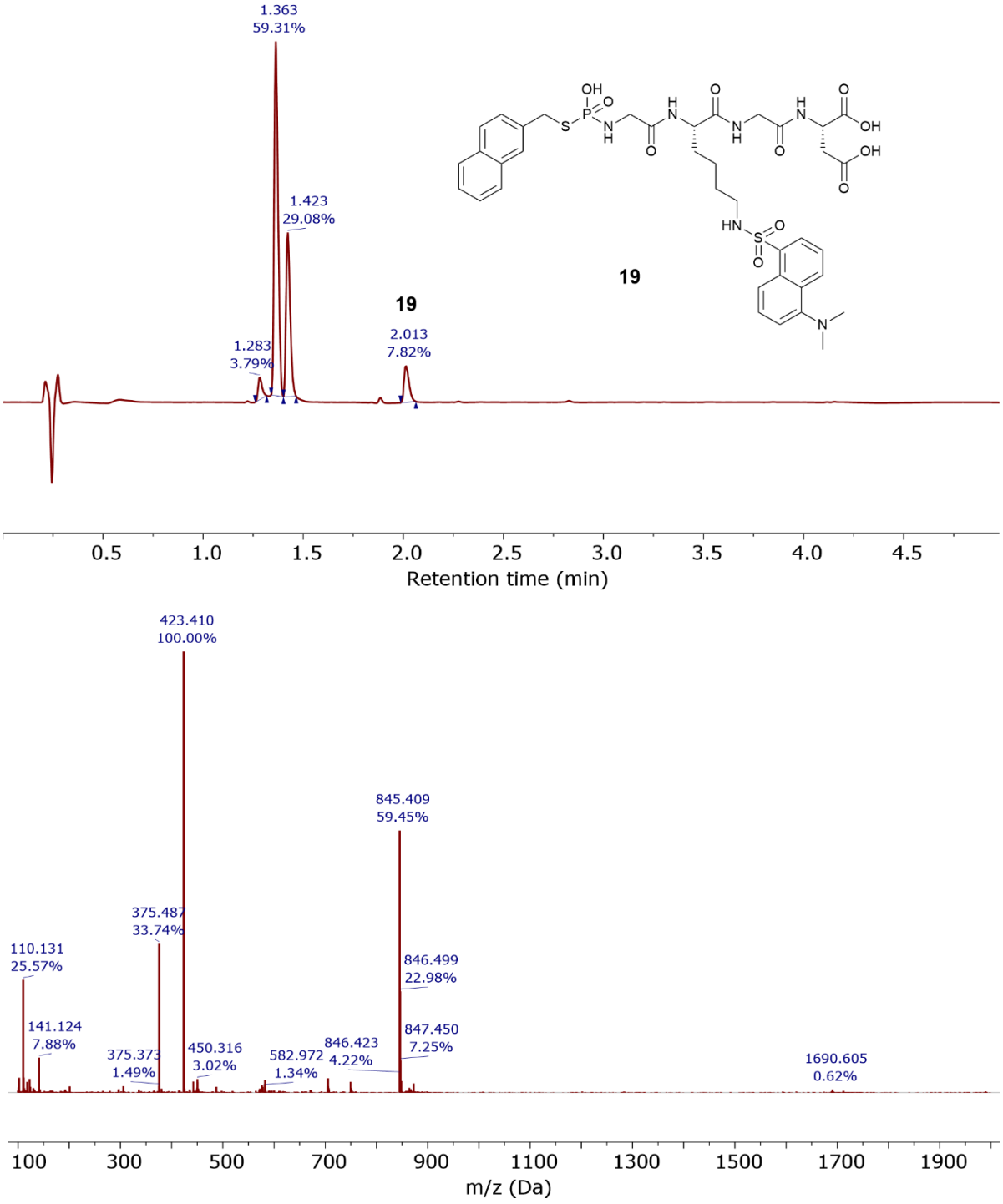


Figure 8.24 LC-MS total absorbance chromatograph and mass spectrum for peak at 2.013 min.

8.2.6. *N*-terminal glycine ε-*N*-thiophosphoramidate **21b** and α-*N*-thiophosphoramidate **21c**

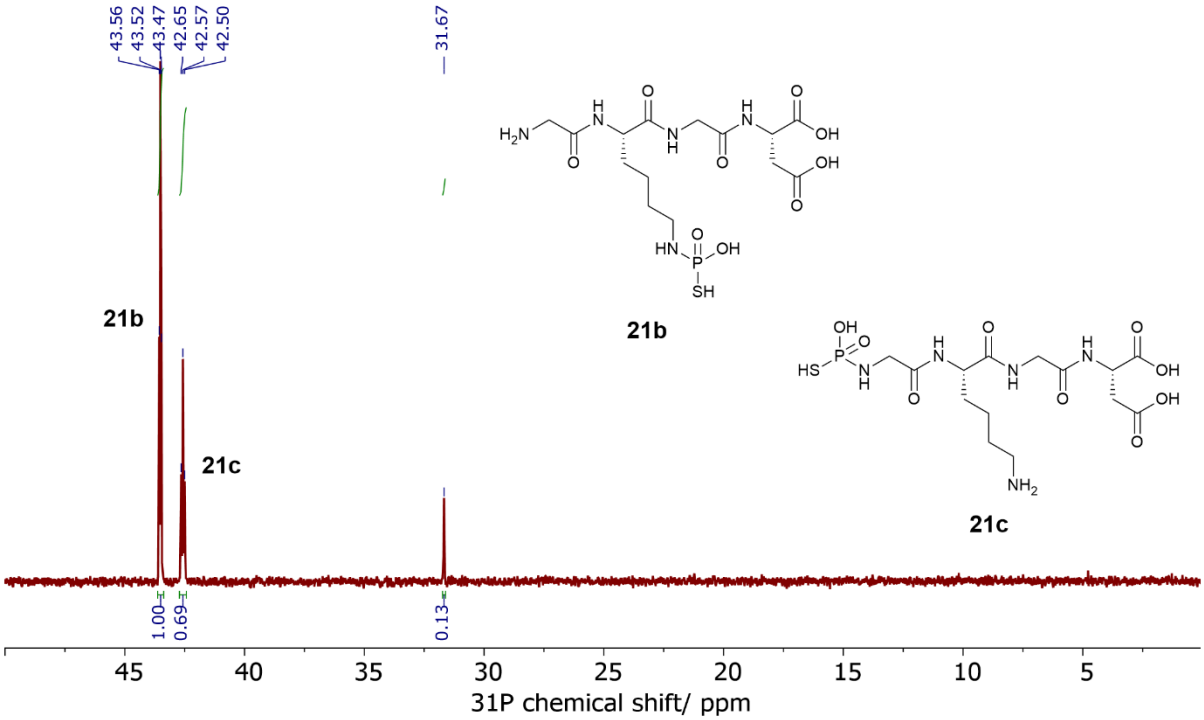


Figure 8.25 ^{31}P NMR spectrum of *N*-terminal glycine ε-*N*-thiophosphoramidate **21b**.

8.2.7. *N*-terminal alanine ε-*N*-thiophosphoramidate **22b** and α-*N*-thiophosphoramidate **22c**

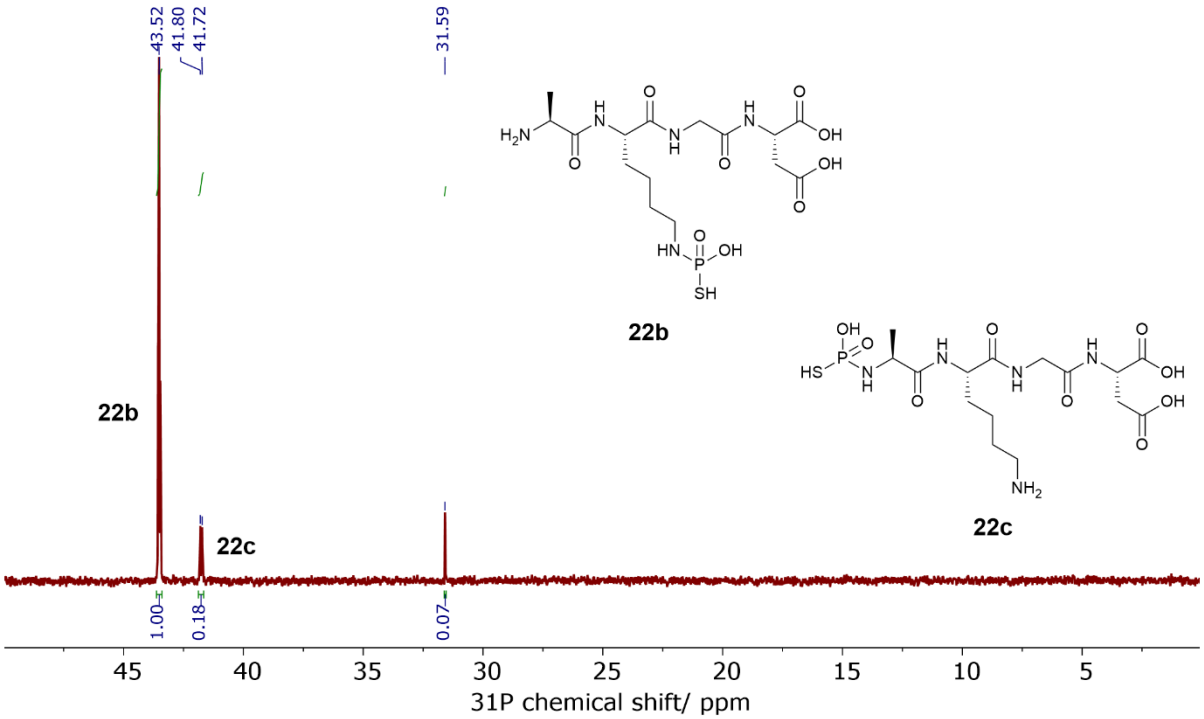


Figure 8.26 ^{31}P NMR spectrum of *N*-terminal alanine ε-*N*-thiophosphoramidate **22b**.

8.2.8. *N*-terminal phenylalanine ε-*N*-thiophosphoramidate **23b** and α-*N*-thiophosphoramidate **23c**

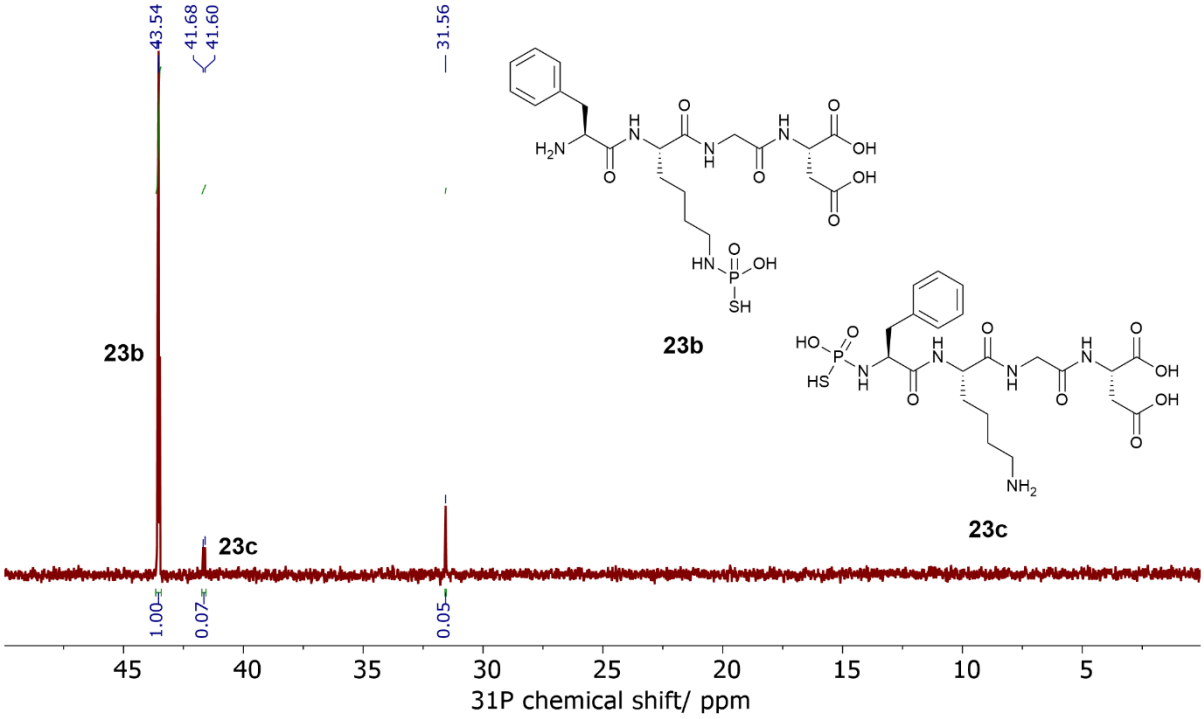


Figure 8.27 ^{31}P NMR spectrum of *N*-terminal phenylalanine ε-*N*-thiophosphoramidate **23b**.

8.2.9. *N*-terminal serine ε-*N*-thiophosphoramidate **24b** and α-*N*-thiophosphoramidate **24c**

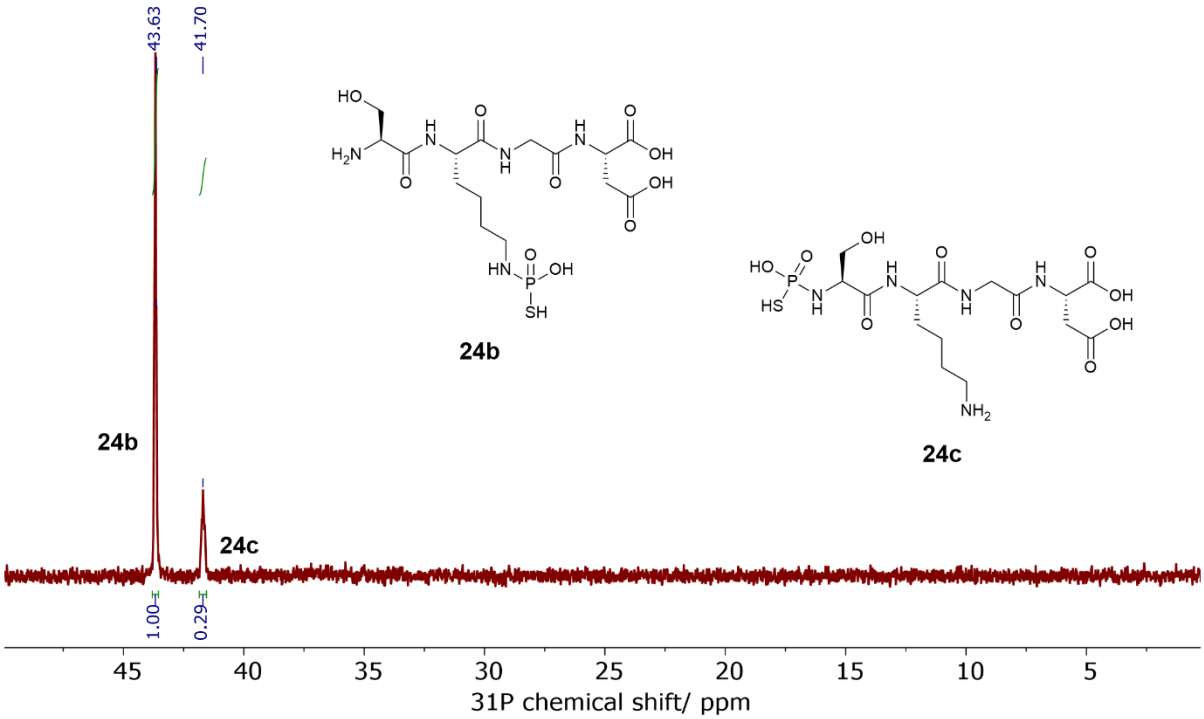


Figure 8.28 ^{31}P NMR spectrum of *N*-terminal serine ε-*N*-thiophosphoramidate **24b**.

8.2.10. *N*-terminal cysteine ε-*N*-thiophosphoramidate **25b** and α-*N*-thiophosphoramidate **25c**

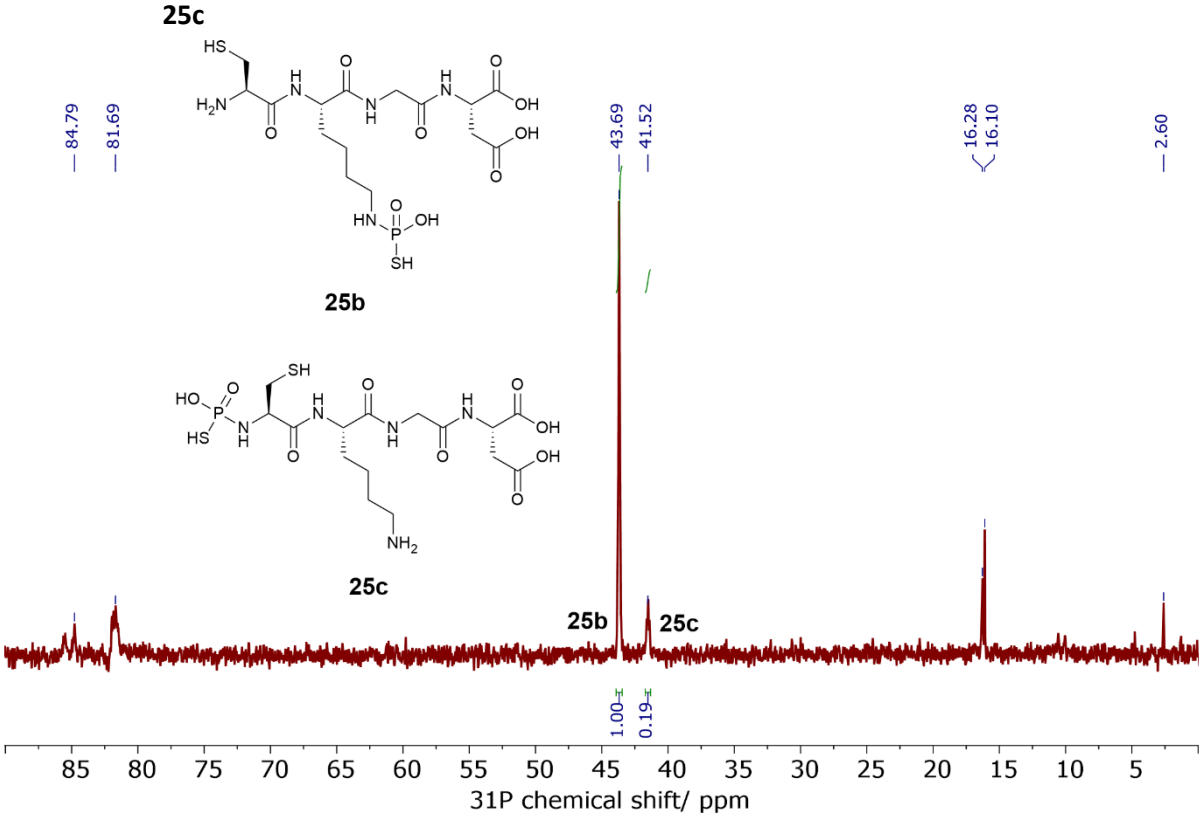


Figure 8.29 ^{31}P NMR spectrum of *N*-terminal cysteine ε-*N*-thiophosphoramidate **25b**.

8.2.11. *N*-terminal histidine ε-*N*-thiophosphoramidate **26b** and α-*N*-thiophosphoramidate **26c**

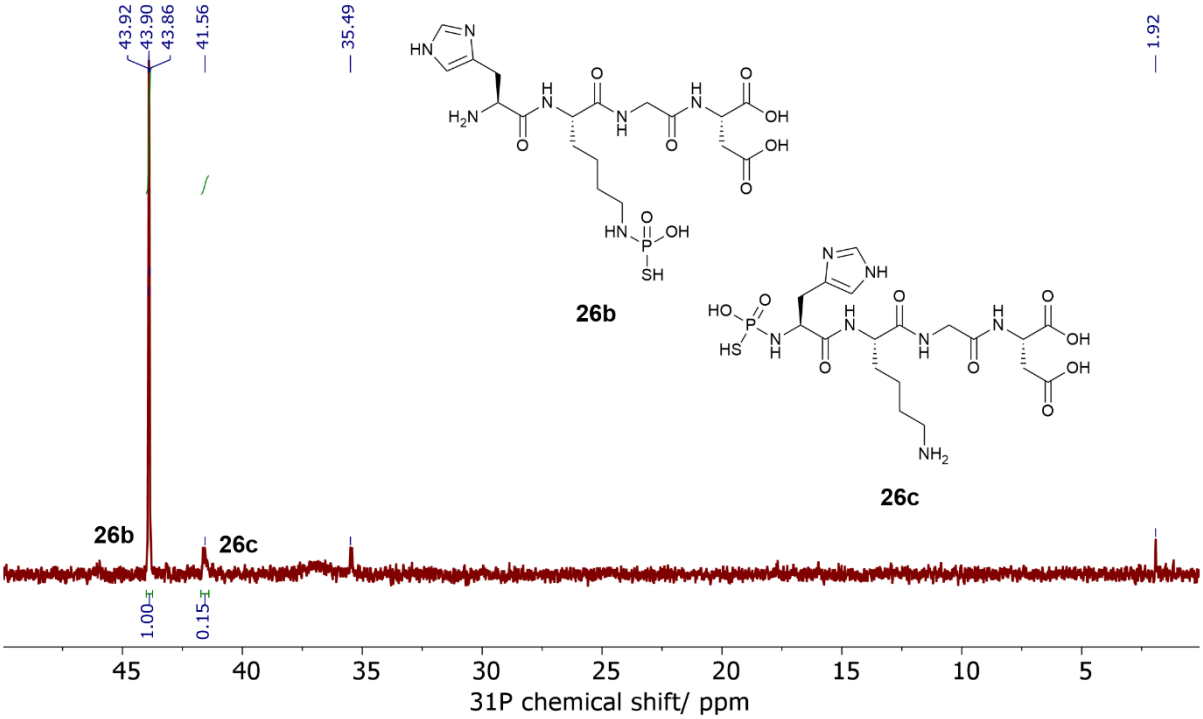


Figure 8.30 ^{31}P NMR spectrum of *N*-terminal histidine ε-*N*-thiophosphoramidate **26b**.

8.2.12. *N*-terminal phenylalanine *S*-naphthyl ε-*N*-thiophosphoramidate **23d**

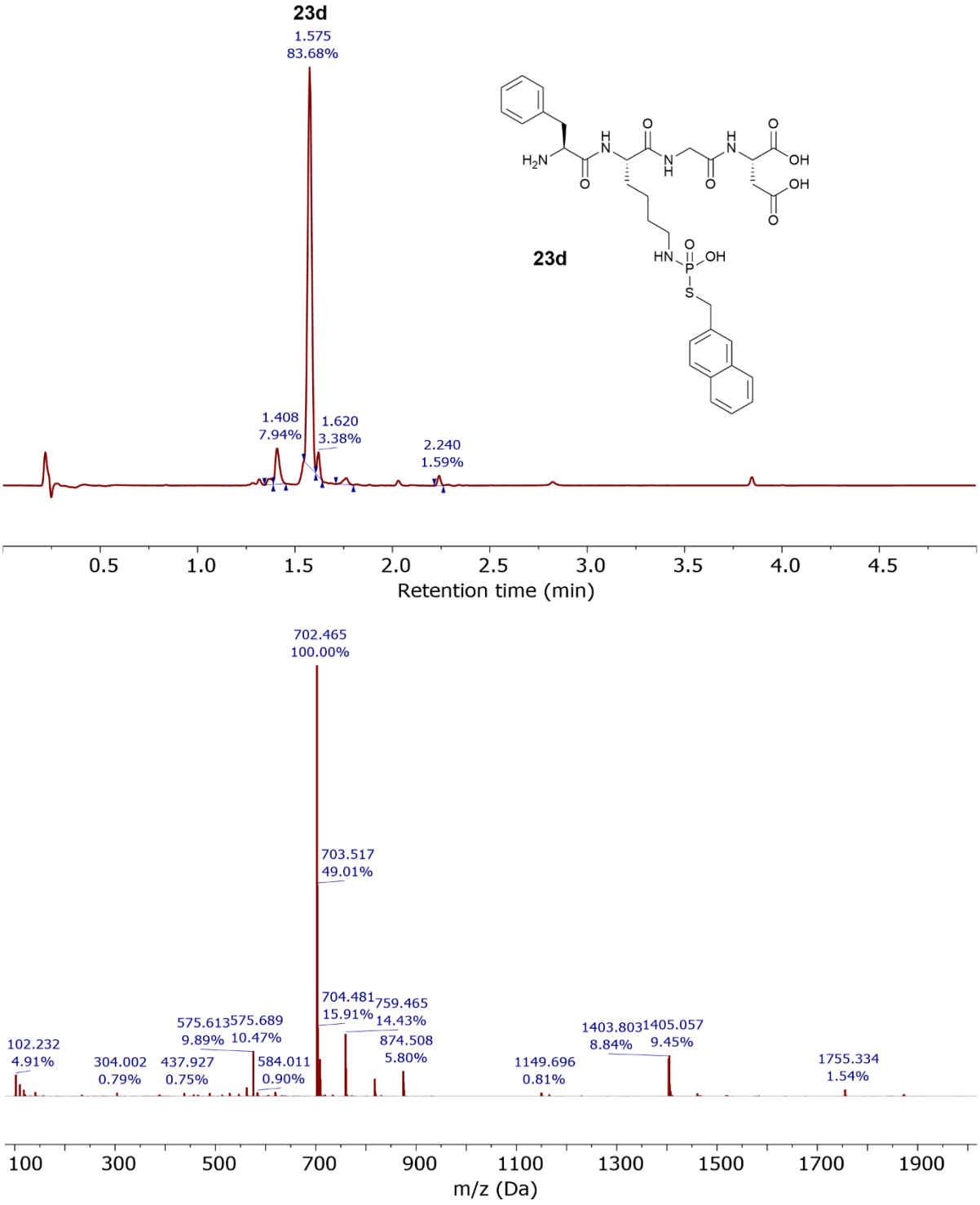


Figure 8.31 LC-MS total absorbance chromatograph and mass spectrum for peak at 1.575 min.

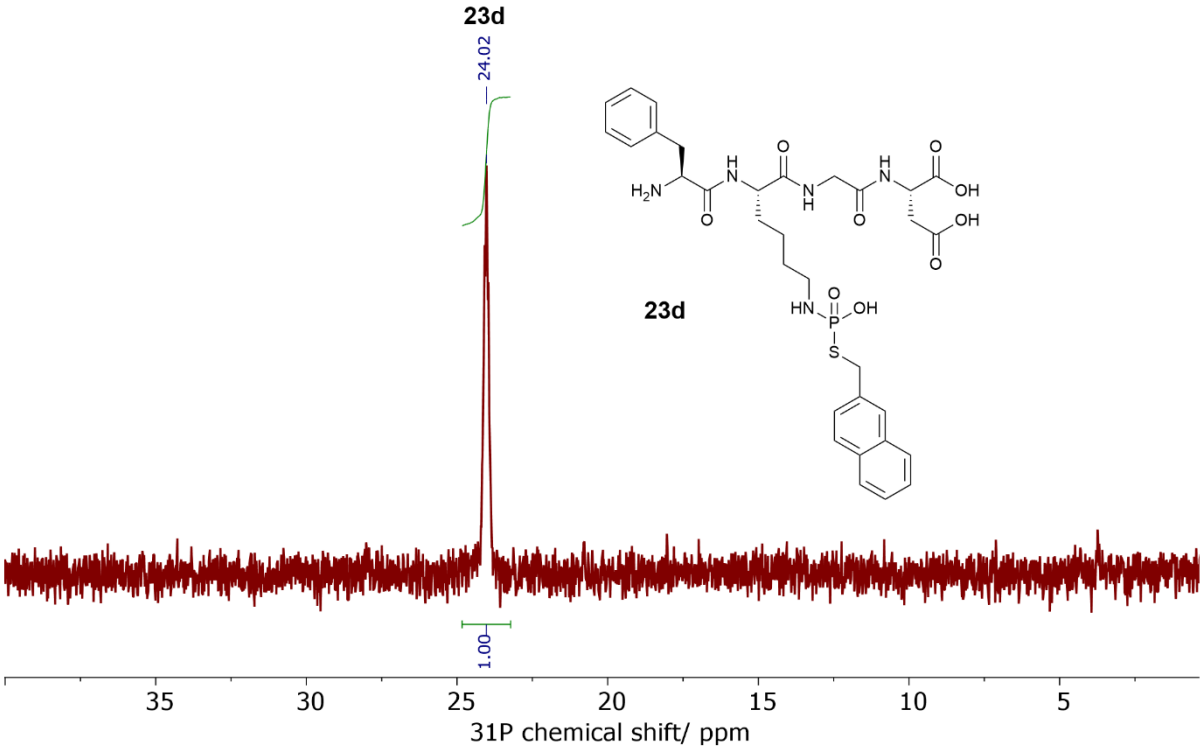


Figure 8.32 ^{31}P NMR spectrum of *N*-terminal phenylalanine *S*-naphthyl ε-*N*-thiophosphoramidate **23d**.

8.3. Chapter 4 compounds

8.3.1. Potassium thiophosphorodichloridate

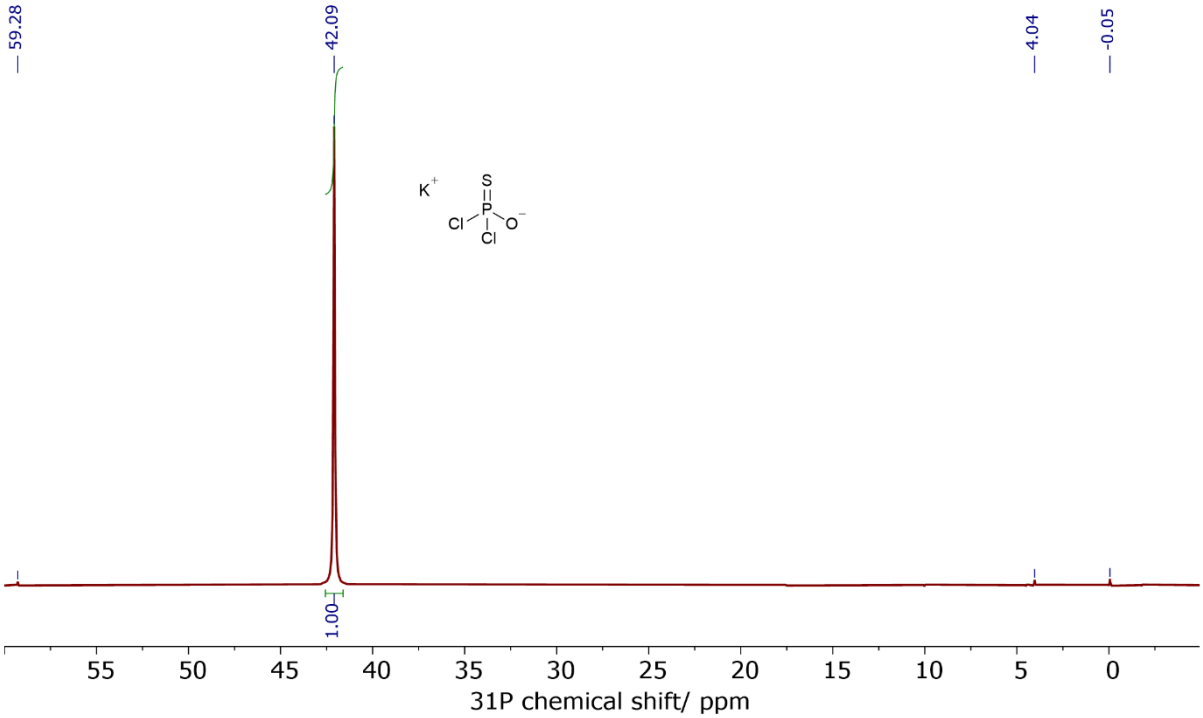


Figure 8.33 ^{31}P NMR spectrum of potassium thiophosphorodichloridate.

8.4. Chapter 5 compounds

8.4.1. α-*N*-dansyl-labelled heptamer **27**

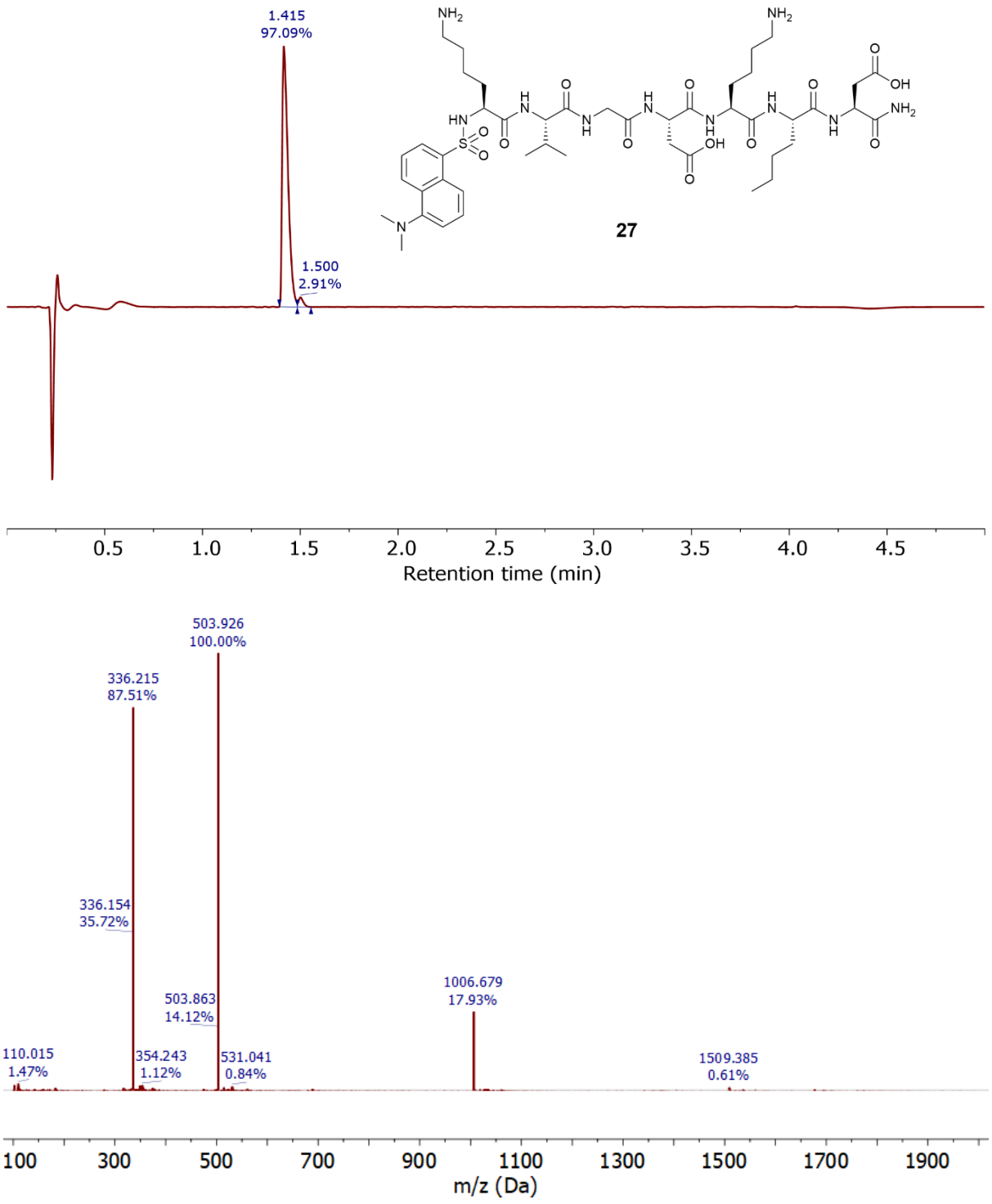


Figure 8.34 LC-MS total absorbance chromatograph and mass spectrum for peak at 1.415 min.

8.4.2. α-*N*-dansyl-labelled di-thiophosphoramidate **28a**

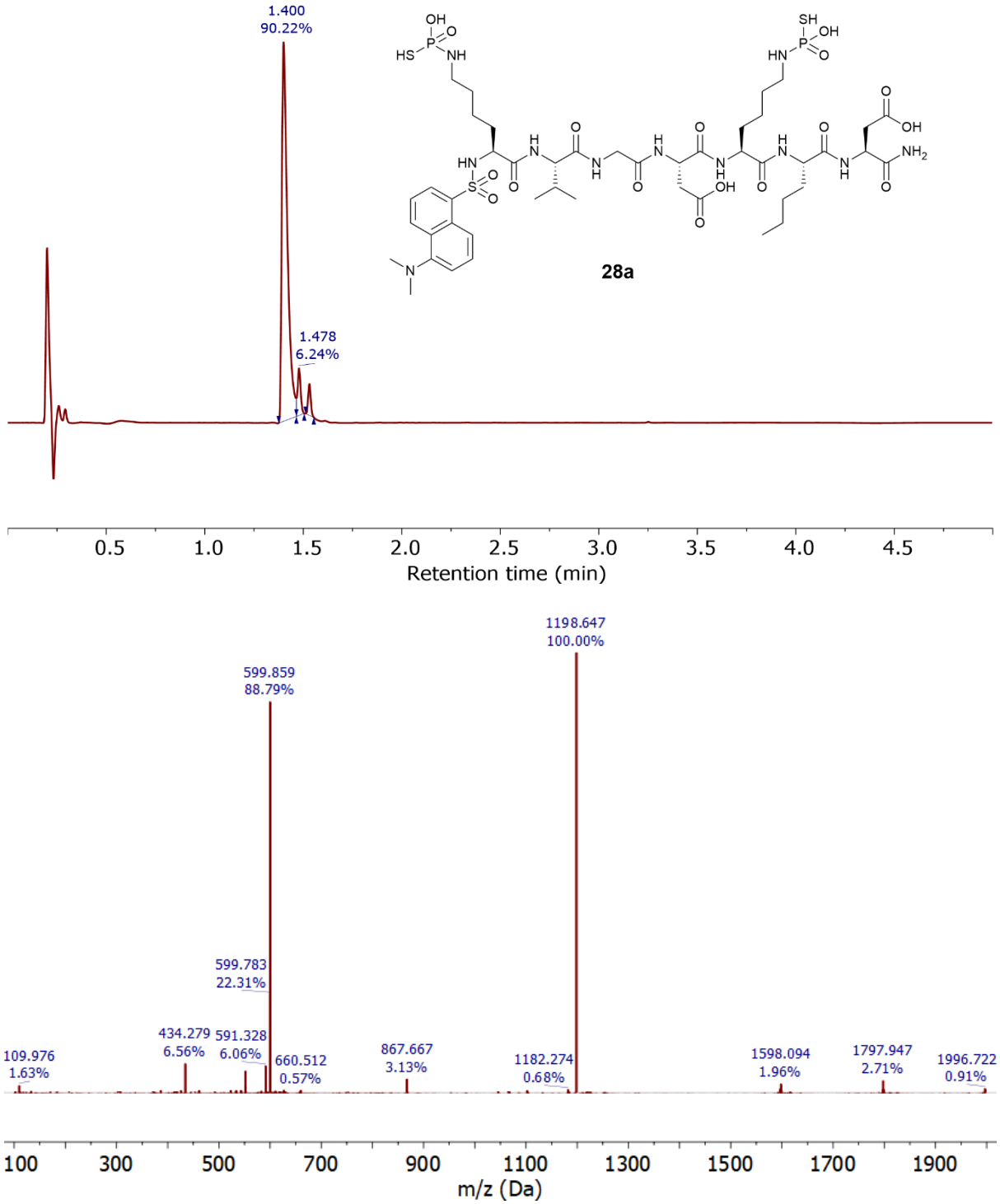


Figure 8.35 LC-MS total absorbance chromatograph and mass spectrum for peak at 1.400 min.

8.4.3. α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**

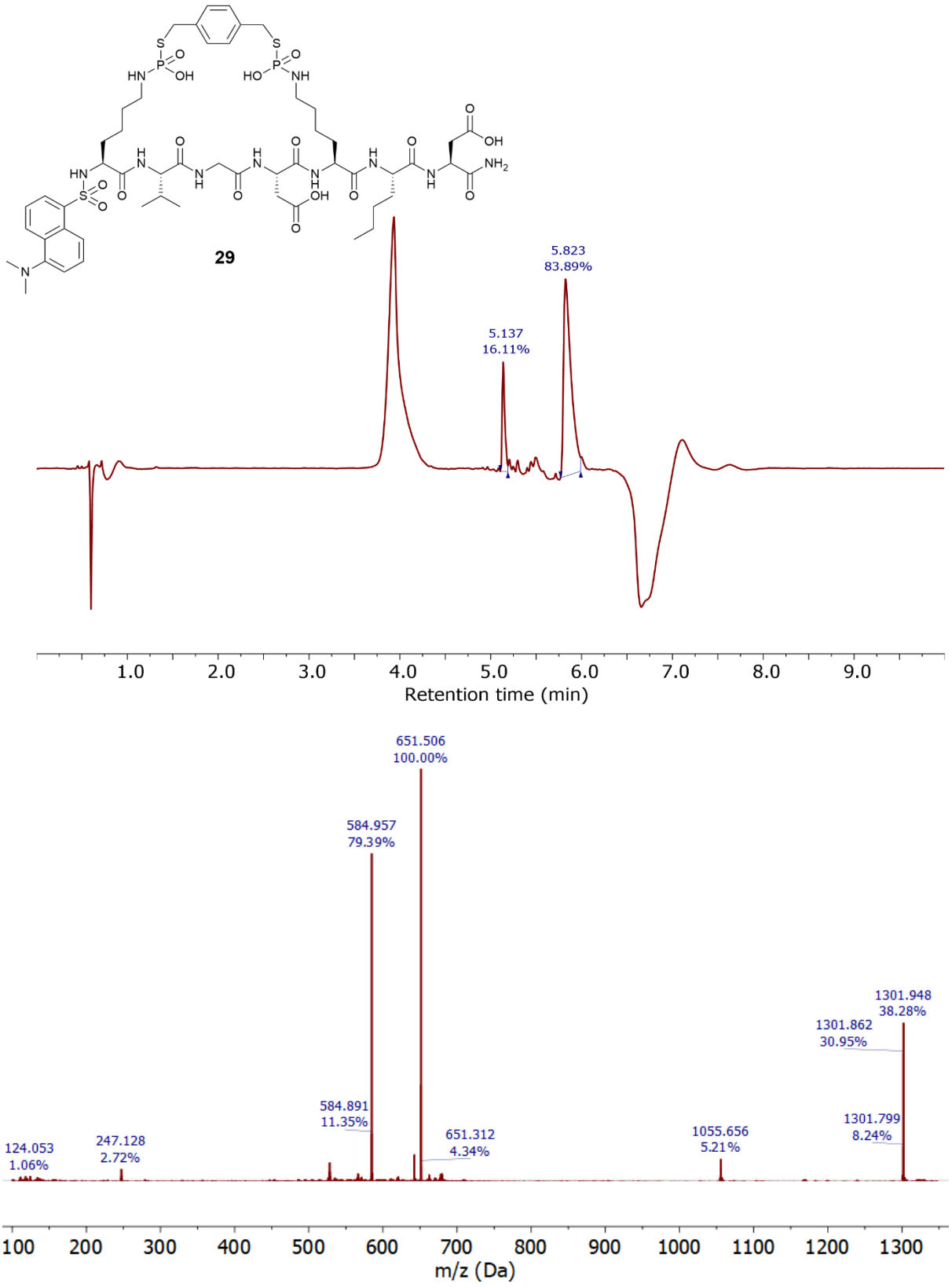


Figure 8.36 LC-MS total absorbance chromatograph and mass spectrum for peak at 5.823 min.

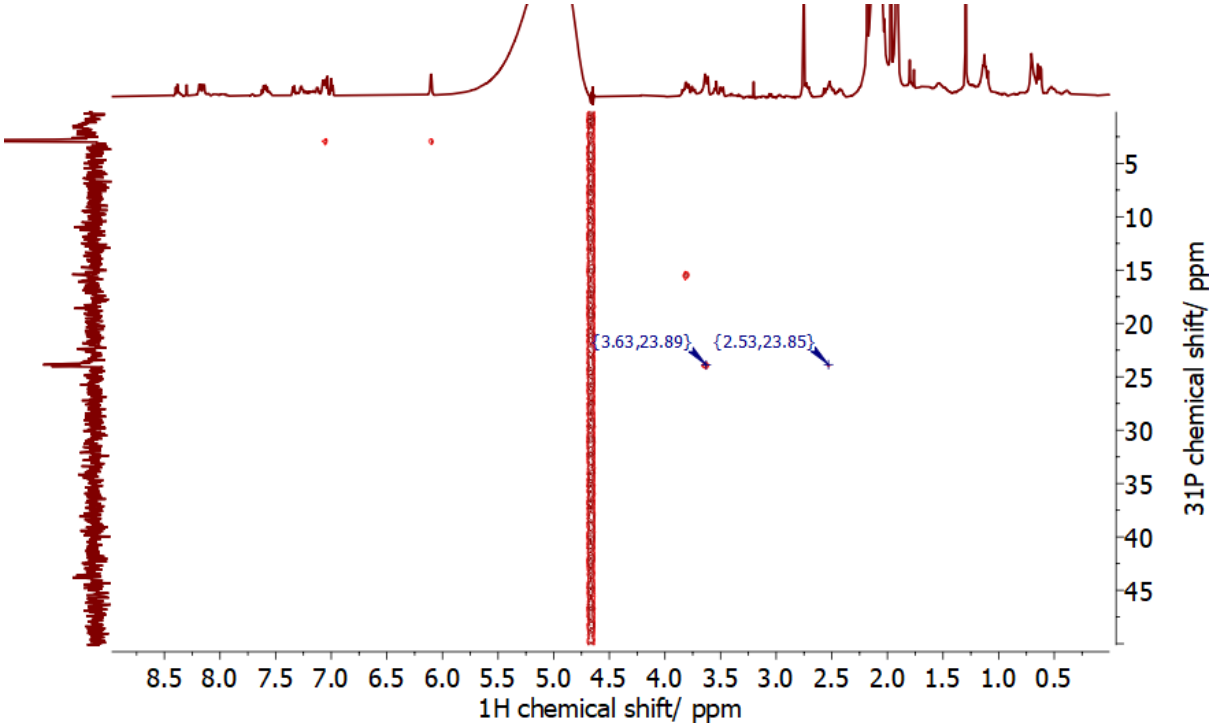


Figure 8.37 ^{1}H-^{31}P HMBC NMR spectrum of α-*N*-dansyl-labelled stapled di-thiophosphoramidate **29**.

9. References

(1) T. Curtius, *J. prakt. Chem.,* 1882, **26**, 145-208.
(2) E. Fischer and E. Fourneau, *Ber. Dtsch. Chem. Ges.,* 1901, **34**, 2868-2877.
(3) R. B. Merrifield, *J. Am. Chem. Soc.,* 1963, **85**, 2149-2154.
(4) J. Lau, P. Bloch, L. Schäffer, I. Pettersson, J. Spetzler, J. Kofoed, K. Madsen, L. B. Knudsen, J. McGuire, D. B. Steensgaard, *J. Med. Chem.,* 2015, **58**, 7370-7380.
(5) M. J. Gomara, Y. Perez, P. Gomez-Gutierrez, C. Herrera, P. Ziprin, J. P. Martinez, A. Meyerhans, J. J. Perez, I. Haro, *Sci. Rep.,* 2020, **10**.
(6) A. C. Conibear, E. E. Watson, R. J. Payne, C. F. W. Becker, *Chem. Soc. Rev.,* 2018, **47**, 9046-9068.
(7) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science,* 1994, **266**, 776-779.
(8) V. Agouridas, O. El Mahdi, V. Diemer, M. Cargoët, J-C. M. Monbaliu, O. Melnyk, *Chem. Rev.,* 2019, **119**, 7328-7443.
(9) P. Thapa, R-Y. Zhang, V. Menon, J-P. Bingham, *Molecules,* 2014, **19**, 14461-14483.
(10) L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.,* 1996, **118**, 5891-5896.
(11) H. Sun, A. Brik, *Acc. Chem. Res.,* 2019, **52**, 3361-3371.
(12) S. Tang, L-J. Liang, Y-Y. Si, S. Gao, J-X. Wang, J. Liang, Z. Mei, J-S. Zheng, L. Lie, *Angew. Chem. Int. Ed.,* 2017, **56**, 13333-13337.
(13) G. G. Kochendoerfer, S-Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, *Science,* 2003, **299**, 884-887.
(14) B. L. Nilsson, M. B. Soellner, R. T. Raines, *Ann. Rev. Biophys. Biomol. Struct.,* 2005, **34**, 91-118.
(15) E. E. Watson, X. Liu, R. E. Thompson, J. Ripoll-Rozada, M. Wu, I. Alwis, A. Gori, C-T. Loh, B. L. Parker, G. Otting, *ACS Cent. Sci.,* 2018, **4**, 468-476.
(16) N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo, *Science,* 2020, **368**, 980-987.
(17) D. J. Craik, D. P. Fairlie, S. Liras, D. Price, *Chem. Biol. Drug Des.,* 2013, **81**, 136-147.
(18) I. Vecchio, C. Tornali, N. L. Bragazzi, M. Martini, *Front. Endocrinol.,* 2018, **9**.
(19) P. Diem, P. H. Ducluzeau, A. Scheen, *Diabet. Epidemiol. Manag.,* 2022, **5**, 100049.
(20) M. Levy, *Insulin Development and Commericalization*. Am. Chem. Soc., 2023. https://www.acs.org/education/whatischemistry/landmarks/insulin.html (accessed 05 October 2024).
(21) N. A. Baeshen, M. N. Baeshen, A. Sheikh, R. S. Bora, M. M. M. Ahmed, H. A. I. Ramadan, K. S. Saini, E. M. Redwan, *Microb. Cell Fact.,* 2014, **13**.
(22) V. P. Martinovich, K. U. Baradzina, *Russ. J. Bioorg. Chem.,* 2022, **48**, 221-232.
(23) V. du Vigneaud, C. Ressler, C. J. M. Swan, C. W. Roberts, P. G. Katsoyannis, S. Gordon, *J. Am. Chem. Soc.,* 1953, **75**, 4879-4880.
(24) V. du Vigneaud, C. Ressler, C. J. M. Swan, C. W. Roberts, P. G. Katsoyannis, *J. Am. Chem.Soc.,* 1954, **76**, 3115-3121.
(25) M. Bodanszky and V. du Vigneaud, *J. Am. Chem. Soc.,* 1959, **81**, 5688-5691.
(26) J. L. Lau and M. K. Dunn, *Bioorg. Med. Chem.,* 2018, **26**, 2700-2707.
(27) L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang, C. Fu, *Signal Transduct. Target.Ther.,* 2022, **7**.
(28) K. Sharma, K. K. Sharma, A. Sharma, R. Jain, *Drug Discov. Today,* 2023, **28**, 103464.
(29) *Peptide Therapeutics Market Size, Share, and Trends 2024 to 2033*. 2024. https://www.precedenceresearch.com/peptide-therapeutics-market#:~:text=The%20global%20peptide%20therapeutics%20market,5.63%25%20from%202024%20to%202033. (accessed 2024 05 October 2024).
(30) G. A. Christou, N. Katsiki, J. Blundell, G. Fruhbeck, D. N. Kiortsis, *Obes. Rev.,* 2019, **20**, 805-815.
(31) A. M. Chao, J. S. Tronieri, A. Amaro, T. A. Wadden, *Trends Cardiovas. Med.,* 2023, **33**, 159-166.
(32) J. P. H. Wilding, R. L. Batterham, S. Calanna, M. Davies, L. F. Van Gaal, I. Lingvay, B. M. McGowan, J. Rosenstock, M. T. D. Tran, T. A. Wadden, *N. Engl. J. Med.,* 2021, **384**, 989-1002.

(33) N. C. Bergmann, M. J. Davies, I. Lingvay, F. K. Knop, *Diabetes, Obes. Metab.,* 2022, **25**, 18-35.
(34) T. Ochi, and H. Roughley, *An exploration of Ozempic*. 2024. https://www.bps.ac.uk/publishing/pharmacology-matters/april-2024/an-exploration-of-ozempic (accessed June 2024).
(35) Z. Zhang and W. Tang, *Acta Pharm. Sin. B,* 2018, **8**, 721-732.
(36) N. Tsomaia, *Eur. J. Med. Chem.,* 2015, **94**, 459-470.
(37) D. J. Craik, and M-W. Kan, *Expert Opin. Drug Discov.,*2021, **16**, 1399-1402.
(38) P. Vlieghe, V. Lisowski, J. Martinez, M. Khrestchatisky, *Drug Discov. Today,* 2010, **15,** 40-56.
(39) S. Sachdeva, *Int. J. Pept. Res. Ther.,* 2017, **23**, 49-60.
(40) K. Fosgerau, and T. Hoffmann, *Drug Discov. Today,* 2015, **20**, 122-128.
(41) C. Recio, F. Maione, A. J. Iqbal, N. Mascolo, V. De Feo, *Front. Pharmacol.,* 2016, **7**.
(42) D. Mcgregor, *Curr. Opin. Pharmacol.,* 2008, **8**, 616-619.
(43) A. C-L. H. Lee, L. Janeele, K. K. Khanna, J-H. Hong, *Int. J. Mol. Sci.,* 2019, **20**, 1-21.
(44) J. K. Dozier and M. D. Distefano, *Int. J. Mol. Sci.,* 2015, **16**, 25831-25864.
(45) L-Q. Pan, H-B. Wang, J. Lai, Y-C. Xu, C. Zhang, S-Q. Chen, *Biomater.,* 2013, **34**, 9115-9123.
(46) F. M. Veronese and G. Pasut, *Drug Discov. Today,* 2005, **10**, 1451-1458.
(47) N. Nischan and C. P. R. Hackenberger, *J. Org. Chem.,* 2014, **79**, 10727-10733.
(48) Y. Gao, M. Joshi, Z. Zhao, S. Mitragotri, *Bioeng. Transl. Med.,* 2023, **9**, e10600.
(49) S. Fishbane, B. Schiller, F. Locatelli, A. C. Covic, R. Provenzano, A. Wiecek, N. W. Levin, M. Kaplan, I. C. Macdougall, C. Francisco, *N. Engl. J. Med.,* 2013, **368**, 307-319.
(50) I. C. Macdougall, R. Provenzano, A. Sharma, R. J. Schmidt, P. E. Pergola, R. I. Zabaneh, S. Tong-Starksen, M. R. Mayo, H. Tang, *N. Engl. J. Med.,* 2013, **368**, 320-332.
(51) H. Schmid, *Drug Eval.,* 2013, **14**, 937-948.
(52) S. M. Hoy, *Drugs,* 2021, **81**, 1423-1430.
(53) M. Griffin, R. Kelly, I. Brindel, L. Maafa, R. Trikha, P. Muus, T. Munir, A. M. Varghese, L. Mitchell, S. Nagumantry, *J. Hematol.,* 2024, **99**, 816-823.
(54) P. Hillmen, J. Szer, I. Weitz, A. Röth, B. Höchsmann, J. Panse, K. Usuki, M. Griffin, J-J. Kiladjian, C. de Castro, *N. Engl. J. Med.,* 2021, **384**, 1028-1037.
(55) A. Deiters, T. A. Cropp, D. Summerer, M. Mukherji, P. G. Schultz, *Bioorg. Med. Chem. Lett.,* 2004, **14**, 5743-5745.
(56) R. Menacho-Melgar, J. S. Decker, J. N. Hennigan, M. D. Lynch, *J. Control. Release,* 2019, **295**, 1-12.
(57) P. Home and P. Kurtzhals, *Drug Eval.,* 2006, **7**, 325-343.
(58) T. M. Chapman and C. M. Perry, *Drugs,* 2004, **64**, 2577-2595.
(59) S. C. L. Gough, S. Harris, V. Woo, M. Davies, *Diabetes Obes. Metab.,* 2012, **15**, 301-309.
(60) S. P. Marso, D. K. McGuire, B. Zinman, N. R. Poulter, S. S. Emmerson, T. R. Pieber, R. E. Pratley, P-M. Haahr, M. Lange, K. Brown-Frandsen, *N. Engl. J. Med.,* 2017, **377**, 723-732.
(61) J. Vora, B. Cariou, M. Evans, J. L. Gross, S. Harris, L. Landstedt-Hallin, A. Mithal, M. R. Rodriguez, L. Meneghini, *Diabetes Res. Clin. Pract.,* 2015, **109**, 19-31.
(62) L. B. Knudsen and J. Lau, *Front. Endocrinol.,* 2019, **10**, 155.
(63) A. Mehta, S. P. Marso, I. J. Neeland, *Obes. Sci. Pract.,* 2016, **3**, 3-14.
(64) D. J. Drucker, A. Dritselis, P. Kirkpatrick, *Nat. Rev. Drug Discov.,* 2010, **9**, 267-268.
(65) S. Dhillon, *Drugs,* 2018, **78**, 275-284.
(66) H. Jiang, X. Zhang, H. Lin, *Sci. Rep.,* 2016, **6**, 24371.
(67) G. Komaniecki, H. Lin, *Front. Cell Dev. Biol.,* 2021, **9**, 717503.
(68) T. Schlatzer, J. Kriegesmann, H. Schröder, M. Trobe, C. Lembacher-Fadum, S. Santner, A. V. Kravchuk, C. F. W. Becker, R. Breinbauer, *J. Am. Chem. Soc.,* 2019, **141**, 14931-14937.
(69) F. L. Zhang and P. J. Casey, *Ann*. *Rev. Biochem.,* 1996, **65**, 241-269.
(70) C. C. Palsuledesai, M. D. Distefano, *ACS Chem. Biol.,* 2014, **10**, 51-62.
(71) H. Jiang, X. Zhang, X. Chen, P. Aramsangtienchai, Z. Tong, H. Lin, *Chem. Rev.,* 2018, **118**, 919-988.
(72) M. D. Resh, *Prog. Lipid Res.,* 2016, **63**, 120-131.

(73) B. Wang, T. Dai, W. Sun, Y. Wei, J. Ren, L. Zhang, M. Zhang, F. Zhou, *Cell. Mol. Immunol.,* 2021, **18**, 878-888.
(74) M. Yuan, Z-H. Song, M-D. Ying, H. Zhu, Q-J. He, B. Yang, J. Cao, *Acta Pharmacol. Sin.,* 2020, **41**, 1005-1015.
(75) D. D. O. Martin, E. Beauchamp, L. G. Bertiaume, *Biochim.,* 2011, **93**, 18-31.
(76) S-H. Yang, P. W. R. Harris, G. M. Williams, M. A. Brimble, *Eur. J. Org. Chem.,* 2016, **15**, 2608-2616.
(77) C. E. Schafmeister, J. Po, G. L. Verdine, *J. Am. Chem.Soc.,* 2000, **122**, 5891-5892.
(78) B. N. Bullock, A. L. Jochim, P. S. Arora, *J. Am. Chem. Soc.,* 2011, **133**, 14220-14223.
(79) A. M. A. Ali, N. Van Oosterwijk, M. R. Groves, A. Dömling, *Comput. Struct. Biotechnol. J.,* 2019, **17**, 263-281.
(80) M. Moiola, M. G. Memeo, P. Quadrelli, *Molecules,* 2019, **24**, 3654.
(81) G. L. Verdine and L. D. Walensky, *Clin. Cancer Res.,* 2007, **13**.
(82) M. M. Madden, A. Muppidi, Z. Li, X. Li, J. Chen, Q. Lin, *Bioorg. Med. Chem. Lett.,* 2011, **21**, 1472-1475.
(83) Y. S. Tan, D. P. Lane, C. S. Verma, *Drug Discov. Today,* 2016, **21**, 1642-1653.
(84) E. Gavathiotis, M. Suzuki, M. L. Davis, K. Pitter, G. H. Bird, S. G. Katz, H-C. Tu, H. Kim, E. H-Y. Cheng, N. Tjandra, *Nature,* 2008, **455**, 1076-1082.
(85) A. O. Frank, B. Vangamudi, M. D. Feldkamp, E. M. Souza-Fagundes, J. W. Luzwick, D. Cortez, E. T. Olejniczak, A. G. Waterson, O. W. Rossanese, W. J. Chazin, *J. Med. Chem.,* 2014, **57**, 2455-2461.
(86) Y-Q. Long, S-X. Huang, Z. Zawahir, Z-L. Xu, H. Li, T. W. Sanchez, Y. Zhi, S. De Houwer, F. Christ, Z. Debyser, *J. Med. Chem.,* 2013, **56**, 5601-5612.
(87) M. L. Stewart, E. Fire, A. E. Keating, L. D. Walensky, *Nat. Chem. Biol.,* 2010, **6**, 595-601.
(88) *Fmoc-Lys-OH*. Merck, 2024. https://www.sigmaaldrich.com/GB/en/product/mm/852023?srsltid=AfmBOopUjF3YPO7E9N_LUbzVBmmWeCvc9pcpOoDd8f_Ct2Ri-atmJRWY (accessed July 2024).
(89) *Fmoc-azidolysine*. Merck, 2024. https://www.sigmaaldrich.com/GB/en/product/mm/852326 (accessed July 2024).
(90) F. R. Smith, D. Meehan, R. C. Griffiths, H. J. Knowles, P. Zhang, H. E. L. Williams, A. J. Wilson, N. J. Mitchell, *Chem. Sci.,* 2024, **15**, 9612-9619.
(91) H. Grus, R. C. Feiner, R. Mseya, D. C. Schroder, M. Jewginski, K. M. Muller, R. Latajka, A. Marion, N. Sewald, *Beilstein J. Org. Chem.,* 2022, **18**, 1-12.
(92) Y-W. Kim, T. N. Grossmann, G. L. Verdine, *Nat. Protoc.,* 2011, **6**, 761-771.
(93) A. L. P. Lourenço, T. B. Rios, Á. P. da Silva, O. L. Franco, M. H. S. Ramada, *Antibiotics,* 2023, **12**, 1400.
(94) L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine, S. J. Korsmeyer, *Science,* 2004, **305**, 1466-1470.
(95) N. N. Danial and S. J. Korsmeyer, *Cell,* 2004, **116**, 205-219.
(96) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.,* 2002, **41**, 2596-2599.
(97) S. Cantel, I. A. Le Chevalier, M. Scrima, J. J. Levy, R. D. Dimarchi, P. Rovero, J. A. Halperin, A. M. D'Ursi, A. M. Papini, M. Chorev, *J. Org. Chem.,* 2008, **73**, 5663-5674.
(98) S. A. Kawamoto, A. Coleska, X. Ran, H. Yi, C-Y. Yang, S. Wang, *J. Med. Chem.,* 2012, **55**, 1137-1146.
(99) M. Scrima, I. A. Le Chevalier, P. Rovero, A. M. Papini, M. Chorev, A. M. D'Ursi, *Eur. J. Org. Chem.,* 2010, **3**, 446-457.
(100) O. Torres, D. Yüksel, M. Bernardina, K. Kumar, D. Bong, *ChemBioChem,* 2008, **9**, 1701-1705.
(101) Y. H. Lau, P. De Andrade, S-T. Quah, M. Rossmann, L. Laraia, N. Sköld, T. J. Sum, P. J. E. Rowling, T. L. Joseph, C. Verma, *Chem. Sci.,* 2014, **5**, 1804-1809.
(102) V. Gaikwad, A. R. Choudhury, R. Chakrabarti, *Biophys. Chem.,* 2024, **307**, 107197.
(103) B. Liu, W. Zhang, S. Gou, H. Huang, J. Yao, Z. Yang, H. Liu, C. Zhong, B. Liu, J. Ni, *J. Pept. Sci.,* 2017, **23**, 824-832.

(104) P. T. Tran, C. Ø. Larsen, T. Røndbjerg, M. De Foresta, M. B. A. Kunze, A. Marek, J. H. Løper, L-E. Boyhus, A. Knuhtsen, K. Lindorff-Larsen, *Chem. Eur. J.,* 2017, **23**, 3490-3495.
(105) A. A. Aimetti, R. K. Shoemaker, C-C. Lin, K. S. Anseth, *Chem. Comm.,* 2010, **46**, 4061-4063.
(106) X. Li, S. Chen, W-D. Zhang, H-G. Hu, *Chem. Rev.,* 2020, **120**, 10079-10144.
(107) C. M. Grison, G. M. Burslem, J. A. Miles, L. K. A. Pilsl, D. J. Yeo, Z. Imani, S. L. Warriner, M. E. Webb, A. J. Wilson, *Chem. Sci.,* 2017, **8**, 5166-5171.
(108) A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y-S. Lin, B. L. Pentelute, *J. Am. Chem. Soc.,* 2013, **135**, 5946-5949.
(109) D. P. Fairlie and A. Dantas De Araujo, *Biopolymers,* 2016, **106**, 843-852.
(110) H. Jo, N. Meinhardt, Y. Wu, S. Kulkarni, X. Hu, K. E. Low, P. L. Davies, W. F. DeGrado, D. C. Greenbaum, *J. Am. Chem. Soc.,* 2012, **134**.
(111) Y. Wang and D. H-C. Chou, *Angew. Chem. Int. Ed.,* 2015, **54**, 10931-10934.
(112) J. C. Phelan, N. J. Skelton, A. C. Braisted, R. S. McDowell, *J. Am. Chem. Soc.,* 1997, **119**, 455-460.
(113) M. G. Ricardo, A. M. Ali, J. Plewka, E. Surmiak, B. Labuzek, C. G. Neochoritis, J. Atmaj, L. Skalniak, R. Zhang, T. A. Holak, *Angew. Chem. Int. Ed.,* 2020, **59**, 5235-5241.
(114) M. G. Ricardo, F. E. Morales, H. Garay, O. Reyes, D. Vasilev, L. A. Wessjohann, D. G. Rivera, *Org. Biomol. Chem.,* 2015, **13**, 438-446.
(115) X. Cai, W. Zheng, X. Shi, L. Chen, Z. Liu, Z. Li, *Mol. Pharmaceutics,* 2018, **15**, 5646-5652.
(116) M. R. Ghadiri and A. K. Fernholtz, *J. Am. Chem. Soc.,* 1990, **112**, 9633-9635.
(117) R. J. Todd, M. E. Van Dam, D. Casimiro, B. L. Haymore, F. H. Arnold, *Proteins,* 1991, **10**, 156-161.
(118) S. Learte-Aymamí, C. Vidal, A. Gutiérrez-González, J. L. Mascareñas, *Angew. Chem. Int. Ed.,* 2020, **59**, 9149-9154.
(119) S. Learte-Aymamí, N. Curado, J. Rodríguez, M. E. Vázquez, J. L. Mascareñas, *J. Am. Chem. Soc.,* 2017, **139**, 16188-16193.
(120) K. M. Makwana and R. Mahalakshmi, *Org. Lett.,* 2015, **17**, 2498-2501.
(121) E. Y-L. Hui, B. Rout, Y. S. Tan, C. S. Verma, K-P. Chan, C. W. Johannes, *Org. Biomol. Chem.,* 2018, **16**, 389-392.
(122) X. Li, W. D. Tolbert, H-G. Hu, N. Gohain, Y. Zou, F. Niu, W-X. He, W. Yuan, J-C. Su, M. Pazgier, *Chem. Sci.,* 2019, **10**, 1522-1530.
(123) F. M. Brunel and P. E. Dawson, *Chem. Comm.,* 2005, 2552-2554.
(124) F. M. Brunel, M. B. Zwick, R. M. F. Cardoso, J. D. Nelson, I. A. Wilson, D. R. Burton, P. E. Dawson, *J. Virol.,* 2006, **80**, 1680-1687.
(125) M. Todorovic, K. D. Schwab, J. Zeisler, C. Zhang, F. Bénard, D. M. Perrin, *Angew. Chem. Int. Ed.,* 2019, **58**, 14120-14124.
(126) M. C. Morejón, A. Laub, B. Westermann, D. G. Rivera, L. A. Wessjohann, *Org. Lett.,* 2016, **18**, 4096-4099.
(127) J. E. Montgomery, J. A. Donnelly, S. W. Fanning, T. E. Speltz, X. Shangguan, J. S. Coukos, G. L. Greene, R. E. Moellering, *J. Am. Chem. Soc.,* 2019, **141**, 16374-16381.
(128) J. W. Taylor, *Biopolymers,* 2002, **66**, 49-75.
(129) G. Osapay and J. W. Taylor, *J. Am. Chem. Soc.,* 1992, **114**, 6966-6973.
(130) N. E. Shepherd, G. Abbenante, D. P. Fairlie, *Angew. Chem. Int. Ed.,* 2004, **43**, 2687-2690.
(131) M. E. Houston, C. L. Gannon, C. M. Kay, R. S. Hodges, *J. Pept. Sci.,* 1995, **1**, 274-282.
(132) F. E. Morales, H. E. Garay, D. F. Muñoz, Y. E. Anugusto, A. J. Otero-González, O. R. Acosta, D. G. Rivera, *Org. Lett.,* 2015, **17**, 2728-2731.
(133) K. Fujimoto, N. Oimoto, K. Katsuno, M. Inouye, *Chem. Comm.,* 2004, 1280-1281.
(134) K. Fujimoto, M. Amano, H. Horibe, M. Inouye, *Org. Lett.,* 2006, **8**, 285-287.
(135) K. Fujimoto, M. Kajino, I. Sakaguchi, M. Inouye, *Chem. Eur. J.,* 2012, **18**, 9834-9840.
(136) M. M. Madden, C. I. R. Vera, W. Song, Q. Lin, *Chem. Comm.,* 2009, **37**, 5588-5590.
(137) A. M. Felix, E. P. Heimer, C-T. Wang, T. J. Lambros, A. Fournier, T. F. Mowles, S. Maines, R. M. Campbell, B. B. Wegrzynski, V. Toome, *Int. J. Pept. Protein Res.,* 1988, **32**, 441-454.
(138) K. Fujimoto, M. Kajino, M. Inouye, *Chem. Eur. J.,* 2008, **14**, 857-863.

(139) H. Li, Y. Hu, Q. Pu, T. He, Q. Zhang, W. Wu, X. Xia, J. Zhang, *J. Med. Chem.,* 2020, **63**, 4081-4089.
(140) Y. Hu, H. Li, R. Qu, T. He, X. Tang, W. Chen, L. Li, H. Bai, C. Li, W. Wang, *J. Med. Chem.,* 2022, **65**, 579-591.
(141) G. Lautrette, F. Touti, H. G. Lee, P. Dai, B. L. Pentelute, *J. Am. Chem. Soc.,* 2016, **138**, 8340-8343.
(142) H. G. Lee, G. Lautrette, B. L. Pentelute, S. L. Buchwald, *Angew. Chem. Int. Ed.,* 2017, **56**, 3177-3181.
(143) M. G. Ricardo, D. Llanes, L. A. Wessjohann, D. G. Rivera, *Angew. Chem. Int. Ed.,* 2018, **58**, 2700-2704.
(144) K. C. A. Garber and E. E. Carlson, *ACS Chem. Biol.,* 2013, **8**, 1671-1676.
(145) S. W. Kwon, S. C. Kim, J. Jaunbergs, J. R. Falck, Y. Zhao, *Mol. Cell. Proteomics,* 2003, **2**, 242-247.
(146) M. Trmčić, F. L. Chadbourne, P. M. Brear, P. W. Denny, S. L. Cobb, D. R. W. Hodgson, *Org. Biomol. Chem.,* 2013, **11**, 2660.
(147) L. P. Conway, R. J. Delley, J. Neville, G. R. Freeman, H. J. Maple, V. Chan, A. J. Hall, D. R. W. Hodgson, *RSC Adv.,* 2014, **4**, 38663.
(148) L. P. Conway, S. Mikkola, A. C. O'Donoghue, D. R. W. Hodgson, *Org. Biomol. Chem.,* 2016, **14**, 7361-7367.
(149) T. Ruman, K. Dlugopolska, A. Jurkiewicz, D. Rut, T. Fraczyk, J. Ciesla, A. Les, Z. Szewczuk, W. Rode, *Bioorg. Chem.,* 2010, **38**, 74-80.
(150) M. Lasker, C. D. Bui, P. G. Besant, K. Sugawara, P. Thai, G. Medzihradszky, C. W. Turck, *Protein Sci.,* 1999, **8**, 2177-2185.
(151) M. C. Pirrung, K. D. James, V. S. Rana, *J. Org. Chem.,* 2000, **65**, 8448-8453.
(152) M. Trmčić and D. R. W. Hodgson, *Chem. Comm.,* 2011, **47**, 6156-6158.
(153) R. J. Delley, A. M. C. O'Donoghue, D. R. W. Hodgson, *J. Org. Chem.,* 2012, **77**, 5829-5831.
(154) D. R. Lide, *Handbook of Chemistry and Physics*; CRC Press, 1991.
(155) A. Isidro-Llobet, M. Álvarez, F. Albericio, *Chem. Rev.,* 2009, **109**, 2455-2504.
(156) R. F. Chen, *Anal. Biochem.,* 1968, **25**, 412-416.
(157) C. Gros and B. Labouesse, *Eur. J. Biochem.,* 1969, **7**, 463-470.
(158) M. M. Achmatowicz, O. R. Thiel, J. T. Colyer, J. Hu, M. V. S. Elipe, J. Tomaskevitch, J. S. Tedrow, R. D. Larsen, *Org. Process Res. Dev.,* 2010, **14**, 1490-1500.
(159) A. T. P. Carvalho, A. M. C. O'Donoghue, D. R. W. Hodgson, S. C. L. Kamerlin, *Org. Biomol. Chem.,* 2015, **13**, 5391-5398.
(160) E. V. Anslyn and D. A. Dougherty, *Modern Physical Organic Chemistry*; University Science Books, 2006.
(161) M. I. Page and W. P. Jencks, *Proc. Natl. Acad. Sci,* 1971, **68**, 1678-1683.
(162) L. M. Rellick and W. J. Becktel, *Biopolymers,* 1996, **42**, 191-202.
(163) J. Pontius, J. Richelle, S. J. Wodak, *J. Mol. Biol.,* 1996, **264**, 121-136.
(164) Y. Zhang, Y. Cao, Y. Chi, S. Chen, X. Zeng, Y. Liu, G. Tang, Y. Zhao, *Adv. Synth. Catal.,* 2022, **364**, 2091-2265.
(165) Y. Segall, G. B. Quistad, S. E. Sparks, J. E. Casida, *Chem. Res. Toxicol.,* 2003, **16**, 350-356.
(166) P. D. Jadhav, J. Shen, P-G. Burnett, J. Yang, R. Sammynaiken, M. J. T. Reaney, *RSC Adv.,* 2021, **11**, 20859-20864.
(167) L. Moroder, *J. Pept. Sci.,* 2005, **11**, 187-214.
(168) D. B. Cowie, G. N. Cohen, E. T. Bolton, H. de Robichon-Szulmajster, *Biochim. Biophys. Acta,* 1959, **34**, 39-46.
(169) H. Cai, Y. Liang, L. Huang, J. Wang, *Spectrochim. Acta A: Mol Biomol Spectrosc,* 2022, *277*.
(170) K. Rosenheck and P. Doty, *Proc. Natl. Acad. Sci.,* 1961, ***47***, 1775-1785.
(171) S. Prasad, I. Mandal, S. Singh, A. Paul, B. Mandal, R. Venkatramani, R. Swaminathan, *Chem. Sci.,* 2017, **8**, 5416-5433.
(172) F. Zhang, O. Sadovski, S. J. Xin, G. A. Woolley, *J. Am. Chem. Soc.,* 2007, **129**, 14154-14155.
(173) D. L. Comins and S. P. Joseph, *N,N*-Dimethylformamide. In *Encyclopedia of Reagents for Organic Synthesis*, John Wiley & Sons, 2001.

(174) E. Nemati-Kande and H. Shekaari, *Fluid Ph. Equilib.,* 2013, **360**, 357-366.
(175) Trmčić, M. The development of thiophosphoryl 'click' chemistry. Durham University, 2009.