Development of Synthetic Peptides as Functional Probes for the Lipoprotein Biosynthesis Pathway in Bacteria

Katherine Bowen, B.A. (Mod.)

June 2021

Trinity College Dublin
The University of Dublin

Based on research carried out under the supervision of
Prof. Eoin Scanlan

A thesis submitted to the School of Chemistry, Trinity College Dublin, The University of Dublin, for the degree of Doctor of Philosophy

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work, other than where acknowledged.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Katherine Bowen

Abstract

Protein lipidation is a critical post-translational modification that occurs in all bacteria. The essential enzymes involved in this modification are exclusive to bacteria and have demonstrated potential as antibiotic targets. This thesis, entitled "Development of Synthetic Peptides as Functional Probes for the Lipoprotein Biosynthesis Pathway in Bacteria", explores the synthesis of functionalised lipopeptide-based probes to investigate the enzymes involved in this pathway.

The work described in this thesis is divided into 7 chapters. Chapter 1 provides a brief historical overview of the antibiotic era and highlights the importance of continual antibiotic development. The structural variations between lipoproteins that are native to Gram-negative and Gram-positive bacteria are presented and the enzymes responsible for these modifications are introduced. Previous investigations into these enzymes as antibiotics targets are reviewed with a focus on the barriers to further development. Following this, existing synthetic and recombinant routes to access bacterial lipoproteins and lipopeptides are summarised. The chapter concludes by detailing advances in the field of peptide science that have provided novel opportunities in the synthesis of functionalised lipoprotein derivatives.

In Chapter 2 we develop a novel Förster resonance energy transfer (FRET)-based lipopeptide probe which is ultimately utilised in a high-throughput assay for lipoprotein signal peptidase II (LspA). This chapter begins by investigating the synthesis of lipopeptides *via* the linear incorporation of a diacylglycerol-modified cysteine residue into a peptide sequence through solid-phase peptide synthesis. This initially involves the production of a lipidated cysteine residue. To this aim, two synthetic routes to this modified amino acid are presented. The introduction of this residue into a peptide sequence is explored, in addition to dual functionalisation of the peptide with multiple fluorophores. Various peptide sequences are prepared and screened for activity with LspA to facilitate the development of a high-throughput FRET-based assay for this endopeptidase.

In the second part of this work, detailed in Chapters 3 and 4, we utilise the synthetic strategy presented in Chapter 2 to access peptides containing an N-terminal diacylglycerol-modified cysteine residue. Such peptides mimic the product of the LspA peptidase reaction and thus are primed to probe the activity of the subsequent enzymes

in the biosynthetic pathway, namely apolipoprotein *N*-acyltransferase (Lnt) and lipoprotein intramolecular transacylase (Lit). In Chapter 3 we investigate the synthesis of lipidated cysteine derivatives to explore the minimal structural requirements for Lnt recognition. These insights facilitate the design of a fluorescent peptide substrate which is used to develop a novel functional assay for this enzyme. In Chapter 4, we utilise synthetic derivatives of this Lnt substrate to probe the mechanism of acyl transfer mediated by Lit. A novel assay is developed for this recently discovered acyltransferase, and we unambiguously characterise the product of the Lit reaction. To achieve this, a deuterium labelled fluorescent substrate of Lit is produced *via* regioselective acylation of the glyceryl functionality and selective transfer of the deuterium labelled *sn*-2 acyl chain by Lit is demonstrated using nuclear magnetic resonance (NMR) analysis. This chapter concludes by utilising paramagnetic NMR to show the capacity of Cu(II) to selectively interact at the N-terminus of apolipopeptides.

In Chapter 5 we detail a novel synthetic route to diacylglycerol-modified peptides. The convergent methodology we report involves the late-stage modification of peptides *via* liposome-mediated thiol-Michael addition of a lipidated thiol to a dehydroalanine containing peptide. This affords a native diacylglycerol modification at the resultant cysteine residue. Additionally, we explore further applications of dehydroalanine, utilising this amino acid in a novel peptide ligation methodology to extend the concept of native chemical ligation. A novel on-resin route to thioester formation through β,γ-*C*,*S* thiol-Michael addition of thioacids to dehydroalanine containing peptides is investigated. This reaction yields a cysteinyl thioester which undergoes *S*-to-*N* acyl transfer through a 5-membered cyclic intermediate to afford a native peptide bond at the ligation site.

In Chapter 6 we report the overall conclusions of this work and outline the future direction of this research.

Finally, in Chapter 7 the experimental procedures used in this work are provided along with the associated characterisation for the compounds synthesised.

Abbreviations

3-NT	3-nitrotyrosine
Abz	2-aminobenzamide
Ac	acetyl
Ac_2O	acetic anhydride
AFPS	automated flow peptide synthesis
Ala; A	alanine
AM	aminomethyl
AMR	antimicrobial resistance
APC	antigen presenting cell
aq.	aqueous
Ar	aromatic
Arg; R	arginine
Asn; N	asparagine
Asp; D	aspartic acid
Boc	*tert*-butyloxycarbonyl
Boc_2O	di-*tert*-butyl dicarbonate
BOP	benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate
Bu	butyl
calcd.	calculated
CF	cystic fibrosis
CMC	critical micelle concentration
d	doublet
Dabsyl	4-[[4′-(*N*,*N*-dimethylamino)phenyl]diazenyl]benzoic acid
DAG	diacylglycerol
DCC	*N,N'*-dicyclohexylcarbodiimide
dd	doublet of doublets
Dha	dehydroalanine
DIC	*N,N'*-diisopropylcarbodiimide
DIPEA	*N*,*N*-diisopropylethylamine

DKP	diketopiperazine
DMAP	4-(dimethylamino)pyridine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethylsulfoxide
DVB	divinylbenzene
EDANS	5-(2-aminoethyl)aminonaphthalene sulfonic acid
EDC·HCl	*N*-(3-dimethylaminopropyl)-*N′*-ethylcarbodiimide hydrochloride
Equiv.	equivalents
ESI	electrospray ionisation
Et	ethyl
EtOAc	ethyl acetate
FC-12	*n*-dodecylphosphocholine
Fmoc	fluorenylmethyloxycarbonyl
Fmoc-OSu	*N*-(9-fluorenylmethoxycarbonyloxy)succinimide
FRET	Förster resonance energy transfer
FSL-1	fibroblast-stimulating lipopeptide
GC	guanine-cytosine
Gln; Q	glutamine
GLP-1	glucagon-like peptide 1
Glu; E	glutamic acid
Gly; G	glycine
HCl	hydrochloric acid
HCTU	*O*-(1*H*-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
hex	hexane
His; H	histidine
HIV	human immunodeficiency virus
Hmb	2-hydroxy-4-methoxybenzyl
HMBC	heteronuclear multiple bond correlation
HMPA	hexamethylphosphoric triamide
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole monohydrate

HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single quantum coherence
HTS	high-throughput screening
ICP	inhibitor of cysteine protease
IFE	inner filter effect
IL	interleukin
Ile; I	isoleucine
IM	inner membrane
IR	infrared
J	coupling constant
Leu; L	leucine
Lgt	lipoprotein diacylglyceryl transferase
Lit	lipoprotein intramolecular transacylase
LMNG	lauryl maltose neopentyl glycol
Lns	lipoprotein *N*-acylation transferase system
Lnt	apolipoprotein *N*-acyltransferase
LntEco	*Escherichia coli* apolipoprotein *N*-acyltransferase
LntPae	*Pseudomonas aeruginosa* apolipoprotein *N*-acyltransferase
Lol	localisation of lipoproteins
LP	lipoprotein
LPS	lipopolysaccharide
LspA	lipoprotein signal peptidase II
LspMrs	methicillin-resistant *Staphylococcus aureus* lipoprotein signal peptidase II
LspPae	*Pseudomonas aeruginosa* lipoprotein signal peptidase II
Lys; K	lysine
Lyso-NBD-PE	1-oleoyl-2-hydroxy-*sn*-glycero-3-phosphoethanolamine-*N*-(7-nitro-2-1,3-benzoxadiazol-4-yl)
M	mass ion
m	multiplet
m/z	mass per unit charge
M59	squalene

MALDI	matrix assisted laser desorption ionization
MDR	multi-drug resistant
MeCN	acetonitrile
MeOH	methanol
MES	2-(*N*-morpholino)ethanesulfonic acid
MHC	major histocompatibility complex
MRSA	methicillin-resistant *Staphylococcus aureus*
NCL	native chemical ligation
NDB-PE	1-oleoyl-2-12-[(7-nitro-2-1,3-benzoxadiazol-4-yl)amino]dodecanoyl-*sn*-glycero-3-phosphoethanolamine
NMM	*N*-methylmorpholine
NMP	*N*-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
M^{pro}	main protease
OAc	acetate
OM	outer membrane
Pal	palmitic
Pam	palmitoyl
Pam_2Cys	2,3-dipalmitoyl-*S*-glycerylcysteine-serine-$(lysine)_4$
PBu_3	tributylphosphine
PE	phosphatidylethanolamine
PEG	polyethylene glycol
PG	phosphatidylglycerol
Phe; F	phenylalanine
ppm	parts per million
Pro; P	proline
PRR	pattern recognition receptor
PTM	post-translational modification
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
qC	quaternary carbon
QM/MM	quantum mechanics/molecular mechanics
Q-Tof	quadrupole time-of-flight

RET	resonance energy transfer
$\boldsymbol{R_f}$	retention factor
RP	reversed phase
rt	room temperature
s	singlet
SAR	structure–activity relationship
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
sat.	saturated
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel
Sec	general secretory
Ser; S	serine
***sn*-1**	stereospecific numbering 1 (glyceryl C-1)
***sn*-2**	stereospecific numbering 2 (glyceryl C-2)
spp	species
SPPS	sold-phase peptide synthesis
t	triplet
TAT	twin-arginine translocation
TB	tuberculosis
TCEP	tris(2-carboxyethyl)phosphine
TES	triethylsilane
TFA	trifluoroacetic acid
TFE	trifluoroethanol
THF	tetrahydrofuran
Thr; T	threonine
TLC	thin layer chromatography
TLR	toll-like receptor
Tmob	(2,4,6-trimethoxyphenyl)methanethiol
TNF	tumor necrosis factor
TOCSY	total correlation spectroscopy
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
Trityl; Trt	triphenylmethane
Trp; W	tryptophan

Tyr; Y	tyrosine
Val; V	valine
WHO	world health organization

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Eoin Scanlan for his invaluable guidance throughout the course of this work. From the initial opportunity to join the lab as an undergraduate summer student in 2015, to establishing collaborations and networking opportunities throughout my PhD, I could not have asked for a more supportive and encouraging mentor during this stage of my career.

I would like to acknowledge the Irish Research Council for providing the funding for this work through the Government of Ireland Postgraduate Scholarship Programme.

The discussion of this work includes experiments performed by our collaborators in the group of Prof. Martin Caffrey in the School of Biochemistry, Trinity College Dublin. I would like to express my sincere gratitude to Prof. Martin Caffrey and the members of his research group, especially Dr. Samir Olatunji, Dr. Dietmar Weichert and Dr. Jonathan Bailey. Their creativity and enthusiasm have made this project possible and defined the direction of my PhD.

I am so grateful for the support of Dr. John O'Brien who has taught me everything I know about NMR. John's expertise, curiosity, and endless enthusiasm have broadened the scope of what was possible with this project. I would also like to thank Dr. Manuel Ruether for providing an abundance of NMR support and Dr. Gary Hessman for his invaluable assistance with MS analysis. I would also like to thank Dr. Rob Elmes from Maynooth University who introduced our lab to SPPS back in 2017 and for the many hours he spent assisting with the HPLC analysis of some nightmare compounds in the early days of this project.

I have had the opportunity to directly collaborate with some incredibly talented members of the Scanlan group during my PhD. To Dr. Rita Petracca, thank you for your invaluable input on the Dha project and for showing me the ways of chemistry while somehow making me fall in love with research in the process. To Len van Dalsen, thank you for sharing the many ups and downs of both the FRET and cotton wool-based projects. Your infectious enthusiasm for all things chemistry made even the spiciest of columns so entertaining. To all of the incredible past and present members of the Scanlan group, Danielle, Lauren, Moss, Helen, Rita, Josh, Dylan, Len, Mark, Susie, Pierre, Alby, Glenna, and Nikitas, I could not have imaged a better bunch of people to share this experience with. Thank you all for providing such a welcoming and wholesome atmosphere where

everyone in our weird and wonderful lab group can thrive. A huge shoutout to the memeing dream team, Le'en and Mark, for providing the many much-needed fits of laughter that brightened up even the bleakest of days. To our much-loved honorary lab members – Susie, thank you for always being ready with a stream of puppy photos (and gin) to fix every bad NMR and for keeping me (almost) sane during lockdown, and Ellen, thank you for always being on hand with an abundance of tea and chats to get me through the late nights of writing.

Finally, I would like to thank my family and friends, particularly my mother and brother, for their unwavering support and encouragement that have made the last few years possible. This thesis is dedicated to my father, Dave Bowen, who inspires my pursuit of science.

Publications

S. Olatunji, K. Bowen, C. Y. Huang, D. Weichert, W. Singh, I. Tikhonova, E. M. Scanlan, V. Olieric, M. Caffrey, Structural basis of the membrane intramolecular transacylase reaction responsible for lyso-form lipoprotein synthesis, *Nat. Commun.*, in press.

S. Olatunji, X. Yu, J. Bailey, C. Y. Huang, M. Zapotoczna, K. Bowen, M. Remškar, R. Müller, E. M. Scanlan, J. Geoghegan, V. Olieric, M. Caffrey, Structures of lipoprotein signal peptidase II from *Staphylococcus aureus* complexed with antibiotics globomycin and myxovirescin, *Nat. Commun.*, 2020, **11**, 140.

R. Petracca, K. Bowen, L. McSweeney, S. O'Flaherty, V. Genna, B. Twamley, M. Devocelle, E. M. Scanlan, Chemoselective Synthesis of N-Terminal Cysteinyl Thioesters *via* β,γ-*C,S* Thiol-Michael Addition, *Org. Lett.*, 2019, **21**, 3281-3285.

Contents

Chapter 2: The Development of Synthetic Substrates of Lipoprotein Signal Peptidase II

Chapter 3: Development of Synthetic Substrates of Apolipoprotein *N*-Acyltransferase

Chapter 1

Introduction

1.0 Introduction

This chapter will outline the urgent need for antibiotic development to address the public health threat caused by multi-drug resistant (MDR) pathogens. The importance of exploring novel antibiotic targets will be discussed, with an emphasis on how synthetic access to bacterial lipopeptides can enhance our ability to probe new bacterial targets. The challenges posed by the synthesis of such lipopeptides and lipoproteins will be summarised and contemporary approaches to the synthesis of these hydrophobic peptides will be introduced.

1.1 Antibiotic Development

Antibiotics are now considered a cornerstone of modern medicine. Prior to the clinical introduction of these therapeutics, infectious diseases were a leading cause of mortality both in Europe and in the United States of America.[1] In addition to the treatment of infectious diseases, the ability to control bacterial infection is critical to the success of several fields of medicine including surgery, immunosuppressant chemotherapy, and organ replacement. The discovery of antibiotics at the start of the 20th century proved to be revolutionary for public health. The prolific era of antibiotic discovery which followed led to the widespread belief that the threat of bacterial infections to human life had been overcome. However, the rapid emergence of antimicrobial resistance greatly diminished the efficacy of many antibiotics. With the exhaustion of the once fruitful antibiotic discovery platforms coupled with the lack of success in recent novel drug discovery, the field of antibiotic development has failed to contain the prolific emergence of resistant microbes. As a result, antibiotic resistant pathogens are now classified as one of the top three threats to public health by the World Health Organisation (WHO).[2] Providing effective treatment options for bacterial infections through novel antibiotic discovery is essential to combat this threat. Overcoming the current stagnation in the field of antibiotic discovery requires the identification of new targets and inhibitors, in addition to the implementation of strategies and policies to slow the emergence of resistance.

1.1.1 Origins of the Antibiotic Era

In 1899, following advances in microbiology by Lister, Koch and Pasteur,[3, 4] the first clinical use of an antibiotic was reported by Emmerich and Löw.[5] Pyocyanase, an extract from *Pseudomonas aeruginosa*, was observed to inhibit the growth of other bacterial species. This extract was applied topically to wounds, however, despite more recent

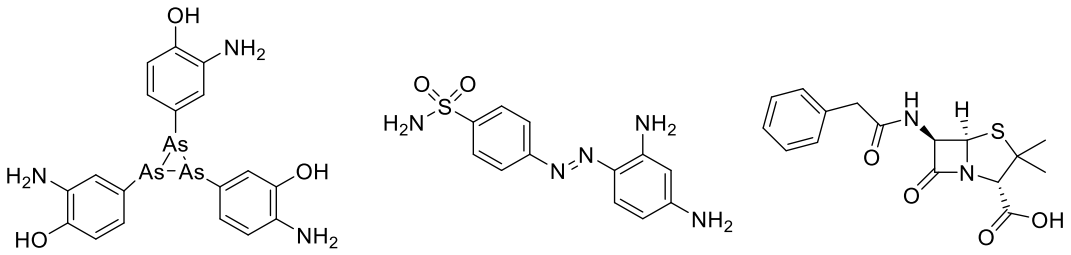
studies which have confirmed that *P. aeruginosa* does produce compounds with antimicrobial activity,[6] this approach was not consistently clinically effective.

The work of Paul Ehrlich in 1904 on targeted medicines, known as the 'magic bullet' concept, triggered the next significant advance in the development of therapeutic antimicrobials. Noting the variation in a dye's affinity for different cellular tissues, Ehrlich proposed that receptors exclusive to microbes could similarly be selectively targeted by a drug. By 1909 Ehrlich, Bertheim, and Hata had developed a screening programme through which the synthetic drug salvarsan (arsphenamine, **1**) was discovered.[7] This anti-syphilis compound selectively targets the pathogenic organism *Treponema pallidum* with limited host toxicity. Despite difficulties with stability and administration, this drug represented a major advance in the treatment of syphilis which had previously required highly toxic and ineffective inorganic mercury salts.[8] Ehrlich's screening concept, consisting of multiple iterations of synthetic structural modification to optimise the activity of a drug, has become the basis of modern medicinal chemistry.[7]

Inspired by Ehrlich's work, Domagk, Mietzsch and Klarer (Bayer, I.G. Farben, Germany) developed prontosil (**2**), a substituted diphenyldiazene with anti-streptococcal activity.[9] This potent antimicrobial was introduced into the clinic in 1935, at which time prontosil was reclassified as a pro-drug with the activity attributed to a metabolite, 4-aminobenzenesulfonamide. The well-established synthesis of this simple compound resulted in the rapid production and testing of thousands of derivatives, there-by establishing the influential sulfa-drug class of antibiotics. Sulfa-drugs became the most commonly prescribed class of drug until the clinical introduction of penicillin in 1944.[10] The extensive early use of sulfonamides has resulted in widespread resistance to this class of antimicrobials, however, they are still used sparingly today.[11] In 1939, Gerhard Domagk was awarded the Nobel Prize in Physiology or Medicine "for the discovery of the antibacterial effects of prontosil".

Concurrently, an influential breakthrough in the field was being made by Alexander Fleming with the discovery of penicillin (**3**) in 1928.[12] Although not the first time inhibition of bacterial growth by mould had been observed,[13] Fleming identified and isolated the mould product penicillin, a β-lactam with activity against Gram-positive organisms.[14] Issues with stability and production scale-up hindered clinical introduction but such limitations were overcome by Chain, Florey and Heatley in 1940, resulting in

large scale bioproduction and widespread clinical use by 1945.[15] Fleming, Chain, and Florey were awarded the Nobel Prize in Physiology or Medicine in 1945 "for the discovery of penicillin and its curative effect in various infectious diseases". The structures of the early antibiotics discussed are shown in **Fig. 1.1**.



Salvasaran **1** | Prontosil **2** | Penicillin G **3**

Figure 1.1. The structure of 3 early antibiotics.

1.1.2 The Golden Era of Antibiotic Discovery

From the 1940s to 1960s novel antibiotic discovery was highly successful. Many new classes were identified, the majority of which were secondary metabolites produced by soil-dwelling actinomycetes. Novel antibiotics were classed by their chemical structure with many classes exhibiting similar mechanisms of action. Most of these drugs acted through one of three modes: shutting down cell wall biosynthesis, RNA gyrase and topoisomerase IV inhibition or ribosome function inhibition.[10]

Streptomycin, produced by the actinomycete *Streptomyces griseus*, was isolated in 1944 by Waksman and Schatz.[16] Active against penicillin resistant Gram-negative bacteria and *Mycobacterium tuberculosis*, streptomycin revolutionised the treatment of tuberculosis (TB), earning Waksman the Nobel Prize in Physiology or Medicine in 1952.

The systematic approach, pioneered by Waksman, of screening actinomycetes against susceptible bacteria was widely adopted by the pharmaceutical industry over the next two decades. This platform yielded the major classes of antibiotics which represent the backbone of the golden era of development.[17] The discovery timeline of the major classes is illustrated in **Fig. 1.2**.

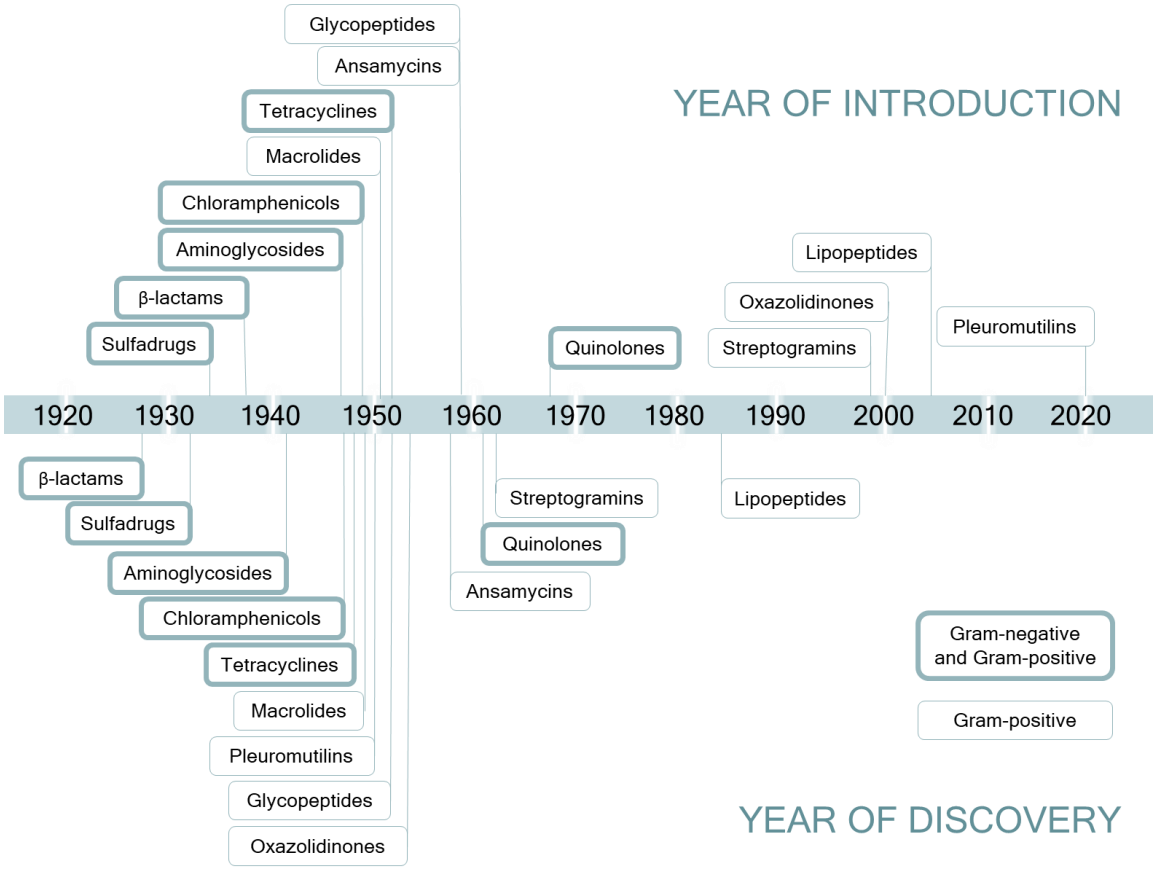


Figure 1.2. **Timeline of the discovery (bottom) and clinical introduction (top) of the major classes of antibiotics.** Broad spectrum antibiotics are emboldened. Adapted from K. Lewis, *Cell*, 2020, **181**, 29-45.[18]

Despite the initial success of the field, the rate of discovery slowed as outputs from the Waksman platform were exhausted. Research emphasis moved to semi- and fully synthetic antibiotic discovery programmes. Much focus was placed on the synthetic modification of existing scaffolds to improve the pharmacokinetic and pharmacodynamic properties of a drug. These platforms provided access to successive generations of each antibiotic class, a critical tool in the arms-race against acquired drug resistance in bacteria. However, the increased difficulty of identifying novel drugs and the inevitable emergence of resistance to all classical antimicrobial drugs has effectively stalled successful development.

1.1.3 Antimicrobial Resistance

The ability of microbes to evolve strategies to evade the effects of antibiotics is an ancient phenomenon. Acquired resistance against clinical antimicrobials was first noted by Fleming[12] and confirmed by Abraham and Chain in 1940 who identified that the presence

of penicillinase in bacteria conferred penicillin-resistance.[19] The accelerated emergence of microorganisms resistant to penicillin following its clinical introduction foreshadowed the global emergence of many MDR pathogens. Resistance now exists against every antibiotic on the market.[10] The unchecked global dissemination of multi-drug resistant 'superbugs' has resulted in the current antimicrobial resistance (AMR) crisis. Based on the current rate of discovery and resistance development, antimicrobial resistant infections are expected to cause 10 million deaths globally per year by 2050.[20]

Bacteria can acquire resistance through both genetic mutation and horizontal gene transfer in which foreign DNA is incorporated into a bacterium's genome. As a result, a wide variety of resistance mechanisms capable of reducing a microbe's susceptibility to a drug have emerged. These mechanisms of resistance can be broadly categorized to include: (i) modifications of the drug target, (ii) enzymatic modification of the drug, (iii) decrease in drug uptake, (iv) increase in drug efflux, and (v) bypassing the drug target through alterations in an enzymatic pathway.[21]

1.1.3.1 Controlling Antimicrobial Resistance

Considering the current stagnation in antibiotic discovery, strategies capable of prolonging the lifespan of existing treatments by slowing the emergence of AMR are of critical importance. The AMR crisis has been accelerated by widespread antibiotic use in clinical and agricultural settings. Inappropriate use over the past 50 years, coupled with the lack of development, has had a significant impact on the current public health crisis. Measures to control use are gradually being introduced. For instance, the use of sub-therapeutic doses of antibiotics to promote livestock growth was outlawed by the European Union in 2006 and by the United States of America in 2017, however, this practice is still commonplace elsewhere in the world.[22]

Strategies which reduce selection pressure on bacteria have the potential to reduce the rate at which AMR develops. In some instances, advances in diagnostics have enabled the rapid identification of the causative pathogen of an infection. This has the potential to facilitate treatment with species-specific drugs, avoiding unnecessary selection pressure on non-problematic microbes by broad-spectrum drugs. Additionally, many drug candidates have been overlooked in the past due to their lack of broad-spectrum activity. This advance in diagnostics may result in modern therapeutics being able to take advantage of this former limitation, enabling discarded leads to be revisited. Furthermore,

combination therapies in which multiple drugs are administered together, as has been demonstrated with TB treatment, has also been reinvestigated. This strategy can extend the lifespan of each antibiotic; however, regulatory restrictions can complicate the introduction of drugs as combinations. Finally, anti-virulence approaches in which the pathogen is not directly killed by a drug but instead disarmed, leaving the microbe vulnerable to attack by the host's immune system and limiting the propagation of an infection, is a further consideration to reduce selection pressure exerted on a pathogen. Although these strategies may have the potential to reduce the rate at which AMR emerges, current drugs will inevitably continue to lose efficacy. To maintain the range of treatment options available to clinicians, novel antibiotic discovery is essential.

1.1.4 Reviving Antibiotic Development

In recent years, the re-examination of some abandoned antibiotic candidates has yielded clinically relevant Gram-positive antimicrobials.[18] However, the lack of a fruitful discovery platform has hindered novel discovery. In 2019, the WHO declared that the antibiotics currently in the clinical pipeline will not be sufficient to tackle the increasing emergence of multi-drug resistant bacteria.[23] Of urgent clinical concern, and highlighted in the WHO critical pathogens list, are the ESKAPE pathogens: *Enterococcus faecium*, *Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumanii, Pseudomonas aeruginosa*, and *Enterobacter* species, of which pan-resistant strains are becoming increasingly prominent in hospital and community settings.[24]

Although structural modifications to existing antibiotics are being actively pursued by pharmaceutical firms, novel development has been abandoned by most of the industry. The scientific challenges and limited profitability associated with antibiotic development render it a commercially unattractive field. Even in successful cases, effective antimicrobials may initially be frequently prescribed, thereby promoting rapid resistance emergence, limiting the drug's lifespan. Alternatively, the prescription of an effective drug may be highly restricted to preserve the drug as a last resort therapy. In either case, sales are likely to generate significantly less income than other therapeutic areas while the associated cost of development is exceptionally high.[25] As such, publicly funded strategies to incentivise antibiotic research, such as the EU Innovative Medicine Initiative's 'New Drugs for Bad Bugs' programme, are essential to bring novel drugs to the market.

Synthetic chemistry-based high-throughput screening (HTS) discovery platforms were heavily pursued following the identification of several essential bacterial gene products during the genomic revolution in the 1990s. Unfortunately, this approach has been met with limited success, particularly in the case of Gram-negative bacteria.[25] These species are particularly difficult to target due to the orthogonal nature of their two-membrane cell envelope and intricate drug efflux mechanisms capable of reducing intracellular drug concentration.[26] This almost impermeable barrier has resulted in the high failure rate associated with translating *in vitro* HTS hits into the clinic.

The bacterial cell-envelope prevents many potential antibiotics from reaching intracellular targets and has been a major impediment in discovery, as illustrated in **Fig. 1.3**. Ascertaining the physicochemical properties of a compound which permit permeation of the Gram-negative bacteria cell envelope may inform and guide antibiotic development in a comparable manner to Lipinski's "rule of five". Lipinski's framework essentially correlates the physicochemical properties of a compound with its ability to permeate gastrointestinal epithelial cells and therefore predicts the oral availability.[27] In the case of bacterial cells, understanding the drug properties required to reach intracellular targets can inform the design of compound libraries specifically tailored for antibiotic discovery. This rationale should increase the likelihood of target-based hits from HTS demonstrating whole-cell activity. Studies to this effect, which characterise the rules of Gram-negative permeation and efflux avoidance have recently been reported. In general, factors which favour permeation through negatively charged porins in the outer membrane (OM) are a molecular weight under 600 Da, low hydrophobicity, a flat planar structure and a positively charged amino functionality ($-NH_3^+$).[28-30] These rules were developed based on the OM of *Escherichia coli* and their generality has yet to be confirmed. The permeability of other bacterial membranes, such as the highly restrictive *P. aeruginosa* OM,[31] may be governed by other factors.

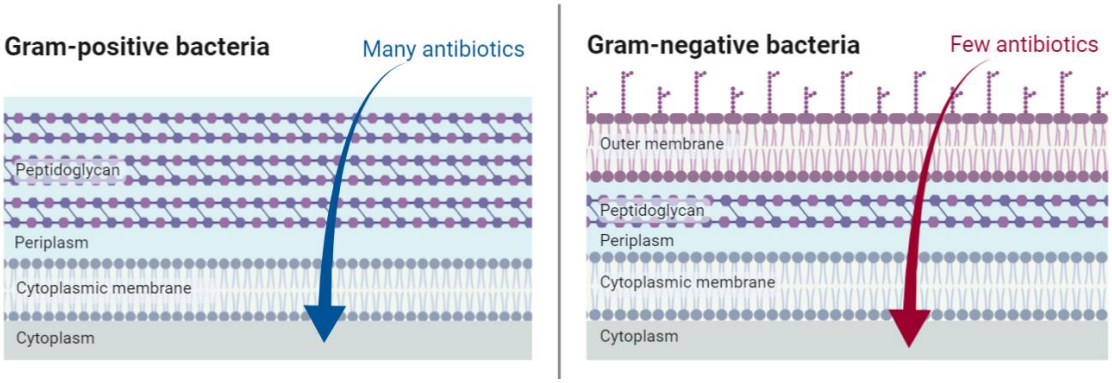


Figure 1.3. **A schematic representation of the Gram-positive and Gram-negative bacterial cell membranes**. Antibiotic permeation through the bacterial cell envelope is highly restricted in Gram-negative bacteria due to the OM. This additional barrier is not present in Gram-positive species.

To streamline the discovery process, antibiotic targets which obviate the need to permeate both the outer and inner membrane (IM) are highly attractive. Specifically, the biosynthesis of bacterial lipoproteins (LPs) involves essential enzymes located on the periplasmic face of the inner membrane. The potential of these enzymes as targets in antibiotic discovery is explored in detail below.

1.2 Post-Translational Modifications of Proteins

The covalent modification of a protein with various functional groups is universally utilised by nature to broaden the scope of a protein's function. Such post-translational modifications (PTMs), including lipidation, glycosylation and phosphorylation, can vastly alter the structure and activity of a protein. Using these modifications an organism can diversify its proteome while maintaining a relatively simple genome.[32]

1.2.1 Lipoproteins

Protein lipidation is ubiquitous in all domains of life. It can be broadly categorised as the covalent attachment of a hydrophobic moiety to a protein's N- or C-terminus or side chain. This modification directs localisation of the protein to a cell membrane.[33] LPs are essential for many fundamental processes, both in the eukaryotic and prokaryotic cell, with roles ranging from cell envelope architecture and nutrient uptake to mediating host inflammatory responses.[34,35]

Between organisms, the structure of the LP lipid moiety is highly variable in terms of acyl chain length, branching, and saturation. Eukaryotic LPs are produced through a

multitude of biosynthetic pathways leading to a range of lipidation patterns including *N*-myristoylation, *S*-palmitoylation, glycosylphosphatidylinositolation and prenylation.[36] Notably, these modifications are distinct from protein lipidation in bacteria and occur through separate biochemical pathways which are not shared with bacteria.

1.2.2 Lipoproteins in Bacteria

All bacterial LPs feature a characteristic N-terminal modification by an acylglycerol modified cysteine residue.[37] LPs in Gram-negative bacteria are triacylated, with an amide-linked acyl chain at the conserved N-terminal cysteine as depicted in compound **4**, **Fig. 1.4**.[35] In Gram-positive species, LP N-terminal structure is varied with a range of cysteine *N*-acylation states exhibited depending on species, strain, and environmental growth conditions.[38, 39] Once believed to produce only diacylated LPs, several non-canonical N-terminal modifications have now been identified in Gram-positive species.[39] Known Gram-positive LP chemotypes include proteins N-terminally functionalised with *N*-acyl-*S*-diacylglyceryl-cysteine (**4**), NH_2-cysteine-*S*-diacylglyceryl (**5**), *N*-acetyl-*S*-diacylglyceryl-cysteine (**6**), *N*-peptidyl-*S*-diacylglyceryl-cysteine (**7**) and *N*-acyl-*S*-monoacylglyceryl-cysteine (**8**)**,** herein referred to as the lyso-form.[39, 40]

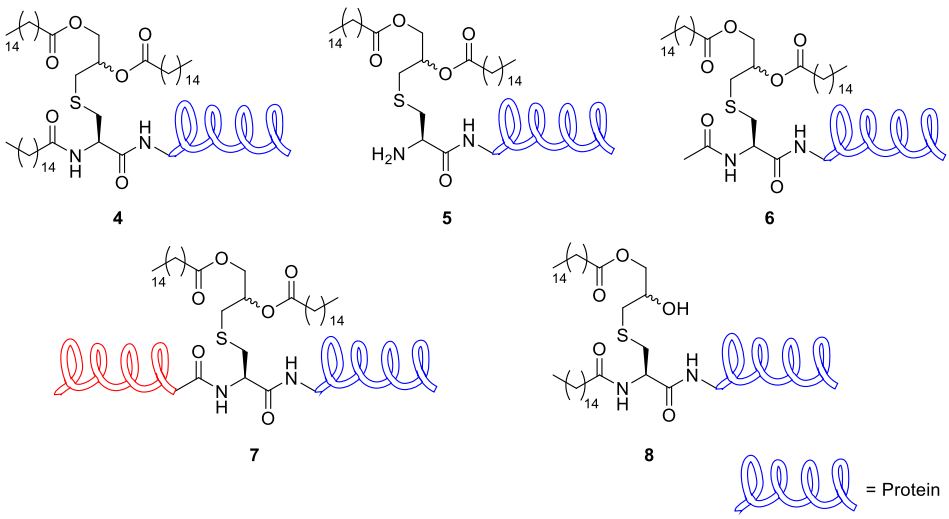


Figure 1.4. **Bacterial LP structure**. Mature Gram-negative LPs are triacylated (compound **4**). Mature Gram-positive LPs are predominantly diacylated (compound **5**) but a wide variety of lipid structures (compounds **4-8**) in these microbes have been identified.

By definition, Gram-negative and Gram-positive bacteria differ in their cell envelope. Gram-negative species possess a double membrane with a thin central peptidoglycan layer whereas Gram-positive bacteria exhibit a single membrane with a thick

peptidoglycan coating as illustrated in **Fig. 1.3**.[41] The cytoplasmic or IM of both classes is composed of a phospholipid bilayer. The Gram-negative OM is an asymmetric bilayer consisting of a phospholipid layer on the inner, periplasmic face and lipopolysaccharide (LPS) on the outer, extracellular side.[42]

Each side of the Gram-negative OM and the periplasmic face of the IM in both classes are heavily decorated with essential LPs which partake in multiple roles.[35] For instance, Braun's LP, the most abundant protein in *E. coli*, is responsible for tethering the OM to the peptidoglycan layer, maintaining the integrity of the cell envelope.[43] Interrupting the biogenesis of such critical biomolecules has repercussions for bacterial viability and virulence. The consequences of targeting and inhibiting the lipidation pathway vary from fatal to detrimental depending on the bacterial class and species.[44] As such, there is potential to exploit a pathogen's reliance on correct lipidation patterns through the development of novel antimicrobials which target this pathway.

1.2.3 Bacterial Lipoprotein Maturation

Bacterial protein lipidation is a PTM which occurs at the periplasmic face of the IM. In general, the process involves three steps, mediated by three enzymes as shown in **Fig 1.5**.[45, 46] Prior to lipidation, preprolipoproteins are produced with a characteristic hydrophobic N-terminal signal peptide which directs the protein, *via* the general secretion (Sec) or twin-arginine translocation (TAT) pathways, to the periplasmic domain of the IM.[47, 48]

Preprolipoproteins are anchored to the IM by the signal peptide.[49] Such signal peptides are generally 20 amino acids in length with the final 4 residues at the C-terminal end of the peptide classified as the 'lipobox'. These 4 amino acids constitute the cleavage region and are most commonly $L^{-3}A^{-2}G^{-1}C^{+1}$. The cysteine residue C^{+1} is denoted as +1 as it will become the first amino acid at the N-terminus of the mature LPs following signal peptide cleavage. Although some variation is permitted in the first 3 residues ($[L/V/I]^{-3}[A/S/T/V/I]^{-2}[G/A/S]^{-1}$), the cysteine residue is conserved across all bacterial LPs.[50]

In Gram-negative and high guanine-cytosine (GC) Gram-positive bacteria (actinobacteria) preprolipoproteins are processed sequentially by three highly conserved enzymes to yield mature triacylated LPs. The first enzyme in the pathway, lipoprotein diacylglyceryl transferase (Lgt), introduces the lipid moiety *via* formation of a thioether bond between the preprolipoprotein cysteine sulfhydryl and a diacylglycerol group

originating from phosphatidylglycerol (PG), yielding the prolipoprotein.[51] The second enzyme, lipoprotein signal peptidase II (LspA), cleaves the signal peptide from the prolipoprotein by hydrolysis of the glycine-cysteine amide bond in the lipobox region.[52] This renders a free amine at the N-terminal dagylated cysteine of the apolipoprotein. This amine is subsequently acylated by the third enzyme in the pathway, apolipoprotein *N*-acyltransferase (Lnt) which utilises a phospholipid as an acyl source.[52] Following this reaction, mature triacylated LPs are translocated to the OM by the localisation of lipoproteins (Lol) pathway or retained at the inner membrane.[53]

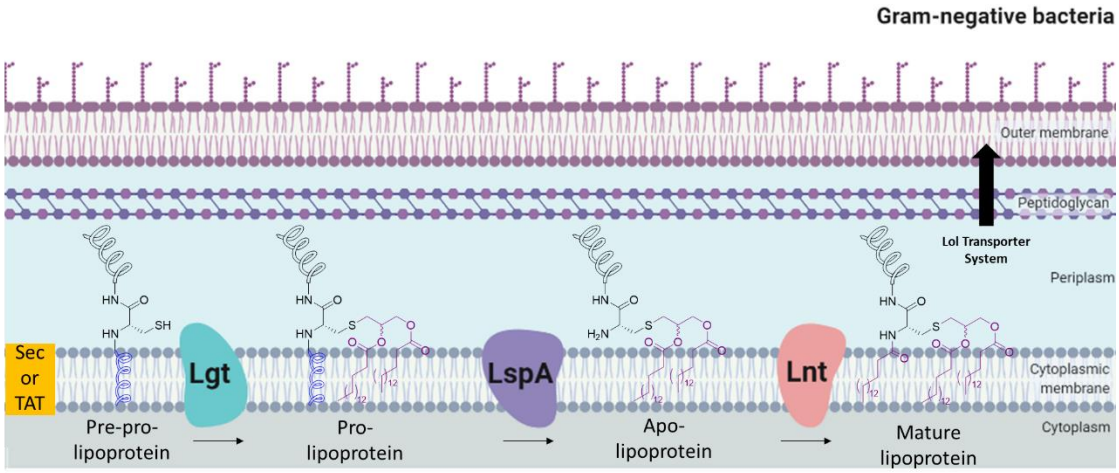


Figure 1.5. The LP maturation pathway in Gram-negative bacteria involves three enzymes: Lgt, LspA and Lnt, which work sequentially to render a triacylated cysteine residue at the N-terminus of mature lipoproteins.

The first two enzymes in this pathway, Lgt and LspA, are conserved in all bacteria. There is variation in the third step of the pathway among bacterial classes.[54-56] In Gram-negative and high GC Gram-positive species (actinobacteria) Lnt yields mature triacylated LPs. The gene encoding Lnt is not present in low GC Gram-positive bacteria (firmicutes) where the majority of LPs are diacylated.[44] However, despite the lack of *lnt*, triacylated LPs and other non-canonical modifications have been identified in some Gram-positive bacteria.[40, 57, 58] The purpose of *N*-acylation was originally believed to be to permit Lol-trafficking of the LP to the OM, and hence, the discovery of triacylated LPs in Gram-positive bacteria, which lack an OM, was unexpected. These findings, in addition to the range of structural N-terminal LP heterogeneity which has recently been identified, confirm that the role of *N*-acylation is more complex than previously recognised.

1.2.3.1 The Role of *N*-acylation

The widespread occurrence of N-terminal acylation and the conservation of such a range of chemotypes indicates that these modifications have been favoured by post-speciation selection pressure. The advantages granted to a bacterium by LP *N*-acylation have yet to be fully unveiled, however, recent advances have shed light on the physiological role of N-terminal chemotype conversion. Meredith and co-workers have recently discovered four enzymes (Lit1, Lit2, LnsA and LnsB) in Gram-positive firmicutes which mediate the conversion of diacylated LPs to lyso LPs and triacylated LPs.[38, 59, 60]

As surface exposed components of the bacterial cell, bacterial LPs play a critical role in alerting the host immune system to infection. LPs are detected by host innate immune cells *via* activation of Toll-like receptors (TLRs). TLRs are pathogen recognition receptors (PRRs) displayed on the surface of antigen presenting cells (APCs) such as dendritic cells. Recognition of an agonist by a TLR leads to cytokine production and downstream activation of the adaptive immune response.[61]

Bacterial LPs are agonists of TLR2.[62-64] Diacylated and triacylated LPs (and lipopeptides[65]) interact with TLRs through two different mechanisms depending on the acylation state of the N-terminus of the LP. LP activation of TLR2 is initiated through TLR2 heterodimer formation with either TLR1 or TLR6.[66, 67] Lee and co-workers have reported the crystal structure of the TLR2-TLR1 heterodimer bound to the synthetic triacylated lipopeptide Pam_3CSK_4 and the TLR2-TLR6 heterodimer bound to the synthetic diacylated peptide Pam_2CSK_4 (**Fig. 1.6**).[68, 69] Structural evaluation of the heterodimer binding sites reveals a hydrophobic binding pocket in TLR2 which coordinates to the thioether bound diacylglyceryl group in both complexes. A hydrophobic pocket in TLR1 accommodates the third acyl chain of *N*-acylated LPs whilst the terminal amino group of diacylated LPs is harboured by TLR6. The selectivity of the TLR2-TLR6 heterodimer for diacylated substrates can be explained by the lack of a hydrophobic pocket in TLR6, which diminishes its affinity for *N*-acylated LPs.

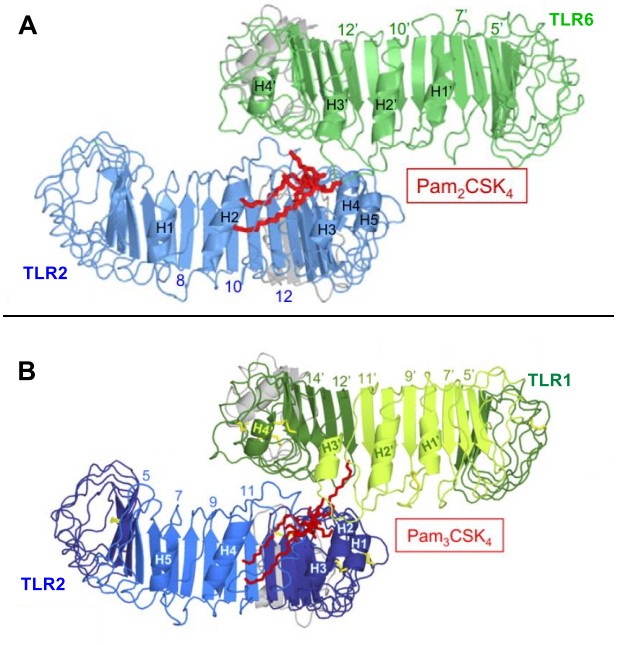


Figure 1.6. **A**: Structure of the TLR2/6 heterodimer in complex with Pam_2CSK_4 (shown in red). Diacylated LPs are recognised by TLR2/6.[68] **B**: Structure of the TLR2/1 heterodimer in complex with Pam_3CK_4 (shown in red). Triacylated LPs are recognised by TLR2/1.[69] Figure adapted from references[68, 69].

Using native LPs produced by *Staphococci* and di- and triacylated synthetic analogues, Nguyen *et al.*[70] have characterised the effect of *N*-acylation on TLR2 activation and cytokine secretion. Diacylated LPs and short (C_2) chain *N*-acetylated LPs produced by *Staphylococcus carnosus* are recognised by TLR2/6. They induce a 10-fold more potent TLR2 response than long (C_{17}) chain *N*-acylated LPs produced by *Staphylococcus aureus* and *Staphylococcus epidermidis*, which are recognised by TLR2/1. Native *N*-acetylated (C_2) LPs were found to induce a similar response to the diacylated synthetic substrate Pam_2CSK_4 and they both induced TLR2/6 activation, suggesting the C_2 *N*-acyl chain is a poor TLR2/1 ligand which is insufficient for TLR2/1 agonisation. Importantly, compared to long *N*-acyl chain LPs, short chain and diacylated *N*-acyl LPs induce higher levels of pro-inflammatory cytokines (a 3-fold increase of IL-12p40, a 5-fold increase of IL-6, and a 10-fold increase of TNF) but lower levels of the anti-inflammatory cytokine IL-8 (a 10-fold reduction). In summary, this work concludes that triacylated LPs induce a weaker pro-inflammatory response than diacylated and *N*-acetylated LPs. This demonstrates that LP *N*-acylation can diminish both the innate and adaptive host immune response to the

microbe, suggesting bacteria have adopted *N*-acylation as a means to evade and suppress the host immune response.[70]

1.2.3.2 *N*-acylation Transferase System

As discussed above, *N*-acylation in Gram-negative bacteria is controlled by Lnt. In 2020, Meredith and co-workers identified two gene products which perform *N*-acylation in *S. aureus*, a Gram-positive firmicute.[60] This two-gene system, termed the lipoprotein *N*-acylation transferase system (Lns), comprises of two enzymes: LnsA and LnsB. The source of the acyl moiety in the reaction and the individual roles of LnsA and LnsB have yet to be determined. There is no sequence homology between LnsAB and Lnt, indicating that these two methods of *N*-acylation have evolved independently. This emphasises the strong selection pressure placed on bacteria to conceal immuno-potent diacylated LPs.

1.2.3.3 Lipoprotein Intramolecular Transacylase

In 2017, Armbruster and Meredith identified a novel enzyme in several Gram-positive firmicutes which acts on the same diacylated apolipoprotein substrate as Lnt.[59] In contrast to the action of Lnt and LnsAB, this enzyme, lipoprotein intramolecular transacylase (Lit) is proposed to facilitate an *O*-to-*N* acyl transfer of a glyceryl O-linked acyl chain to the N-terminus to yield *N*-acyl-*S*-monoacylglyceryl-cysteine functionalised LPs, (lyso-form LPs: structure **8**, **Fig. 1.4**). The purpose of this acyl transfer and the function of the lyso-product was previously unknown, however, using a TLR2/1/6 reporter cell line to evaluate LP activation of TLR2, the authors demonstrate a marked difference in TLR2 activation between the diacylated and lyso-forms. This indicates that this acyl transfer has a role in modulation of the host immune response. Lyso LPs were shown to signal preferentially through the TLR2/6 complex but interact with much lower affinity than diacylated LPs. In comparison to triacylated LPs, a significantly attenuated immune response was induced by lyso-LPs. Correspondingly, a synthetic lyso-lipopeptide required a concentration of 1.0 ng/well to elicit a TLR2 response compared to 0.01 ng/well with a synthetic diacylated lipopeptide. Based on this work, these three LP *N*-chemotypes can be ranked in terms of TLR2 activation potency from the most potent diacylated LPs to the lesser triacylated LPs and the least potent lyso-form LPs, as illustrated in **Fig. 1.7**.[59]

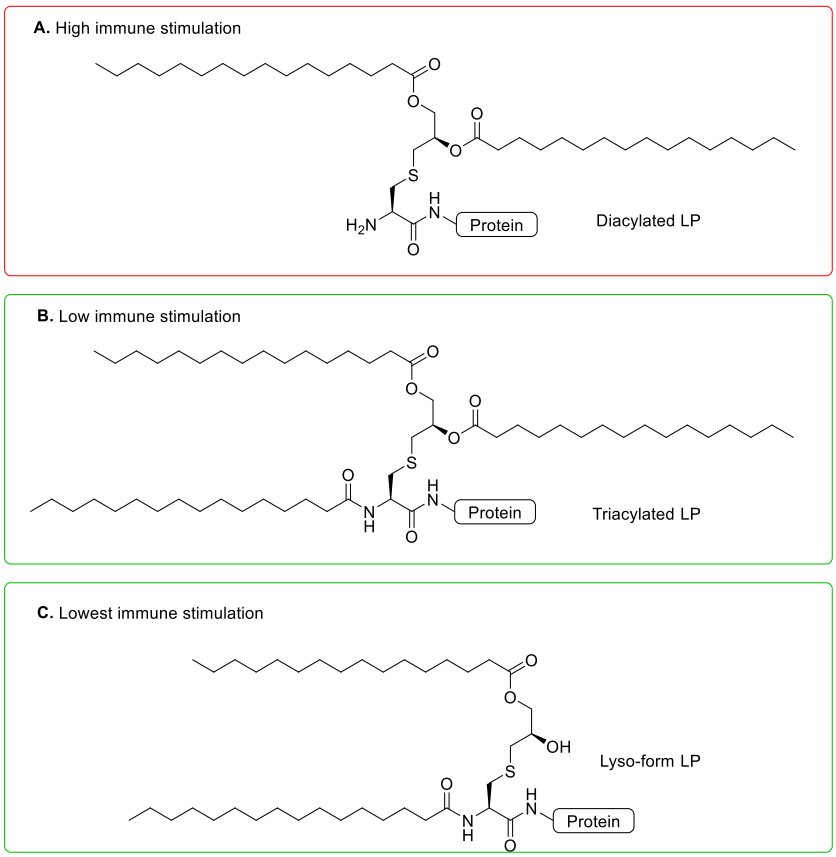


Figure 1.7. The immunostimulating potential of bacterial LPs is dependent on their *N*-acylation state. **A**: Diacylated LPs, predominantly present in Gram-positive bacteria, have the highest affinity for TLR2, causing the strongest immune responses. **B**: Triacylated LPs, primarily present in Gram-negative species, produce an attenuated response. **C**: Gram-positive lyso–LPs elicit the weakest response.[70]

In 2019, the Meredith group also reported a second Lit-type protein, termed Lit2. This *lit* orthologue was identified in environmental isolates of *Enterococcus* spp. and *Listeria monocytogen*es.[38] Critically, the expression of this acyltransferase is promoted by environmental exposure to copper(II). Following the growth of bacteria in media containing copper(II), the conversion of LPs from diacylated to lyso-form was confirmed using MALDI-MS/MS. As a result of this acyl shift, growth in copper(II) containing media resulted in reduced TLR2 activation. *Lit2* is encoded within a known copper resistance operon on a mobile plasmid.

Due to the extensive use of copper(II) as an antimicrobial throughout history, it is unsurprising that bacteria have evolved resistance strategies.[71] The copper(II)-induced expression of Lit2 may be such a strategy. In addition to tempering the immune response, Meredith *et al.*[38] hypothesise that lyso-form LPs may be involved in nutrient uptake at

the cell membrane. It is conceivable, but unproven, that such LPs may have a reduced affinity for copper(II) compared to diacylated LPs due to the now masked amino functionality. Similar effects have been previously reported with Lnt depleted bacteria, in which diacylated LPs are produced and increased intracellular copper(II) concentrations are observed as a result of membrane trafficking defects.[72] It is possible this N-terminal remodelling *via* acyl transfer impacts copper(II) uptake, imparting protection to the microbe from copper(II).[38]

1.2.4 Inhibition of Bacterial Lipoprotein Maturation

The enzymes involved in bacterial protein lipidation are exclusive to bacteria. The lack of eukaryotic analogues makes these targets attractive for antibiotic development as off-target effects on eukaryotic cells are less likely, minimising the side effects of treatment. Mutation studies have revealed how the loss of Lgt, LspA and Lnt function affects viability in several bacterial species. Considering the multitude of roles performed by LPs it is unsurprising that inhibition of their synthesis has a wide range of pleiotropic effects for the bacterial cell. The extent of these effects varies between bacterial classes due to the differences in LP function across bacterial species.

In most Gram-negative bacteria Lgt, LspA and Lnt are essential.[73-78] As these bacteria rely on functional, triacylated LPs to assemble their OM, inhibiting the production of such LPs is fatal.[79] Compromising the OM also makes the bacterium more vulnerable to other antibiotics which are unable to penetrate the OM suggesting this strategy may have potential for application in combination therapies.

The situation is more diverse in Gram-positive bacteria. For some Gram-positive bacteria, *lgt* deletion has been shown to have little effect on growth.[80-83] Additionally, the effect of *lgt* deletion on Gram-positive virulence appears to vary depending on the bacterial strain.[83-85] As this disparity is not yet fully understood, the potential of Lgt as an antibiotic target remains under investigation.

LspA is not necessary for the growth of all Gram-positive bacteria, however, it is crucial for full virulence.[78, 80] For example, in 2020 Olatunji *et al*.[77] demonstrated that *lspA* deficient mutants of methicillin-resistant *Staphylococcus aureus* (MRSA) are unable to survive in whole human blood. Consequently, this enzyme is an attractive antibiotic target, even in bacteria where LspA activity is not essential. In these instances, LspA inhibition may constitute an anti-virulence approach which favourably reduces the

selection pressure exerted on the bacterium. Additionally, the structural variation between LspA orthologues of two human pathogens, *P. aeruginosa* (LspPae) and *S. aureus* (LspMrs), has recently been revealed through X-ray crystallography.[77] It may be possible to exploit these variations through the design of inhibitors specific to each pathogen, promoting a species-specific inhibition approach to minimise antibiotic side effects arising from inadvertent targeting of benign microbes.

Despite their potential as drug targets, the location of protein lipidation enzymes at the IM makes these enzymes inherently difficult to access in Gram-negative bacteria. The low permeability of the OM and the multitude of resistance mechanisms exhibited by these species provide extensive protection to the microbe. Somewhat fortunately, the active sites of LspA, Lnt, and Lit reside on the outer side of the IM, exposed to the periplasm.[86] A successful broad-spectrum drug candidate must be capable of reaching a sufficient periplasmic concentration by permeating the OM in Gram-negative bacteria and overcoming resistance mechanisms such as drug efflux.

Previous drug development efforts targeting the lipidation pathway have focussed primarily on LspA. The potential of this enzyme as an antibiotic target has been recognised for decades. Two naturally occurring antibiotics, myxovirescin (**9**) and globomycin (**10**), which inhibit LP biosynthesis were discovered in the 1970s and 1980s.[87, 88] Although LspA was identified as the target of globomycin in 1978,[89] it was not until 2012 that myxovirescin was found to selectively target LspA.[90] Produced by *Streptomyces hagronensis*, globomycin is a cyclic depsipeptide while myxovirescin, produced by *Myxococcus xanthus*, is a macrocyclic lactone.[91, 92] The structure of each compound is shown in **Fig. 1.8**. Poor *in vivo* activity has prevented the translation of these antibiotics into the clinic. Unfavourable physicochemical properties result in poor OM permeation and low plasma stability, which is likely a result of esterase activity.[93-95] For example, the adhesive nature of myxovirescin halted its development as it tends to localise at the administration site without achieving measurable blood concentration.[94] In contrast to LspA, no specific inhibitor of Lnt or Lit has been identified.

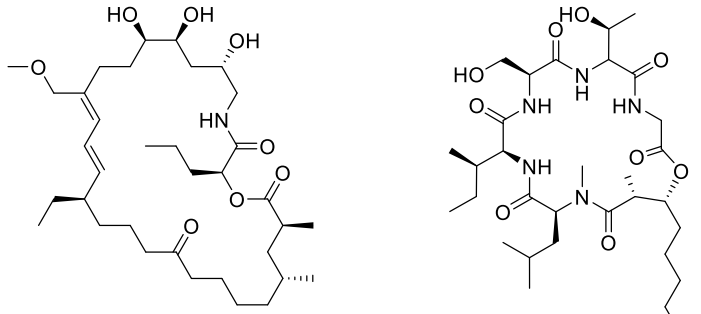


Figure 1.8. The structure of two naturally occurring inhibitors of LspA, myxovirescin (**9**) and globomycin (**10**).

Many of the factors which have stalled the development of LspA inhibitors have been overcome in recent years. Globomycin shows activity only against Gram-negative bacteria but cannot permeate the OM sufficiently to have clinical application. Recently, synthetic derivatives have been reported with anti-Gram-positive activity where there is no need to permeate an OM.[95] In 2020, Genentech reported the synthesis and evaluation of cyclic ether analogues of globomycin which are stable to esterase activity and hence exhibit improved plasma stability.[66] The physicochemical properties required for OM permeation are also now more fully understood.[28-30] This is likely to guide the design of future broad-spectrum derivatives. The mechanism of action of myxovirescin is now known and the crystal structure of this inhibitor complexed with LspMrs has been reported, in addition to the complex of globomycin with both LspPae and LspMrs.[77, 86] The detailed structural information this provides about active site morphology, coupled with *in silico* modelling, will inform rational inhibitor design.

With two promising 'hit' compounds, LspA is perfectly poised for further drug development. However, the screening of optimised inhibitors is challenging due to the lack of robust, time efficient methods for assaying the activity of this enzyme. An assay with appropriate throughput will accelerate the search for globomycin and myxovirescin derivatives with clinical potential. Additionally, a high-throughput assay will permit HTS which has the potential for novel inhibitor identification.

As discussed in **Section 1.1.4**, HTS in antibiotic discovery has been met with limited success in the past.[25] However, combinatorial and synthetic chemistry advances have expanded the diversity of compound libraries used in HTS, improving the probability of identifying an inhibitor which retains activity *in vivo*. Furthermore, the development and

miniaturisation of automated screening have lowered the cost and increased the speed of HTS. Progress in the field of drug delivery, notably the use of nanoparticulate carriers such as liposomes or polymeric nanoparticles, can mediate the impact of traditionally unfavourable pharmacokinetic properties of a drug.[96] Ultimately, when coupled with guidance by the emerging 'rules of permeation' these scientific developments warrant the continued investment in antibiotic HTS campaigns.

Identifying potent inhibitors remains a key step in the drug development process. The lack of efficient *in vitro* assays for LspA, Lnt and Lit have stalled the discovery of novel inhibitors. As such, a major aim of this project focused on developing synthetic peptides for application in the development of HTS-compatible assays for LspA, Lnt and Lit.

1.3 Bacterial Lipoproteins in Vaccine Design

As controlling infectious diseases with antibiotics is becoming increasingly challenging, alternative approaches to combat bacterial outbreaks are in high demand. One such strategy is vaccination. Vaccines, which can be prophylactic or therapeutic, have been introduced against many bacterial diseases including TB, diphtheria, tetanus and typhoid.[97-100] Furthermore, the use of vaccines in cancer treatment is becoming increasingly prominent following the recent clinical success and approval of immunotherapeutics such as sipuleucel-T.[101, 102]

In contrast to whole pathogen or live-attenuated vaccines, several modern vaccines consist of a purified antigen which is co-administered with an adjuvant. In these rationally designed vaccines, an antigenic pathogen-derived fragment, or subunit, is used to elicit a well-defined immune response. With the capability to precisely characterise this immune response, coupled with the inability of a subunit to cause infection through a reversion to virulence, subunit vaccines are generally considered safer than tradition live-attenuated vaccines. However, unlike whole pathogens, antigens alone often elicit an insufficient immune response.[103] Their low affinity to host immune cells necessitates the inclusion of an adjuvant to first boost the immune response by mimicking an infection. Adjuvants, defined as any compound capable of enhancing the immunogenicity of an antigen, are currently used in more than 30 approved vaccines.[104]

The most commonly used adjuvants are alum (potassium aluminium sulphate or other aluminium salts) and MF59 (squalene). However, their mechanisms of action remain unclear despite widespread use.[105] Alternative adjuvants, such as TLR agonists, are now

receiving heightened attention. Upon recognition of an agonist by TLRs, inflammatory cytokines are produced, and major histocompatibility complex (MHC) presentation of the co-administered antigen is prompted. T cells specific to the antigen displayed by MHC are activated, with this adaptive immune response leading to long term immunity.[106] The ability of bacterial LPs to provoke a host immune response by engaging TLR2 with limited apparent host toxicity has resulted in the extensive investigation of recombinant LPs and synthetic lipopeptides as adjuvants.[107-109]

The structural features of bacterial LP derivatives which are essential for activation of the human TLR2 signalling pathway have been extensively investigated to enhance immunogenic activity while maintaining low host toxicity. Using synthetic lipopeptides, these structure-activity relationship (SAR) studies have revealed the influence of structural modifications of both the central glyceryl component and lipid chains on immunogenic activity. These modifications and their effect on activity in comparison to compound **11** are summarised in **Fig. 1.9**. Variation of the lipid moiety by Buwitt-Beckmann *et al.*[110] (compounds **12** and **13**) showed that activity was dependent on glyceryl *O*-acyl chain length. A minimum chain length of C_8 was required for TLR2 activation, with activity increasing with chain length, reaching a maximum at C_{16}. Agnihotri *et al.*[111] reported that introducing unsaturation into the acyl chains *via* substitution of palmitic acid with *cis,cis*-9,12-octadecadienoic acid (as in compound **14**) slightly reduced activity. Substitution of ester-linked acyl groups for amide-linked acyl chains by Seyberth *et al.*[112] and Schromm *et al.*[113] (as in compound **15**) resulted in a total loss of activity. Agnihotri *et al.*[111] showed that loss of one lipid group had surprisingly little effect, with the stereospecific numbered C-1 (*sn*-1) monoacylated analogue **16** retaining moderate activity.

Through modification of the glyceryl component, Takeuchi *et al.*[114] found that the (*R*)-glyceryl stereoisomer (compound **17**) was >100 times more active than the (*S*)-glyceryl isomer (compound **18**). Additionally, Agnihotri *et al.*,[111] Wu *et al.*[115] and Metzger *et al.*[116] have shown that cysteine thioether substitution by either a methylene (compound **19**) or an ether group (compound **20**) abolished activity while substitution with a selenoether increased activity slightly (compound **21**). Agnihotri *et al.*[111] found that increasing the number of carbon atoms between the cysteine α-carbon and the cysteine thioether (as in compound **22**), resulted in a substantial decrease in activity. Brimble and co-workers have recently reported glyceryl modifications which improve TLR2 activity.[117] Incorporating

an additional methylene group between the two lipids did not impact the agonist potency of compound **23** in comparison to compound **11**. Furthermore, this group demonstrated that acetylation of the N-terminus (as in compounds **24** and **25**) was well tolerated, while palmitoylation (as in **26**) was not, in line with the results from Nguyen *et al.*[70] discussed in **Section 1.2.3.1**.

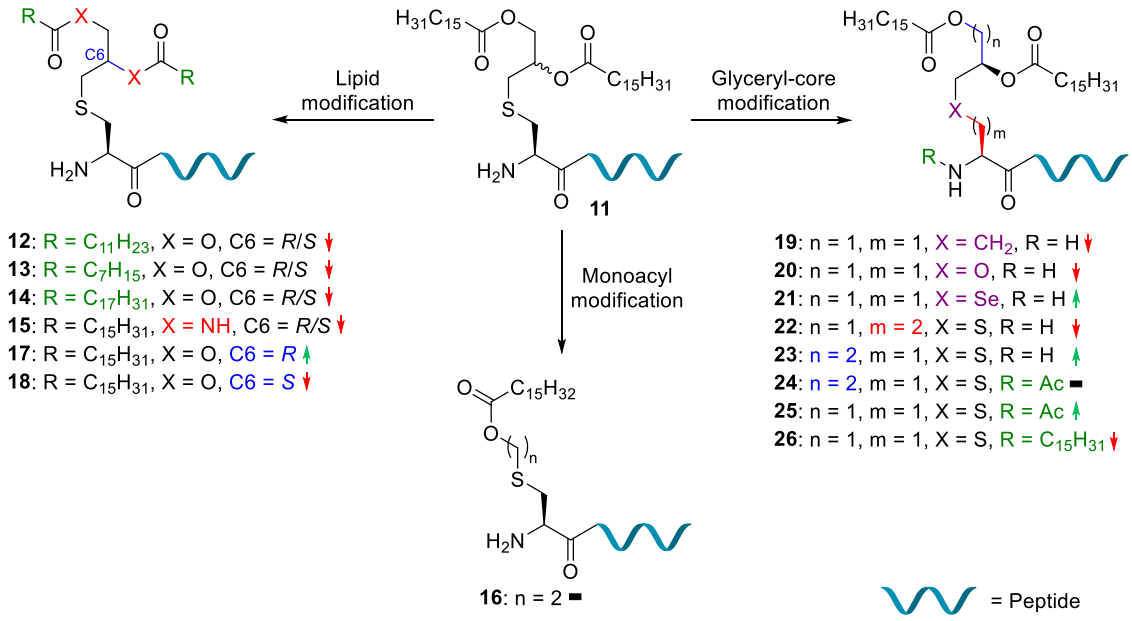


Figure 1.9. Summary of SAR studies of TLR2 agonists with synthetic derivation of both the lipid groups and glyceryl-core. Modifications resulting in a decrease in TLR2 activity when compared to **11** are marked with a red arrow. Modifications resulting in an increase in activity when compared to **11** are marked with a green arrow. Modifications with a negligible change in activity when compared to **11** are marked with a black dash.

The lipopeptides used in these SAR studies were synthesised using Fmoc/*t*Bu solid-phase peptide synthesis (SPPS). Primarily, the lipid modification was introduced *via* coupling of Fmoc-Cys(2,3-di(palmitoyloxy)-propyl)-OH (Fmoc-Cys(Pam)$_2$-OH), or a corresponding analogue, at the N-terminus of the peptide followed by deprotection and resin cleavage. However, on-resin acylation of 2,3-dihydroxylpropyl cysteine-functionalised peptides has also been reported.[118] In either case, the hydrophobic nature of these bacteria-derived lipopeptides is problematic for both synthesis and purification. To overcome this, a solubility tag, (Lys)$_4$, is often incorporated into the peptide sequence in addition to selective consideration of the complete amino acid sequence to promote aqueous (aq.) solubility. A water-soluble peptide is desirable for multiple reasons. Peptides with multiple polar residues exhibit reduced aggregation during peptide

assembly on the resin, maximising reaction yields and minimising side product formation. Purification by reversed-phase (RP) high performance liquid chromatography (HPLC) is also possible with water-soluble peptides. Furthermore, solubility under physiological conditions is advantageous in many biological applications which may not be tolerant to dimethylsulfoxide (DMSO) addition.

Although the incorporation of $(Lys)_4$ does not disrupt TLR2 binding, not all applications of lipopeptides are amenable to this restriction in amino acid sequence. Synthetic strategies which give greater flexibility in the potential amino acid sequences are required to increase the scope of synthetically accessible lipidated peptides.

In an alternative approach, several LP adjuvant candidates have been produced recombinantly. However, despite *E. coli* expression systems being equipped with the machinery (Lgt, LspA and Lnt) to post-translationally modify any lipobox containing protein, expression in *E. coli* often leads to a complete lack of or incomplete protein lipidation.[119] Extensive optimisation of growth conditions for the *E. coli* strain is required to promote lipidation and LPs are often difficult to isolate due to problematic contamination with LPS. A highly heterogeneous population of LPs, with regards to acyl chain length and saturation, is obtained due to the large variation in phospholipid acyl source present in the bacterial cell. *In vitro* enzymatic modification of LP precursors with each of the three enzymes sequentially suffers from similar limitations. The activity of each enzyme requires highly specific conditions. Performing the 3-step modification in one-pot is hindered by the variation in these reaction conditions, with each of the enzymes performing optimally with a different buffer, detergent, temperature, and pH for example. Additionally, the expression and purification of these membrane enzymes are not trivial, making this costly approach impractical for accessing the large quantities of LP required for clinical applications.[120] Chemical ligation has the potential to provide improved access to bacterial LPs and lipopeptides while avoiding the issues associated with recombinant expression and linear SPPS strategies.

1.4 Chemical Protein Synthesis

The application of synthetic chemistry to study the effects of PTMs on protein function is a rapidly growing field of research. Synthetic approaches can provide access to precisely modified proteins for both therapeutic application and SAR studies to advance the understanding of specific biological systems.

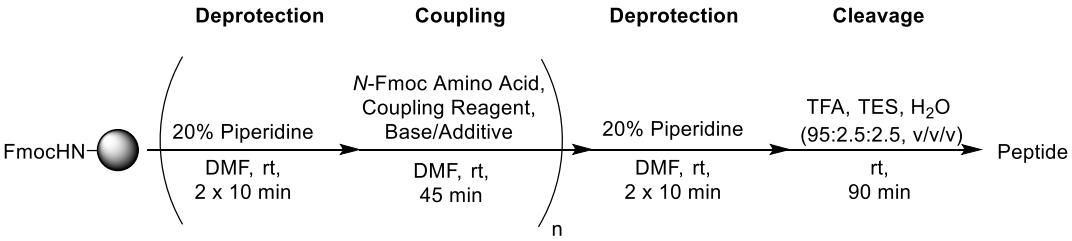
Proteins which are modified post-translationally are often difficult to access *via* recombinant means due to the influence of growth conditions on the expression system. The activity and substrates of enzymes used to introduce PTMs can be impacted by these conditions. As a result, expressed proteins may suffer from a wide range of heterogeneity in the PTM making it impossible to carry out accurate SAR studies while also hindering the regulatory process if the material is destined for clinical use. Additionally, in the production of human therapeutics the use of mammalian expression systems, such as Chinese hamster ovary cells, can result in the incorporation of non-human epitopes which may trigger an undesirable immune response.[121] In some cases the enzymes required for PTM are not present in the expression system, such is the case with *E. coli* expression of human *N*-linked glycoproteins.[122]

Recombinant expression of LPs is challenging. Access to correctly lipidated proteins would permit the study of the role of LPs as immunomodulators, while access to pro- or apo-LP precursors would enable the study of the LP biosynthesis pathway. Chemical synthesis is uniquely situated to introduce residues such as fluorophores and functional probes into synthetic lipopeptides and LPs through SPPS and peptide ligation.

1.4.1 Peptide Synthesis

The amide functionality is ubiquitous in pharmaceuticals, in part attributable to the growing popularity of peptide-based therapeutics such as glucagon-like peptide (GLP)-1 receptor agonists. As such, efficient and economical methods for amide formation are highly sought after. Amide bond formation typically involves the reaction of an amine with an activated acyl source, such as a carboxylic acid derivative. Traditional solution-phase peptide synthesis requires the protection of other reactive groups in both the amine and carboxylic acid reactants. This necessitates the use of extensive orthogonal protecting group strategies and involves time-consuming chromatographic purification after each amino acid coupling. Despite the time investment required, this approach is viable for the synthesis of short peptides of up to ~5-6 residues. The maximum peptide length is, however, limited by the reducing solubility of the growing protected peptide in organic solvents. The insolubility of the peptide will result in the failure of both coupling and deprotection reactions when attempting to further manipulate or extend the peptide in addition to preventing effective purification. The revolutionary solid-phase approach pioneered by Merrifield in 1963[123] in which the peptide is immobilised and assembled on

a solid support has overcome this limitation and enabled the synthesis of peptides with up to ~50 amino acids.[124] Merrifield's strategy involves the anchoring of an *N*-protected amino acid to a solid support *via* the C-terminus through an amide or ester bond. Iterative *N*-deprotection and amino acid coupling steps permit the elongation of the resin-bound peptide as depicted in **Scheme 1.1**. Following the peptide assembly, resin cleavage and global side-group deprotection yield the unprotected peptide which is then subject to chromatographic purification.



Scheme 1.1. A general Fmoc SPPS strategy on an amine functionalised resin.

Peptides are normally assembled in the C-to-N direction as the alternative direction involves the C-terminal activation of an N-terminally anchored peptide, a species which is prone to 5(4*H*)-oxazolone or diketopiperazine (DKP) formation.[125] These side reactions can result in epimerisation of the C-terminal residue or chain capping as illustrated in **Fig. 1.10**.[126]

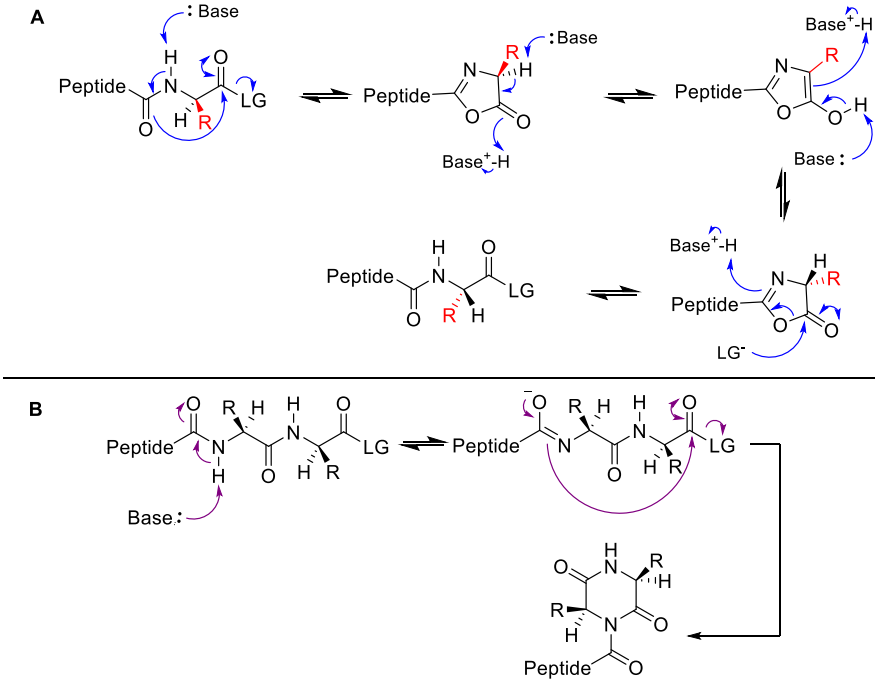


Figure 1.10. Common side reactions which occur during SPPS. A. Oxazolone formation can lead to a loss of stereo-integrity of the activated C-terminal residue. **B**. Diketopiperazine formation terminates chain elongation.

1.4.1.1 *tert*-Butyloxycarbonyl (Boc) SPPS

In 1984, Merrifield was awarded the Nobel Prize for his seminal work on SPPS. This technique has become the cornerstone of the field of peptide science and remains the fundamental method used today to synthesise peptides. Merrifield's original work reports the synthesis of a tetrapeptide using Boc-protected amino acids with a chloride functionalised 2% cross-linked vinylbenzene-divinylbenzene (DVB) copolymer resin.[123] Referred to as Boc SPPS, this method exploits the difference in acid lability between the Boc protecting group on the N-terminal amino acid, side chain protecting groups (such as *O*-benzyl) and the Merrifield resin linker. N-Terminal Boc deprotection is performed using neat trifluoroacetic acid (TFA) while side chain deprotection and resin cleavage are effected by anhydrous hydrogen fluoride (HF).[127] However, the use of highly toxic HF is unfavourable due to safety concerns and as it requires specialised plastic-ware. Additionally, repeated *N*-Boc deprotection with neat TFA results in minor amounts of side chain deprotection and resin cleavage, reducing crude peptide purity and yield.

1.4.1.2 Fluorenylmethyloxycarbonyl (Fmoc) SPPS

Since the introduction of SPPS, several advances have permitted the use of more favourable conditions. Notably, the introduction of the Fmoc *N*-protecting group in 1970 by Carpino and Han[128] has resulted in the widespread uptake of the Fmoc-*t*Bu strategy, developed by Chang and Meienhofer.[129, 130] Fmoc is a mild-base labile N-terminal protecting group commonly deprotected using 20% (v/v) piperidine/*N*,*N*-dimethylformamide (DMF). Advantageously, Fmoc-deprotection yields a UV active flourene byproduct, the absorbance of which can be used for inline quantification of the reaction progress in automated SPPS, providing indirect information regarding the success of the previous coupling.[131]

Fmoc SPPS generally employs an acid labile resin linker, such as the hydroxymethylphenoxy-based linkers, and utilises acid labile *tert*-butyl protection on many amino acid side chains. Resin cleavage and side group deprotection are orthogonal to N-terminal deprotection, meaning the peptide remains intact during repeated Fmoc deprotection reactions. The popularity of Fmoc SPPS has resulted in widespread commercial availability of cheap, high quality Fmoc-protected amino acids with acid labile side chain protecting groups. The global deprotection and resin cleavage steps are performed under acidic conditions. The concentration of the acid required varies from

1% to 95% TFA, for instance, depending on the resin linker. A wide range of linkers have been developed which are compatible with Fmoc SPPS and are utilised in numerous applications. Common examples are depicted in **Fig. 1.11** and include *p*-alkoxybenzyl alcohol (Wang) resin, a hydroxyl functionalised resin which upon cleavage with 95% TFA, gives access to peptide carboxylic acids.[132] 4-((2,4-Dimethoxyphenyl)(Fmoc-amino)methyl)phenoxyalkyl (Rink amide) functionalised resin yields peptide carboxyamides following cleavage with 95% TFA[133] and 2-chlorotrityl chloride resin, a chloride functionalised resin which upon cleavage with 1% TFA, can yield the protected peptide carboxylic acid.[134]

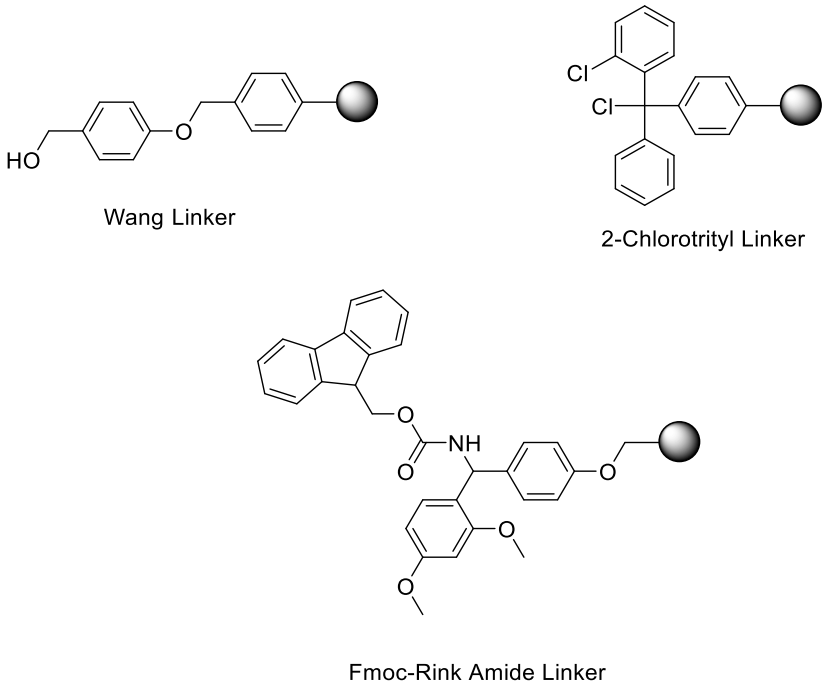


Figure 1.11. Structures of common resin-bound linkers used in SPPS.

1.4.1.3 Limitations of SPPS

SPPS has revolutionised peptide science, however, the length of peptide which can be synthesised remains a major limitation of the methodology. The maximum peptide length obtainable is highly dependent on the peptide sequence. Under ideal conditions, peptides of up to ~50 amino acids can be synthesised, however, SPPS is commonly restricted to shorter sequences of less than ~30 residues.[124] This restriction arises from inter- and intra-chain aggregation of the resin-bound peptide, with the resultant lack of solvation of the N-terminus hindering solvent accessibility and thus suppressing the desired deprotection or coupling reaction.[135, 136] Successful SPPS relies on each coupling and deprotection reaction proceeding in near quantitative yields. To achieve this with short reaction times

of under 1 h, a large excess (>3 equivalents (equiv.)) of the activated amino acid is used in high concentration (>0.2 M) for coupling reactions. As aggregation can reduce the rate of the reaction, side reactions such as epimerisation of the activated carboxyl species can become increasingly problematic, particularly under the basic conditions used for Fmoc SPPS. The factors affecting the physical stability of resin-bound peptides remain poorly understood due to the challenges of determining the secondary structure of peptides bound to a heterogeneous solid matrix.[137] As such, the identification of "difficult sequences" which have a propensity for aggregation remains difficult. Generally, however, peptides containing multiple hydrophobic residues are expected to have a greater tendency to aggregate.[138]

1.4.1.4 Extending the Scope of SPPS

1.4.1.4.1 Backbone Protection

The difficulties of handling hydrophobic peptides impact each stage of the synthetic process including the synthesis on resin, purification, and analysis. Methods to improve the scope of SPPS and circumvent aggregation include the introduction of temporary protecting groups on backbone amide NH functionalities. Despite the addition of further hydrophobic groups, peptide solubility and synthesis success is greatly enhanced due to disruption of the hydrogen bonding network between peptide chains.[139, 140] Similarly, the use of serine, threonine or cysteine pseudoproline derivatives, which can be readily deprotected during 95% aq. TFA global cleavage, induces 'kinking' of the peptide and prevents aggregation, improving the yield and purity of the crude peptide.[141, 142] Johnson *et al.*[143] have reported the use of acetylated 2-hydroxy-4-methoxybenzyl (AcHmb) which yields a TFA stable derivative of the popular Hmb backbone protecting group. This enables the backbone protected free peptide to be obtained after resin cleavage which exhibits improved aq. solubility, permitting purification and characterisation. Deacetylation of Hmb can be performed with 20% piperidine/DMF, restoring the acid lability of Hmb, which can be cleaved to give the fully deprotected peptide using 95% aq. TFA.[143]

1.4.1.4.2 Temperature

The use of conventional or microwave assisted heating is an alternative method of suppressing peptide aggregation during the synthesis of large, difficult sequences.[144] In comparison to room temperature, an increase in temperature by 60 °C has been reported

to increase the rate of coupling and deprotection reactions by 50-fold.[145] With amino acid coupling complete within 10 min and Fmoc-deprotection in 2.5 min, these expedited reaction rates provide rapid access to a broader scope of challenging peptides in synthetically viable purity. Caution is required with cysteine or histidine containing peptides due to the reduced level of stereocontrol at higher temperatures with these racemisation prone residues.[146] Furthermore, aspartimide formation becomes increasingly prevalent at high temperatures with aspartic acid containing peptides.[147] Additionally, a loss of peptide loading has been observed at elevated temperatures, particularly with trityl-based linkers.[148]

1.4.1.4.3 Polymer Choice

The appropriate choice of polymer-support for SPPS is critical to achieving a high level of solvent accessibility to reactive sites. Traditionally, polystyrene resins have been employed in SPPS, used in combination with DMF and CH_2Cl_2. However, the hydrophobicity of this resin can promote peptide aggregation while also being incompatible with aq. conditions.[149] To extend the scope of on-resin reactions to include those carried out in aq. solution, polyethylene glycol (PEG) containing resins have been developed.[150] The superior swelling properties of PEG-based resins across a range of solvents, when compared to polystyrene resins, results in improved crude purity and yield.[151] For example, Albericio *et al.* have reported the straightforward synthesis of the notoriously aggregation prone β-amyloid (1-42) peptide with a crude purity of 91% using the fully PEG-based ChemMatrix resin.[151] In addition, PEG-based resins can also permit the use of "greener" solvents, such as MeCN and anisole, with an improved environmental impact and reduced safety concerns when compared to the traditional SPPS solvents CH_2Cl_2 and DMF.[152]

1.4.1.4.5 Automated Flow Peptide Synthesis

The application of flow chemistry to SPPS, in which solvent carrying reagent is pumped through a stationary column of functionalised resin, was first investigated in 1970.[153, 154] Using a customised flow system, the classical advantages of flow chemistry can be applied to SPPS. In particular, the high level of control over physical parameters, such as temperature and flow rate, can be utilised to maximise reaction rate and yield while minimising side product formation. In 2020, Hartrampf *et al.*[155] reported the application of automated fast-flow peptide synthesis (AFPS) to prepare 9 functional proteins,

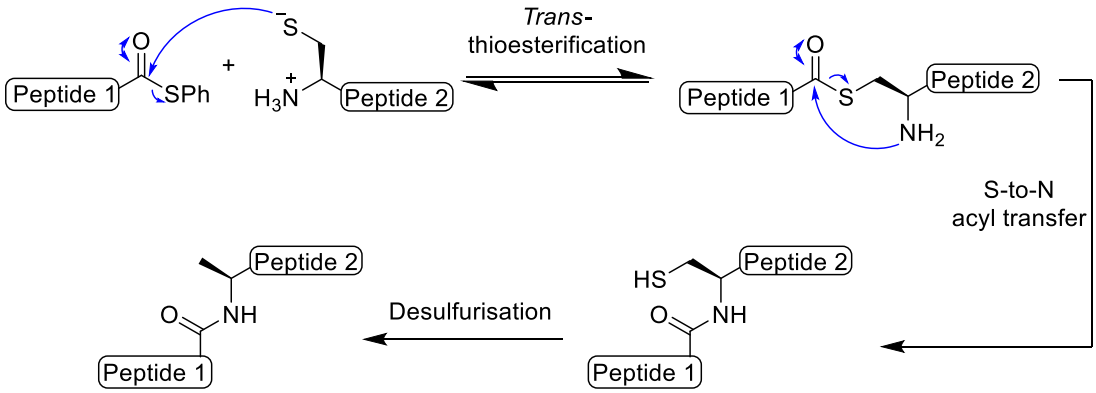
including HIV-1 protease (99 amino acids) and Sortase A (59-206) (164 amino acids), synthesised in hours *via* SPPS performed with a customised flow instrument. Due to the number of consecutive reactions (up to 327) performed without chromatographic purification, producing proteins through AFPS in synthetically viable crude purity requires essentially quantitative, consistent yields for every reaction. Hartrampf *et al.* demonstrate that this is achievable under flow conditions. Following rigorous optimisation of reaction conditions, including temperature, amino acid pre-activation time and temperature, flow rate, coupling reagent and base, Fmoc-deprotection base and side-chain protecting groups, the production of purified proteins in milligram quantities with 1-5% overall yields was realised. Notably, coupling times of 40 sec are reported with each complete coupling-deprotection cycle achieved in 2.5 min.[155] This represents a significant advance in what can be achieved using SPPS technology. However, the limited ability to monitor the structure of the resin-bound peptide throughout the synthesis may hinder the uptake of this methodology. Beyond indications arising from UV analysis of Fmoc-deprotection reactions, structural irregularities and side product formation cannot be detected prior to analysis of the cleaved protein. With peptides of this size, it is likely that microheterogeneities in the structure, such as epimerisation of a single chiral centre, may be below the limit of detection of the analytical methods currently available. To account for this, the authors perform functional assays demonstrating the comparable activity of synthetic and expressed proteins. The generality of AFPS for protein synthesis, however, particularly with difficult, aggregation prone sequences, remains to be determined.

Despite the attractive simplicity of AFPS, until analytical methods reach the capability required to monitor and effectively troubleshoot AFPS, peptide ligation will likely remain the most prevalent route to access synthetic proteins. By synthesising proteins *via* ligation of several peptide fragments, structural integrity can be monitored and confirmed at multiple points throughout the synthesis, while also offering improved atom economy in comparison to AFPS, where 60 equiv. of each amino acid are employed in coupling reactions.

1.4.2 Peptide Ligation

In the mid-1990s the field of chemical protein synthesis was transformed following the seminal report of native chemical ligation (NCL) by Kent and co-workers.[156] For the first time, the total synthesis of native proteins could be achieved by the chemo-selective ligation of unprotected peptide fragments. The high level of structural control which can be attained with chemical protein synthesis offers unrivalled flexibility in terms of the introduction of unnatural amino acids, homogeneous PTMs and modifications on the protein backbone. Since its introduction in 1994, NCL has been applied to the synthesis of numerous proteins, including hydrophobic LPs.[157, 158]

The proposed mechanism of NCL involves a two-step process, shown in **Scheme 1.2**. The first step, *trans*-thioesterification of a C-terminal peptide thioester with an N-terminal cysteine on a second peptide, is reversible and rate-determining. In the second step, a spontaneous intramolecular *S*-to-*N* acyl shift occurs through a 5-membered cyclic transition state which yields a native peptide bond between the two original fragments.[159]



Scheme 1.2. NCL is a two-step process shown here followed by desulfurisation to yield an alanine residue at the ligation site.

Due to its low natural abundance, the requirement for cysteine at the ligation site can be regarded as a limitation in NCL as it is likely that the target protein will not contain an appropriately located central cysteine residue. However, the combined NCL-desulfurisation approach developed by Yan and Dawson overcomes this as the cysteine residue can be converted to alanine, a more abundant residue, post-ligation.[160] This approach has been furthered by metal-free desulfurisation chemistry developed by Kan and Danishefsky,[161] and extended beyond alanine to incorporate ligations performed at junctions containing a β- or γ-thiol-functionalised amino acid. Such amino acids can be

desulfurised to their corresponding native residue post-ligation.[162] In contrast to ligation at cysteine, the use of more sterically hindered thiol-amino acids has a detrimental effect on the reaction rate. As such, thiol-additives, such as (4-carboxylmethyl)thiophenol (MPAA), are routinely added to increase the rate of reaction *via* nucleophilic catalysis.[159]

Favourably, NCL is chemo-selective and compatible with aq. conditions, permitting the use of water soluble, unprotected peptides. Secondary and tertiary structure formation, which can be detrimental during ligation, can be interrupted using NCL-compatible denaturants, such as guanidinium chloride, to promote aq. solubility. However, with lipopeptides and long, apolar peptide fragments, the requirement for water solubility can be problematic. Unsurprisingly, poor fragment solubility has a detrimental effect on ligation rate with no or low levels of conversion observed after long reaction times of up to 1 week, making ligation impractical in these cases.[157] Due to the demand for hydrophobic synthetic membrane proteins, several methods have been developed to realise such ligations.

Firstly, organic co-solvents including DMSO, DMF and trifluoroethanol (TFE) can be well tolerated in NCL and have been used to assist in solubilising fragments.[157, 163] Additionally, NCL of hydrophobic peptide fragments has been demonstrated in anhydrous organic solvents such as DMF in the presence of triethylamine, as reported by Engelhard and co-workers.[164]

In cases where organic solvent addition is not useful or not well tolerated, such as with DMSO insoluble fragments, solubility issues have been tackled through the use of detergents and lipid matrices.[165, 166] When used at a concentration below the critical micelle concentration (CMC), detergents such as the non-ionic *n*-octyl-β-D-glucopyranoside have been employed to promote the formation of a water soluble detergent-peptide complex, as reported in the synthesis of the transmembrane protein NS4A through NCL by Bianchi *et al*.[167]

When used at concentrations above the CMC, micellar or liposomal assemblies (**Fig 1.12**) of detergent encapsulate the reagents, with the ligation reaction occurring in the lipidic environment in which reagents are present in high, localised concentrations. As reported by Boons and co-workers in the synthesis of a tripalmitoylated glyco(lipo)peptide, lipophilic reagents are sequestered into the lipid phase, increasing their local concentration and accelerating the ligation rate.[158]

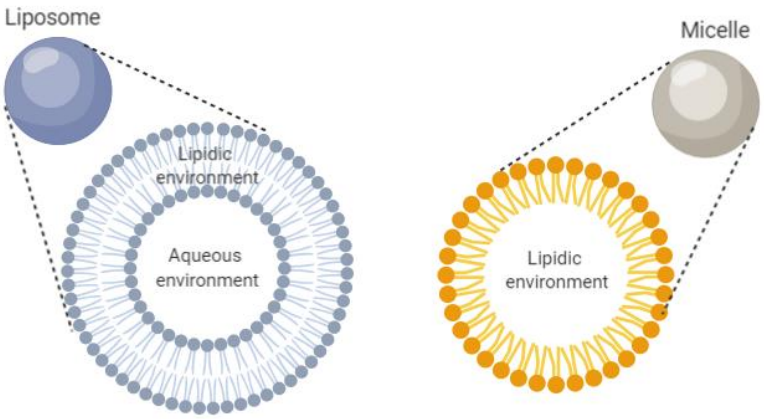


Figure 1.12. Cross-sectional view of a liposome and micelle.

In related efforts to improve the rate of NCL, Jin *et al.*[168] recently reported lipid-facilitated peptide ligation, a mixed micelle approach carried out in aq. buffer which obviates the need for thiol additives and external detergents. Both the thioester and cysteine fragment are temporarily lipidated to promote self-assembly into a mixed micelle, bringing the reactive groups into close proximity. The approach, which was illustrated with C-terminal ubiquitin derivatization and with the difficult sequence magainin 2, a 23-mer peptide, utilises a C_8 alkyl thioester on one peptide and an N-terminal cysteinyl peptide featuring a N_ε-lysine linked photolabile lipid proximal to the N-terminus on the second peptide. A significant rate enhancement was observed when compared to the non-lipidated controls. Notably, the reaction of the ubiquitin thioester Ubi-αC(O)SC_8 with the short peptide CK(PhC_{16})ANK to yield the ligated product Ubi-CK(PhC_{16})ANK was reported with 71% conversion after 5 h in the absence of a thiol catalyst. Contrastingly, the ligation of Ubi-αC(O)SC_8 with a non-lipidated control substrate, CKANK, achieved only 6% conversion in 5 h.[168]

Despite the challenges posed by the synthesis of hydrophobic peptides and proteins, such targets can be realised through the application of several extensions of the fundamental SPPS and NCL methodologies. These extended methods continue to be developed and improved upon, broadening the scope of peptide and protein synthesis to include ever more challenging biological targets.

1.5 Work Described Within This Thesis

As discussed throughout this chapter, the development of novel therapeutics is essential to combat the mounting threat of antimicrobial resistance. Bacterial lipoprotein biosynthesis involves multiple essential enzymes that have been established as potential antibiotic targets. Further development of these targets requires a comprehensive understanding of their catalytic function. By providing access to functionalised substrates of these enzymes, synthetic chemistry is uniquely poised to deliver novel tools to probe the mechanism and function of these enzymes. Building on previous work within the field of peptide synthesis, we aim to develop synthetic strategies to access a broad range of functionalised lipopeptide derivatives. In this thesis, we detail our interdisciplinary efforts to probe the function of 3 enzymes: LspA, Lnt and Lit, *via* the synthesis and application of fluorescently labelled lipopeptide substrates of these enzymes.

Chapter 2 details a linear synthetic strategy towards peptides containing an internal dagylated cysteine residue which act as substrates of LspA. We first present the preparation of a lipidated cysteine precursor which is subsequently incorporated into a peptide *via* SPPS. Additionally, the modification of such peptides with Förster resonance energy transfer (FRET) pairs is investigated. In collaboration with the group of Prof. Martin Caffrey, the resulting unimolecular FRET probes are screened for activity with LspA to facilitate the development of an optimised FRET-functionalised substrate of the enzyme. The identified optimal FRET probe is utilised in the development of a high-throughput assay for LspA, ultimately culminating in the screening of a major compound library to identify novel inhibitors of this enzyme.

Utilising the synthetic strategy developed in Chapter 2, Chapter 3 presents the synthesis of a peptide N-terminally labelled with a dagylated cysteine residue. Due to the unprotected amine present on the N-terminus, such peptides mimic the product of the LspA reaction and thus are primed for *N*-acylation by Lnt. The target peptide is utilised in a thin layer chromatography (TLC)-based assay for Lnt and attempts to extend the application of the peptide to a FRET-based assay are detailed. Furthermore, we investigate the synthesis of monomeric dagylated cysteine derivatives to explore the minimal structural requirements for Lnt recognition.

In Chapter 4, derivatives of the Lnt substrate identified in Chapter 3 are prepared to probe the mechanism of Lit. We investigate the regioselective acylation of the glyceryl

functionality to produce a Lit substrate which contains deuterium labelled acyl chain on the glyceryl stereospecific numbered C-2 (*sn*-2) position. We explore Lit-mediated acyl transfer of this compound and utilise NMR spectroscopy to unambiguously characterise the product of the Lit reaction, thus providing evidence of the structure of lyso LPs. Additionally, we present evidence of the Cu(II) binding capacity of diacylated lipopeptides through the application of paramagnetic NMR.

Chapter 5 details an alternative synthetic route to DAG-modified peptides. This novel convergent methodology involves late-stage modification of peptides *via* liposome-mediated thiol-Michael addition of a DAG-functionalised thiol to afford a native DAG modification. Furthermore, we explore the application of dehydroalanine in a novel peptide ligation methodology through β,γ-*C,S* thiol-Michael addition of thioacids to dehydroalanine containing peptides.

The overall conclusions from this work are detailed in Chapter 6 which also outlines the future direction of this research.

Finally, Chapter 7 details the experimental procedures and the associated characterisation of compounds synthesised in this work.

Chapter 2

The Development of Synthetic Substrates of Lipoprotein Signal Peptidase II

2.1 Introduction

Lipoprotein signal peptidase II (LspA) is a membrane-embedded peptidase responsible for cleaving a signal peptide from prolipoproteins during bacterial lipoprotein maturation. LspA is an essential enzyme in Gram-negative bacteria, necessary for full virulence of Gram-positive species, and crucial for the survival of some Gram-positive strains under physiological conditions.[77] With no eukaryotic equivalent, this enzyme represents a promising antibiotic target. Despite being under investigation for decades, identification of inhibitors has been hindered by the lack of efficient HTS methods for this membrane protein. This project aims to develop synthetic substrates of LspA from *P. aeruginosa* (LspPae) and LspA from MRSA (LspMrs) to facilitate the development of HTS assays to identify and optimise of novel inhibitors of these enzymes.

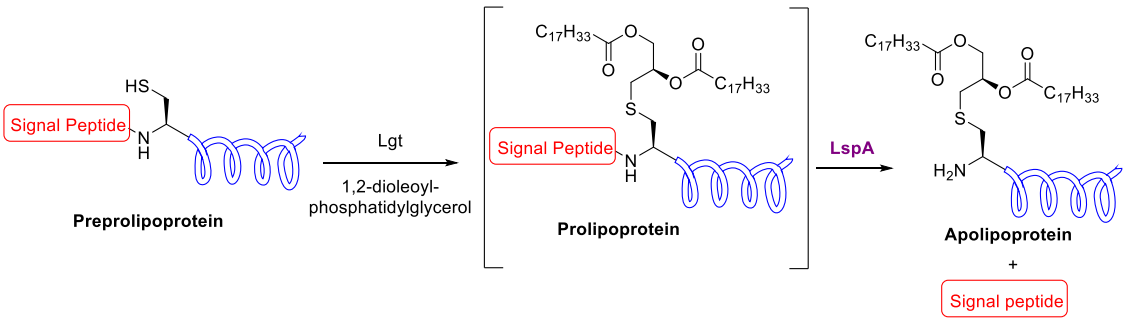
2.1.1 LspA as an Antibiotic Target

Two natural antibiotics which target LspA have been identified: globomycin and myxovirescin.[92, 169] Synthetic derivatives of both compounds have shown potent activity against Gram-negative and Gram-positive bacteria. However, translation of these antibiotics into the clinic has been prevented by low *in vivo* activity coupled with stability, production, and scalability issues.[93-95] In Gram-negative bacteria, the position of LspA at the IM makes this enzyme challenging to target as the microbe is protected by the OM. However, as the AMR crisis continues to worsen, targets previously classified as difficult to target must be revisited. Advances in protein crystallography and both computational and synthetic chemistry have inspired the re-exploration of targets such as LspA. Developments in membrane protein crystallisation, for example, have enabled the crystallisation of multiple LspA orthologues. The structures of LspPae in complex with globomycin, and LspMrs in complex with globomycin and with myxovirescin have recently been reported by Caffrey and co-workers.[77, 86] The structural information these studies provide about the binding pocket of the enzyme, coupled with *in silico* modelling, has informed the rational design of both inhibitors and synthetic substrates of LspA.

2.1.2 An Existing LspA Assay

Vogeley *et al.*[86] have previously reported a gel-shift assay for LspPae and LspMrs. To monitor the activity of LspA, this assay utilises a prolipoprotein native to *P. aeruginosa* as a substrate of the enzyme. Both the consumption of this substrate by LspA and the resulting product formation can be visualised by sodium dodecyl sulfate-polyacrylamide

gel electrophoresis (SDS-PAGE), providing a means through which reaction progress can be monitored. The LspA substrate, pro-inhibitor of cysteine protease (pro-ICP), is produced *in situ* by the Lgt catalysed reaction of recombinant pre-pro-ICP with a phospholipid. Pre-pro-ICP is first expressed and purified, then treated with Lgt. The Lgt reaction introduces diacylglycerol (DAG), a moiety which is essential for LspA recognition, onto the thiol of cysteine in pre-pro-ICP, to yield pro-ICP. After 1 h, LspA is added to the Lgt reaction. The two steps occurring in this one-pot process are depicted in **Scheme 2.1**. After a fixed time, the enzymatic reactions are stopped by SDS addition. The products of the LspA reaction are separated from the reaction media by SDS-PAGE. Staining and imaging of the gel permit quantification so that LspA activity can be measured.



Scheme 2.1. A previously reported Lgt coupled LspA assay involves *in situ* production of a LspA substrate.[86] Proteolysis of this substrate by LspA results in apolipoprotein and signal peptide production, which are separated and quantified using SDS-PAGE to reveal LspA activity.

This assay has been used to measure the activity of LspA mutants, to estimate the specific activities of LspPae and LspMrs, and to determine IC_{50} values for globomycin and myxovirescin.[86] Advantageously, the native substrate used in this assay is an accurate biomimetic, meaning the behaviour of the enzyme in its native environment is likely to be accurately represented. However, this assay is complex and has several limitations. Two enzymatic reactions are occurring sequentially in one-pot: lipidation by Lgt, followed by proteolysis by LspA. Resultantly, it is possible that the LspA substrate is being continuously generated throughout the assay at an undetermined rate. As the exact concentration of substrate is therefore unknown, quantitative kinetic information cannot readily be obtained. Additionally, the assay is slow to perform, does not provide a continuous readout and is not suitable for high-throughput applications. It was therefore of considerable interest to develop an assay which overcame these limitations and was suitable for both HTS and kinetic analysis.

An activity assay must meet certain criteria to be considered viable for HTS.[170] It must be simple to run and generate a readout which is readily detected and interpreted, such as a fluorometric response. The assay needs to be robust and reliable to minimise the detection of false positives and negatives. A signal to background ratio of >3 and a Z-prime of >0.6 should be established, with tolerance to DMSO addition demonstrated. Affordability is a critical consideration because scalability is essential. Therefore, the assay should use a small volume (<30 μL) and require minimal material to limit cost. These materials must be accessible on a suitable scale. To meet these criteria, a FRET-based assay for LspA was envisioned.

2.1.3 Förster resonance energy transfer

Resonance energy transfer (RET) is a photophysical process through which energy emitted by an excited donor species is transferred to an acceptor species. In systems which utilise a fluorescent donor, RET is classified as Fluorescence Resonance Energy Transfer or FRET. FRET is a mechanism of fluorescence quenching arising from energy transfer between two proximal chromophores. These two chromophores are paired due to their spectral overlap. In a FRET process, relaxation of an electron into a lower energy state in the fluorophore (donor) does not result in fluorescence, instead, this energy is transferred to the quencher (acceptor) *via* a virtual photon (**Fig. 2.1**). The wavelength of emission of the donor must match the absorbance wavelength of the acceptor for efficient energy transfer to occur.[171] FRET is a distance-dependent phenomenon, occurring only when the donor and complimentary acceptor are within close proximity of one other. As a result of its short-range R^{-6} distance-dependence described by Förster, FRET is a useful tool for measuring the spatial separation of a donor and acceptor in chemical and biological systems.[172] FRET-based assays have been used extensively to monitor the activity of various protease enzymes for example.[173] Recently, the application of FRET probes as tools for investigating glycosidase enzyme activity has been reviewed.[174]

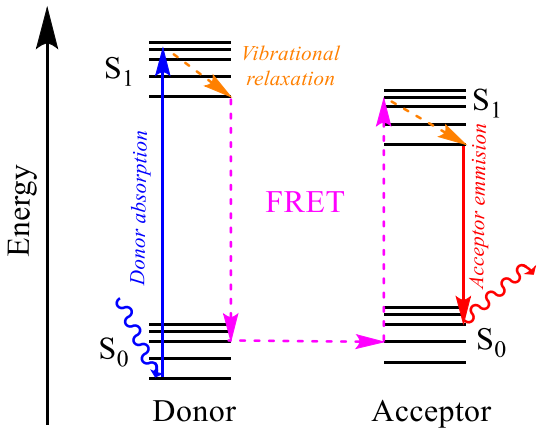


Figure 2.1. Jablonski diagram showing energy transitions occurring in a FRET system. Donor absorption is shown in blue, vibrational relaxation shown in orange, the coupled transitions between donor emission and acceptor absorbance shown in pink, and acceptor emission is shown in red.

In the single-molecule FRET probe envisioned for this project, both the donor and acceptor would be incorporated into the same molecule, tethered together by a peptide to limit their spatial separation. Excitation of this molecule at the characteristic absorption wavelength of the donor would result in FRET, with minimal energy released *via* donor fluorescence. The probe would act as a substrate of LspA, with LspA mediated hydrolysis of the peptide leading to increased separation of donor and acceptor. Peptide hydrolysis would disrupt FRET and result in an increased fluorescence signal at the donor's characteristic emission wavelength. As such, LspA activity could be quantified by measuring the change in fluorescence, allowing for kinetic analysis and determination of inhibitor IC_{50} values.

2.2 Aims of this work

This work is part of a collaborative project led by the biochemistry group of Prof. Martin Caffrey (TCD). The ultimate aim of this project is to identify novel small-molecule and peptide-based LspA inhibitors which have potential for clinical development. To achieve this, we aimed to develop efficient synthetic strategies to access LspA substrates to facilitate the development of high-throughput FRET-based assays for LspPae and LspMrs. Multiple peptide-based probes for LspPae and LspMrs would be synthesised to develop scalable synthetic access to novel FRET probes for HTS and permit investigation of the substrate specificity of each enzyme This chapter will describe the synthesis of LspA FRET probes and explore the synthetic challenges associated with accessing such lipopeptides.

The initial focus of this project was to design and synthesise FRET probes for LspPae. *P. aeruginosa* is a Gram-negative bacterium classified by the WHO as one of the top three most critical drug resistant pathogens. This ubiquitous microbe can cause fatal chronic infections in immunocompromised patients, with such infections being the major cause of fatality in cystic fibrosis (CF) patients.[175] The availability of effective anti-pseudomonal drugs is of particular importance for Ireland, the country with the highest incidence of CF *per capita*.[176] The staggering ability of this bacterium to evade antimicrobials has resulted in a major shortage of treatment options for patients infected with MDR *P. aeruginosa*.[177] Novel antibiotics that target *P. aeruginosa* are urgently required. Research later focused on FRET probes for LspMrs. MRSA is a Gram-positive ESKAPE 'superbug' responsible for several epidemics across healthcare and community settings worldwide. MRSA is highly prevalent in nosocomial environments and is commonly associated with surgical wound infections and ventilator-associated pneumonia.[178] In the community setting, MRSA is the causative pathogen of the vast majority of cutaneous infections in the United States. Of these infections, 5-10% become invasive and can lead to life-threatening sepsis and necrotising pneumonia.[179] The worldwide prevalence of MRSA is a growing threat which requires the development of further treatment options to control outbreaks of such drug resistant *S. aureus* strains.

2.3 LspPae FRET Assay – First-Generation FRET Probe

2.3.1 Probe Design

LspA is an endopeptidase which selectively cleaves the amide bond on the N-side of the essential dipalmitoylated cysteine ($Cys(Pam)_2$) residue in prolipoproteins. This yields a cysteine amino group at the N-terminus of the resulting apolipoprotein. The proposed mechanism of peptide hydrolysis by LspA, an aspartyl protease, is shown in **Fig. 2.2**.[180] Analysis of the amino acid sequence of native LPs from *P. aeruginosa* has revealed the prolipoprotein consensus sequence around the essential $Cys(Pam)_2$ residue.[181] The scissile glycine-cysteine bond in the prolipoprotein is flanked on the N-terminus by hydrophobic residues of the membrane-embedded signal peptide, the most abundant of which are $[L]^{-3}[A]^{-2}[G]^{-1}[C(Pam)_2]^{+1}$.[50] C-terminally, cysteine is commonly neighboured by $[S]^{+2}[S]^{+3}$.[182] The structure of the target peptide probe **27** was based on this consensus sequence to promote substrate recognition by LspPae.

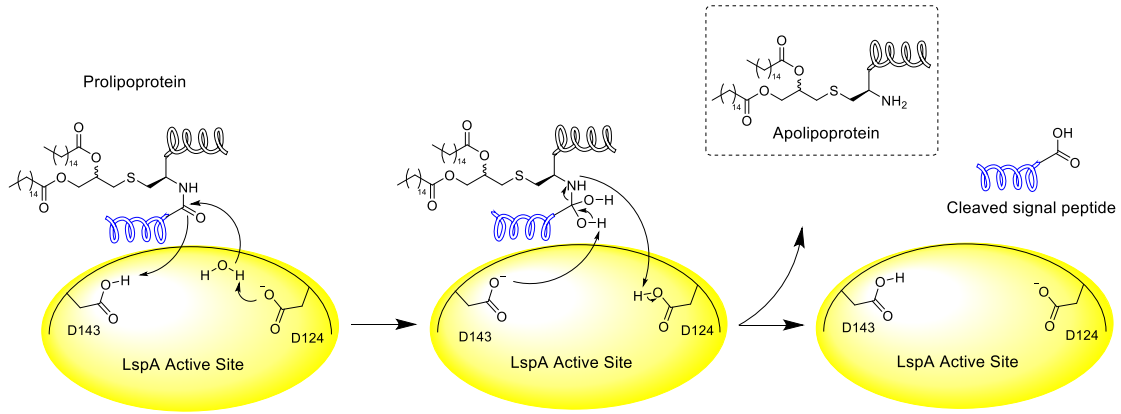


Figure 2.2. LspA mediated hydrolysis of the glycine-cysteine amide bond yields the apolipoprotein and the cleaved signal peptide. The proposed mechanism of an aspartyl protease is shown here with the proposed catalytic dyad of LspPae, Asp143 and Asp124, labelled.

Several factors governed the selection of FRET pair for this probe. The popular 2-aminobenzamide (Abz)/3-nitrotyrosine (3-NT) pair which was selected has appropriate spectral overlap between the emission of the Abz donor (λ_{ex}: 320 nm, λ_{em} : 420 nm) and absorbance of the 3-NT acceptor (basic pH λ_{ex}: 422 nm, acidic pH λ_{ex}: 355 nm,[183] pK_a = 6.74[184]). The Abz/3-NT Förster radius of 31 Å makes this FRET pair suitable for use with small-medium sized peptides with a molecular separation of 15-45 Å.[185] Additionally, these chromophores are small amino acids which, based on molecular modelling, are unlikely to destabilise the formation of the enzyme-substrate complex through steric hindrance or repulsive electronic interactions. Furthermore, 3-NT, the acceptor, is essentially non-fluorescent which minimises potential background fluorescence arising from direct excitation of the acceptor. Finally, both Abz and 3-NT can be readily introduced into a peptide sequence through SPPS following facile protection of their amino functionalities. Based on these principles, peptide **27** (**Fig. 2.3**) was identified as a promising FRET substrate.

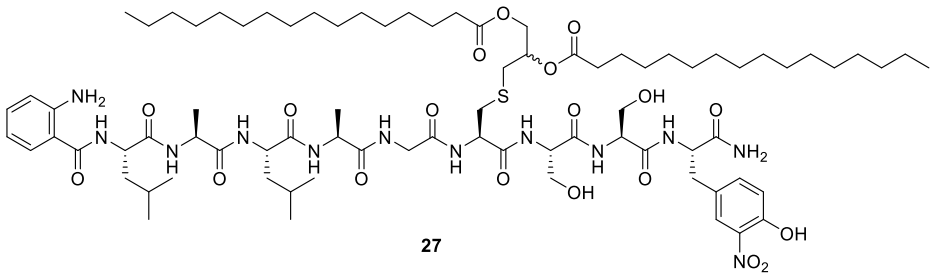


Figure 2.3. Structure of target FRET peptide **27**.

The potential of peptide **27** to act as a substrate of LspPae was supported by *in silico* modelling, the results of which are shown in **Fig. 2.4**.[a] This model shows the peptide docked with the scissile glycine-cysteine amide bond favourably positioned between the carboxyl groups of the proposed catalytic residues: Asp124 and Asp143 residues (shown in purple in **Fig. 2.4**, inset). This promisingly highlighted the potential for enzyme recognition, justifying the synthesis and testing of this peptide.

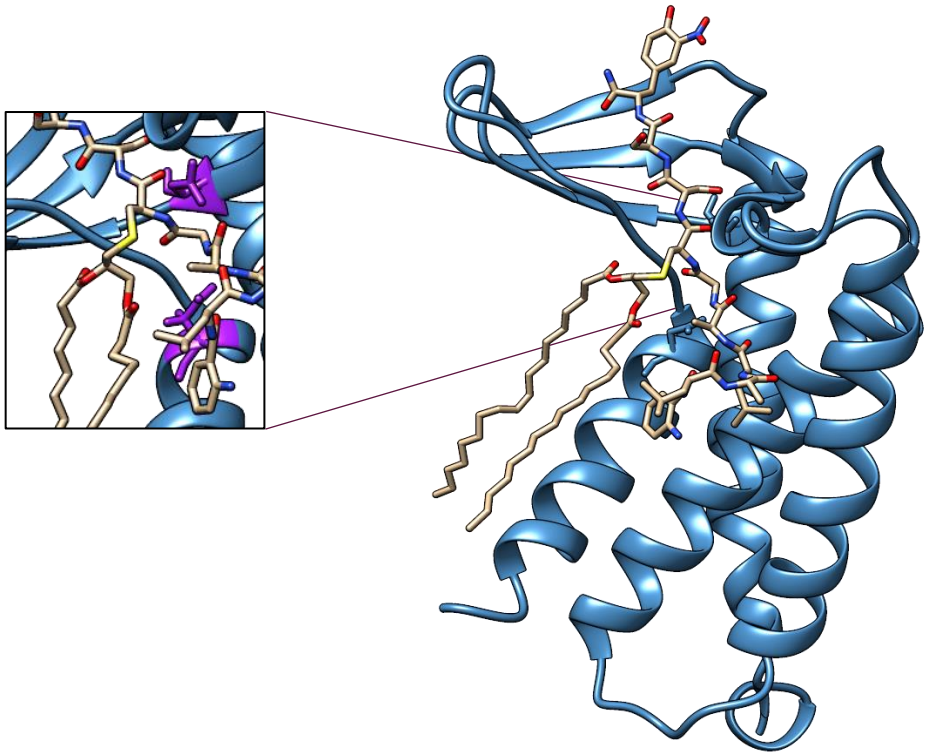


Figure 2.4. A model of LspPae (blue) interacting with the FRET lipopeptide **27** (grey)|. Inset: The scissile glycine-cysteine amide bond nestles perfectly between the two aspartyl residues (Asp124 and Asp143) which construct the proposed catalytic dyad of the active site.

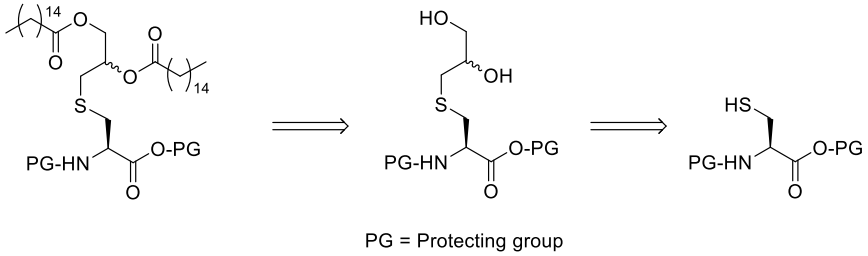
SPPS is an ideal strategy to access peptides of this size. A solution-phase approach would require an impractical number of coupling and chromatographic purification steps. As such, a solid phase synthetic strategy was utilised with the linear incorporation of Fmoc-Cys[(*R*/*S*)Pam$_2$]-OH (**28**) during the synthesis. Although the use of SPPS in modern peptide synthesis is widespread, the technique still has many limitations. One major drawback is the requirement of a large excess (3 equivalents (equiv.)) of protected amino acids in coupling reactions to force reactions to near completion in under 1 h. In many cases, this lack of atom economy is not prohibitive as standard Fmoc amino acids are

[a] Computational studies were performed by the group of Prof. Phillip Stansfeld, University of Oxford.

inexpensive and commercially available. However, the introduction of a modified amino acid requires the preparation of a substantial amount of the modified material. As such, incorporation of Cys(Pam)$_2$ into the target peptide first necessitated the development of an efficient synthetic strategy to access gram-scale quantities of Fmoc-Cys[(*R*/*S*)Pam$_2$]-OH (**28**).

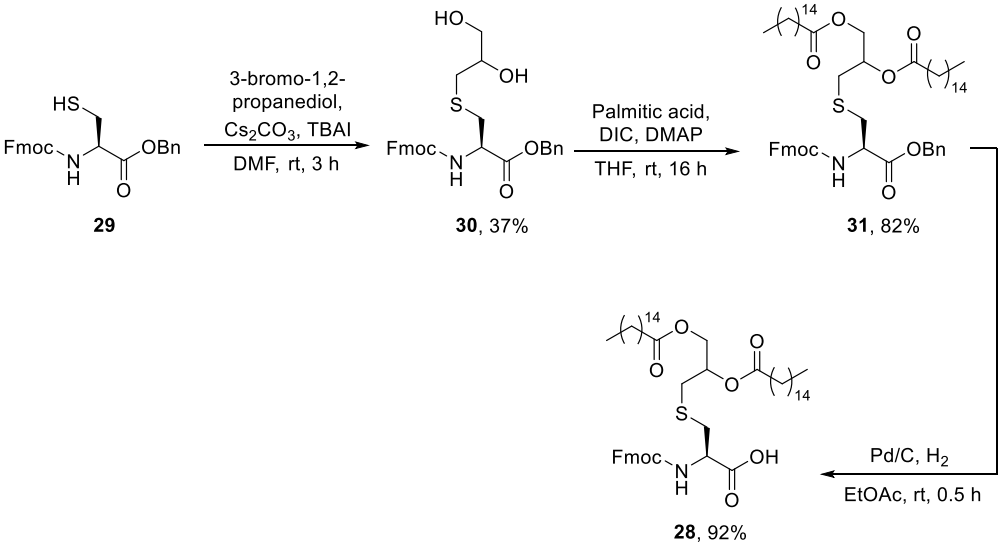
2.3.2 Synthesis of Fmoc-Cys[(*R*/*S*)Pam$_2$]-OH (28)

A synthetic strategy was developed which allowed comprehensive access to Cys(Pam)$_2$ with various protecting groups on both the N- and C-terminus. A strategy offering flexibility in the final protecting groups was highly desirable as it allowed for the application of Cys(Pam)$_2$ in both SPPS and solution phase synthesis. The key disconnections of this route are shown in the retrosynthetic analysis shown in **Scheme 2.2**.



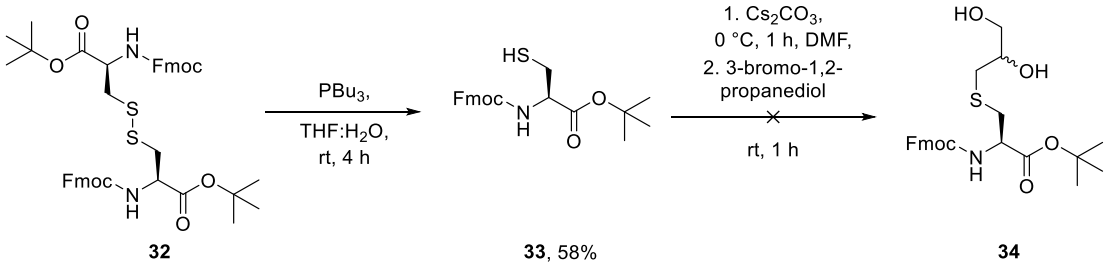
Scheme 2.2. Retrosynthetic analysis highlighting the key disconnections used in the synthesis of Fmoc-Cys[(*R*/*S*)Pam$_2$]-OH (**28**).

An initial synthetic strategy towards compound **28** was based on the work of Liu *et al.*[186] who reported the synthesis from Fmoc-Cys-OBn (**29**) in three steps with an overall yield of 30% as shown in **Scheme 2.3**.



Scheme 2.3. Synthesis of compound **28** reported by Liu *et al.*[186]

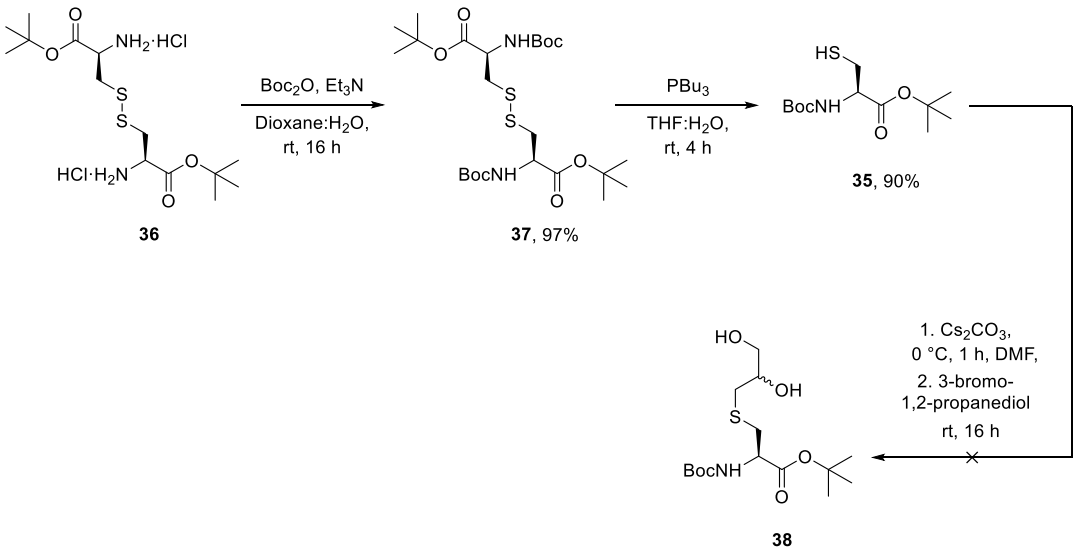
Due to reports of Fmoc deprotection under the reducing conditions required for debenzylation,[187] a *tert*-butyl ester was used to protect the carboxylic acid in place of the reported benzyl ester. The sensitivity of *tert*-butyl esters to TFA permits selective deprotection in the presence of the acid-stable Fmoc group. The use of Fmoc as the initial amine protecting group was preferable to avoid late-stage protecting group manipulation. As depicted in **Scheme 2.4,** tributylphosphine (PBu_3) mediated reduction of commercial (Fmoc-Cys-O*t*Bu)$_2$ (**32**) yielded Fmoc-Cys-O*t*Bu (**33**) in a yield of 58% after 4 h. However, *S*-alkylation of thiol **33** with 3-bromo-1,2-propanediol proved unsuccessful as the use of Cs_2CO_3 in DMF to generate the thiolate of compound **33** resulted in near quantitative deprotection of the Fmoc group after 1 h at 0 °C. No desired product was detected by TLC throughout the reaction.



Scheme 2.4. Preparation of thiol **33** and attempted S_N2 reaction with 3-bromo-1,2-propanediol.

Considering the sensitivity of Fmoc to basic conditions, the strategy was revised to introduce Fmoc late in the synthesis. Boc was chosen as the most appropriate initial amine protecting group due to its stability to basic reaction conditions and facile protection and

deprotection reactions. To prepare the suitably protected cysteine Boc-Cys-O*t*Bu (**35**), Boc was introduced onto the amine groups of commercially sourced disulfide (H-Cys-O*t*Bu)$_2$·2HCl (**36**) upon treatment with di-*tert*-butyl decarbonate (Boc_2O) and triethylamine (Et_3N) affording the product in a 97% yield. This was followed by the reduction of disulfide **37** using PBu_3, as shown in **Scheme 2.5**. Despite altering the amine protecting group, the S_N2 mediated *S*-alkylation of thiol **35** with 3-bromo-1,2-propanediol and Cs_2CO_3 yielded no detectable reaction by TLC after 16 h.

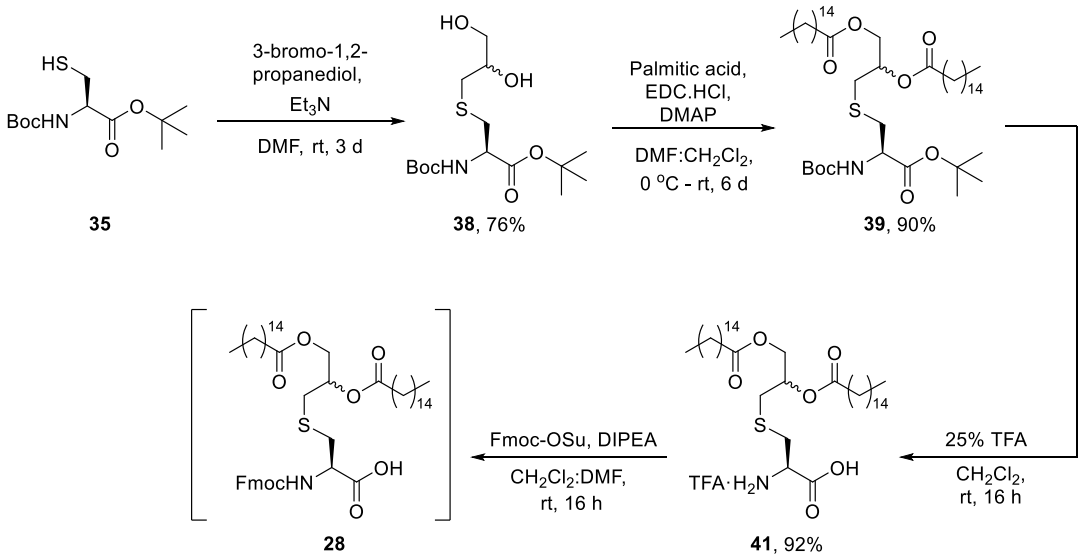


Scheme 2.5. Preparation and attempted *S*-alkylation of thiol **35**.

The base was subsequently changed to Et_3N (10 equiv.), an organic soluble base which allowed access to a higher concentration of base in solution. As a result, the desired alkylation reaction was observed by TLC but proceeded slowly, requiring 3 days to reach completion. During this time a minor amount of disulfide formation was observed by TLC, with this side product isolated by column chromatography in a yield of 8%. However, despite the slow reaction rate, diol **38** was isolated in an acceptable yield of 76% as shown in **Scheme 2.6**. Esterification of diol **38** was performed under Steglich conditions. Carboxylic acid activation using carbodiimides is a mild strategy commonly used in the preparation of both amides and esters. The *O*-acylisourea initially formed from the reaction of carbodiimide and carboxylic acid is susceptible to nucleophilic attack from amines but not alcohols. As Steglich reported in 1978, 3-10 mol% 4-dimethylaminopyridine (DMAP) is critical to promote the reaction of alcohols.[188] Nucleophilic catalysts, such as DMAP, react with the *O*-acylisourea, yielding a charged amide intermediate. This highly reactive amide is susceptible to nucleophilic attack from

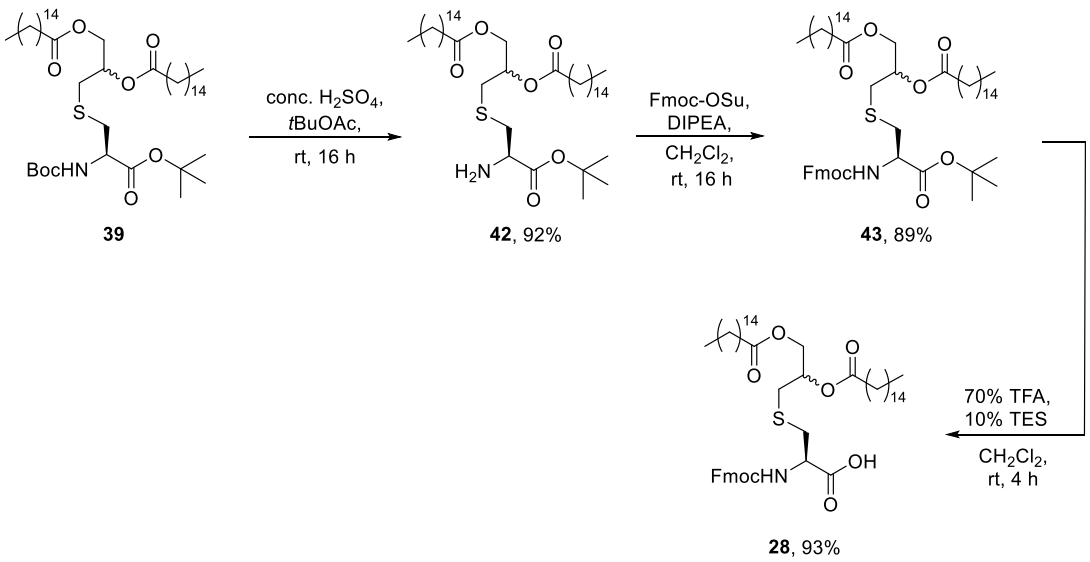
alcohols. In this case, palmitic acid was activated with *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (EDC·HCl) and esterification was catalysed by DMAP. The use of the classical carbodiimide coupling reagent *N*,*N*'-dicyclohexylcarbodiimide (DCC) was avoided due to the notorious purification issues which arise due to the poor solubility of the resulting dicyclohexylurea in both organic and aq. solutions. Additionally, attempted esterification in this case with *N*,*N*'-diisopropylcarbodiimide (DIC) was problematic due to the co-elution of the organic soluble diisopropylurea by-product with the product **39** during silica column chromatography. As such, EDC·HCl was used as the resulting urea is water soluble and readily removed by either aq. work up or silica column chromatography. Complete esterification of both glyceryl hydroxyl groups required 6 days to afford product **39** in a yield of 90%. The slow rate was attributed to the poor solubility of diol **38** in CH_2Cl_2:DMF (1:1 v/v). This was further indicated by the much faster rate observed with the CH_2Cl_2 soluble Fmoc protected derivative diol **40** (see **Section 2.5.1**). Removal of both Boc and *tert*-butyl ester protecting groups was performed with 25% (v/v) TFA in CH_2Cl_2. ^{1}H NMR analysis confirmed the dipalmitoyl esters were stable to TFA under these conditions. As TFA stability is a critical requirement for Rink-amide Fmoc SPPS, a further study was performed which revealed that compound **41** was stable to neat TFA for 24 h at room temperature (rt). The reaction was monitored by TLC. Similarly, the stability of this compound to 20% (v/v) piperidine in DMF over 24 h at rt was demonstrated with no change over this time detected by TLC.

Fmoc protection of amine **41** was attempted using 9-fluorenylmethyl *N*-succinimidyl carbonate (Fmoc-OSu). In the majority of cases, Fmoc-OSu has become the reagent of choice for Fmoc protection of *N*- and *C*-unprotected amino acids. The use of Fmoc-OSu minimises side reactions which are commonly observed with the more reactive Fmoc chloroformate (Fmoc-Cl), namely dipeptide formation.[189] Although TLC analysis revealed the reaction of amine **41** and Fmoc-OSu proceeded to completion, challenges in the separation of the product from the by-product *N*-hydroxysuccinimide and excess Fmoc-OSu rendered this route non-viable. The limited solubility of carboxylic acid **28** in organic solvents (CH_2Cl_2, EtOAc, Et_2O) prevented both aq. work up and column chromatography from yielding sufficient quantities of pure material.



Scheme 2.6. The synthesis of carboxylic acid **28** from thiol **35** shown here yielded the product, however, the desired compound was not isolatable following the Fmoc protection reaction, limiting the utility of this route.

An alternative synthesis of carboxylic acid **28** from Boc-protected **39** is shown in **Scheme 2.7**. Selective removal of the Boc group of compound **39** in the presence of the *tert*-butyl ester was performed using conc. H_2SO_4 in *t*BuOAc, a strategy first reported by Lin *et al.*[190] After 16 h, amine **42** precipitated from the reaction mixture and, following neutralisation, was readily isolated by filtration in a yield of 92%. The amine was re-protected using Fmoc-OSu to yield the readily organic soluble compound **43** in 89% yield following column chromatography. Finally, the *tert*-butyl ester was removed using 70% (v/v) TFA in CH_2Cl_2 yielding the target compound Fmoc-Cys[(*R*/*S*)Pam_2]-OH (**28**) with an overall yield of 51% over 6 steps from the commercially available cysteine derivative (H-Cys-O*t*Bu)$_2$·2HCl (**36**). This synthetic route was readily scalable to produce multi-gram quantities of carboxylic acid **28**.



Scheme 2.7. Protecting group manipulations of Boc-protected **39** to yield the target compound Fmoc-Cys[(*R*/*S*)Pam$_2$]-OH (**28**).

2.3.3 Synthesis of Abz-LALAGC(Pam$_2$)SSY(NO$_2$) (27)

With compound **28** in hand, manual SPPS of the target FRET peptide **27** was performed on Rink amide (aminomethyl)polystyrene resin with an initial loading of 0.7 mmol/g. A Rink-amide resin was used as peptide cleavage can be carried out under acidic conditions (95% aq. TFA) to which peptide **27** was anticipated to be stable. Furthermore, Rink amide cleavage generates a primary amide at the C-terminus of the peptide. This functionality was desirable as it more accurately mimics the native protein than a carboxylic acid.

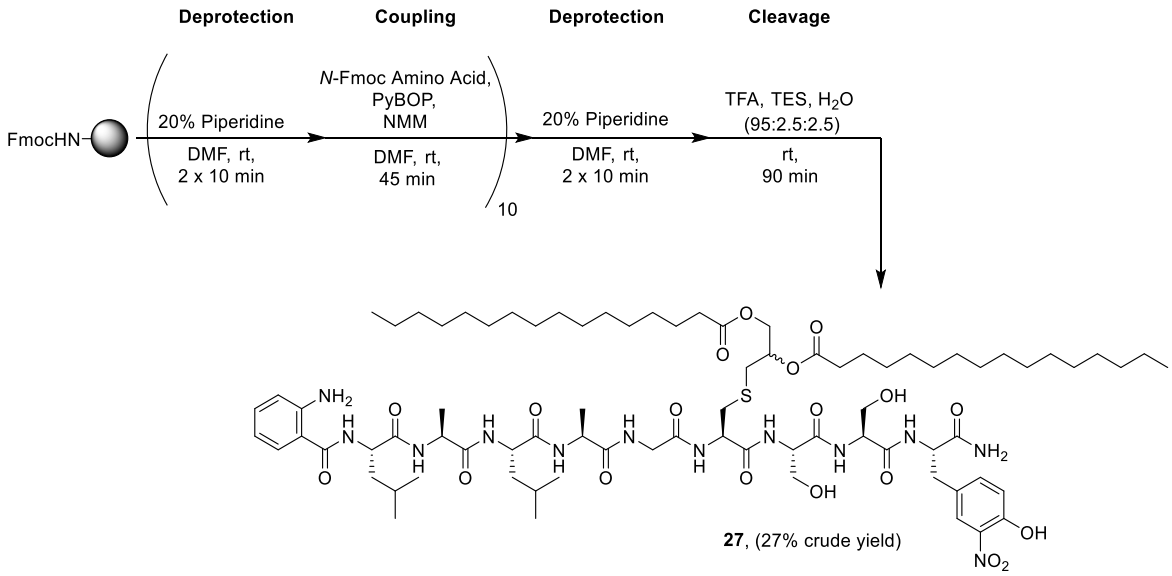
A resin consisting of polystyrene cross-linked with 1% DVB was used for this synthesis as reactions were performed in either DMF or CH_2Cl_2. Polystyrene (1% DVB) swells sufficiently in these solvents, ensuring reactive sites are solvent accessible.[191] In an attempt to mitigate aggregation of the resin-bound peptide, a relatively low resin loading (0.7 mmol/g) was used. A high resin loading (>1 mmol/g) may have promoted aggregation of this hydrophobic peptide, preventing reagents from accessing resin-bound functionalities and shutting down the efficiency of both coupling and deprotection reactions. The resultant truncated peptide impurities would be difficult to remove from the target compound, reducing the isolated yield.

Fmoc protected amino acids were incorporated through activation with benzotriazolyloxy-tris[pyrrolidino]-phosphonium hexafluorophosphate (PyBOP) and *N*-methylmorpholine (NMM) in DMF and coupled for 45 min. PyBOP is a powerful

phosphonium based coupling reagent which generates hydroxybenzotriazole (HOBt) *in situ* and is used extensively in SPPS.[192] PyBOP, a pyrrolidino derivative of benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate (BOP), was developed to overcome the safety issues associated with BOP which generates hexamethylphosphoric triamide (HMPA), a relatively volatile carcinogen, during coupling. With PyBOP, such hazards are avoided while the efficient coupling power of BOP is retained.[193] The use of basic conditions during the coupling reaction can lead to the occurrence of side reactions, such as Fmoc deprotection. Tertiary bases, such as NMM, have low reaction rates and hence minimise the formation of side products.[194] Despite the use of tertiary base, coupling times were limited to 45 min to minimise undesirable Fmoc deprotection.

Throughout the synthesis of peptide **27**, coupling reactions were monitored qualitatively using bromophenol blue in CH_2Cl_2 (0.15 mM) as described by Krchnak *et al.*[195] This indicator detects resin-bound amino groups, turning the resin blue in their presence. Following complete coupling of amino acids to the resin, the resin appears yellow when exposed to the indicator. When incomplete couplings were detected by the presence of blue resin, a second coupling was performed prior to Fmoc deprotection. Although less popular than the widely used Kaiser test,[196] this method avoids the need for toxic KCN.

Ten iterative rounds of Fmoc deprotection and amino acid coupling were performed to assemble the peptide on the resin as shown in **Scheme 2.8**. Fmoc deprotection was carried out by agitation of the resin in 20% (v/v) piperidine in DMF for 2 x 10 min. 3-NT was introduced as the C-terminal residue by coupling of Fmoc-Tyr(3-NO_2)-OH (**44**) to Rink amide resin. The serine residues were coupled with hydroxyl side groups protected by the acid sensitive *tert*-butyl ether group. Abz was introduced Boc protected as Boc-Abz-OH (**45**), prepared as described by Broomfield *et al.*[197] These protecting groups were selected due to their lability to TFA, which is used in the final resin cleavage step. Prior to resin cleavage, a final treatment of the peptide resin with 20% (v/v) piperidine in DMF (2 x 10 min) was carried out to cleave the potential phenolic ester on the unprotected 3-NT hydroxyl group. The formation of an ester at this position during SPPS of 3-NT containing peptides has been previously described by Singh *et al.*[198] who also reported its hydrolysis under treatment with 20% (v/v) piperidine in DMF. Global deprotection and resin cleavage were effected upon treatment with TFA:TES:H_2O (95:2.5:2.5 v/v/v).



Scheme 2.8. SPPS of peptide **27** on Rink amide resin.

Following cleavage from the resin, the peptide containing solution was concentrated and the peptide was precipitated in Et_2O to give peptide **27** in a crude yield of 27%. Purification of this peptide proved to be problematic. The highly amphiphilic compound **27** was poorly soluble in both aq. and organic solvent systems, with the exception of neat TFA and neat DMSO. Routine analysis and purification of the peptide by RP-HPLC produced unsatisfactory results despite extensive method optimisation.

Solubilisation of the crude peptide in MeCN (+1% TFA) was limited to a maximum concentration of ~0.01 mg/mL. Analytical HPLC was performed on the crude peptide using a C_{18} column with a gradient of MeCN:H_2O (1:1 v/v) to MeCN over 50 min with a photodiode array (PDA) detector monitoring wavelengths of 280 and 360 nm. The major component of the mixture was eluted at t = 28.3 min (**Fig. 2.5**). Considering the limited solubility of this compound in the mobile phase and its high affinity to the stationary phase, it is unsurprising that the use of a C_{18} column resulted in a poor peak shape as shown in **Fig 2.5**. From the resulting chromatogram it is not possible to determine if the broad, shouldered peak at 28.3 min indicates poor resolution of multiple compounds or streaking of a single compound. Attempts to overcome this by reducing sample loading and changing the stationary phase to C_8, C_5 and C_4 did not improve peak shape or significantly alter retention to the column. Considering these limitations, the purity of the crude peptide was approximated to be 65% with rt = 28.3 min.

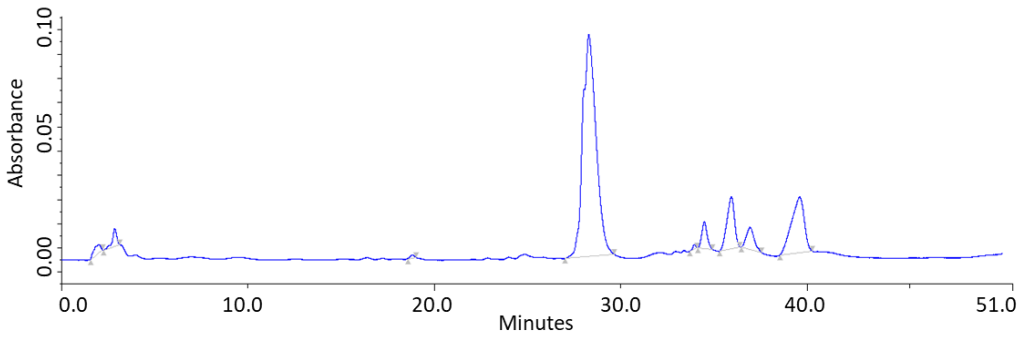


Figure 2.5. Chromatogram of crude peptide **27** visualised at a wavelength of 280 nm.

Attempts to transfer the optimised analytical HPLC method to semi-preparative scale using either a C_{18} (250 x 10 mm) and a C_5 (250 x 10 mm) column were unsuccessful. Following sample loading, no compound was detected eluting from either column despite extensive multi-hour run times with 100% MeCN. At this larger scale, following mixing of the injected sample with the initial mobile phase of MeCN:H_2O (1:1 v/v), it is feasible that the sample precipitated or adhered to the guard column. The same result was observed with an initial mobile phase of 100% MeCN (+0.1% TFA). Purification of the peptide was subsequently performed using a C_{18} (150 x 4.6 mm) analytical column, however, due to the limited concentration per injection, only microgram quantities of the peptide could be obtained despite many iterative runs. MALDI-MS analysis confirmed the identity of the isolated target peptide. However, the quantity of isolated material was insufficient for biological analysis with LspA.

An attempt to improve crude purity through *N*-acetylation of unreacted amine groups following each amino acid coupling during SPPS was unsuccessful. Treatment of the resin with acetic anhydride (Ac_2O):pyridine (1:5 v/v) for 3 x 4 min prior to each Fmoc deprotection reaction was carried out during an attempted synthesis of peptide **27**. However, ^{1}H NMR analysis of the crude peptide which was obtained following resin cleavage revealed the absence of palmitoyl signals. Four characteristic signals of the palmitoyl groups (δ_H 2.26, 1.67, 1.23, 0.87 ppm), present in the ^{1}H NMR of the crude peptide **27** previously synthesised without capping, were missing and no product was identified by MS in the crude mixture. This suggests the palmitoyl esters were unstable to the capping conditions used in this synthesis and thus this strategy was not pursued.

2.3.4 Testing with LspA[b]

In the first instance, testing with LspA was performed with the crude peptide **27**. This preliminary work aimed to determine if peptide **27** was a substrate of LspPae and to assess its suitability as a FRET probe. The effect of LspPae was investigated by monitoring fluorescence following incubation of peptide **27** both in the presence and absence of LspPae. An excitation wavelength of 320 nm, the characteristic absorption wavelength of donor (Abz), was used.

The fluorescence spectrum of peptide **27** following incubation in buffer without LspPae at 37 °C for 18 h is shown in red in **Fig 2.6**. A minor amount of background fluorescence was detected, with a maximum emission wavelength at 420 nm, the characteristic emission wavelength of Abz. This likely arises due to incomplete internal quenching of the peptide or the presence of Abz containing impurities. Gratifyingly, incubation of peptide **27** with LspPae at 37 °C for 18 h resulted in intense emission at 420 nm (**Fig. 2.6**, black). The approximately 10-fold increase in emission intensity indicated that fluorescence quenching had ceased, confirming peptidolysis of peptide **27** by LspPae.

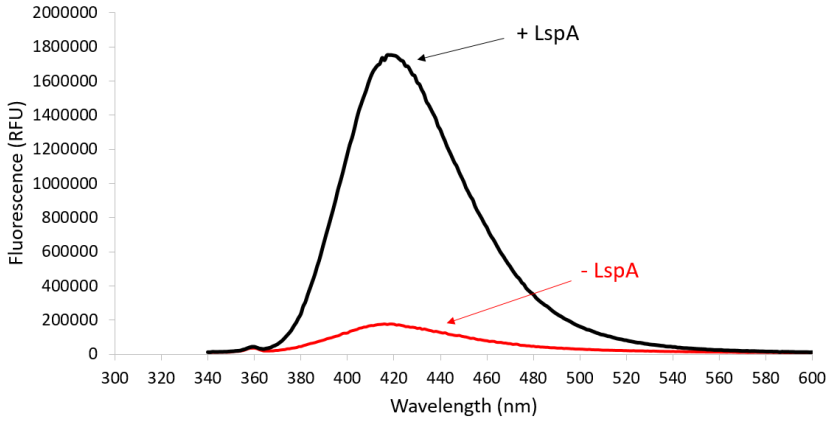


Figure 2.6.[b] A 10-fold increase in fluorescence intensity of the FRET peptide **27** upon excitation at 320 nm was detected at 420 nm following the incubation of peptide with LspPae.

Two reaction progress curves are shown in **Fig. 2.7** which depict the change in peptide fluorescence as a function of time. As expected, in the absence of enzyme (green data points) no change in fluorescence was detected. Gratifyingly, when LspPae was included in the reaction (grey, red and blue data points), a sharp increase in fluorescence was detected which plateaued as the concentration of peptide **27** decreased. These

[b] Biological work involving LspA and the assays discussed in this section were performed and developed by Dr. Jonathan Bailey and Dr. Samir Olatunji (Caffrey Group, Biochemistry, TCD).

experiments confirmed that the change in fluorescence intensity was a result of LspPae activity, thus proving that this peptide is a substrate of LspPae. The significant change in emission intensity caused by LspPae hydrolysis of peptide **27** highlighted the suitability of the compound as a FRET probe.

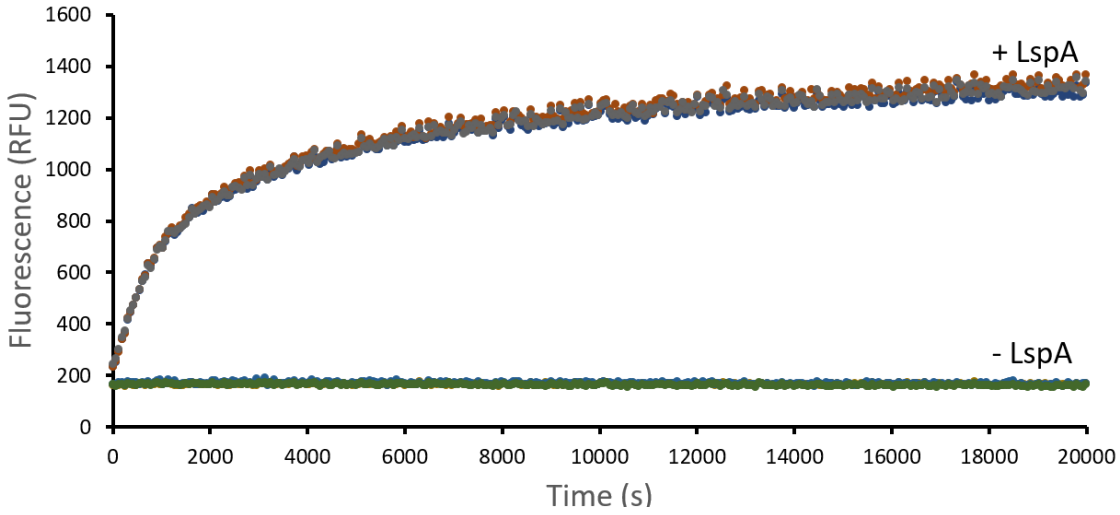


Figure 2.7.[b] The increase of fluorescence intensity of FRET probe **27** over 5.5 h. Measurements (performed in triplicate) were taken every 30 s for 5.5 h with an excitation wavelength of 320 nm and a detection wavelength of 420 nm.

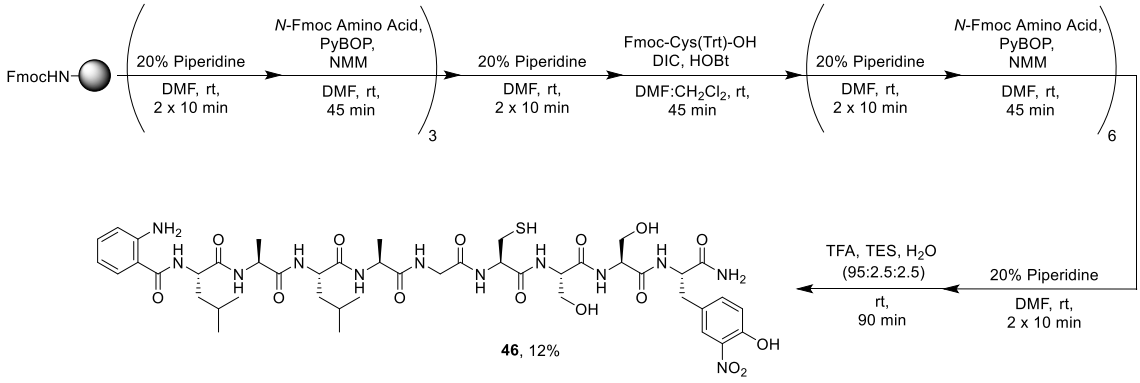
2.3.5 Requirement for Lipid Component

To investigate the importance of the DAG group in substrate recognition by LspPae, a peptide was synthesised which lacked the DAG modification at cysteine. As the processing of native preprolipoproteins by Lgt is essential before LspA can cleave the signal peptide,[52] we aimed to confirm that cysteine palmitoylation of the substrate is necessary for enzyme processing.

2.3.5.1 Synthesis of Abz-LALAGCSSY(NO_2) (46)

This peptide was synthesised using Fmoc SPPS in a similar manner to that previously described for peptide **27**. However, following the deprotection of the second serine residue, cysteine was introduced *via* coupling of Fmoc-Cys(Trt)-OH, as shown in **Scheme 2.9**. The trityl protecting group is acid labile and is deprotected during TFA-mediated resin cleavage. As base mediated activation of Fmoc-Cys(Trt)-OH with PyBOP and NMM is notoriously prone to racemization at the cysteine α-proton,[199] a base free activation approach was used to couple this amino acid, with activation *via* DIC and HOBt. The final six amino acids of the peptide were coupled using PyBOP and NMM as

previously described. Following a final 20% (v/v) piperidine/DMF treatment of the resin, the peptide was cleaved, and precipitated in Et_2O. The crude peptide was purified by semi-preparative C_{18} HPLC and the identity of the isolated peptide was confirmed by MS. The solubility and purification issues encountered in the purification of peptide **27** did not arise in this case, indicating the adverse effect of the DAG functionality on the purification of a peptide.



Scheme 2.9. Synthesis of non-lipidated peptide Abz-LALAGCSSY(NO_2) (**46**).

Incubation of peptide **46** with LspPae resulted in no change of fluorescence over 18 h.[c] This result, in contrast with the clear fluorescence change observed with peptide **27**, shows that this compound is not a LspPae substrate, confirming that DAG modification of cysteine is essential for recognition by LspPae. Discerning the essential nature of the hydrophobic interactions between the enzyme and the peptide aided in the design of later generations of FRET probes.

2.3.6 Synthetic Peptidase Products

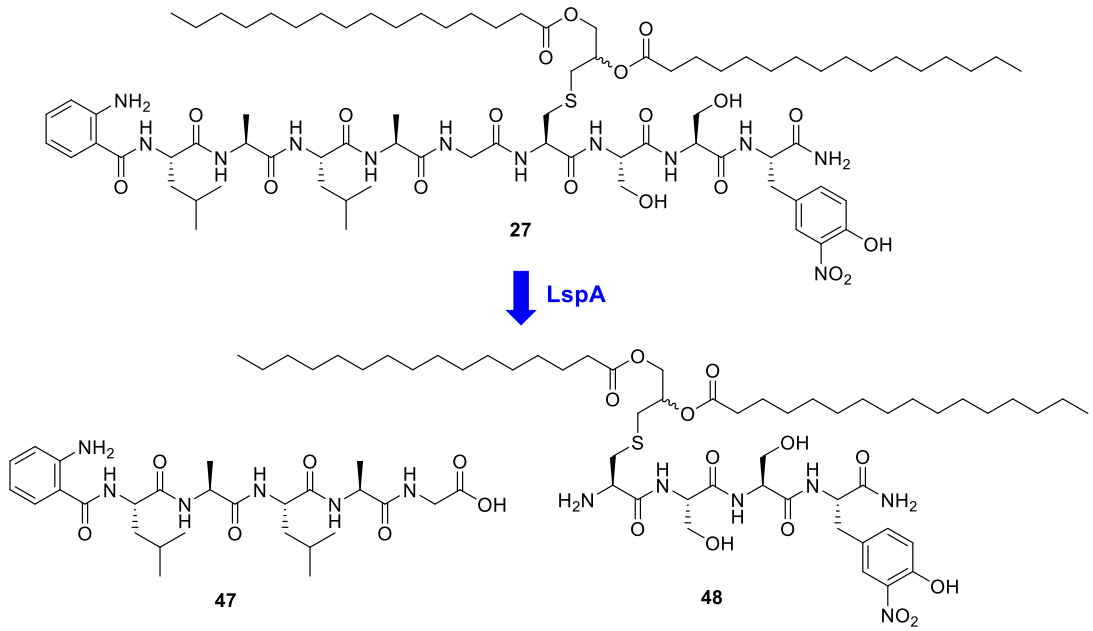
To determine the specific activity of LspPae with probe **27** it was necessary to quantify the fluorescence change in terms of moles of substrate consumed or products formed. To achieve this, a calibration curve was developed using the synthetic proteolysis products of the enzymatic reaction which are shown in **Scheme 2.10**.

The hexapeptide **47** was synthesised by SPPS on Wang resin to provide the peptide carboxylic acid. Initial attempts to synthesise this peptide using PyBOP and NMM to activate the first amino acid, Fmoc-Gly-OH, to couple to the resin, followed by standard deprotection and PyBOP/NMM coupling of the subsequent five residues resulted in a low crude peptide yield of 4%. Altering the loading conditions for the coupling of Fmoc-Gly-

[c] Experiment performed by Dr. Jonathan Bailey (Caffrey Group, Biochemistry, TCD).

OH to the resin by including DMAP (0.1 equiv.) as a nucleophilic catalyst was essential to achieve satisfactory loading to the resin, due to the poor nucleophilicity of the hydroxyl functionalised Wang resin. Using these optimised loading conditions, peptide **47** was synthesised and purified by silica column chromatography in an overall yield of 12%. The structure of the peptide was confirmed by MS and NMR analysis.

The tetrapeptide **48** was synthesised in an identical manner to FRET peptide **27**, until the coupling of Fmoc-Cys[(*R*/*S*)Pam$_2$]-OH (**28**), at which stage a final resin treatment with 20% piperidine in DMF was performed and followed by standard TFA cleavage from the resin. The peptide was precipitated in Et_2O and purified by silica column chromatography affording tetrapeptide **48** in an overall yield of 13%. The identity of the peptide was confirmed by MS and NMR analysis.



Scheme 2.10. LspA mediated hydrolysis of peptide **27** yields hexapeptide **47** and tetrapeptide **48**. To access sufficient quantities of pure compounds **47** and **48**, both peptides were synthesised through SPPS.

To determine the specific activity of the enzyme with peptide **27**, a standard curve of fluorescence as a function of moles of product was created using equimolar amounts of the two synthetic products **47** and **48**.[d] This confirmed linear fluorescence emission at the substrate concentration used in the assay (30 µM). However, the observed fluorescence resulting from the reaction of peptide **27** with LspPae was much lower than the expected

[d] Performed by Dr. Jonathan Bailey (Caffrey Group, Biochemistry, TCD).

fluorescence based on the synthetic products. It is feasible that this discrepancy may arise from the enzymatic reaction not proceeding to completion, potentially due to unidentified SPPS side products interfering with enzymatic activity, while Abz containing impurities may also impact fluorometric analysis. To confirm the fluorescence change can be attributed to hydrolysis of peptide **27** and to enable accurate investigation of the enzyme kinetics, the structure of the FRET peptide was redesigned with the aim of retaining LspPae recognition while enabling purification.

2.4 LspPae FRET Assay – Second-Generation FRET Probe

As biological experiments with compound **27** were limited by peptide purity, the structure of the probe was modified to that shown in **Fig. 2.8** to facilitate purification and access to sufficient quantities of material for assays. Two N-terminal hydrophobic residues were removed in an effort to alter the solubility and improve the purification of the peptide.

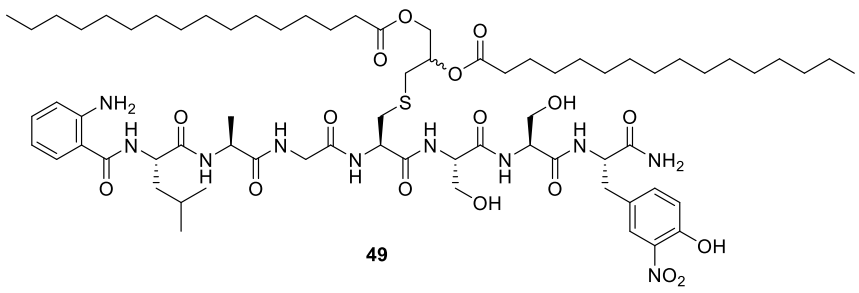


Figure 2.8. Structure of redesigned FRET probe **49**.

2.4.1 Synthesis of Abz-LAGC(Pam$_2$)SSY(NO$_2$) (49)

Peptide **49** was synthesised under the same conditions previously described for peptide **27** using Rink amide resin with PyBOP/NMM amino acid activation. The final resin-bound peptide was cleaved and precipitated in Et_2O. Although insoluble in aq. conditions, peptide **49** exhibited solubility in CH_2Cl_2:MeOH (1:1 v/v) unlike the initial FRET probe **27**. This permitted analysis by TLC (MeOH:CH_2Cl_2 1:9 v/v) and purification by normal phase silica column chromatography which afforded the target compound in an overall yield of 10%. Following purification, the structure of peptide **49** was confirmed by MS and homonuclear and heteronuclear 2D NMR analysis. Sufficient peptide purity was demonstrated by TLC and both ^{1}H and ^{13}C NMR.

2.4.2 Testing with LspA[e]

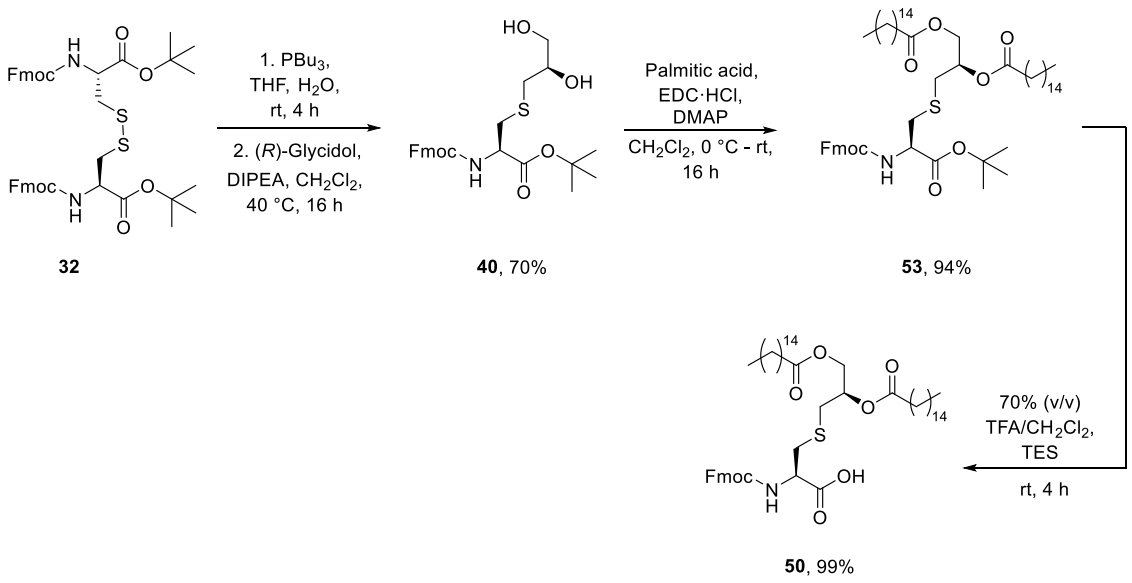
Gratifyingly, peptide **49** was found to be a substrate of LspPae, behaving comparably to peptide **27** under identical conditions with LspPae. However, despite the improved purity, incomplete hydrolysis of **49** was still evident. It was hypothesised that the enzyme may recognise only one of the two diastereomers present. To test this hypothesis, access to each glyceryl diastereomer (referred to here as (*R*)-DAG and (*S*)-DAG) was necessary. Separating the mixture of diastereomers in peptide **49** was deemed to be impractical as both compounds have identical R_f values on normal phase silica in MeOH:CH_2Cl_2 (1:9 v/v). Therefore, we aimed to synthesise each peptide separately. This first required synthetic access to Fmoc-Cys[(*R*)Pam_2]-OH (**50**) and Fmoc-Cys[(*S*)Pam_2]-OH (**51**).

2.5 LspPae FRET Assay – Third-Generation FRET Probe

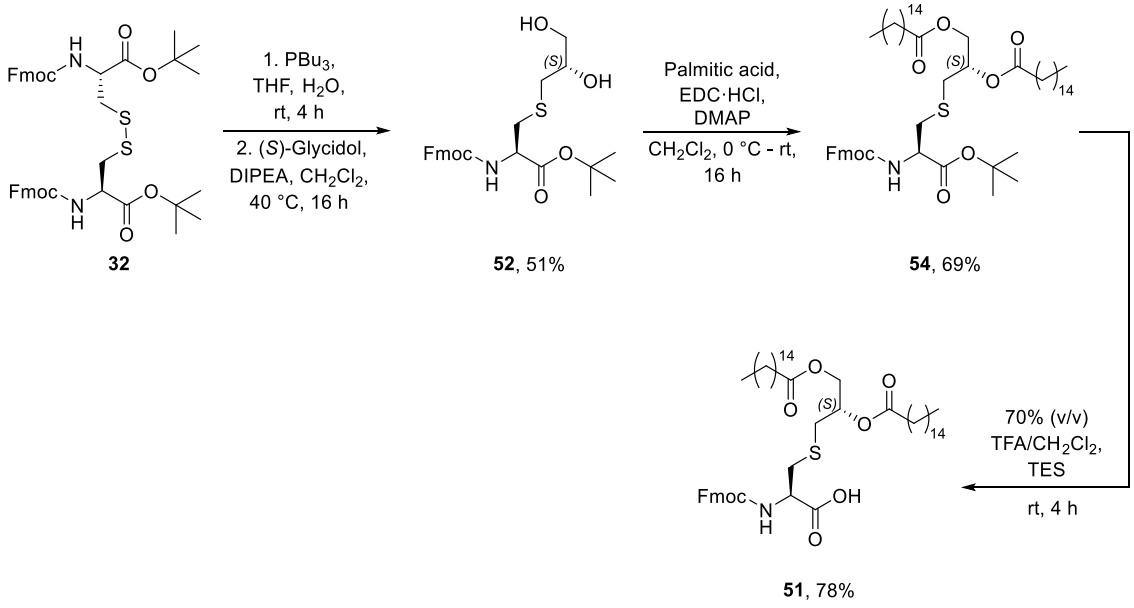
2.5.1 Synthesis of Fmoc-Cys[(*R*)Pam_2]-OH (50) and Fmoc-Cys[(*S*)Pam_2]-OH (51)

As the individual enantiomers of 3-bromo-1,2-propandiol were not commercially available, the synthesis of Fmoc-Cys[(*R*/*S*)Pam_2]-OH (**28**), as detailed in **Section 2.3.2**, was revised accordingly. Based on work by Metzger *et al.*[118] the synthesis was carried out starting from commercially sourced (Fmoc-Cys-O*t*Bu)$_2$ (**32**) in 3 steps as shown in **Schemes 2.11-2.12**. Both diastereomers were synthesised under the same conditions. The reduction of disulfide (Fmoc-Cys-O*t*Bu)$_2$ (**32**) was carried out using PBu_3 with quantitative conversion observed by TLC after 4 h. The solvents were evaporated, and the thiol was alkylated using the epoxide (*R*)-glycidol or (*S*)-glycidol yielding diols **40** and **52** respectively. As this reaction was carried out in one-pot, the re-oxidation of the thiol during alkylation was prevented by the presence of excess PBu_3. Initial alkylation attempts using 1.5 equiv. of glycidol resulted in a low isolated yield of 12% of diol **40**, despite TLC analysis revealing complete consumption of starting material. The crude NMR revealed a characteristic singlet at 6.05 ppm and doubling of the aromatic CH Fmoc signals at 7.73, 7.59, 7.37 and 7.28 ppm arising from 9-methylene-9*H*-fluorene, a product of Fmoc deprotection. To counteract this side reaction, the rate of the desired reaction was increased through the use of 5 equiv. of glycidol, enabling the isolation of diol **40** in a 70% yield and diol **52** in a 51% yield.

[e] Performed by Dr. Jonathan Bailey (Caffrey Group, Biochemistry, TCD).



Scheme 2.11. Synthesis of Fmoc-Cys[(*R*)Pam$_2$]-OH (**50**) from (Fmoc-Cys-O*t*Bu)$_2$ (**32**) in 3 steps with an overall yield of 65%.



Scheme 2.12. Synthesis of Fmoc-Cys[(*S*)Pam$_2$]-OH (**51**) from (Fmoc-Cys-O*t*Bu)$_2$ (**32**) in 3 steps with an overall yield of 27%.

Steglich esterification of diols **40** and **52** was carried out using palmitic acid activated with EDC·HCl and DMAP. In contrast to the similar reaction carried out with the Boc-protected derivative diol **38** described in **Section 2.3.2**, Fmoc-protected diols **40** and **52** were readily soluble in CH_2Cl_2 reducing the reaction time from 6 days required for diol **38** to 16 h and negating the use of DMF. Difficulties were encountered in the removal of excess palmitic acid from this reaction due to co-elution during column chromatography.

This was overcome through the addition of excess Amberlyst A21 resin to the reaction after 16 h when TLC analysis had indicated complete consumption of diol. This basic resin sequestered unreacted palmitic acid, with complete removal of palmitic acid from the solution observed after 1 h of stirring at rt. The Amberlyst-bound palmitic acid was subsequently removed by filtration and compound **53** and compound **54** were readily isolated from the filtrate by column chromatography. Both diastereomers **53** and **54** have the identical R_f values (0.60, EtOAc:hexane 1:4 v/v), however, the ^{1}H NMR spectra reveal a useful variation in the NH signal of each compound, with a difference of 0.02 ppm between the NH signals as shown in **Fig. 2.9**. Although this NH proton is not proximal to the altered stereocentre, the NMR spectrum shows the effect of the glyceryl-CH stereochemistry on the chemical environment of this characteristic NH resonance. The apparent triplet in the spectrum of the mixture of diastereomers **43** (synthesised in **Section 2.3.2**) arises from the overlap of these two doublets.

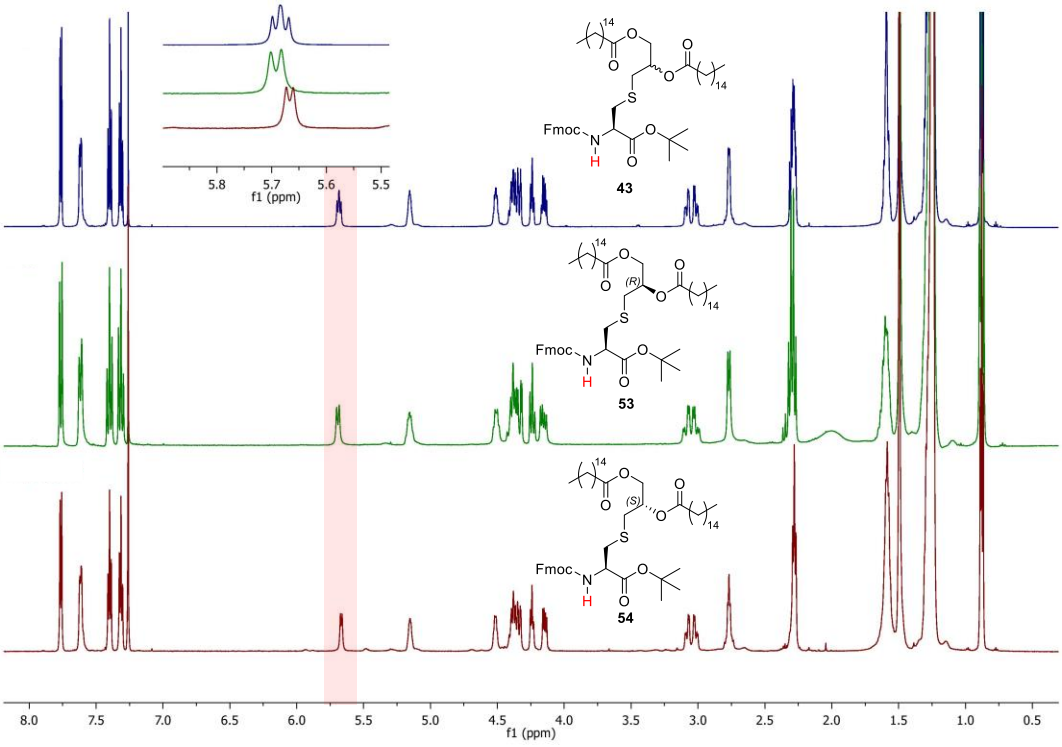


Figure 2.9. ^{1}H NMR (600 MHz, $CDCl_3$) of compound **43** (blue, top), **53** (green, middle) and **54** (red, bottom) with the NH signal highlighted. Inset: Expanded region showing the change in chemical shift between the NH signal in each diastereomer.

Similarly, comparison of the ^{13}C NMR spectra of compounds **43**, **53** and **54** reveals a slight shift of 0.08 ppm between the glyceryl-CH ^{13}C signal of each diastereomer as

shown in **Fig. 2.10**. Two signals are visible in the spectrum of the diastereomeric mixture **43**. This change indicates the electronic environment, and thus the chemical shift, of this carbon signal is dependent on the stereochemistry of this position. Along with the characteristic NH signal exhibited by each diastereomer, the single peak in the sample of **53** and **54** confirms the diastereomeric purity of each compound.

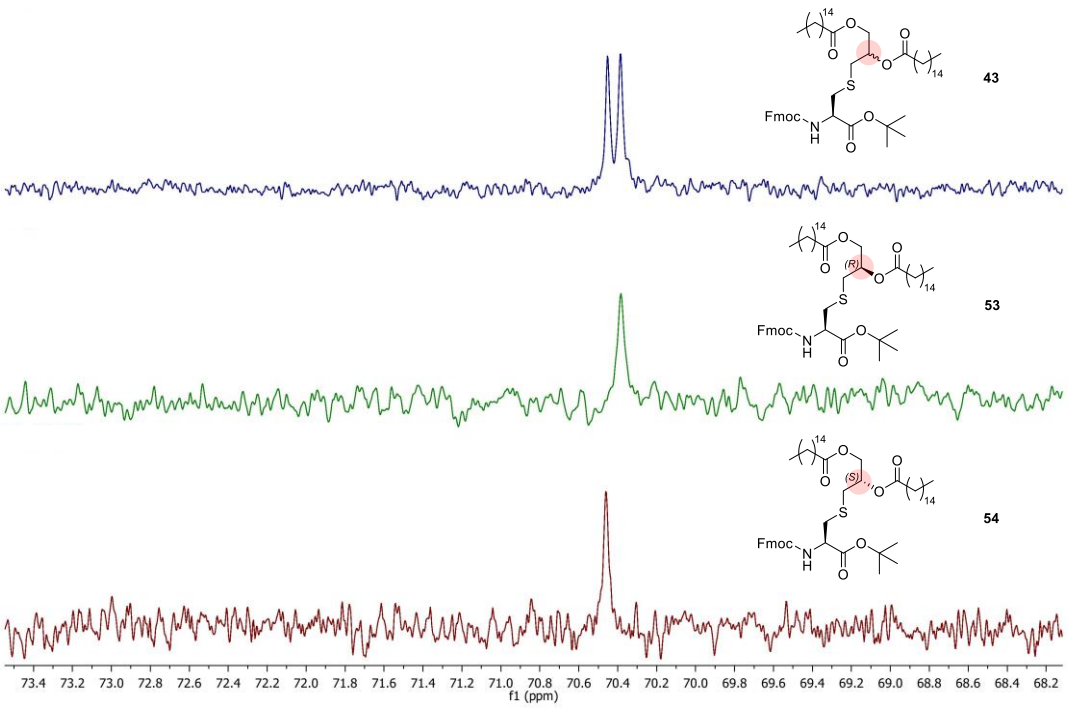
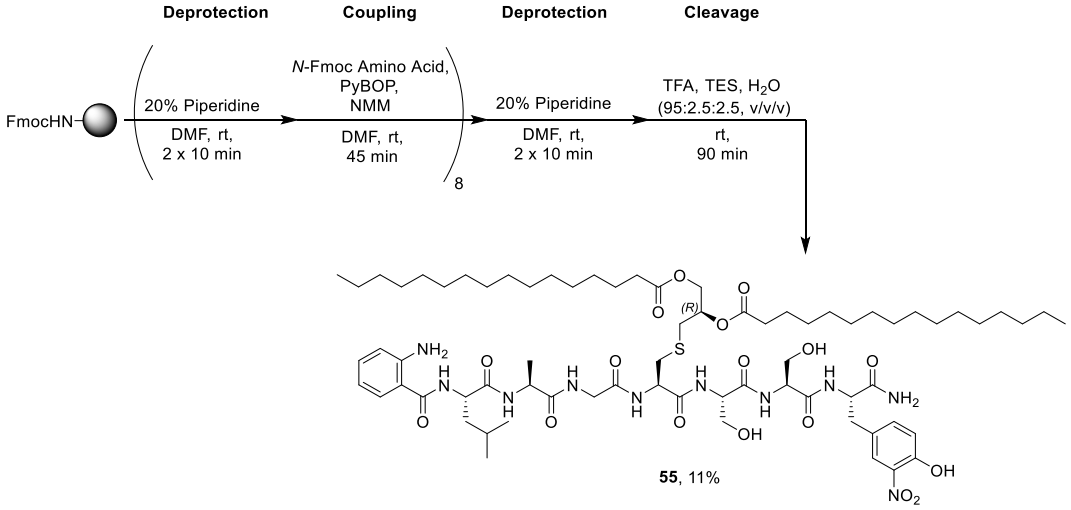


Figure 2.10. Expanded region of ^{13}C NMR (151 MHz, $CDCl_3$) comparison of compound **43** (blue, top), **53** (green, middle) and **54** (red, bottom) showing the glyceryl-CH signal (highlighted). The change in stereochemistry at this position influences of the chemical shift of this carbon resonance.

The *tert*-butyl esters of compound **53** and **54** were deprotected over 4 h using TFA:CH_2Cl_2:TES (7:2:1 v/v/v). TES was included as a hydride source to scavenge the *tert*-butyl cation released during deprotection. Due to the poorly resolved NMR signals of the resulting carboxylic acid **50**, a recrystallisation was performed, with potentially incomplete recovery from the mother liquor resulting in a yield of 78%. In contrast, the evaporation of volatiles yielded carboxylic acid **51** in a quantitative yield with NMR analysis confirming further purification was not necessary. This synthetic route proved to be scalable with moderate yields maintained at a larger scale at which 11.69 g of Fmoc-Cys[(*R*)Pam_2]-OH (**50**) was ultimately produced starting from 11.90 g of (Fmoc-Cys-O*t*Bu)$_2$ (**38**).

2.5.2 Synthesis of Abz-LAGC[(*R*)Pam$_2$]SSY(NO$_2$) (55)

The synthesis and purification of peptide **55** was performed under the same conditions discussed previously for peptide **49** (**Section 2.4.1**) and as shown in **Scheme 2.13**. The peptide was isolated in an 11% yield following normal phase silica column chromatography and the structure of the peptide was confirmed by 1D and 2D NMR and MS. This synthesis proved to be readily scalable to 0.20 mmol of resin. Above this scale, SPPS efficiency was reduced with increased side-product formation hindering effective purification. Isolated yields reduced to 2% at a 0.35 mmol scale. Through multiple parallel syntheses performed at 0.20 mmol, 250 mg of purified peptide was ultimately produced for HTS applications.



Scheme 2.13. SPPS of FRET probe **55** isolated in a yield of 11%.

Complete assignment of the ^{1}H and ^{13}C NMR spectra was possible using ^{1}H-^{13}C HSQC (**Fig. 2.11**), TOCSY, HMBC and ^{1}H-^{15}N HSQC NMR experiments. The diastereomeric purity of peptide **55** was confirmed by ^{13}C NMR. As depicted in **Fig. 2.12**, the glyceryl-CH signal (69.9 ppm) is a single signal in the diastereomeric mixture **49** and so this signal cannot be used to determine stereochemical purity. However, each diastereomer has a characteristic peak for the neighbouring glyceryl-CH_2-S signal (31.63 and 31.57 ppm). Only a single peak at 31.63 ppm is exhibited in the ^{13}C NMR spectrum of **55**, confirming diastereomeric purity.

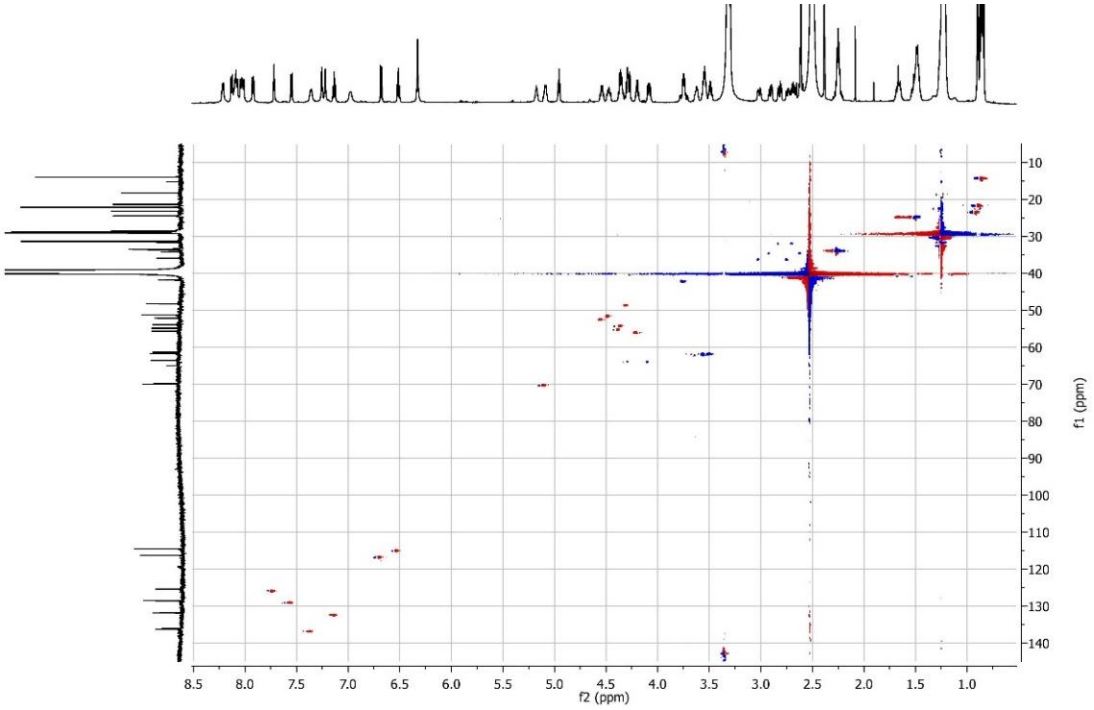


Figure 2.11. HSQC (600 MHz, DMSO-d_6) of peptide **55**.

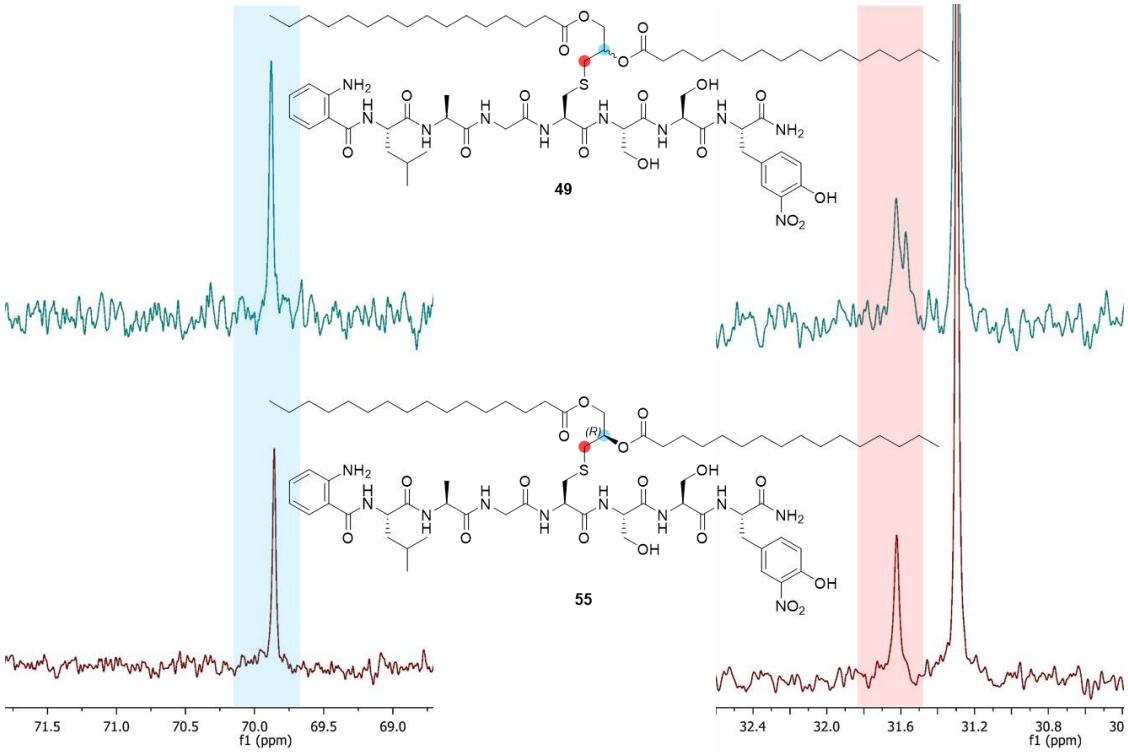


Figure 2.12. Selected peaks from the ^{13}C NMR (600 MHz, DMSO-d_6) spectrum of the diastereomeric mixture **49** (blue, top) and the single (*R*)-DAG diastereomer **55** (red, bottom). There is no distinction in the ^{13}C glyceryl-CH resonances (highlighted in blue), however, each diastereomer exhibits a distinct glyceryl-CH_2-S signal (highlighted in red: 31.63 and 31.57 ppm) providing a means to confirm the diastereomeric purity of **55**.

2.5.3 Testing with LspA[f]

As shown in the reaction progress curve in **Fig. 2.13**, peptide **55** is a substrate of LspPae. Addition of LspPae to **55** resulted in a linear increase in fluorescence shown here between 0-30 min (purple and green data points), signifying the release of a fluorescent product *via* peptide hydrolysis. When the enzyme was not included in the reaction (red and blue data points) there was no change in fluorescence, confirming that peptide hydrolysis requires LspPae. Reaction progress curves enable the initial velocity region, where <10% of the substrate has been processed, to be identified. Subsequent kinetic measurements to determine V_{max} and K_M of the enzyme were made in the 2-10 min range.

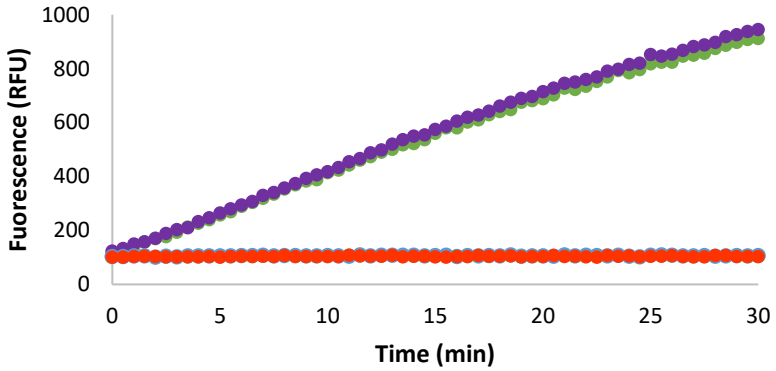


Figure 2.13.[f] Reaction progress curves showing the change in peptide fluorescence as a function of time. **Purple and green data points**: Addition of LspPae to peptide **55** results in an increase in fluorescence (experiment performed in duplicate). **Red and blue data points**: LspPae-free controls – no change in fluorescence occurs in the absence of the enzyme (experiment performed in duplicate). Measurements were recorded every 30 s for 30 min with an excitation wavelength of 320 nm and a detection wavelength of 420 nm.

The complete consumption of this peptide by LspPae was observed which, when contrasted to the incomplete hydrolysis of peptide **49**, suggests that LspPae is stereoselective, with the stereochemistry at the glyceryl-CH position impacting on enzyme activity. These results are in line with a simultaneous 2018 study by Kitamura and Wolan who reported a similar FRET assay for LspA from *E. coli* (LspEco) using a synthetic lipopeptide (Dabsyl-VTGC(Pam_2)AKD(EDANS)-$CONH_2$).[200] This report illustrates that LspEco selectively recognises the (*R*)-DAG diastereomer, with no substrate recognition for the (*S*)-DAG isomer. With an active (*R*)-DAG based probe in

[f] Assays shown in **Fig. 2.13** were performed by Dr. Samir Olatunji (Caffrey Group, Biochemistry, TCD).

hand, and with the publication of the concept elsewhere, the synthesis of (*S*)-DAG isomer was not performed in this case.

The effect of LspA inhibitors globomycin and myxovirescin on the progress of the peptidase reaction of peptide **55** is shown in **Fig. 2.14** and **Fig. 2.15** respectively.[g] Addition of a 5-fold molar excess of either compound inhibited the peptidase reaction within minutes. This rapid effect is readily visualised by the fluorescence output which remains unchanged following antibiotic addition, indicating complete enzyme inhibition. The fluorescence reading is stable for at least 4 h, a favourable property for HTS applications where practical constraints may delay the fluorescence measurement.

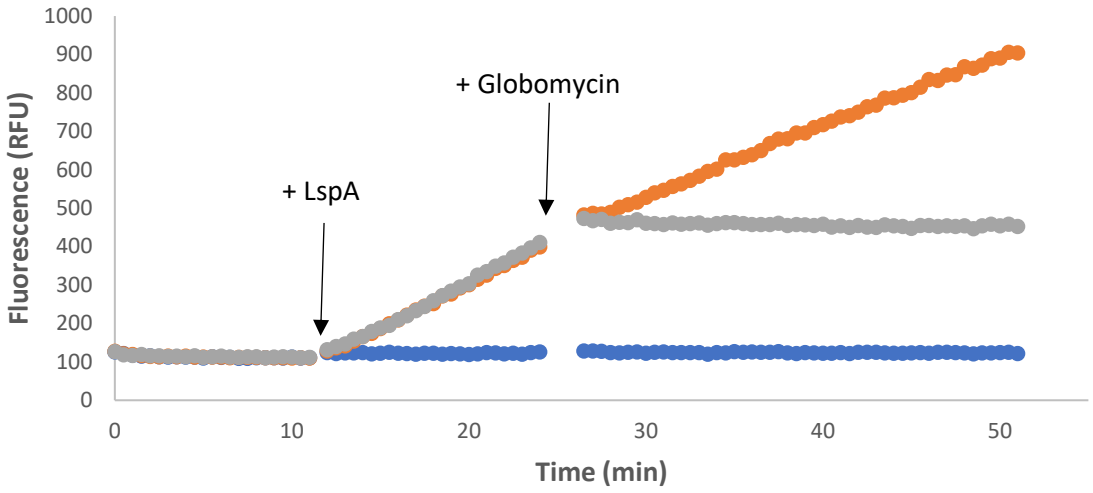


Figure 2.14.[g] The three progress curves shown here reveal the effect of LspPae on peptide **55** and the result of globomycin addition. **Blue data points**: LspPae-free control. **Orange data points**: addition of LspPae to the FRET substrate at t = 11 min initiates the peptidase reaction. **Grey data points:** globomycin addition to the reaction of peptide **55** and LspPae at t = 25 min inhibits the peptidase reaction. The gaps in each progress curve arise from the need to pause measurements to add enzyme and antibiotic to the fluorescence plate reader.

[g] Experiments shown in **Fig 2.14** and **Fig 2.15** were performed by Dr. Jonathan Bailey and Dr. Samir Olatunji (Caffrey Group, Biochemistry, TCD)

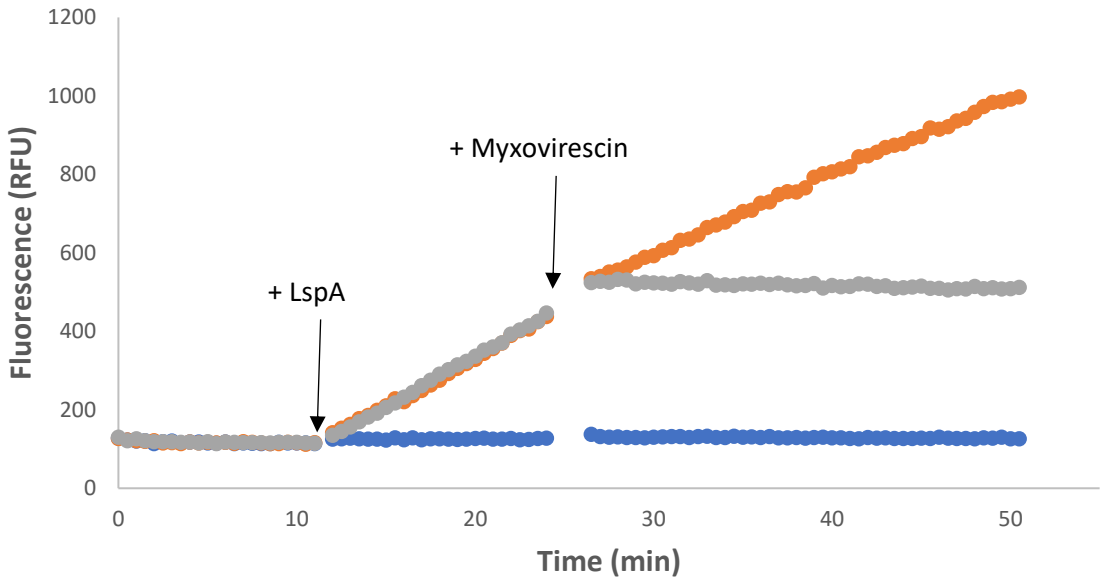


Figure 2.15.[g] The three progress curves shown here illustrate the action of LspPae on peptide **55** and the effect of myxovirescin addition. **Blue data points**: LspPae-free control. **Orange data points**: addition of LspPae to the FRET substrate at t = 11 min. **Grey data points**: Myxovirescin addition to the reaction of peptide **55** and LspPae at t = 25 min.

Using the previously developed standard fluorescence curve to convert relative fluorescence units (RFU) to moles of product, the rate of reaction and kinetic parameters V_{max} and K_M could be determined.[h] From the substrate saturation analysis shown in **Fig. 2.16**, V_{max} was determined to be 107 nmol/(mg.min) and K_M 10.2 µM at an enzyme concentration of 0.1 µM. K_M, the concentration of substrate at which half V_{max} is reached, is an essential factor to determine during assay design as the concentration of FRET substrate used in the assay will impact the type of inhibitors that are detected. At a FRET substrate concentration below the K_M, competitive inhibitors are more likely to be detected. Such inhibitors bind to the active site where the substrate also interacts. As FRET substrate concentration decreases, the active site in increasingly likely to be unoccupied by substrate, increasing the probability of inhibitor binding. At a FRET substrate concentration above the K_M, uncompetitive inhibitors, compounds which bind to the enzyme-substrate complex, are more likely to be detected. The concentration of enzyme-substrate complex increases at higher substrate concentrations, increasing the probability of an uncompetitive inhibitor binding. Non-competitive inhibitors bind at an allosteric site which is unaffected by substrate binding. As such, substrate concentration

[h] Experiments and analysis performed by Dr. Jonathan Bailey (Caffrey Group, Biochemistry, TCD)

does not affect the probability of detecting such inhibitors. Ideally, the substrate concentration used should be close to K_M value to identify all classes of inhibitors.[201]

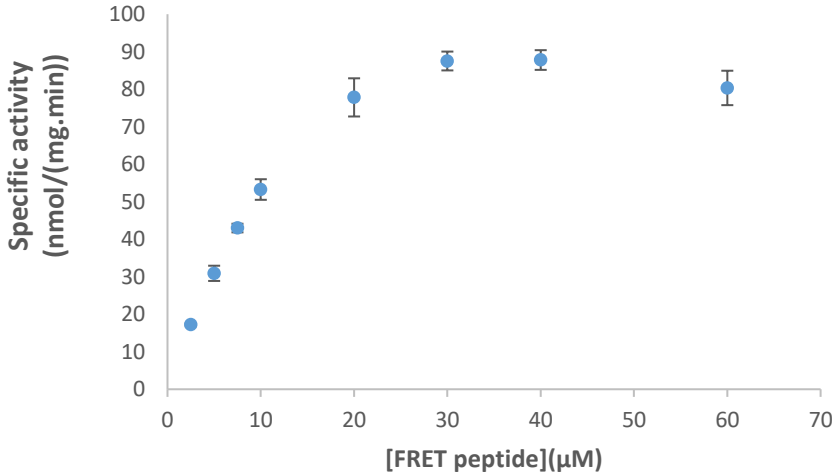


Figure 2.16.[h] Substrate saturation analysis of LspPae at an enzyme concentration of 0.1 µM. Data points represent the mean value ± standard deviation (n = 2).

With optimised FRET assay conditions in hand, and the suitability of the assay for HTS determined, this FRET assay was used by the European Lead Factory to screen 535,000 compounds against LspPae. ~5000 compounds displayed activity and are currently undergoing additional screening to identify the most potent inhibitors which will be selected for further development.

2.6 LspMrs FRET Assay

As previously discussed, LspA is essential for MRSA survival under physiological conditions.[77] Due to the potential of this enzyme as an antibiotic target in MRSA, developing an assay to identify novel inhibitors of this enzyme was also of interest. The sequence identity between LspPae and LspMrs is 31%, however, these enzymes are structurally similar with particularly homologous substrate binding sites.[77] Based on this, the FRET probe **55** designed for LspPae was tested with LspMrs. The peptide was determined to be a substrate of LspMrs, however, in comparison to LspPae, the specific activity of LspMrs with peptide **55** was much lower. At an enzyme concentration of 0.3 µM, V_{max} was determined to be 2.54 nmol/(mg.min) and K_M 46.7 µM based on a simple Michaelis–Menten model applied to the substrate saturation curve shown in **Fig. 2.17**.[i]

[i] Experiments and analysis performed by Dr. Jonathan Bailey (Caffrey Group, Biochemistry, TCD)

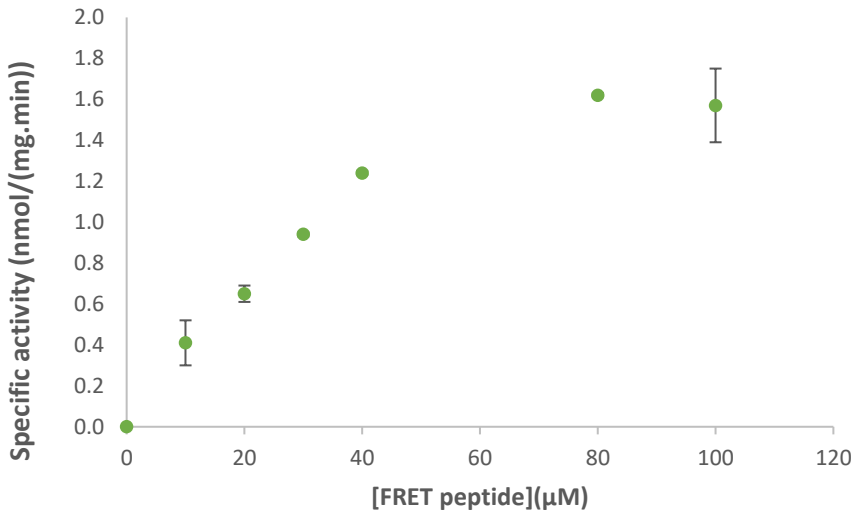
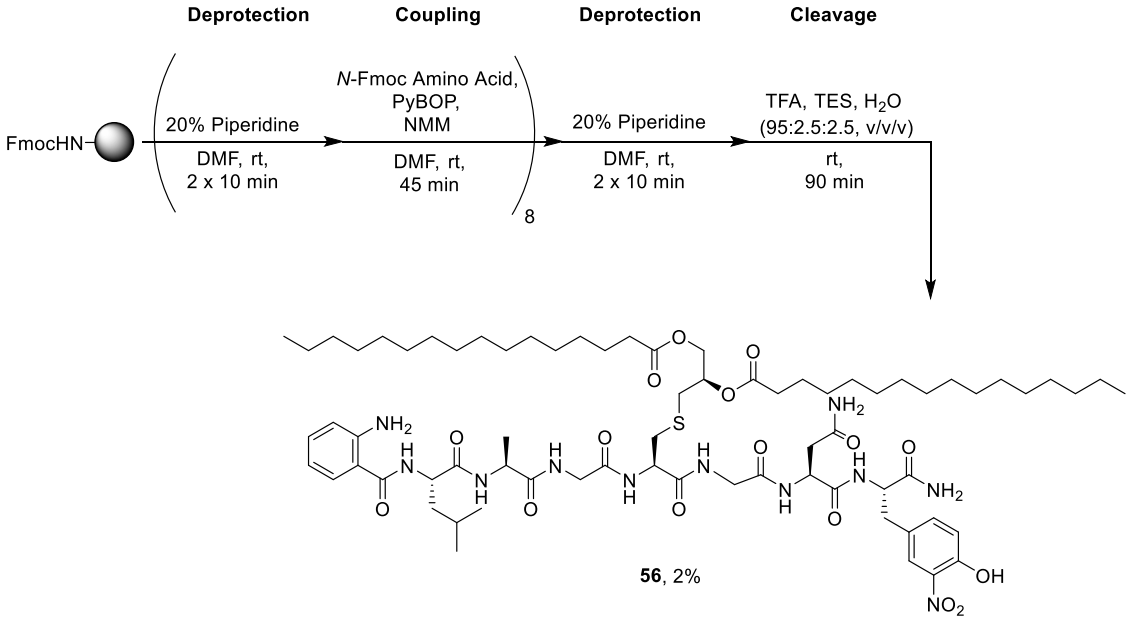


Figure 2.17.[i] Substrate saturation analysis of LspMrs at an enzyme concentration of 0.3 µM. Data points represent the mean value ± standard deviation (n = 2).

2.6.1 Altering the Peptide Sequence

As a result of the low specific activity of LspMrs with peptide **55**, a high enzyme concentration is required to generate a robust signal in the assay. At such concentrations, the assay has reduced sensitivity towards inhibitors, in particular to those such as globomycin which exhibit tight binding, and IC_{50} values below [E]/2 cannot be measured.[77, 202] To determine accurate kinetic parameters and to effectively screen inhibitors, optimisation of the FRET substrate was attempted. The peptide sequence of peptide **55** is based on the consensus sequence of native *P. aeruginosa* lipoproteins, not those of MRSA. Consequently, the sequence was changed to the MRSA consensus sequence of LAGCGN. Peptide **56**, shown in **Scheme 2.14**, was synthesised under standard conditions as previously described using SPPS at a scale of 0.35 mmol. The asparagine residue was introduced by PyBOP/NMM mediated coupling of Fmoc-Asn(Trt)-OH. The target peptide **56** was isolated in a yield of 2%. This low yield is attributable to the low crude purity resulting from the formation of multiple significant side products impeding chromatographic purification. A similar result was observed with peptide **55** at this scale. Nevertheless, 9.0 mg of peptide **56** was isolated for characterisation and testing with LspMrs. The structure was confirmed by NMR and MS. Unfortunately, peptide **56** was not a substrate of LspMrs and was not processed by the enzyme under the conditions previously optimised for LspMrs and peptide **55**. Hence, neither of the Abz/3-NT functionalised peptides **55** or **56** are suitable substrates for HTS

screening of LspMrs. As altering the peptide sequence resulted in a total loss of substrate recognition, future efforts to develop an improved FRET substrate for LspMrs will initially focus on screening alternative FRET pairs with a peptide which retains the core LAGCSS sequence.



Scheme 2.14. Synthesis of peptide **56** *via* SPPS yielded the product in an isolated yield of 2%.

2.7 Conclusions and Future Work

In conclusion, the synthesis of fluorescently labelled LspA substrates has been developed. The purification of these hydrophobic, water-insoluble peptides has been achieved with normal phase silica column chromatography, yielding peptides in high purity suitable for biochemical analysis. These substrates have facilitated the development of a high-throughput, continuous, FRET-based assay for LspPae which has been utilised in the screening of a major compound library and has expedited efforts to identify novel inhibitors. Additionally, the substrate specificities of both LspPae and LspMrs have been explored using various synthetic substrates. The stereoselectivity of LspA has been demonstrated, along with the requirement of DAG functionalisation for LspA recognition.

To develop a more efficient FRET assay for LspMrs which is suitable for HTS it necessary to generate a reliable signal at low enzyme concentrations. This requires either improved substrate recognition by LspMrs or increased sensitivity arising from a substrate which generates a greater change in fluorescence intensity following peptide

hydrolysis. Changing the FRET pair has the potential to improve substrate recognition while also allowing for the use of a fluorophore which exhibited higher fluorescence intensity than Abz, arising from a higher fluorescence quantum yield (ϕ_F). To meet these criteria, the 5-(2-aminoethyl)aminonaphthalene sulfonic acid (EDANS) / 4-[[4′-(*N*,*N*-dimethylamino)phenyl]diazenyl]benzoic acid (Dabsyl) FRET pair is currently under investigation. Previously reported measurements of the fluorescence quantum yields enable the comparison of both fluorophores, with EDANS ϕ_F = 0.69[203] while Abz ϕ_F = 0.60[204] in EtOH. The efficient donor-quencher energy transfer which occurs with this widely used FRET pair is a result of the spectral overlap of EDANS (donor) emission (λ_{max} = 490 nm) and Dabsyl (quencher) absorbance (λ_{max} = 472 nm) as summarised in **Fig. 2.18**.[205]

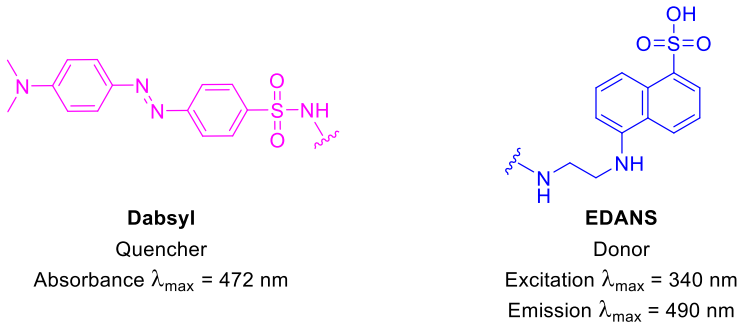


Figure 2.18. EDANS/Dabsyl spectral properties.

Promisingly, the EDANS/Dabsyl FRET pair has been previously utilised in a LspEco FRET assay described by Wolan and co-workers, indicating substrate recognition of an EDANS/Dabsyl containing peptide by the homologous LspMrs may be feasible.[93] However, inner filter effects (IFEs) have previously been reported in assays utilising this FRET pair. Primary IFEs arise from attenuation of incoming light intensity as it travels through the optically dense sample due to absorption by the high concentration of substrate. Secondary IFEs arise from reabsorption of the light emitted by EDANS.[206] Though IFEs cannot be avoided, they can be minimised by using low substrate concentrations.[207, 208] Although not reported in the LspEco EDANS/Dabsyl assay developed by Wolan and co-workers which uses a FRET substrate concentration 20 μM,[93] such an effect has been described in an EDANS/Dabcyl peptide FRET assay for the main protease (M^{pro}) of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In this case, much higher substrate concentrations of 5-640 μM were ultilised.[209] Consequently, a successful EDANS/Dabsyl peptide probe for LspMrs must be an

excellent substrate for the enzyme to permit the use of a low substrate concentration. Based on the poor substrate recognition of Abz/3-NT peptide **55** by LspMrs, changing only the FRET pair may not accomplish this. Therefore, further optimisation of a peptide probe for LspMrs will likely require extensive investigation into the substrate scope of the enzyme. This would require alterations at many sites of the peptide and include the exploration of further FRET pairs, different peptide sequences and various lipid lengths and degrees of saturation. The linear SPPS strategy utilised in this chapter and the purification of such hydrophobic peptides is excessively time consuming and low yielding, representing a major bottleneck in the screening effort. As such, accessing a large number of peptide variants through a linear SPPS approach is not realistic. To comprehensively explore the substrate scope of LspMrs, more efficient ways of synthesising and purifying these substrates is required. To this aim, a convergent synthetic methodology permitting late-stage lipid functionalisation of cysteine-containing peptides was developed. This work is further detailed in Chapter 5.

Chapter 3

Development of Synthetic Substrates of Apolipoprotein *N*-Acyltransferase

3.1 Introduction

Apolipoprotein *N*-acyltransferase (Lnt) catalyses the acylation of the N-terminal amine of diacylated LPs to yield triacylated, mature LPs (**Fig. 3.1**). Triacylation is critical for the correct localisation of these proteins within the bacterial cell, enabling LPs to perform critical functions ranging from maintaining cell envelope integrity to the modulation of the host immune response, as discussed in **Section 1.2.1**.

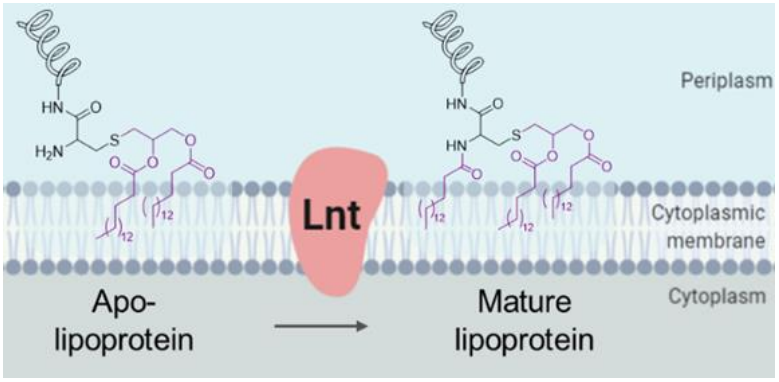


Figure 3.1. In Gram-negative bacteria, Lnt introduces a third acyl group onto the N-terminal cysteine of diacylated LPs (apolipoproteins) to yield mature *N*-acylated LPs.

The crystal structures of Lnt from *P. aeruginosa* (LntPae) and *E. coli* (LntEco) have recently been reported.[181, 210, 211] These structures provide structural insights into both the putative active site and mechanism of catalysis *via* a Glu-Lys-Cys catalytic triad, consistent with the nitrilase family of enzymes. Nitrilase enzymes are responsible for nitrile and amide bond hydrolysis and, as in this case, condensation.[212] Lnt catalyses amide bond formation through a proposed two-step mechanism.[181] The first step involves the formation of a thioester intermediate *via* acylation of the active site cysteine sulfhydryl. The second step involves *S*-to-*N* acyl transfer of the acyl chain from the acylated-Lnt intermediate to the diacylated LP substrate. These steps are illustrated schematically in **Fig. 3.2**. As shown, the Lnt reaction requires two substrates. The first is a phospholipid, usually phosphatidylethanolamine (PE), which acts as an acyl-donor in the initial O-to-S acyl transfer reaction in which the *sn*-1 acyl chain of PE is transferred to the enzyme. The second substrate is a diacylated LP containing an amino group at the N-terminal dagylated cysteine. This amino group is acylated by the enzyme in the second step, yielding the desired triacylated LP product.

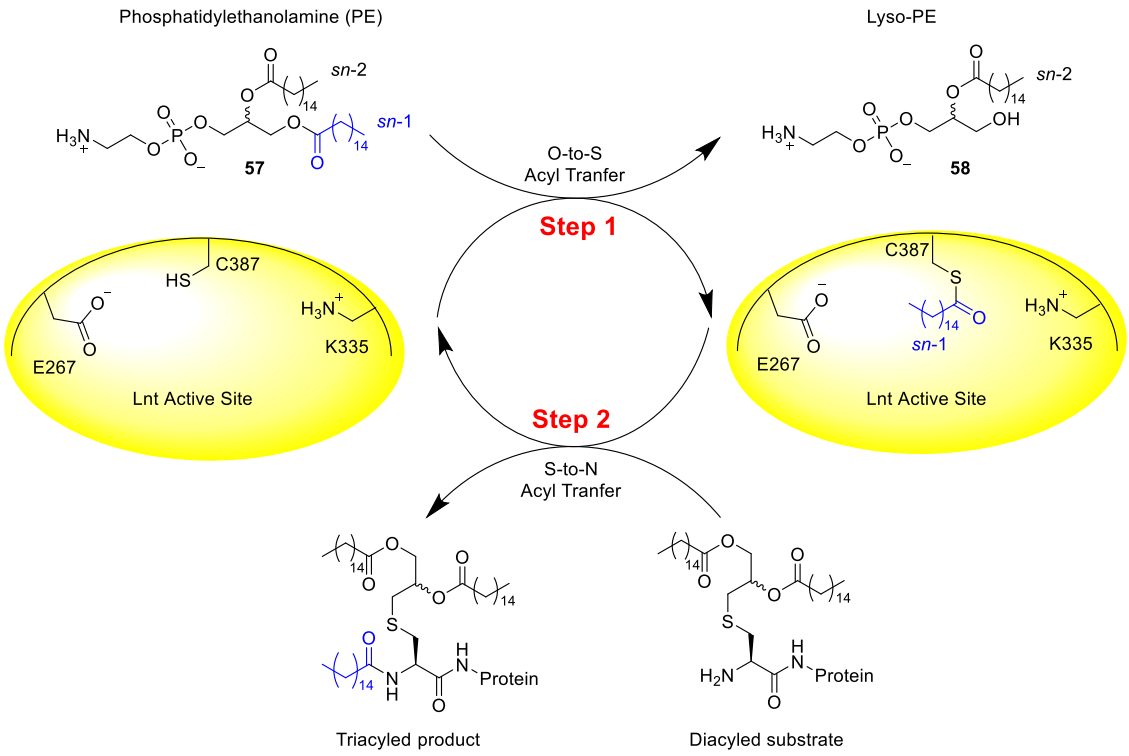


Figure 3.2. The proposed Lnt reaction consists of two steps and proceeds through a thioester intermediate. Firstly, the *sn*-1 acyl chain of PE (**57**) is transferred to the active site cysteine of Lnt. Next, the newly formed thioester, which links the acyl chain to Lnt, undergoes nucleophilic attack by the N-terminal amine of a diacylated LP. This regenerates the enzyme and produces a triacylated LP.

Lnt recognises the same substrate (diacylated LPs) as the TLR2/6 heterodimer present on mammalian innate immune cells. A complete understanding of the Lnt apolipoprotein binding site will assist in the design of specific inhibitors of Lnt which can differentiate between the TLR2/6 binding pocket and the Lnt binding pocket. As the crystal structure of the TLR2/6 heterodimer bound to a diacylated lipopeptide has been previously reported,[213] capturing the step of the Lnt reaction which involves *S*-acylated Lnt and a diacylated substrate *via* crystallography would allow for comparison and identification of important interactions between the enzyme and substrate. This first requires scalable access to diacylated substrates of Lnt which act as accurate biomimetics of diacylated LPs.

3.1.1 An Existing Assay for Lnt

An *in vitro* activity assay for LntEco and LntPae which monitors the first step of the Lnt reaction has previously been reported by Caffrey and co-workers.[181] This thin-layer chromatography (TLC) based assay utilises a commercial fluorescent phospholipid substrate, 1-oleoyl-2-12-[(7-nitro-2-1,3-benzoxadiazol-4-yl)amino]dodecanoyl-*sn*-glycero-3-phosphoethanolamine (NBD-PE, **59**), containing a fluorophore labelled *sn*-2 acyl chain. In the presence of a diacylated LP substrate, Lnt activity is indicated by the consumption of NBD-PE and production of 1-oleoyl-2-hydroxy-*sn*-glycero-3-phosphoethanolamine-*N*-(7-nitro-2-1,3-benzoxadiazol-4-yl) (lyso-NBD-PE, **60**). The difference in polarity between NBD-PE and lyso-NBD-PE enables separation and visualisation by TLC. Following EtOH/$CHCl_3$ extraction of the enzymatic reaction, a TLC of the LntEco reaction with NBD-PE and a commercial, biotin labelled diacylated lipopeptide, fibroblast stimulating ligand (FSL)-1 biotin ($Cys(Pam)_2$GDPKHPKSF-biotin) is performed with the output from a time course of such reactions shown in **Fig. 3.3**. This assay provides indirect evidence of the activity of Lnt and has enabled the activity of both wild-type and mutant Lnt to be screened to identify essential residues. A minor amount of lyso-NBD-PE is produced in the absence of the LP substrate (**Fig. 3.3**). Although the enzyme is used in catalytic quantities, the formation of the Lnt cysteinyl thioester from NBD-PE (step 1) will occur regardless of LP presence. This would produce a minor and likely undetectable quantity of lyso-NBD-PE. However, thioesters are relatively labile and can be hydrolysed under aq. conditions. The background quantity lyso-NBD-PE observed in the absence of LP likely arises from successive cycles of formation and subsequent hydrolysis of this thioester. As a result, the signal to background ratio of the assay is limited. Furthermore, it is not optimal for inhibitor screening as it is not continuous nor high-throughput, and it does not provide direct evidence of Lnt activity.

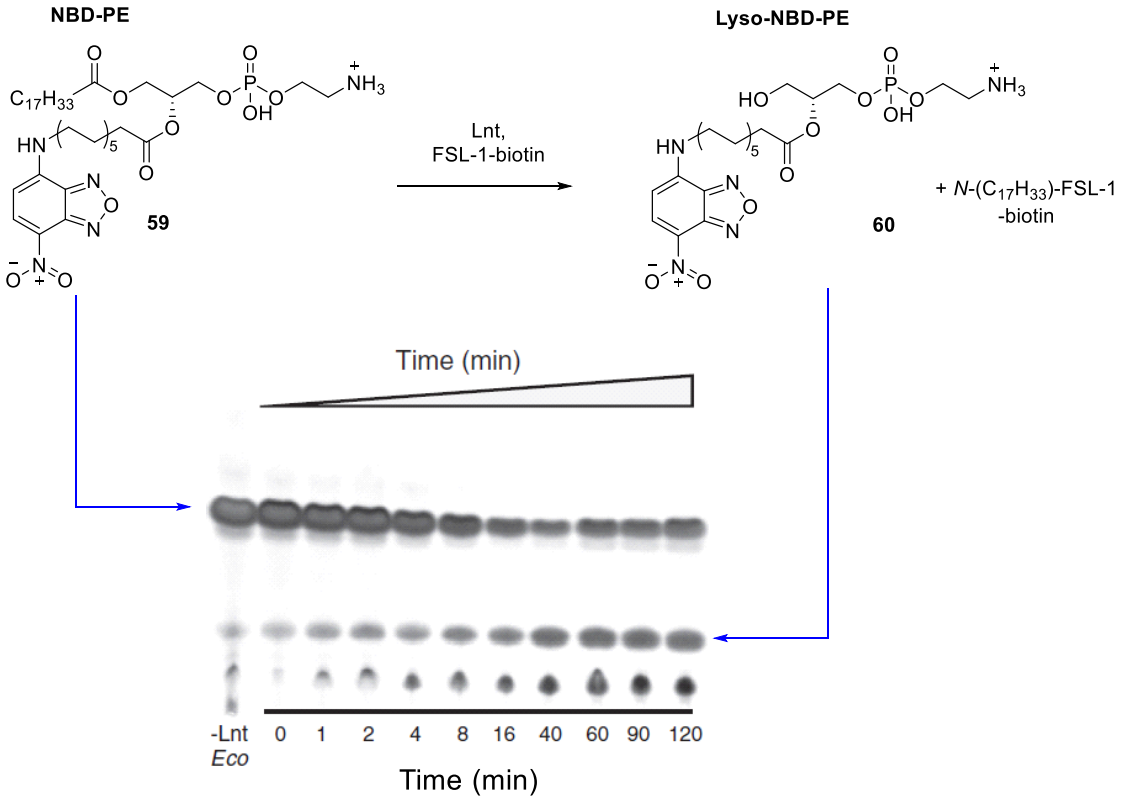


Figure 3.3. A 2 h TLC time course of the Lnt reaction between NBD-PE (**59**) and FSL-1-biotin. A minor quantity of lyso-NBD-PE (**60**) can be observed in the absence of the enzyme (lane 1).[181]

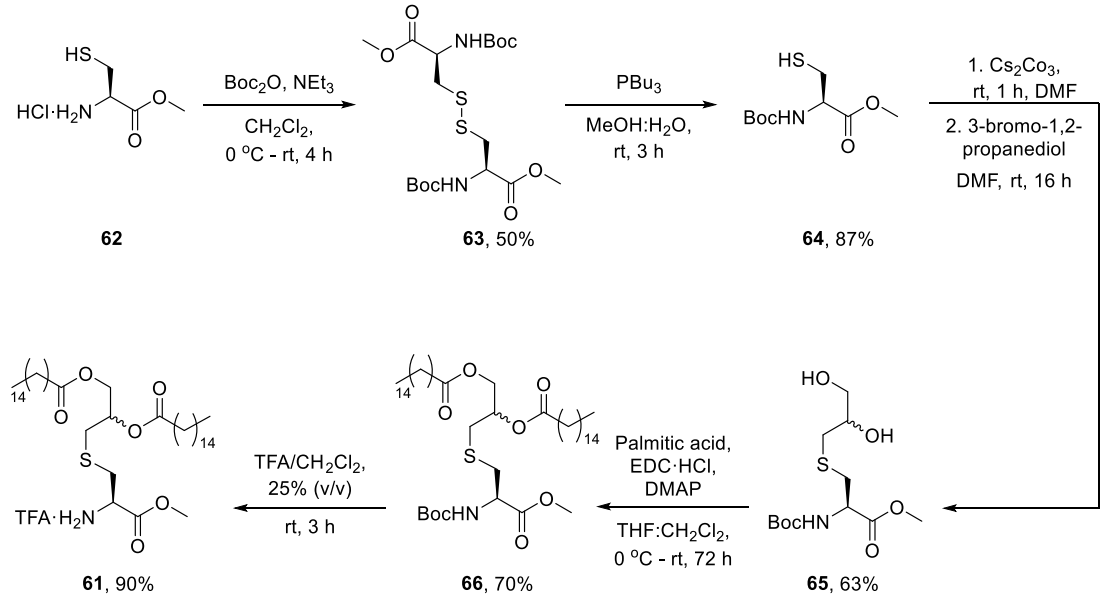
3.2 Aims of this work

In collaboration with the group of Prof. Martin Caffrey, this study sought to investigate the substrate specificity of Lnt and develop a FRET-based assay for Lnt suitable for HTS to ultimately identify specific inhibitors of the enzyme. Although commercial lipopeptides, such as FSL-1, have been identified as Lnt substrates,[181] the prohibitive cost of such compounds (~€2500 per mg) renders them unsuitable for application in HTS and *in meso* protein co-crystallisation which both require substantial quantities of material. Additionally, these unprotected peptides are not amenable to further in-house functionalisation, limiting their application. As such, this project aimed to develop a facile synthesis of Lnt substrates to facilitate further investigation into the function of this enzyme.

3.3 Lnt Substrate Scope

3.3.1 Synthesis of Lnt Substrate H-Cys[(*R*/*S*)Pam]$_2$-OMe (61)

The initial objective of this project was to establish the minimal structural requirements for Lnt recognition. As such, the small molecule H-Cys[(*R*/*S*)Pam]$_2$-OMe (**61**) was synthesised as shown in **Scheme 3.1** and screened for activity with Lnt.



Scheme 3.1. Synthesis of Lnt substrate H-Cys[(*R*/*S*)Pam]$_2$-OMe (**61**).

Boc protection of commercially sourced H-Cys-OMe·HCl (**62**) was performed using Boc_2O and Et_3N under aerobic conditions to yield compound **63** as a disulfide. Reduction of disulfide **63** using PBu_3 afforded thiol **64** in a yield of 87%. The S_N2 reaction of thiol **64** with 3-bromo-1,2-propanediol and Cs_2CO_3 was successful in generating diol **65** in 16 h in an acceptable yield of 63%. Interestingly, under identical conditions this reaction did not proceed with diol **35**, the *tert*-butyl protected analogue of diol **65** as discussed in **Section 2.3.2**. Since both compounds were readily soluble in DMF it is feasible that the dramatic increase in reaction rate observed in this case is a result of reduced steric bulk of the C-terminal protecting group, despite the reaction occurring at a distal reaction site. A similar trend of improved reaction rate was observed in the esterification of diol **65** when compared to the esterification of diol **35**. Diacylation of diol **65** to yield compound **66** under Steglich conditions with palmitic acid, EDC·HCl and DMAP reached completion after 3 days *via* the formation and subsequent consumption of the *sn*-1 monoacylated intermediate. Standard TFA mediated Boc deprotection of compound **66** afforded the desired amine **61** as a TFA salt.

3.3.2 Functional Analysis of Lnt with Compounds 61 and 66[j]

The processing of amine **61** and the Boc protected analogue **66** by LntPae was investigated using the reported NBD-PE TLC-based assay. Improved substrate recognition by the enzyme increases the rate of the enzymatic reaction and thereby increasing the amount of lyso-NBD-PE produced over a fixed time. As such, this assay indicates enzyme turnover number, enabling direct comparison of substrates. The functional analysis of various substrates of Lnt is shown in **Fig. 3.4**. Lanes 9 and 10 show the background level of lyso-NBD-PE which is produced when no substrate is present. Lanes 11 and 12 demonstrate that the background production of lyso-NBD-PE is eliminated in LntPae C387A, an inactive Lnt mutant in which the active site cysteine has been mutated to alanine. This confirms the background level of lyso-NBD-PE produced in the absence of a LP substrate arises from repeated Lnt *S*-acylation and hydrolysis as discussed above. Lanes 1, 2, 7 and 8 show the enzymatic reaction mixture after incubation with commercial substrates FSL-1 and Pam_2Cys ($Cys(Pam)_2SK_4$). This confirms that the presence of a substrate for step 2 of the Lnt reaction enhances the amount of NBD-lyso-PE produced and that Lnt exhibits higher activity with FSL-1 than with Pam_2Cys. Lanes 5 and 6 show Lnt mediated NBD-lyso-PE production in the presence of amine **61**. This shows that Lnt generates comparable quantities of NBD-lyso-PE with both amine **61** and FSL-1, revealing that this small molecule amine **61** is a substrate of the enzyme.

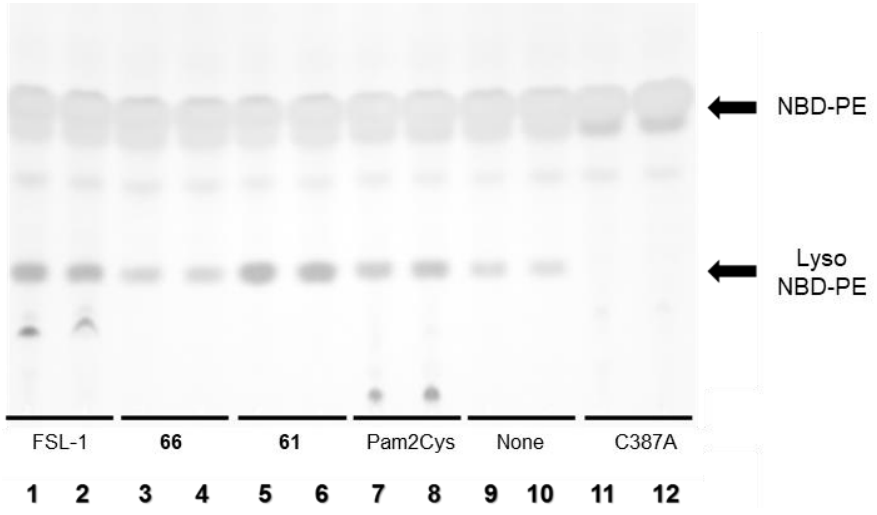
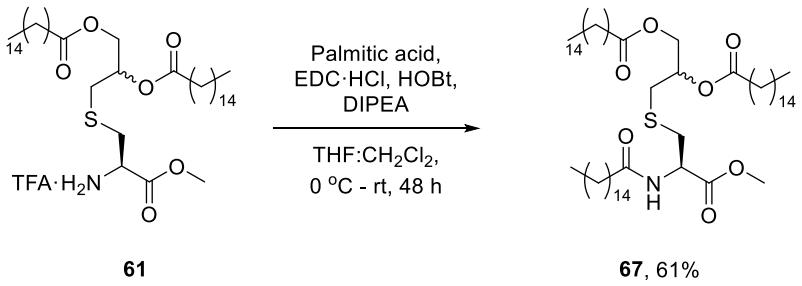


Figure 3.4.[j] Functional analysis of Lnt substrates using the NBD-PE TLC assay.

[j] Lnt assays shown in the section were developed and performed by Dr. Dietmar Weichert (Caffrey Group, Biochemistry, TCD).

Lnt is proposed to introduce a third acyl group onto the free N-terminal amino group of LPs. To confirm this, a substrate analogue in which this amino group was Boc protected, compound **66**, was tested for activity. Gratifyingly, the resultant lyso-NBD-PE production (lanes 3 and 4) is comparable to that of the control (lanes 9 and 10), indicating that the Boc protected compound is not a substrate. To further establish that Lnt processing of amine **61** involves *N*-acylation, the expected product of the Lnt reaction with amine **61** was synthesised. As shown in **Scheme 3.2**, triacylated **67** was synthesised through carbodiimide-mediated coupling of palmitic acid in a yield of 61% following purification by column chromatography.



Scheme 3.2. Synthetic *N*-acylation of amine **61** to yield triacylated **67**.

The R_f of this compound was compared with that of the products of the Lnt reaction, as depicted in **Fig. 3.5**. As shown in lanes 1 and 2, the reaction produces a compound of equal R_f to the synthetic triacylated product **67** (lane 5). A control reaction, shown in lanes 3 and 4, confirms that the quantity of this compound produced is significantly reduced in the absence of amine **61**. This result further suggests that Lnt processes **61** *via* palmitoylation of the N-terminal amine.

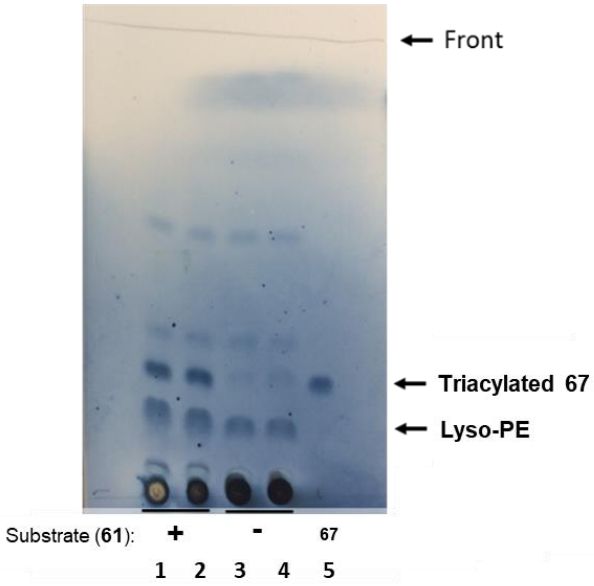


Figure 3.5.[j] TLC analysis of the enzymatic reaction mixture (lanes 1-4) following incubation of the LntPae with substrate **61** (lanes 1-2) and without substrate (lanes 3-4) and comparison to the R_f of synthetically produced Pam-Cys[(*R*/*S*)Pam]$_2$-OMe (**67**), lane 5. Plate stained with ammonium molybdate (0.26 M) in aq. H_2SO_4 (1 M).

The R_f comparison shown in **Fig. 3.5**, coupled with the previously discussed functional analysis indicates that Lnt recognition and *N*-acylation of amine **61** is comparable to Lnt processing of a native LP substrate. Unlike previously reported substrates, substrate **61** is readily accessible in gram-scale quantities, facilitating the addition of this substrate to the enzyme in high concentration during co-crystallisation trails.

3.4 A FRET Assay for Lnt

Lnt catalyses the formation of an amide bond. As this reaction involves the condensation of two individual fragments, the activity of the enzyme has the potential to be monitored using a bimolecular FRET system. The corresponding peptide probe would involve a fluorophore (donor) labelled diacylated lipopeptide substrate. The fluorescence of this peptide could be quenched *via* FRET following Lnt-mediated *N*-acylation with a lipid containing a quencher. This quencher would originate from a phospholipid labelled with a quencher in the *sn*-1 position. In step 1 of the Lnt reaction, the quencher-functionalised acyl chain of the phospholipid would be transferred to C387 of Lnt. During step 2 of the Lnt reaction this lipid would then be transferred to the N-terminus of a fluorescent lipopeptide substrate, quenching the fluorescence of the peptide by FRET. These steps are depicted in **Fig. 3.6**.

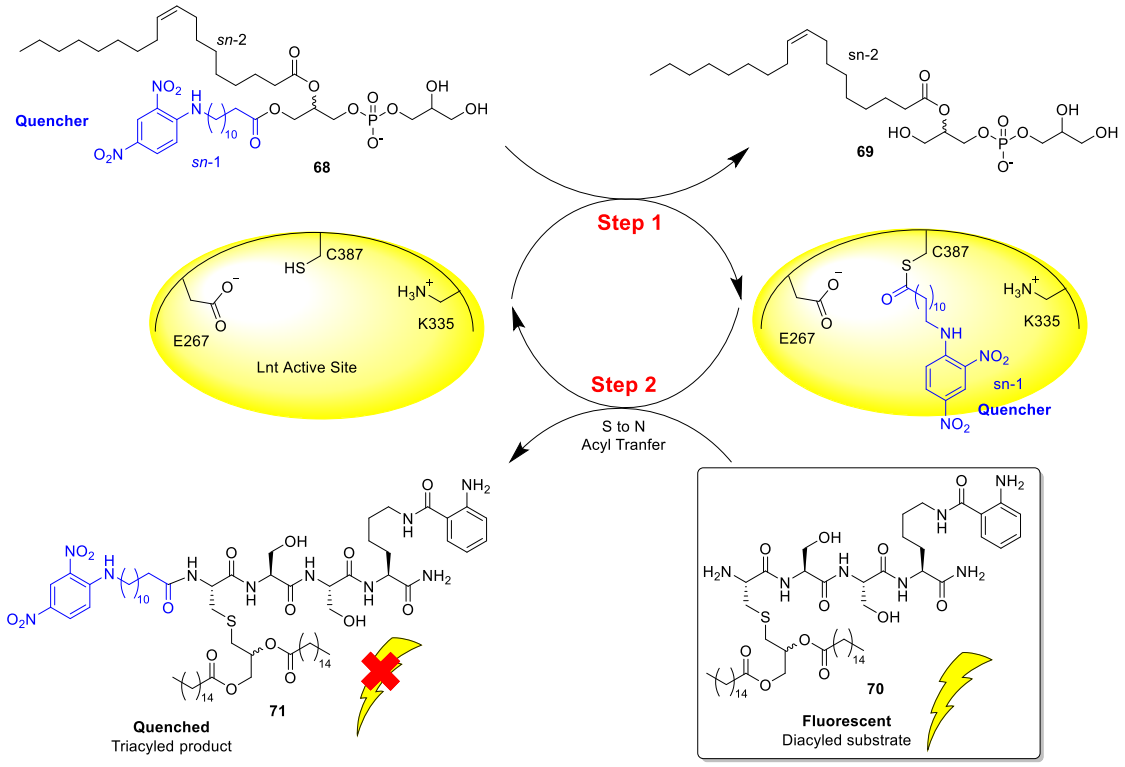


Figure 3.6. A proposed bimolecular FRET assay for LntPae.

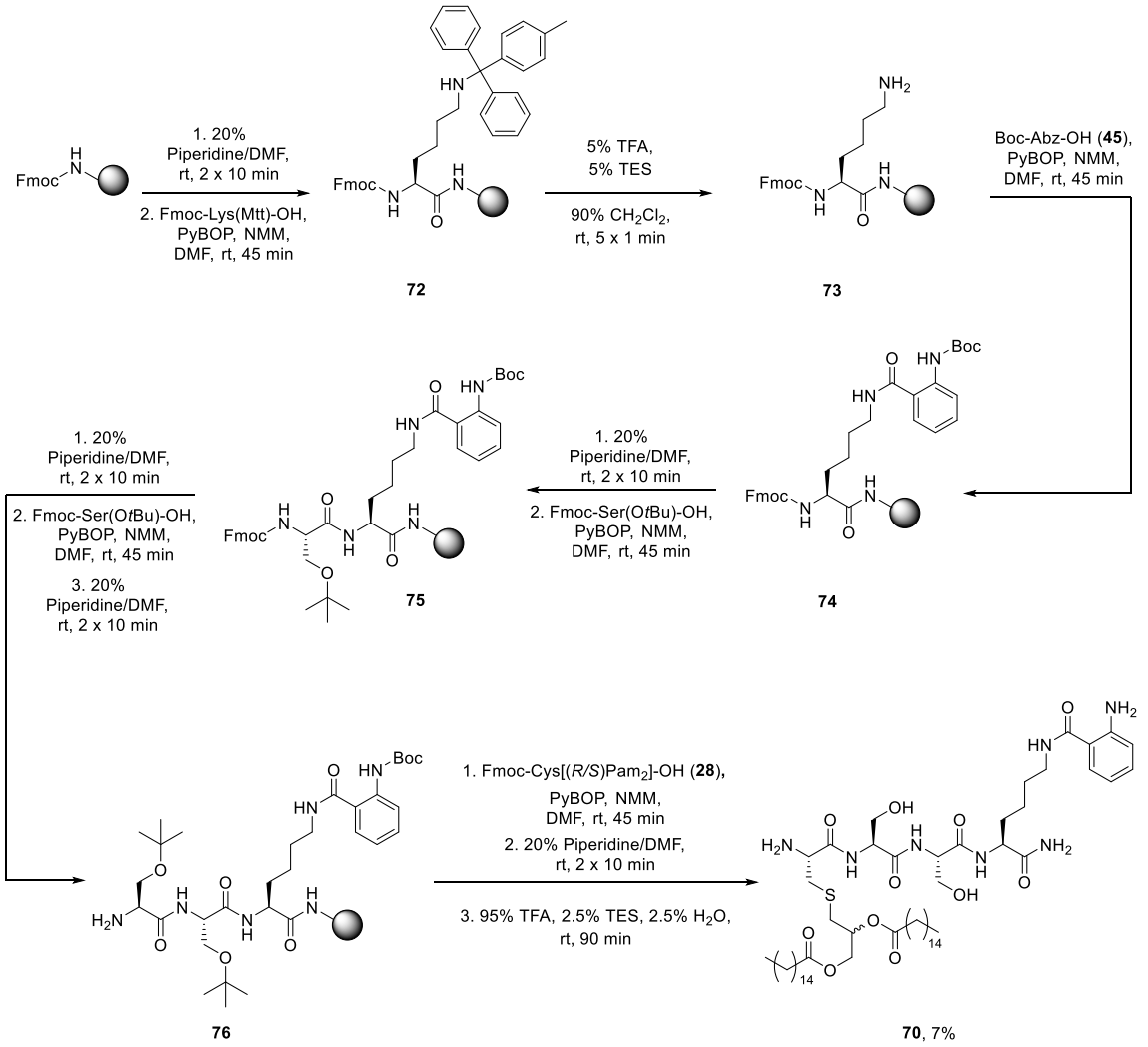
To realise this objective, a fluorophore-labelled Lnt substrate **70** was designed and synthesised. Abz was selected as the fluorescent donor as this amino acid can be readily introduced by SPPS. Additionally, due to the small size of this fluorophore, when located at a distal region to the essential DAG recognition element it is less likely to impede substrate recognition than a bulkier fluorophore. Abz was therefore incorporated onto a C-terminal lysine ε-NH_2 in the peptide substrate.

Quencher design incurs more of a challenge as an unnatural lipid must be transferred to the enzyme. 2,4-Dinitrophenyl (2,4-DNP), a small hydrophobic quencher, was selected due to its appropriate spectral overlap with Abz (2,4-DNP λ_{ex} = 400 nm).[214] The quencher **68**, a 2,4-DNP *sn*-1 labelled derivative of phosphatidylglycerol was synthesised by Gerard Reid (Caffrey lab, Biochemistry, TCD).

3.4.1 Peptide Synthesis

Peptide **70** was synthesised on Rink amide resin as shown in **Scheme 3.3**. Fmoc-amino acid coupling was performed *via* PyBOP/NMM activation. Fmoc deprotection was effected upon treatment with 20% (v/v) piperidine in DMF. An on-resin strategy was envisioned for the introduction of Boc-Abz-OH (**45**) to the lysine ε-NH_2 side chain. To achieve this, 4-methyltrityl (Mtt) was used as a lysine ε-NH_2 protecting group, introduced

through the use of commercial Fmoc-Lys(Mtt)-OH. Selective Mtt deprotection by 5% (v/v) TFA in the presence of the Rink amide linker has been previously reported by Weller *et al*.[215] Following Fmoc-resin deprotection, Fmoc-Lys(Mtt)-OH was coupled to the resin under standard PyBOP/NMM conditions. Mtt deprotection was performed with 5% (v/v) TFA, 5% (v/v) TES in CH_2Cl_2 for 5 x 1 min. Advantageously, the deprotection of Mtt by 5% TFA proceeds rapidly, with a transient colourless to bright yellow colour change observed in solution immediately after TFA addition. This colour arises from the Mtt cation which is subsequently quenched *via* hydride transfer from TES. Complete quenching is indicated after 5-10 s by a solution colour change from yellow to colourless. During iterative TFA additions to deprotect Mtt, the yellow colour change indicates Mtt is still present and the reaction is not yet complete. The deprotection reaction was repeated for 5 x 1 min, through the addition and subsequent draining of 5 portions of TFA/TES/CH_2Cl_2 after which no colour change occurred following further TFA addition, indicating the completion of the reaction. This colour change enabled the extent of the reaction to be monitored qualitatively in real time. This is particularly advantageous for an on-resin strategy as typically analysis to would require cleavage of a small portion of the resin followed by NMR and MS analysis. Following Mtt deprotection, coupling of Boc-Abz-OH (**45**) was then performed using PyBOP/NMM activation. Lysine Fmoc deprotection was followed by two iterative cycles of Fmoc-Ser(O*t*Bu)-OH coupling and Fmoc deprotection. Fmoc-Cys[(*R*/*S*)Pam_2]-OH (**28**) was synthesised as previously described (**Section 2.3.2**) and coupled with PyBOP/NMM. Finally, the cysteine Fmoc group was deprotected. The peptide was cleaved from the resin using 95% (v/v) TFA, 2.5% (v/v) TES, 2.5% (v/v) H_2O, concentrated, and precipitated in Et_2O. The crude peptide, which was soluble in MeOH/CH_2Cl_2 (1:1 v/v), was purified by normal phase silica column chromatography with the product isolated in an overall yield of 7%. The structure and purity of the peptide were confirmed by MS, TLC, 1D and 2D NMR.



Scheme 3.3. The synthesis of peptide **70** was achieved by the on-resin introduction of Boc-Abz onto the lysine side chain and sequential SPPS under standard conditions to yield the target peptide.

3.4.2 Functional Analysis of Lnt with Peptide 70[k]

Initially, peptide **70** was investigated as a potential substrate of Lnt by monitoring the reaction of the peptide with Lnt and a known phospholipid substrate (PE). The progress of the reaction, following EtOH/$CHCl_3$ extraction, was tracked by TLC analysis as shown in **Fig. 3.7**. After 30 min of incubation of peptide **70** with Lnt and PE, the formation of a new fluorescent compound was observed by TLC. After 3 h TLC analysis revealed complete consumption of peptide **70**. The reaction did not occur in the absence of Lnt, confirming that product formation is dependent on Lnt activity.

[k] Lnt assays shown in the section were developed and performed by Dr. Dietmar Weichert (Caffrey Group, Biochemistry, TCD).

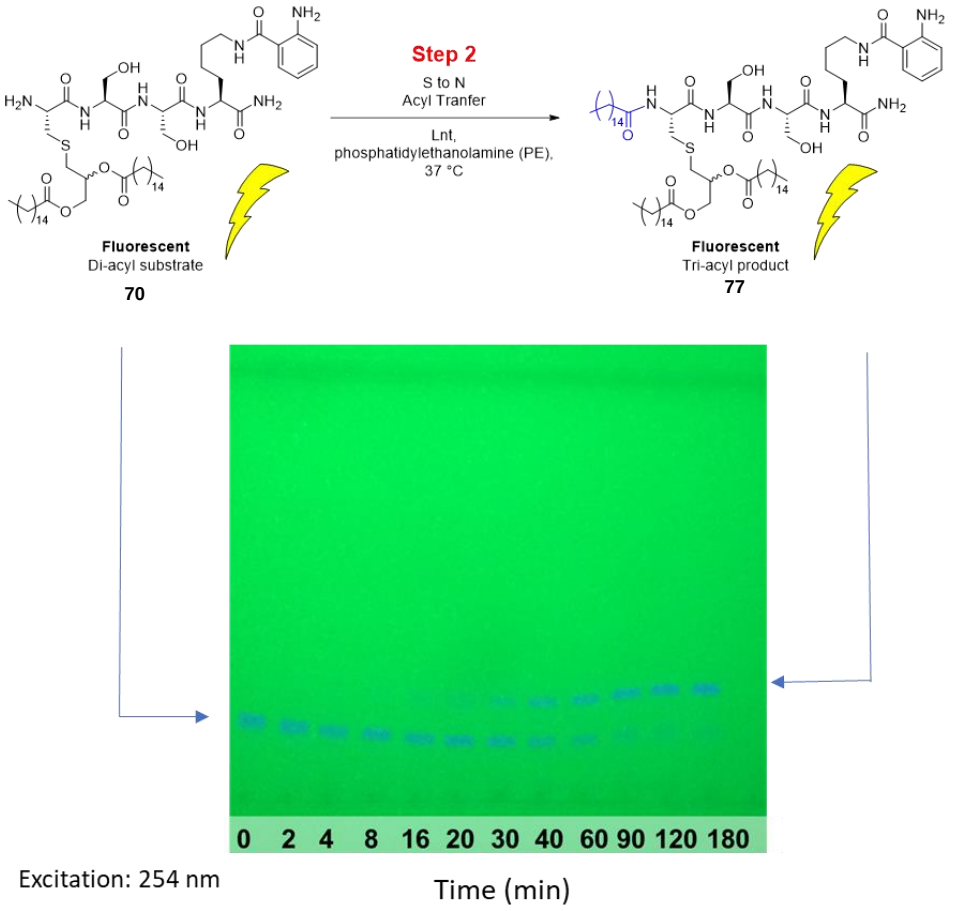


Figure 3.7.[k] TLC analysis of the reaction peptide **70** with Lnt in the presence of PE over 3 h. TLC developed in MeOH:$CHCl_3$:NH_4OH 19:80:1 v/v/v and visualised at 245 nm.

The time-course shown in **Fig. 3.7** demonstrated that peptide **70** is a substrate of LntPae. The peptide, which is a mixture of two diastereomers due to the unresolved glyceryl stereochemistry, is completely consumed in the reaction showing that Lnt is not selective for a single diastereomer of peptide **70** and is capable of processing both the (*R*) and (*S*) DAG isomers.

Next, to explore FRET assay development with Lnt, the 1,2-DNP-labelled phospholipid **68** was screened as a Lnt substrate. However, the reaction of peptide **70** with Lnt did not proceed when carried out using the modified phospholipid **68** in place of PE. This result indicates that the *sn*-1 acyl chain of phospholipid **68** is not transferred to Lnt as this modified phospholipid is not a substrate of the enzyme. This highlights the sensitivity of Lnt activity to changes in the *sn*-1 acyl chain group.

3.5 Conclusions and Future Work

In this study, two novel substrates of LntPae have been identified. An efficient synthetic route towards the dagylated cysteine monomer **61** has been established and this small molecule has been demonstrated to be sufficient for substrate recognition Lnt. This has highlighted the significant role of the glyceryl-ester linked groups in docking of the substrate correctly in the active site. Based on this preliminary work, a short peptide substrate of Lnt was designed. Conditions for the on-resin functionalisation of the lysine side chain have been optimised and the synthesis of a fluorophore-labelled peptide was developed. Synthetic access to this peptide substrate of Lnt prompted investigation into the stereoselectivity of the enzyme, revealing that Lnt can process substrates which feature both (*R*) and (*S*) glyceryl stereochemistry.

Attempts to develop a Lnt FRET assay has thus far been unsuccessful and will require further screening of quencher labelled phospholipid derivatives for Lnt activity. Due to the synthetic challenge of selectively modifying the *sn*-1 acyl chain of a phospholipid, *in silico* modelling of potential substrates with Lnt could provide valuable guidance in substrate design prior to further synthetic efforts. Despite the limited success in FRET assay development, peptide **70** has been successfully repurposed for use in a novel TLC-based assay for Lnt. As shown in **Fig. 3.4**, the Lnt reaction contains multiple organic soluble compounds including detergents and phospholipid substrate and products. When using a non-fluorescent substrate, such as in **Fig. 3.4**, monitoring of the reaction by TLC requires staining of the TLC plate. The challenge of differentiating the substrate and product from other organic components of this reaction impedes reaction tracking. However, the fluorescent substrate **70** has enabled facile tracking of the reaction as this substrate and corresponding product can be selectively visualised under short wave UV illumination (254 nm). This short peptide yields a definitive change in polarity following Lnt mediated *N*-acylation which provides an unambiguous indication of Lnt activity. This represents an advance on previously reported Lnt assays as no background signal is present and the second step of the reaction can be monitored directly, ultimately furthering the understanding of Lnt and providing a means to screen potential enzyme inhibitors for antibiotic discovery.

Chapter 4

Development of Synthetic Substrates of Lipoprotein Intramolecular Transacylase

4.1 Introduction

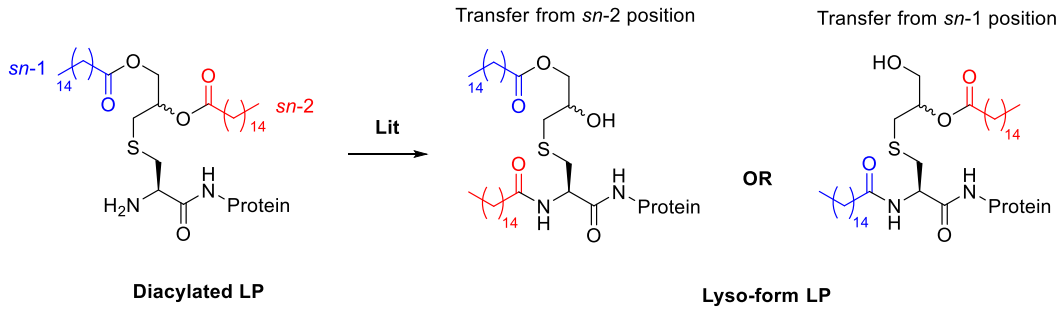
4.1.1 Discovery of Lit

Gram-positive bacteria, unlike their Gram-negative counterparts, express neither Lnt nor a *lnt* orthologue. Resultantly, the majority of Gram-positive LPs are diacylated, not triacylated. However, in 2012, Kurokawa *et al*.[40] identified a novel class of Gram-positive LP in *Enterococcus faecalis*, *Bacillus cereus*, *Lactobacillus bulgaricus*, and *Streptococcus sanguinis*. These low GC Gram-positive firmicutes reside in the human gastrointestinal tract and many are pathogens of clinical importance. The novel LP structure features an N-terminal *N*-acyl-*S*-monoacylglyceryl cysteine residue. Termed lyso LPs, these newly discovered biomolecules are implicated in modulation of the host immune response as weaker agonists of TLR2 than their diacylated counterparts. The enzyme responsible for this lyso modification was identified in 2017 by Armbruster and Meredith in both *Enterococcus faecalis* and *Bacillus cereus.*[59] This enzyme, Lipoprotein intramolecular transacylase (Lit), is proposed to facilitate an *O*-to-*N* acyl transfer yielding lyso-form LPs from diacylated LPs. Additionally, in 2019, Armbruster *et al*.[38] discovered a second Lit type enzyme in *Listeria monocytogenes* and *Enterococcus* spp., named Lit2. These bacteria are pathogenic, with *L. monocytogenes* causing listeriosis,[206] a serious food-borne illness while *Enterococci* can cause many serious nosocomial infections, including meningitis.[216] Critically, Lit2 is expressed only in the presence of environmental copper(II) and is genetically encoded within a copper resistance operon. This suggests Lit activity may play a role in imparting copper resistance to the microbe and that use of copper as an antimicrobial may result in Lit expression and subsequently reduced innate immune activation. Due to the ubiquity of copper(II) in healthcare settings, further investigation of this hypothesis is necessary to determine the clinical implications of Lit.

4.1.2 Mechanism of Lit

Lit is known to produce *N*-acyl-*S*-monoacylglyceryl LPs from diacylated LPs, however, the mechanism of Lit-mediated acyl transfer has not been described. Two methods of acyl transfer have been proposed. The first involves an intermolecular two-step process in which the ester-linked acyl chain is first transferred onto the enzyme and then transferred onto the N-terminus, in a Lnt-like reaction. The second involves the intramolecular transfer of the glyceryl ester-linked acyl chain directly to the N-terminus

through a cyclic tetrahedral intermediate. The position of the ester-linked acyl chain on the glyceryl core in the lyso-form product has not been determined and it is not known which ester undergoes acyl transfer. Hence, the remaining ester may be present in the product on either the *sn*-1 or *sn*-2 position as shown in **Scheme 4.1**. Depending on which ester-linked chain is transferred, intramolecular acyl transfer would involve either an 8- or 9-membered cyclic intermediate. Although cyclic transition states (TSs) of this size are thermodynamically highly disfavoured due to ring strain,[217] *O*-to-*N* acyl transfer over TSs of this size has previously been reported in rare instances in the synthetic chemistry literature and are discussed below.



Scheme 4.1. Lit is proposed to mediate an *O*-to-*N* acyl transfer. The two possibilities of ester-linked acyl chain position in the product of the Lit reaction are shown here.

4.1.2.1 *O*-to-*N* Intramolecular Acyl Transfer

Intramolecular *O*-to-*N* acyl transfer is an amide bond-forming reaction that has gained particular attention in the field of chemical protein synthesis where several methodologies utilise a ligation reaction followed by an intramolecular *O*-to-*N* acyl transfer to yield a native peptide bond.[218-221] To this end, several groups have systematically studied *O*-to-*N* acyl transfer and have reported several factors which dictate the efficacy of the shift, including: TS cycle size,[222-226] favourable orientation of ester and amine functionalities through preorganisation,[225] nucleophilicity of the amine,[227] electrophilicity of the ester,[225] pH,[228] and solvent effects.[223] These studies emphasise that to undergo *O*-to-*N* intramolecular acyl transfer a molecule must be capable of adopting a conformation with a short (< 5 Å) N-C geometrical separation.[225] As this reaction proceeds through a cyclic TS the ring strain, and therefore the ring size, of the cyclic intermediate has an important influence on the positioning of amine relative to the ester and impacts reaction rate.

O-to-*N* Acyl transfer has been demonstrated over a wide range of TS sizes from 5-34 membered rings.[219, 229, 230] As minimal ring strain exists in a 5 membered cyclic

intermediate,[217] it is unsurprising that *O*-to-*N* acyl transfers though 5 membered TS are the most widely reported.[230-232] This reaction can occur spontaneously under mild aq. conditions at neutral pH and is accelerated at basic pH.[228] During the development of thiol capture ligation, Kemp *et al.*[219, 233] studied *O*-to-*N* acyl shift over both small (5 and 6) and medium (9, 11, 12 and 13) sized cyclic intermediates. Despite the particularly severe ring-strain present in medium-sized cyclic intermediates, this work demonstrated that acyl transfer can be promoted over these ring sizes using a rigid scaffold between amine and ester to place their nucleophilic and electrophilic sites in a favourable conformation. By utilising this strategy to provide a high level of preorganisation, acyl transfer over a 9 membered cyclic TS was surprisingly found to proceeded quantitatively in DMSO with a $t_{1/2}$ of 1.3 s.[223] However, structural alterations to the scaffold were not well tolerated and resulted in a significant decrease in reaction rate.[234] This highlights the dramatic effect of molecule conformation on the kinetics of this acyl shift. This high susceptibility to changes in structure makes it impossible to predict the outcome of an *O*-to-*N* acyl intramolecular transfer reaction based solely on TS cycle size. The authors demonstrate that successful acyl transfer requires minimal steric repulsion, a rigid structural scaffold and stabilising hydrogen bonding interactions.[218]

Katrisky *et al.* later reported *O*-to-*N* acyl shift of Ser/Thr O-acyl isotripeptides over 8 membered cyclic intermediates.[224-226, 235] Once again, reactivity was shown to be strongly affected by preorganisation and conformation. The reaction was highly susceptible to steric bulk around the ester and variation of the O-acyl peptide structure was not well tolerated. These reactions did not proceed spontaneously and required microwave conditions in DMF:piperidine to promote product formation. The authors highlight the contrast between this successful *O*-to-*N* acyl transfer reaction and their earlier report of failed *S*-to-*N* acyl transfers with structurally similar *S*-acylated compounds over an 8 membered TS. In the case of *S*-to-*N* acyl transfer, the 8 membered TS was highly disfavoured for the range of the compounds reported and no acyl transfer was detected.[236, 237]

These studies highlight that intramolecular *O*-to-*N* acyl transfer can only occur when ester and amine are favourably positioned. Intramolecular acyl transfers over a highly strained 8 and 9 membered cyclic TSs have been demonstrated in rare, specific cases when an appropriate level of structural preorganisation is present. Therefore, if Lit can position the ester and amine in the required conformation, it is conceivable that the acyl

transfer from either the *sn*-1 or *sn*-2 position could take place in an intramolecular fashion.

4.2 Aims of this work

The recent discovery of Lit may provide considerable opportunities for antibiotic discovery. Lyso LPs may have evolved to shield the microbe from the host immune system. By interrupting the production of these virulence factors, immunopotent diacylated LPs are instead produced and are presented to immune cells. Consequently, Lit inhibition may promote recognition and clearance of the infection of by the immune system of the host. However, before the therapeutic potential of targeting this enzyme can be evaluated, a thorough understanding of the physiological implications of lyso LPs and hence Lit action is required.

In collaboration with the Caffrey Group, this study aims to develop synthetic substrates of Lit to characterise both the product and the mechanism of the Lit reaction. Discerning the mechanism of the reaction first requires the identification of amino acid residues that are involved in catalysis. Towards this aim, synthetic substrates will be utilised to develop an activity assay for Lit to screen the activity of Lit analogues in which highly conserved residues have been mutated. Using site-directed mutagenesis coupled with activity profiling to identify essential residues, a putative active site can be classified. Supporting evidence for this can be sought from protein crystallography, particularly in the form of capturing a substrate or product in the putative active site. Development of the activity assay first requires investigation of the substrate scope of the enzyme which was the focus of initial efforts in this project.

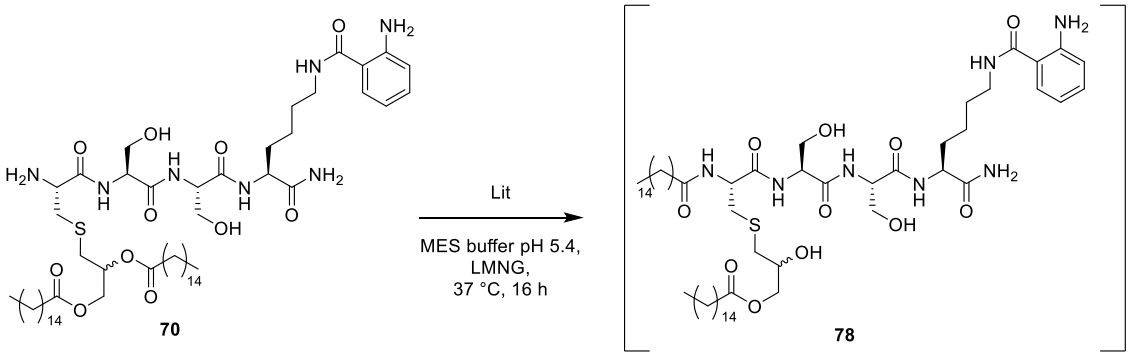
4.3 Functional Analysis of Lit with Peptide 70[l]

As detailed in **Chapter 3**, previous efforts have identified a fluorescent diacylated lipopeptide (peptide **70**) which is a substrate of Lnt. As Lit acts on the same native substrates as Lnt, peptide **70** was screened for activity with Lit. Due to the low molecular weight of this peptide (1092.53 g/mol) it was proposed that the minor change in polarity resulting from acyl transfer, in which an amino group is acylated and a hydroxyl group is produced, would be sufficient for separation of substrate and product by TLC. This

[l] Lit assays and TLC analysis were performed by Dr. Dietmar Weichert and Dr. Samir Olatunji, Caffrey Group, Biochemistry, TCD.

would enable the tracking of the reaction by TLC in a similar manner to that described with Lnt.

To investigate Lit activity with peptide **70**, the peptide was incubated in buffer with the enzyme for 16 h. The reaction was extracted with EtOH:$CHCl_3$ and the extract loaded onto a normal phase silica TLC plate. The subsequent development of the plate revealed a new spot at higher R_f, indicating the formation of a single product during the enzymatic reaction. A proposed acyl-transfer reaction accounting for the production of this new compound is shown in **Scheme 4.2**. This promising result indicates enzyme activity with peptide **70**, however, a significant amount of unreacted starting material was evident from TLC analysis.

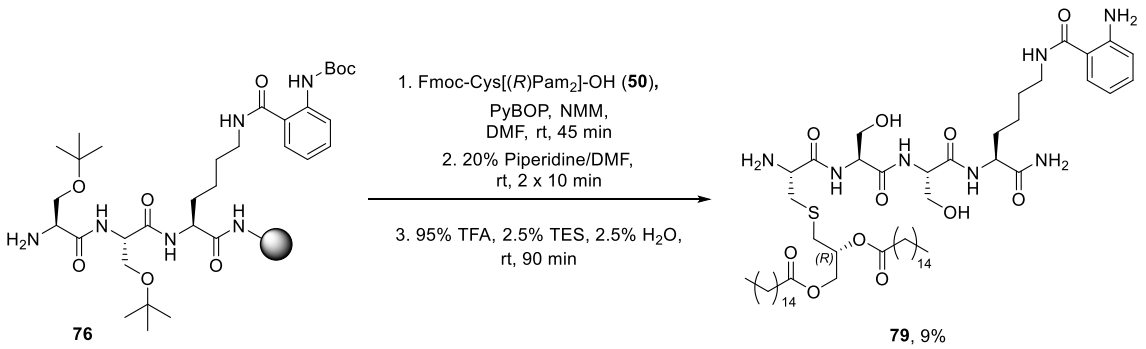


Scheme 4.2. The reaction of peptide **70** with Lit yielded the potential lyso-peptide product **78**. Acyl transfer is shown in this instance from the *sn*-2 position.

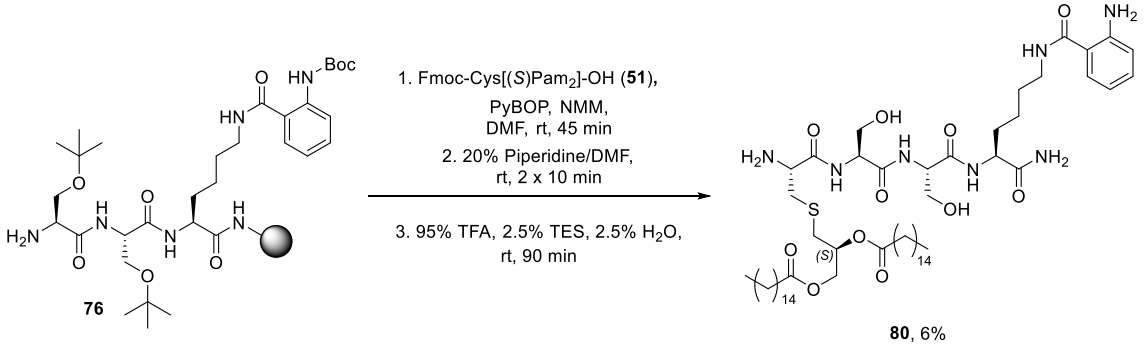
4.4 Lit Activity with Individual Diastereomers 79 and 80

4.4.1 Synthesis and Characterisation of Peptides 79 and 80

Due to the propensity of lipoprotein processing enzymes to be stereoselective, as observed with LspA, it is possible that the Lit only recognises one of the diastereomers of peptide **70**. To test this hypothesis, both the (*R*)-DAG and (*S*)-DAG isomer of this peptide were synthesised separately. Both peptides were synthesised by SPPS under the same conditions previously described for peptide **70** (**Section 3.4.1**). The common resin-bound intermediate **76** was functionalised with Fmoc-[(*R*)Pam_2]-OH (**50**) to yield peptide **79**, and Fmoc-[(*S*)Pam_2]-OH (**51**) to afford peptide **80**, as shown in **Schemes 4.3** and **4.4** respectively. Each peptide was isolated by normal phase silica column chromatography in an overall yield of 9% for peptide **79** and 6% for peptide **80**.



Scheme 4.3. Synthesis of peptide **79** from common resin intermediate **76** which was synthesised as described in **Scheme 3.3**.



Scheme 4.4. Synthesis of peptide **80** from common resin intermediate **76** which was synthesised as described in **Scheme 3.3**.

4.4.2 NMR Analysis of Substrate

To preface later studies in which NMR analysis was used to characterise the product of the Lit reaction (**Section 4.5.4**), the NMR experiments used to support the identity of the peptide substrate are discussed here in detail using the diastereomeric mixture **70** as an example. Both MS and NMR analysis were used to confirm the structure of each peptide. However, while MS provides the molecular weight of the peptide, further structural information cannot be attained from this spectrometric technique in the absence of tandem MS experiments. Instead, the peptide structure was unambiguously confirmed by NMR analysis. The NMR data from each peptide was readily interpretable due to the short amino acid chain resulting in minimal resonance overlap in the 1D spectra.

The ^{1}H NMR spectra of each of the 3 samples **70**, **79**, and **80** are consistent with each other (**Fig. 4.1**). However, a minor difference between the ^{13}C spectra of **70**, **79**, and **80** is observed which can be used to identify the diastereomer present. As with the LspA FRET probe **49**, each diastereomer exhibits a distinct ^{13}C signal arising from the

methylene group neighbouring the glyceryl-CH stereocentre. In the mixture **70**, the signals are separated by 0.06 ppm as shown in **Fig. 4.2**. Each diastereomer exhibits a single signal in this region confirming the diastereomeric purity of peptides **79** and **80**.

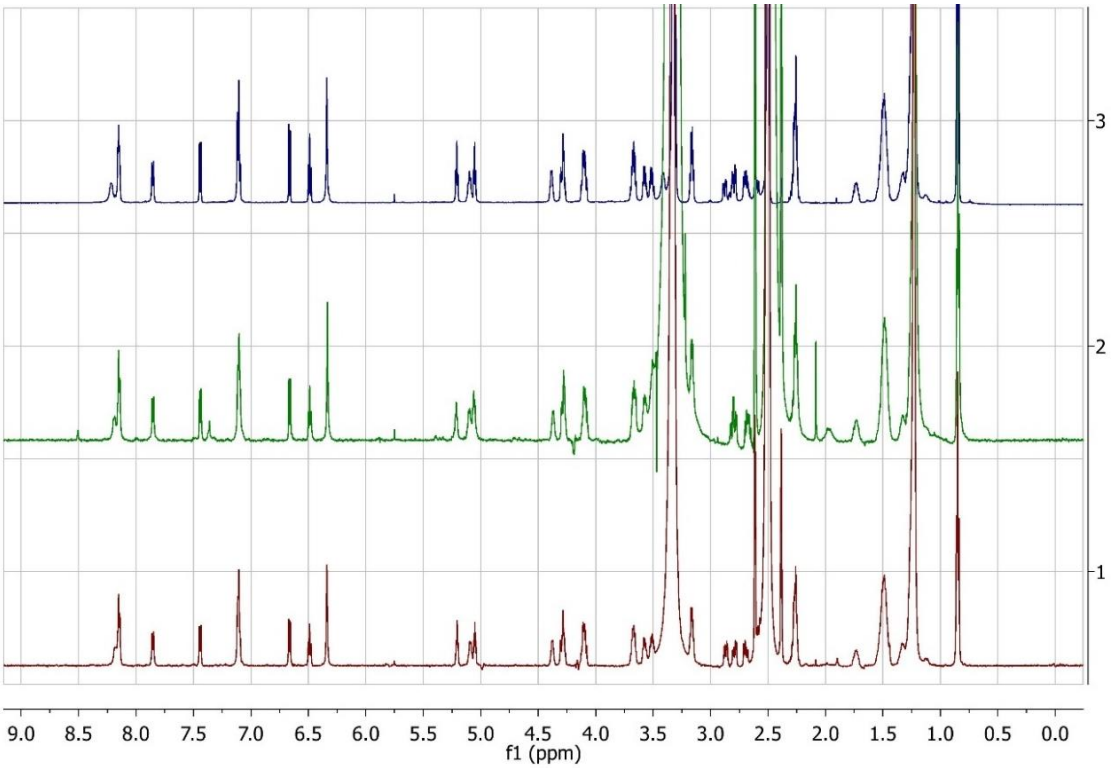


Figure 4.1. ^{1}H NMR (600 MHz, DMSO-d_6) comparison of the diastereomeric mixture **70** (top, blue) with each isomer: **80** (*S*-DAG isomer – middle, green) and **79** (*R*-DAG - bottom, red).

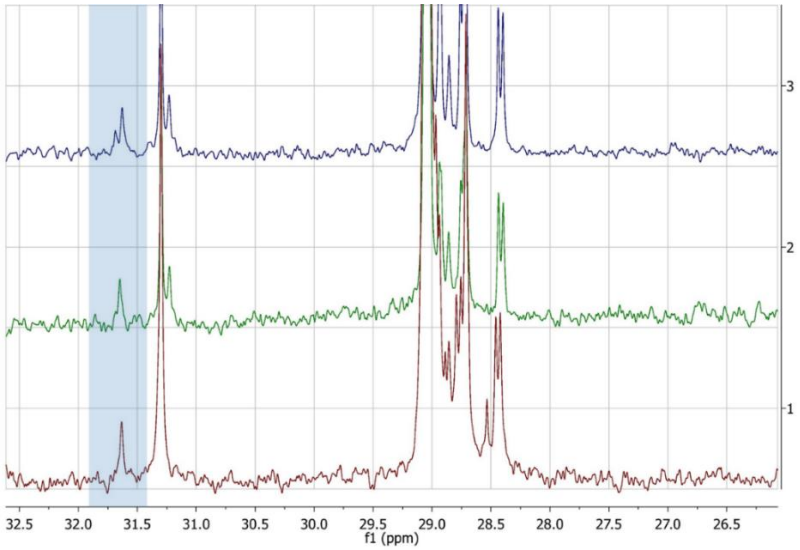


Figure 4.2. Expanded view of ^{13}C NMR (151 MHz, DMSO-d_6) comparison of the diastereomeric mixture **70** (top, blue) with each isomer: **80** (*S*-DAG isomer – middle, green) and **79** (*R*-DAG - bottom, red).

The ^{1}H and ^{13}C resonances of each residue and the connectivity of the amino acids are clearly defined by 2D NMR data. The ^{1}H-^{13}C HSQC spectrum of diastereomeric mixture **70** (**Fig. 4.3**) shows four resonances between (7.50, 135.0) – (6.40, 110.0) ppm (highlighted in blue) which arise from four aromatic CH groups, confirming the presence of the Abz spin system. A further four resonances between (4.50, 57.0) – (3.40, 51.0) ppm arise from the four α-CH groups (highlighted in red) confirming the presence of four α-amino acids. Additionally, the functionalised cysteine side chain has multiple components which are colour coded in the spectrum and structure in **Fig. 4.3**. The glyceryl core and cysteine β-CH_2 signals are highlighted in purple while the four lipid-associated signals (2.40, 35.0) – (0.75, 12.0) ppm are highlighted in green. The two serine side chain hydroxyl group ^{1}H signals are readily identifiable at 5.21 ppm and 5.05 ppm due to their lack of HSQC cross-peak. No correlation is present as these protons are not bound to a carbon atom.

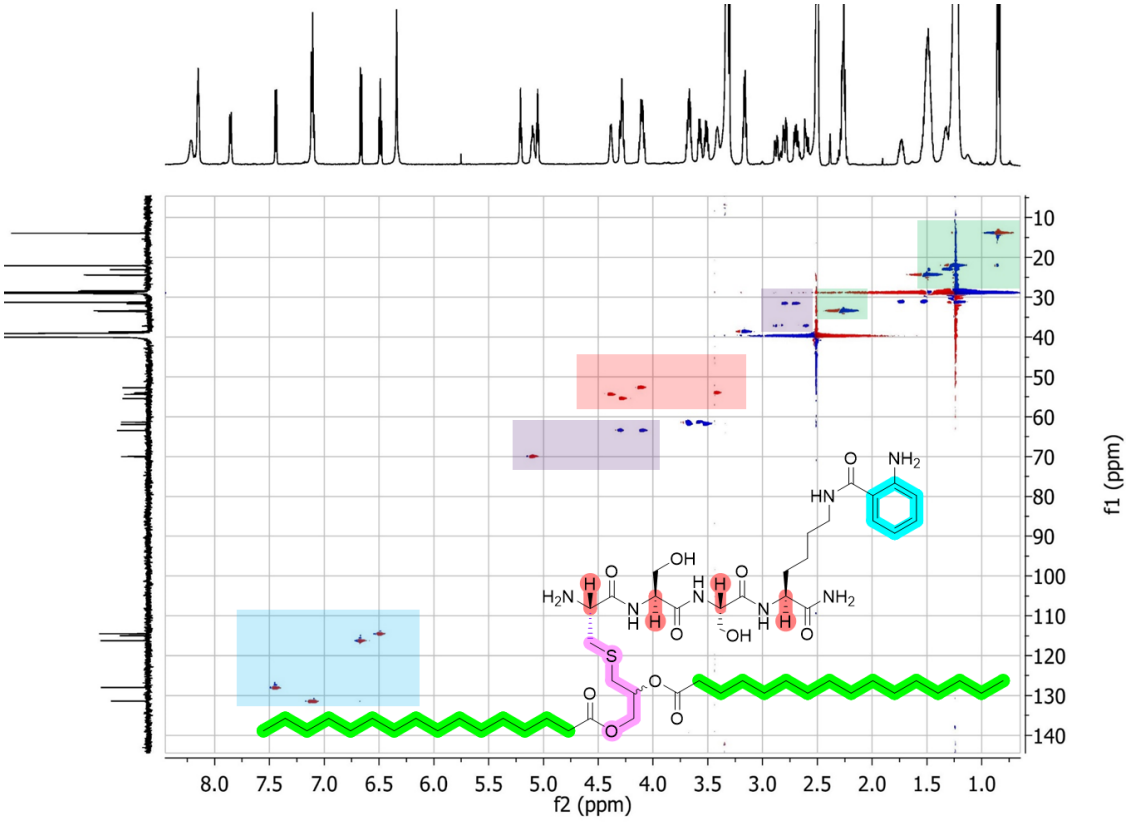


Figure 4.3. ^{1}H-^{13}C HSQC (DMSO-d_6) of compound **70**. Regions of the peptide are colour coded with the corresponding signals in the spectrum.

The ^{15}N-^{1}H HSQC reveals six NH signals as shown in **Fig. 4.4**. However, compound **70** contains seven distinct NH environments. The signal arising from the N-terminal cysteine NH_2 group is absent. Furthermore, the cysteinyl carbonyl does not exhibit a signal in the ^{13}C spectrum. Detection of the target peptide molecular ion by MS confirms this residue is present. It is feasible that peptide self-assembly in DMSO solution during NMR analysis has led to an increase in the relaxation rate of spins located at the N-terminal end of the peptide, broadening and effectively quenching these two signals. Additional resonances arising from the cysteine residue, such as the signal of the α-CH (3.41 ppm) and β-CH_2 (2.87 and 2.80 ppm), appear as broad, or weak, signals in the ^{1}H-^{13}C HSQC and ^{1}H spectrum, supporting this hypothesis.

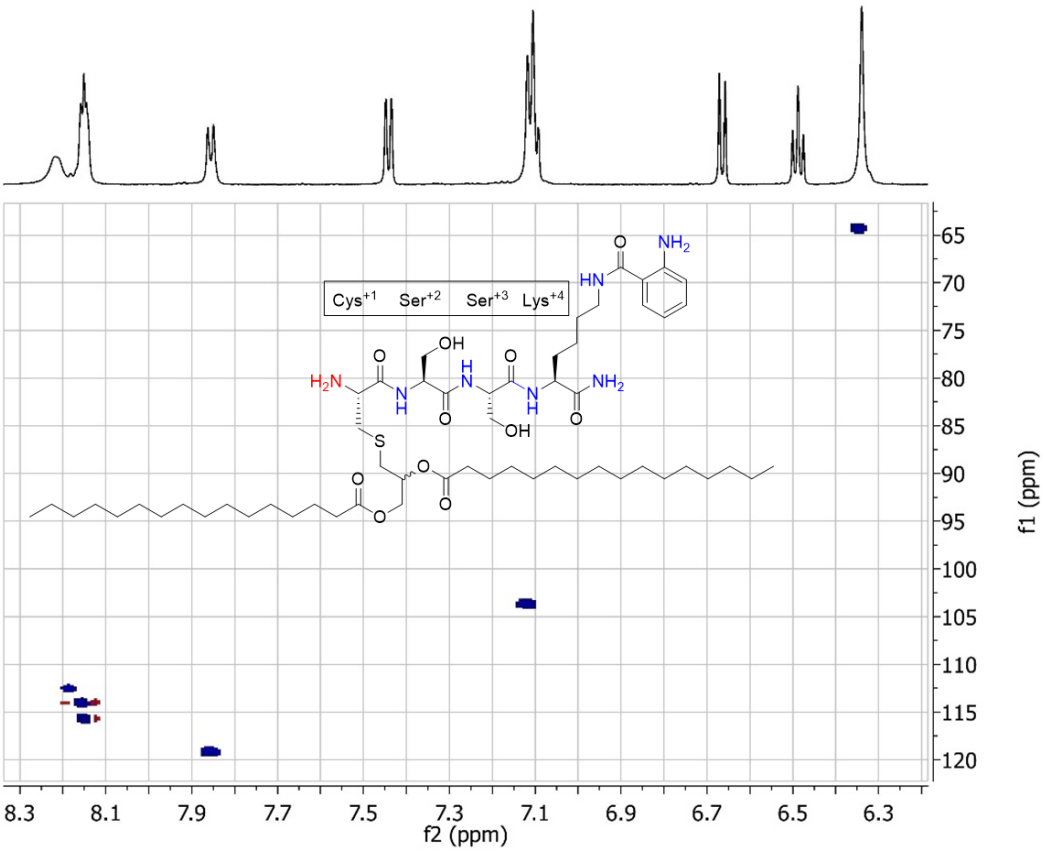


Figure 4.4. ^{15}N-^{1}H HSQC of peptide **70** showing six NH signals. The seven NH regions in the peptide are highlighted in colour with the cysteine NH_2 group that does not produce a signal highlighted in red.

Selective 1D TOSCY experiments (**Fig. 4.5**) were used to isolate the ^{1}H resonances associated with each amino acid. Through selective excitation of each serine hydroxyl ^{1}H resonance (5.21 and 5.05 ppm), each associated β-CH_2, α-CH and NH resonance was revealed (row 3 and 4, **Fig. 4.5**). Similarly, excitation of the lysine α-NH signal (7.86

ppm) revealed the lysine α-CH, side chain CH_2 and NH signals (row 1). Excitation of the glyceryl-CH signal (5.10 ppm) shows the CH_2 signals arising from methylene groups positioned on either side of the CH moiety (row 2).

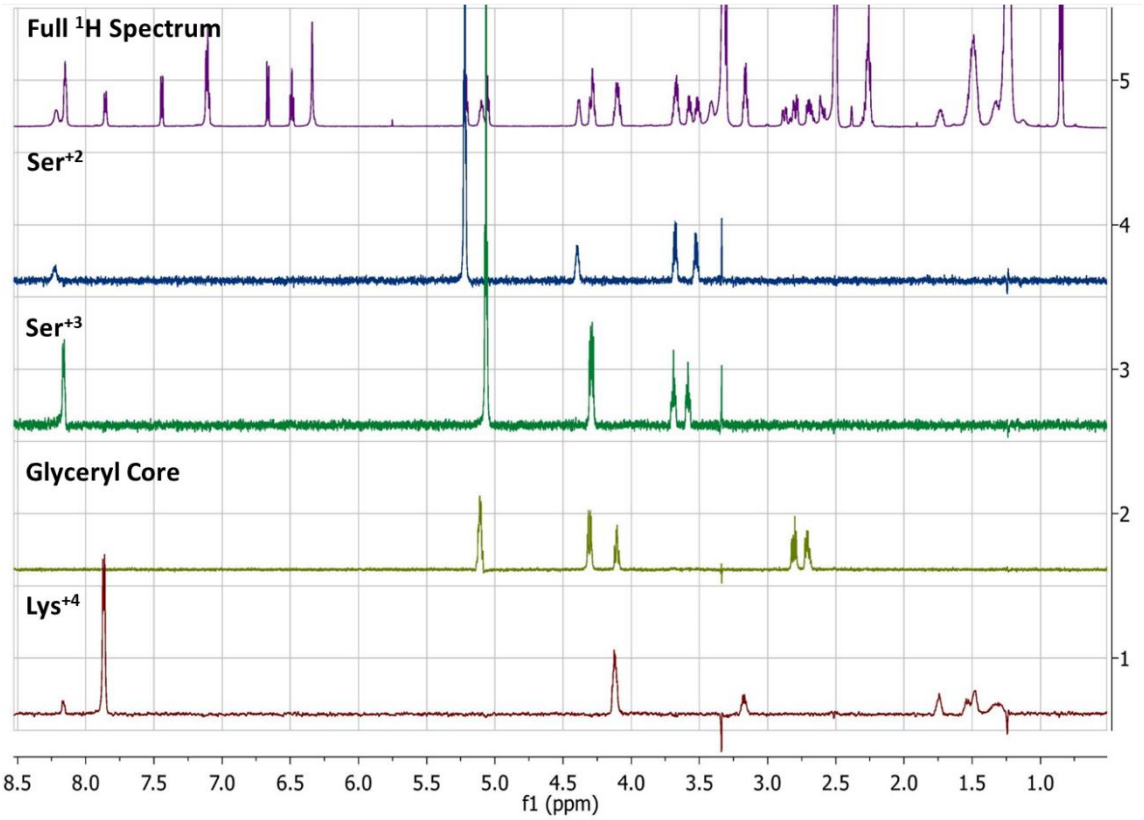


Figure 4.5. Selective 1D TOCSYs reveal each resonance associated with serine, lysine and glyceryl residues in peptide **70**.

Selective excitation of a signal arising from cysteine was not possible due to the overlap of the broad α-CH signal (3.41 ppm) with residual water in the sample. Additionally, the resonance frequencies of the β-CH_2 group partially overlap with both the glyceryl CH_2 group and the DMSO solvent satellite peak (2.61 ppm). Resultantly, a 2D TOCSY was performed in which the correlations of this cysteine residue could be observed. The 2D TOCSY of peptide **70** is shown in **Fig. 4.6**. In this spectrum, a cross peak correlation between the cysteine α-CH signal (3.41 ppm) and β-CH_2 (2.87 and 2.80 ppm) is highlighted in blue. This correlation reaffirms the presence of the cysteine residue. Furthermore, the lack of a correlating cysteine NH_2 signal confirms this resonance relaxes at a rate which cannot be detected.

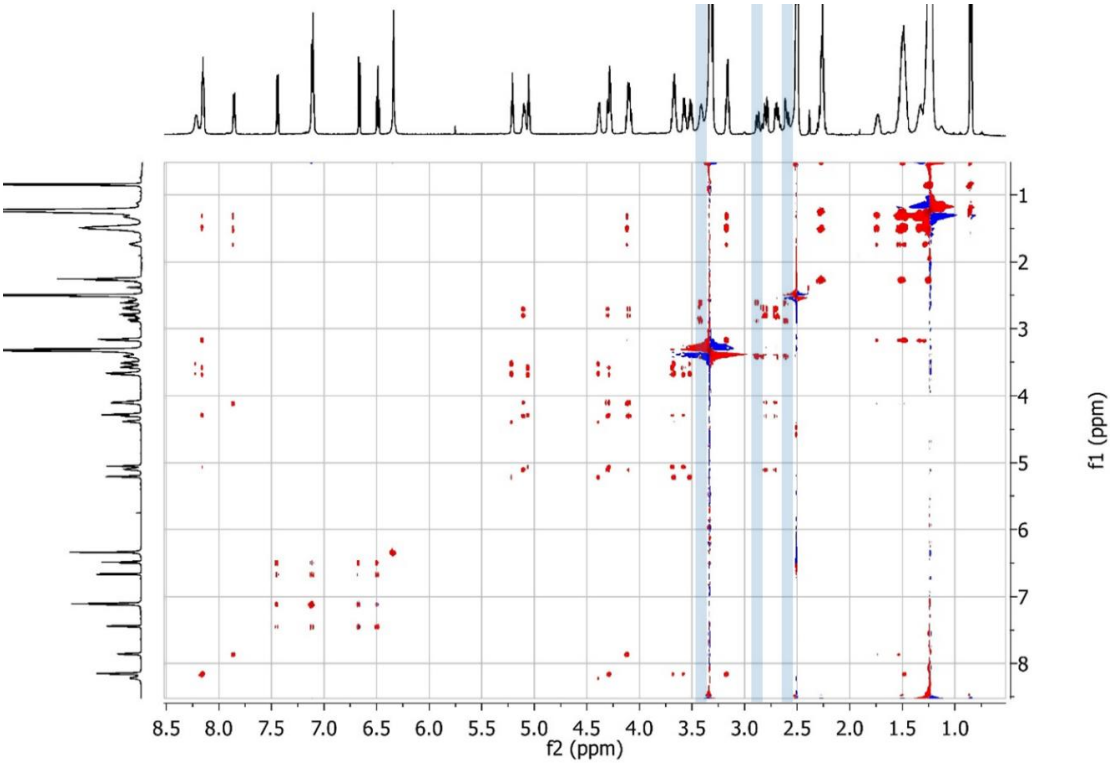


Figure 4.6. 2D TOCSY of peptide **70** showing through bond ^{1}H-^{1}H correlations. The notable correlations of the cysteine α-CH and β-CH_2 signals are highlighted in blue.

Differentiation of the two serine spin systems was possible through a HMBC experiment. Typically, this long-range ^{1}H-^{13}C experiment can readily detect correlations between ^{1}H and ^{13}C resonances separated by 2 to 3 bonds, including through a heteroatom. This makes it possible to elucidate the connectivity of the amino acids. For example, the lysine α-NH signal (7.86 ppm) correlates to one carbonyl ^{13}C signal at 169.7 ppm as highlighted in blue (**Fig. 4.7**). This carbon resonance, in turn, correlates to ^{1}H resonances of one CH and one CH_2 group, highlighted in blue. As such, it can be concluded that these ^{1}H signals belong to the serine residue neighbouring lysine. In-keeping with LP nomenclature, this residue is referred to here as serine^{+3}.

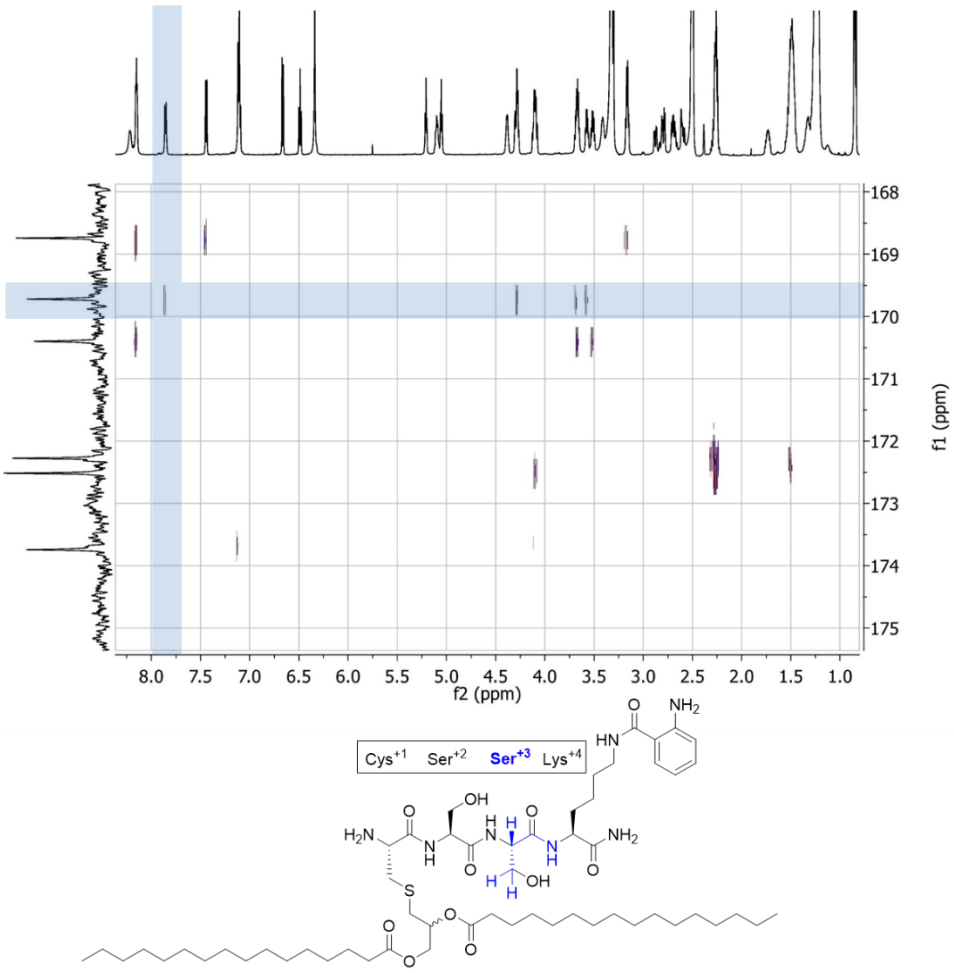


Figure 4.7. Expanded view of the ^{1}H-^{13}C HMBC of peptide **70** showing the correlations between resonances of the lysine α-NH, serine^{+3} carbonyl, α-CH, and βCH_2.

4.4.3 Testing with Lit[m]

On testing of each of the diastereomers for activity with Lit, the (*R*)-DAG isomer **79**, was found to be processed with approximately 70% conversion to the product after 30 min, as shown in **Fig. 4.8**, lanes 3 and 4. Contrastingly, 0% conversion of the (*S*)-DAG isomer **80**, was observed after 16 h, as shown in lanes 11 and 12. This result shows that Lit is stereospecific, only recognising a LP mimetic with (*R*)-stereochemistry at the glyceryl C2 position. This specificity indicates that the mechanism of the enzyme is highly sensitive to the spatial arrangement of the substrate's *sn*-2 acyl chain in the active site, a key finding which has implications for inhibitor design.

[m] Lit assays and TLC analysis were performed by Dr. Dietmar Weichert and Dr. Samir Olatunji, Caffrey Group, Biochemistry, TCD.

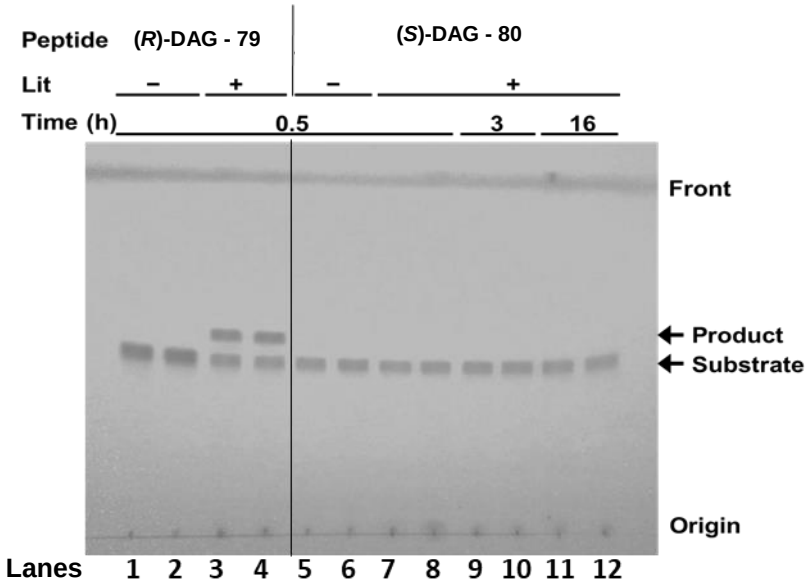


Figure 4.8.[n] TLC analysis of Lit reaction with (*R*)-DAG peptide **79** (lanes 1-4) and (*S*)-DAG peptide **80** (lane 5-12). With both compounds, no product is formed in the absence of Lit (lanes 1,2, 5, 6). With (*R*)-DAG peptide **79** product formation is observed after 0.5 h (lanes 3 and 4). With (*S*)-DAG peptide **80**, no product is observed despite an extended reaction time of up to 16 h with Lit (lanes 5-12).

With an optimised substrate in hand, the TLC-based assay was quantified using a calibration curve, with substrate concentration ranging from 0 to 10 µg (**Fig. 4.9A**).[n] Following the loading of the substrate, the plate was developed and photographed. The signal was quantified by densitometric analysis of each spot using ImageJ. **Fig. 4.9C** shows the time course of the reaction from 0 - 240 min, with the data plotted as a reaction progress curve in **Fig. 4.9D**. As shown, more than 90% of the (*R*)-DAG substrate **79** was converted to product after 240 min.

[n] Performed by Dr. Dietmar Weichert and Dr. Samir Olatunji, Caffrey Group, Biochemistry, TCD.

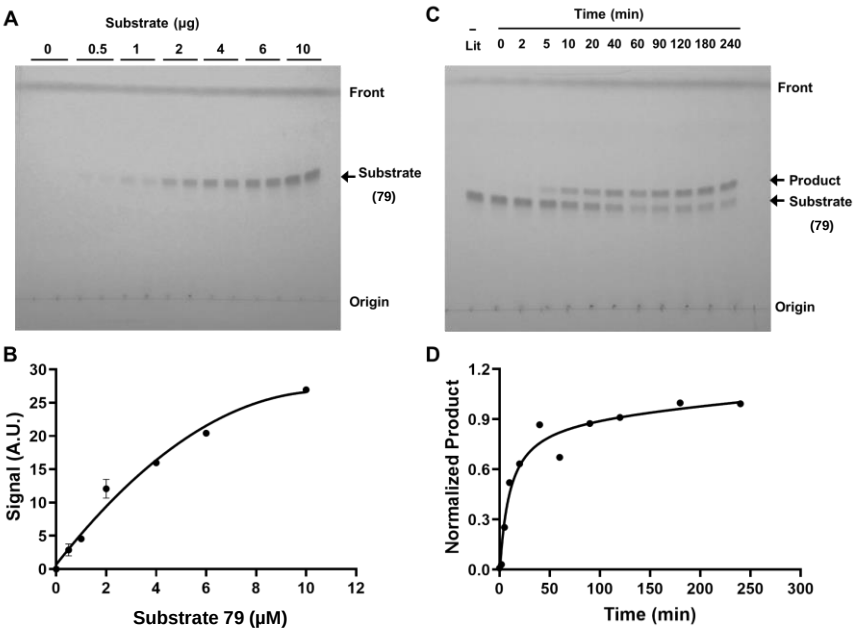


Figure 4.9.[n] **A**: Substrate **79** concentrations of 0-10 µg used to develop the calibration curve shown in **B**. A digital image of the UV-illuminated TLC plate was taken and the substrate band intensity was quantified by densitometric analysis of the data in **A** using ImageJ. The results were plotted using Prism 8 (GraphPad) and shown in **B**. **C**: TLC showing the consumption of substrate **79** and formation of product during the Lit reaction between 0-240 min. This data is plotted as a reaction progress curve in **D**.

4.5 Lit transfers the *sn*-2 acyl chain to the N-terminus

To characterise the action of Lit, we aimed to isolate and characterise the product of the enzymatic reaction. As this proposed reaction involves acyl transfer, both the substrate and proposed product have the same molecular weight. Resultantly, tandem MS is necessary to provide a structural comparison of the substrate and product. MS/MS was used by both Kurokawa *et al*.[40] and Armbruster *et al*.[59] to differentiate lyso LPs from diacylated LPs. In these previous reports of MS/MS analysis of lyso LPs, diagnostic fragmentation between the *sn*-1 and *sn*-2 position has never been observed. The dagylated cysteinyl residue is prone to fragment to form the dehydroalanyl species. This provides evidence of the mass of the N-terminal amide-linked acyl chain but does not provide evidence about the position of the ester-linked acyl chain. It is therefore unknown if the ester-linked acyl chain resides on the *sn*-1 or *sn*-2 position. Determining which acyl chain is transferred is key to enabling computation studies of the mechanism of the enzyme. If this reaction proceeds intramolecularly, the size and thus the stability of the cyclic intermediate TS is dependent on the location of the ester. As such, using computational

studies to identify an energetically feasible reaction mechanism requires knowledge of which acyl chain is transferred

4.5.1 Differentiating the Acyl Chains

In 2017, Armbruster *et al.*[59] reported putative evidence indicating *sn*-2 acyl transfer. The enzymes involved in phospholipid biosynthesis which acylate the *sn*-1 and *sn*-2 phospholipid positions (PlsB and PlcC, respectively) have characteristic substrate specificities.[238] Acyl chain length and saturation impact the activity of these enzymes. PlsB processes $C_{16:0}$ and $C_{18:1}$ acyl chains[o] but does not typically process $C_{16:1}$ acyl chains.[239] Therefore the *sn*-1 position of a phospholipid, and consequently a diacylated LP, will typically not contain an unsaturated $C_{16:1}$ acyl chain. Armbruster *et al*. isolated a LP from *E. coli* which had been *N*-acylated by Lnt. Lnt is known to acylate the N-terminus with an acyl chain originating from the *sn*-1 position of a phospholipid.[240, 241] Mass analysis of the fragment ion arising from the N-terminal dehydroalanyl fragment of the LP revealed a $C_{16:0}$ *N*-linked acyl chain. When the authors introduced Lit into *E. coli*, mass analysis of the N-terminal dehydroalanyl fragment ion of the LP revealed a $C_{16:1}$ *N*-linked acyl chain. As both reactions were performed in *E. coli*, both Lit and Lnt had access to the same *E. coli* phospholipid pool. If Lit transferred the *sn*-1 acyl chain, the N-terminus of lyso LP should feature a $C_{16:0}$ acyl chain. As this is not the case, it is likely that the N-linked acyl chain in lyso LPs does not originate from the *sn*-1 position. Therefore, Lit-mediated transfer of the *sn*-2 acyl chain is indicated.

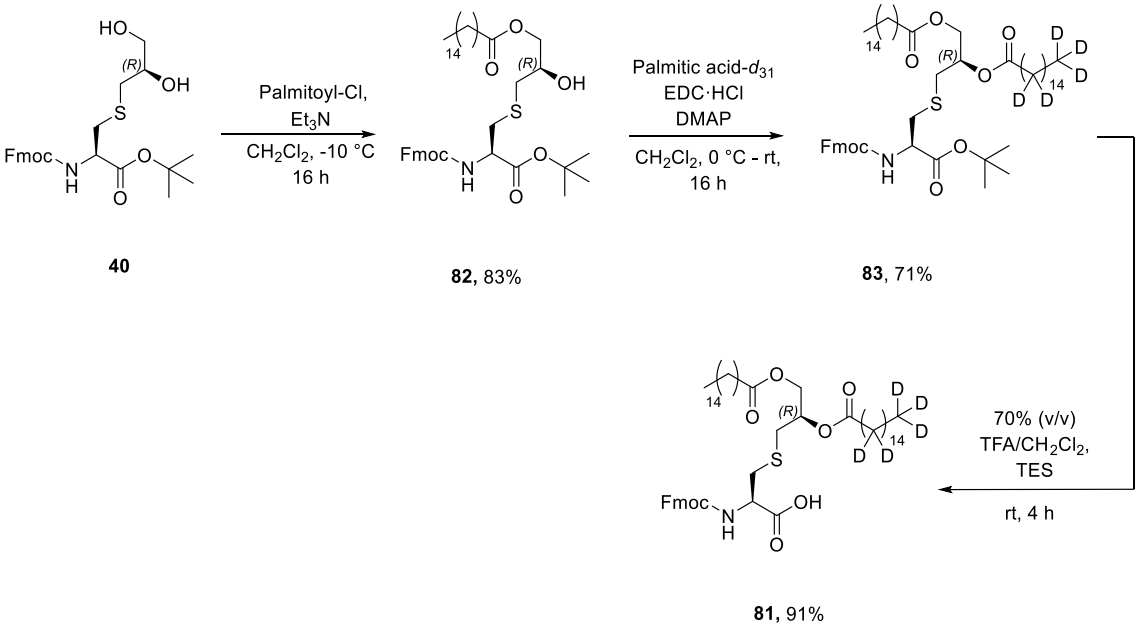
To directly determine which acyl chain is transferred by Lit by MS/MS, we aimed to selectively label one of the acyl chains in the substrate. To achieve this, a perdeuterated palmitoyl group was introduced at the *sn*-2 position of peptide **79**. A deuterated analogue enables differentiation of the acyl chains by MS/MS without impacting substrate recognition by the enzyme. This study aimed to detect the mass of a fragment containing only one of the two acyl chains, taking advantage of the reported tendency of these molecules to form the dehydroalanyl fragment. In this way, it could be determined whether the amide-linked acyl chain was perdeuterated or protiated, and hence which acyl chain had been transferred. The transfer of the substrate's perdeuterated *sn*-2 acyl chain would result in a perdeuterated amide-linked acyl chain in the product, while the

[o] Shorthand fatty acid nomenclature ($C_{x:y}$) is used to denote the number of carbons in the acyl chain (x) and the number of unsaturated bonds (y).

transfer of the *sn*-1 acyl chain would result in a protiated acyl chain becoming amide-linked.

4.5.2 Synthesis of Deuterium-Labelled Substrate

Prior to SPPS of the deuterium-labelled analogue, the corresponding deuterium-containing cysteine residue was first synthesised. Fmoc-Cys[(*R*)Pam,Pam-d_{31}]-OH (**81**) was synthesised from diol **40** in 3 steps with an overall yield of 54%, as shown in **Scheme 4.5**. In the initial step, the difference in reactivity between the primary and secondary alcohols of diol **40** was exploited to enable regioselective esterification of the primary hydroxyl using palmitoyl chloride according to a procedure reported by Arai *et al*.[242] The use of reported conditions, which utilised a temperature of 0 °C - rt, afforded product **82** in a yield of 38%. This is comparable to the reported yield of 44%. However, minor amounts of both diacylated and *sn*-2 monoacylated product were observed by TLC throughout the course of the reaction, despite the use of 0.9 equiv. of palmitoyl chloride. Consequently, the temperature was decreased to -10 °C to reduce the rate of these undesirable side reactions. Under these conditions, no diacylated nor *sn*-2 acylated product was observed by TLC. The monoacylated product **82** was formed cleanly and isolated by column chromatography in a yield of 83%. NMR analysis confirmed the *sn*-1 positioning of the acyl chain in the isolated product. The perdeuterated palmitoyl chain was introduced *via* EDC·HCl and DMAP mediated Steglich esterification of monoacylated **82** with palmitic acid-d_{31} which afforded the product in a yield of 71%. The ^{1}H NMR of diacylated **83** in comparison to starting material **82** (**Fig. 4.10**) revealed a large downfield shift of the glyceryl *sn*-2 C<u>H</u> signal from 3.93 ppm to 5.16 ppm, confirming the acylation of the *sn*-2 hydroxyl group. Supporting evidence was provided by ^{2}H NMR which, along with MS analysis, confirmed the presence of deuterium. The ^{2}H NMR spectrum (**Fig. 4.11**) reveals four signals which are characteristic of the palmitoyl acyl chain. Finally, standard TFA mediated deprotection of the *tert*-butyl ester of diacylated **83** proceeded in a yield of 91% to afford Fmoc-Cys[(*R*)Pam,Pam-d_{31}]-OH (**81**).



Scheme 4.5. The synthesis of Fmoc-Cys[(*R*)Pam,Pam-d_{31}]-OH (**81**) over 3 steps from diol (**40**).

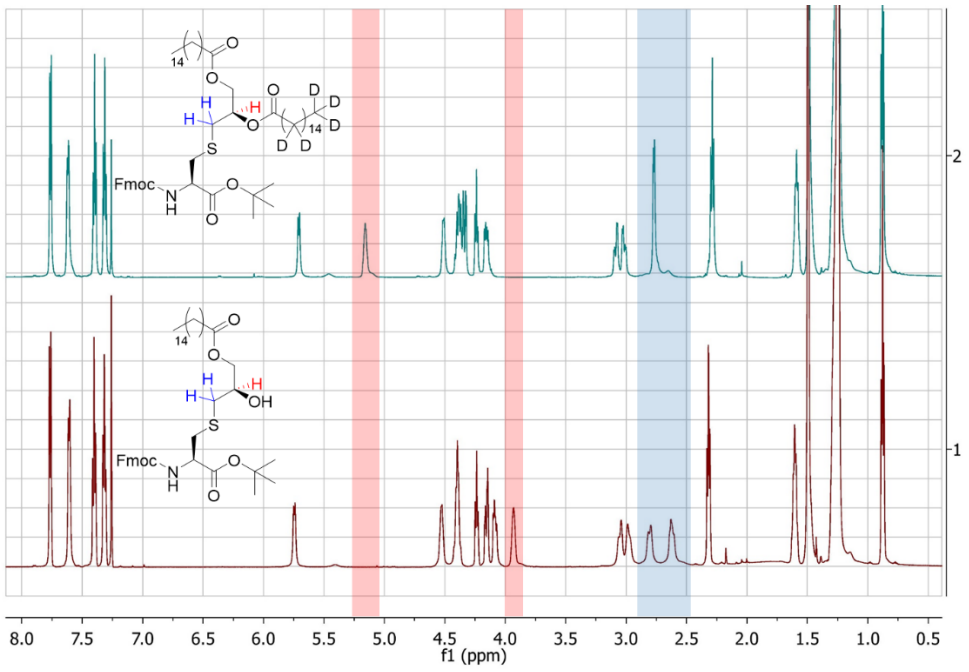


Figure 4.10. ^{1}H NMR (600 MHz, $CDCl_3$) comparison of diacylated product **83** (blue, top) and monoacylated starting material **82** (red, bottom) confirms acylation of the *sn*-2 position by palmitic acid-d_{31}. The large shift of the *sn*-2 glyceryl C<u>H</u> signal is highlighted in red. The signals of the neighbouring S-C<u>H</u>$_2$ group undergo a minor shift as highlighted in blue.

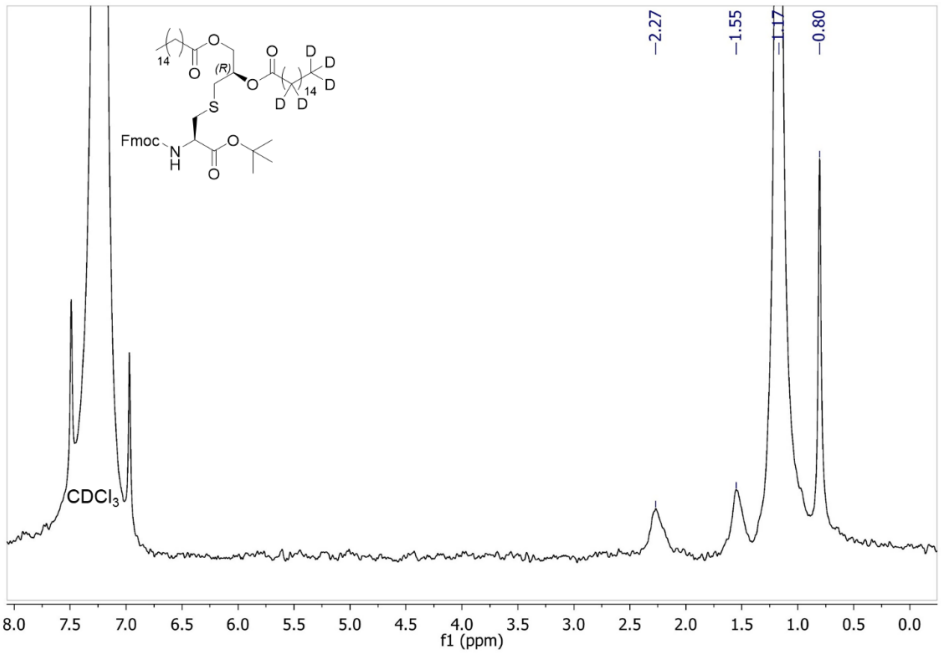
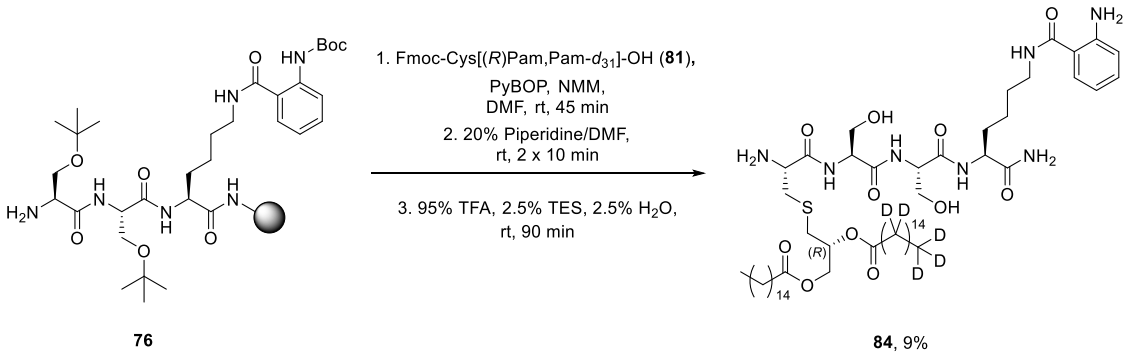


Figure 4.11. Deuterium ^{2}H NMR (400 MHz, $CDCl_3$) of compound **83** confirms the presence of a deuterated palmitoyl group.

The target deuterium-labelled peptide **84** was prepared by SPPS from resin-bound intermediate **76** with PyBOP/NMM mediated coupling of Fmoc-Cys[(*R*)Pam,Pam-d_{31}]-OH (**81**). Following global deprotection and resin cleavage, the peptide was isolated by column chromatography in an overall yield of 9% as depicted in **Scheme 4.6**.



Scheme 4.6. Synthesis of deuterium-labelled peptide **84**.

The NMR of perdeuterated peptide **84** was consistent with the protiated analogue **79** with the exception of the lipid signals. In particular, the α-CH_2 signal at 2.27 ppm is a clearly defined triplet in deuterated peptide **84** as only one acyl chain is protiated and is hence clearly visible by ^{1}H NMR. Conversely, in protiated peptide **79**, the overlap of two α-CH_2 signals results in the multiplet shown in **Fig. 4.12** (inset).

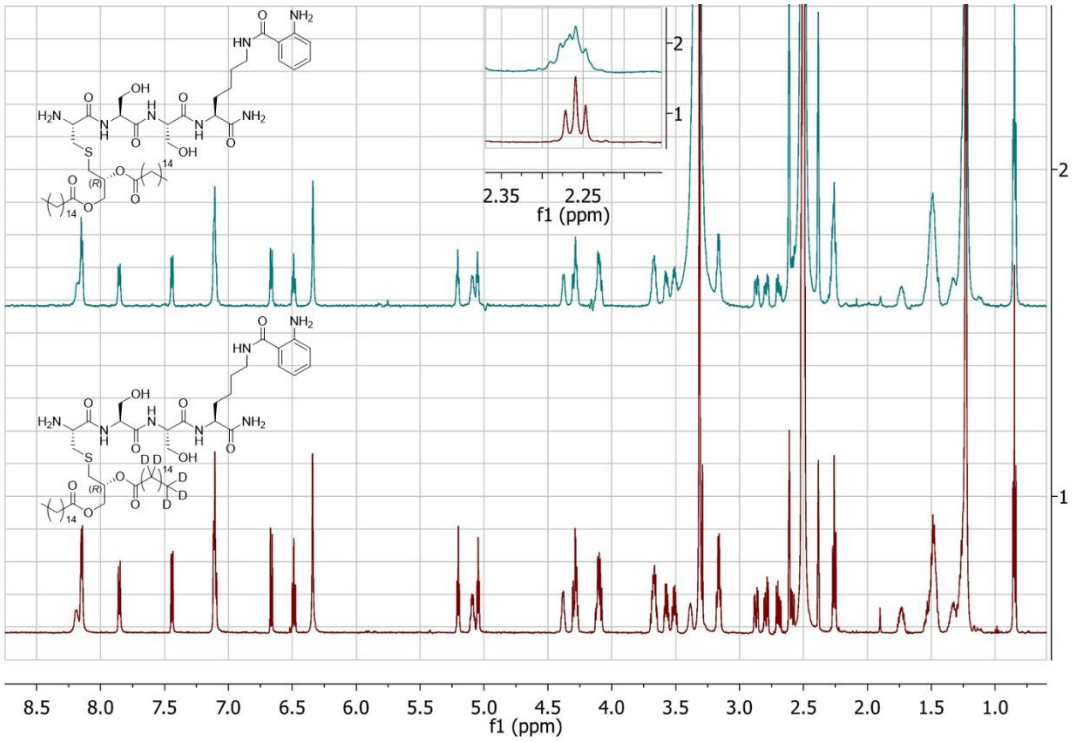


Figure 4.12. ^{1}H NMR (600 MHz, DMSO-d_6) comparison of protiated substrate **79** (blue, top) and deuterated substrate **84** (red, bottom). **Inset**: Expanded view of the region between 2.35 – 2.15 ppm showing the signals arising from the acyl chain αCH_2 groups in each compound.

MS and ^{2}H NMR analysis were used to confirm the presence of deuterium. The limited solubility of this peptide necessitated the use of DMSO-d_6 for NMR analysis. The large excess of DMSO-d_6 results in a broad signal at 4.25 – 1.50 ppm arising from the CD_3 groups of DMSO-d_6 which engulfs the majority of the lipid CD_2 signals. However, signals arising from the peptide are visible at 1.19 and 0.80 ppm as shown in **Fig. 4.13**.

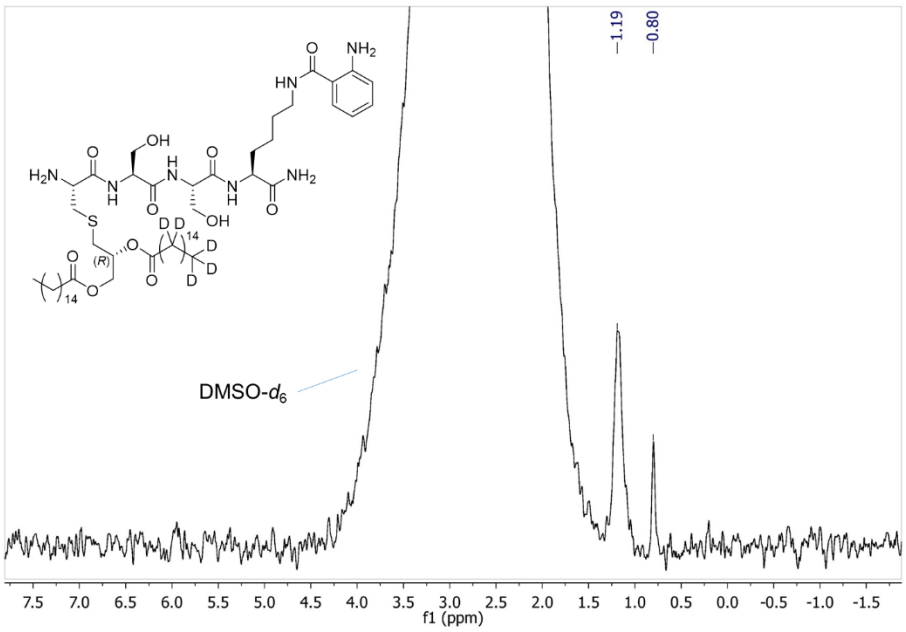


Figure 4.13. Deuterium ^{2}H NMR (400 MHz, DMSO-d_6) of deuterated peptide **84**.

4.5.3 MS Analysis of the Lit Product

As expected, the deuterated peptide **84** was processed by Lit in an comparable manner to the protiated analogue **79**. The reaction time was increased to 16 h to maximise product formation and the reaction was subsequently extracted with $CHCl_3$/EtOH.[p] The product, peptide **85**, was isolated from other lipophilic reaction constituents, including MS incompatible detergents, by preparative TLC, and subjected to ESI MS/MS analysis. A diagnostic result required the fragmentation of the molecular ion between the amide and ester-linked acyl chains. However, following fragmentation, no diagnostic mass signal was obtained. The mass spectra resulting from fragmentation of both the substrate **84** (green, top) and the product **85** (blue, bottom) molecular ions are shown in **Fig. 4.14**. Both substrate and product fragment in an identical fashion. No dehydroalanyl fragment ([M] = 776.67 Da) ion is observed. The problematic fragmentation at the cysteine-serine amide bond ($[M+H]^+$ = 439.22 Da) reported by Armbruster *et al*.[59] was also detected in this case, as shown in **Fig. 4.14**. Armbruster *et al*. overcome this issue in MALDI MS/MS analysis by adjusting the analyte solution pH with aq. $NaHCO_3$, promoting formation of the sodium salt of the peptide. This was not feasible in the current case due to a restriction to MS/MS performed with an ESI source which is incompatible with such non-volatile additives.

[p] Performed by Dr. Dietmar Weichert and Dr. Samir Olatunji, Caffrey Group, Biochemistry, TCD.

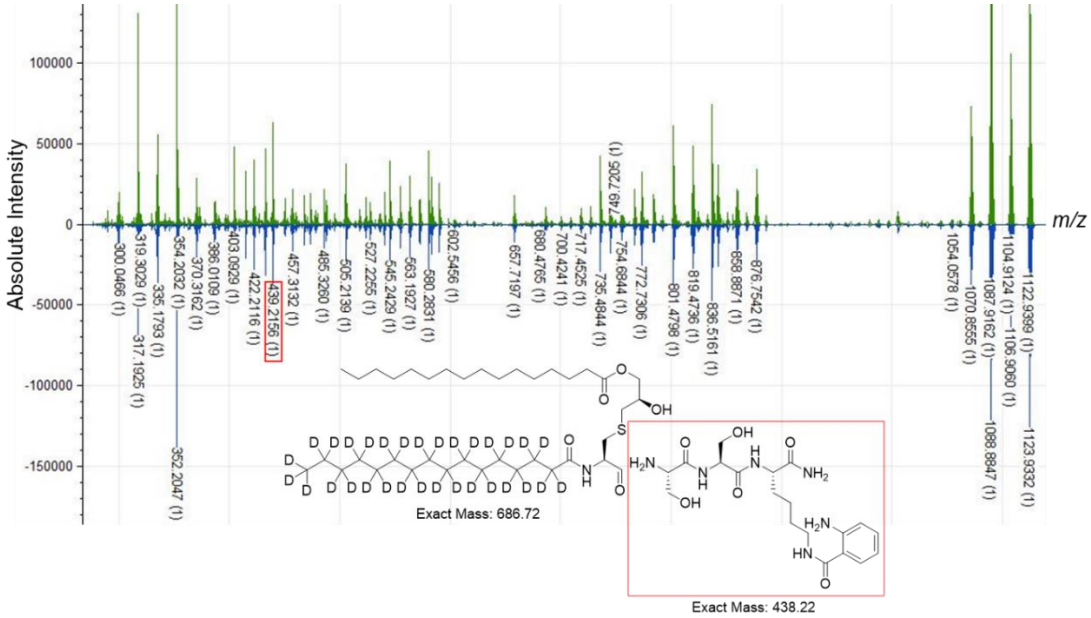


Figure 4.14. The comparison of ESI MS/MS spectra of substrate **84** (top, green) and product **85** (bottom, blue). No difference exists between the identity of peaks present in the two spectra. No monoacylated fragment was observed in either spectrum meaning no structural insight regarding acyl chain position can be gained. Fragmentation of the cysteine-serine amide bond yields a product of mass 438.22 Da. A peak corresponding to $[M+H]^+$ of this fragment is highlighted in red in both the structure and the spectrum. This signal arises in both the substrate and product.

4.5.4 NMR Analysis of the Lit Product

In the absence of conclusive data from MS/MS analysis, structural insights were instead sought from NMR analysis, a technique which was initially discounted due to the higher limit of detection of NMR (600 MHz) in comparison to MS. Hence, NMR analysis required a higher quantity of material. This was achieved by performing 60 parallel Lit reactions, each containing 8.4 µg of substrate **84**.[q] The product was extracted as previously described, isolated by preparative TLC, and then subjected to NMR analysis in DMSO-d_6. A sufficient quantity of material was isolated, allowing for the use of 1D and 2D NMR (^{1}H, HSQC and TOCSY) to analyse the product. This analysis provided unambiguous evidence that the single remaining ester-linked acyl chain in the product resides at the *sn*-1 position. This conclusion is drawn from the analysis of six changes in the ^{1}H spectrum of the product **85** in comparison to the substrate **84**. Each change is highlighted in **Fig. 4.15**.

[q] Performed in collaboration with Dr. Samir Olatunji, Caffrey Group, Biochemistry, TCD.

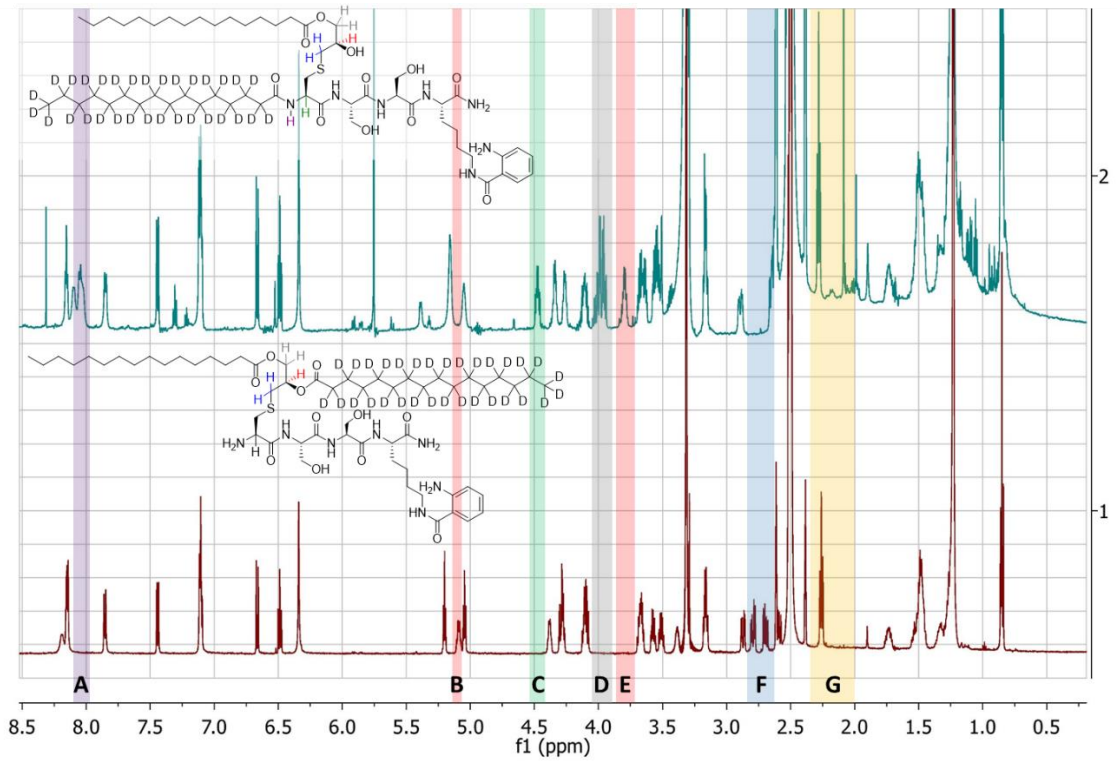


Figure 4.15. ^{1}H NMR (600 MHz, DMSO-d_6) comparison of deuterated product **85** (blue, top) and deuterated substrate **84** (red, bottom). The changes in the spectra are highlighted and colour coded within the structural illustration of the product.

In what follows, each region of change in the spectra is examined to provide evidence which supports this conclusion. Firstly, a new signal is visible in region A (purple) in the product spectrum. This amide NH signal at 8.06 ppm arises from the newly formed amide bond in the product, confirming that the N-terminus has been acylated.

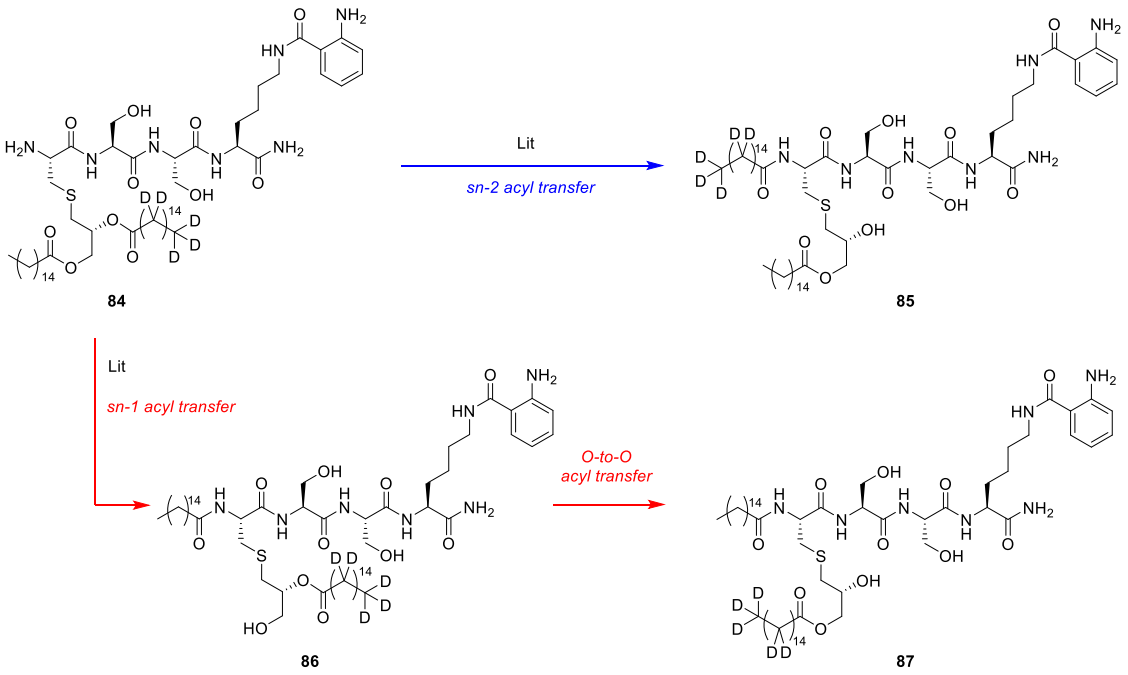
Next, a significant upfield shift of the signal at 5.10 ppm (region B, red) to 3.80 ppm (region E, red) is evident in going from substrate to product. This signal arises from the glyceryl-C<u>H</u> (red) adjacent to the *sn*-2 oxygen. The chemical shift of this signal depends on whether it is in ester linkage to an acyl chain as in the substrate (5.10 ppm) or bound to a hydroxyl group as in the product (3.80 ppm). This large upfield shift is characteristic of the deacylation of an adjacent ester. An upfield shift of this magnitude (1.3 ppm) indicates that the proton is now significantly more shielded, a result of the removal of the electron-withdrawing ester. This confirms the secondary alcohol at the glyceryl-C<u>H</u> group is not acylated in the product.

In region C (green) a new signal is visible at 4.48 ppm in the product. This signal belongs to the α-C<u>H</u> of cysteine (green). In the substrate this signal appears at 3.37 ppm (partially concealed in this instance by residual H_2O). The signal has shifted downfield in the product as the neighbouring NH_2 group has been acylated, withdrawing electron density from the αC-H bond which has the effect of deshielding the proton.

In region D (grey) a small upfield shift of the substrate signals at 4.30 ppm and 4.11 ppm to 3.98 ppm is observed in the product. This signal arises from the glyceryl-O-C<u>H</u>$_2$ (grey) which is slightly more shielded in the product, indicating removal of the *sn*-2 ester (4 bonds away). This upfield shift of 0.32 ppm is not large enough to signify deacylation of the *sn*-1 position, indicating that the adjacent *sn*-1 oxygen remains acylated in the product. Note that this CH_2 group in the substrate appears as two signals, one at 4.30 ppm, the other at 4.11 ppm. The signals in the substrate overlap with the resonances of Ser-αCH (4.30 ppm) and Lys-αCH (4.11 ppm).

In going from substrate to product, a small upfield shift of the signals at 2.80 ppm and 2.70 ppm to 2.63 ppm is seen in region F (blue). These signals belong to the glyceryl-S-C<u>H</u>$_2$ group (blue) which is slightly more shielded in the product than in the substrate, indicating the removal of the *sn*-2 ester (4 bonds away). Note that in the product, this CH_2 signal overlaps with the satellite peak of the residual non-deuterated DMSO NMR solvent which appears as a sharp peak at 2.62 ppm.

The presented NMR analysis supports the hypothesis that the *sn*-1 position remains acylated in the product. However, this does not conclusively prove which acyl chain was transferred as there are two potential pathways to an *sn*-1 acylated product. Both pathways are shown in **Scheme 4.7**. The first pathway (blue arrow) involves direct transfer of the *sn*-2 acyl chain to the N-terminus. The second (red arrows) involves transfer of the *sn*-1 acyl chain to the N-terminus followed by O-to-O acyl transfer of the *sn*-2 acyl chain to the *sn*-1 position, which would proceed over an energetically favourable 6-membered cyclic intermediate (compound **86**). Although intermediate **86** was not detected by TLC throughout the Lit reaction, to conclusively determine which acyl chain is transferred, evidence showing if the amide linked acyl chain is deuterated is required.



Scheme 4.7. The two potential pathways which result in an *sn*-1 acylated product.

This evidence is provided in region G (**Fig. 4.15**; yellow) which shows the resonance of the α-CH_2 group of the single protiated ester chain which is at the *sn*-1 position in product **85**. The sharp singlet at 2.09 ppm in the product arises from residual acetone.[243] The signal at 2.27 ppm has not changed in going from substrate to product. This reveals that these protons experience the same electronic environment in both substrate and product, suggesting the protiated acyl chain has remained ester-linked at the *sn*-1 position. If this acyl chain was transferred by Lit it would become amide-linked in the product (compound **87**), as opposed to ester-linked as in the substrate. In such a case, it would likely exhibit an altered chemical shift due to the reduced electron-withdrawing power of an amide in comparison to an ester. A similar perturbation of the α-CH_2 chemical shift to this effect has been observed previously with related compounds diacylated **66** and triacylated **67**. As shown in **Fig. 4.16**, the α-CH_2 of the *N*-palmitoyl group in triacylated **67** yields a distinct resonance at 2.25 ppm, upfield shifted in comparison to the α-CH_2 signals of the ester-linked acyl chains at 2.31 ppm This confirms that different electronic environments are induced by an amide and ester bond. However, these spectra were recorded in $CDCl_3$. Confirmation that an amide-linked protiated acyl chain would exhibit a similarly shifted resonance in the Lit product and in DMSO-d_6 solution was necessary. To this end, the isolation and analysis of the protiated product from the Lit reaction with **79** was performed.

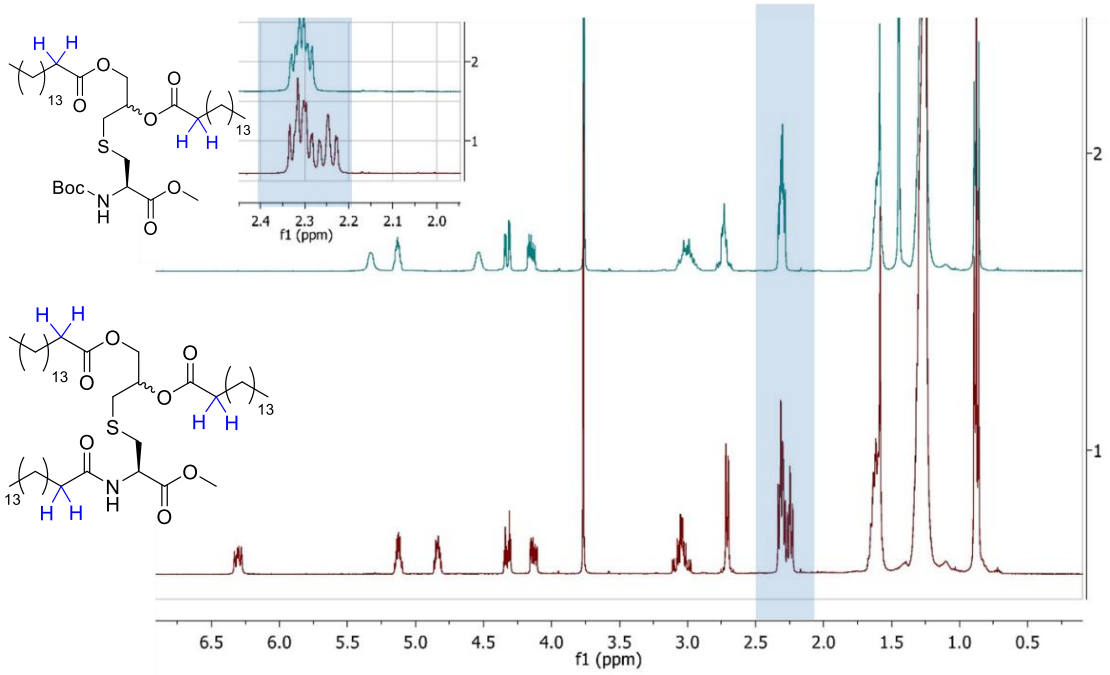


Figure 4.16. ^{1}H NMR (600 MHz, $CDCl_3$) comparison of diacylated **66** (blue, top) and triacylated **67** (red, bottom). An α-CH_2 signal at 2.25 ppm arises from the amide-linked acyl chain in **67**. This signal is shifted upfield in comparison to the ester-linked acyl chain α-CH_2 signals.

The protiated product **88** was produced and isolated from the Lit reaction with peptide **79** in an identical manner to that described for the deuterated product **85**.[r] The shifts observed in the ^{1}H NMR spectrum of the protiated product **88** in comparison to the protiated substrate **79** are shown in **Fig. 4.17**. These changes are consistent with the shifts observed with the deuterated analogue with one key exception. Unlike the deuterated product **85**, the protiated product produces a signal in region G at 2.12 ppm. This signal arises from the α-CH_2 of the amide-linked acyl chain. This reveals that the difference in electronic environment between ester and amide linked acyl chain in the products can be detected by ^{1}H NMR. As the characteristic α-CH_2 signal at 2.12 ppm is not present in the deuterated product **85** it can be concluded that the amide linked acyl chain contains the ^{1}H NMR silent α-CD_2 group and therefore this acyl chain is deuterated. Hence, Lit has catalysed the transfer of the *sn*-2 linked acyl chain to the N-terminus.

[r] Lit reactions were performed by Dr. Dietmar Weichert, Caffrey Group, Biochemistry, TCD.

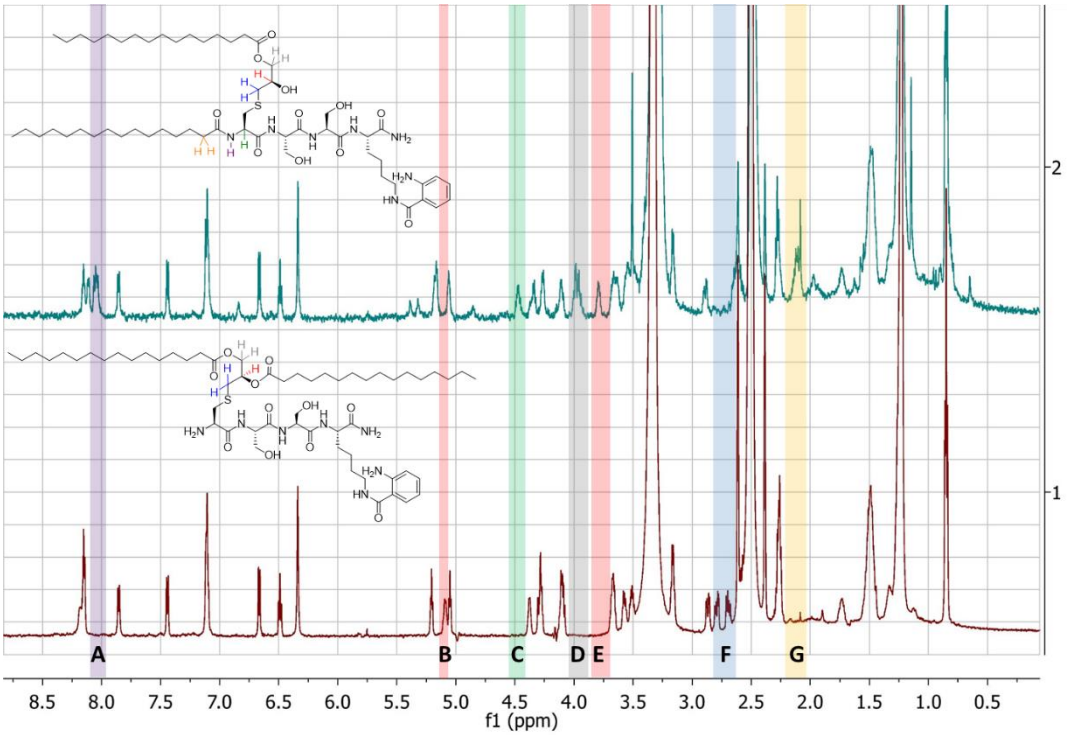


Figure 4.17. ^{1}H NMR (600 MHz, DMSO-d_6) comparison of protiated product **88** (blue, top) and protiated substrate **79** (red, bottom).

The ^{1}H-^{13}C HSQC data for each substrate and product support these conclusions. Notably, the distinction of the acyl chains in the protiated product **88** is clearly defined and is highlighted in **Fig. 4.18**.

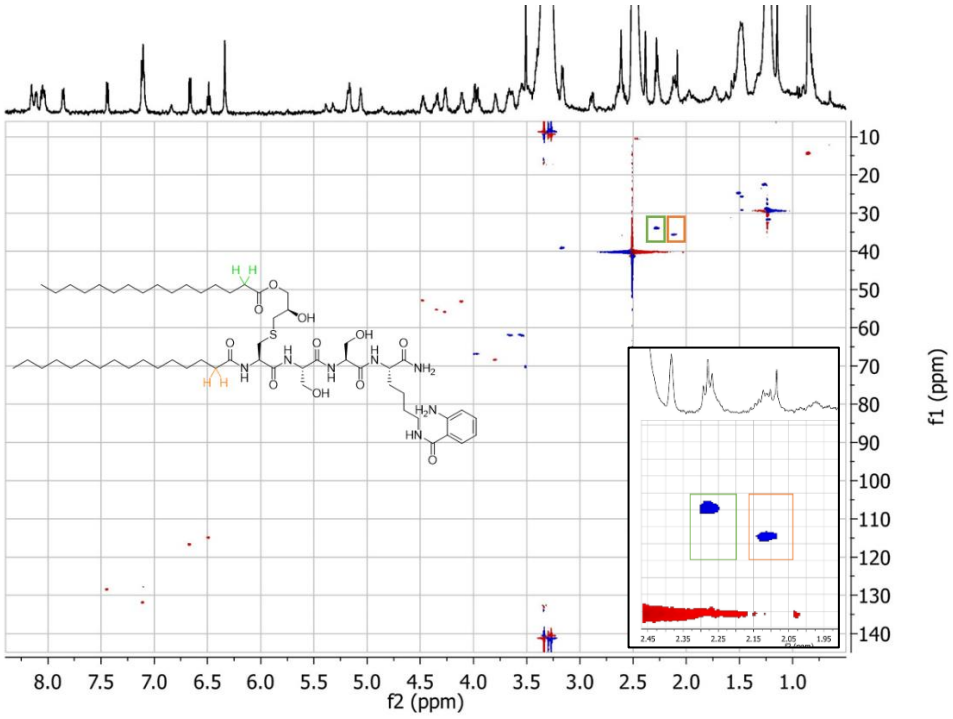


Figure 4.18. HSQC (DMSO-d_6) of protiated product **88**. Inset: Expanded view of the α-CH_2 region.

The characteristic amide-linked acyl chain α-CH_2 signal is not present in the HSQC of the deuterated product **85** as shown in **Fig. 4.19**. This illustrates that the *sn*-1 linked acyl chain has not shifted and that the amide-linked acyl chain in product **85** is deuterated.

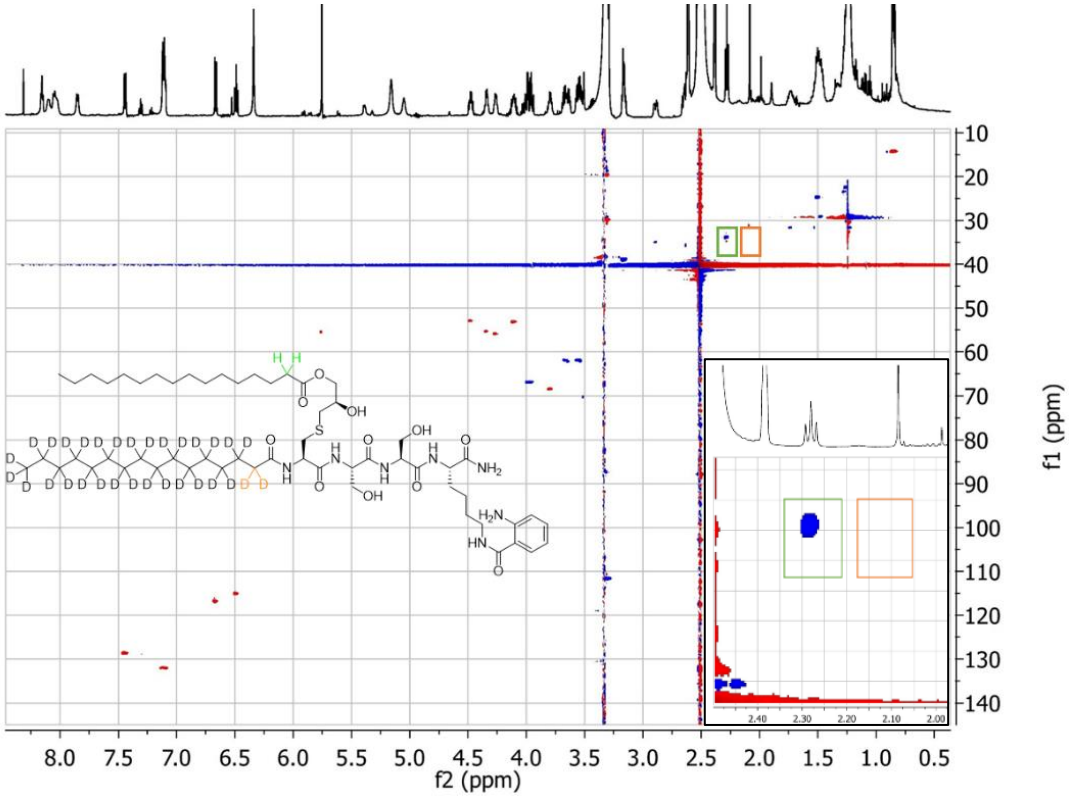


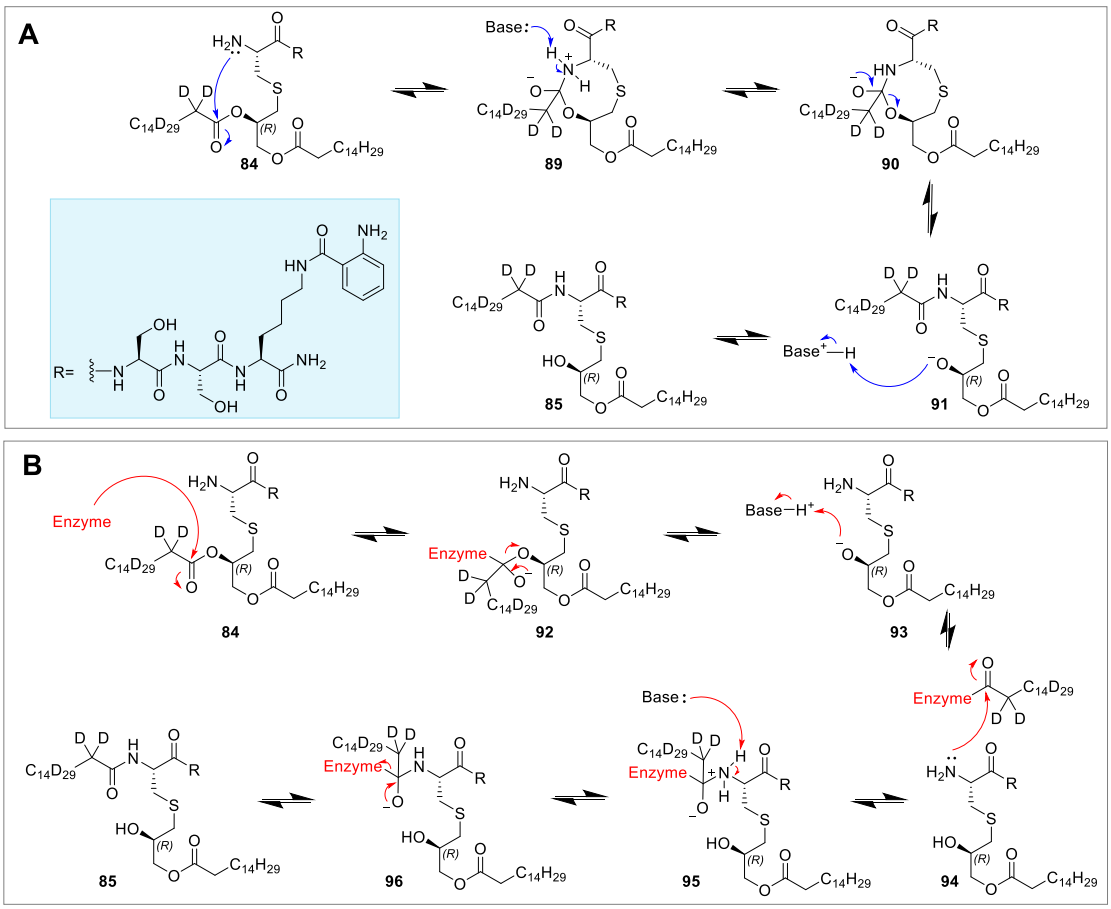
Figure 4.19. HSQC (DMSO-d_6) of deuterated product **85**. **Inset**: Expanded view of the α-CH_2 region.

In summary, we have demonstrated that the α-CH_2 of the amide linked and ester-linked acyl chains exhibit a distinct chemical shift due to the difference in their electronic environments. This difference has been exploited to prove that the amide-linked acyl chain in the Lit product **85** originated from the *sn*-2 position. Supporting evidence involving shifts of five other ^{1}H resonances confirms that the *sn*-2 position has been deacylated.

4.6 Intramolecular *vs.* Intermolecular Acyl Transfer

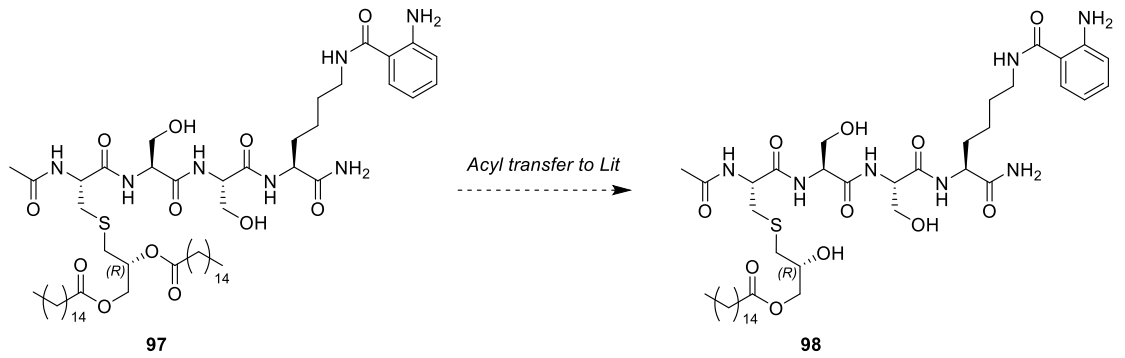
There are two potential mechanisms of Lit-mediated transfer of the *sn*-2 acyl chain. Both possibilities are depicted with substrate **84** in **Scheme 4.8**. Presently, intramolecular transfer is the preferred hypothesis in the literature as it has been deemed unlikely that a protein of Lit's small size (~220 amino acids) is capable of performing the bifunctional roles required to facilitate intermolecular acyl transfer.[59] In contrast to Lit, the

bifunctional Lnt is much larger, containing ~512 amino acids.[181] Beyond this assumption, no direct evidence has been reported in support of either mechanism.



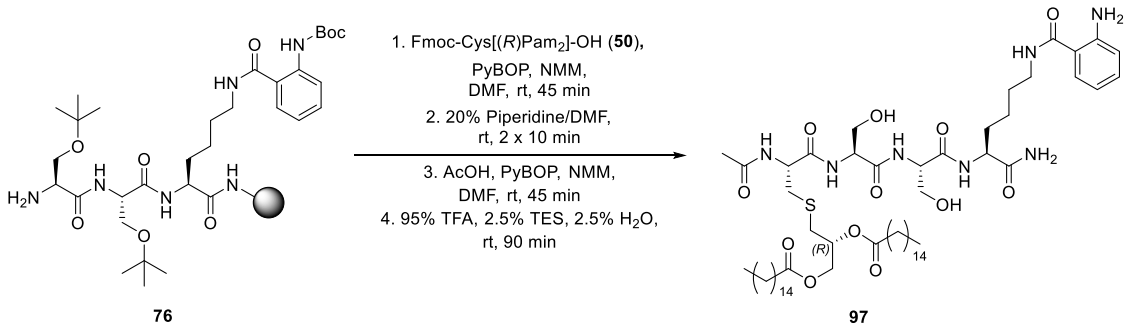
Scheme 4.8. Proposed mechanisms of Lit-mediated acyl transfer of substrate **84**. **A**: Intramolecular acyl transfer. **B**: Intermolecular acyl transfer.

As discussed, intramolecular *O*-to-*N* acyl transfer over an 8 membered cyclic intermediate has previously been reported in the synthetic chemistry literature. These reports suggest that intramolecular transfer is possible, provided that Lit can constrain the substrate into an appropriate conformation with a short geometric separation of amine and ester. To probe the possibility of intermolecular acyl transfer, a potential substrate in which the N-terminus was blocked with an acetyl group (peptide **97**) was synthesised. This study aimed to explore if such a compound was capable of acylating the enzyme. In such a scenario, the *sn*-2 deacylated compound would be produced in detectable quantities with the use of a stoichiometric amount of enzyme. Due to the difference in polarity of these compounds, it was proposed that TLC could be used to monitor the consumption of substrate **97** and formation of deacylated product **98**, the reaction indicated in **Scheme 4.9.**



Scheme 4.9. A potential substrate of Lit which was used to probe potential acylation of Lit and the product which would result if peptide **97** was capable of acting as an acyl source.

Peptide **97** was synthesised *via* SPPS as depicted in **Scheme 4.10**. PyBOP mediated coupling of Fmoc-Cys[(*R*)Pam$_2$]-OH (**50**) to the common resin intermediate **76** was followed by Fmoc deprotection. Acetic acid was subsequently coupled to the N-terminus prior to resin cleavage. Resin cleavage and global deprotection yielded the crude peptide which was purified by normal phase column chromatography to give peptide **97** in an overall yield of 11%



Scheme 4.10. Synthesis of peptide **97**.

Peptide **97** was treated with Lit and the reaction was extracted and analysed by TLC as previously described.[s] No change in R_f was detected by TLC, indicating that this compound is not processed by the enzyme. Hence, this study which had the potential to provide evidence in support of an intermolecular transfer did not do so. However, although this compound is not capable of acylating the enzyme, limited conclusions can be drawn. The lack of deacylation of peptide **97** does not rule out an intermolecular transfer mechanism. The *N*-terminal structure of the proposed substrate **97** closely resembles that of the lyso product of the Lit reaction. *N*-acylated compounds, such as

[s] Lit reaction and TLC analysis was performed by Dr. Samir Olatunji, Caffrey Group, Biochemistry, TCD.

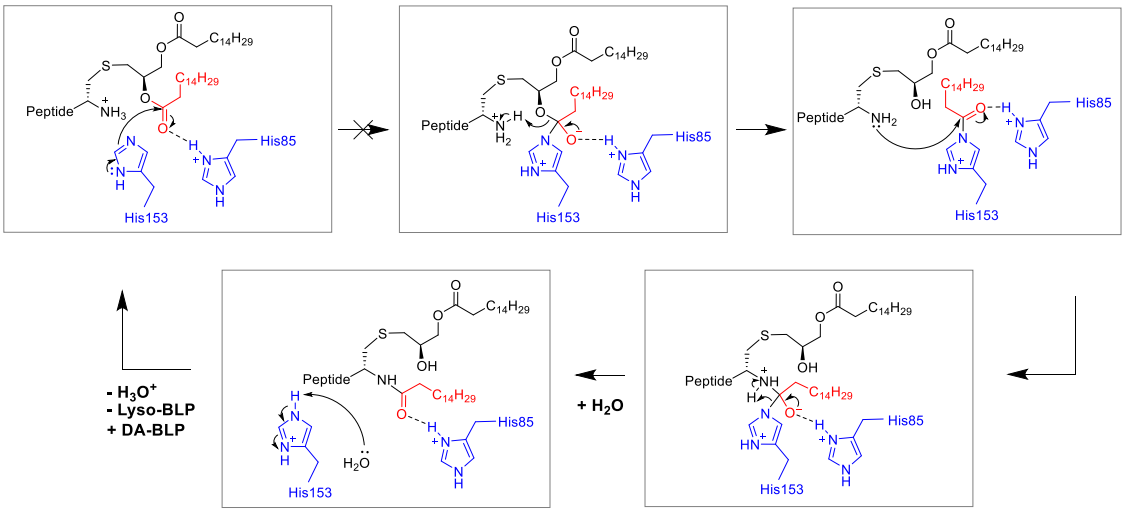
lyso-form LPs, are conceivably ejected from the active site by a conformational change of the enzyme, preparing the active site for another round of catalysis. In these circumstances, despite the short (C_2) *N*-acyl chain of peptide **97**, enzyme confirmation may prohibit entry of the *N*-acylated **97** to the active site, concealing any mechanistic information.

4.6.1 *In Silico* Modelling[t]

Due to the challenge of probing inter- *vs.* intra-molecular acyl transfer mechanisms experimentally, quantum mechanics/molecular mechanics (QM/MM) studies were utilised to provide mechanistic insights.[s] Such studies required the identification of residues potentially involved in catalysis. Highly conserved residues were identified through sequence analysis and include two polar residues: His153 and His 85. Use of the TLC activity assay, coupled with site-directed mutagenesis revealed His85, His153 and Phe157 are essential for activity.[s] This suggests acyl transfer is potentially mediated by a catalytic histidine dyad involving His85 and His153. Histidine residues generally exhibit a pK_a of ~ 6.[244] The environment of the active site can modulate this, giving histidine a pK_a close to neutral. Consequently, this residue is capable of acting as both an acid and base and hence it is commonly employed by nature in a catalytic role. Despite the ubiquity of histidine in enzymatic catalysis, histidine dyads are rare and furthermore histidine does not commonly act as a nucleophile.[244]

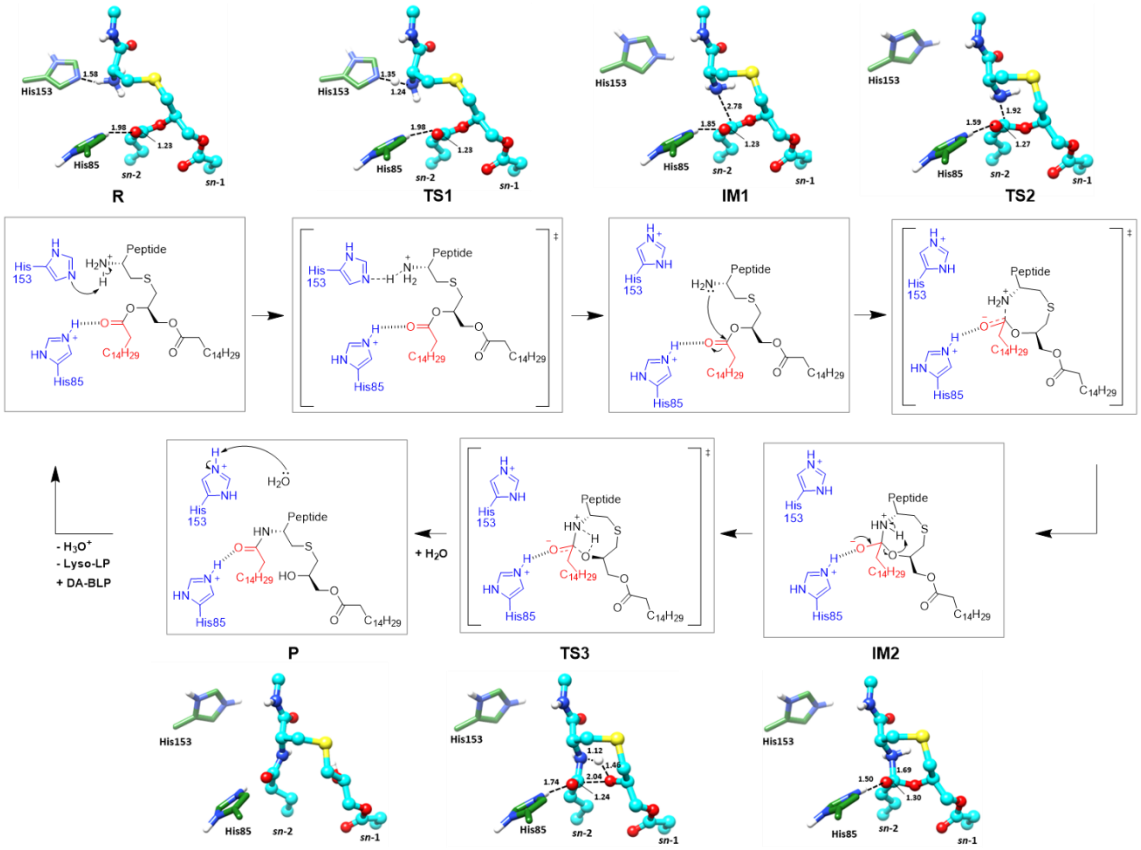
Assuming a catalytic histidine dyad, intermolecular acyl transfer would involve nucleophilic attack of a carbonyl by a histidine residue as shown in **Scheme 4.11**. However, acyl-histidine intermediates are not common.[244, 245] QM/MM studies using the peptide sequence NH_3^+-Cys[(*R*)-Pam_2]-Ser-Ser-Lys-$CONH_2$ as the substrate concluded that this intermolecular mechanism is not energetically feasible. The modelled approach of the His153 ε-nitrogen towards the *sn*-2 carbonyl during nucleophilic attack resulted in increasing potential energy as the distance decreased. No intermediate nor corresponding transition state was identified.

[t] Site-directed mutagenesis was performed by Dr. Samir Olatunji.
Computational work was performed by Dr. Warispreet Singh and Dr. Irena Tikhonova, School of Pharmacy, QUB.



Scheme 4.11. Intermolecular acyl transfer mediated by the two histidine residues of the proposed catalytic histidine dyad. QM/MM revealed that nucleophilic attack by His153 was not energetically feasible.

The alternative mechanism shown in **Scheme 4.12** involves intramolecular acyl transfer in which the histidine residues act in a general acid-base capacity. The predicted conformation of the peptide throughout the reaction is depicted in the outer panels of **Scheme 4.12**. As shown in intermediate 1 (IM1), in the binding pocket of Lit the peptide adopts a predicted conformation in which the amine and *sn*-2 carbonyl carbon are proximal to each other, separated by just 2.78 Å. As such, the energy barrier to formation of the 8-membered cyclic transition state (TS2) is low and is depicted in the reaction profile shown in **Fig. 4.20**. QM/MM calculations predict this overall mechanism has a thermodynamically feasible energy barrier of 21.7 kcal/mol. The corresponding predicted rate constant of the reaction would therefore be 0.19 min^{-1}.[s] This is comparable to the observed experimental turnover number of 0.25 min^{-1} for substrate **79**. Therefore, it is feasible that Lit catalysed acyl transfer proceeds intramolecularly *via* an 8-memebered cyclic intermediate.



Scheme 4.12. A proposed mechanism of intramolecular acyl transfer. His85 remains protonated throughout and exhibits a hydrogen bonding interaction that activates the *sn*-2 carbonyl (red) towards nucleophilic attack. His153 functions in the first instance as a base, deprotonating the N-terminal ammonium group of the substrate to yield a nucleophilic amine. Nucleophilic attack yields the 8-membered cyclic intermediate (IM2). The triggering of tetrahedral intermediate collapse of IM2 is shown here by intramolecular proton transfer *via* transition state 3 (TS3) to yield the *N*-acylated product (P) which then exists the binding pocket and deprotonation of His153 by water regenerates the enzyme for another round of catalysis. Inner panels: histidine residues are shown in blue, the *sn*-2 acyl chain that undergoes acyl transfer is shown in red, the modelled peptide (NH_3^+-Cys[(*R*)-Pam_2]-Ser-Ser-Lys-$CONH_2$) is shown in black. Outer panel: 3D models of the peptide in the putative binding pocket of Lit. Histidine residues are shown with green carbon atoms, the peptide is shown with blue carbon atoms. Distances between relevant atoms are shown in Å.

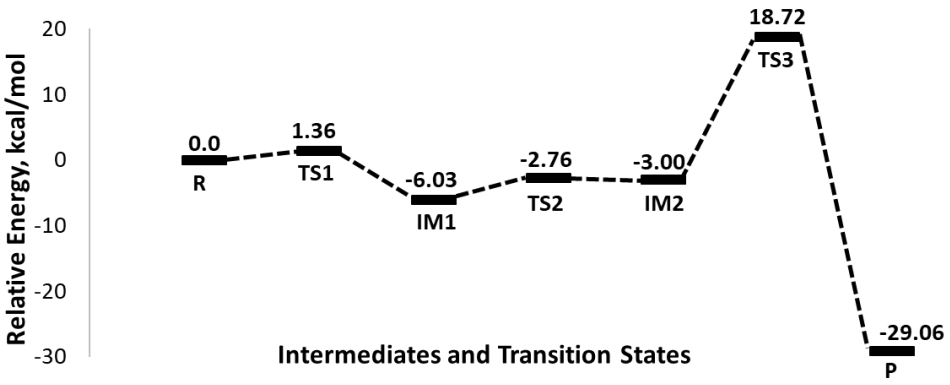


Figure 4.20.[s] Energy profile of the proposed Lit reaction shown in **Scheme 4.12** which proceeds through an 8-membered cyclic intermediate (IM2).

4.7 The Affinity of Diacylated LPs for Cu(II)

Finally, we aimed to explore how Lit activity impacts the interaction of a LP mimetic with Cu(II). LPs are involved in multiple roles within the Gram-positive bacterial cell, including nutrient uptake.[246] Armbruster *et al.*[38] have hypothesised that in the lyso-form LPs may have a reduced affinity for Cu(II). In this case, Cu(II) uptake into the cell by LPs involved in nutrient uptake may be reduced, protecting the cell from the toxic effects of the metal ion.

Paramagnetic ions, such as Cu(II), contain an unpaired electron which can broaden and induce shifts in NMR resonances of nuclei which are in close proximity to the metal ion.[247] Peak broadening in the presence of a paramagnetic ion is the result of accelerated nuclear spin relaxation. This paramagnetic relaxation enhancement is a distance dependent phenomenon arising from the interaction of the nuclei with the electron spin of a paramagnetic ion.[248] Relaxation enhancement effects can be observed up to 20 Å away from the electronic spin with bleaching of the signal occurring at distances of less than 5-10 Å.[249] Consequently, NMR spectroscopy can be used to detect Cu(II) binding at specific regions of a protein or peptide.[250] This property of Cu(II) was used here to investigate the possibility that copper binds at the N-terminus of diacylated LPs. To this end, the effect of Cu(II) addition on the ^{1}H and ^{13}C NMR resonances of Lit substrate **84** was examined. The addition of $CuSO_4$ (0.1 equivalents) to peptide **84** in DMSO-d_6 resulted in selective quenching of five ^{13}C signals of **84**, as shown in **Fig. 4.21**. The affected signals belong to residues from the N-terminal end of the peptide: Cys^{+1}, Ser^{+2} and the glyceryl core.

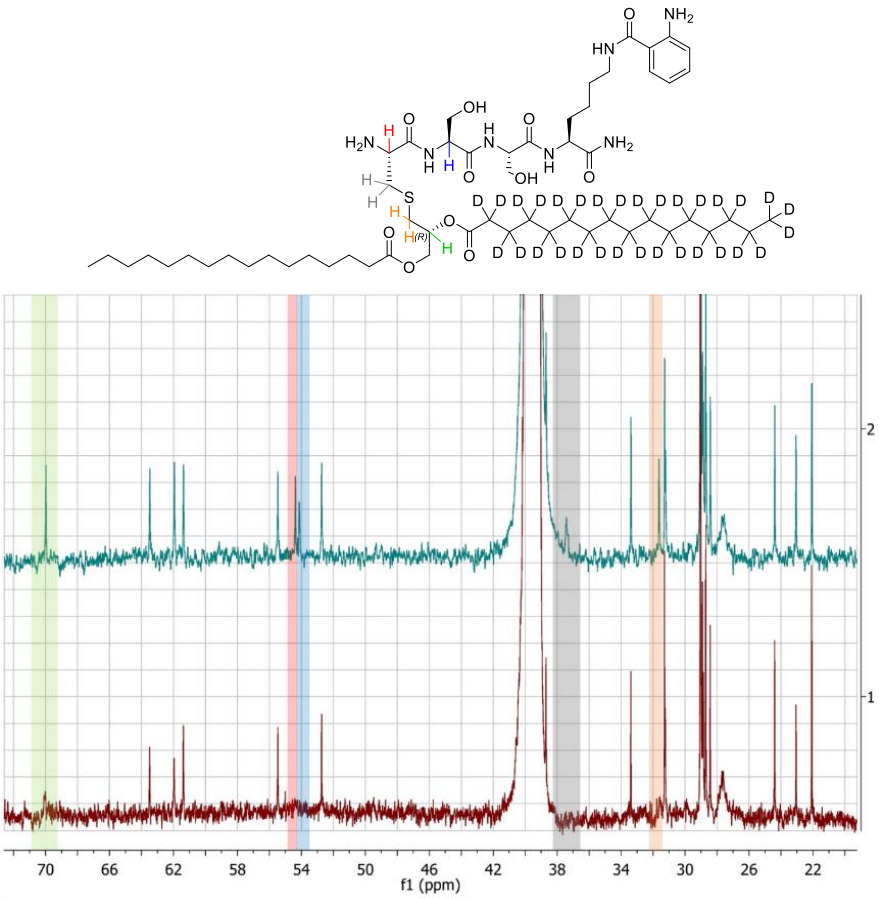


Figure 4.21. Expanded view of the ^{13}C NMR (151 MHz, DMSO-d_6) comparison of peptide **84** in the absence (blue, top) and presence (red, bottom) of Cu(II). In the chemical structure the hydrogen atom which neighbours the affected carbon atom is coloured to highlight the identidy of the perturbed carbon signals.

The comparison of the ^{1}H NMR spectra of peptide **84** before and after Cu(II) addition is shown in **Fig. 4.22**. As with ^{13}C NMR, quenching of N-terminal resonances is observed. Affected signals are highlighted in blue. These changes are reaffirmed in the ^{1}H-^{13}C HSQC spectrum (**Fig. 4.23**) which shows the absence of signals from N-terminal residues confirming a selective interaction of Cu(II) at the N-terminal region of the lipopeptide.

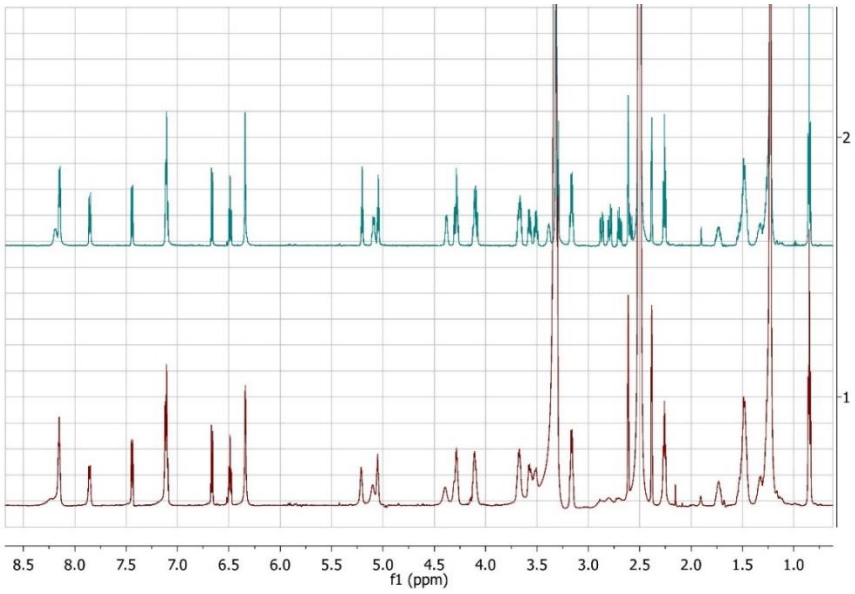


Figure 4.22. ^{1}H NMR (600 MHz, DMSO-d_6) comparison of deuterated peptide **84** in the absence of Cu(II) (blue, top) and deuterated peptide **84** in the presence of Cu(II) (red, bottom).

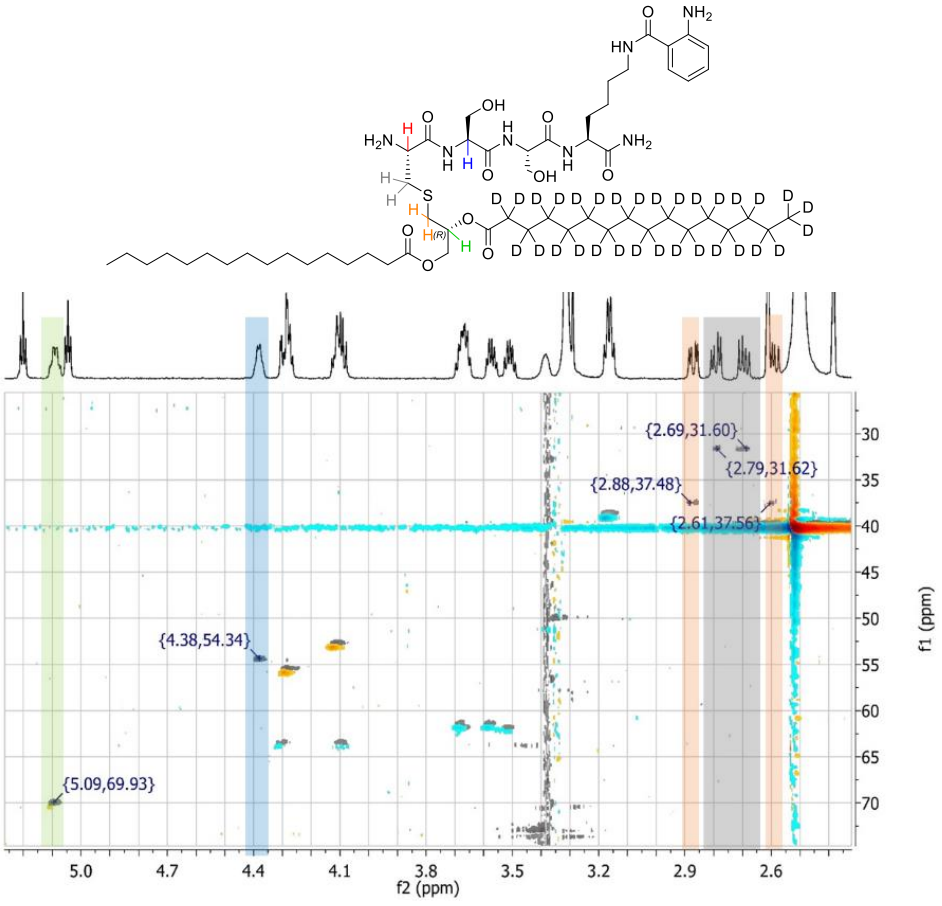


Figure. 4.23. Expanded view of the superimposition of the HSQC (DMSO-d_6) of peptide **84** in the absence of Cu(II) (greyscale cross-peaks) and in the presence of Cu(II) (blue and yellow cross-peaks). The horizontal trace shows the ^{1}H spectrum in the absence of Cu(II) for reference. The signals which are perturbated or quenched in the presence of Cu(II) are highlighted in colour.

Although this experiment holds little physiological relevance due to the non aq. conditions employed, the results interestingly indicate a selective impact of $CuSO_4$ addition on the N-terminal end of peptide **84** in DMSO-d_6 solution. Attempts to perform NMR analysis of **84** in aq. solution, using 5% (v/v) D_2O in H_2O, or up to 20% (v/v) DMSO-d_6 in H_2O, were unsuccessful due to the insolubility of the peptide under such conditions. Resultantly, to gain physiological insight into the possibility of diacylated LPs binding Cu(II) a water soluble diacylated lipopeptide is required. It is possible that aq. solubility could be achieved *via* incorporation of a Lys_4 tag at the C-terminus of the peptide. Furthermore, isolation or synthesis of the corresponding lyso-form product is necessary for comparison. The preliminary results presented here highlight the potential of paramagnetic NMR to reveal the copper binding capacity of diacylated and lyso lipopeptides and warrant continued exploration of this effect.

4.8 Conclusions and Future Work

In conclusion, a novel diacylated lipopeptide substrate of Lit has been identified. The synthesis, purification, and characterisation of this substrate and several analogues has been established. The peptide substrate **79** is efficiently converted to a lyso lipopeptide (**84**) by Lit, with this reaction reaching ~90% completion within 4 h. The enzyme turnover number with this substrate of 0.25 min^{-1} matches well to the computationally predicted rate constant of 0.19 min^{-1}, calculated based on the energy barrier to a proposed, thermodynamically feasible, intramolecular process. Furthermore, facile monitoring of the reaction by TLC has allowed for the functional analysis of Lit under a multitude of conditions. Following the identification of optimal conditions, the development of this new functional assay for Lit has enabled the screening of Lit proteins with site-selective mutations. Through this process, several highly conserved residues which are essential for enzymatic activity have been identified. Supported by the crystal structure of the protein, this has led to the classification of a putative active site involving a catalytic histidine-histidine dyad.

Additionally, the product of the Lit has been isolated from the Lit reaction and characterised by NMR spectroscopy. This has unambiguously confirmed that Lit selectively transfers the *sn*-2 acyl chain of a diacylated lipopeptide to the N-terminus. In 2020, prior to publication of this study, Armbruster *et al.*[251] reached a similar conclusion through the use of a recombinantly produced diacylated LP containing a deuterated acyl

chain at the *sn*-2 position. Upon treatment with Lit, this protein substrate underwent acyl transfer to yield a lyso-product which was shown by MS/MS analysis to contain a deuterated *N*-linked acyl chain at the N-terminus. This reported study was performed with a near-native substrate of Lit which is a highly accurate biomimetic. Gratifyingly, both of these studies provide complementary evidence which supports lyso-form LP production *via* Lit-mediated transfer of the *sn*-2 acyl chain. The analogous conclusions drawn in each case supports the physiological relevance of the work reported here, suggesting the short, non-native peptide **79** behaves with Lit in a similar manner to a native LP substrate.

An investigation into the hypothesis by Armbruster *et al.*[38] that lyso LPs may have reduced affinity for Cu(II) and thus that their production may result in avoidance of copper toxicity by the microbe has been initiated. Using paramagnetic NMR spectroscopy we have demonstrated that Cu(II) selectively interacts with the N-terminus of a diacylated lipopeptide in DMSO. Due to the challenge of extracting a sufficient quantity of the lyso-products **85** and **88** from the Lit reaction, it will likely be necessary to synthesise the lyso-product to further explore this hypothesis. In the first instance, this will allow for a comparable qualitative paramagnetic NMR study to be performed on the lyso product in DMSO, revealing if a similar effect of quenching of N-terminal resonances occurs with an *N*-acylated peptide. Ultimately, to yield a physiologically relevant result, these experiments should be performed under aq. conditions. This will require the use of a water soluble diacylated- and lyso-lipopeptides. NMR titration with $CuSO_4$ would allow for the quantification of the binding constant of Cu(II) to the peptide. Furthermore, the use of complementary analytical techniques such as monitoring the absorbance of Cu(II) by UV/vis spectroscopy following titration with each of the lipopeptides will allow for quantification of the binding constant of both the diacylated- and lyso-lipopeptide.

Chapter 5

Dehydroalanine in Peptide Lipidation and Ligation

5.0 Dehydroalanine in Peptide Lipidation and Ligation

5.1 Peptide Dagylation *via* Dehydroalanine

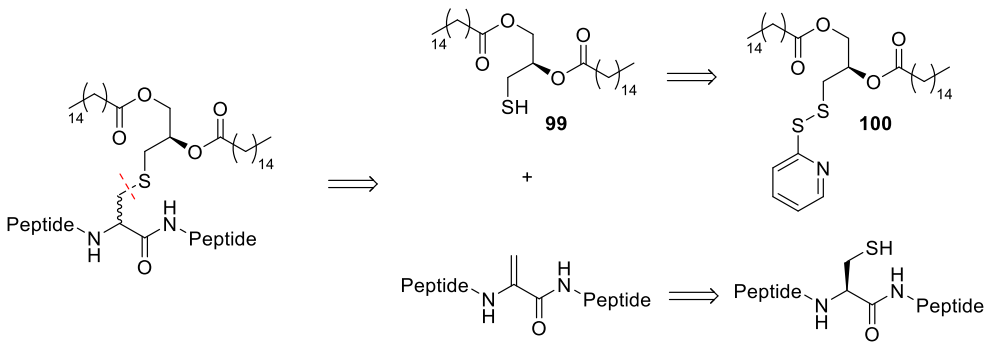
5.1.1 Introduction

As discussed in Chapter 2, the inclusion of a DAG group in a peptide sequence reduces the aq. solubility of the compound. This property restricts effective purification of the peptide by RP-HPLC, ultimately resulting in poor isolated yields and time-consuming purification. We aimed to develop an alternative synthetic route to DAG-functionalised peptides which obviated such limitations. A convergent approach was envisioned, with late-stage introduction of the DAG group by addition to a dehydroalanine (Dha) containing peptide sequence. The functionalisation of Dha is widely used to yield cysteine modifications on proteins and peptides.[252] This α,β-dehydroamino acid can act as a Michael acceptor and undergo 1,4-conjugate addition with thiols. The complex reactivity of Dha arises from the conjugated π-electron system involving the planar sp^2 hybridised alkene and the flanking amide nitrogen and carbonyl groups. The electronic environment of this residue facilitates numerous types of reaction, including both nucleophilic and radical addition,[253, 254] cycloaddition,[255] and transition metal catalysed cross-coupling.[256] The site-selective introduction of native PTMs and non-native analogues into peptides and proteins *via* Dha functionalisation has been demonstrated with protein phosphorylation,[257] glycosylation,[258] and prenylation,[259] among other examples which highlight the biocompatibility of the conjugate addition reaction.[252] A multitude of methods have been developed to introduce Dha into both peptides and proteins through chemical synthesis,[253] mutagenesis,[260] and biosynthetic approaches.[261]

5.1.2 Aims of this work

To our knowledge, the investigation of peptide and protein dagylation through Dha has not previously been reported. We aimed to apply thiol-Michael addition to Dha to yield DAG functionalised peptides, suitable for biological studies. Advantageously, this strategy would first require the SPPS of a non-lipidated, cysteine containing precursor that would exhibit improved aq. solubility in comparison to the DAG modified derivative. Thus, a critical HPLC purification step could be readily performed prior to dehydroalanine formation and lipidation to enable the removal of SPPS side products, including truncated products which often exhibit similar properties to the target compound. With ready access to purified non-lipidated precursors, peptide dagylation *via*

this methodology has the potential to provide access to a wider range of dagylated peptides in a shorter time frame in comparison to linear SPPS. The retrosynthetic analysis shown in **Scheme 5.1** depicts the envisioned strategy towards late-stage lipidation using Dha.



Scheme 5.1. Retrosynthetic analysis towards DAG containing peptides.

One limitation of the use of Dha is the loss of stereochemical control of the cysteine α-position. Nevertheless, we proceeded with this route with the ultimate aim of investigating how the stereochemistry at this position impacts on LspA activity. To this end, we aimed to use this approach to synthesise the LspA FRET probe **101** (**Fig. 5.1**) to facilitate comparison of the activity the peptide obtained from both synthetic strategies. Low or no activity of an epimer may not preclude the biological application of the molecule as such a species may be suitable in the facilitation of protein co-crystallisation.

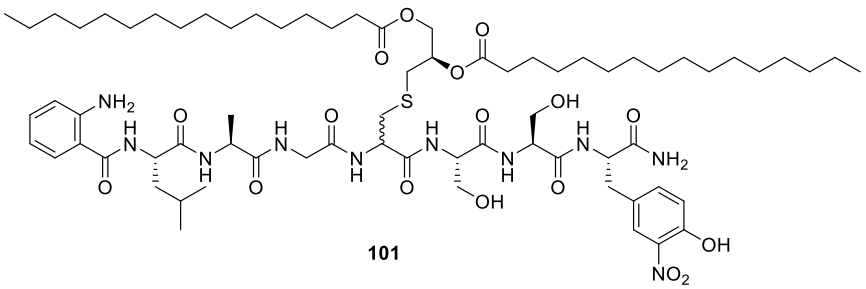
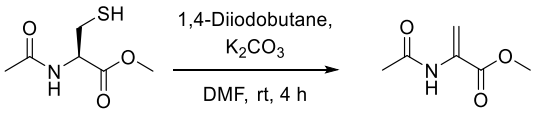


Figure 5.1. Target compound to demonstrate this methodology.

5.1.3 Thiol-Addition to Ac-Dha-OMe (102)

Initial efforts focused on identifying optimal conditions for thiol-Michael addition to a monomeric model acceptor Ac-Dha-OMe (**102**). Compound **102** was synthesised by bis-alkylation and elimination of Ac-Cys-OMe (**103**) with 1,4-diiodobutane and K_2CO_3 as reported by Chalker *et al*.[262] and as shown in **Scheme 5.2.**

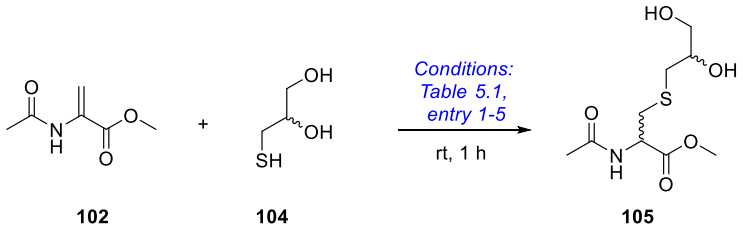


103 **102**, 60%

Scheme 5.2. Synthesis of Ac-Dha-OMe (**102**).

To identify reaction conditions compatible with thiol-Michael addition, the addition of thioglycerol (**104**) to Ac-Dha-OMe (**102**) was screened as depicted in **Table 5.1**. Despite the solubility of both reagents in all solvents screened, no reaction was observed in H_2O, EtOAc and DMF (entries 1-3 respectively) after 1 h, while reaction in THF (entry 4) resulted in decomposition of Dha **102**. To promote thiolate formation, catalytic or stoichiometric equivalents of Et_3N were added to examples performed in organic solvent (entries 2-4), however, this did not promote the desired reaction after 1 h. Based on conditions reported by Wright *et al*.,[253] albeit for radical addition to Dha, the reaction was then attempted in aq. NH_4OAc (1.0 M) solution (pH 6) (entry 5). Despite the low pH of this solvent, the formation of the desired compound was surprisingly observed in aq. NH_4OAc solution after 1 h. After this time, complete consumption of the alkene and formation of a new αCH signal were observed by NMR confirming production of the desired compound which was further established by identification of the product by MS. Following aq. work up the product was isolated in a yield of 84%. Due to the requirement for biocompatible aq. conditions and the fast reaction rate observed, the use of aq. NH_4OAc solution was ideal and subsequently applied to the addition to DAG-functionalised thiol **99** to Dha **102**. This first required synthesis of the required thiol **99**.

Table 5.1. Optimisation of conditions for the thiol-Michael addition of thiol **104** to Dha **102**.



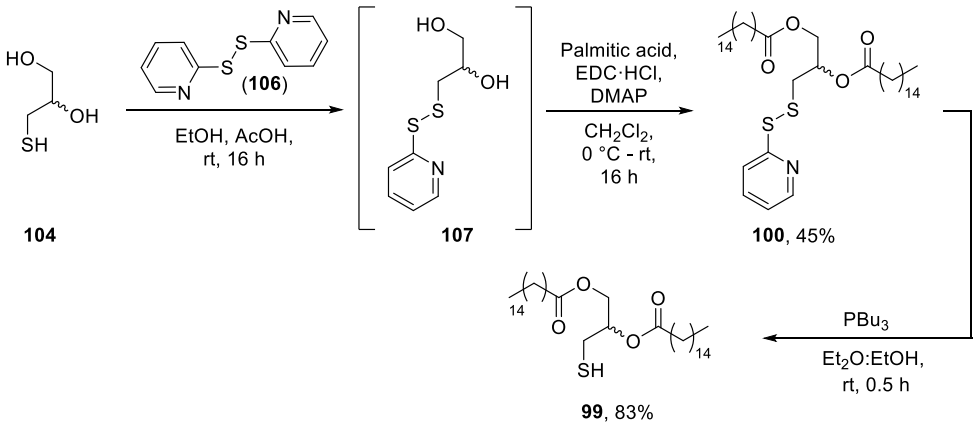
102 **104** **105**

Entry	Solvent	NMR conversion
1	H_2O	No conv.[a]
2	DMF	No conv.[a]
3	EtOAc	No conv.[a]
4	THF	No conv.[b]
5	Aq. NH_4OAc (1.0 M)	quant.

All reactions performed at rt for 1 h using thiol **104** (4 equiv.) and Dha **102** (1 equiv.). Reactions were degassed with argon following solvent addition. [a]Crude ^{1}H NMR and TLC analysis indicated no reaction of **102** or **104**. [b]Crude ^{1}H NMR and TLC analysis show the conversion of the starting acceptor **102** into degradation products.

5.1.4 Synthesis of Dipalmitoylthioglycerol (99)

To access dipalmitoylthioglycerol (**99**), the thiol functionality of thioglycerol (**104**) was protected as a mixed disulfide *via* reaction with the disulfide exchange reagent 2,2'-dipyridyldisulfide (**106**), as illustrated in **Scheme 5.3**. A thiol protecting group which avoided acidic or basic deprotection conditions was selected due to the propensity of the final product **99** to undergo acid and base catalysed O-to-S acyl transfer. During this protection reaction, the consumption of thioglycerol and formation of disulfide **107** was observed by TLC after 16 h. Due to the solubility of the product in aq. conditions hindering the yield of purification, the crude material was used directly in the next step. Following evaporation of the solvent to yield crude disulfide **107**, Steglich esterification with palmitic acid provided the bis-acylated compound **100** in a yield of 45% over two steps following column chromatography. Finally, PBu_3 mediated reduction of the disulfide **100** afforded thiol **99** in a yield of 83%. The solubility of thiol **99** in cyclohexane permitted facile removal of the pyridine-2-thiolate by-product of this reaction by washing with immiscible MeCN. As thioglycerol (**104**) is commercially available only as a racemic mixture, optimisation of this synthetic strategy towards dagylated peptides was performed with the enantiomeric mixture **99**.



Scheme 5.3. Synthesis of racemic mixture **99** from thioglycerol (**104**).

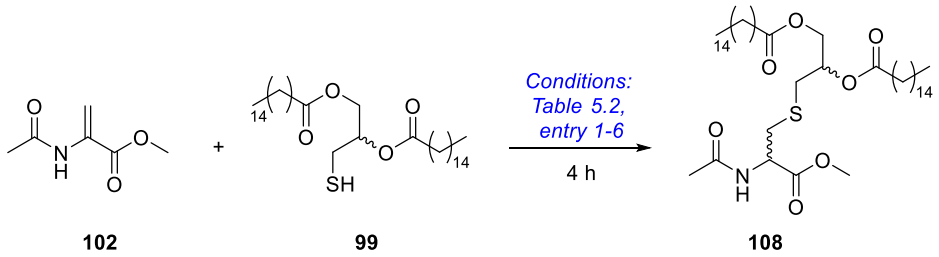
5.1.5 Thiol-Michael Addition of Thiol 99 to Dha 102

With thiol **99** in hand, conditions to promote the addition of this thiol to Dha **102** were investigated (**Table 5.2**). Unlike with the water-soluble thiol **104**, the use of aq. NH_4OAc (1.0 M) solution as solvent, in this case, resulted in no product formation (entry 1). Unsurprisingly, the highly hydrophobic thiol **99** was insoluble under these aq. conditions. Use of organic solvents in which both reagents were soluble resulted in no reaction (EtOAc, entry 2), or decomposition of the Dha containing compound (THF, entry 4). Addition of organic base (Et_3N, 1.0 equiv.) as in entry 3 resulted in decomposition of the thiol, likely arising from base-catalysed O-to-S acyl transfer and subsequent thioester hydrolysis.

This thiol-Michael addition reaction is sensitive to solvent choice, requiring the use of aq. conditions to proceed cleanly. The thiol substrate **99**, however, is unavoidably insoluble under such conditions. Inspired by the work of Boons *et al*.,[158] the phospholipid dodecylphosphocholine (FC-12) was added to the reaction to promote incorporation of the poorly water soluble reagents into liposomes, thereby increasing the local concentration of reagents and facilitating the desired reaction. To drive liposome formation, the concentration of FC-12 used was above the CMC of the detergent (1.1 mM).[263] Inclusion of FC-12 (8.0 mM, 4 equiv.) with Dha **102** (1 equiv.) and thiol **99** (1 equiv.) in an aq. solution of NH_4OAc (1.0 M) afforded a new product observable by TLC after 4 h at 37 °C. The compound was isolated by extraction and column chromatography which afforded the desired conjugate addition product in a yield of 48%. Increasing the equivalents of thiol used to 4.0 equiv. (entry 6) provided compound **108** in an increased yield of 64%. As this reaction does not require the use of base, side reactions involving

base-catalysed intra- or intermolecular acyl transfer of thiol **99**, are unlikely to occur at a rapid rate. However, other side reactions involving the thiol, namely disulfide formation, likely consume some of this reagent accounting for the increase in yield observed with excess thiol. Modulating the rate of such side reactions cannot be readily achieved by adjusting the concentration of thiol **99** due to the influence of liposome sequestration on the local concentration of thiol. With the aim of investigating whether the altered solubility of a peptidic acceptor in comparison to that of monomeric Dha **102**, would impact the rate of conjugate addition, the use of a peptidic acceptor was next attempted

Table 5.2. Optimisation of conditions for the thiol-Michael addition of thiol **99** to Dha **102**.



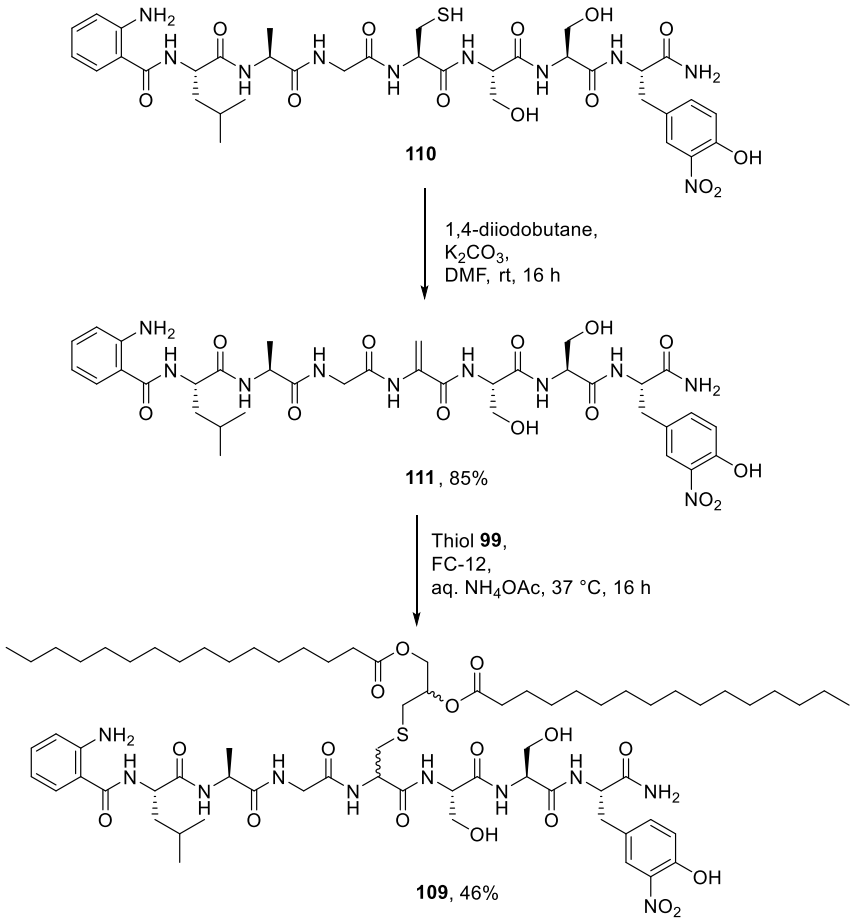
Entry	Solvent	Temp. (°C)	Thiol (equiv.)	Base (equiv.)	Isolated yield
1	Aq. NH_4OAc (1.0 M)	37	1.0	-	No conv.[a]
2	EtOAc	rt	1.0	-	No conv.[a]
3	EtOAc	rt	1.0	Et_3N (1.0)	No conv.[b]
4	THF	rt	1.0	-	No conv.[b]
5	Aq. NH_4OAc (1.0 M) + FC-12	37	1.0	-	48%
6	Aq. NH_4OAc (1.0 M) + FC-12	37	4.0	-	64%

All reactions performed for 4 h using thiol **99** and Dha **102**. Reactions were degassed with argon following solvent addition. [a]Crude ^{1}H NMR and TLC analysis indicated no reaction of **99** or **102**. [b]Crude ^{1}H NMR and TLC analysis show the conversion of the starting materials into degradation products.

5.1.6 Thiol-Michael Addition of Thiol 99 to Peptide 110

This reaction was applied to the synthesis of peptide **109**. The non-lipidated peptide precursor **110** was first produced by SPPS and purified. Transformation of the cysteine residue was performed by bis-alkylation of the thiol and subsequent elimination by 1,4-

diiodobutane and K_2CO_3 in DMF as shown in **Scheme 5.4**. Next, the Dha containing peptide **111** (1 equiv.), thiol **99** (4 equiv.), and the detergent FC-12 were solubilised in EtOH:$CHCl_3$, evaporated to a thin film, and suspended in aq. NH_4OAc (1.0 M) solution. The reaction was monitored by TLC and full consumption of the starting material was observed after 16 h. After this time, excess thiol was extracted with CH_2Cl_2 and the peptide product **109** was isolated from FC-12 by a silica plug. Gratifyingly, the desired product was identified by NMR and MS analysis and afforded in a yield of 46%. This reaction did not proceed in the absence of FC-12.



Scheme 5.4. Synthesis of non-lipidated precursor **110**, conversion to Dha, and functionalisation with DAG *via* thiol-Michael addition to yield peptide **109**.

The rate of reaction was slower with peptidic acceptor **111**, in comparison to the monomer **102**, requiring 16 h to reach completion. This rate decrease may arise from the difference in solubility of these compounds. It is feasible that the amphiphilic nature of the peptide impacts how the compound partitions between the aq. and hydrophobic phases of the liposomal formulation. Coupled with potential secondary structure formation of the peptide, this may reduce the accessibility of the electrophilic site of the

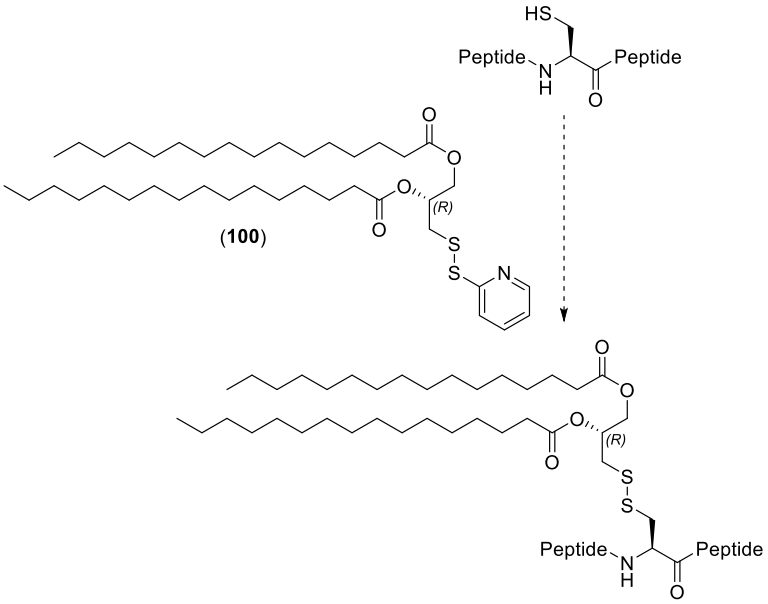
alkenyl peptide. The resulting decrease in reaction rate allows increased time for potential thiol decomposition, including through disulfide formation. Therefore, the use of further equivalents of thiol could result in a further increase in yield. Due to the high molecular weights of larger peptide and protein-based acceptors, the requirement for a large excess of thiol can be readily achieved and is unlikely to be prohibitive.

5.1.7 Conclusions and Future Work

In this work, the addition of dipalmitoylthioglycerol (**99**) to a Dha functionalised peptide has been demonstrated. Additionally, a facile and scalable synthesis of the thiol **99** required for this strategy has been developed. Combined with the well-established literature procedures for the introduction of Dha into peptides and proteins,[262] this work has the potential to facilitate access to a broad range of dagylated peptides and proteins. Investigation of the scope of this reaction with a range of peptidic substrates would enable the effect of amino acid sequence on reaction rate and yield to be elucidated.

A mixture of four diastereomers is obtained from thiol-Michael addition of dipalmitoylthioglycerol (**99**) to peptide **111** due to the use of thiol. Screening of this diastereomeric mixture for LspA activity would not be informative due to the presence of the (*S*)-DAG stereoisomer which is known to be inactive with LspA.[200] Therefore, to determine the effect of cysteine αCH stereochemistry on LspA activity, access to enantiomerically pure dipalmitoylthiolglycerol is first required. As reported by Lo *et al.*,[264] (*R*)-thioglycerol can be synthesised by epoxide ring opening of (*R*)-glycidol by H_2S in the presence of $Ba(OH)_2$. With this compound in hand, the synthesis of enantiomerically pure dipalmitoylthioglycerol could be achieved using the synthetic route established here, facilitating the preparation of (*R*)-DAG functionalised peptides for LspA screening. Efforts to this end are ongoing within the Scanlan lab.

Additionally, the dipalmitoyl disulfide intermediate **100** synthesised in course of this work can be applied to orthogonal peptide dagylation strategies. Mixed disulfides, such as compound **100**, are commonly employed in cysteine specific bioconjugation reactions.[265] With this reagent, the formation of a disulfide-linked DAG group to a peptide or protein, as depicted in **Scheme 5.5**, can be readily realised. Preliminary work towards this goal, in addition to the screening of the resultant disulfide modified peptides for LspA and Lit activity, is ongoing.



Scheme 5.5. The activated disulfide **100** can be applied in the site-selective labelling cysteine containing peptides and proteins.

A related strategy was reported in 2019 by Fuente-Herreruela *et al.*[266] in which 1,2-dipalmitoyl-*sn*-glycero-3-phospho-thio-ethanol was reacted with 2,2'-dipyridyldisulfide (**106**) to form an activated disulfide **112**, shown in **Fig. 5.2**. This disulfide was then reacted with a cysteine containing peptide to introduce a non-native DAG derivative.

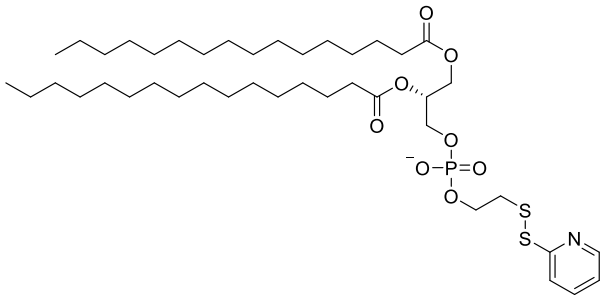
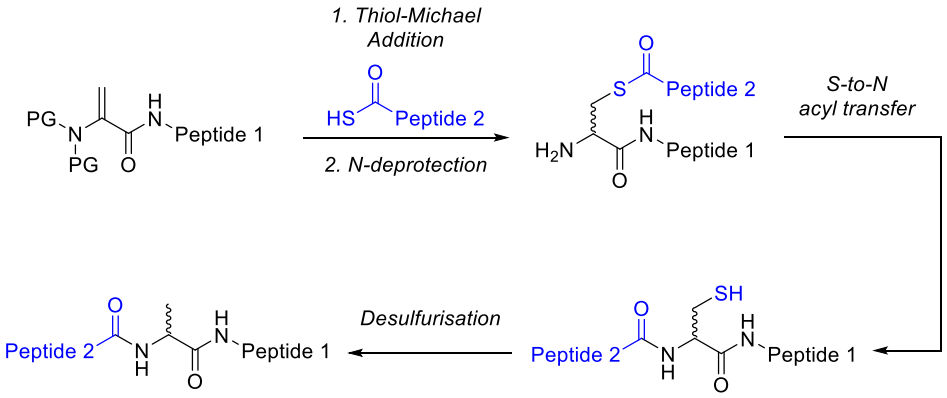


Figure 5.2. Structure of DAG derived mixed disulfide **112** used by Fuente-Herreruela *et al*. in the preparation of lipopeptide bioconjugates.[266]

The phosphate linker involved in this reported work would likely negatively impact recognition by LspA, Lnt and Lit as these enzymes act on the proximal cysteine residue of the lipidated peptide or protein. The charged phosphate group would be required to approach and dock close to the hydrophobic active site of the enzyme which is improbable. Hence, the strategies described in this chapter, which involve native or near-native DAG functionalisation which are more accurate biomimetics offer a distinct advantage in the synthesis of lipopeptides destined for the study of lipoprotein processing enzymes.

5.2 Peptide Ligation *via* Dehydroalanine

The use of dehydroamino acids in peptide ligation is an active area of research within the Scanlan Group. In this approach, acyl thiol-ene chemistry is applied in the synthesis of peptidic thioesters which subsequently undergo intramolecular *S*-to-*N* acyl transfer to yield a native peptide bond between two peptide fragments in a manner analogous to native chemical ligation of peptides developed by Kent *et al*. The structure of the dehydroamino acid used determines the size of the cyclic intermediate through which acyl transfer occurs. Desulfurisation of the resultant thiol containing residue has been demonstrated to yield native residues at the ligation junction. Towards the aim of peptide ligation, the conjugate addition of thioacids to Dha was investigated. In this work, the formation of cysteinyl thioesters through β,γ-*C*,*S* Michael addition of a thioacid to Dha is followed by intramolecular *S*-to-*N* acyl transfer to the N-terminus over an ideal 5-membered intermediate to yield a native amide bond, as illustrated in **Scheme 5.6**.

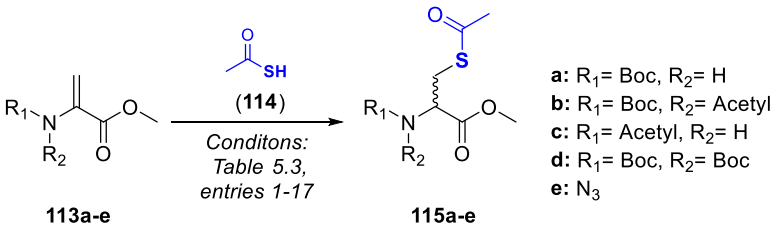


Scheme 5.6. Amide bond formation *via* *S*-to-*N* acyl transfer of a cysteinyl thioester prepared by β,γ-*C*,*S* Michael addition of a thioacid to Dha.

5.2.1 Previous Work Performed in the Scanlan Group

Previous work within the Scanlan Group by Dr. Lauren McSweeney and Dr. Rita Petracca has demonstrated the potential of this methodology with amino acid and dipeptide substrates. Optimisation of this methodology was first performed with Dha derivatives **113a-e** and thioacetic acid (**114**) to determine the influence of Dha protecting groups on the electrophilicity of the Michael acceptor, as shown in **Table 5.3**. High yielding reaction required bis-Boc protection of Dha (entry 13) or monoacetylated Dha (entry 17). As the *S*-to-*N* acyl transfer requires a free N-terminus, the labile Bis-Boc protection route was selected for further optimisation. Based on work by Dr. Petracca and Dr. McSweeney, DMF:phosphate buffer pH 8 was identified as the optimal solvent.

Table 5.3. Optimisation of β,γ-*C*,*S* thiol-Michael addition conditions for the addition of thioacetic acid (**114**) to Dha (**113a-e**)

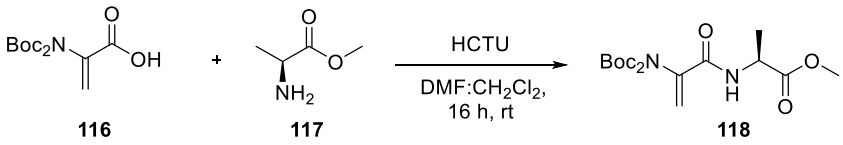


Entry	Acceptor (113)	114 equiv.	Base (equiv.)	Solvent	Time (h)	Temp (°C)	Yield (%)
1	a	4.0	-	$CHCl_3$	12	rt	no conv.[a]
2	a	4.0	K_2CO_3 (4)	$CHCl_3$	24	rt	no conv.[a]
3	a	1.2	Et_3N (0.1)	DMF	96	rt	no conv.[a]
4	a	4.0	Et_3N (4)	H_2O	96	rt	no conv.[a]
5	a	10	-	toluene	16	reflux	no conv.[b]
6	a	4.0	-	DMF: buffer pH 8	4	37	47[d]
7	b	10	-	toluene	16	reflux	50[c]
8	b	20	-	toluene	20	reflux	87[c]
9	b	4.0	Et_3N (0.1)	DMF	16	rt	84[c]
10	b	6.0	Et_3N (6)	H_2O	16	rt	53[c]
11	b	4.0	Et_3N (0.1)	THF	5	rt	89[c]
12	c	4.0	Et_3N (0.1)	DMF	16	rt	no conv.[b]
13	c	4.0	-	DMF: buffer pH 8	4	37	98[d]
14	d	10	Et_3N (0.1)	toluene	16	reflux	90[d]
15	d	4.0	Et_3N (0.1)	THF	5	rt	76[d]
16	e	4.0	Et_3N (0.1)	THF	5	rt	no conv.[b]

17	d	4.0	-	DMF: buffer pH 8	4	37	98[d]

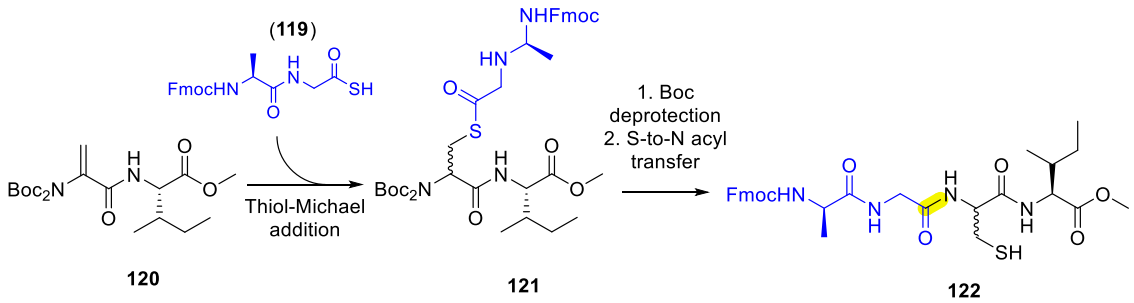
[a]Crude ^{1}H NMR shows only the presence of the starting acceptor **113**. [b]Crude ^{1}H NMR shows the conversion of the remaining starting acceptor **113** into a degradation product. [c] ^{1}H NMR conversion of the purified thioester as an inseparable mixture with starting material. [d]Isolated yield.

Expansion of the scope of this reaction was initially demonstrated through the addition of Fmoc-amino thioacids to **113d**. Moderate to high yields were obtained for a range of polar, apolar and aromatic amino thioacids, including glycine, alanine, valine, isoleucine, leucine, histidine, phenylalanine, and tyrosine.[267] Next, Dha-containing dipeptides were prepared by the *O*-(1*H*-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) mediated coupling of Boc_2-Dha-OH (**116**) to NH_2-Ala-OMe (**117**) as shown in **Scheme 5.7**.



Scheme 5.7. Optimised of coupling conditions for Boc_2-Dha-OH (**116**) and NH_2-Ala-OH (**117**) performed previously in the Scanlan lab.

The previous work culminated in the dipeptide-dipeptide coupling of Fmoc-Ala-Gly-SH (**119**) to Boc_2-Dha-Ile-OMe (**120**) shown in **Scheme 5.8**. Boc deprotection of the ligated product **121** yielded the TFA salt of the amine which was desalted using a basic Amberlyst A21 resin, affording the free amine which underwent *S*-to-*N* acyl transfer to yield the tetrapeptide **122**.



Scheme 5.8. The preparation of a tetrapeptide **122** by dipeptide-dipeptide β,γ-*C*,*S* thiol-Michael addition and *S*-to-*N* acyl transfer. The newly formed amide bond is highlighted in yellow.

5.2.2 Aims of this work

In the current work we aimed to extend the methodology to the ligation of larger peptide fragments. GLP-1 (7-36) was selected as a model peptide to explore the scope of this ligation. This peptide is a product of proglucagon processing, an incretin hormone responsible for stimulating a reduction in blood glucose.[268] Multiple analogues of GLP-1, which act as GLP-1 receptor agonists, are approved treatments for Type 2 diabetes.[269] This 30-mer contains 3 alanine residues which offer 3 potential options for ligation sites which are shown in **Fig. 5.3**.

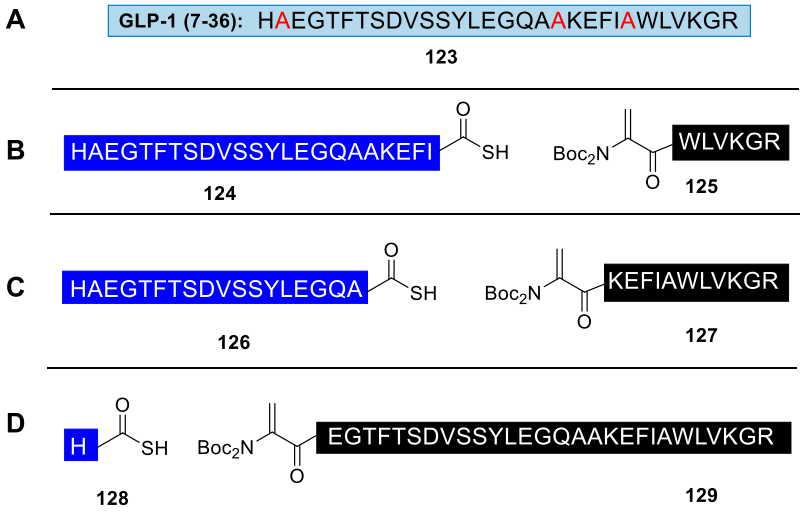
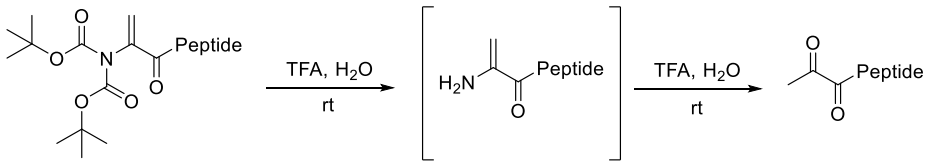


Figure 5.3. GLP-1 (1-36) (**A**) contains three potential ligation junctions depicted in **B**, **C** and **D**.

The ligation site shown in **Fig. 5.3** (**B**) was first investigated. This required a 7-mer peptide functionalised with Dha (**125**) as the acceptor and a 23-mer peptide C-terminal thioacid (**124**). Initial efforts focused on the synthesis of the 7-mer acceptor. Previous work within the group, in addition to computational studies,[267] have shown that *N*-bis-Boc protection of Dha is necessary to impart the required electrophilicity of the Michael acceptor to drive thioester formation. The length of the target peptide necessitated synthesis on solid phase, however, the introduction of bis-Boc-Dha into a peptide *via* SPPS proved problematic as both Fmoc and Boc SPPS strategies require an acidic resin cleavage step and bis-Boc is highly acid labile.[270] Additionally, with a free N-terminal amine, Dha undergoes hydrolysis in aq. acidic conditions to the corresponding pyruvoyl derivative, as depicted in **Scheme 5.9**.[271] This prevents preparation of suitably protected compounds *via* strategies involving reprotection post resin cleavage.

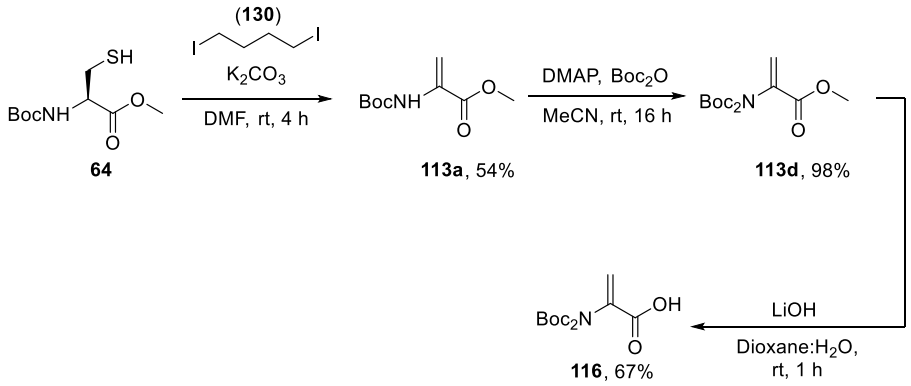


Scheme 5.9. Following N-terminal deprotection, NH_2-Dha functionalised peptides undergo spontaneous tautomerization to the corresponding imine which is followed by imine hydrolysis to yield the pyruvoyl peptide derivative.

It was hypothesised that the bis-Boc protected compound may be obtainable in solution from assembly and subsequent cleavage of the peptide from 2-chlorotrityl chloride resin. Resin cleavage from this linker can be achieved with 0.5-1% (v/v) TFA in CH_2Cl_2.[272] Peptide side chain protecting groups will remain in place under such conditions which may negatively impact aq. and organic solubility of the compound but as ligation had been demonstrated in DMF:H_2O the use of DMF as an organic co-solvent could potentially overcome such issues.

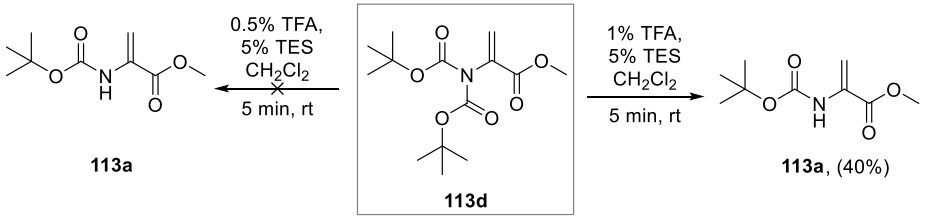
5.2.3 Dha Stability

To investigate the stability of bis-Boc-Dha, compound **116** was first synthesised as described by Chalker *et al*.[262] and as shown in **Scheme 5.10**.



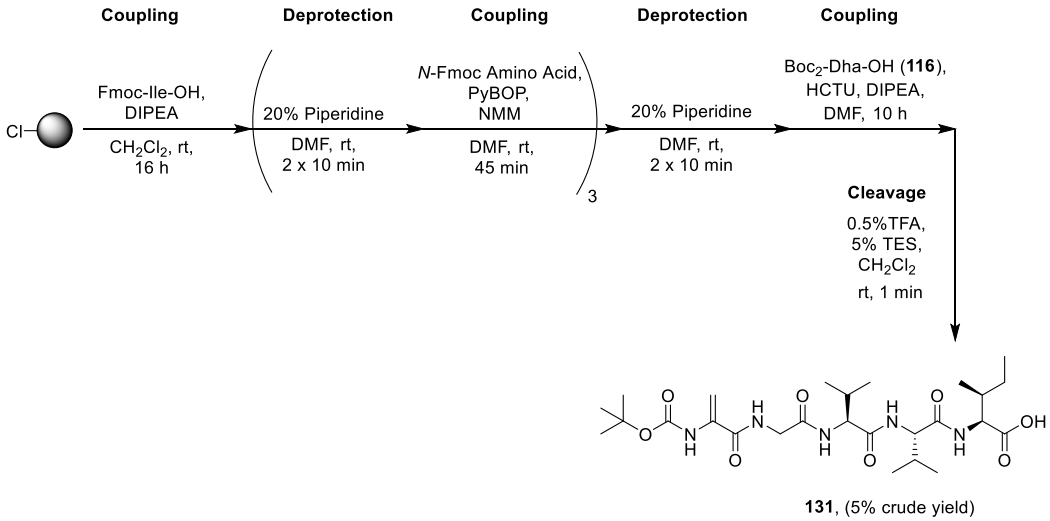
Scheme 5.10. Synthesis of Boc_2-Dha-OH (**116**) according to literature precedent.

To investigate the stability of Boc_2-Dha-OMe (**113d**) under acidic conditions, the compound was treated with 1% TFA, 5% TES in CH_2Cl_2 for 5 min. NMR analysis of the resultant product showed partial deprotection of the second Boc group under such conditions. Promisingly, no deprotection occurred with 0.5% TFA after 5 min, as summarised in **Scheme 5.11**.



Scheme 5.11. The second Boc of Boc_2-Dha-OMe (**113d**) is stable to 0.5% TFA for 5 min, however, this protecting group is partially removed with 1% TFA.

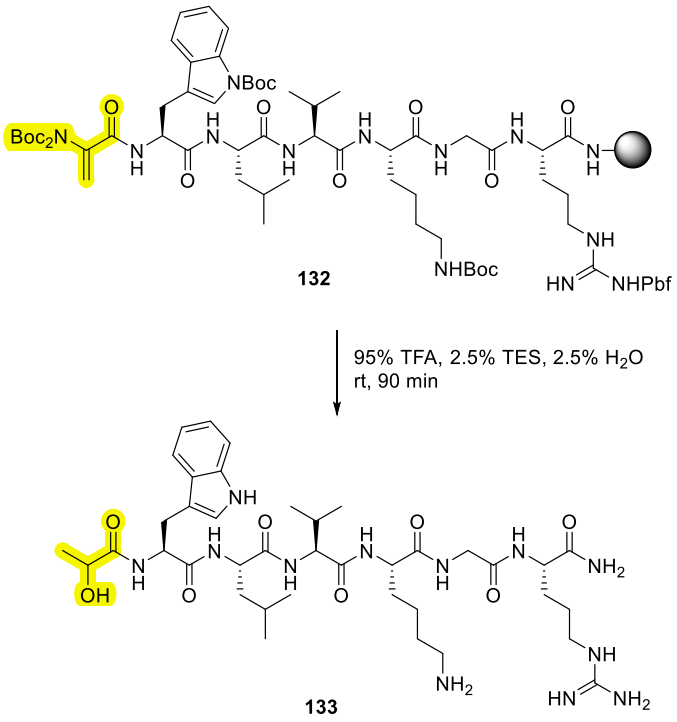
Thus, the synthesis of a bis-Boc protected Dha containing peptide was attempted on 2-chlorotrityl chloride resin. As shown in **Scheme 5.12**, a pentapeptide was assembled on 2-chlorotrityl chloride resin under standard SPPS conditions. Following N-terminal Fmoc deprotection, Boc_2-Dha-OH (**116**) was coupled using HCTU and DIPEA in DMF for 10 h. These conditions were previously optimised in the group for solution-phase coupling of Boc_2-Dha-OH (**116**) as described above. Bromophenol blue indicator was used to confirm the consumption of the amine during coupling. The peptide was subsequently cleaved with 0.5% TFA, 5% TES in CH_2Cl_2 for 1 min, however, the low crude yield of 5% obtained suggests incomplete cleavage from the resin after this time. Additionally, MS analysis revealed that the compound obtained was mono-Boc protected. This shows that the second Boc group has been deprotected during resin cleavage. The difference in acid lability between the trityl-resin linker and the second Boc protecting group of Dha is thus not sufficient to allow for selective cleavage of the peptide while keeping the bis-Boc-Dha protected peptide intact.



Scheme 5.12. Synthesis of a Dha functionalised peptide on resin followed by resin cleavage with 0.5% TFA for 1 min resulted in the deprotection of the second Boc group, yielding a mono-Boc protected peptide **131** which is not suitable for use in thiol-Michael addition reactions with thioacids.

5.2.4 On-Resin Ligation

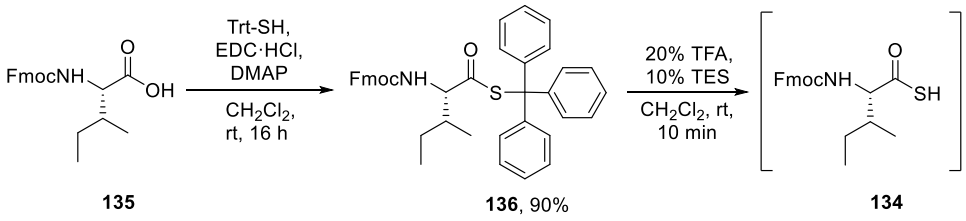
Due to the challenge of obtaining a bis-Boc protected peptide in solution, coupled with the limitations of this protected peptide approach, an alternative on-resin ligation strategy was explored. First, a Rink-amide bound hexapeptide (GLP-1 (31-36)) was synthesised under standard SPPS conditions. Boc_2-Dha-OH (**116**) was coupled to the N-terminus using HCTU as previously described. To confirm Dha coupling, a small amount of peptide resin **132** was cleaved using 95% TFA, 2.5% TES and 2.5% H_2O and analysed by MS. This revealed that the N-terminus of the desired peptide was functionalised with an α-hydroxy-α-methyl group, as shown in **Scheme 5.13**, not the expected pyruvoyl derivative. It is feasible that this product results from TES mediated reduction of the pyruvoyl terminal carbonyl to the corresponding alcohol. Reduction of carbonyl containing compounds in acidic media by TES has previously been reported by West *et. al*..[273] Nevertheless, the identification of the compound indirectly confirms that Dha coupling to the resin-bound peptide was successful.



Scheme 5.13. Under acidic conditions with a free N-terminus, Dha hydrolyses to a pyruvoyl group which then undergoes TES mediated reduction to the alcohol functionality highlighted in peptide **133**.

As a trial ligation, the addition of Fmoc-Ile-COSH (**134**) to the resin-bound **132** was attempted. Triphenylmethanethiol (Trt-SH) coupling to Fmoc-Ile-OH (**135**) yielded the

thioacid precursor **136**, protected as a thioester. Trityl deprotection was performed by 20% (v/v) TFA, 10% (v/v) TES in CH_2Cl_2 to yield the thioacid **134** as shown in **Scheme 5.14**. Due to the propensity of thioacids to hydrolyse under atmospheric conditions, thioacid **134** was dried under high vacuum to remove traces of TFA and TES and then used immediately in the ligation.



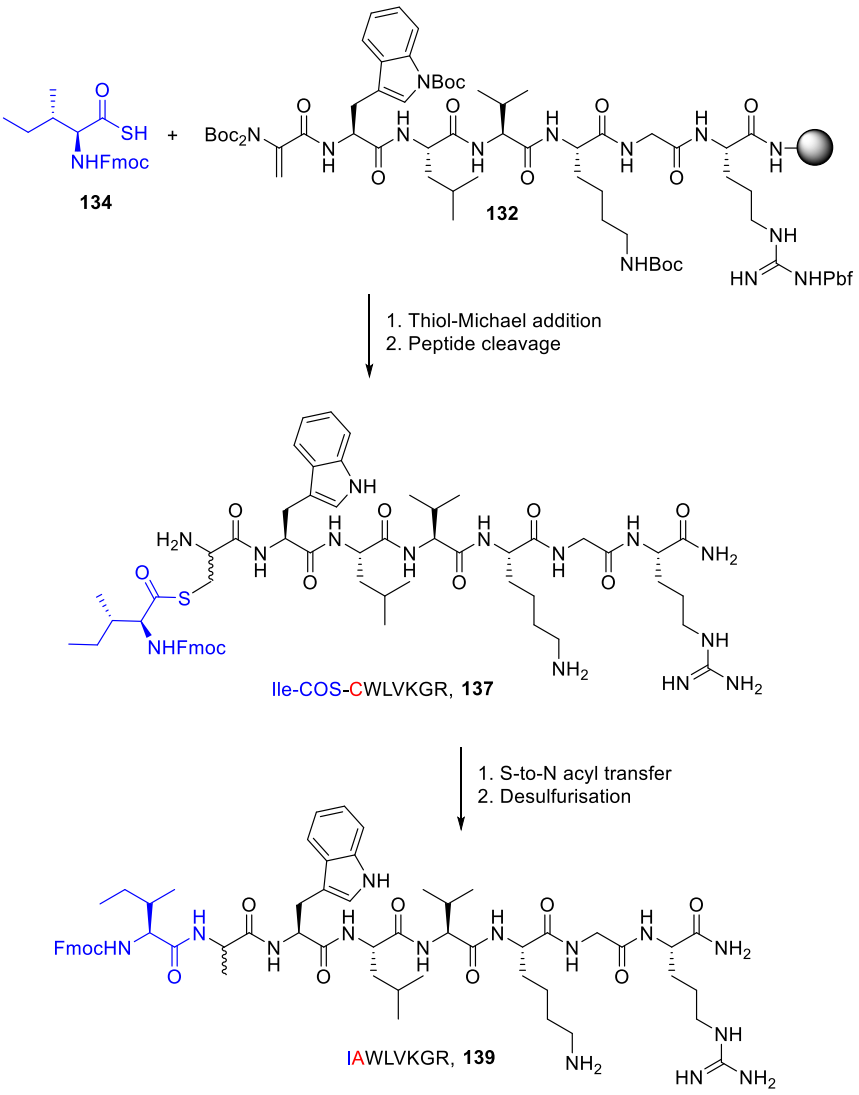
135 **136**, 90% **134**

Scheme 5.14. Coupling of Trt-SH to Fmoc-Ile-OH (**135**) to yield the thioester **136** which was deprotected to afford thioacid **134** that was used directly in the ligation reaction.

The sensitivity to this ligation reaction to solvent had previously been investigated within the group. As such, conditions identical to the optimised solution ligation conditions were used in this approach and are shown in **Scheme 5.15**. Thioacid **134** (1.5 equiv.) was dissolved in DMF and added to a syringe containing resin-bound **132**. Phosphate buffer (pH 8) was then added and the reaction was agitated for 16 h at 37 °C. Subsequently, the resin was washed and then cleaved with 95% (v/v) TFA, 2.5% (v/v) TES and 2.5% (v/v) H_2O. Satisfyingly, the presence of a thioester in the resultant peptide **137** was confirmed by MS, indicating that thiol-Michael addition had occurred. Analytical HPLC revealed 45% conversion of Dha peptide **132** to the thioester **137**, indicating that the ligation reaction did not proceed to completion. This low conversion may be due to thioacid hydrolysis occurring in the aq. reaction environment. The use of higher thioacid equivalents is undesirable due to the challenge of preparing sufficient quantities of large peptide thioacids. By increasing the rate of the thiol-Michael addition reaction, the impact of potential thioacid hydrolysis on the reaction yield could be minimalised. The use of aq. conditions in this reaction are not highly compatible with the polystyrene resin on which the Dha-peptide **132** is bound. Although pre-swelled in DMF, it is likely that the addition of H_2O results in contraction of the resin throughout the reaction. This would conceal reactive sites, slowing the reaction and limiting the yield. The reduced rate of reaction this would impart provides more time for thioacid decomposition. To increase the rate of reaction, the use of a resin which swells sufficiently in both organic and aq. solvents is required. Sequential peptide coupling and thiol-Michael addition performed

on a PEG based resin, such as ChemMatrix, functionalised with Rink amide linker may lead to an improvement in yield.

The obtained thioester **137** was subsequently stirred in an aq. solution of NH_4HCO_3 (0.5 M) to promote *S*-to-*N* acyl shift to yield a cysteine residue at the ligation site after 16 h. As this compound has the same molecular weight as the thioester starting material (**137**), desulfurisation of the resultant cysteine sulfhydryl (**138**) was performed to confirm acyl transfer had occurred. The cysteine of the product was photochemically desulfurised under conditions reported by Chisholm *et al*.[274] Glutathione, tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl) and Et_3N were added to peptide **138** solubilised in DMF and the reaction was irradiated at 254 nm for 3 h to afford the desired desulfurised product **139** which was identified by MS. Gratifying, this result indicates that both the acyl transfer and desulfurisation reactions occurred. As this project ultimately aims to utilise this reaction for peptide ligation in chemical protein synthesis, the identification of optimal ligation conditions required synthetic access to peptide thioacids.



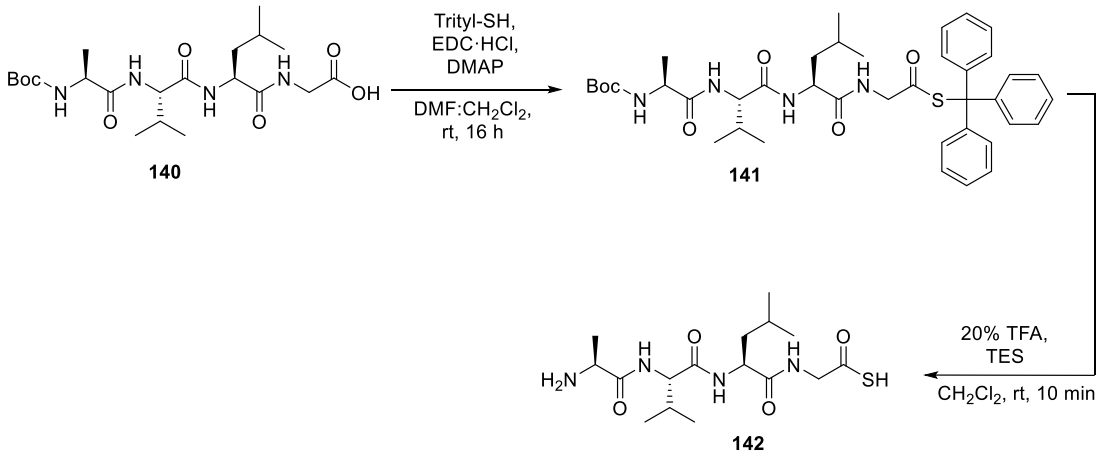
Scheme 5.15. Thiol-Michael addition of thioacid **134** to resin-bound Dha functionalised acceptor **132** following by peptide cleavage, *S*-to-*N* acyl transfer and desulfurisation to successfully yield peptide **139** with alanine at the ligation site.

5.2.5 Peptide Thioacid Synthesis

5.2.5.1 Solution Phase Synthesis of Peptide Thioacids

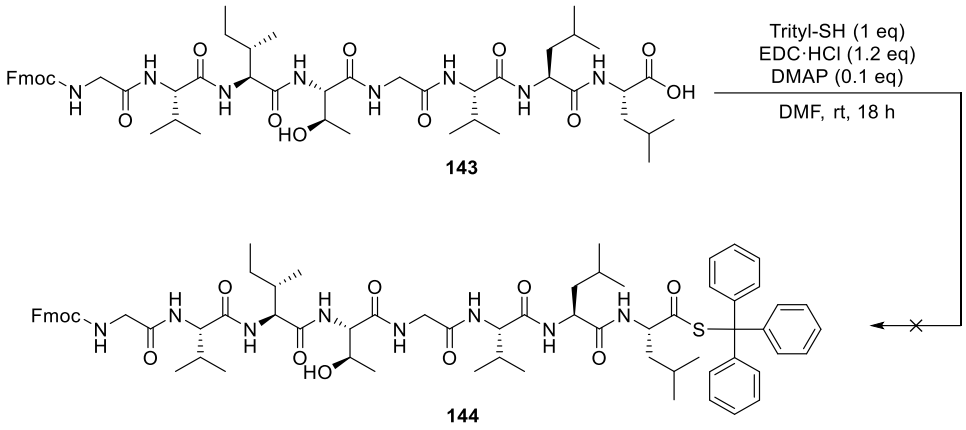
The Scanlan Group have previously extensively investigated the synthesis of α-amino thioacids *via* carbodiimide mediated thioesterification of Fmoc-amino acids with Trt-SH to produce trityl thioesters as masked thioacids. Trityl deprotection of these bench stable thioesters with 20% (v/v) TFA yields the corresponding thioacid in 10 min.[267] This method has been applied to the synthesis of peptide thioacids up to 4 residues in length, as illustrated in **Scheme 5.16**. However, the C-terminal residue of these peptides is restricted to glycine as C-terminal activation which is required to couple Trt-SH can result

in epimerisation of the stereocentre of the C-terminal amino acid. However, strategies to limit racemisation in thioester synthesis have been reported by Kajihara *et al.*,[275] and as protected peptides can be readily obtained from 2-chlorotrityl chloride resin, this strategy was further investigated as part of this work. The ligation reaction we aimed to apply these thioacids in utilises DMF as a co-solvent, as such, it was hypothesised that the limited solubility of protected peptides could potentially be overcome.



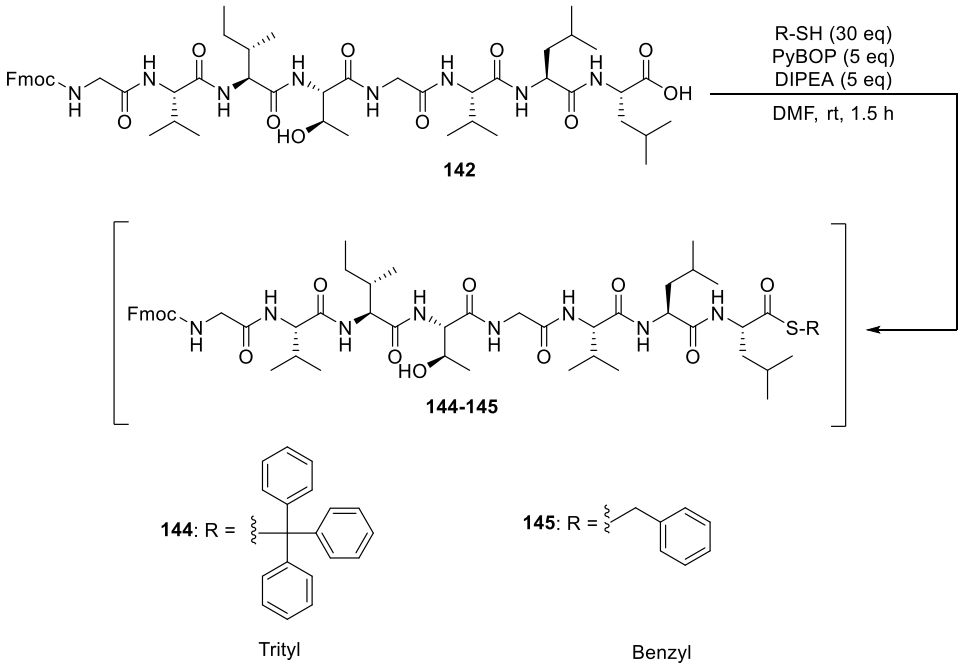
Scheme 5.16. Previous work by Dr. Lauren McSweeney towards thioacid synthesis in the Scanlan Group.

To identify conditions suitable for the synthesis the 23-mer thioacid **124** that was required for GLP-1 synthesis, the application of previously optimised conditions to peptides of medium length (8 residues) was first investigated. However, this strategy proved to be problematic. As shown in **Scheme 5.17**, the thioesterification of the carboxylic acid of a model peptide **143** was attempted with Trt-SH and standard EDC·HCl/DMAP mediated activation in DMF. Despite encountering no solubility issues, no reaction was observed by TLC after 18 h. Addition of a further 2 equiv. of Trt-SH and EDC·HCl did not promote reaction. The lack of reactivity may be caused by the peptide adopting a problematic conformation in DMF solution in which the carboxylic acid is not solvent accessible.



Scheme 5.17. Unsuccessful thioesterification of peptide **143** with Trt-SH in DMF using EDC·HCl/DMAP mediated activation.

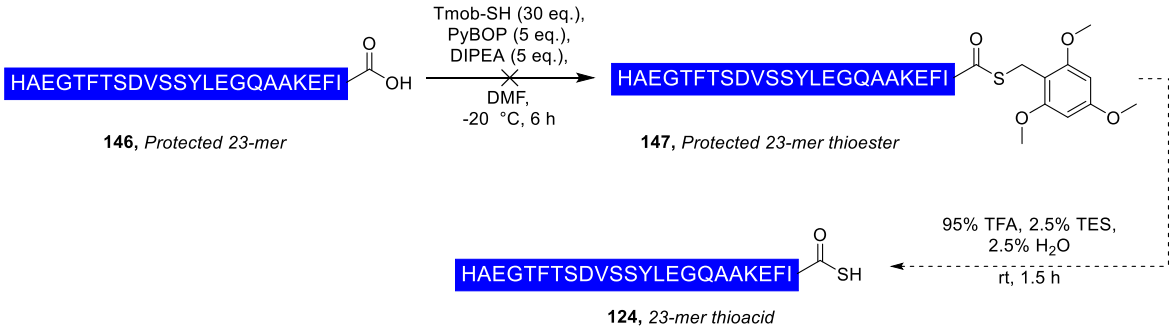
A procedure for the synthesis of *S*-benzyl thioesters of glycopeptides which was published by Kajihara *et al.*[275] was later used by Wilkinson *et al.*[276] in the preparation peptide alkyl thioesters. These conditions employ a large excess of thiol (30 equiv.), PyBOP activation and a temperature of -20 °C to limit racemization. Despite the low isolated yields of 4-9% reported, these coupling conditions were applied to peptide **143** with Trt-SH as shown in **Scheme 5.18**. In contrast to the EDC·HCl/DMAP mediated coupling, TLC analysis revealed the peptide was consumed after 1.5 h. It was not possible to observe product formation by TLC, potentially due to the large excess of reagents in the R_f region of the expected product. A trace amount of product was observed by MS in the crude mixture however isolation of the product by precipitation or column chromatography was not possible. No thioester carbonyl signal was detected in ^{13}C NMR and HMBC analysis of the crude material. Coupled with indications from MS analysis, this suggests product formation was minor. Despite being soluble at rt, Trt-SH exhibits low solubility in DMF at -20 ^{o}C. The reaction was consequently repeated with benzyl-SH which was soluble in DMF at -20 ^{o}C. MS analysis confirmed the presence of the desired product and a HMBC cross peak at (4.07, 201.7) ppm indicates the presence of a thioester carbonyl. Detection of this carbonyl by NMR suggests a higher concentration of desired product was formed with benzyl-SH than with trityl-SH, likely due to improved solubility of the thiol. However, as with the trityl analogue, it was not possible to isolate this compound due to the similar polarities and solubilities of excess reagents present.



Scheme 5.18. Poor yielding thioesterification of carboxylic acid **143** with both Trt-SH and benzyl-SH.

Previous reports in which these conditions are applied to the thioesterification of protected peptides isolate the benzyl thioester by precipitation in Et_2O. This is followed by global deprotection with 95% TFA to yield a water-soluble peptide thioester which was subsequently purified by RP-HPLC. In an attempt to emulate this approach, esterification of the C-terminal carboxylic acid of a protected 23-mer peptide (**146**) with 2,4,6-trimethoxyphenyl)methanethiol (Tmob) was attempted. Tmob is an acid labile thiol protecting group which can be deprotected using 30% TFA, 0.5% TES in CH_2Cl_2.[277] In this case, deprotection of the Tmob-peptide thioester would yield the 23-mer peptide thioacid (**124**) required for ligation to the resin-bound 7-mer acceptor peptide (**132**) (**Fig. 5.3, B**). The protected 23-mer **146** was synthesised successfully *via* manual SPPS from 2-chlorotrityl chloride resin. Following resin cleavage by 1% (v/v) TFA in CH_2Cl_2, the peptide carboxylic acid **146** was confirmed by MS. Unsurprisingly, RP-HPLC purification of the crude peptide was not possible due to the poor aq. solubility of the protected peptide. Thioesterification was hence attempted with the crude peptide. Tmob-SH coupling *via* PyBOP activation of peptide carboxylic acid **146** in DMF was performed as shown in **Scheme 5.19**. Despite extending the reaction time up to 6 h no product was identified in the crude reaction by MS although the starting material was consumed. The restrictions of this approach, including the inability to purify the starting material prior to thioesterification, the potential for epimerisation of the C-terminal residue during

activation and the requirement for final purification to be performed on the aq. soluble deprotected peptide thioacid resulted in this route not being further pursued. During side group deprotection, the TFA sensitive *S*-Tmob group will also be deprotected, yielding the unstable thioacid. This moiety is unlikely to survive both treatment with 95% aq. TFA for 1.5 h required for side group deprotection and chromatographic purification due to the propensity for thioacid oxidation to the corresponding diacyl-disulfide dimer under aerobic conditions and hydrolysis to the corresponding carboxylic acid.[278] Coupled with the challenging and low yielding nature of this route, solid-phase synthetic strategies towards thioacid synthesis were instead investigated.



146, *Protected 23-mer*

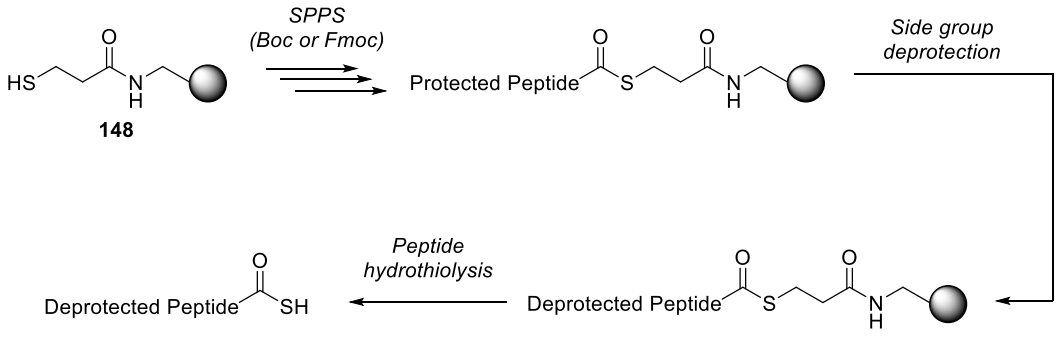
147, *Protected 23-mer thioester*

124, *23-mer thioacid*

Scheme 5.19. Unsuccessful solution-phase Tmob-SH coupling to protected peptide **146**.

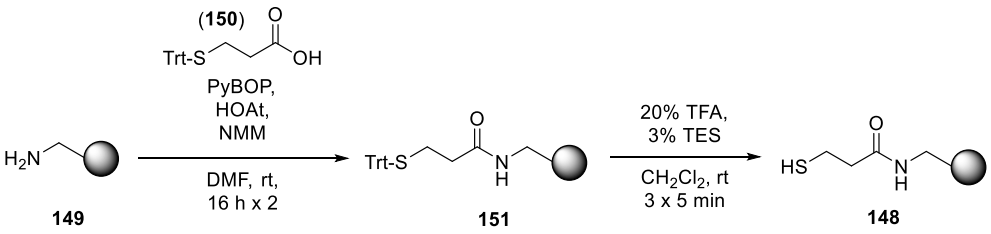
5.2.5.2 Solid-Phase Synthesis of Peptide Thioacids

The synthesis of peptide thioacids has been reported *via* multiple solid-phase strategies.[278] Solid-phase approaches towards thioacid synthesis do not require C-terminal activation of a peptide and therefore circumvent potential epimerisation issues. A methodology developed by Zhang *et al*.[279] involves a thioester linker bound to PEG-based ChemMatrix resin. The assembly of a peptide on resin with this linker has been demonstrated by the authors with both Boc and Fmoc SPPS strategies. Following peptide assembly, on-resin deprotection of amino acid side group protecting groups is followed by hydrothiolysis of the thioester linker to yield the deprotected peptide thioacid in aq. solution as depicted in **Scheme 5.20**.



Scheme 5.20. Synthetic strategy towards peptide thioacids reported by Zhang *et al*.[279]

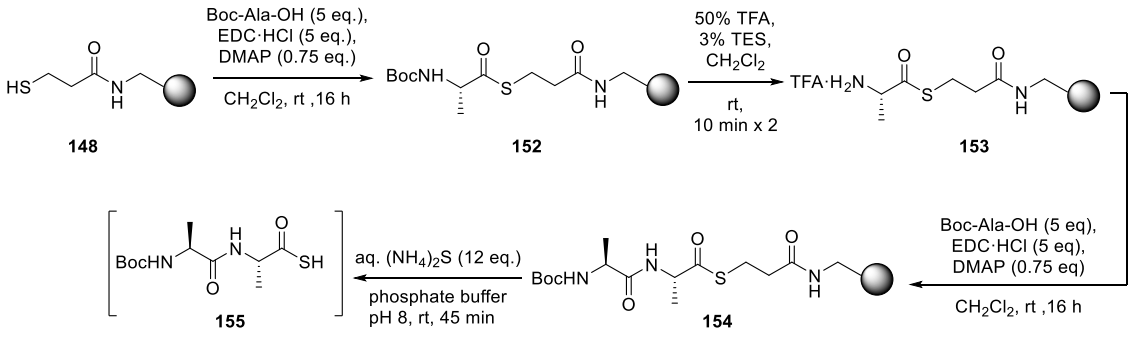
We attempted to apply this approach in thioacid synthesis. Initially, the preparation of thiol functionalised resin **148** was optimised. A PEG-based resin which swells in H_2O was required as the final hydrothiolysis step is performed under aq. conditions. Thus, aminomethyl functionalised ChemMatrix resin (**149**) was coupled to 3-(tritylthio)propanoic acid (**150**) *via* activation with PyBOP and hydroxy-7-azabenzotriazole (HOAt) with NMM in DMF for 16 h. The coupling was repeated for a further 16 h as a bromophenol blue test of the resin revealed the presence of uncapped amine groups. Following the repeated coupling the resin was capped using Ac_2O and DIPEA in DMF to ensure the full consumption of resin-bound amines groups. The *S*-trityl protecting group was then removed from the resin using 20% (v/v) TFA, 5% (v/v) TES in CH_2Cl_2 for 3 x 5 min as shown in **Scheme 5.21**. The presence of thiol on resin **148** was confirmed by the addition of a solution of 5,5'-dithiobis-(2-nitrobenzoic acid) (Ellman's reagent) in MeOH with DIPEA to a few resin beads. This disulfide indicator reacts with the thiol-functionalised resin, forming a mixed disulfide on resin with the release of an intensely coloured thiolate into solution. The resultant bright yellow colour observed in solution confirmed successful coupling and deprotection of the linker.



Scheme 5.21. Preparation of Liu linker on ChemMatrix resin.

Although the authors present both a Boc and Fmoc strategy for peptide assembly, the Boc approach was used in this work to prevent base mediated hydrolysis of the thioester linker during repeated Fmoc deprotection. On-resin Fmoc deprotection in the presence of a

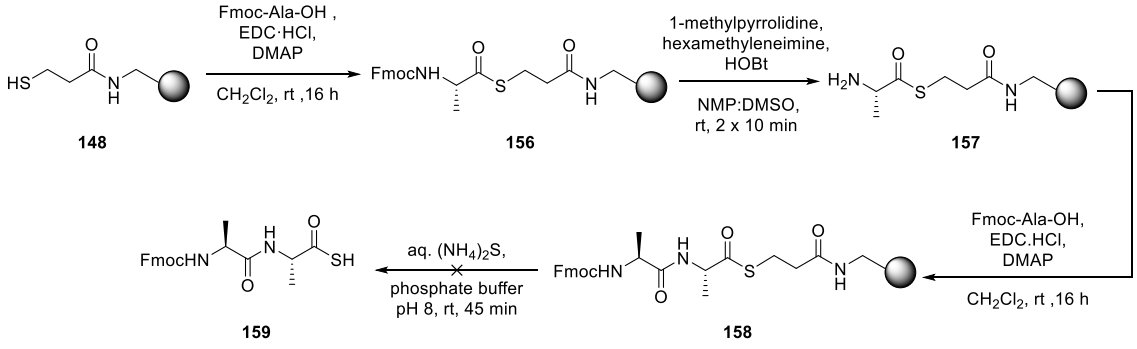
thioester has previously been investigated,[280] however, the minimised use of base in the Boc strategy was deemed desirable to limit premature resin cleavage. Coupling of the first amino acid to resin **148** was based on conditions which had been optimised within the Scanlan Group for solution-phase coupling of Trt-SH to *N*-protected amino acids in thioester synthesis. Thus, Boc-Ala-OH was dissolved in CH_2Cl_2, activated with EDC·HCl and DMAP, and then added to the resin for 1 h as shown in **Scheme 5.22**. Following this reaction, the addition of a few resin beads to Ellman's reagent resulted in no colour change, showing free thiol was not present and thus confirming completion of the reaction. Boc deprotection of resin **152** was then effected by 50% (v/v) TFA, 3% (v/v) TES in CH_2Cl_2 for 2 x 10 min. To the resin-bound free amine **153**, Boc-Ala-OH was coupled using EDC·HCl, DMAP, and DIPEA for 45 min. DIPEA was included to freebase the resin-bound TFA-salted amine arising from the previous Boc deprotection step. To analyse the progress of this synthesis, the dipeptide was cleaved from resin **154** after coupling of this second amino acid. Resin cleavage was performed by treatment of the resin **154** with aq. ammonium sulfide solution in phosphate buffer at pH 8 for 45 min. The resultant solution was then acidified to pH 2 and degassed with argon to remove H_2S which would interfere in the ligation reaction. MS analysis revealed the presence of the desired thioacid, however, TLC and NMR analysis of the extracted of the reaction mixture revealed a large number of impurities.



Scheme 5.22. Synthesis of thioacid **155**.

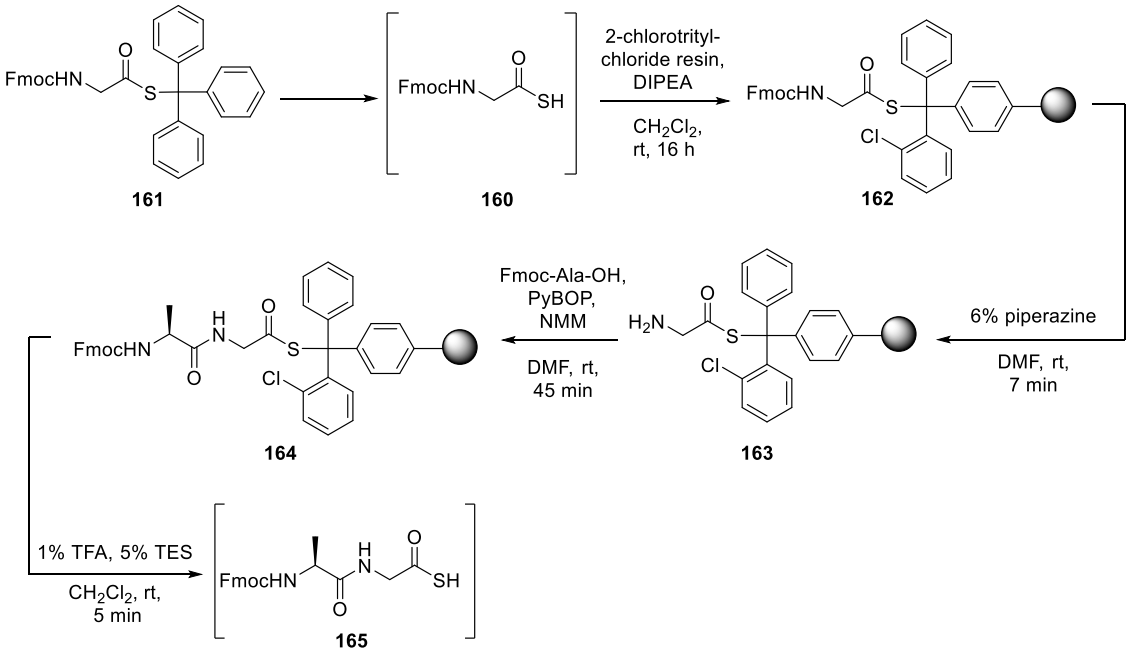
The number of impurities formed during thioacid synthesis suggests poor efficiency of the reactions carried out up to this stage. Base-catalysed hydrolysis of the thioester linker may be occurring during amino acid coupling steps, regenerating the thiol-functionalised resin which can subsequently undergo further coupling. Such side reactions would result in low crude purity of the resultant peptide mixture, particularly with peptides containing a large number of residues due to the extensive number of coupling reactions performed.

As we aimed to use the thioacid directly in ligation, optimisation of the synthesis was thus attempted to avoid chromatographic purification of the unstable intermediate. The second amino acid coupling was hence repeated without DMAP to prevent the potential nucleophilic attack of the thioester cleaving the linker, however, no improvement in crude purity was observed. Next, the synthesis was attempted using a Fmoc-based SPPS strategy. Fmoc-Ala-OH was coupled to the resin-bound linker **148** and the Fmoc group was deprotected using 1-methylpyrrolidine, hexamethyleneimine and HOBt in *N*-methyl-2-pyrrolidone (NMP) and DMSO. These conditions were reported by Li *et al.*[280] for the deprotection of the Fmoc group on solid phase in the presence of a peptide thioester. Subsequently, a second Fmoc-Ala-OH residue was coupled to the deprotected resin and the peptide was then cleaved by hydrothiolysis as shown in **Scheme 5.23**. MS analysis of the cleaved material did not identify the desired product. The synthesis was repeated using 6% (v/v) piperazine in DMF to effect Fmoc deprotection as reported by Schçwe *et al.*[281] in the presence of a thioester, however, this strategy did not yield the desired product. Thus, the Fmoc strategy, which employs a greater number of steps involving base, was less successful than the Boc SPPS equivalent. This comparison is a further indication that base-mediated thioester hydrolysis may be occurring throughout the synthesis. The total avoidance of base in an SPPS strategy is not feasible. Fmoc strategies inherently rely on basic conditions for Fmoc deprotection and Boc strategies require base to freebase the TFA salt of the N-terminus following Boc deprotection. Although SPPS strategies for thioester and thioacid synthesis using peptides linked to a resin *via* various thioester linkers have been reported in the literature,[278] it is possible that the nature of this alkyl linker yields a highly electrophilic thioester which is prone to nucleophilic hydrolysis.



Scheme 5.23. Attempted optimisation of thioacid synthesis using a Fmoc based approach was not successful.

Inspired by successful thioacid synthesis in the Scanlan lab utilising *S*-trityl protection in solution, we attempted to transfer this approach to an on-resin strategy. This method aimed to introduce the first residue as an amino thioacid, as depicted in **Scheme 5.24** to modify to a 2-chlorotrityl linker and yield a thioester. As such, Fmoc-Gly-SH (**160**) was prepared from Fmoc-Gly-STrt (**161**) and added to 2-chlorotrityl chloride resin for 16 h. Following extensive washing of the resin to remove residual reactants, resin cleavage was performed on a small portion of the resin **162** with 1% (v/v) TFA, 5% (v/v) TES in CH_2Cl_2. Promisingly, the desired thioacid Fmoc-Gly-SH (**160**) was afforded which was identified by MS. The synthesis on-resin was continued with piperazine mediated Fmoc-deprotection of resin **162**, followed by coupling of Fmoc-Ala-OH and subsequent cleavage of the thioacid dipeptide from resin **165**. Advantageously, mild acidic conditions (1% (v/v) TFA in CH_2Cl_2) can be used to cleave the peptide in minutes, minimalizing potential thioacid decomposition Although MS and NMR analysis confirmed the presence of the desired compound, the crude material **165** was again prohibitively impure, as with the previous strategy. Competing hydrolysis of the thioester linker during amino acid coupling is again a potential cause of the low crude purity obtained. To provide access to peptide thioacids in synthetically viable yields and purity through this strategy, the development of Fmoc-deprotection and coupling conditions which minimalize competing linker hydrolysis are required.



Scheme 5.24. Synthesis of a dipeptide thioacid on 2-chlorotrityl chloride resin.

5.2.6 Conclusions and Future Work

In conclusion, the synthesis of Dha-functionalised resin-bound peptides has been achieved. This has enabled the scope of the β,γ-*C*,*S* thiol-Michael addition of thioacids to Dha to be expanded to include peptidic acceptors. The potential of this reaction in peptide ligation has been shown by the *S*-to-*N* acyl transfer of the thioester over a 5-membered cyclic intermediate to yield a native peptide bond. Additionally, a novel application of 2-chlorotrityl chloride resin in this synthesis of thioacids has been identified. The limiting factor in the goal of this project towards peptide ligation is the challenge of the efficient synthesis of peptide thioacids. If an on-resin strategy which yields a C-terminal thioacid after resin cleavage is to be employed, it is likely that the thioacid will decompose during purification. Therefore, it is desirable to react the crude peptide thioacid directly with the Dha containing fragment. In this situation, a high purity of crude thioacid is required to enable the successful synthesis, characterisation, and purification of the resultant ligated product. This requires the identification of SPPS conditions which are compatible with thioester linkers to minimise side product formation during the assembly of the peptide on-resin.

Chapter 6

Overall Conclusions

6.1 Overall Conclusions

In this work, a range of lipopeptide based probes have been prepared through two synthetic strategies. The first approach, which is utilised to synthesise the probes described in Chapters 2-4, employs a linear method in which a dagylated cysteine residue is incorporated *via* SPPS. This strategy has been successfully implemented to access peptide probes for application in the functional screening of LspA, Lnt and Lit. Biological screening of the peptides synthesised in this thesis was performed by collaborators in the group of Prof. Martin Caffrey. The second synthetic strategy, which is described in Chapter 5, is based on the liposome-mediated late-stage lipidation of Dha containing peptides. Implementation of this methodology will provide the means to overcome several limitations of the linear approach that were encountered in the course of this work.

In Chapter 2, we described the synthesis of lipopeptide-based FRET probes and prepare an optimal FRET substrate of LspA. The synthetic challenges encountered in this chapter, including the purification of highly hydrophobic, organic and water insoluble peptides were mediated by alteration of the peptide sequence and use of normal phase silica-based chromatography systems. Ultimately, the functionalised peptides prepared in this chapter have been utilised to explore the substrate scope of the enzyme and have facilitated the development of a high-throughput assay for LspPae. Optimisation of the synthetic route provided access to the purified target FRET probe in quantities up to 250 mg, a sufficient amount to facilitate a high-throughput screening campaign of a major compound library which has expedited efforts to identify novel inhibitors of LspA. Furthermore, the 1D and 2D NMR analysis of the peptides synthesised provided a foundation for the work detailed in Chapter 4 in which NMR spectroscopy was used to probe the mechanism of the Lit reaction.

In Chapter 3 our efforts towards the synthesis of Lnt substrates are detailed. The efficient synthesis of a dagylated cysteine monomeric substrate (compound **61**) was established. Subsequently, a short lipopeptide featuring a lysine ε*N*-linked fluorophore was designed and utilised in a TLC-based assay for this enzyme. This has enabled the second step of the Lnt reaction to be monitored directly, however, this assay does not provide a continuous readout and is not suitable for high-throughput application. Attempts to transfer this substrate to a FRET-based assay for Lnt to overcome these limitations have been investigated. However, a complimentary Lnt substrate which is capable of acylating

the Lnt active site cysteine with a quencher labelled acyl chain has yet to be identified. Due to the synthetic challenge of selectively modifying the *sn*-1 acyl chain of a phospholipid, *in silico* modelling of potential substrates with Lnt could provide valuable guidance in substrate design prior to further synthetic efforts.

In Chapter 4 we have established the synthesis, purification, and characterisation of fluorescent and deuterium-labelled Lit substrates. Processing of these substrates by Lit yielded the corresponding lyso-peptides which were isolated from the enzymatic reaction and characterised by NMR spectroscopy. This provided detailed insights into the acyl transfer reaction which is catalysed by Lit. Furthermore, we utilised the paramagnetic properties of Cu(II) to investigate the interaction of this metal ion at the N-terminal region of diacylated lipopeptides *via* NMR spectroscopy. We provide preliminary data regarding the selective interaction of Cu(II) at the N-terminus of peptide **84** in DMSO solution. In the future, a comparative study is necessary to analyse the effect of LP *N*-acylation on Cu(II) binding.

In Chapter 5 we explore two applications of thiol-Michael addition to Dha containing peptides. Firstly, the addition of dipalmitoylthioglycerol (**99**) to a Dha functionalised peptide was demonstrated. This biocompatible methodology is performed in aq. conditions and yields a native DAG functionality at a cysteine residue of a peptide. The inevitable orthogonal solubility of the two reactants of this reaction was overcome *via* incorporation of the reagents into FC-12 based liposomes. This accelerated the rate of reaction and provided access to the product in synthetically viable yields. Future work on this strategy will first require the synthesis of the enantiomerically pure dagylated thiol to facilitate the preparation of potential LspA substrates. Furthermore, building on the established methods of Dha introduction into a protein, we envision that this biocompatible reaction could have applications in protein dagylation to provide semi-synthetic access to native bacterial LPs. Chapter 5 concluded by detailing the use of Dha in a novel peptide ligation methodology. The preparation of resin-bound N-terminal cysteinyl peptide thioesters was demonstrated *via* β,γ-*C,S* thiol-Michael addition of thioacids to Dha containing peptides. Following resin cleavage of the resulting peptide thioester, *S*-to-*N* acyl transfer of the thioester over a 5-membered cyclic intermediate to yield a native peptide bond was demonstrated, highlighting the potential applications of this reaction in peptide ligation. To further the scope of this methodology, future work will require the development of robust synthetic routes towards peptide thioacids.

Progress towards this aim was made through a novel application of 2-chlorotrityl chloride resin, however, if an on-resin strategy is to be employed, it is necessary to thoroughly optimise each SPPS reaction to develop conditions under which the labile thioester linker is stable to prevent premature peptide cleavage from the resin. Ultimately, with access to a range of appropriately functionalised peptide fragments, the scope of this reaction can be fully explored.

Chapter 7

Experimental

7.0 Experimental

7.1 Instrumental and General Considerations

^{1}H NMR and ^{13}C NMR spectra were recorded on a 400 MHz Bruker Avance spectrometer (^{1}H, 400.13 MHz; ^{13}C, 100.6 MHz) or a 600 MHz Bruker Avance II spectrometer (^{1}H, 600.13 MHz; ^{13}C, 150.6 MHz). Resonances δ, are in ppm calibrated using residual nondeuterated solvent (^{1}H NMR) or the deuterated solvent (^{13}C NMR) as internal reference standards. Deuterated solvents were purchased from Merck, VWR and Fluorochem. NMR data was analysed using MestReNova v6.0.2-5475. Infrared (IR) spectra were recorded on a Perkin Elmer spectrometer. Mass spectrometry analysis was performed with a Waters Premier quadrupole time-of-flight (Q-ToF) mass spectrometer equipped with Z–spray ESI and MALDI sources. *trans*-2-[3-(4-*tert*-Butylphenyl)-2-methyl-2-propenylidene]malononitrile) was used as the MALDI matrix. Silica gel Florisil (200 mesh; Merck) was used for column chromatography. TLC was performed using Merck 60 F254 silica gel plates (pre-coated, 0.2 mm thick) and visualised by UV light (λ_{max} = 254 nm) and ammonium molybdate staining (ammonium molybdate (0.26 M) in aq. H_2SO_4 (1 M)). Unless otherwise stated, protected amino acids for peptide synthesis and all other reagents were purchased from an industrial supplier (Fluorochem, Merck, VWR, Bachem or Avanti Polar Lipids). Rt is classified in this work as 19.5 – 20.5 °C. Manual SPPS reactions were performed in polypropylene syringe reaction vessels (5 or 10 mL; Torviq, MI, USA) fitted with a polystyrene frit. SPPS was performed manually unless otherwise stated. Manual SPPS reactions were carried out under continuous agitation at rt unless otherwise stated. Automated SPPS reactions were performed with a CEM Liberty Blue microwave peptide synthesiser. HPLC or peptide synthesis grade solvents were used for SPPS. RP-HPLC was performed on a Shimadzu Prominence system. For analytical HPLC, either a C_{18} Phenomenex Gemini 5 µm, 110 Å, 250 x 4.6 mm LC column or a C_{18} Phenomenex Clarity 5 µm Oligo-RP 150 x 4.60 mm LC column was used, as stated, with a flow rate of 1 mL/min. For semi-preparative HPLC, a C_{18} Phenomenex Gemini 5 µm, 110 Å, 250 x 10 mm LC column was used with a flow rate of 4 mL/min. UV absorption signals were detected with a PDA detector at wavelengths of 280 nm, 254 nm and 214 nm.

7.2 General Experimental Procedures

Procedure A: Disulfide reduction

To a solution of disulfide (1.0 equiv.) in anhydrous THF under argon was added PBu_3 (1.0 - 2.0 equiv. as stated). The reaction was stirred at rt for 2 min before H_2O was added. The reaction was then stirred at rt for 4 h and the solvents were subsequently removed *in vacuo*. The crude product was either purified by column chromatography (SiO_2) to yield the corresponding thiol or used directly in the next step as stated.

Procedure B: *tert*-Butyl ester deprotection

A solution of protected amino acid (0.15 M) in TFA:TES:CH_2Cl_2 (7:2:1 v/v/v) was stirred at rt for 4 h. Toluene (10 mL) was added and the solvents were removed *in vacuo* to yield the corresponding product.

Procedure C: Preparation of *S*-trityl thioesters

To a solution of the carboxylic acid (1.0 equiv) in anhydrous CH_2Cl_2 under argon was added Trt-SH (1.0 equiv), DMAP (0.1 equiv), and EDC·HCl (1.2 equiv). The reaction was stirred at rt for 16 h and the solvent was then removed *in vacuo.* The product was purified by column chromatography (SiO_2).

Procedure D: *S*-Trityl deprotection

To the trityl-thioester under argon was added TFA:TES:CH_2Cl_2 (2:1:7). The reaction was stirred at rt for 10 min and volatiles were subsequently removed *in vacuo* to yield the thioacid which was used directly without further purification.

7.2.1 SPPS General Produces

Procedure E: First amino acid coupling to Rink amide resin

To a SPPS syringe was added Rink amide AM resin (0.7 mmol/g, 100-200 mesh, 1.0 equiv.) and DMF (5 mL). The syringe was agitated for 20 min, then drained. 20% (v/v) Piperidine in DMF (5 mL) was added to the resin for 2 x 10 min to effect Fmoc deprotection. The deprotected resin was washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL). PyBOP (4.0 equiv.) and NMM (8.0 equiv.) were added to a solution of Fmoc protected amino acid (4.0 equiv.; 0.2 M) in DMF. This solution was transferred to the resin in the syringe and agitated for 45 min. Excess reagents were

drained from the reaction vessel and the resin was washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL).

Procedure F: First amino acid coupling to Wang resin

To a SPPS syringe was added Wang resin (1.4 mmol/g, 100-200 mesh, 1.0 equiv.) and DMF (5 mL). The syringe was agitated for 20 min, then drained. PyBOP (4.0 equiv.), NMM (8.0 equiv.) and DMAP (0.1 equiv.) were added to a solution of Fmoc protected amino acid (4.0 equiv.; 0.2 M) in DMF. This solution was transferred to the resin in the syringe and agitated for 45 min. Excess reagents were drained from the reaction vessel and the resin was washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL).

Procedure G: First amino acid coupling to 2-chlorotrityl chloride resin

To a SPPS syringe was added 2-chlorotrityl chloride resin (1.1 mmol/g, 100-200 mesh, 1.0 equiv.) and anhydrous CH_2Cl_2 (5 mL) under argon. The syringe was agitated for 20 min, then drained. DIPEA (7.5 equiv.) was added to a solution of Fmoc protected amino acid (3.0 equiv.; 0.2 M) in anhydrous CH_2Cl_2 under argon. This solution was transferred to the resin in the syringe and agitated for 16 h under argon. Excess reagents were then drained from the reaction vessel and the resin was washed with CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (5 mL). A solution of CH_2Cl_2:DIPEA:MeOH (17:2:1 v/v/v; 5 mL) was added to the syringe for 2 x 10 min and then the resin was washed with two rounds of CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL).

Procedure H: Bromophenol blue monitoring of coupling reactions

The extent of each coupling reaction was qualitatively monitored by treatment of 5-10 resin beads with a solution of bromophenol blue (BPB) in CH_2Cl_2 (0.15 mM; 0.1 mL). In the case of incomplete coupling, the coupling procedure was repeated prior to Fmoc deprotection.

Procedure I: Second and subsequent amino acid Fmoc deprotection and coupling

Each subsequent amino acid coupling cycle consisted of (i) Fmoc deprotection of the resin-bound peptide by the addition of 20% (v/v) piperidine in DMF (5 mL) to the resin for 2 x 10 min, (ii) resin washes with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL), (iii) peptide coupling with addition of PyBOP (3.0 equiv.), NMM (6.0 equiv.)

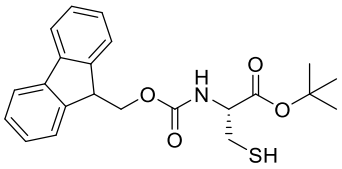
and *N*-protected amino acid (3.0 equiv.; 0.2 M) in DMF to the peptide resin for 45 min, (iv) resin washes with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL), (v) qualitative BPB test as per procedure H. Following the final coupling, the resin was treated with 20% (v/v) piperidine in DMF (2 x 10 min; 5 mL) and the resin was washed with 2 rounds of DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL), followed by drying of the resin under reduced pressure.

Procedure J: Resin cleavage and global deprotection

To the dry peptide resin was added CH_2Cl_2 (5 mL). The syringe was agitated for 20 min, then drained. The cleavage cocktail (TFA:TES:H_2O; 95:2.5:2.5 v/v/v; 5 mL) was added to the syringe which was tightly capped and agitated for 90 min. The syringe was drained, and the filtrate was collected. The resin was washed with cleavage cocktail (2 x 2.5 mL) and the washings were combined with the initial filtrate. The combined solution was concentrated under a flow of N_2, followed by precipitation of the peptide with Et_2O (10 mL) at 0 °C. The crude peptide suspension was centrifuged and the supernatant decanted. The peptide pellet was washed again with Et_2O (2 x 10 mL) at 0 °C, centrifuged and collected. The crude solid was dried *in vacuo*.

7.3 Experimental Details for Chapter 2

tert-Butyl (((9*H*-fluoren-9-yl)methoxy)carbonyl)-L-cysteinate (33)



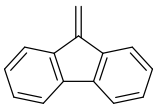
Prepared according to procedure A using (Fmoc-Cys-O*t*Bu)$_2$ **32** (3.38 g, 4.24 mmol), PBu_3 (2.12 mL, 8.50 mmol), THF (50 mL) and H_2O (5 mL). Purified by column chromatography (SiO_2, EtOAc:hex 1:4 v/v) to give thiol **33** as a colourless oil (1.96 g, 58%). The isolated compound was in good agreement with the literature.[282]

R_f = 0.46 (EtOAc:hex 1:4 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.61 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.41 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.33 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 5.69 (d, *J* = 6.9 Hz, 1H, NH), 4.59 – 4.52 (m, 1H, Cys αCH), 4.47 – 4.36 (m, 2H, Fmoc CH_2), 4.24 (t, *J* = 6.9 Hz, 1H, Fmoc CH), 3.04 – 2.94 (m, 2H, Cys βCH_2), 1.51 (s, 9H, CH_3 x 3) ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{22}H_{25}NO_4SNa$ = 422.1397 $(M + Na)^+$. Found = 422.1396.

9-methylene-9*H*-fluorene (166)

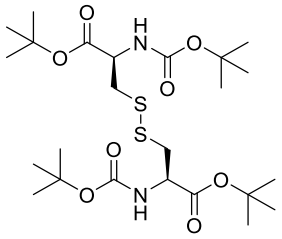


Obtained from the attempted synthesis of *tert*-butyl *N*-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-*S*-(2,3-dihydroxypropyl)-L-cysteinate (**34**): To a solution of thiol **33** (500 mg, 1.25 mmol) in anhydrous DMF (5 mL) under argon at 0 °C was added Cs_2CO_3 (408 mg, 1.25 mmol). The reaction was stirred at 0 °C for 1 h. A solution of 3-bromo-1,2-propanediol (214 mg, 1.39 mmol) in DMF (5 mL) was added to the reaction which was then stirred 0 °C for 1 h. The reaction was diluted with H_2O (50 mL) and extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with sat. aq. NaCl solution (3 x 100 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the product obtained from this reaction was isolated by column chromatography (SiO_2, EtOAc:hex. 1:1 v/v) which yielded the by-product 9-methylene-9*H*-fluorene (147 mg, 66%);

R_f = 0.80 (EtOAc:hex, 1:1 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.78 – 7.74 (m, 2H, Fmoc Ar-CH), 7.71 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.39 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.32 (td, *J* = 7.4, 0.9 Hz, 2H, Fmoc Ar-CH), 6.09 (s, 2H, C=CH_2) ppm.

Di-*tert*-butyl-3,3')-disulfanediyl(2*R*,2'*R*)-bis(2-((*tert*-butoxycarbonyl)amino)propanoate) (37)



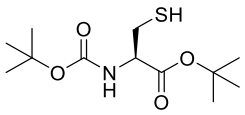
To a solution of (H-Cys-O*t*Bu)$_2$·2HCl **36** (3.00 g, 7.05 mmol) in H_2O:dioxane (1:1 v/v; 40 mL) was added Et_3N (3.93 mL, 28.2 mmol) then Boc_2O (3.39 g, 15.5 mmol). The reaction was stirred at rt for 16 h. Dioxane was evaporated under reduced pressure and the aq. layer was acidified to pH 3 using aq. HCl solution (1 M) and extracted with EtOAc (3 x 50 mL). The combined extracts were washed with sat. aq. NaCl solution (3 x 50 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* to yield a pale yellow viscous oil (3.79 g, 97%). The isolated compound was in good agreement with the literature.[283]

R_f = 0.45 (EtOAc:hex, 1:9 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.34 (d, *J* = 6.9 Hz, 2H, NH), 4.48 - 4.40 (m, 2H, Cys αCH), 3.20 (dd, *J* = 13.9 Hz, *J* = 5.0 Hz, 2H, Cys βC$\underline{H_a}H_b$), 3.11 (dd, *J* = 13.9 Hz, *J* = 5.0 Hz, 2H, Cys βC$H_a\underline{H_b}$), 1.47 (s, 18H, Boc CH_3 x 3), 1.44 (s, 18H, O*t*Bu CH_3 x 3) ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{24}H_{44}N_2O_8S_2Na$ = 575.2437 $(M + Na)^+$. Found 575.2431.

***tert*-Butyl (*tert*-butoxycarbonyl)-L-cysteinate (35)**



Prepared according to Procedure A using disulfide **37** (3.89 g, 7.04 mmol), PBu_3 (2.85 mL, 14.1 mmol), THF (40 mL) and H_2O (5 mL). The crude product was purified by

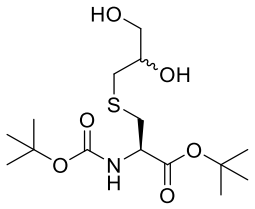
column chromatography (SiO_2, CH_2Cl_2) to yield a white solid (3.51 g, 90%). The isolated compound was in good agreement with the literature.[284]

R_f = 0.23 (EtOAc:hex. 1:9 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.38 (d, *J* = 6.0 Hz, 1H, NH), 4.52 - 4.37 (m, 1H, Cys αCH), 2.94 (dd, *J* = 8.8 Hz, *J* = 4.0 Hz, 2H, Cys βCH_2), 1.47 (s, 9H, Boc-CH_3 x 3), 1.44 (s, 9H, O*t*Bu-CH_3 x 3) ppm;

HRMS (*m/z* ESI$^-$) calcd. for $C_{12}H_{22}NO_4S$ = 276.1270 (M - H)$^-$. Found = 276.1275.

tert-Butyl _N_-(_tert_-butoxycarbonyl)-_S_-(2,3-dihydroxypropyl)-L-cysteinate (38)



To a solution of thiol **35** (2.48 g, 8.95 mmol) in anhydrous DMF (25 mL) under argon was added Et_3N (12.5 mL, 89.5 mmol) and 3-bromo-1,2-propanediol (4.37 g, 13.4 mmol). The reaction was stirred under argon at rt for 3 d until TLC analysis indicated completion by consumption of compound **35** (R_f = 0.23; EtOAc:hex, 1:9 v/v). DMF was evaporated under reduced pressure and the crude material was dissolved in CH_2Cl_2 (50 mL) and washed with H_2O (50 mL). The aq. layer was back extracted with CH_2Cl_2 (5 x 20 mL) and the combined organic layers were washed with sat. aq. NaCl solution (50 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude product was purified by column chromatography (SiO_2, EtOAc:hex. 50:50 v/v) to yield a colourless oil (2.37 g, 76%). The isolated compound was in good agreement with the literature.[285]

R_f = 0.55 (EtOAc:hex, 8:2 v/v);

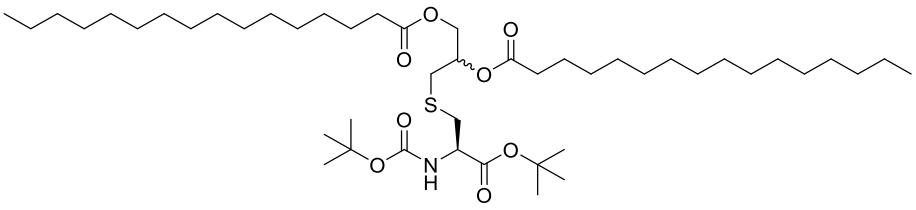
^{1}H NMR (600 MHz, DMSO-d_6) two diastereoisomers δ 7.13 (d, *J* = 7.7 Hz, 1H, NH), 4.77 (d, *J* = 4.9 Hz, 1H, S-glyceryl CH-O$\underline{H}$), 4.58 – 4.53 (m, 1H, S-glyceryl CH_2-O$\underline{H}$), 4.01 – 3.93 (m, 1H, Cys αCH), 3.59 – 3.53 (m, 1H, S-glyceryl C$\underline{H}$), 3.36 – 3.32 (m, 2H, S-glyceryl OC$\underline{H_2}$), 2.91 – 2.82 (m, 1H, Cys βC$\underline{H_a}H_b$), 2.77 – 2.72 (m, 1H, Cys βC$H_a\underline{H_b}$), 2.68 – 2.61 (m, 1H, S-glyceryl C$\underline{H_a}H_b$), 2.49 – 2.44 (m, 1H, S-glyceryl C$H_a\underline{H_b}$), 1.40 (s, 9H, Boc CH_3 x 3), 1.39 (s, 9H, O*t*Bu CH_3 x 3) ppm;

^{13}C NMR (151 MHz, DMSO-d_6) two diastereoisomers δ 170.3 (αCH-<u>C</u>O), 155.3 (NH-CO), 80.7 (Boc qC), 78.2 (O*t*Bu qC), 71.3 (S-glyceryl CH), 64.4 (S-glyceryl OCH_2), 54.6 (Cys αC), 35.7 (S-glyceryl CH_2), 33.5 (Cys βCH_2), 28.1 (O*t*Bu CH_3), 27.6 (Boc CH_3) ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{15}H_{29}NO_6SNa$ = 374.1613 $(M + Na)^+$. Found = 374.1600;

ν_{max} (thin film)/cm^{-1} 3368 (NH), 2928 (CH), 1732 (C=O), 1054 (C-O).

3-(((*R*)-3-(*tert*-butoxy)-2-((*tert*-butoxycarbonyl)amino)-3-oxopropyl)thio)propane-1,2-diyl dipalmitate (39)



To a solution of palmitic acid (4.32 g, 16.7 mmol) in anhydrous DMF:CH_2Cl_2 (100 mL, 1:1 v/v) at 0 °C under argon was added EDC·HCl (3.88 g, 20.2 mmol) and DMAP (0.123 g, 1.35 mmol). The reaction was stirred at 0 °C for 1 h. Diol **38** (2.37 g, 6.74 mmol) was dissolved in anhydrous CH_2Cl_2 (50 mL) was added to the reaction which was then stirred at rt under argon for 6 d when TLC analysis indicated completion of the reaction following consumption of the monoacylated intermediate (R_f = 0.32; EtOAc:hex 1:9 v/v). The solvents were removed *in vacuo* and the crude product was purified by column chromatography (EtOAc:hex, 1:9 v/v) to yield the diacylated product as a white solid (5.05 g, 90%).

R_f = 0.60 (EtOAc:hex, 1:9 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) two diastereoisomers δ 5.39 – 5.30 (m, 1H, NH), 5.23 – 5.10 (m, 1H, S-glyceryl CH), 4.46 – 4.39 (m, 1H, Cys αCH), 4.39 – 4.32 (m, 1H, S-glyceryl OCH_aH_b), 4.22 – 4.13 (m, 1H, S-glyceryl OCH_aH_b), 3.09 – 2.94 (m, 2H, Cys βCH_2), 2.86 – 2.70 (m, 2H, S-glyceryl CH_2), 2.36 – 2.30 (m, 4H, Pal αCH_2 x 2), 1.70 – 1.55 (m, 4H, Pal βCH_2 x 2), 1.49 (s, 9H, Boc CH_3 x 3), 1.47 (s, 9H, O*t*Bu CH_3 x 3), 1.35 – 1.20 (m, 48H, Pal CH_2 x 24), 0.90 (t, *J* = 6.7 Hz, 6H, Pal CH_3 x 2) ppm;

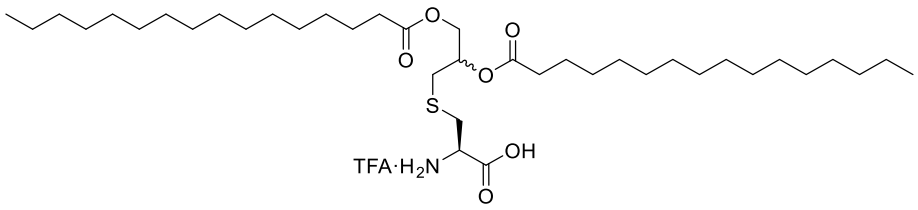
^{13}C NMR (101 MHz, $CDCl_3$) two diastereoisomers δ 173.3 (αCH-<u>C</u>O), 172.9 (Pal C=O), 173.0 (Pal C=O), 169.8 (NH-CO), 82.7 (Boc qC), 79.9 (O*t*Bu qC)) 70.4 (S-glyceryl CH),

63.5 (S-glyceryl OCH_2), 53.9 (Cys αC), 35.5 (Cys βCH_2), 34.1 (Pal αCH_2), 33.3 (S-glyceryl CH_2), 29.7, 29.5, 29.4, 29.3 (Pal CH_2), 28.3 (O*t*Bu CH_3), 28.0 (Boc CH_3), 24.9 (Pal βCH_2), 14.1 (Pal CH_3) ppm;

ν_{max} (thin film)/cm^{-1} 3391 (NH), 2919 (CH), 1736 (Pal C=O), 1689 (NHC=O);

HRMS (*m/z* ESI^+) calcd. for $C_{47}H_{89}NO_8SNa$ = 850.6207 $(M + Na)^+$. Found = 850.6209.

(2*R*)-3-((2,3-bis(Palmitoyloxy)propyl)thio)-2-((2,2,2-trifluoroacetyl)-λ^4-azaneyl)propanoic acid (41)



Prepared according to procedure B using compound **39** (1.01 g, 1.23 mmol). The solvents evaporated *in vacuo* to yield a while solid (0.86 g, 92%). The isolated compound was in good agreement with the literature.[286]

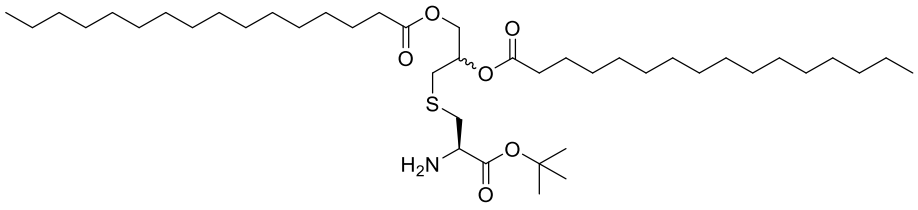
^1H NMR (400 MHz, $CDCl_3$) two diastereoisomers δ 5.21 – 5.13 (m, 1H, S-glyceryl CH), 4.39 – 4.31 (m, 1H, Cys αCH), 4.19 – 4.07 (m, 2H, S-glyceryl OCH_2), 3.33 – 3.17 (m, 1H, Cys βCH_aH_b), 3.14 – 2.94 (m, 1H, Cys βCH_aH_b), 2.89 – 2.68 (m, 2H, S-glyceryl CH_2), 2.36 – 2.26 (m, 4H, Pal αCH_2 x 2), 1.65 – 1.54 (m, 4H, Pal βCH_2 x 2), 1.36 – 1.19 (m, 48H, Pal CH_2 x 24), 0.88 (t, *J* = 6.8 Hz, 6H, Pal CH_3 x 2) ppm;

^{19}F NMR (376 MHz, $CDCl_3$) two diastereoisomers δ -75.56 ppm;

ν_{max} (thin film)/cm^{-1} 2955 (COOH), 2850 (CH), 1740 (Pal C=O), 1604 (N-H bend);

HRMS (*m/z* ESI^+) calcd. for $C_{38}H_{74}NO_6S$ = 672.5237 $(M + H)^+$. Found = 672.5231.

3-(((*R*)-2-Amino-3-(*tert*-butoxy)-3-oxopropyl)thio)propane-1,2-diyl dipalmitate (42)



Compound **39** (3.00 g, 3.62 mmol) was suspended in *t*BuOAc (40 mL) and conc. H_2SO_4 (0.58 mL, 10.9 mmol) was added, resulting in a pale-yellow solution. The reaction was

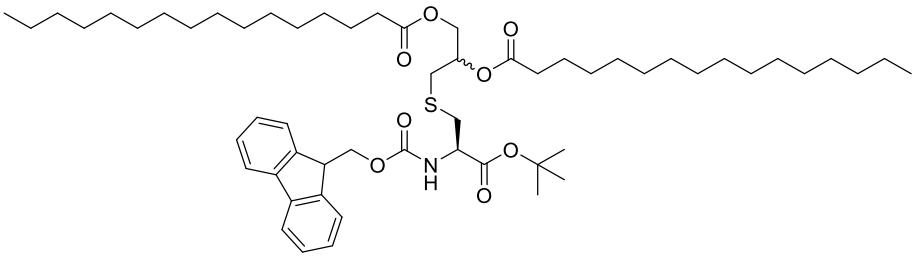
stoppered and stirred at rt for 16 h resulting in the formation of a white precipitate. The reaction mixture was neutralised to pH 7 with sat. aq. $NaHCO_3$ solution. The solid was filtered, washed with H_2O (5 x 50 mL) and dried *in vacuo* to yield a white solid (2.76 g, 92%). The isolated compound was in good agreement with the literature.[287]

R_f = 0.30 (toluene:acetone, 9:1 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) two diastereoisomers δ 5.23 – 5.18 (m, 1H, S-glyceryl CH), 4.34 (dd, *J* = 12.0 Hz, *J* = 3.5 Hz, 1H, S-glyceryl OC$\underline{H_a}H_b$), 4.16 (dd, *J* = 12.0 Hz, *J* = 3.5 Hz, 1H, S-glyceryl OC$H_a\underline{H_b}$), 3.52 (m, 1H, Cys αCH), 2.96 – 2.85 (m, 1H, Cys βC$\underline{H_a}H_b$), 2.82 – 2.77 (m, 1H, Cys βC$H_a\underline{H_b}$), 2.76 (d, *J* = 6.4 Hz, 2H, S-glyceryl CH_2), 2.32 – 2.35 (m, 4H, Pal αCH_2 x 2), 1.64 – 1.55 (m, 4H, Pal βC$\underline{H_2}$ x 2), 1.46 (s, 9H, O*t*Bu CH_3 x 3), 1.33 – 1.19 (m, 48H, Pal CH_2 x 24), 0.86 (t, *J* = 6.7 Hz, 6H, Pal CH_3 x 2) ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{42}H_{82}NO_6S$ = 728.5863 $(M + H)^+$. Found = 728.5847.

3-(((*R*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(*tert*-butoxy)-3 oxopropyl)thio)propane-1,2-diyl dipalmitate (43)



To a solution of compound **42** (2.75 g, 3.33 mmol) in CH_2Cl_2 (70 mL) was added Fmoc-OSu (1.26 g, 3.33 mmol) and DIPEA (1.16 mL, 6.66 mmol) and the reaction was stirred at rt for 16 h. The solvents were evaporated *in vacuo* and the crude product was purified by column chromatography (SiO_2, EtOAc:hex, 1:9 v/v) to yield a white solid (2.81 g, 89%). The isolated compound was in good agreement with the literature.[288]

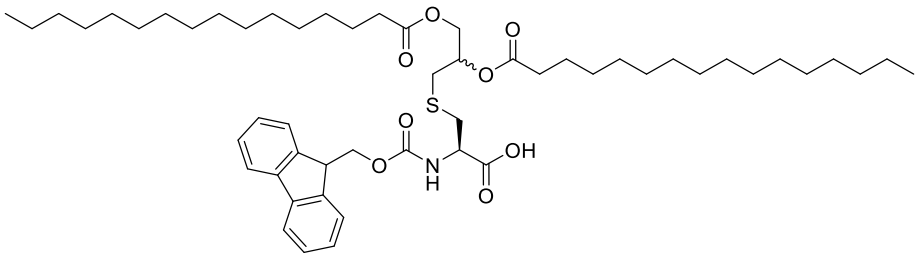
R_f = 0.43 (EtOAc:hex, 2:8 v/v);

^{1}H NMR (600 MHz, $CDCl_3$) two diastereoisomers δ 7.78 (d, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.64 (d, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.42 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.34 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 5.76 – 5.63 (m, 1H, NH), 5.20 – 5.15 (m, 1H, S-glyceryl CH), 4.55 – 4.51 (m, 1H, Cys αCH), 4.46 – 4.33 (m, 3H, Fmoc-CH_2, S-glyceryl OC$\underline{H_a}H_b$), 4.26 (t, *J* = 8.2 Hz, 1H, Fmoc-CH), 4.19 – 4.15 (m, 1H, *S*-glyceryl OC$H_a\underline{H_b}$), 3.13 – 3.01

(m, 2H, Cys βCH$_2$), 2.85 – 2.73 (m, 2H, S-glyceryl CH$_2$), 2.34 – 2.28 (m, 4H, Pal αCH$_2$ x 2), 1.68 – 1.55 (m, 4H, Pal βCH$_2$ x 2), 1.51 (s, 9H, O*t*Bu CH$_3$ x 3), 1.34 – 1.24 (m, 48H, Pal CH$_2$ x 24), 0.90 (t, *J* = 7.0 Hz, 6H, Pal CH$_3$ x 2) ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{57}H_{91}NO_8SNa$ = 972.6363 $(M + Na)^+$. Found = 972.6342.

_N_-(((9_H_-Fluoren-9-yl)methoxy)carbonyl)-_S_-(2,3-bis(palmitoyloxy)propyl)-L-cysteine (28)



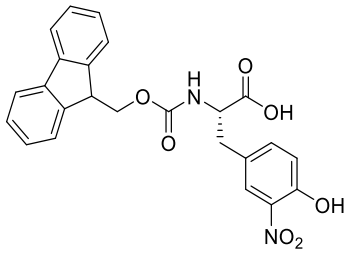
Prepared according to procedure B using compound **43** (1.32 g, 1.30 mmol). The solvents were removed *in vacuo* to yield a white solid (1.08 g, 93%). The isolated compound was in good agreement with the literature.[288]

R_f = 0.53 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) two diastereoisomers δ 7.76 (d, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.61 (d, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.40 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.31 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 5.79 – 5.73 (m, 1H, NH), 5.20 – 5.12 (m, 1H, S-glyceryl CH), 4.71 – 4.61 (m, 1H, Cys αCH), 4.47 – 4.30 (m, 3H, Fmoc-CH$_2$, S-glyceryl OC$\underline{H_a}H_b$), 4.24 (t, *J* = 7.0 Hz, 1H, Fmoc-CH), 4.20 – 4.02 (m, 1H, S-glyceryl OCH$_a\underline{H_b}$), 3.22 – 3.04 (m, 2H, Cys βCH$_2$), 2.86 – 2.64 (m, 2H, S-glyceryl CH$_2$), 2.38 – 2.24 (m, 4H, Pal αCH$_2$ x 2), 1.64 – 1.55 (m, 4H, Pal βCH$_2$ x 2), 1.34 – 1.18 (m, 48H, Pal CH$_2$ x 24), 0.88 (t, *J* = 6.7 Hz, 6H, Pal CH$_3$ x 2) ppm;

HRMS (*m/z* APCI$^+$) calcd. for $C_{53}H_{84}NO_8S$ = 894.5912 $(M + H)^+$. Found = 894.5921.

(*S*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-hydroxy-3-nitrophenyl)propanoic acid (44)



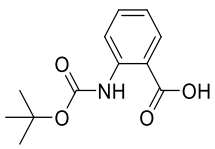
To a suspension of 3-nitrotyrosine (2.00 g, 8.85 mmol) in H_2O (150 mL) at 0 °C was added a solution of Fmoc-OSu (2.99 g, 8.85 mmol) in acetone (100 mL). The pH of the reaction was adjusted to pH 9 using aq. $NaHCO_3$ solution (10% w/v). The reaction was stirred at rt for 16 h resulting in the formation of a clear yellow solution. Acetone was removed under reduced pressure and the aq. layer was washed with Et_2O (3 x 100 mL). The aq. layer was acidified to pH 3 resulting in the precipitation of a yellow solid. This was extracted with EtOAc (5 x 50 mL) and the combined organic layers were washed with water (3 x 100 mL) and sat. aq. NaCl solution (3 x 100 mL), dried over $MgSO_4$, and filtered. The solvents were removed *in vacuo* yielding the product as a yellow solid (2.96 g, 77%). The isolated compound was in good agreement with the literature.[289]

R_f = 0.22 (CH_2Cl_2:MeOH, 9:1 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.48 (s, 1H, Phenol-OH), 7.91 (s, 1H, Tyr Ar-CH), 7.76 (d, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.54 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.39 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.34 – 7.27 (m, 3H, Fmoc Ar-CH x 2, Tyr Ar-CH), 7.06 (d, *J* = 8.4 Hz, 1H, Tyr Ar-CH), 5.25 (d, *J* = 7.7 Hz, 1H, NH), 4.70 – 4.64 (m, 1H, Tyr αCH), 4.54 – 4.35 (m, 2H, Fmoc CH_2), 4.19 (t, *J* = 6.4 Hz, 1H, Fmoc CH), 3.25 – 3.00 (m, 2H, Tyr βCH_2) ppm;

m/z HRMS (ESI$^+$) calcd. for $C_{24}H_{20}N_2NaO_7$ = 471.1163 $(M + Na)^+$. Found = 471.1167.

2-((*tert*-Butoxycarbonyl)amino)benzoic acid (45)



To a solution of 2-aminobenzoic acid (5.00 g, 36.5 mmol) in THF:H_2O (1:1 v/v, 80 mL) was added aq. NaOH solution (2 M) until pH 10 was reached. Boc_2O (8.75 g, 40.1 mmol)

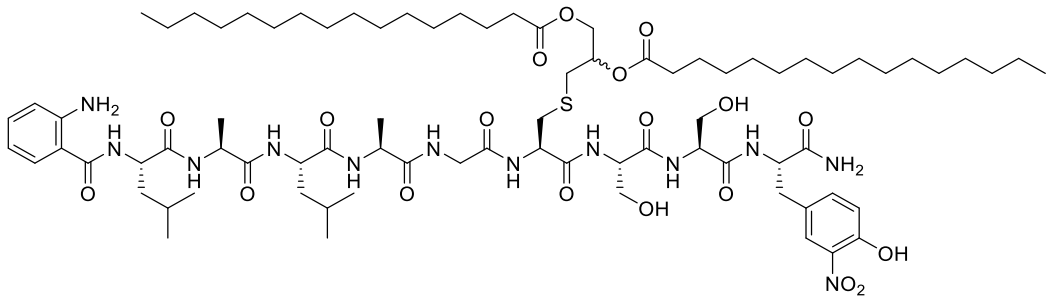
was added and the reaction was stirred at rt for 16 h. THF was evaporated under reduced pressure and the aq. layer acidified to pH 4 using aq. citric acid solution (15% w/v) resulting in the formation of a white precipitate. The precipitate was filtered and washed with H_2O (3 x 50 mL). The solid was dissolved in EtOAc (100 mL) and washed with H_2O (3 x 100 mL) and sat. aq. NaCl solution (100 mL). The combined organic layers were dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* to give the product as a white solid (6.41 g, 74%). The isolated compound was in good agreement with the literature.[197]

R_f = 0.56 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.00 (s, 1H, NH), 8.45 (d, *J* = 8.6 Hz, 1H, Ar H6), 8.10 (dd, *J* = 8.0 Hz, *J* = 1.5 Hz, 1H, Ar H3), 7.55 (dt, *J* = 8.0 Hz, *J* = 1.5 Hz, 1H, Ar H4), 7.06 – 6.97 (m, 1H, Ar H5), 1.54 (s, 9H, Boc CH_3 x 3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{12}H_{15}NNaO_4$ = 260.0893$(M + Na)^+$. Found = 260.0898.

(3*S*,6*S*,9*S*,12*S*,18*R*)-18-(((*S*)-1-(((*S*)-1-(((*S*)-1-Amino-3-(4-hydroxy-3-nitrophenyl)-1-oxopropan-2-yl)amino)-3-hydroxy-1-oxopropan-2-yl)amino)-3-hydroxy-1-oxopropan-2-yl)carbamoyl)-1-(2-aminophenyl)-3,9-diisobutyl-6,12-dimethyl-1,4,7,10,13,16-hexaoxo-20-thia-2,5,8,11,14,17-hexaazatricosane-22,23-diyl dipalmitate (27)



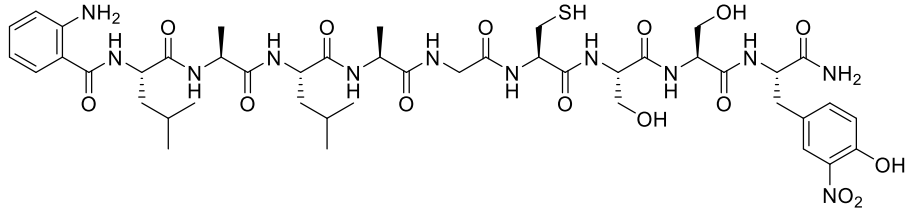
Fmoc-Tyr(3-NO_2)-OH **44** (251 mg, 0.56 mmol) was loaded to Rink amide resin (200 mg, 0.140 mmol) as per procedure E. The coupling of Fmoc-Ser(*t*Bu)-OH (161 mg, 0.42 mmol) x 2, Fmoc-Cys[(*R*/*S*)Pam_2]-OH **28** (376 mg, 0.42 mmol), Fmoc-Gly-OH (125 mg, 0.42 mmol), Fmoc-Ala-OH (131 mg, 0.42 mmol) x 2, Fmoc-Leu-OH (148 mg, 0.42 mmol) x 2, and Boc-Abz-OH **45** (100 mg, 0.42 mmol) was performed as per procedure I. The resin was treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 15 min prior to cleavage. The decapeptide was cleaved from the resin using procedure J affording a

yellow solid (60 mg, 27% crude yield). Analytical HPLC (C_{18} Phenomenex Clarity Column 5 µm Oligo-RP 150 x 4.60 mm; gradient = 50% MeCN (+0.1% TFA) with 50% H_2O (0.1% TFA) to 100% MeCN (+0.1% TFA) over 50 min; detection wavelength = 280 nm; retention time = 28.30 min) revealed a crude purity of 65%. The peptide was purified under the same HPLC conditions yielding a yellow solid (0.8 mg, 0.3%).

Retention time = 27.90 min;

HRMS (*m/z* MALDI$^+$) calcd. for $C_{80}H_{132}N_{12}O_{19}NaS$ = 1619.9350 $(M + Na)^+$. Found = 1619.9396.

2-Amino-*N*-((2*S*,5*S*,8*S*,11*R*,17*S*,20*S*,23*S*,26*S*)-1-amino-2-(4-hydroxy-3-nitrobenzyl)-5,8-bis(hydroxymethyl)-20-isobutyl-11-(mercaptomethyl)-17,23,28-trimethyl-1,4,7,10,13,16,19,22,25-nonaoxo-3,6,9,12,15,18,21,24-octaazanonacosan-26-yl)benzamide (46)

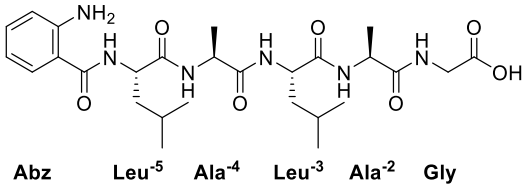


Fmoc-Tyr(3-NO_2)-OH **44** (514 mg, 1.15 mmol) was loaded to Rink amide AM resin (409 mg, 0.29 mmol) as per procedure E. Fmoc-Ser(*t*Bu)-OH (330 mg, 0.86 mmol) x 2 was coupled as per procedure I. Fmoc-Cys(Trt)-OH (503 mg, 0.86 mmol) was coupled using DIC (133 µL, 0.86 mmol) and HOBt (116 mg, 0.86 mmol) in DMF:CH_2Cl_2 (5 mL, 1:1 v/v). The resin was then treated with 20% (v/v) piperidine/DMF (5 mL) for 2 x 10 min followed by coupling of Fmoc-Gly-OH (256 mg, 0.86 mmol), Fmoc-Ala-OH (268 mg, 0.86 mmol) x 2, Fmoc-Leu-OH (304 mg, 0.86 mmol) x 2 and Boc-Abz-OH **45** (204 mg, 0.86 mmol) as per procedure I. The resin was treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 15 min prior to cleavage. The decapeptide was cleaved from the resin using procedure J yielding a yellow solid. The peptide was purified by C_{18} semi-preparative HPLC with a gradient of 5% MeCN (+0.1% TFA) with 95% H_2O (0.1% TFA) to 95% MeCN (+0.1% TFA) with 5% H_2O (0.1% TFA) over 43 min; detection wavelength = 280 nm; retention time = 22.5 min) to give **46** as a yellow solid (36 mg, 12%). The compound was analysed using with the same conditions on a C_{18} Phenomenex Gemini 5 µm, 110 Å, 250 x 4.6 mm LC column.

Retention time = 29.30 min;

HRMS (*m/z* ESI$^-$) calcd. for $C_{45}H_{65}N_{12}O_{15}S$ = 1045.4419 $(M - H)^-$. Found = 1045.4395.

(2-Aminobenzoyl)-L-leucyl-L-alanyl-L-leucyl-L-alanylglycine (47)



Fmoc-Gly-OH (292 mg, 1.78 mmol) was loaded to Wang resin (318 mg, 0.45 mmol) as per procedure F. Fmoc-Ala-OH (416 mg, 1.34 mmol) x 2, Fmoc-Leu-OH (471 mg, 1.34 mmol) x 2 and Boc-Abz-OH **45** (317 mg, 1.34 mmol) were coupled as per procedure I. The peptide was cleavage form the resin as per procedure J and purified by column chromatography (SiO_2, EtOAc:EtOH:H_2O 7:2:1 v/v/v) to yield the product as a white solid (30.0 mg, 12%).

R_f = 0.33 (EtOAc:EtOH:H_2O 7:2:1 v/v/v);

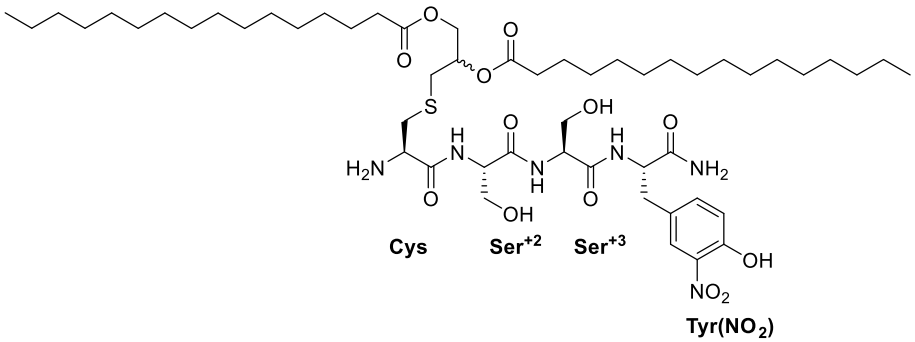
^{1}H NMR (600 MHz, MeOD) δ 7.56 (dd, *J* = 7.9, 1.0 Hz, 1H, Ar H6), 7.21 (t, *J* = 7.9 Hz, 1H, Ar H5), 6.77 (d, *J* = 7.9 Hz, 4H, Ar H3), 6.66 (t, *J* = 7.9 Hz, 1H, Ar H4), 4.54 (dd, *J* = 9.5 Hz, *J* = 5.4 Hz, 1H, Leu^{-5} αCH), 4.44 – 4.34 (m, 3H, Leu^{-3} αCH, Ala^{-4} αCH, Ala^{-2} αCH), 3.79 (d, *J* = 3.5 Hz, 2H, Gly CH_2), 1.81 – 1.72 (m, 2H, Leu^{-5} βC$\underline{H_a}H_b$, Leu^{-5} CH), 1.72 – 1.60 (m, 4H, Leu^{-5} βC$H_a\underline{H_b}$, Leu^{-3} βCH_2, Leu^{-3} CH), 1.40 (dd, *J* = 7.1, 2.4 Hz, 6H, Ala^{-4} CH_3, Ala^{-2} CH_3), 1.00 (dd, *J* = 16.0, 6.4 Hz, 6H, Leu^{-5} CH_3 x 2), 0.94 – 0.89 (m, 6H, Leu^{-3} CH_3 x 2) ppm;

^{13}C NMR (151 MHz, MeOD) δ 174.5 (Leu^{-5} CO), 174.1 (Leu^{-3} CO), 173.5 (Ala CO), 173.4 (Ala CO), 171.2 (Abz CO), 170.6 (Gly CO), 149.5 (Ar qC-NH_2), 132.4 (Ar C5), 128.3 (Ar C6), 117.1 (Ar C3), 116.3 (Ar C4), 115.9 (Ar q$\underline{C}$-CO), 52.9 (Leu^{-5} αCH), 52.4 (Leu^{-3} αCH), 49.7 (Ala αCH), 49.4 (Ala αCH), 42.7 (Gly CH_2), 40.5 (Leu CH_2 x 2), 25.1 (Leu^{-5} CH), 24.8 (Leu^{-3} CH), 22.5 (Leu^{-5} CH_3 x1), 22.3 (Leu^{-3} CH_3 x1), 21.2 (Leu^{-3} CH_3 x1), 20.8 (Leu^{-5} CH_3 x1), 17.0 (Ala CH_3), 16.5 (Ala CH_3) ppm;

HRMS (*m/z* MALDI$^+$) calcd. for $C_{27}H_{42}N_6O_7Na$ = 585.3013 $(M + Na)^+$. Found = 585.3041;

ν_{max} (thin film)/cm^{-1} 3280 (NH), 2956 (CH), 1627 (amide C=O), 1159 (C-O).

(2*S*,5*S*,8*S*,11*R*)-1,11-Diamino-2-(4-hydroxy-3-nitrobenzyl)-5,8-bis(hydroxymethyl)-1,4,7,10-tetraoxo-13-thia-3,6,9-triazahexadecane-15,16-diyl dipalmitate (48)



Fmoc-Tyr(3-NO_2)-OH **44** (270 mg, 0.60 mmol) was loaded to Rink amide AM resin (216 mg, 0.15 mmol) as per procedure E. Fmoc-Ser(*t*Bu)-OH (174 mg, 0.45 mmol) x 2 and Fmoc-Cys[(*R*/*S*)Pam]$_2$-OH **28** (405 mg, 0.45 mmol) were coupled as per procedure I. The resin was treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 15 min prior to cleavage. The tetrapeptide was cleaved from the resin as per procedure J. The crude peptide was purified using column chromatography (SiO_2, EtOAc:EtOH 8:2 v/v) affording the peptide as a yellow solid (21.0 mg, 13%).

R_f = 0.33 (EtOAc:EtOH:H_2O 7.5:2:0.5 v/v/v);

^{1}H NMR (600 MHz, DMSO-d_6) two diastereoisomers δ 8.28 – 8.18 (m, 1H, Ser NH), 8.18 – 8.13 (m, 1H, Ser NH), 7.94 (d, *J* = 7.9 Hz, 1H, Tyr NH), 7.75 (s, 1H, Tyr Ar-CH), 7.40 (d, *J* = 7.1 Hz, 1H, Tyr Ar-CH), 7.29 – 7.20 (m, 2H, Cys NH_2), 7.03 (d, *J* = 8.4 Hz, 1H, Tyr Ar-CH), 5.27 – 5.17 (m, 1H, Ser OH), 5.14 – 5.08 (m, 1H, S-glyceryl CH), 4.97 (t, *J* = 5.7 Hz, 1H, Ser OH), 4.42 – 4.33 (m, 2H, Ser αCH, Tyr αCH), 4.31 (d, *J* = 12.4 Hz, 1H, S-glyceryl OC$\underline{H_a}H_b$), 4.24 – 4.20 (m, 1H, Ser^{+2} αCH), 4.15 – 4.07 (m, 1H, S-glyceryl OC$H_a\underline{H_b}$), 3.70 – 3.65 (m, 1H, Ser^{+2} C$\underline{H_a}H_b$), 3.58 – 3.48 (m, 3H, Ser^{+2} C$H_a\underline{H_b}$, Ser^{+3} CH_2), 3.45 – 3.41 (m, 1H, Cys αCH), 3.05 (dd, *J* = 14.3, 4.5 Hz, 1H, Tyr C$\underline{H_a}H_b$), 2.92 – 2.88 (m, 1H, Cys C$\underline{H_a}H_b$), 2.93 – 2.84 (m, 1H, S-glyceryl C$\underline{H_a}H_b$), 2.83 – 2.74 (m, 1H, Tyr C$H_a\underline{H_b}$), 2.73 – 2.66 (m, 2H, S-glyceryl C$H_a\underline{H_b}$, Cys C$H_a\underline{H_b}$), 2.30 – 2.25 (m, 4H, Pal αCH_2 x 2), 1.56 – 1.46 (m, 4H, Pal βCH_2 x 2), 1.32 – 1.12 (m, 48H, Pal CH_2 x 24), 0.87 (t, *J* = 6.9 Hz, 6H, Pal CH_3 x 2) ppm;

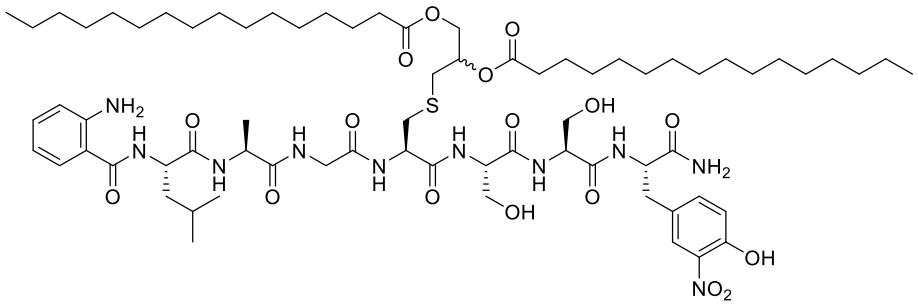
^{13}C NMR (151 MHz, DMSO-d_6) two diastereoisomers δ 173.0, 173.0 (Pal C=O), 172.8 (Cys C=O), 170.9 (Tyr C=O), 170.1 (Ser C=O, Ser C=O), 151.7 (Tyr qC-ArOH), 136.7 (Tyr Ar CH), 136.5 (Tyr qC-Ar), 129.5 (Tyr qC-Ar), 125.8 (Tyr Ar-CH), 119.5 (Tyr Ar-

CH), 70.6 (S-glyceryl CH), 64.0 (S-glyceryl OCH_2), 62.4 (Ser CH_2), 61.7 (Ser CH_2), 56.2 (Ser αCH), 54.9 (Tyr αCH), 54.6 (Cys αCH), 54.4 (Ser αCH), 37.7 (Cys CH_2), 36.3 (Tyr CH_2), 34.1, 33.9 (Pal CH_2), 31.8 (S-glyceryl CH_2), 29.6, 29.5, 29.4, 29.2 (Pal CH_2), 28.9 (Pal CH_2), 25.0, 24.9 (Pal βCH_2), 22.6 (Pal CH_2), 14.4 (Pal CH_3) ppm;

HRMS (*m/z* ESI^+) calcd. for $C_{53}H_{93}N_6O_{13}S$ = 1053.6516 $(M + H)^+$. Found = 1053.6530;

ν_{max} (thin film)/cm^{-1} 3393 (NH), 1655 (Amide C=O), 1023 (Ester C-O).

(3*S*,6*S*,12*R*)-12-(((*S*)-1-(((*S*)-1-(((*S*)-1-Amino-3-(4-hydroxy-3-nitrophenyl)-1-oxopropan-2-yl)amino)-3-hydroxy-1-oxopropan-2-yl)amino)-3-hydroxy-1-oxopropan-2-yl)carbamoyl)-1-(2-aminophenyl)-3-isobutyl-6-methyl-1,4,7,10-tetraoxo-14-thia-2,5,8,11-tetraazaheptadecane-16,17-diyl dipalmitate (49)



Fmoc-Tyr(3-NO_2)-OH **44** (86 mg, 0.19 mmol) was loaded to Rink amide resin (69 mg, 0.048 mmol) as per procedure E. The coupling of Fmoc-Ser(*t*Bu)-OH (55 mg, 0.14 mmol) x 2, Fmoc-Cys[(*R*/*S*)Pam_2] **28** (107 mg, 0.14 mmol), Fmoc-Gly-OH (43 mg, 0.14 mmol), Fmoc-Ala-OH (45 mg, 0.14 mmol), Fmoc-Leu-OH (51 mg, 0.14 mmol) and Boc-Abz-OH **45** (46 mg, 0.14 mmol) was performed as per procedure I. The resin was treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 15 min prior to cleavage. The peptide was cleaved from the resin using procedure J affording a yellow solid which was purified by column chromatography (SiO_2, MeOH:CH_2Cl_2 1:9 v/v) to yield the product as a yellow solid (5 mg, 8%).

R_f = 0.32 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (600 MHz, DMSO-d_6) two diastereoisomers δ 8.23 – 8.19 (m, 1H, Ser^{+2} NH), 8.16 – 8.08 (m, 3H, Leu NH, Gly NH, Cys NH), 8.07 – 8.04 (m, 1H, Ser^{+3} NH), 8.01 – 7.98 (m, 1H, Ala NH), 7.91 (d, *J* = 8.7 Hz, 1H, Tyr NH), 7.70 – 7.66 (m, 1H, Tyr Ar-CH), 7.55 (d, *J* = 7.8 Hz, 1H, Abz Ar-CH), 7.27 – 7.24 (m, 1H, Tyr Ar-CH), 7.23 - 7.20 (m, 2H, $CONH_2$) 7.13 (t, *J* = 7.6 Hz, 1H, Abz Ar-CH), 6.68 (d, *J* = 8.5 Hz, 1H, Abz Ar-

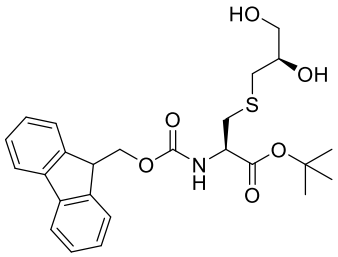
CH), 6.51 (t, *J* = 7.4 Hz, 1H, Abz Ar-CH), 6.32 (s, 2H, Abz NH_2), 5.27 – 5.15 (m, 1H, Ser^{+2} OH), 5.15 – 5.02 (m, 1H, S-glyceryl CH), 5.00 – 4.91 (m, 1H, Ser^{+3} OH), 4.58 – 4.50 (m, 1H, Cys αCH), 4.50 – 4.42 (m, 1H, Leu αCH), 4.39 – 4.26 (m, 4H, Ser^{+2} αCH, Tyr αCH, Ala αCH, S-glyceryl O$C\underline{H_a}H_b$), 4.22 – 4.18 (m, 1H, Ser^{+3} αCH), 4.13 – 4.05 (m, 1H, S-glyceryl O$CH_a\underline{H_b}$), 3.80 – 3.72 (m, 2H, Gly CH_2), 3.67 – 3.58 (m, 1H, Ser^{+2} $C\underline{H_a}H_b$), 3.58 – 3.52 (m, 2H, Ser^{+2} $CH_a\underline{H_b}$, Ser^{+3} $C\underline{H_a}H_b$), 3.52 – 3.46 (m, 1H, Ser^{+2} $CH_a\underline{H_b}$), 3.05 – 2.94 (m, 1H, Tyr $C\underline{H_a}H_b$), 2.94 – 2.87 (m, 1H, Cys $C\underline{H_a}H_b$), 2.87 – 2.77 (m, 2H, Cys $CH_a\underline{H_b}$, S-glyceryl $C\underline{H_a}H_b$), 2.74 – 2.65 (m, 2H, Tyr $CH_a\underline{H_b}$, S-glyceryl $CH_a\underline{H_b}$), 2.30 – 2.22 (m, 4H, Pal αCH_2 x 2), 1.74 – 1.61 (m, 3H, Leu CH_2, Leu $C\underline{H}$-CH_3), 1.57 – 1.43 (m, 7H, Pal βCH_2 x 2, Ala CH_3), 1.32 – 1.13 (m, 48H, Pal CH_2 x 24), 0.95 – 0.77 (m, 12H, Leu CH_3 x 2, Pal CH_3 x 2) ppm;

^{13}C NMR (151 MHz, DMSO-d_6) two diastereoisomers δ 172.7, 172.5 (Pal C=O), 172.3 ($CONH_2$), 172.2 (Leu C=O), 172.1 (Ala C=O), 170.2 (Ser^{+2} C=O), 170.1 (Cys C=O), 169.6 (Ser^{+3} C=O), 168.9 (Abz C=O), 168.7 (Gly C=O), 149.6 (Abz qC NH_2), 136.1 (Tyr Ar-CH), 135.8 (Tyr Ar-CH), 131.8 (Abz Ar-CH), 128.5 (Abz Ar-CH), 125.5 (Tyr Ar-CH), 116.3 (Abz Ar-CH), 114.9 (Abz Ar-CH), 114.5 (Abz Ar-CH), 69.9 (S-glyceryl CH), 63.5 (S-glyceryl OCH_2), 61.6 (Ser^{+2} βCH_2), 61.4 (Ser^{+3} βCH_2), 55.6 (Ser^{+3} αCH), 54.9 (Ser^{+2} αCH), 54.0 (Tyr αCH), 52.2 (Cys αCH), 52.1 (Cys αCH), 51.3 (Leu αCH), 48.3 (Ala αCH), 41.8 (Gly CH_2), 36.0 (Tyr βCH_2), 34.2 (Cys βCH_2), 33.9 (Pal αCH_2), 33.6 (Pal αCH_2), 33.4 (Pal CH_2), 31.6 (S-glyceryl CH_2 [diastereoisomer A]), 31.6 (S-glyceryl CH_2 [diastereoisomer B]), 29.1, 29.0, 28.9, 28.8, 28.8, 28.7, 28.5, 28.4 (Pal CH_2), 24.5 (Leu CH), 24.4 (Pal CH_2), 23.2 (Leu CH_3), 22.1 (Pal CH_2), 21.3 (Leu CH_3), 18.3 (Ala CH_3), 13.9 (Pal CH_3);

HRMS (*m/z* MALDI$^+$) calcd. for $C_{71}H_{116}N_{10}O_{17}SNa$ = 1435.8138 $(M + Na)^+$. Found = 1435.8182;

ν_{max} (thin film)/cm^{-1} 3282 (NH), 2921 (CH_2), 1746, 1629 (Amide I: CO, CN), 1533 (Amide II: CN, NH), 1370 (Amide III: CN, CO, CC), 1214 (CO).

***tert*-Butyl-*N*-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-*S*-((*R*)-2,3-dihydroxypropyl)-L-cysteinate (40)**



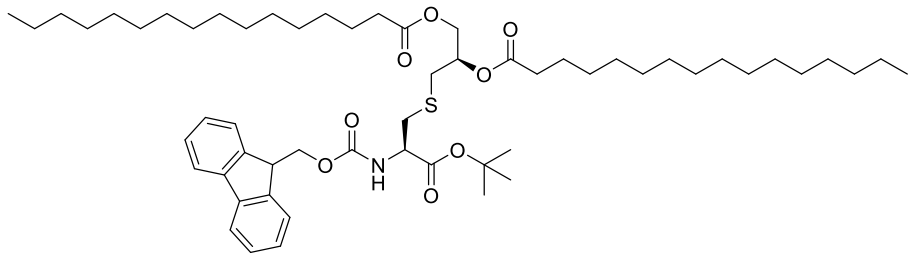
Disulfide reduction was performed as per procedure A using (Fmoc-Cys-O*t*Bu)$_2$ **32** (11.90 g, 14.97 mmol), THF (200 mL), PBu_3 (3.68 mL, 14.93 mmol) and H_2O (40 mL). After 4 h the solvents were evaporated under reduced pressure and anhydrous CH_2Cl_2 (200 mL) was added under argon. To this solution was added (*R*)-(+)-glycidol (9.96 mL, 149.3 mmol) and DIPEA (5.20 mL, 29.86 mmol). The reaction was stirred at 40 °C for 16 h. The reaction was washed with H_2O (3 x 100 mL) and sat. aq. NaCl solution (3 x 100 mL) and the organic layer was dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the product was purified using column chromatography (SiO_2, EtOAc:hex 1:1 – 4:1 v/v) to afford a viscous, colourless oil (9.85 g, 65%). The isolated compound was in good agreement with the literature.[290]

R_f = 0.36 (EtOAc:hex 4:1 v/v);

^{1}H NMR (400 MHz, acetone-d_6) δ 7.86 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.73 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.42 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.33 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 6.87 (d, *J* = 8.3 Hz, 1H, NH), 4.42 – 4.30 (m, 3H, Fmoc CH_2, Cys αCH), 4.30 – 4.21 (m, 1H, Fmoc CH), 4.00 (d, *J* = 4.4 Hz, 1H, CH-O<u>H</u>), 3.84 –3.76 (m, 1H, S-glyceryl CH), 3.74 – 3.66 (m, 1H, CH_2-O<u>H</u>), 3.62 – 3.51 (m, 2H, S-glyceryl OCH_2), 3.08 (dd, *J* = 13.3, 5.0 Hz, 1H, Cys βC<u>H$_a$</u>H$_b$), 2.97 (dd, *J* = 13.3, 5.0 Hz, 1H, Cys βCH$_a$<u>H$_b$</u>), 2.87 – 2.82 (m, 1H, S-glyceryl C<u>H$_a$</u>H$_b$), 2.66 (dd, *J* = 13.7, 7.9 Hz, 1H, S-glyceryl CH$_a$<u>H$_b$</u>), 1.47 (s, 9H, *t*Bu CH_3 x 3) ppm;

HRMS (*m/z* APCI$^-$) calc. $C_{25}H_{31}ClNO_6S$ = 508.1566 (M + Cl)$^-$. Found = 508.1551.

(*R*)-3-(((*R*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(*tert*-butoxy)-3-oxopropyl)thio)propane- 1,2-diyl dipalmitate (53)



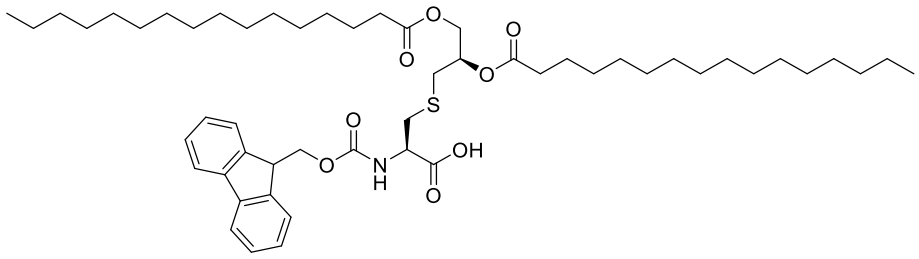
To a solution of palmitic acid (0.88 g, 3.42 mmol) in anhydrous CH_2Cl_2 (50 mL) under argon at 0 °C was added EDC·HCl (0.66 g, 3.42 mmol) and DMAP (0.033 g, 0.27 mmol). The reaction was stirred at 0 °C for 1 h. Diol **40** (0.69 g, 1.37 mmol) was dissolved in anhydrous CH_2Cl_2 (10 mL) and added to the reaction which was stirred at rt under argon for 16 h. The solvent was removed *in vacuo* and the product was purified using column chromatography (SiO_2, hex - EtOAc:hex 1:9 v/v) to yield a white solid (1.22 g, 94%). The isolated compound was in good agreement with the literature.[290]

R_f = 0.60 (EtOAc:hex 1:4 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.62 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.40 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.31 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 5.69 (d, *J* = 7.6 Hz, 1H, NH), 5.20 - 5.12 (m, 1H, *S*-glyceryl CH), 4.54 – 4.47 (m, 1H, Cys αCH), 4.44 – 4.30 (m, 3H, Fmoc-CH_2, S-glyceryl OC$\underline{H_a}H_b$), 4.24 (t, *J* = 7.1 Hz, 1H, Fmoc CH), 4.16 (dd, *J* = 11.8, 5.9 Hz, 1H, S-glyceryl OC$H_a\underline{H_b}$), 3.11 (dd, *J* = 13.7, 4.9 Hz, 1H, Cys βC$\underline{H_a}H_b$), 3.04 (dd, *J* = 13.7, 4.9 Hz, 1H, Cys βC$H_a\underline{H_b}$), 2.77 (d, *J* = 6.3 Hz, 2H, S-glyceryl CH_2), 2.33 – 2.26 (m, 4H, Pal αCH_2 x 2), 1.69 – 1.54 (m, 4H, Pal CH_2 x 2), 1.49 (s, 9H, *t*Bu CH_3 x 3), 1.35 – 1.20 (m, 48H, Pal CH_2 x 24), 0.88 (t, *J* = 6.8 Hz, 6H, Pal CH_3 x 2) ppm;

HRMS (*m/z* APCI$^+$) calcd. for $C_{57}H_{92}NO_8S$ = 950.6538 $(M + H)^+$. Found = 950.6556.

***N*-(((9*H*-Fluoren-9-yl)methoxy)carbonyl)-*S*-((*R*)-2,3-bis(palmitoyloxy)propyl)-L-cysteine (50)**



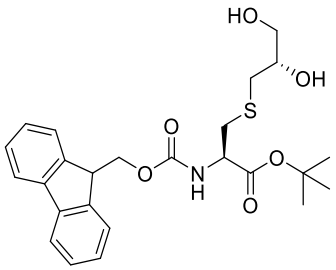
Prepared as per procedure B using compound **53** (1.22 g, 1.28 mmol). The solvents were removed *in vacuo* to yield the product as a white solid (1.13 g, 99%). The isolated compound was in good agreement with the literature.[290]

R_f = 0.53 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.60 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.40 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.31 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 5.76 (d, *J* = 7.8 Hz, 1H, NH), 5.21 – 5.13 (m, 1H, S-glyceryl CH), 4.69 – 4.62 (m, 1H, Cys αCH), 4.41 (d, *J* = 7.0 Hz, 2H, Fmoc CH_2), 4.35 (dd, *J* = 11.8, 3.4 Hz, 1H, S-glyceryl OC$\underline{H_a}H_b$), 4.24 (t, *J* = 7.0 Hz, 1H, Fmoc CH), 4.15 (dd, *J* = 11.8, 5.9 Hz, 1H, S-glyceryl OC$H_a\underline{H_b}$), 3.16 (dd, *J* = 13.4, 4.9 Hz, 1H, Cys βC$\underline{H_a}H_b$), 3.08 (dd, *J* = 13.4, 4.9 Hz, 1H, Cys βC$H_a\underline{H_b}$), 2.83 – 2.69 (m, 2H, S-glyceryl CH_2), 2.34 – 2.27 (m, 4H, Pal αCH_2 x 2), 1.69 – 1.54 (m, 4H, Pal βCH_2 x 2), 1.34 – 1.21 (m, 48H, Pal CH_2 x 24), 0.88 (t, *J* = 6.8 Hz, 6H, Pal CH_3 x 2) ppm;

m/z HRMS (APCI+) calcd. for $C_{53}H_{84}NO_8S$ = 894.5912 $(M + H)^+$. Found = 894.5927.

***tert*-Butyl-*N*-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-*S*-((*S*)-2,3-dihydroxypropyl)-L-cysteinate (52)**



Disulfide reduction was performed as per procedure A using (Fmoc-Cys-O*t*Bu)$_2$ **32** (0.600 g, 0.750 mmol), THF (10 mL), PBu_3 (0.187 mL, 0.750 mmol) and H_2O (2 mL).

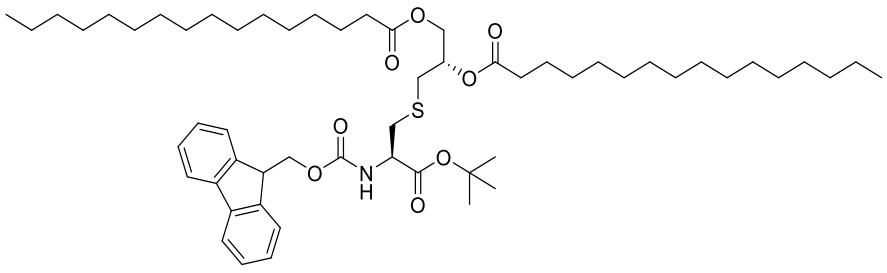
After 4 h the solvents were evaporated under reduced pressure, followed by addition of anhydrous CH_2Cl_2 (10 mL) under argon. To this solution was added (*S*)-(-)-glycidol (1.11 mL, 15.0 mmol) and DIPEA (0.260 mL, 1.50 mmol) under argon. The reaction was stirred at 40 °C for 16 h. The reaction was cooled to rt and washed with H_2O (3 x 15 mL) and sat. aq. NaCl solution (3 x 15 mL) and the organic layer was dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the product was purified using column chromatography (SiO_2, EtOAc:hex 1:1 – 4:1 v/v) to afford the product as a viscous, colourless oil (0.385 g, 51%). The isolated compound was in good agreement with the literature.[290]

R_f = 0.36 (EtOAc:hex 4:1 v/v);

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.61 (d, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.41 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 7.32 (t, *J* = 7.4 Hz, 2H, Fmoc Ar-CH), 5.70 (d, *J* = 7.7 Hz, 1H, NH), 4.53 – 4.47 (m, 1H, Cys αCH), 4.47 – 4.34 (m, 2H, Fmoc CH_2), 4.27 – 4.24 (t, *J* = 7.3 Hz, 1H, Fmoc CH), 3.83 – 3.78 (m, 1H, S-glyceryl CH), 3.69 (dd, *J* = 11.1, 3.2 Hz, 1H, S-glyceryl OC$\underline{H}_aH_b$), 3.56 (dd, *J* = 11.1, 5.6 Hz, 1H, S-glyceryl OC$H_a\underline{H}_b$), 3.05 – 2.99 (m, 1H, Cys βC$\underline{H}_aH_b$), 2.97 – 2.92 (m, 1H, Cys βC$H_a\underline{H}_b$), 2.82 – 2.76 (m, 1H, S-glyceryl C$\underline{H}_aH_b$), 2.70 – 2.64 (m, 1H, S-glyceryl C$H_a\underline{H}_b$), 1.50 (s, 9H, *t*Bu CH_3 x 3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{25}H_{31}NNaO_6S$ = 496.1764 $(M + Na)^+$. Found = 496.1772.

(*S*)-3-(((*R*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(*tert*-butoxy)-3-oxopropyl)thio)propane-1,2-diyl dipalmitate (54)



To a solution of palmitic acid (0.387 g, 1.51 mmol) in anhydrous CH_2Cl_2 (10 mL) under argon at 0 °C was added EDC·HCl (0.270 g, 1.51 mmol) and DMAP (0.015 g, 0.12 mmol). The reaction was stirred at 0 °C for 1 h. Diol **52** (0.305 g, 0.605 mmol) was dissolved in anhydrous CH_2Cl_2 (10 mL) and added to the reaction which was stirred at rt under argon for 16 h. The solvent was removed *in vacuo* and the product was purified

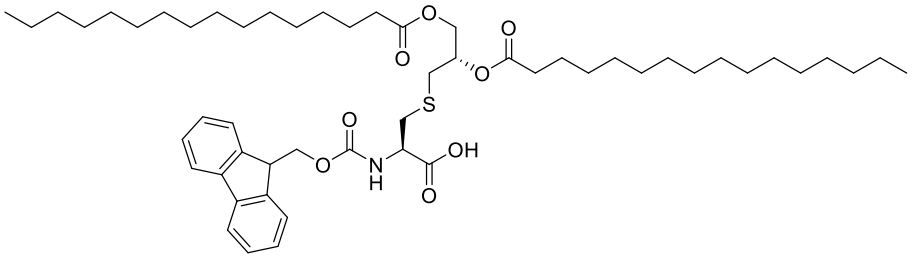
using column chromatography (SiO_2, hex - EtOAc:hex 1:9 v/v) to yield the product as a white solid (0.395 g, 69%). The isolated compound was in good agreement with the literature.[290]

R_f = 0.60 (EtOAc:hex 1:4 v/v);

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.3 Hz, 2H, Fmoc Ar-CH), 7.62 (d, *J* = 7.3 Hz, 2H, Fmoc Ar-CH), 7.40 (t, *J* = 7.3 Hz, 2H, Fmoc Ar-CH), 7.31 (t, *J* = 7.3 Hz, 2H, Fmoc Ar), 5.67 (d, *J* = 7.3 Hz, 1H, NH), 5.18 – 5.13 (m, 1H, S-glyceryl CH), 4.54 – 4.48 (m, 1H, Cys αCH), 4.43 – 4.30 (m, 3H, Fmoc CH_2, S-glyceryl OC$\underline{H}_aH_b$), 4.24 (t, *J* = 7.1 Hz, 1H, Fmoc CH), 4.15 (dd, *J* = 11.9, 5.9 Hz, 1H, S-glyceryl OC$H_a\underline{H}_b$), 3.08 (dd, *J* = 13.5, 4.2 Hz, 1H, Cys βC$\underline{H}_aH_b$), 3.02 (dd, *J* = 13.5, 5.1 Hz, 1H, Cys βC$H_a\underline{H}_b$), 2.82 – 2.72 (m, 2H, S-glyceryl CH_2), 2.28 (t, *J* = 7.3 Hz, 4H, Pal αCH_2 x 2), 1.61 – 1.56 (m, 4H, Pal βCH_2 x 2), 1.49 (s, 9H, *t*Bu CH_3 x 3), 1.32 – 1.22 (m, 48H, Pal CH_2 x 24), 0.88 (t, *J* = 6.8 Hz, 6H, Pal CH_3 x 2) ppm;

m/z HRMS (ESI^+) calcd. for $C_{57}H_{91}NNaO_8S$ = 972.6358 $(M + Na)^+$. Found = 972.6359.

_N_-(((9_H_-Fluoren-9-yl)methoxy)carbonyl)-_S_-((_S_)-2,3-bis(palmitoyloxy)propyl)-L-cysteine (51)



Prepared as per procedure B using compound **54** (0.390 g, 0.411 mmol). The solvents were removed *in vacuo* and the product recrystallised from a minimum amount of CH_2Cl_2 (5 mL) layered with MeOH (50 mL) for 16 h. The obtained white solid was filtered and dried *in vacuo* to yield the product and as a white solid (0.286 g, 78%). The isolated compound was in good agreement with the literature.[290]

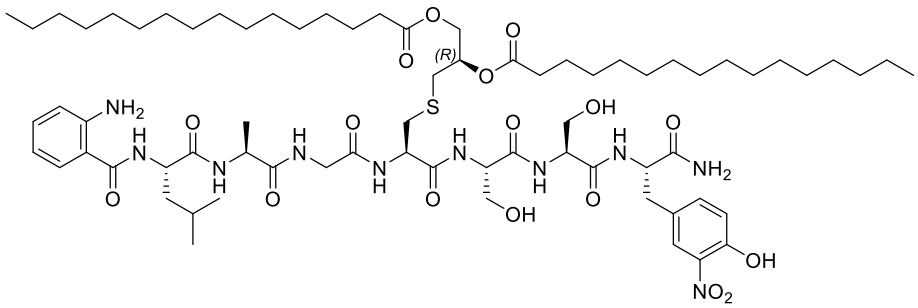
R_f = 0.53 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.63 – 7.59 (m, 2H, Fmoc Ar-CH), 7.40 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.31 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 5.76 (d, *J* = 7.4 Hz, 1H, NH), 5.19 – 5.14 (m, 1H, S-glyceryl CH), 4.71 – 4.66

(m, 1H, Cys αCH), 4.42 – 4.36 (m, 3H, S-glyceryl OC$\underline{H}_aH_b$, Fmoc CH_2), 4.24 (t, *J* = 7.0 Hz, 1H, Fmoc CH), 4.13 (dd, *J* = 11.8, 6.1 Hz, 1H, S-glyceryl $OCH_a\underline{H}_b$), 3.17 (dd, *J* = 13.4, 3.3 Hz, 1H, Cys βC$\underline{H}_aH_b$), 3.10 (dd, *J* = 13.4, 4.3 Hz, 1H, Cys β$CH_a\underline{H}_b$), 2.79 (dd, *J* = 13.8, 6.5 Hz, 1H, S-glyceryl C$\underline{H}_aH_b$), 2.73 (dd, *J* = 13.8, 5.9 Hz, 1H, S-glyceryl $CH_a\underline{H}_b$), 2.33 – 2.28 (m, 4H, Pal αCH_2 x 2), 1.63 – 1.56 (m, 2H, Pal βCH_2 x 2), 1.35 – 1.21 (m, 48H, Pal CH_2 x 24), 0.88 (t, *J* = 6.9 Hz, 6H, Pal-CH_3 x 2) ppm;

m/z HRMS (ESI^+) calcd. for $C_{53}H_{83}NNaO_8S$ = 916.5732 $(M + Na)^+$. Found = 916.5727.

(3*S*,6*S*,12*R*,16*R*)-12-(((*S*)-1-(((*S*)-1-(((*S*)-1-Amino-3-(4-hydroxy-3-nitrophenyl)-1-oxopropan-2-yl)amino)-3-hydroxy-1-oxopropan-2-yl)amino)-3-hydroxy-1-oxopropan-2-yl)carbamoyl)-1-(2-aminophenyl)-3-isobutyl-6-methyl-1,4,7,10-tetraoxo-14-thia-2,5,8,11-tetraazaheptadecane-16,17-diyl dipalmitate (55)



Fmoc-Tyr(3-NO_2)-OH **44** (358 mg, 0.80 mmol) was loaded to Rink amide resin (286 mg, 0.20 mmol) as per procedure E. The coupling of Fmoc-Ser(*t*Bu)-OH (230 mg, 0.60 mmol) x 2, Fmoc-Cys[(*R*)Pam_2] **50** (536 mg, 0.60 mmol), Fmoc-Gly-OH (178 mg, 0.60 mmol), Fmoc-Ala-OH (178 mg, 0.60 mmol), Fmoc-Leu-OH (187 mg, 0.60 mmol) and Boc-Abz-OH **45** (142 mg, 0.60 mmol) was performed as per procedure I. The resin was treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 15 min prior to cleavage. The peptide was cleaved from the resin using procedure J affording a yellow solid which was purified by column chromatography (SiO_2, MeOH:CH_2Cl_2 1:9 v/v) to yield the product as a yellow solid (29.7 mg, 11%).

R_f = 0.32 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (600 MHz, DMSO-d_6) δ 8.22 (d, *J* = 7.5 Hz, 1H, Ser^{+2} NH), 8.13 (d, *J* = 8.0 Hz, 1H, Leu NH), 8.11 – 8.06 (m, 2H, Gly NH, Cys NH), 8.04 (d, *J* = 7.2 Hz, 1H, Ser^{+3} NH), 8.02 (d, *J* = 7.2 Hz, 1H, Ala NH), 7.93 (d, *J* = 8.1 Hz, 1H, Tyr NH), 7.73 – 7.69 (m, 1H, Tyr Ar-CH), 7.55 (d, *J* = 7.5 Hz, 1H, Abz Ar-CH), 7.36 (d, 1H, *J* = 8.0 Hz, Tyr Ar-CH),

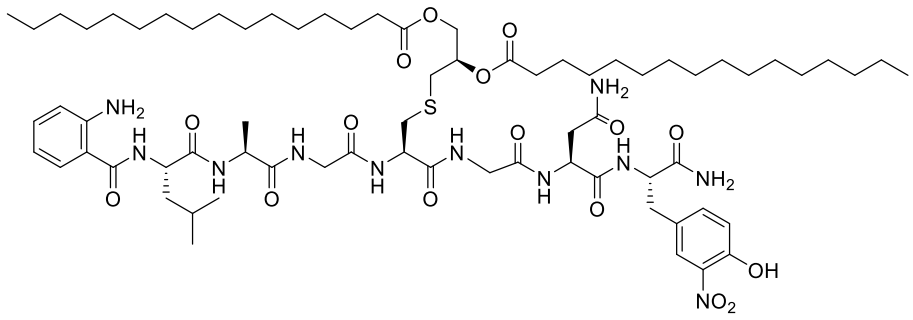
7.27 – 7.21 (m, 2H, $CONH_2$), 7.13 (t, *J* = 8.2 Hz, 1H, Abz Ar-CH), 6.98 (d, *J* = 8.0 Hz, 1H, Tyr Ar-CH), 6.69 (d, *J* = 8.2 Hz, 1H, Abz Ar-CH), 6.54 – 6.49 (m, 1H, Abz Ar-CH), 6.35 – 6.31 (m, 2H, Abz NH_2), 5.20 – 5.15 (m, 1H, Ser^{+2} OH), 5.12 – 5.07 (m, 1H, S-glyceryl CH), 4.96 (t, *J* = 5.5 1H, Ser^{+3} OH), 4.57 – 4.52 (m, 1H, Cys αCH), 4.50 – 4.45 (m, 1H, Leu αCH), 4.39 – 4.33 (m, 2H, Ser^{+2} αCH, Tyr αCH), 4.31 – 4.26 (m, 2H, Ala αCH, S-glyceryl $OC\underline{H_a}H_b$), 4.22 – 4.18 (m, 1H, Ser^{+3} αCH), 4.11 – 4.05 (m, 1H, S-glyceryl $OCH_a\underline{H_b}$), 3.80 – 3.71 (m, 2H, Gly CH_2), 3.66 – 3.60 (m, 1H, Ser^{+2} $C\underline{H_a}H_b$), 3.59 – 3.52 (m, 2H, Ser^{+2} $CH_a\underline{H_b}$, Ser^{+3} $C\underline{H_a}H_b$), 3.52 – 3.47 (m, 1H, Ser^{+3} $CH_a\underline{H_b}$), 3.04 – 2.98 (m, 1H, Tyr $C\underline{H_a}H_b$), 2.91 (dd, *J* = 13.5, 4.6 Hz, 1H, Cys $C\underline{H_a}H_b$), 2.82 (dd, *J* = 14.0, 5.4 Hz, 1H, S-glyceryl $C\underline{H_a}H_b$), 2.76 – 2.71 (m, 1H, Tyr $CH_a\underline{H_b}$), 2.70 – 2.65 (m, 1H, S-glyceryl $CH_a\underline{H_b}$), 2.63 – 2.61 (m, 1H, Cys $CH_a\underline{H_b}$), 2.30 – 2.21 (m, 4H, Pal αCH_2 x 2), 1.71 – 1.62 (m, 3H, Leu CH_2, Leu $\underline{C}H\text{-}CH_3$), 1.56 – 1.45 (m, 7H, Pal CH_2 x 2, Ala CH_3), 1.31 – 1.18 (m, 48H, Pal CH_2 x 24), 0.93 – 0.81 (m, 12H, Leu CH_3 x 2, Pal CH_3 x 2) ppm;

^{13}C NMR (151 MHz, DMSO-d_6) δ 173.0 (Pal C=O), 173.0 (Pal C=O), 172.9 ($CONH_2$), 172.7 (Leu C=O), 172.7 (Ala C=O), 170.7 (Ser^{+2} C=O), 170.6 (Cys C=O), 170.1 (Ser^{+3} C=O), 169.3 (Abz C=O), 169.2 (Gly C=O), 150.1 (Abz-qC NH_2), 136.7 (Tyr Ar-CH), 136.5 (Tyr Ar-CH), 132.3 (Abz Ar-CH), 129.0 (Abz Ar-CH), 125.8 (Tyr Ar-CH), 116.7 (Abz Ar-CH), 115.0 (Abz Ar-CH), 114.9 (Abz qC), 70.3 (S-glyceryl CH), 64.0 (S-glyceryl OCH_2), 62.1 (Ser^{+2} βCH_2), 61.8 (Ser^{+3} βCH_2), 56.1 (Ser^{+3} αCH), 55.3 (Ser^{+2} αCH), 54.3 (Tyr αCH), 52.6 (Cys αCH), 51.7 (Leu αCH), 48.7 (Ala αCH), 42.3 (Gly CH_2), 36.3 (Tyr βCH_2), 34.6 (Cys βCH_2), 34.0 (Pal αCH_2), 33.9 (Pal αCH_2), 32.1 (Pal CH_2), 31.8 (S-glyceryl CH_2), 29.5, 29.4, 29.5, 29.3, 29.2, 29.2, 28.9 (Pal CH_2), 25.0 (Leu CH), 24.9 (Pal CH_2), 24.9, 23.7 (Leu CH_3), 22.6 (Pal CH_2), 21.8 (Leu CH_3), 18.7 (Ala CH_3), 14.4 (Pal CH_3) ppm;

HRMS (*m/z* MALDI$^+$) calcd. for $C_{71}H_{116}N_{10}O_{17}SNa$ = 1435.8138 $(M + Na)^+$. Found = 1435.8124;

ν_{max} (thin film)/cm^{-1} 3283 (NH), 2924 (CH_2), 2852 (CH_2), 1746, 1631 (Amide I: CO, CN), 1536 (Amide II: CN, NH), 1369 (Amide III: CN, CO, CC), 1210 (CO).

(3*S*,6*S*,12*R*,16*R*)-12-((2-(((*S*)-4-Amino-1-(((*S*)-1-amino-3-(4-hydroxy-3-nitrophenyl)-1-oxopropan-2-yl)amino)-1,4-dioxobutan-2-yl)amino)-2-oxoethyl)carbamoyl)-1-(2-aminophenyl)-3-isobutyl-6-methyl-1,4,7,10-tetraoxo-14-thia-2,5,8,11-tetraazaheptadecane-16,17-diyl dipalmitate (56)



Fmoc-Tyr(3-NO_2)-OH **44** (634 mg, 1.41 mmol) was loaded to Rink amide resin (508 mg, 0.35 mmol) as per procedure E. The coupling of Fmoc-Asn(Trt)-OH (634 mg, 1.06 mmol), Fmoc-Gly-OH (315 mg, 1.06 mmol) x 2, Fmoc-Cys[(*R*)Pam_2] **50** (633 mg, 1.062 mmol), Fmoc-Ala-OH (330 mg, 1.06 mmol), Fmoc-Leu-OH (375 mg, 1.06 mmol) and Boc-Abz-OH **45** (252 mg, 1.06 mmol) was performed as per procedure I. The resin was treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 15 min prior to cleavage. The peptide was cleaved from the resin using procedure J affording a yellow solid which was purified by column chromatography (SiO_2, MeOH:CH_2Cl_2 3:47 v/v) to yield the product as a yellow solid (9.0 mg, 2%).

R_f = 0.47 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (600 MHz, DMSO-d_6) δ 8.40 – 8.29 (m, 1H, Gly^{+2} NH), 8.19 – 8.08 (m, 3H, Leu NH, Gly^{-1} NH, Cys NH), 8.08 – 7.99 (m, 3H, Asn NH, Tyr NH, Ala NH), 7.73 (s, 1H, Tyr Ar-CH), 7.54 (d, *J* = 7.1 Hz, 1H, Abz Ar-CH), 7.37 (s, 2H, Tyr Ar-CH, CON$\underline{H_a}H_b$), 7.19 – 7.10 (m, 1H, Abz Ar-CH), 7.02 – 6.95 (m, 1H, Tyr-Ar CH), 6.93 (s, 1H, CON$H_a\underline{H_b}$), 6.68 (d, 1H, *J* = 7.9 Hz, Abz Ar-CH), 6.51 (t, 1H, *J* = 7.4 Hz, Abz Ar-CH), 6.32 (s, 2H, Abz NH_2), 5.24 – 5.01 (m, 1H, S-glyceryl CH), 4.55 – 4.40 (m, 3H, Cys αCH, Leu αCH, Tyr αCH), 4.38 – 4.22 (m, 3H, Ala αCH, Asn αCH, S-glyceryl OC$\underline{H_a}H_b$), 4.16 – 4.05 (m, 1H, S-glyceryl OC$H_a\underline{H_b}$), 3.77 (d, *J* = 5.3 Hz, 2H, Gly^{+2} CH_2), 3.68 (dd, *J* = 16.8, 5.8 Hz, 2H, Gly^{-1} CH_2), 3.07 – 3.00 (m, 1H, Asn βC$\underline{H_a}H_b$), 2.93 (dd, *J* = 13.6, 4.9 Hz, 1H, Cys βC$\underline{H_a}H_b$), 2.85 – 2.74 (m, 2H, Asn βC$H_a\underline{H_b}$, S-glyceryl C$\underline{H_a}H_b$), 2.72 – 2.62 (m, 2H, Cys βC$H_a\underline{H_b}$, S-glyceryl C$H_a\underline{H_b}$), 2.38 – 2.31 (m, 1H, Tyr βC$\underline{H_a}H_b$) 2.28 – 2.19 (m, 4H, Pal αCH_2 x 2), 1.72 – 1.60 (m, 2H, Leu βCH_2), 1.57 – 1.43 (m, 5H,

Leu CH, Pal CH_2 x 2), 1.31 – 1.15 (m, 51H, Pal CH_2 x 24, Ala CH_3), 0.94 – 0.79 (m, 12H, Leu CH_3 x 2, Pal CH_3 x 2) ppm;

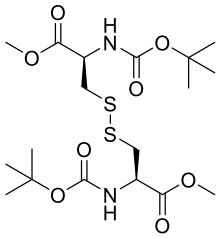
^{13}C NMR (151 MHz, DMSO-d_6) δ 172.6, 172.5, 172.5, 172.3, 171.7, 170.6, 170.5, 168.9, 168.9, 168.4 (C=O), 149.6 (Abz qC-NH_2), 136.4 (Tyr Ar-CH), 136.0 (Tyr Ar-CH), 131.8 (Abz Ar-CH), 128.5 (Abz Ar-CH), 125.3 (Tyr Ar-CH), 116.3 (Abz Ar-CH), 114.5 (Abz Ar-CH), 114.5 (Abz qC), 69.8 (S-glyceryl CH), 63.5 (S-glyceryl OCH_2), 53.9 (Asn αCH), 52.3 (Cys αCH), 51.2 (Leu αCH), 49.7 (Tyr αCH), 48.3 (Ala αCH), 42.1 (Gly CH_2), 41.8 (Gly-CH_2), 36.9 (Tyr βCH_2), 35.6 (Asn βCH_2), 33.8 (Cys βCH_2), 33.6 (Pal αCH_2), 33.4 (Pal αCH_2), 31.6 (S-glyceryl CH_2), 31.3, 29.1, 29.0, 29.0, 28.9, 28.8, 28.7, 28.5, 28.4 (Pal CH_2), 24.5 (Leu CH_2), 24.4 (Leu CH), 23.2 (Leu CH_3), 21.3 (Leu CH_3), 18.2 (Ala CH_3), 13.9 (Pal CH_3) ppm;

HRMS (*m/z* MALDI$^+$) calcd. for $C_{71}H_{115}N_{11}O_{16}SNa$ = 1432.8146 $(M + Na)^+$. Found = 1432.8146;

ν_{max} (thin film)/cm^{-1} 3280 (NH), 2919 (CH_2), 2848 (CH_2), 1635 (Amide I: CO, CN), 1532 (Amide II: CN, NH), 1376 (Amide III: CN, CO, CC), 1207 (CO).

7.4 Experimental Details for Chapter 3

Dimethyl-3,3'-disulfanediyl(2*R*,2'*R*)-bis(2-(*tert*-butoxycarbonyl)amino)propanoate) (63)



To a solution of $(NH_2$-Cys-OMe$)_2$·HCl **62** (4.80 g, 28.0 mmol) in CH_2Cl_2 (150 mL) at 0 °C was added Et_3N (8.58 mL, 61.5 mmol) dropwise over 10 min. Boc_2O (7.32 g, 33.6 mmol) was dissolved in CH_2Cl_2 (30 mL) and added dropwise to the reaction. The reaction was allowed to warm to rt and stirred for 4 h. TLC analysis revealed formation of disulfide R_f = 0.30 (EtOAc:hex 3:7 v/v) and corresponding thiol R_f = 0.60 (EtOAc:hex 3:7 v/v). The reaction was washed with H_2O (3 x 100 mL) and sat. aq. NaCl solution (3 x 100 mL) and the organic layers were dried over $MgSO_4$. The solvent was removed in *vacuo*. The disulfide product was purified using column chromatography (SiO_2,

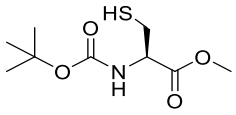
EtOAc:hex 2:8 v/v) yielding a white solid (3.25 g, 50%). The isolated compound was in good agreement with literature.[291]

R_f = 0.30 (EtOAc:hex 3:7 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.41 (d, *J* = 7.5 Hz, 2H, NH x 2), 4.66 – 4.58 (m, 2H, αCH x 2), 3.79 (s, 6H, OCH_3 x 2), 3.18 (d, *J* = 5.1 Hz, 4H, βCH_2 x 2), 1.47 (s, 18H, Boc CH_3 x 6) ppm;

HRMS (*m/z* MALDI$^+$) calcd. for $C_{18}H_{32}N_2O_8NaS_2$ = 491.1498 $(M + Na)^+$. Found 491.1498.

Methyl (*tert*-butoxycarbonyl)-L-cysteinate (64)



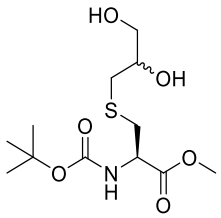
To a solution of disulfide **63** (2.55 g, 5.43 mmol) in MeOH (100 mL) and H_2O (20 mL) was added PBu_3 (2.71 mL, 10.9 mmol). After stirring at rt for 3 h the reaction was diluted with Et_2O (100 mL) and H_2O (50 mL). Following separation of the organic layer, the aq. layer was extracted with Et_2O (3 x 100 mL). The combined organic layers were washed with H_2O (3 x 100 mL) and sat. aq. NaCl solution (3 x 100 mL), dried over $MgSO_4$ and filtered. The solvent was removed in *vacuo*. The product was purified using column chromatography (SiO_2, EtOAc:hex 3:17 v/v) to afford a viscous, colourless oil (2.23 g, 87%). The isolated compound was in good agreement with literature.[262]

R_f = 0.60 (EtOAc:hex 3:7 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.47 – 5.34 (m, 1H, NH), 4.61 – 4.42 (m, 1H, αCH), 3.74 (s, 3H, OCH_3), 3.00 – 2.87 (m, 2H, βCH_2), 1.41 (s, 9H, Boc CH_3 x 3) ppm;

HRMS (*m/z* APCI$^-$) calcd. for $C_9H_{16}NO_4S$ = 234.0806 $(M - H)^-$. Found 234.0802.

Methyl *N*-(*tert*-butoxycarbonyl)-*S*-(2,3-dihydroxypropyl)-L-cysteinate (65)



To a solution of thiol **64** (1.46 g, 6.23 mmol) in anhydrous DMF (10 mL) under argon was added Cs_2CO_3 (2.03 g, 6.23 mmol). The reaction was stirred at rt for 1 h. 3-Bromo-1,2-propandiol (1.06 g, 6.85 mmol) dissolved in anhydrous DMF (5 mL) under argon and added to the reaction which was subsequently stirred at rt for 16 h. H_2O (100 mL) was added, and the product was extracted with CH_2Cl_2 (3 x 100 mL). The combined organic layers were washed with sat. aq. NaCl solution (5 x 100 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude product was purified by column chromatography (SiO_2, EtOAc:hex, 8:2 v/v) to yield a yellow oil (1.22 g, 63%).

R_f = 0.25 (EtOAc:hex 8:2 v/v);

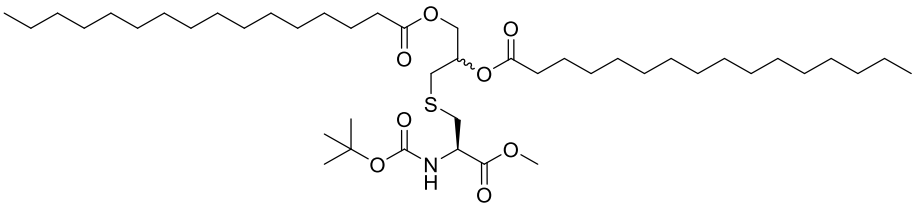
^{1}H NMR (400 MHz, $CDCl_3$) two diastereoisomers (1:1) δ 5.43 (d, *J* = 8.1 Hz, 1H, NH), 4.65 – 4.54 (m, 1H, αCH), 3.88 – 3.81 (m, 1H, S-glyceryl CH), 3.80 (s, 3H, OCH_3), 377 – 3.70 (m, 1H, S-glyceryl $OC\underline{H_a}H_b$), 3.63 – 3.54 (m, 1H, S-glyceryl $OCH_a\underline{H_b}$), 3.13 – 2.98 (m, 1H, $\beta C\underline{H_a}H_b$), 2.97 – 2.88 (m, 1H, $\beta CH_a\underline{H_b}$), 2.86 – 2.74 (m, 1H, S-glyceryl $C\underline{H_a}H_b$) 2.72 – 2.58 (m, 1H, S-glyceryl $CH_a\underline{H_b}$), 1.47 (s, 9H, Boc CH_3 x 3) ppm;

^{13}C NMR (101 MHz, $CDCl_3$); two diastereoisomers δ 171.6 (C=O), 155.5 (NH-C=O), 80.4 (Boc qC), 70.7 (S-glyceryl CH), 65.1 (S-glyceryl OCH_2), 53.4 (αC), 52.6 (OCH_3), 36.2 (S-glyceryl CH_2), 35.4 (βCH_2), 28.3 (Boc CH_3 x 3) ppm;

HRMS (*m*/*z* ESI$^+$) calcd. for $C_{12}H_{23}NNaO_6S$ = 332.1138 $(M + Na)^+$. Found = 332.1133;

ν_{max} (thin film)/cm^{-1} 3373 (OH), 2977 (CH), 1741 (ester C=O), 1689 (Boc C=O).

3-(((*R*)-2-((*tert*-butoxycarbonyl)amino)-3-methoxy-3-oxopropyl)thio)propane-1,2-diyl dipalmitate (66)



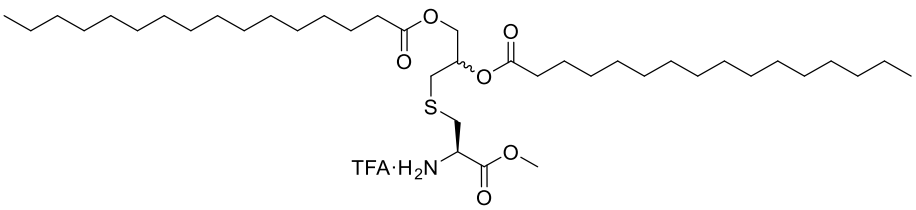
To a solution of palmitic acid (518 mg, 2.02 mmol) in anhydrous THF:CH_2Cl_2 (2:3 v/v, 20 mL) at 0 °C under argon was added EDC·HCl (465 mg, 2.43 mmol) and DMAP (10.0 mg, 0.081 mmol). The reaction was stirred at 0 °C for 1 h. Diol **65** (250 mg, 0.808 mmol) dissolved in anhydrous THF (5 mL) was added to the reaction which was stirred at rt under argon for 3 d. The solvents were remove *in vacuo* and the crude product was purified by column chromatography (SiO_2, EtOAc:hex, 3:7 v/v) to yield the diacylated compound **66** as a white solid (444 mg, 70%). The isolated compound was in good agreement with the literature.[115]

R_f = 0.64 (EtOAc:hex 35:65 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) two diastereoisomers (1:1) δ 5.36 – 5.27 (m, 1H, NH), 5.15 – 5.08 (m, 1H, αCH), 4.56 – 4.49 (m, 1H, S-glyceryl CH), 4.33 – 4.28 (m, 1H, S-glyceryl $OC\underline{H_a}H_b$), 4.16 – 4.10 (m, 1H, S-glyceryl $OCH_a\underline{H_b}$), 3.74 (s, 3H, OCH_3), 3.11 – 2.89 (m, 2H, βCH_2), 2.77 – 2.66 (m, 2H, S-glyceryl CH_2), 2.38 – 2.20 (m, 4H, Pal αCH_2 x 2), 1.63 – 1.54 (m, 4H, Pal βCH_2 x 2), 1.43 (s, 9H, Boc CH_3 x 3), 1.33 – 1.18 (m, 48H, Pal CH_2 x 24), 0.86 (t, *J* = 6.8 Hz, 6H, Pal CH_3 x 2) ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{44}H_{83}NO_8SNa$ = 808.5741 $(M + Na)^+$. Found = 808.5732.

3-(((*R*)-3-methoxy-3-oxo-2-((2,2,2-trifluoroacetyl)-λ^4-azaneyl)propyl)thio)propane-1,2-diyl dipalmitate (61)



Compound **66** (62 mg, 0.0789 mmol) was dissolved in CH_2Cl_2 (3 mL) and TFA (1 mL) was added. The reaction was stirred at rt for 16 h and the solvents removed *in vacuo* to

yield a pale yellow solid (57 mg, 90%). The isolated compound was in good agreement with the literature.[115]

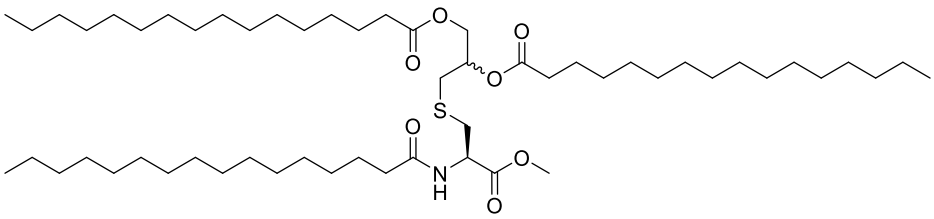
R_f = 0.26 (EtOAc:hex 1:1 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) two diastereoisomers δ 5.22 – 5.13 (m, 1H, S-glyceryl CH), 4.43 – 4.29 (m, 2H, αCH, S-glyceryl OC$\underline{H_a}H_b$), 4.21 – 4.10 (m, 1H, S-glyceryl OC$H_a\underline{H_b}$), 3.88 (s, 3H, OCH_3), 3.39 – 3.14 (m, 2H, βCH_2), 2.85 – 2.72 (m, 2H, S-glyceryl CH_2), 2.41 – 2.30 (m, 4H, Pal αCH_2 x 2), 1.69 – 1.57 (m, 4H, Pal βCH_2 x 2), 1.38 – 1.19 (m, 48H, Pal CH_2 x 24), 0.90 (t, *J* = 6.9 Hz, 6H, Pal CH_3 x 2) ppm;

^{19}F NMR (377 MHz, $CDCl_3$) two diastereoisomers δ -75.91 ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{39}H_{76}NO_6S$ = 686.5388 $(M + H)^+$. Found = 686.5380.

3-(((*R*)-3-methoxy-3-oxo-2-palmitamidopropyl)thio)propane-1,2-diyl dipalmitate (67)



To a solution of palmitic acid (48 mg, 0.187 mmol) in anhydrous THF:CH_2Cl_2 (1:2 v/v, 15 mL) at 0 °C under argon was added EDC·HCl (36 mg, 0.19 mmol), HOBt (28 mg, 0.19 mmol) and DIPEA (87 μL, 0.500 mmol). The reaction was stirred at 0 °C for 1 h. Compound **61** (100 mg, 0.13 mmol) was added to the reaction which was stirred at rt under argon for 2 d. The solvents were removed *in vacuo* and the crude product was purified by flash chromatography (hex - EtOAc:hex 2:8 v/v) to yield the triacylated compound as a white solid (70.0 mg, 61%). The isolated compound was in good agreement with the literature.[292]

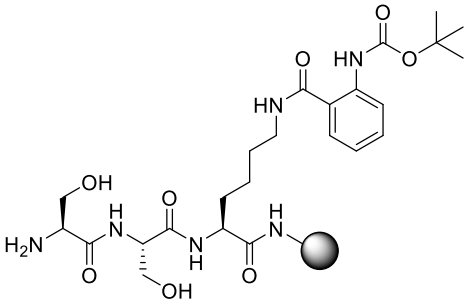
R_f = 0.32 (CH_2Cl_2:MeOH, 99:1 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) two diastereoisomers δ 6.34 – 6.28 (m, 1H, NH), 5.14 – 5.08 (m, 1H, αCH), 4.91 – 4.72 (m, 1H, S-glyceryl CH), 4.34 – 4.28 (m, 1H, S-glyceryl OC$\underline{H_a}H_b$), 4.15 – 4.09 (m, 1H, S-glyceryl OC$H_a\underline{H_b}$), 3.79 (s, 3H, OCH_3), 3.16 – 2.99 (m, 2H, S-glyceryl CH_2), 2.73 (d, *J* = 6.4 Hz, 2H, βCH_2), 2.38 – 2.26 (m, 4H, O-Pal

αCH$_2$ x 2), 2.25 – 2.22 (m, 2H, N-Pal αCH$_2$) 1.87 – 1.55 (m, 6H, Pal βCH$_2$ x 3), 1.39 – 1.21 (m, 72H, Pal CH$_2$ x 36), 0.90 (t, *J* = 6.8 Hz, 9H, Pal CH$_3$ x 9);

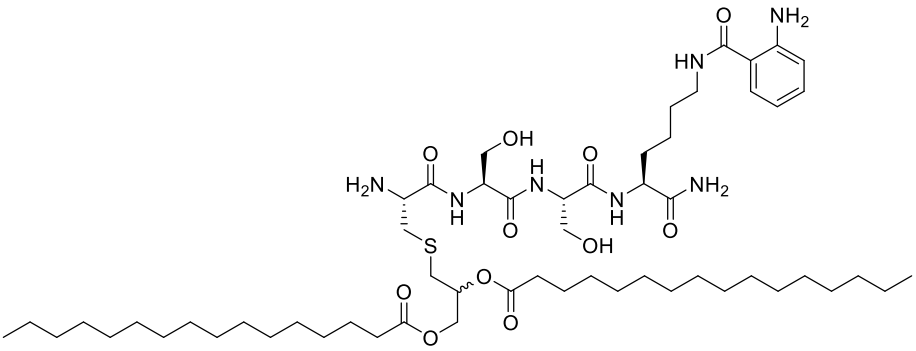
m/z HRMS (ESI$^+$) calcd. for $C_{55}H_{105}NNaO_7S$ = 946.7504 $(M + Na)^+$. Found 946.7500.

Common resin intermediate (76)



Fmoc-Lys(Mtt)-OH (365 mg, 0.580 mmol) was loaded to Rink amide resin (209 mg, 0.150 mmol) as per procedure E. Mtt deprotection of the Lys side chain was effected by a solution of 5% (v/v) TFA, 5% (v/v) TES, 90% (v/v) CH_2Cl_2 (5 mL) which was added to the resin for 5 x 1 min. The resin was drained and washed with CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL). Boc-Abz-OH **45** (104 mg, 0.438 mmol) was coupled as per procedure I and followed by Lys Fmoc deprotection using a solution of 20% (v/v) piperidine in DMF (5 mL) which was added to the resin for 2 x 10 min. The resin was washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) then DMF (3 x 5 mL) and Fmoc-Ser(*t*Bu)-OH (168 mg, 0.438 mmol) x 2 was coupled to the resin as per procedure I. The resin was drained and washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL) and used directly in the next step.

(7*S*,10*S*,13*S*,16*R*)-16-Amino-1-(2-aminophenyl)-7-carbamoyl-10,13-bis(hydroxymethyl)-1,9,12,15-tetraoxo-18-thia-2,8,11,14-tetraazahenicosane-20,21-diyl dipalmitate (70)



Fmoc-Cys[(*R*/*S*)Pam$_2$]-OH **28** (380 mg, 0.425 mmol) was coupled to peptide resin **76** (0.142 mmol) according to procedure I. The resin was treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 10 min to remove the Cys Fmoc group and washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL). The peptide was subsequently cleaved from the resin as per procedure J, purified by column chromatography (SiO_2, CH_2Cl_2 – CH_2Cl_2:MeOH 9:1 v/v), and dried *in vacuo* at 0 °C to yield the target peptide as a white solid (10.8 mg, 7%).

R_f = 0.31 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (600 MHz, DMSO-d_6) two diastereoisomers δ 8.22 (s, 1H, Lys εNH), 8.18 – 8.12 (m, 2H, Ser^{+2} NH, Ser^{+3} NH), 7.86 (d, *J* = 8.0 Hz, 1H, Lys NH), 7.44 (dd, *J* = 7.9, 1.1 Hz, 1H, Abz Ar-CH), 7.17 – 7.05 (m, 3H, $CONH_2$, Abz Ar-CH), 6.66 (dd, *J* = 7.9, 0.8 Hz, 1H, Abz Ar-CH), 6.52 – 6.45 (m, 1H, Abz Ar-CH), 6.34 (s, 2H, Abz-NH_2), 5.21 (t, *J* = 5.3 Hz, 1H, Ser^{+2} OH), 5.15 – 5.07 (m, 1H, S-glyceryl CH), 5.05 (t, *J* = 5.5 Hz, 1H, Ser^{+3} OH), 4.42 – 4.35 (m, 1H, Ser^{+2} αCH), 4.33 – 4.25 (m, 2H, Ser^{+3} αCH, S-glyceryl OC$\underline{H_a}H_b$), 4.14 – 4.06 (m, 2H, Lys αCH, S-glyceryl OC$H_a\underline{H_b}$), 3.72 – 3.64 (m, 2H, Ser^{+2} C$\underline{H_a}H_b$, Ser^{+3} C$\underline{H_a}H_b$), 3.60 – 3.55 (m, 1H, Ser^{+3} C$H_a\underline{H_b}$), 3.54 – 3.49 (m, 1H, Ser^{+2} C$H_a\underline{H_b}$), 3.45 – 3.39 (m, 1H, Cys αCH), 3.16 (dd, *J* = 12.9, 7.0 Hz, 1H, Lys εCH_2), 2.88 (dd, *J* = 13.3, 4.4 Hz, 1H, Cys C$\underline{H_a}H_b$), 2.80 (dd, *J* = 14.0, 5.5 Hz, 1H, S-glyceryl C$\underline{H_a}H_b$), 2.73 – 2.66 (m, 1H, S-glyceryl C$H_a\underline{H_b}$), 2.63 – 2.57 (m, 1H, Cys C$H_a\underline{H_b}$), 2.32 – 2.22 (m, 4H, Pal αCH_2 x 2), 1.79 – 1.64 (m, 1H, Lys βC$\underline{H_a}H_b$), 1.57 – 1.39 (m, 7H, Lys βC$H_a\underline{H_b}$, Pal-CH_2 x 2, Lys *δ*CH_2), 1.39 – 1.32 (m, 2H, Lys γCH_2), 1.31 - 1.16 (m, 48H, Pal CH_2 x 24), 0.85 (t, *J* = 7.0 Hz, 6H, Pal CH_3 x 2) ppm;

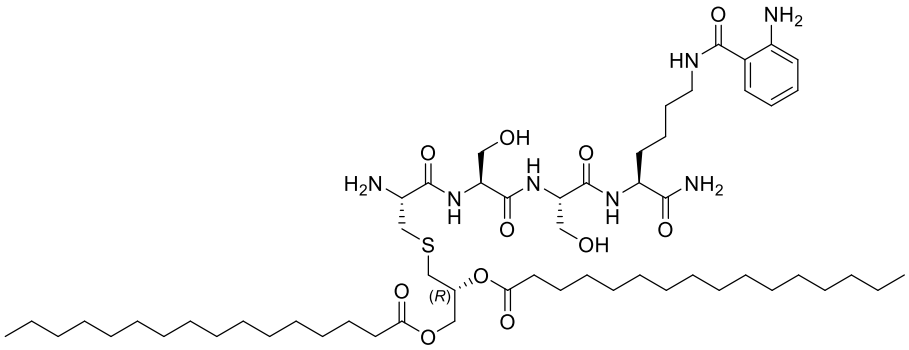
^{13}C NMR (151 MHz, DMSO-d_6) two diastereoisomers δ 173.6 ($CONH_2$), 172.5 (Pal C=O), 172.3 (Pal C=O), 170.1 (Ser^{+2} C=O), 169.6 (Ser^{+3} C=O), 168.7 (Abz C=O), 149.3 (Abz qC-NH_2), 131.4 (Abz Ar-CH), 116.1 (Abz Ar-CH), 115.0 (Abz qC), 114.3 (Abz Ar-CH), 70.0 (S-glyceryl CH), 63.6 (S-glyceryl OCH_2), 61.8 (Ser^{+2} βCH_2), 61.5 (Ser^{+3} βCH_2), 55.7 (Ser^{+3} αCH), 54.6 (Ser^{+2} αCH), 54.1 (Cys αCH), 52.8 (Lys αCH), 38.7 (Lys εCH_2), 37.3 (Cys βCH_2), 33.6 33.4 (Pal αCH_2), 31.7 (S-glyceryl CH_2; diastereoisomer A), 31.6 (S-glyceryl CH_2; diastereoisomer B), 31.3 (Lys βCH_2), 31.2, 29.0, 28.9, 28.8, 28.7, 28.6, 28.4, 28.3 (Pal CH_2), 24.5, 24.4 (Pal βCH_2), 23.0 (Lys γCH_2), 22.1 (Pal CH_2), 13.9 (Pal CH_3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{57}H_{102}N_7O_{11}SNa$ = 1114.7177 $(M + Na)^+$. Found = 1114.7131;

ν_{max} (thin film)/cm^{-1}: 3282 (Amide A: NH), 2923 (CH_2), 1670 (Amide I: CO, CN), 1529 (Amide II: CN, NH), 1279 (Amide III: CN, CO), 1201 (CO).

7.5 Experimental Details for Chapter 4

(7*S*,10*S*,13*S*,16*R*,20*R*)-16-Amino-1-(2-aminophenyl)-7-carbamoyl-10,13-bis(hydroxymethyl)-1,9,12,15-tetraoxo-18-thia-2,8,11,14-tetraazahenicosane-20,21-diyl dipalmitate (79)



Fmoc-Cys[(*R*)Pam_2]-OH **50** (392 mg, 0.438 mmol) was coupled to peptide resin **76** (0.146 mmol) according to procedure I. The resin was then treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 10 min to remove the Cys Fmoc group and washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL). The peptide was subsequently cleaved from the resin as per procedure J, purified by column chromatography (SiO_2, CH_2Cl_2 – CH_2Cl_2:MeOH 9:1 v/v), and dried *in vacuo* at 0 °C to yield the target peptide as a white solid (14.0 mg, 9%).

R_f = 0.31 (CH_2Cl_2:MeOH 9:1 v/v);

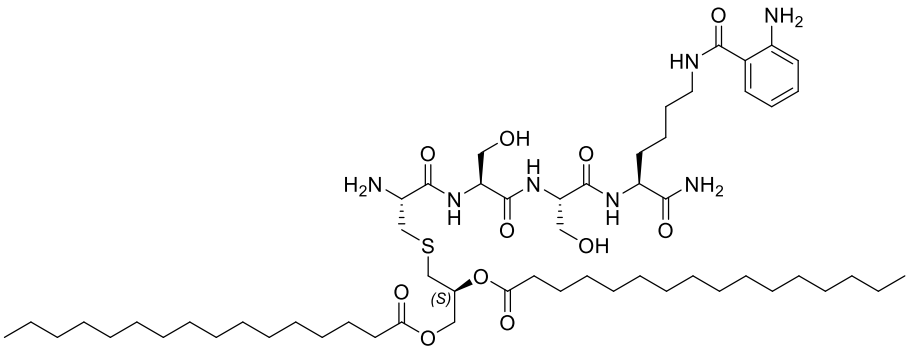
^{1}H NMR (600 MHz, DMSO-d_6) δ 8.21 – 8.13 (m, 3H, Lys εNH, Ser^{+2} NH, Ser^{+3} NH), 7.87 (d, *J* = 7.9 Hz, 1H, Lys NH), 7.45 (d, *J* = 8.0, Hz, 1H, Abz Ar-CH), 7.16 – 7.07 (m, 3H, $CONH_2$, Abz Ar-CH), 6.68 (d, *J* = 8.0 Hz, 1H, Abz Ar-CH), 6.53 – 6.47 (m, 1H, Abz Ar-CH), 6.35 (s, 2H, Abz NH_2), 5.24 – 5.19 (m, 1H, Ser^{+2} OH), 5.14 – 5.08 (m, 1H, S-glyceryl CH), 5.08 – 5.04 (m, 1H, Ser^{+3} OH), 4.41 - 4.36 (m, 1H, Ser^{+2} αCH), 4.34 – 4.26 (m, 2H, Ser^{+3} αCH, S-glyceryl O$C\underline{H_a}H_b$), 4.15 – 4.07 (m, 2H, Lys αCH, S-glyceryl O$CH_a\underline{H_b}$), 3.72 – 3.64 (m, 2H, Ser^{+2} $C\underline{H_a}H_b$, Ser^{+3} $C\underline{H_a}H_b$), 3.61 – 3.55 (m, 1H, Ser^{+3} $CH_a\underline{H_b}$), 3.55 – 3.49 (m, 1H, Ser^{+2} $CH_a\underline{H_b}$), 3.41 – 3.38 (m, 1H, Cys αCH), 3.17 (dd, *J* = 12.0, 5.6 Hz, 2H, Lys εCH_2), 2.88 (dd, *J* = 13.3, 4.3 Hz, 1H, Cys $C\underline{H_a}H_b$), 2.80 (dd, *J* = 14.1, 5.7 Hz, 1H, S-glyceryl $C\underline{H_a}H_b$), 2.70 (dd, *J* = 14.1, 7.3 Hz, 1H, S-glyceryl $CH_a\underline{H_b}$), 2.61 – 2.57 (m, 1H, Cys $CH_a\underline{H_b}$), 2.31 – 2.24 (m, 4H, Pal αCH_2 x 2), 1.78 – 1.71 (m, 1H, Lys β$C\underline{H_a}H_b$), 1.57 – 1.42 (m, 7H, Lys β$CH_a\underline{H_b}$, Pal CH_2 x 2, Lys *δ*CH_2), 1.38 – 1.32 (m, 2H, Lys γCH_2), 1.31 – 1.17 (m, 48H, Pal CH_2 x 24), 0.86 (t, *J* = 6.9 Hz, 6H, Pal CH_3 x 2) ppm;

^{13}C NMR (151 MHz, DMSO-d_6) δ 173.8 ($CONH_2$), 172.5 (Pal C=O), 172.3 (Pal C=O), 170.5 (Ser^{+2} C=O), 169.7 (Ser^{+3} C=O), 168.8 (Abz C=O), 149.5 (Abz qC-NH_2), 131.4 (Abz Ar-CH), 128.0 (Abz Ar-CH), 116.3 (Abz Ar-CH), 115.0 (Abz qC), 114.5 (Abz Ar-CH), 70.0 (S-glyceryl CH), 63.5 (S-glyceryl OCH_2), 62.0 (Ser^{+2} βCH_2), 61.4 (Ser^{+3} βCH_2), 55.5 (Ser^{+3} αCH), 54.4 (Ser^{+2} αCH), 54.2 (Cys αCH), 52.8 (Lys αCH), 38.7 (Lys εCH_2), 37.5 (Cys βCH_2), 33.6 (Pal αCH_2), 33.4 (Pal αCH_2), 31.7 (S-glyceryl CH_2), 31.3 (Lys βCH_2), 31.2, 29.1, 29.0, 28.9, 28.8, 28.7, 28.6, 28.4, 28.4 (Pal CH_2), 24.5, 24.4 (Pal βCH_2), 23.1 (Lys γCH_2), 22.1 (Pal-CH_2), 13.9 (Pal-CH_3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{57}H_{102}N_7O_{11}S$ = 1092.7353 $(M + H)^+$. Found = 1092.7350;

ν_{max} (thin film)/cm^{-1}: 3287 (Amide A: NH), 2919 (CH_2), 2851 (CH_2), 1674 (Amide I: CO, CN), 1534 (Amide II: CN, NH), 1440, 1303 (Amide III: CN, CO), 1203 (CO).

(7*S*,10*S*,13*S*,16*R*,20*S*)-16-Amino-1-(2-aminophenyl)-7-carbamoyl-10,13-bis(hydroxymethyl)-1,9,12,15-tetraoxo-18-thia-2,8,11,14-tetraazahenicosane-20,21-diyl dipalmitate (80)



Fmoc-Cys[(*S*)Pam$_2$]-OH **51** (114 mg, 0.219 mmol) was coupled to peptide resin **76** (0.073 mmol) according to procedure I. The resin was treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 10 min to remove the Cys Fmoc group and washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL). The peptide was subsequently cleaved from the resin as per procedure J, purified by column chromatography (SiO_2, CH_2Cl_2 – CH_2Cl_2:MeOH 9:1 v/v), and dried *in vacuo* at 0 °C to yield the target peptide as a white solid (5.0 mg, 6%).

R_f = 0.31 (CH_2Cl_2:MeOH 9:1 v/v);

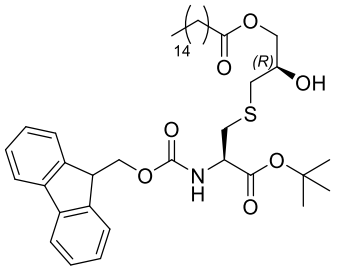
^{1}H NMR (600 MHz, DMSO-d_6) δ 8.27 – 8.13 (m, 3H, Lys εNH, Ser^{+2} NH, Ser^{+3} NH), 7.86 (d, *J* = 7.7 Hz, 1H, Lys NH), 7.44 (dd, *J* = 8.2, 1.0 Hz, 1H, Abz Ar-CH), 7.19 – 7.00 (m, 3H, $CONH_2$, Abz Ar-CH), 6.66 (dd, *J* = 8.2, 1.0 Hz, 1H, Abz Ar-CH), 6.56 – 6.44 (m, 1H, Abz Ar-CH), 6.34 (s, 2H, Abz NH_2), 5.21 (t, *J* = 5.3 Hz, 1H, Ser^{+2} OH), 5.15 – 5.08 (m, 1H, S-glyceryl CH), 5.05 (t, *J* = 5.5 Hz, 1H, Ser^{+3} OH), 4.44 – 4.34 (m, 1H, Ser^{+2} αCH), 4.35 – 4.22 (m, 2H, Ser^{+3} αCH, S-glyceryl OC$\underline{H_a}H_b$), 4.16 – 4.03 (m, 2H, Lys αCH, S-glyceryl OC$H_a\underline{H_b}$), 3.73 – 3.61 (m, 2H, Ser^{+2} C$\underline{H_a}H_b$, Ser^{+3} C$\underline{H_a}H_b$), 3.60 – 3.55 (m, 1H, Ser^{+3} C$H_a\underline{H_b}$), 3.54 – 3.47 (m, 1H, Ser^{+2} C$H_a\underline{H_b}$), 3.45 – 3.39 (m, 1H, Cys αCH), 3.16 (dd, *J* = 12.8, 7.0 Hz, 2H, Lys εCH_2), 2.88 (dd, *J* = 13.2, 4.4 Hz, 1H, Cys C$\underline{H_a}H_b$), 2.80 (dd, *J* = 13.9, 5.5 Hz, 1H, S-glyceryl C$\underline{H_a}H_b$), 2.73 – 2.65 (m, 1H, S-glyceryl C$H_a\underline{H_b}$), 2.65 – 2.57 (m, 1H, Cys C$H_a\underline{H_b}$), 2.29 – 2.23 (m, 4H, Pal αCH_2 x 2), 1.77 – 1.69 (m, 1H, Lys βC$\underline{H_a}H_b$), 1.58 – 1.44 (m, 7H, Lys βC$H_a\underline{H_b}$, Pal CH_2 x 2, Lys CH_2), 1.36 – 1.31 (m, 2H, Lys γCH_2), 1.31 – 1.17 (m, 48H, Pal CH_2 x 24), 0.85 (t, *J* = 6.9 Hz, 6H, Pal CH_3 x 2);

^{13}C NMR (151 MHz, DMSO-d_6) δ 173.7 ($CONH_2$), 172.5 (Pal C=O), 172.3 (Pal C=O), 170.4 (Ser^{+2} C=O), 169.7 (Ser^{+3} C=O), 168.7 (Abz C=O), 149.5 (Abz qC-NH_2), 131.4 (Abz Ar-CH), 128.0 (Abz Ar-CH), 116.2 (Abz Ar-CH), 115.0 (Abz qC), 114.5 (Abz Ar-CH), 70.0 (S-glyceryl CH), 63.5 (S-glyceryl OCH_2), 61.9 (Ser^{+2} βCH_2), 61.4 (Ser^{+3} βCH_2), 55.5 (Ser^{+3} αCH), 54.4 (Ser^{+2} αCH), 54.0 (Cys αCH), 52.7 (Lys αCH), 38.7 (Lys εCH_2), 37.2 (Cys βCH_2), 33.6, 33.4 (Pal αCH_2), 31.6 (S-glyceryl CH_2), 31.3 (Lys βCH_2), 29.1, 29.0, 28.9, 28.9, 28.8, 28.7, 28.4 (Pal CH_2), 24.5, 24.4 (Pal βCH_2), 23.1 (Lys γCH_2), 22.1 (Pal CH_2), 13.9 (Pal CH_3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{57}H_{102}N_7O_{11}S$ = 1092.7353 $(M + H)^+$. Found = 1092.7345;

ν_{max} (thin film)/cm^{-1}: 3354 (Amide A: NH), 2920 (CH_2), 2855 (CH_2), 1636 (Amide I: CO, CN), 1560 (Amide II: CN, NH), 1223 (Amide III: CN, CO), 1040 (CO).

(*R*)-3-(((*R*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(*tert*-butoxy)-3-oxopropyl)thio)-2-hydroxypropyl palmitate (82)



To a solution of *tert*-butyl-*N*-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-*S*-((*R*)-2,3-dihydroxypropyl)-L-cysteinate **40** (0.500 g, 0.992 mmol) in anhydrous CH_2Cl_2 (35 mL) under argon at -10 °C was added Et_3N (0.207 mL, 1.49 mmol) and a solution of palmitoyl chloride (0.451 mL, 1.49 mmol) in anhydrous CH_2Cl_2 (15 mL) at -10 °C dropwise *via* syringe. The reaction was stirred for 4 h with the temperature strictly maintained at -10 °C. MeOH (5 mL) was added and the solvents were removed *in vacuo* at 0 °C. The product was purified using column chromatography (SiO_2, hex - EtOAc:hex 3:17 v/v) and dried *in vacuo* at 0 °C to yield the desired product as a white solid (0.586 g, 83%). The isolated compound was in good agreement with the literature.[242]

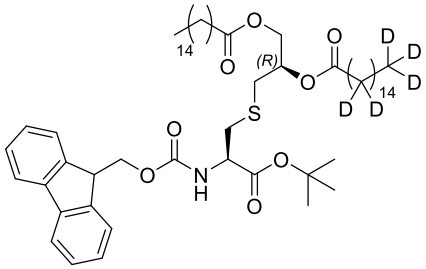
R_f = 0.34 (EtOAc:hex 3:7 v/v);

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.78 (d, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.63 (d, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.42 (t, *J* = 7.5 Hz, 2H, Fmoc Ar-CH), 7.34 (t, *J* = 7.5 Hz, 2H,

Fmoc Ar-CH), 5.76 (d, J = 7.0 Hz, 1H, NH), 4.58 – 4.52 (m, 1H, Cys αCH), 4.42 (d, J = 6.5 Hz, 2H, Fmoc CH_2), 4.26 (t, J = 6.5 Hz, 1H, Fmoc CH), 4.18 (dd, J = 11.4, 3.7 Hz, 1H, S-glyceryl $OC\underline{H}_aH_b$), 4.13 – 4.07 (m, 1H, S-glyceryl $OCH_a\underline{H}_b$), 3.97 – 3.92 (m, 1H, S-glyceryl CH), 3.11 – 2.95 (m, 2H, βCH_2), 2.87 – 2.79 (m, 1H, S-glyceryl $C\underline{H}_aH_b$), 2.68 – 2.60 (m, 1H, S-glyceryl $CH_a\underline{H}_b$), 2.34 (t, J = 7.5 Hz, 2H, Pal αCH_2), 1.67 – 1.58 (m, 2H, Pal CH_2), 1.52 (s, 9H, *t*Bu CH_3 x 3), 1.34 – 1.13 (m, 24H, Pal CH_2 x 12), 0.90 (t, J = 6.9 Hz, 3H, Pal CH_3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{41}H_{61}NNaO_7S$ = 734.4061 $(M + Na)^+$. Found = 734.4063.

(5*R*,9*R*)-5-(*tert*-Butoxycarbonyl)-1-(9*H*-fluoren-9-yl)-3,12-dioxo-2,11-dioxa-7-thia-4-azaheptacosan-9-yl hexadecanoate-d_{31} (83)



To a solution of palmitic acid-d_{31} (0.220 g, 0.764 mmol) in anhydrous CH_2Cl_2 (20 mL) under argon at 0 °C was added EDC·HCl (0.146 g, 0.764 mmol) and DMAP (12 mg, 0.102 mmol). After 1 h of stirring at 0 °C, compound **82** (0.363 g, 0.509 mmol) was added. The reaction was stirred at rt for 16 h and the product was purified using column chromatography (SiO_2, EtOAc:hex 1:9) to yield the product as a white solid (0.354 g, 71%).

R_f = 0.60 (EtOAc:hex 1:4 v/v);

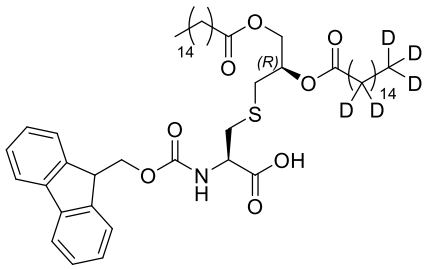
^{1}H NMR (600 MHz, $CDCl_3$) δ 7.76 (d, J = 7.4 Hz, 2H, Fmoc Ar-CH), 7.62 (d, J = 7.4 Hz, 2H, Fmoc Ar-CH), 7.40 (t, J = 7.4 Hz, 2H, Fmoc Ar-CH), 7.31 (t, J = 7.4 Hz, 2H, Fmoc Ar-CH), 5.71 (d, J = 7.5 Hz, 1H, NH), 5.18 – 5.13 (m, 1H, S-glyceryl CH), 4.54 – 4.48 (m, 1H, Cys αCH), 4.44 – 4.28 (m, 3H, Fmoc CH_2, S-glyceryl $OC\underline{H}_aH_b$), 4.24 (t, J = 7.1 Hz, 1H, Fmoc CH), 4.18 – 4.12 (m, 1H, S-glyceryl $OCH_a\underline{H}_b$), 3.09 (dd, J = 13.5, 4.5 Hz, 1H, Cys $\beta C\underline{H}_aH_b$), 3.02 (dd, J = 13.5, 4.5 Hz, 1H, Cys $\beta CH_a\underline{H}_b$), 2.77 (d, J = 6.0 Hz, 2H, S-glyceryl CH_2), 2.29 (t, J = 7.7 Hz, 2H, Pal αCH_2), 1.67 – 1.56 (m, 2H, Pal βCH_2), 1.49 (s, 9H, *t*Bu CH_3 x 3), 1.33 – 1.16 (m, 24H, Pal CH_2 x 12), 0.88 (t, J = 7.0 Hz, 3H, Pal CH_3) ppm;

^{13}C NMR (151 MHz, $CDCl_3$) δ 173.3 (Pal C=O), 173.1 (Pal C=O), 169.5 (Cys C=O), 155.7 (Fmoc C=O), 143.8 (Fmoc qC), 141.3 (Fmoc qC), 127.7 (Fmoc Ar-CH), 127.1 (Fmoc Ar-CH), 125.1 (Fmoc Ar-CH), 120.0 (Fmoc Ar-CH), 83.0 (*t*Bu qC), 70.2 (S-glyceryl CH), 67.2 (Fmoc CH_2), 63.5 (S-glyceryl OCH_2), 54.3 (Cys αCH), 47.1 (Fmoc CH), 35.3 (Cys βCH_2), 34.1 (Pal αCH_2), 33.3 (S-glyceryl CH_2), 31.9, 29.7, 29.6, 29.5, 29.3, 29.2, 29.1 (Pal CH_2), 28.0 (*t*Bu-CH_3), 24.9 (Pal βCH_2), 22.7 (Pal CH_2), 14.1 (Pal CH_3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{57}H_{60}D_{31}NNaO_8S$ = 1003.8303 $(M + Na)^+$. Found = 1003.8310;

ν_{max} (thin film)/cm^{-1}: 2925 (CH_2), 2854 (CH_2), 1737 (CO), 1467 (CH_2), 1369 (CH_3), 1248 (CO), 1155 (CO).

_N_-(((9_H_-Fluoren-9-yl)methoxy)carbonyl)-_S_-((_R_)-2-((hexadecanoyl-_d_$_{31}$)oxy)-3-(palmitoyloxy)propyl)-L-cysteine (81)



Prepared as per procedure B using compound **83** (0.524 g, 0.535 mmol). The solvents were removed *in vacuo* and the product recrystallised from a minimum amount of CH_2Cl_2 (5 mL) layered with MeOH (50 mL) for 16 h to yield the product as a white solid (0.450 g, 91%).

R_f = 0.53 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.3 Hz, 2H, Fmoc Ar-CH), 7.61 (d, *J* = 7.3 Hz, 2H, Fmoc Ar-CH), 7.40 (t, *J* = 7.3 Hz, 2H, Fmoc Ar-CH), 7.32 (t, *J* = 7.3 Hz, 2H, Fmoc Ar-CH), 5.74 (d, *J* = 7.3 Hz, 1H, NH), 5.20 – 5.14 (m, 1H, S-glyceryl CH), 4.68 – 4.62 (m, 1H, Cys αCH), 4.41 (d, *J* = 7.0 Hz, 2H, Fmoc CH_2), 4.35 (dd, *J* = 12.0, 2.8 Hz, 1H, S-glyceryl OC$\underline{H}_aH_b$), 4.24 (t, *J* = 7.0 Hz, 1H, Fmoc CH), 4.15 (dd, *J* = 12.0, 5.7 Hz, 1H, S-glyceryl OC$H_a\underline{H}_b$), 3.16 (dd, *J* = 13.8, 4.1 Hz, 1H, Cys βC$\underline{H}_aH_b$), 3.08 (dd, *J* = 13.8, 4.5 Hz, 1H, Cys βC$H_a\underline{H}_b$), 2.82 – 2.71 (m, 2H, S-glyceryl CH_2), 2.30 (t, *J* = 7.6 Hz,

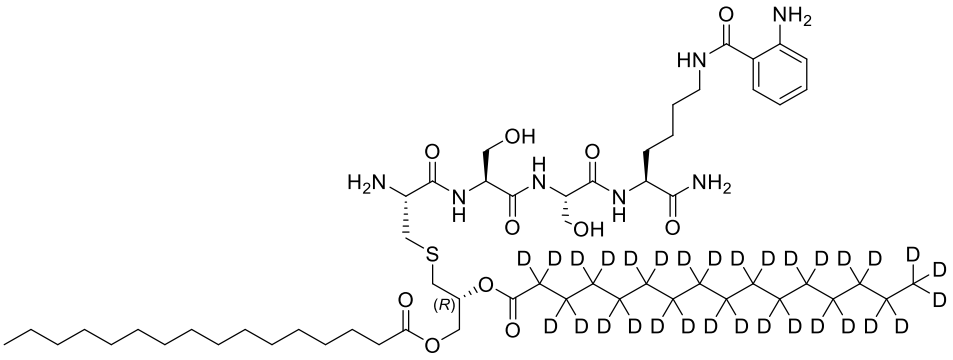
2H, Pal αCH$_2$), 1.66 – 1.51 (m, 2H, Pal βCH$_2$), 1.39 – 1.07 (m, 24H, Pal CH$_2$ x 12), 0.88 (t, *J* = 7.0 Hz, 3H, Pal-CH$_3$) ppm;

^{13}C NMR (151 MHz, CDCl$_3$) δ 173.7 (Pal C=O), 173.0 (Pal C=O), 156.0 (Cys C=O), 143.8 (Fmoc qC), 141.5 (Fmoc qC), 127.9 (Fmoc Ar-CH), 127.3 (Fmoc Ar-CH), 125.3 (Fmoc Ar-CH), 120.2 (Fmoc Ar-CH), 70.4 (S-glycery CH), 67.6 (Fmoc CH$_2$), 63.7 (S-glyceryl OCH$_2$), 53.6 (Cys αCH), 47.2 (Fmoc CH), 34.7 (Cys βCH$_2$), 34.3 (Pal CH$_2$), 33.1 (S-glycery CH$_2$), 32.1, 29.9, 29.8, 29.7, 29.6, 29.5, 29.4, 29.3, 25.0, 22.9 (Pal CH$_2$), 14.3 (Pal CH$_3$) ppm;

m/z HRMS (ESI$^+$) calcd. for $C_{53}H_{52}D_{31}NO_8SNa$ = 947.7677 $(M + Na)^+$. Found = 947.7669;

ν_{max} (thin film)/cm^{-1}: 1741 (CO), 1498 (CH$_2$), 1352 (CH$_3$), 1270 (CO).

(7*S*,10*S*,13*S*,16*R*,20*R*)-16-Amino-1-(2-aminophenyl)-7-carbamoyl-10,13-bis(hydroxymethyl)-1,9,12,15,23-pentaoxo-22-oxa-18-thia-2,8,11,14-tetraazaoctatriacontan-20-yl hexadecanoate-*d*$_{31}$ (84)



Fmoc-Cys[(*R*)Pam,Pam-*d*$_{31}$]-OH **81** (494 mg, 0.534 mmol) was coupled to peptide resin **76** (0.178 mmol) according to procedure I. The resin was then treated with 20% (v/v) piperidine in DMF (5 mL) for 2 x 10 min to remove the Cys Fmoc group and washed with DMF (3 x 5 mL), CH$_2$Cl$_2$ (3 x 5 mL) and DMF (3 x 5 mL). The peptide was subsequently cleaved from the resin as per procedure J, purified by column chromatography (SiO$_2$, CH$_2$Cl$_2$ – CH$_2$Cl$_2$:MeOH 9:1 v/v), and dried *in vacuo* at 0 °C to yield the target peptide as a white solid (18.0 mg, 9%).

R_f = 0.31 (CH$_2$Cl$_2$:MeOH 9:1 v/v);

^{1}H NMR (600 MHz, DMSO-*d*$_6$) δ 8.23 – 8.11 (m, 3H, Lys-εNH, Ser^{+2} NH, Ser^{+3} NH), 7.85 (d, *J* = 7.8 Hz, 1H, Lys NH), 7.44 (d, *J* = 8.0 Hz, 1H, Abz Ar-CH), 7.15 – 7.04 (m,

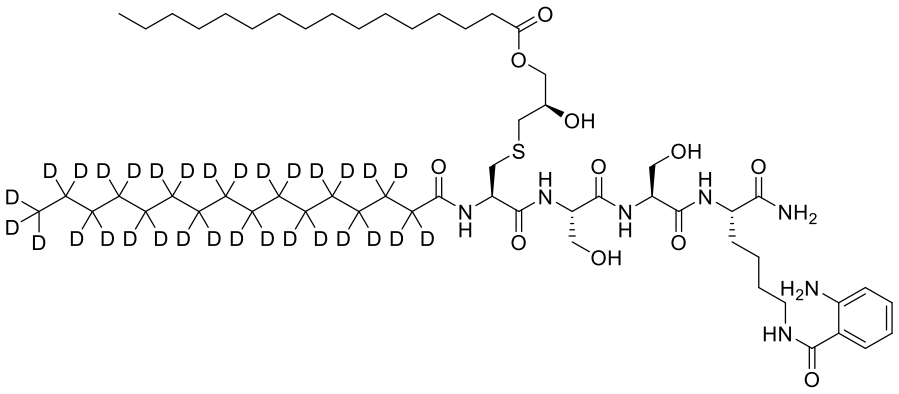
3H, $CONH_2$, Abz Ar-CH), 6.66 (d, *J* = 8.0, Hz, 1H, Abz Ar-CH), 6.56 – 6.43 (m, 1H, Abz Ar-CH), 6.34 (s, 2H, Abz NH_2), 5.20 (t, *J* = 5.6 Hz, 1H, Ser^{+2} OH), 5.12 – 5.07 (m, 1H, S-glyceryl CH), 5.05 (t, *J* = 5.6 Hz, 1H, Ser^{+3} OH), 4.42 – 4.34 (m, 1H, Ser^{+2} αCH), 4.34 – 4.24 (m, 2H, Ser^{+3} αCH, S-glyceryl $OC\underline{H_a}H_b$), 4.20 – 4.03 (m, 2H, Lys αCH, S-glyceryl $OCH_a\underline{H_b}$), 3.75 – 3.61 (m, 2H, Ser^{+2} $C\underline{H_a}H_b$, Ser^{+3} $C\underline{H_a}H_b$), 3.61 – 3.55 (m, 1H, Ser^{+3} $CH_a\underline{H_b}$), 3.55 – 3.47 (m, 1H, Ser^{+2} $CH_a\underline{H_b}$), 3.41 – 3.37 (m, 1H, Cys αCH), 3.16 (dd, *J* = 13.0, 7.1 Hz, 2H, Lys εCH_2), 2.87 (dd, *J* = 13.0, 4.2 Hz, 1H, Cys $C\underline{H_a}H_b$), 2.79 (dd, *J* = 14.0, 5.9 Hz, 1H, S-glyceryl $C\underline{H_a}H_b$), 2.69 (dd, *J* = 14.0, 7.3 Hz, 1H, S-glyceryl $CH_a\underline{H_b}$), 2.59 (dd, *J* = 13.0, 8.1 Hz, 1H, Cys $CH_a\underline{H_b}$), 2.26 (t, *J* = 7.3 Hz, 2H, Pal αCH_2), 1.79 – 1.69 (m, 1H, Lys $\beta C\underline{H_a}H_b$), 1.56 – 1.43 (m, 5H, Lys $\beta CH_a\underline{H_b}$, Pal CH_2, Lys δCH_2), 1.37 – 1.30 (m, 2H, Lys γCH_2), 1.29 – 1.17 (m, 24H, Pal CH_2 x 12), 0.85 (t, *J* = 6.9 Hz, 3H, Pal CH_3) ppm;

^{13}C NMR (151 MHz, DMSO-d_6) δ 173.7 ($CONH_2$), 172.5 (Pal C=O), 172.3 (Pal C=O), 170.5 (Ser^{+2} C=O), 169.7 (Ser^{+3} C=O), 168.7 (Abz C=O), 149.5 (Abz qC-NH_2), 131.4 (Abz Ar-CH), 128.0 (Abz Ar-CH), 116.2 (Abz Ar-CH), 115.0 (Abz qC), 114.5 (Abz Ar-CH), 70.0 (S-glyceryl CH), 63.5 (S-glyceryl OCH_2), 61.9 (Ser^{+2} βCH_2), 61.4 (Ser^{+3} βCH_2), 55.5 (Ser^{+3} αCH), 54.5 (Ser^{+2} αCH), 54.2 (Cys αCH), 52.8 (Lys αCH), 38.7 (Lys εCH_2), 37.4 (Cys βCH_2), 33.4 (Pal αCH_2), 31.7 (S-glyceryl CH_2), 31.3 (Lys βCH_2), 29.1, 28.9, 28.9, 28.9, 28.7, 28.7, 28.4 (Pal CH_2), 24.4 (Pal βCH_2), 23.1 (Lys γCH_2), 22.1 (Pal CH_2), 13.9 (Pal-CH_3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{57}H_{71}D_{31}N_7O_{11}S$ = 1123.9226 $(M + H)^+$. Found = 1123.9332;

ν_{max} (thin film)/cm^{-1}: 3287 (Amide A: NH), 2920 (CH_2), 2851 (CH_2), 1676 (Amide I: CO, CN), 1545 (Amide II: CN, NH), 1261 (Amide III: CN, CO), 1021 (CO).

(7*S*,10*S*,13*S*,16*R*,20*R*)-1-(2-Aminophenyl)-7-carbamoyl-16-(hexadecanamido-2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,12,12,13,13,14,14,15,15,16,16,16-d_{31})-20-hydroxy-10,13-bis(hydroxymethyl)-1,9,12,15-tetraoxo-18-thia-2,8,11,14-tetraazahenicosan-21-yl palmitate (85)



60 parallel reactions, each containing deuterated substrate **84** (8.4 µg, 150 µM), Lit (10 µM), MES/NaOH pH 5.4 (25 mM), NaCl (200 mM), LMNG (0.02% w/v) at a final reaction volume of 50 µL were incubated at 37 °C for 18 h without shaking and then stopped by flash-freezing in liquid N_2. EtOH:H_2O (50 µL; 7:3 v/v) was added to each reaction which was then vortexed for 10-20 s at 20 °C. $CHCl_3$ (30 µL) was added and the mixture vortexed for a further 60 s. Centrifugation for 2 min at 13,000 × *g* and 20 °C was performed to facilitate phase separation. The organic layer was collected and centrifuged for 2 min at 13,000 × *g* and 20 °C, and then the collected organic phase from each reaction was spotted on a silica gel F_{254} TLC plate. The plate was dried *in vacuo* to remove residual solvents and then developed using a mobile phase consisting of $CHCl_3$:MeOH:28% NH_4OH_{aq} (8:2:0.1 v/v/v). The product band from each reaction, visualised on the TLC plate under UV light at 254 nm, was collected. The compound was extracted from SiO_2 by addition of $CHCl_3$:MeOH (19:1 v/v; 2 x 5 mL) and the silica was pelleted by centrifugation 1,000 x *g* for 2 min. The supernatant was collected, and the solvent was evaporated under a stream of N_2. Residual solvents were removed *in vacuo* to yield a white solid which was dissolved in DMSO-d_6 (0.6 mL) for NMR analysis.

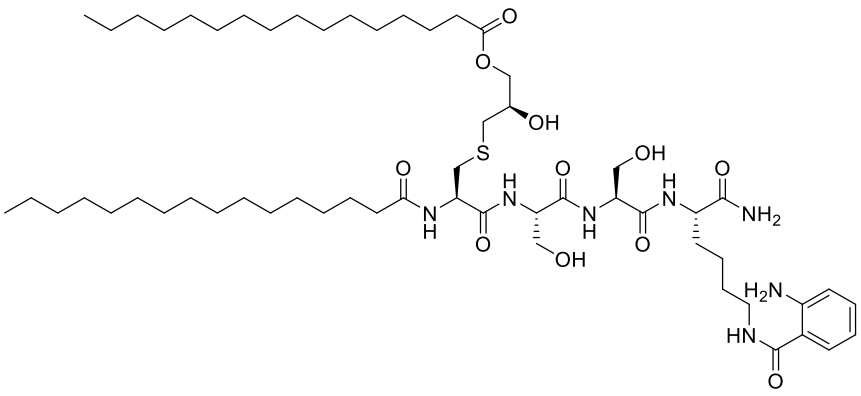
R_f = 0.71 ($CHCl_3$:MeOH:NH_4OH 8:2:0.1);

^{1}H NMR (600 MHz, DMSO-d_6) δ 8.18 – 8.14 (m, 1H, Lys εNH), 8.11 – 8.08 (m, 1H, Ser NH), 8.07 – 8.01 (m, 2H, Cys NH, Ser NH), 7.85 (d, *J* = 7.8 Hz, 1H, Lys NH), 7.44 (d, *J* = 7.3 Hz, 1H, Abz Ar-CH), 7.13 – 7.09 (m, 3H, $CONH_2$, Abz Ar-CH), 6.66 (d, *J* = 8.1 Hz, 1H, Abz Ar-CH), 6.51 – 6.47 (m, 1H, Abz Ar-CH), 6.34 (s, 2H, Abz-NH_2), 5.19 –

5.13 (m, 2H, Ser OH, S-glyceryl CH-O<u>H</u>), 5.07 – 5.03 (m, 1H, Ser OH), 4.50 – 4.45 (m, 1H, Cys αCH), 4.36 – 4.31 (m, 1H, Ser αCH), 4.29 – 4.24 (m, 1H, Ser αCH), 4.14 – 4.08 (m, 1H, Lys αCH), 4.05 – 3.93 (m, 2H, S-glyceryl OCH_2), 3.83 – 3.77 (m, 1H, S-glyceryl CH), 3.71 – 3.61 (m, 2H, Ser C<u>H</u>$_a$H$_b$, Ser C<u>H</u>$_a$H$_b$), 3.59 – 3.52 (m, 2H, Ser CH$_a$<u>H</u>$_b$, Ser CH$_a$<u>H</u>$_b$), 3.19 – 3.14 (m, 2H, Lys εCH_2), 2.92 – 2.86 (m, 1H, Cys C<u>H</u>$_a$H$_b$), 2.68 – 2.62 (m, 3H, Cys CH$_a$<u>H</u>$_b$, S-glyceryl CH_2), 2.31 – 2.25 (m, 2H, O-Pal αCH_2), 1.77 – 1.70 (m, 1H, Lys βC<u>H</u>$_a$H$_b$), 1.54 – 1.43 (m, 5H, Lys βCH$_a$<u>H</u>$_b$, Pal CH_2, Lys *δ*CH_2), 1.30 – 1.19 (m, 24H, Pal CH_2 x 12), 0.85 (t, *J* = 6.9 Hz, 3H, Pal CH_3) ppm;

m/z HRMS (ESI^+) calcd. for $C_{57}H_{71}D_{31}N_7O_{11}S$ = 1123.9226 $(M + H)^+$. Found = 1123.9385.

(7*S*,10*S*,13*S*,16*R*,20*R*)-1-(2-Aminophenyl)-7-carbamoyl-20-hydroxy-10,13-bis(hydroxymethyl)-1,9,12,15-tetraoxo-16-palmitamido-18-thia-2,8,11,14-tetraazahenicosan-21-yl palmitate (88)



14 parallel reactions, each containing protiated peptide **79** (6.9 µg, 150 µM), Lit (10 µM), MES/NaOH pH 5.4 (25 mM), NaCl (200 mM), LMNG (0.02% w/v) at a final volume of 50 µL were incubated at 37 °C for 18 h without shaking and then stopped by flash-freezing in liquid N_2. EtOH:H_2O (50 µL; 7:3 v/v) was added to each reaction which was then vortexed for 10-20 s at 20 °C. $CHCl_3$ (30 µL) was added and the mixture vortexed for a further 60 s. Centrifugation for 2 min at 13,000 × *g* and 20 °C was performed to facilitate phase separation. The organic layer was collected and centrifuged for 2 min at 13,000 × *g* and 20 °C, and then the collected organic phase from each reaction was spotted on a silica gel F_{254} TLC plate. The plate was dried *in vacuo* to remove residual solvents and then developed using a mobile phase consisting of $CHCl_3$:MeOH:28% NH_4OH_{aq} (8:2:0.1 v/v/v). The product band from each reaction, visualised on the TLC plate under UV light at 254 nm, was collected. The compound was extracted from SiO_2 by addition

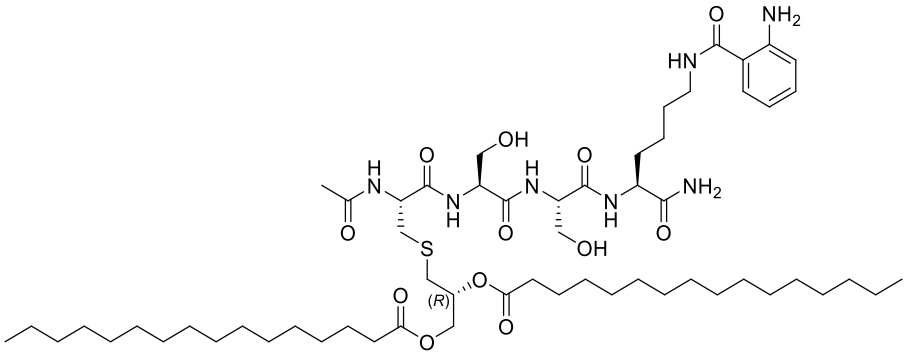
of $CHCl_3$:MeOH (19:1 v/v; 2 x 5 mL) and the silica was pelleted by centrifugation 1,000 x g for 2 min. The supernatant was collected, and the solvent was evaporated under a stream of N_2. Residual solvents were removed *in vacuo* to yield a white solid which was dissolved in DMSO-d_6 (0.6 mL) for NMR analysis.

R_f = 0.71 ($CHCl_3$:MeOH:NH_4OH 8:2:0.1);

^{1}H NMR (600 MHz, DMSO-d_6) δ 8.18 – 8.13 (m, 1H, Lys εNH), 8.14 – 8.08 (m, 1H, Ser NH), 8.08 – 8.01 (m, 2H, Cys NH, Ser NH), 7.86 (d, *J* = 8.1 Hz, 1H, Lys NH), 7.44 (d, *J* = 7.7 Hz, 1H, Abz Ar-CH), 7.14 – 7.07 (m, 3H, $CONH_2$, Abz Ar-CH), 6.66 (d, *J* = 8.1 Hz, 1H, Abz Ar-CH), 6.52 – 6.46 (m, 1H, Abz Ar-CH), 6.34 (s, 2H, Abz-NH_2), 5.21 – 5.11 (m, 2H, Ser-OH, S-glyceryl CH-O<u>H</u>), 5.11 – 5.01 (m, 1H, Ser OH), 4.52 – 4.43 (m, 1H, Cys αCH), 4.41 – 4.31 (m, 1H, Ser αCH), 4.31 – 4.21 (m, 1H, Ser αCH), 4.17 – 4.07 (m, 1H, Lys αCH), 4.02 – 3.91 (m, 2H, S-glyceryl OCH_2), 3.83 – 3.75 (m, 1H, S-glyceryl CH), 3.71 – 3.59 (m, 2H, Ser C<u>H$_a$</u>H$_b$, Ser C<u>H$_a$</u>H$_b$), 3.59 – 3.45 (m, 2H, Ser CH$_a$<u>H$_b$</u>, Ser CH$_a$<u>H$_b$</u>), 3.19 – 3.13 (m, 2H, Lys εCH_2), 2.93 – 2.85 (m, 1H, Cys C<u>H$_a$</u>H$_b$), 2.68 – 2.55 (m, 3H, Cys CH$_a$<u>H$_b$</u>, S-glyceryl CH_2), 2.34 – 2.23 (m, 2H, O-Pal αCH_2), 2.17 – 2.06 (m, 2H, NH-Pal αCH_2), 1.80 – 1.69 (m, 1H, Lys βC<u>H$_a$</u>H$_b$), 1.62 – 1.39 (m, 7H, Lys βCH$_a$<u>H$_b$</u>, Pal CH_2 x 2, Lys *δ*CH_2), 1.39 – 1.30 (m, 2H, Lys γCH_2), 1.30 – 1.10 (m, 48H, Pal CH_2 x 24), 0.89 – 0.79 (m, 6H, Pal CH_3 x 2) ppm;

m/z HRMS (ESI^+) calcd. for $C_{57}H_{102}N_7O_{11}S$ = 1092.7353 $(M + H)^+$. Found = 1092.7365.

(7*S*,10*S*,13*S*,16*R*,20*R*)-16-Acetamido-1-(2-aminophenyl)-7-carbamoyl-10,13-bis(hydroxymethyl)-1,9,12,15-tetraoxo-18-thia-2,8,11,14-tetraazahenicosane-20,21-diyl dipalmitate (97)



Fmoc-Cys[(*R*)Pam$_2$]-OH **50** (175 mg, 0.196 mmol) and then AcOH (0.011 mL, 0.196 mmol) were coupled to peptide resin **76** (0.065 mmol) according to procedure I. The

peptide was cleaved from the resin as per procedure J, purified by column chromatography (SiO_2, CH_2Cl_2 – CH_2Cl_2:MeOH 9:1 v/v), and dried *in vacuo* at 0 °C to yield the target peptide as a white solid (8.0 mg, 11%).

R_f = 0.51 (CH_2Cl_2:MeOH 9:1 v/v);

^{1}H NMR (600 MHz, DMSO-d_6) δ 8.24 – 8.10 (m, 3H, Lys εNH, Ser^{+2} NH, Ser^{+3} NH), 8.04 (d, *J* = 7.1 Hz, 1H, Cys NH), 7.86 (d, *J* = 8.1 Hz, 1H, Lys NH), 7.44 (d, *J* = 7.7 Hz, 1H, Abz Ar-CH), 7.23 – 6.99 (m, 3H, $CONH_2$, Abz Ar-CH), 6.67 (d, *J* = 8.1 Hz, 1H, Abz Ar-CH), 6.49 (t, *J* = 7.3 Hz, 1H, Abz Ar-CH), 6.34 (s, 2H, Abz NH_2), 5.16 (t, *J* = 5.1 Hz, 1H, Ser^{+2} OH), 5.14 – 5.07 (m, 1H, S-glyceryl CH), 5.04 (t, *J* = 5.2 Hz, 1H, Ser^{+3} OH), 4.53 – 4.47 (m, 1H, Cys αCH), 4.39 – 4.35 (m, 1H, Ser^{+2} αCH), 4.32 – 4.21 (m, 2H, Ser^{+3} αCH, S-glyceryl OCH_aH_b), 4.16 – 4.04 (m, 2H, Lys αCH, S-glyceryl OCH_aH_b), 3.72 – 3.60 (m, 2H, Ser^{+2} CH_aH_b, Ser^{+3} CH_aH_b), 3.62 – 3.48 (m, 2H, Ser^{+3} CH_aH_b, Ser^{+2} CH_aH_b), 3.21 – 3.13 (m, 2H, Lys εCH_2), 2.89 (dd, *J* = 13.8, 4.7 Hz, 1H, Cys CH_aH_b), 2.82 (dd, *J* = 14.0, 5.4 Hz, 1H, S-glyceryl CH_aH_b), 2.71 (dd, *J* = 14.0, 7.2 Hz, 1H, S-glyceryl CH_aH_b), 2.62 – 2.58 (m, 1H, Cys CH_aH_b), 2.31 – 2.21 (m, 4H, Pal αCH_2 x 2), 1.85 (s, 3H, CH_3CONH), 1.81 – 1.66 (m, 1H, Lys βCH_aH_b), 1.55 – 1.41 (m, 7H, Lys βCH_aH_b, Pal CH_2 x 2, Lys δCH_2), 1.39 – 1.34 (m, 2H, Lys γCH_2), 1.34 – 1.17 (m, 48H, Pal CH_2 x 24), 0.85 (t, *J* = 6.7 Hz, 6H, Pal CH_3 x 2) ppm;

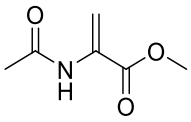
^{13}C NMR (151 MHz, DMSO-d_6) δ 173.7 ($CONH_2$), 172.5 (Pal C=O), 172.3 (Pal C=O), 170.5 (Ser^{+2} C=O), 170.2 (CH_3CONH), 169.7 (Ser^{+3} C=O), 169.4 (Cys C=O), 168.7 (Abz C=O), 149.5 (Abz qC-NH_2), 131.4 (Abz Ar-CH), 128.0 (Abz Ar-CH), 116.2 (Abz Ar-CH), 115.0 (Abz qC), 114.5 (Abz Ar-CH), 69.9 (S-glyceryl CH), 63.5 (S-glyceryl OCH_2), 61.6 (Ser βCH_2), 61.4 (Ser βCH_2), 55.4 (Ser^{+3} αCH), 54.9 (Ser^{+2} αCH), 52.7 (Lys αCH), 52.1 (Cys αCH), 38.7 (Lys εCH_2), 34.1 (Cys βCH_2), 33.6, 33.4 (Pal αCH_2), 31.6 (S-glyceryl CH_2), 31.3 (Lys βCH_2), 30.7, 29.1, 29.0, 28.9, 28.8, 28.7, 28.4 (Pal CH_2), 24.5, 24.4 (Pal βCH_2), 23.0 (Lys γCH_2), 22.5 (CH_3CONH), 22.1 (Pal CH_2), 13.9 (Pal CH_3) ppm;

m/z HRMS (ESI$^+$) calcd. for $C_{59}H_{104}N_7O_{12}S$ = 1134.7458 $(M + H)^+$. Found = 1134.7455;

ν_{max} (thin film)/cm^{-1} 2920 (CH_2), 2854 (CH_2), 1741 (Amide I: CO, CN), 1497 (Amide II: CN, NH), 1234 (Amide III: CN, CO), 1047 (CO).

7.6 Experimental Details for Chapter 5

Methyl 2-acetamidoacrylate (102)



To a solution of methyl acetyl-L-cysteinate **103** (3.00 g, 16.9 mmol) in DMF (40 mL) was added 1,4-diiodobutane (3.35 mL, 25.4 mmol) and K_2CO_3 (4.68 g, 33.9 mmol). The reaction mixture was stirred at rt for 4 h. The solution was diluted with H_2O (100 mL) and subsequently extracted with EtOAc (5 x 50 mL). The organic layers were combined and washed with sat. aq. NaCl solution (5 x 50 mL), dried over $MgSO_4$, filtered, and the solvent was evaporated under reduced pressure. The crude compound purified by column chromatography (SiO_2, hex – EtOAc:hex 1:1 v/v) to afford the product as a white sold (1.45 g, 60%). The isolated compound was in good agreement with the literature.[293]

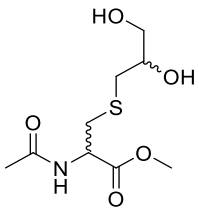
R_f = 0.43 (EtOAc/hex 1:1 v/v);

^{1}H NMR (400 MHz, DMSO-d_6) δ 9.35 (bs, 1H, NH), 6.12 (s, 1H, C=C$\underline{H_a}H_b$), 5.66 (s, 1H, C=C$H_a\underline{H_b}$), 3.75 (s, 3H, COOCH_3), 2.02 (s, 3H, NHCOC$\underline{H_3}$) ppm;

^{13}C-NMR (101 MHz, DMSO-d_6): δ 169.3 (NH$\underline{C}$OCH_3), 164.0 ($\underline{C}$OOCH_3), 133.0 (αC), 109.2 (C=$\underline{C}H_2$), 52.5 (COO$\underline{C}H_3$), 23.5 (NHCO$\underline{C}H_3$) ppm;

HRMS (*m*/*z* ESI$^+$) calcd. for $C_6H_{10}NO_3$ = 144.0655 $(M + H)^+$. Found 144.0629.

Methyl *N*-acetyl-*S*-(2,3-dihydroxypropyl)cysteinate (105)



Dha **102** (100 mg, 0.70 mmol) was dissolved in in aq. NH_4OAc solution (7.2 mL, 1.0 M) which was degassed by bubbling argon through the solution for 10 min and subsequently 3-mercapto-1,2-propanediol **104** (0.091 mL, 1.05 mmol) was added. The reaction was stirred at rt for 1 h and subsequently extracted with EtOAc (3 x 10 mL). The combined organic layers were washed with sat. aq. NaCl (2 x 20 mL), dried over $MgSO_4$ and filtered

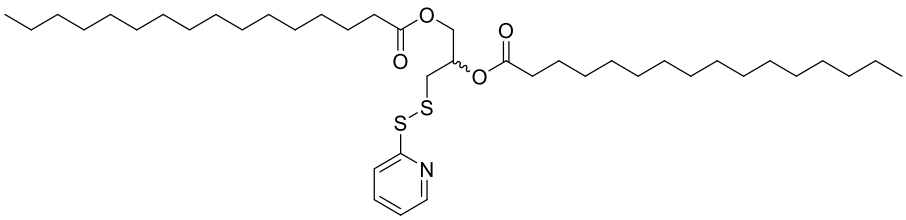
to yield the product as a white solid (148 mg, 84%). The isolated compound was in good agreement with the literature.[294]

R_f = 0.31 (EtOAc);

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.83 (t, *J* = 8.3 Hz, 1H, NH), 4.80 (ddd, *J* = 11.5, 7.4, 4.7 Hz, 1H, Cys αCH), 3.81 – 3.77 (m, 1H, S-glyceryl CH), 3.74 (s, 3H, $COOCH_3$), 3.72 – 3.62 (m, 1H, S-glyceryl $OC\underline{H_a}H_b$), 3.61 – 3.43 (m, 1H, S-glyceryl $OCH_a\underline{H_b}$), 3.07 – 2.90 (m, 2H, Cys $\beta C\underline{H_a}H_b$, S-glyceryl $C\underline{H_a}H_b$), 2.78 – 2.53 (m, 2H, Cys $\beta CH_a\underline{H_b}$, S-glyceryl $CH_a\underline{H_b}$), 2.04 (s, 3H, $NHCOC\underline{H_3}$) ppm;

HRMS (*m*/*z* ESI$^-$) calcd. for $C_9H_{16}NO_5S$ = 250.0827 $(M - H)^-$. Found 250.0823.

3-(Pyridin-2-yldisulfaneyl)propane-1,2-diyl dipalmitate (100)



To a solution of 2,2-dipyridyldisulfide **106** (500 mg, 2.27 mmol) in anhydrous EtOH (20 mL) under argon at rt was added 3-mercapto-1,2-propanediol **104** (0.16 mL, 1.89 mmol) and acetic acid (0.13 mL, 2.27 mmol). The reaction was stirred at rt for 16 h and the solvents were evaporated *in vacuo* to yield crude **107** as a solid yellow residue which was used directly in the next step. A solution of palmitic acid (1.07 g, 4.16 mmol), EDC·HCl (0.97 g, 4.16 mmol) and DMAP (46 mg, 0.38 mmol) in CH_2Cl_2 (50 mL) was stirred under argon at 0 °C for 30 min. Crude diol **107** was dissolved in CH_2Cl_2 (50 mL) and added to the reaction under argon. The reaction was stirred at rt for 16 h and subsequently washed with aq. HCl (1 M; 3 x 50 mL) and sat. aq. $NaHCO_3$ (3 x 50 mL) and sat. aq. NaCl (3 x 50 mL). The organic layer was dried over $MgSO_4$, filtered and the solvents were removed *in vacuo*. The crude material was purified by column chromatography (SiO_2, EtOAc:hex 1:9 v/v) to yield a white solid (596 mg, 45%).

R_f = 0.60 (EtOAc/hex 2:8 v/v);

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.48 (d, *J* = 4.3 Hz, 1H, Ar-CH), 7.80 – 7.46 (m, 2H, Ar-CH x 2), 7.10 (t, *J* = 5.6 Hz, 1H, Ar-CH), 5.31 (dd, *J* = 9.8, 5.3 Hz, 1H, S-glyceryl CH), 4.38 (dd, *J* = 12.0, 5.3 Hz, 1H, S-glyceryl $OC\underline{H_a}H_b$), 4.21 (dd, *J* = 12.0, 5.3 Hz, 1H, S-

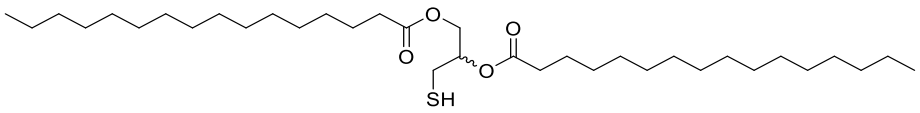
glyceryl $OCH_a\underline{H_b}$), 3.18 – 2.85 (m, 2H, S-glyceryl CH_2), 2.29 (t, J = 7.5 Hz, 4H, Pal αCH_2 x 2), 1.85 – 1.47 (m, 4H, Pal βCH_2 x 2), 1.47 – 1.08 (m, 48H, Pal CH_2 x 24), 0.88 (t, J = 6.9 Hz, 6H, Pal CH_3) ppm;

^{13}C NMR (151 MHz, $CDCl_3$) δ 173.4, 173.0 (C=O), 159.4 (Ar-qC), 150.0 (Ar-CH), 137.1 (Ar-CH), 121.1 (Ar-CH), 120.3 (Ar-CH), 70.0 (S-glyceryl CH), 63.6 (S-glyceryl OCH_2), 39.9 (S-glyceryl CH_2), 34.4 (Pal αCH_2), 34.2 (Pal αCH_2), 32.1, 29.9, 29.8, 29.7, 29.6, 29.5, 29.4, 29.4, 29.3, 29.2, 25.0, 22.8 (Pal CH_2), 14.3 (Pal CH_3) ppm;

HRMS (m/z ESI^+) calcd. for $C_{40}H_{72}NO_4S_2$ = 694.4897 $(M + H)^+$. Found 694.4897;

ν_{max} (thin film)/cm^{-1} 2916 (NH), 2849 (CH_2), 1977 (Ar CH), 1735 (ester C=O), 1155 (ester C-O), 751 (1,2-disubstitued Ar CH).

3-Mercaptopropane-1,2-diyl dipalmitate (99)



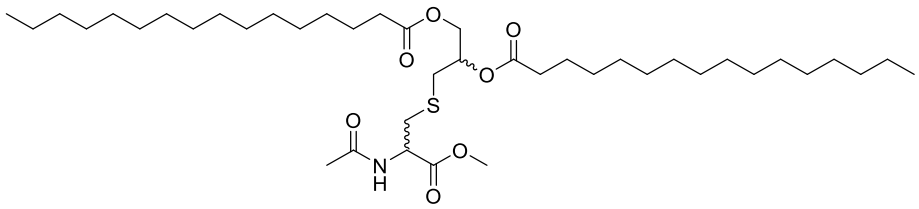
To a solution of disulfide **100** (450 mg, 0.65 mmol) in Et_2O:EtOH (30 mL; 5:1 v/v) at rt under argon was added PBu_3 (0.31 mL, 0.65 mmol) resulting in the immediate formation of a yellow solution. The reaction was stirred at rt for 30 min and solvents were subsequently evaporated. Cyclohexane (30 mL) was added, and this solution was washed with MeCN (3 x 50 mL). The cyclohexane layer was collected and dried *in vacuo* to yield a white solid (314 mg, 83%). The isolated compound was in good agreement with the literature.[295]

R_f = 0.79 (EtOAc/hex 2:8 v/v);

^1H NMR (600 MHz, $CDCl_3$) δ 5.12 – 5.01 (m, 1H, S-glyceryl CH), 4.35 (dd, J = 11.9, 3.9 Hz, 1H, S-glyceryl $OC\underline{H_a}H_b$), 4.23 (dd, J = 11.9, 5.7 Hz, 1H, S-glyceryl $OCH_a\underline{H_b}$), 2.80 – 2.68 (m, 2H, S-glyceryl CH_2), 2.37 – 2.27 (m, 4H, Pal αCH_2 x 2), 1.69 – 1.57 (m, 4H, Pal βCH_2 x 2), 1.37 – 1.21 (m, 48H, Pal CH_2 x 24), 0.88 (t, J = 7.0 Hz, 6H, Pal CH_3 x 2) ppm;

HRMS (m/z ESI^+) calcd. for $C_{35}H_{68}O_4SNa$ = 607.4731 $(M + Na)^+$. Found 607.4724.

3-((2-Acetamido-3-methoxy-3-oxopropyl)thio)propane-1,2-diyl dipalmitate (108)



Thiol **99** (20 mg 0.035 mmol), Dha **102** (5 mg, 0.035 mmol) and dodecylphosphocholine (49 mg, 0.14) were dissolved in $CHCl_3$:EtOH (20 mL; 1:1 v/v). The solvents were evaporated to yield a thin film. Aq. NH_4OAc (1 M; 16 mL) which had been degassed by bubbling argon through the solution for 10 min was subsequently added and the reaction was shaken at 37 °C for 4 h. EtOH (2 mL) was added and the reaction was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with sat. aq. NaCl (3 x 40 mL), dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo*. The crude residue was purified by column chromatography (SiO_2, EtOAc:hex 3:7 v/v) to yield the product as a white solid (12 mg, 48%). The product was isolated as a mixture of diastereomers.

R_f = 0.29 (EtOAc/hex 3:7 v/v);

^{1}H NMR (600 MHz, $CDCl_3$) mixture of 4 diastereomers δ 6.41 (d, *J* = 7.4 Hz, 0.5H, NH; diastereomer A+B), 6.36 (d, *J* = 7.4 Hz, 0.5H, NH; diastereomer C+D), 5.21 – 5.11 (m, 1H, S-glyceryl CH; diastereomer A-D), 4.93 – 4.78 (m, 1H, Cys αCH; diastereomer A-D), 4.40 – 4.27 (m, 1H, S-glyceryl OC$\underline{H}_aH_b$; diastereomer A-D), 4.13 (dd, *J* = 11.9, 6.1 Hz, 1H, S-glyceryl OC$H_a\underline{H}_b$; diastereomer A-D), 3.77 (s, 3H, $COOCH_3$; diastereomer A-D), 3.17 – 2.97 (m, 2H, Cys βCH_2; diastereomer A-D), 2.71 (d, *J* = 6.5 Hz, 4H, S-glyceryl CH_2; diastereomer A-D), 2.38 – 2.26 (m, 4H, Pal αCH_2 x 2; diastereomer A-D), 2.08 (s, 3H, NHCOC$\underline{H}_3$; diastereomer A-D), 1.66 – 1.57 (m, 4H, Pal βCH_2 x 2; diastereomer A-D), 1.36 – 1.20 (m, 48H, Pal CH_2 x 24; diastereomer A-D), 0.88 (t, *J* = 7.0 Hz, 6H, Pal CH_3 x 2; diastereomer A-D) ppm;

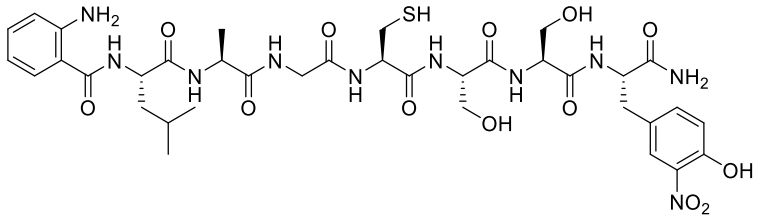
^{13}C NMR (151 MHz, $CDCl_3$) mixture of 4 diastereomers δ 173.6 (NH$\underline{C}OCH_3$; diastereomer A), 173.5 (NH$\underline{C}OCH_3$; diastereomer B), 173.4 (NH$\underline{C}OCH_3$; diastereomer C), 173.3 (NH$\underline{C}OCH_3$; diastereomer D), 171.3 ($\underline{C}OOCH_3$; diastereomer A), 171.2 ($\underline{C}OOCH_3$; diastereomer B), 170.2 ($\underline{C}OOCH_3$; diastereomer C), 170.1 ($\underline{C}OOCH_3$; diastereomer D), 70.4 (S-glyceryl CH; diastereomer A+B), 70.3 (S-glyceryl CH;

diastereomer C+D), 63.7 (S-glyceryl OCH_2; diastereomer A+B), 63.6 (S-glyceryl OCH_2; diastereomer C+D), 52.9 (αC; diastereomer A-D), 52.2 ($NHCO\underline{C}H_3$; diastereomer A+B), 52.0 ($NHCO\underline{C}H_3$; diastereomer A+B), 35.1 (S-glyceryl CH_2; diastereomer A+B), 35.0 (S-glyceryl CH_2; diastereomer C+D), 34.4 (Pal αCH; diastereomer A+B), 34.3 (Pal αCH; diastereomer A +B), 34.2 (Pal αCH; diastereomer C+D), 34.1 (Pal αCH; diastereomer C+D), 32.1 (Cys βCH_2; diastereomer A-D), 29.9, 29.8, 29.8, 29.8, 29.8, 29.7, 29.6, 29.5, 29.4, 29.3 (Pal CH_2; diastereomer A-D), 25.1, 25.0 (Pal CH_2; diastereomer A-D), 23.2 ($COO\underline{C}H_3$; diastereomer A+B), 23.1 ($COO\underline{C}H_3$; diastereomer C+D), 22.8 (Pal CH_2; diastereomer A-D), 14.3 (Pal CH_3; diastereomer A-D) ppm;

HRMS (*m*/*z* ESI^+) calcd. for $C_{41}H_{78}NO_7S$ = 728.4594 $(M + H)^+$. Found 728.5490;

ν_{max} (thin film)/cm^{-1} 3414 (NH), 2918 (CH), 2851 (CH_2), 1740 (ester C=O), 1251 (C-O).

2-Amino-*N*-((2*S*,5*S*,8*S*,11*R*,17*S*,20*S*)-1-amino-2-(4-hydroxy-3-nitrobenzyl)-5,8-bis(hydroxymethyl)-11-(mercaptomethyl)-17,22-dimethyl-1,4,7,10,13,16,19-heptaoxo-3,6,9,12,15,18-hexaazatricosan-20-yl)benzamide (110)



Fmoc-Tyr(3-NO_2)-OH **44** (628 mg, 1.40 mmol) was loaded to Rink amide AM resin (500 mg, 0.35 mmol) as per procedure E. Fmoc-Ser(*t*Bu)-OH (403 mg, 1.05 mmol) x 2 was coupled as per procedure I. Fmoc-Cys(Trt)-OH (615 mg, 1.05 mmol) was coupled using DIC (163 µL, 1.05 mmol) and HOBt (142 mg, 1.05 mmol) in DMF:CH_2Cl_2 (5 mL, 1:1 v/v) for 45 min. The resin then treated with 20% (v/v) piperidine/DMF (5 mL) for 2 x 10 min followed by coupling of Fmoc-Gly-OH (312 mg, 1.05 mmol), Fmoc-Ala-OH (327 mg, 1.05 mmol) x 2, Fmoc-Leu-OH (371 mg, 1.05 mmol) x 2 and Boc-Abz-OH **45** (249 mg, 1.05 mmol) as per procedure I. The resin was treated with 20% (v/v) piperidine/DMF (5 mL) for 2 x 15 min prior to cleavage. The peptide was cleaved from the resin using procedure J. The peptide was purified by column chromatography (SiO_2, CH_2Cl_2 – MeOH:CH_2Cl_2 1:9 v/v) yielding the target compound as a yellow solid (37 mg, 12%).

R_f = 0.26 (MeOH:CH_2Cl_2 1:9 v/v);

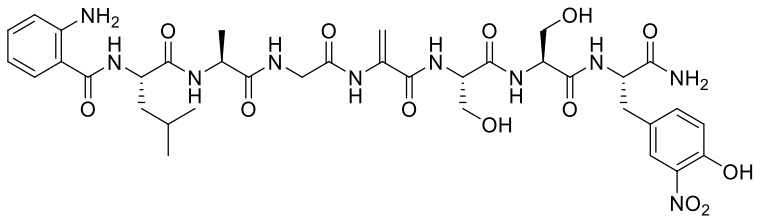
^{1}H NMR (600 MHz, DMSO-d_6) δ 10.73 (br s, 1H, Tyr OH), 8.21 – 8.09 (m, 3H, Ser NH, Leu NH, Gly NH), 8.06 – 7.92 (m, 4H, Ser NH, Ala NH, Tyr NH, Cys NH), 7.78 – 7.71 (m, 1H, Tyr Ar-CH), 7.55 (d, *J* = 7.8 Hz, 1H, Abz Ar-CH), 7.39 (d, *J* = 6.5 Hz, 1H, Tyr Ar-CH), 7.28 (s, 1H, CON$\underline{H_a}H_b$), 7.23 (s, 1H, CONH$_a\underline{H_b}$), 7.17 – 7.10 (m, 1H, Abz Ar-CH), 7.02 (d, *J* = 8.5 Hz, 1H, Tyr Ar-CH), 6.68 (d, *J* = 8.1 Hz, 1H, Abz Ar-CH), 6.61 – 6.44 (m, 1H, Abz Ar-CH), 6.32 (s, 2H, Abz NH_2), 5.16 – 5.10 (m, 1H, Ser OH), 5.00 – 4.93 (m, 1H, Ser OH), 4.53 – 4.41 (m, 2H, Leu αCH, Cys αCH), 4.41 – 4.32 (m, 2H, Tyr αCH, Ser αCH), 4.30 – 4.25 (m, 1H, Ala αCH), 4.25 – 4.17 (m, 1H, Ser αCH), 3.84 – 3.68 (m, 2H, Gly CH_2), 3.66 – 3.60 (m, 1H, Ser βC$\underline{H_a}H_b$), 3.60 – 3.52 (m, 2H, Ser βC$\underline{H_a}H_b$, Ser βCH$_a\underline{H_b}$), 3.52 – 3.44 (m, 1H, Ser βCH$_a\underline{H_b}$), 3.08 – 2.97 (m, 1H, Tyr βC$\underline{H_a}H_b$), 2.82 – 2.69 (m, 3H, Tyr βCH$_a\underline{H_b}$, Cys βCH_2), 1.73 – 1.59 (m, 2H, Leu CH_2), 1.58 – 1.46 (m, 1H, Leu CH), 1.24 (d, *J* = 7.0 Hz, 3H, Ala CH_3), 0.90 (d, *J* = 6.2 Hz, 3H, Leu CH_3), 0.87 (d, *J* = 6.2 Hz, 3H, Leu CH_3) ppm;

^{13}C NMR (151 MHz, DMSO-d_6) δ 172.6 (Leu C=O), 172.5 (Ala C=O), 172.3 (Tyr C=O), 170.2 (Ser C=O), 169.7 (Ser C=O), 169.6 (Cys C=O), 168.9 (Gly C=O), 168.7 (Abz C=O), 150.8 (Tyr qC), 149.5 (Abz-qC), 136.3 (Tyr qC), 136.1 (Tyr Ar-CH), 131.8 (Abz Ar-CH), 129.1 (Tyr qC), 128.5 (Abz Ar-CH), 125.3 (Tyr Ar-CH), 118.9 (Tyr Ar-CH), 116.3 (Abz Ar-CH), 114.6 (Abz Ar-CH), 61.6 (Ser βCH_2), 61.4 (Ser βCH_2), 55.6 (Ser αCH), 55.0 (Ser αCH), 54.7 (Cys αCH), 53.8 (Tyr αCH), 51.3 (Leu αCH), 48.3 (Ala αCH), 41.9 (Gly CH_2), 35.8 (Tyr βCH_2), 26.4 (Cys βCH_2), 24.5 (Leu CH), 23.2 (Leu CH_3), 21.3 (Leu CH_3), 18.2 (Ala CH_3) ppm;

HRMS (*m*/*z* ESI$^+$) calcd. for $C_{36}H_{51}N_{10}O_{13}S$ = 863.3280 $(M + H)^+$. Found 863.3284;

ν_{max} (thin film)/cm^{-1} 3410 (NH), 748 (Ar CH), 1681 (primary amide C=O), 1603 (C=O).

2-amino-*N*-((2*S*,5*S*,8*S*,17*S*,20*S*)-1-amino-2-(4-hydroxy-3-nitrobenzyl)-5,8-bis(hydroxymethyl)-17,22-dimethyl-11-methylene-1,4,7,10,13,16,19-heptaoxo-3,6,9,12,15,18-hexaazatricosan-20-yl)benzamide (111)



To a solution of peptide **110** (10 mg, 0.011 mmol) in DMF (0.5 mL) was added 1,4-diiodobutane (4.6 µL, mmol, 0.035) and K_2CO_3 (6.4 mg, 0.046mmol). The reaction was stirred at rt for 16 h and heptane (20 mL) was subsequently added, resulting in the precipitation of a yellow solid. The solid was filtered and washed with Et_2O (3 x 10 mL), acetone (3 x 10 mL), H_2O (3 x 10 mL) and acetone (3 x 10 mL) to yield **111** as a yellow solid (7.7 mg, 85%).

R_f = 0.41 (MeOH:CH_2Cl_2 1:9 v/v);

^{1}H NMR (600 MHz, DMSO-d_6) δ 10.74 (br s, 1H, Tyr OH), 9.13 (s, 1H, Dha NH), 8.32 – 8.18 (m, 2H, Gly NH, Ser NH), 8.18 – 8.04 (m, 2H, Leu NH, Ser NH), 8.04 – 7.90 (m, 2H, Tyr NH, Ala NH), 7.71 (d, *J* = 1.9 Hz, 1H, Tyr Ar-CH), 7.55 (d, *J* = 7.6 Hz, 1H, Abz Ar-CH), 7.48 (d, *J* = 8.7 Hz, 1H, Tyr Ar-CH), 7.31 (s, 1H, $CONH_a\underline{H_b}$), 7.29 – 7.19 (m, 2H, $CONH_a\underline{H_b}$, Tyr Ar-CH), 7.14 (t, *J* = 7.7 Hz, 1H, Abz Ar-CH), 6.68 (d, *J* = 8.2 Hz, 1H, Abz Ar-CH), 6.52 (t, *J* = 7.5 Hz, 1H, Abz Ar-CH), 6.33 (s, 2H, Abz NH_2), 6.11 (s, 1H, C=C$\underline{H_a}H_b$), 5.57 (s, 1H, C=C$H_a\underline{H_b}$), 5.20 – 5.11 (m, 1H, Ser OH), 5.08 – 4.92 (m, 1H, Ser OH), 4.50 – 4.42 (m, 2H, Leu αCH, Ser αCH), 4.42 – 4.26 (m, 2H, Ala αCH, Tyr αCH), 4.26 – 4.18 (m, 1H, Ser αCH), 3.97 – 3.70 (m, 2H, Gly CH_2), 3.66 (d, *J* = 5.7 Hz, 2H, Ser βCH_2), 3.61 – 3.53 (m, 1H, Ser βC$\underline{H_a}H_b$), 3.53 – 3.45 (m, 1H, Ser βC$H_a\underline{H_b}$), 3.06 (dd, *J* = 14.0, 4.0 Hz, 1H, Tyr βC$\underline{H_a}H_b$), 2.85 – 2.73 (m, 1H, Tyr βC$H_a\underline{H_b}$), 1.72 – 1.59 (m, 2H, Leu CH_2), 1.59 – 1.41 (m, 1H, Leu CH), 1.27 (d, *J* = 7.0 Hz, 3H, Ala CH_3), 0.90 (d, *J* = 6.3 Hz, 3H, Leu CH_3), 0.87 (d, *J* = 6.3 Hz, 3H, Leu CH_3) ppm;

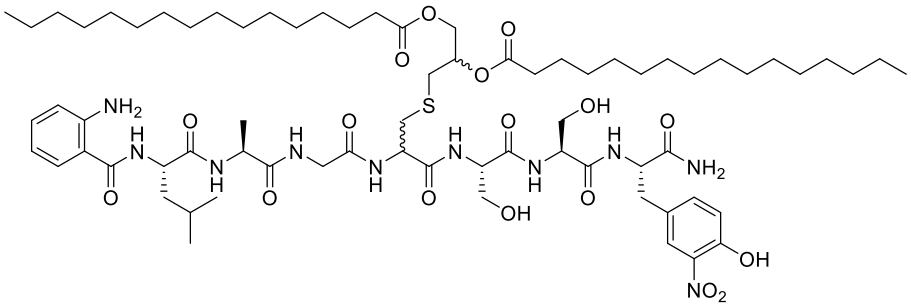
^{13}C NMR (151 MHz, DMSO-d_6) δ 172.8 (Leu C=O), 172.5 (Ala C=O), 172.2 (Tyr C=O), 170.1 (Ser C=O), 169.7 (Ser C=O), 168.8 (Gly C=O), 168.3 (Abz C=O), 163.8 (Dha C=O), 150.8 (Tyr qC), 149.5 (Abz-qC), 139.0 ($\underline{C}$=CH_2), 136.3 (Tyr qC), 136.1 (Tyr Ar-CH), 131.8 (Abz Ar-CH), 129.1 (Tyr qC), 128.5 (Abz Ar-CH), 125.4 (Tyr Ar-CH), 118.9

(Tyr Ar-CH), 116.3 (Abz Ar-CH), 114.6 (Abz Ar-CH), 103.8 (C=$\underline{C}H_2$), 68.0 (Ser βCH_2), 61.5 (Ser βCH_2), 55.6 (Ser αCH), 53.8 (Tyr αCH), 51.3 (Leu αCH), 49.6 (Ser αCH), 48.1 (Ala αCH), 42.8 (Gly CH_2), 35.8 (Tyr βCH_2), 25.1 (Leu CH), 24.5 (Leu CH), 23.2 (Leu CH_3), 21.3 (Leu CH_3), 18.4 (Ala CH_3) ppm;

HRMS (*m*/*z* ESI^+) calcd. for $C_{36}H_{49}N_{10}O_{13}$ = 829.3475 $(M + H)^+$. Found 829.3480;

ν_{max} (thin film)/cm^{-1} 3289 (NH), 2164 (C=C-H), 1649 (C=O), 1628 (C=C).

(3*S*,6*S*)-12-(((*S*)-1-(((*S*)-1-(((*S*)-1-amino-3-(4-hydroxy-3-nitrophenyl)-1-oxopropan-2-yl)amino)-3-hydroxy-1-oxopropan-2-yl)amino)-3-hydroxy-1-oxopropan-2-yl)carbamoyl)-1-(2-aminophenyl)-3-isobutyl-6-methyl-1,4,7,10-tetraoxo-14-thia-2,5,8,11-tetraazaheptadecane-16,17-diyl dipalmitate (109)



Thiol **99** (6.7 mg, 0.011 mmol), peptide **111** (6.4 mg, 0.0077 mmol) and dodecylphosphocholine (11 mg, 0.031 mmol) were dissolved in $CHCl_3$:EtOH (20 mL; 1:1 v/v). The solvents were evaporated to yield a thin film. Aq. NH_4OAc (1 M; 3.75 mL) which had been degassed by bubbling argon through the solution for 10 min was subsequently added and the reaction was shaken at 37 °C for 16 h. EtOH (2 mL) was added and the reaction was washed with CH_2Cl_2 (3 x 5 mL). The aq. layer was adsorbed onto SiO_2 and then product was isolated by column chromatography (SiO_2, MeOH:CH_2Cl_2 1:9 v/v) to yield a yellow solid (5.0 mg, 46%). The target compound was isolated as a mixture of diastereomers.

R_f = 0.32 (MeOH:CH_2Cl_2 1:9 v/v);

^{1}H NMR (600 MHz, DMSO-d_6) mixture of 4 diastereomers δ 8.21 (d, *J* = 7.3 Hz, 1H, Ser^{+2} NH), 8.17 – 7.97 (m, 5H, Leu NH, Gly NH, Cys NH, Ser^{+3} NH, Ala NH), 7.92 (d, *J* = 8.4 Hz, 1H, Tyr NH), 7.72 (s, 1H, Tyr Ar-CH), 7.55 (d, *J* = 7.3 Hz, 1H, Abz Ar-CH), 7.36 (s, 1H, Tyr Ar-CH), 7.29 – 7.18 (m, 2H, $CONH_2$), 7.13 (dd, *J* = 11.2, 4.2 Hz, 1H, Abz Ar-CH), 6.97 (s, 1H, Tyr Ar-CH), 6.68 (d, *J* = 8.4 Hz, 1H, Abz Ar-CH), 6.51 (t, *J* =

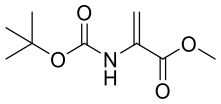
7.2 Hz, 1H, Abz Ar-CH), 6.33 (s, 2H, Abz NH_2), 5.24 – 5.13 (m, 1H, Ser^{+2} OH), 5.13 – 5.03 (m, 1H, S-glyceryl CH), 5.03 – 4.92 (m, 1H, Ser^{+3} OH), 4.66 – 4.51 (m, 1H, Cys αCH), 4.51 – 4.42 (m, 1H, Leu αCH), 4.42 – 4.33 (m, 2H, Ser^{+2} αCH, Tyr αCH), 4.33 – 4.24 (m, 2H, Ala αCH, S-glyceryl $OC\underline{H_a}H_b$), 4.24 – 4.16 (m, 1H, Ser^{+3} αCH), 4.14 – 4.03 (m, 1H, S-glyceryl $OCH_a\underline{H_b}$), 3.86 – 3.69 (m, 2H, Gly CH_2), 3.69 – 3.60 (m, 1H, Ser^{+2} $C\underline{H_a}H_b$), 3.60 – 3.51 (m, 2H, Ser^{+2} $CH_a\underline{H_b}$, Ser^{+3} $C\underline{H_a}H_b$), 3.51 – 3.44 (m, 1H, Ser^{+3} $CH_a\underline{H_b}$), 3.07 – 2.97 (m, 1H, Tyr $C\underline{H_a}H_b$), 2.90 (dd, *J* = 13.8, 4.6 Hz, 1H, Cys $C\underline{H_a}H_b$), 2.82 (dd, *J* = 13.9, 5.5 Hz, 1H, S-glyceryl $C\underline{H_a}H_b$), 2.77 – 2.64 (m, 2H, Tyr $CH_a\underline{H_b}$, S-glyceryl $CH_a\underline{H_b}$), 2.34 – 2.16 (m, 4H, Pal αCH_2 x 2), 1.79 – 1.58 (m, 3H, Leu CH_2, Leu $C\underline{H}$-CH_3), 1.58 – 1.42 (m, 7H, Pal CH_2 x 2, Ala CH_3), 1.35 – 1.11 (m, 48H, Pal CH_2 x 24), 0.98 – 0.75 (m, 12H, Leu CH_3 x 2, Pal CH_3 x 2) ppm;

^{13}C NMR (151 MHz, DMSO-d_6) mixture of 4 diastereomers δ 172.6 (Pal C=O), 172.5 (Pal C=O), 172.4 ($CONH_2$), 172.3 (Leu C=O), 170.2 (Ala C=O), 170.1 (Ser^{+2} C=O), 169.7 (Cys C=O), 168.9 (Ser^{+3} C=O), 168.8 (Abz C=O), 168.7 (Gly C=O), 150.9 (Abz-qC NH_2), 149.7 (qC), 136.3 (Tyr Ar-CH), 136.1 (Tyr Ar-CH), 131.8 (Abz Ar-CH), 129.1 (Abz Ar-CH), 125.3 (Tyr Ar-CH), 118.9 (Abz Ar-CH), 116.3 (Abz Ar-CH), 114.5 (Abz qC), 69.9 (S-glyceryl CH), 63.5 (Ser^{+2} βCH_2), 61.7 (Ser^{+3} βCH_2), 55.6 (Ser^{+3} αCH), 54.8 (Ser^{+2} αCH), 53.8 (Tyr αCH), 52.2 (Cys αCH), 51.3 (Leu αCH), 48.8 (Ala αCH), 41.8 (Gly CH_2), 35.8 (Tyr βCH_2), 34.0 (Pal αCH_2), 33.6 (Cys βCH_2 [diastereomer A+B]), 33.4 (Cys βCH_2 [diastereomer C+D]), 31.6 (Pal CH_2), 31.6 ((S-glyceryl CH_2), 29.1, 29.0, 28.9, 28.8, 28.7, 28.5, 28.4 (Pal CH_2), 24.4 (Leu CH)), 24.1 (Pal CH_2), 24.0 (Leu CH_3), 23.2 (Leu CH_3), 22.1 (Pal CH_2), 21.5 (Leu CH_3), 18.2 (Ala CH_3), 13.9 (Pal CH_3) ppm;

HRMS (*m/z* MALDI$^+$) calcd. for $C_{71}H_{116}N_{10}O_{17}SNa$ = 1435.8138 $(M + Na)^+$. Found = 1435.8108;

ν_{max} (thin film)/cm^{-1} 3408, 3280 (NH), 2929 (CH_2), 2848 (CH_2), 1638 (Amide I: CO, CN), 1520 (Amide II: CN, NH), 1363 (Amide III: CN, CO, CC), 1209 (CO), 748 (Ar CH).

Methyl 2-((*tert*-butoxycarbonyl)amino)acrylate (113a)



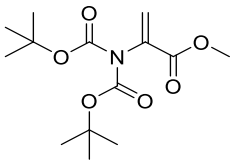
To a solution of Boc-Cys-OMe **64** (3.50 g, 14.70 mmol) in DMF (100 mL) was added 1,4-diiodobutane (3.0 mL, 22.05 mmol) and K_2CO_3 (4.10 g, 29.40 mmol). The reaction mixture was stirred at rt for 4 h. The solution was diluted with H_2O (100 mL) and subsequently extracted with EtOAc (3 x 75 mL). The organic layers were combined and washed with sat. aq. NaCl (3 x 100 mL), dried over $MgSO_4$, filtered, and the solvent was evaporated to dryness. Purification by column chromatography (SiO_2, hex:EtOAc 98:2 – 95:5 v/v) gave compound **113a** as a white solid (1.60 g, 54%). The isolated compound was in good agreement with the literature.[262]

R_f = 0.30 (hex:EtOAc 95:5 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.01 (bs, 1H, NH), 6.16 (s, 1H, C=C$\underline{H_a}H_b$), 5.72 (d, *J* = 1.5 Hz, 1H, C=C$H_a\underline{H_b}$), 3.83 (s, 3H, $COOCH_3$), 1.48 (s, 9H, Boc CH_3 x 3) ppm;

^{13}C NMR (101 MHz, $CDCl_3$) δ 165.3 (C=O), 152.7 (Boc C=O), 131.3 (C=CH_2), 105.0 (C=CH_2), 80.6 (Boc qC), 52.8 ($COOCH_3$), 28.2 (Boc CH_3) ppm.

Methyl 2-(bis(*tert*-butoxycarbonyl)amino)acrylate (113d)



To a solution of Boc-Dha-OMe **113a** (1.60 g, 7.80 mmol) in MeCN (80 mL), Boc_2O (3.70 g, 17.17 mmol) and DMAP (0.19 g, 1.56 mmol) were added. The reaction was stirred at rt for 16 h. The solvent was removed *in vacuo* and the crude material was purified by column chromatography (SiO_2, EtOAc:hex 1:9 v/v) to furnish compound **113d** as a white solid (2.30 g, 98%). The isolated compound was in good agreement with the literature.[296]

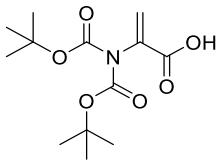
R_f = 0.79 (hex:EtOAc 95:5 v/v);

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.32 (s, 1H, C=C$\underline{H_a}H_b$), 5.62 (s, 1H, C=C$H_a\underline{H_b}$), 3.78 (s, 3H, $COOCH_3$), 1.44 (s, 18H, Boc CH_3 x 6) ppm;

^{13}C NMR (101 MHz, $CDCl_3$) δ 164.1 (C=O), 150.8 (Boc C=O), 136.2 (<u>C</u>=CH_2), 124.8 (C=<u>C</u>H_2), 83.3 (Boc qC), 52.5 ($COOCH_3$), 28.1 (Boc CH_3) ppm;

HRMS (*m*/*z* ESI$^+$) calcd. for $C_{14}H_{23}NNaO_6$ 324.1422 $(M + Na)^+$. Found 324.1421.

2-(Di(*tert*-butoxycarbonyl)amino)acrylic acid (116)



To a solution of Boc_2-Dha-OMe **113d** (6.80 g, 15.65 mmol) in dioxane:H_2O (150 mL; 1:1 v/v) was added LiOH (0.94 g, 39.12 mmol) and the mixture was stirred at rt for 1 h. The organic solvent was removed *in vacuo* and the aq. solution was diluted with H_2O (80 mL), acidified with aq. HCl (6 M) to pH 3 and the crude carboxylic acid was extracted with EtOAc (3 x 50 mL). The organic layers were combined and washed with sat. aq. NaCl (3 x 50 mL), dried over $MgSO_4$, filtered, and concentrated to furnish compound **116** as a white solid (3.00 g, 67%).

R_f = 0.35 (EtOAc:hex 1:9 v/v);

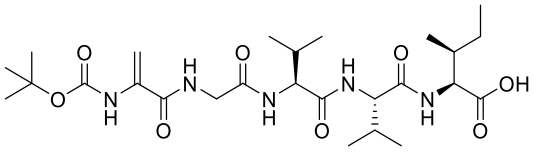
^{1}H NMR (400 MHz, $CDCl_3$) δ 6.46 (s, 1H, C=C<u>H$_a$</u>H$_b$), 5.77 (s, 1H, C=CH$_a$<u>H$_b$</u>), 1.47 (s, 18H, Boc CH_3 x 6) ppm;

^{13}C NMR (101 MHz, $CDCl_3$) δ 168.2 (C=O), 150.8 (Boc C=O), 135.6 (<u>C</u>=CH_2), 126.7 (C=<u>C</u>H_2), 83.7 (Boc qC), 28.1 (Boc CH_3) ppm;

HRMS (*m*/*z* ESI$^+$) calcd. for $C_{13}H_{21}NNaO_6$ = 310.1268 $(M + Na)^+$. Found 310.1261;

ν_{max} (film)/cm^{-1} 1723 (C=O), 2981 (C=C).

(2-((tert-butoxycarbonyl)amino)acryloyl)glycyl-L-valyl-L-valyl-L-isoleucine (131)

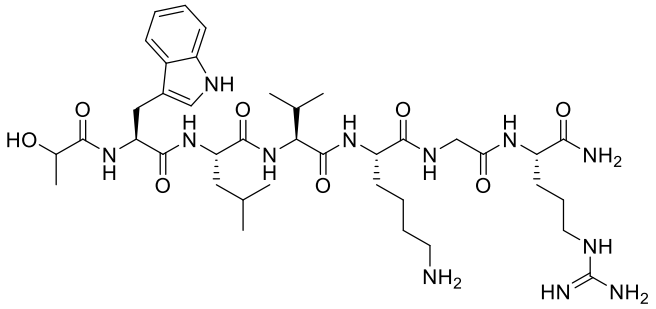


Fmoc-Ile-OH (530 mg, 1.50 mmol) was loaded to 2-chlorotrityl chloride resin (340 mg, 0.50 mmol) as per procedure G. Fmoc-Val-OH (505 mg, 1.50 mmol) x 2 and Fmoc-Gly-OH (456 mg, 1.50 mmol) were coupled according to procedure I. A solution of Boc_2-Dha-OH **116** (452 mg, 1.50 mmol), HCTU (619 mg, 1.50 mmol) and DIPEA (0.52 mL,

3.00 mmol) in DMF (5 mL) was preactivated for 30 min at 0 °C before addition of this solution to the resin which was subsequently agitated for 10 h at rt. Subsequent resin washes with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL) were followed by drying of the resin under reduced pressure. CH_2Cl_2 (5 mL) was added to the peptide resin which was agitated for 20 min and then drained. A solution of TFA (0.025 mL), TES (0.25 mL) and CH_2Cl_2 (4.75 mL) was added to the resin, which was agitated at rt for 1 min, then drained. The filtrate was concentrated under a flow of N_2, followed by precipitation of the peptide with Et_2O (10 mL) at 0 °C. The crude peptide suspension was centrifuged and the supernatant decanted. The peptide pellet was washed again with Et_2O (2 x 10 mL) at 0 °C, centrifuged and collected. The crude material was dried *in vacuo* to yield a while solid (14 mg, 5%).

HRMS (*m*/*z* ESI^-) calcd. for $C_{26}H_{45}N_5O_8$ = 555.3268 $(M - H)^-$. Found 555.3248.

(*S*)-2-((*S*)-2-((*S*)-2-((*S*)-3-(1*H*-Indol-3-yl)-2-(2-oxopropanamido)propanamido)-4-methylpentanamido)-3-methylbutanamido)-6-amino-*N*-(2-(((*S*)-1-amino-5-guanidino-1-oxopentan-2-yl)amino)-2-oxoethyl)hexanamide (133)



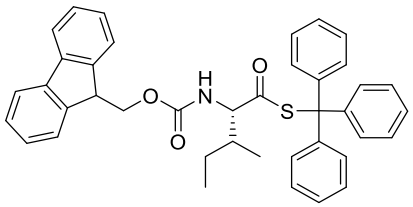
Peptide **133** was synthesised by coupling of Fmoc-Arg(Pbf)-OH (0.73 g, 1.12 mmol) to Rink amide resin (0.40 g, 0.28 mmol) as per procedure E. Fmoc-Gly-OH (0.25 g, 0.84 mmol), Fmoc-Lys(Boc)-OH (0.61 g, 0.84 mmol), Fmoc-Val-OH (0.28 g, 0.84 mmol), Fmoc-Leu-OH (0.30 g. 0.84 mmol) and Fmoc-Trp(Boc)-OH (0.44 g, 0.84 mmol) were then coupling according to procedure I. A solution of Boc_2-Dha-OH **116** (0.24 g, 0.84 mmol), HCTU (0.34 g, 0.84 mmol) and DIPEA (0.29 mL, 1.68 mmol) in DMF (2 mL) was preactivated for 30 min at 0 °C before addition of this solution to the resin which was subsequently agitated for 10 h at rt. Resin washes with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL) were followed by drying of the resin under reduced pressure. A test cleave of a few resin beads was performed according to

procedure J. The analytical data of the resultant white solid was consistent with **133**. The Dha functionalised resin **132** was used directly in thiol-Michael addition reactions.

HRMS (*m/z* MALDI$^+$) calcd. for $C_{39}H_{64}N_{12}O_8$ = 829.4970 $(M + H)^+$. Found = 829.5096;

Analytical HPLC: retention time = 17.32 min (gradient = 5% MeCN (0.1% TFA):95% H_2O (0.1% TFA) to 45% of MeCN (0.1% TFA) over 22 min; column = C_{18} Phenomenex Gemini 5 µm, 110 Å, 250 x 4.6 mm).

_S_-trityl (2_S_,3_S_)-2-((((9_H_-Fluoren-9-yl)methoxy)carbonyl)amino)-3-methylpentanethioate (136)



To a solution of Fmoc-Ile-OH **135** (1.00 g, 2.83 mmol) in anhydrous CH_2Cl_2 (50 mL) under argon was added EDC·HCl (1.08 g, 5.66 mmol) and triphenylmethanethiol (0.94 g, 3.40 mmol) and DMAP (0.034 g, 0.28 mmol). The reaction was stirred for 16 h and then subjected to column chromatography (SiO_2, EtOAc:hex 15:85 v/v) to yield the product a white solid (1.60 g, 90%).

R_f = 0.70 (EtOAc:hex 1:9 v/v);

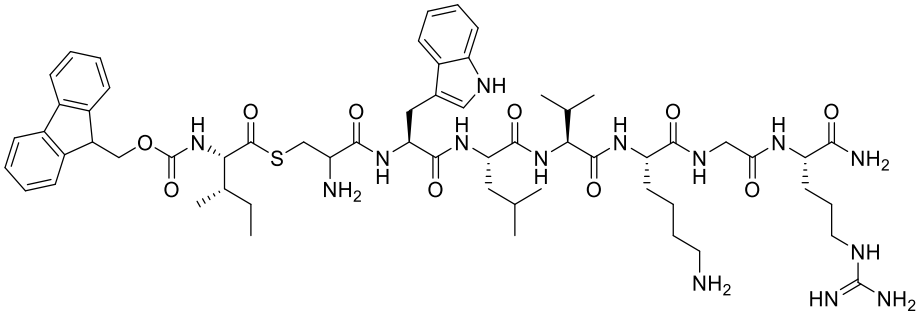
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (d, *J* = 7.4 Hz, 2H, Ar-CH), 7.63 (d, *J* = 7.4 Hz, 2H, Ar-CH), 7.43 (t, *J* = 7.4 Hz, 2H, Ar-CH), 7.34 - 7.25 (m, 17H, Ar-CH), 5.20 (d, *J* = 9.3 Hz, 1H, NH), 4.53 - 4.48 (m, 1H, Fmoc C$\underline{H_a}H_b$), 4.43 - 4.38 (m, 2H, Fmoc C$H_a\underline{H_b}$, Ile αCH), 4.28 (t, *J* = 7.0 Hz, 1H, Fmoc CH), 1.94 - 1.88 (m, 1H, Ile βCH), 1.26 - 1.18 (m, 1H, Ile C$\underline{H_a}H_b$), 1.05 - 0.85 (m, 7H, Ile C$H_a\underline{H_b}$, Ile CH_3 x 2) ppm;

^{13}C NMR (101 MHz, $CDCl_3$) δ 197.4 (SC=O), 156.0 (C=O), 143.9, 143.7, 143.5, 141.4 (Ar-qC), 129.8, 127.8, 127.2, 127.1, 125.1, 125.0, 120.0 (Ar-CH), 70.9 (qC), 67.3 (Fmoc CH_2), 65.1 (Ile αCH), 47.3 (Fmoc CH), 38.2 (Ile CH), 24.0 (Ile CH_2), 15.7 (Ile CH_3), 11.7 (Ile CH_3) ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{40}H_{41}N_2O_3S$ = 629.2803 $(M + H)^+$. Found 629.2932;

ν_{max} (thin film)/cm^{-1} 1686 (C=O), 696 (CH).

***S*-((6*S*,12*S*,15*S*,18*S*,21*S*)-21-((1*H*-Indol-3-yl)methyl)-1,24-diamino-12-(4-aminobutyl)-6- carbamoyl-1-imino-18-isobutyl-15-isopropyl-8,11,14,17,20,23-hexaoxo-2,7,10,13,16,19,22- heptaazapentacosan-25-yl) (2*S*,3*R*)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3- methylpentanethioate (137)**

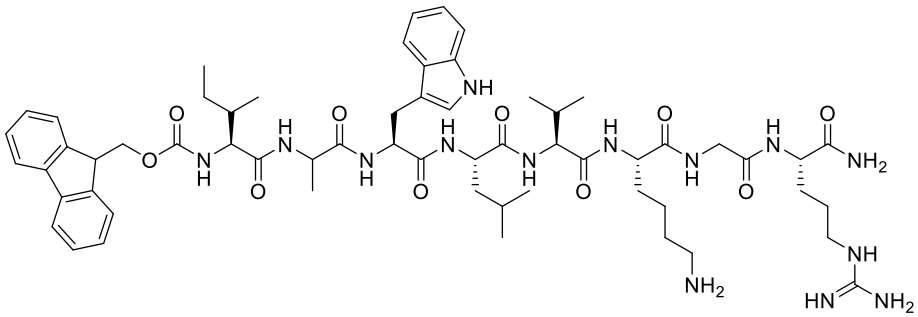


DMF (5 mL) was added to Dha functionalised resin **132** (0.093 mmol) for 20 min in a polypropylene syringe reaction vessel and DMF was then drained. Fmoc-Ile-SH **134** (0.14 mmol, 1.5 equiv) was prepared *via* procedure D and dissolved in DMF (1.5 mL) and added to the resin followed by phosphate buffer (50 mM, pH 8; 1.5 mL). The suspension was agitated at 37 °C for 16 h. The solvent was drained and the resin was washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL). The peptide was cleaved as per procedure J to yield a white solid. 48% Thioester conversion calculated from analytical HPLC.

HRMS (*m*/*z* MALDI$^+$) calcd. for $C_{60}H_{87}N_{14}O_{10}S$ = 1195.6450 $(M + H)^+$. Found = 1195.6414.

Analytical HPLC: retention time = 20.95 min (gradient of 5% MeCN (0.1% TFA):95% H_2O (0.1% TFA) to 45% of MeCN (0.1% TFA) over 22 min; column = C_{18} Phenomenex Gemini 5 µm, 110 Å, 250 x 4.6 mm).

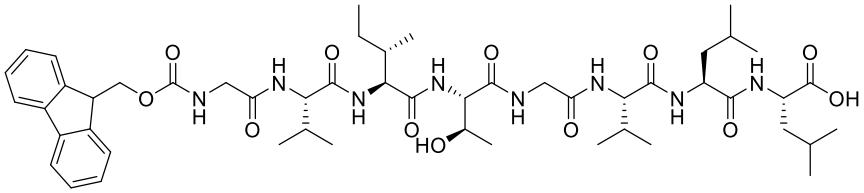
(9*H*-Fluoren-9-yl)methyl ((6*S*,12*S*,15*S*,18*S*,21*S*,27*S*,28*S*)-21-((1*H*-indol-3-yl)methyl)-1-amino-12-(4-aminobutyl)-6-carbamoyl-1-imino-18-isobutyl-15-isopropyl-24,28-dimethyl-8,11,14,17,20,23,26-heptaoxo-2,7,10,13,16,19,22,25-octaazatriacontan-27-yl)carbamate (139)



To peptide thioester **137** (0.017 mmol) was added aq. NH_4HCO_3 solution (5 mL; 0.5 M). The solution was stirred at rt for 16 h to promote acyl transfer. Direct mass analysis of this solution showed the presence of target peptide **138**; HRMS (*m*/*z* MALDI$^+$) calcd. for $C_{60}H_{87}N_{14}O_{10}S$ = 1195.6450 $(M + H)^+$, found = 1195.6475. To this solution was added DMF (1 mL) and then solution was degassed with argon. TCEP·HCl (192 mg, 0.67 mmol) and glutathione (41 mg, 0.13 mmol) were added and the pH was adjusted to pH 8 using neat Et_3N. The suspension was irradiated at 254 nm for 3 h. The solution was analysed by MS which confirmed the formation of the alanine analogue **139**.

HRMS (m/z MALDI$^+$) calcd. for $C_{60}H_{87}N_{14}O_{10}$ = 1163.6730 $(M + H)^+$. Found = 1163.6738.

(((9*H*-Fluoren-9-yl)methoxy)carbonyl)glycyl-L-valyl-L-isoleucyl-L-threonylglycyl-L-valyl-L-leucyl-L-leucine (143)



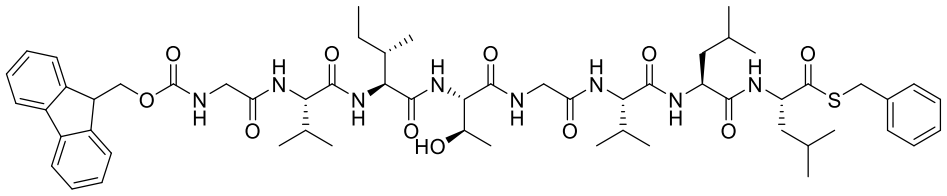
Peptide **143** was synthesised by coupling of Fmoc-Leu-OH (0.792 g, 2.24 mmol) to Wang resin (0.40 g, 0.56 mmol) as per procedure F. Fmoc-Leu-OH (0.59 g, 1.68 mmol), Fmoc-Val-OH (0.57 g, 1.68 mmol) x 2, Fmoc-Gly-OH (0.49 g, 1.68 mmol), Fmoc-Thr-OH (0.67 g, 1.68 mmol), Fmoc-Ile-OH (0.59 g, 1.68 mmol), Fmoc-Gly-OH (0.49 g, 1.68 mmol). The peptide was cleaved from the resin according to procedure J to and purified

by column chromatography (SiO_2, CH_2Cl_2:MeOH 9:1 v/v) yield a white solid (160 mg, 30%) which was used directly in the next step.

R_f = 0.26 (CH_2Cl_2:MeOH 9:1 v/v);

HRMS (*m*/*z* ESI^+) calcd. for $C_{51}H_{77}N_8O_{12}$ = 993.5583 $(M + H)^+$. Found = 993.5591.

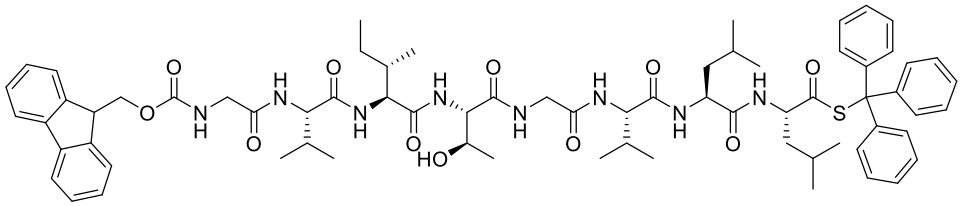
_S_-benzyl (8_S_,11_S_,14_S_,20_S_,23_S_,26_S_)-11-((_S_)-sec-butyl)-1-(9_H_-fluoren-9-yl)-14-((_R_)-1-hydroxyethyl)-23,26-diisobutyl-8,20-diisopropyl-3,6,9,12,15,18,21,24-octaoxo-2-oxa-4,7,10,13,16,19,22,25-octaazaheptacosane-27-thioate (145)



Attempted synthesis: To a solution of peptide **143** (50 mg, 0.05 mmol) in DMF (3 mL) under argon at -20 °C was added benzyl mercaptan (177 µL, 1.51 mmol), PyBOP (131 mg, 0.25 mmol) and DIPEA (43 µL, 0.25 mmol). The reaction was stirred at -20 °C for 1.5 h. DMF was removed under reduced pressure and the solid was triturated with Et_2O (2 x 10 mL) to yield a white solid. MS analysis revealed only a trace amount of the desired product which was not isolated.

HRMS (*m*/*z* ESI^+) calcd. for $C_{58}H_{82}N_8NaO_{11}S$ = 1121.5716 $(M + H)^+$. Found = 1121.5725.

_S_-trityl (8S,11_S_,14_S_,20_S_,23_S_,26_S_)-11-((_S_)-sec-butyl)-1-(9_H_-fluoren-9-yl)-14-((_R_)-1-hydroxyethyl)-23,26-diisobutyl-8,20-diisopropyl-3,6,9,12,15,18,21,24-octaoxo-2-oxa-4,7,10,13,16,19,22,25-octaazaheptacosane-27-thioate (144)

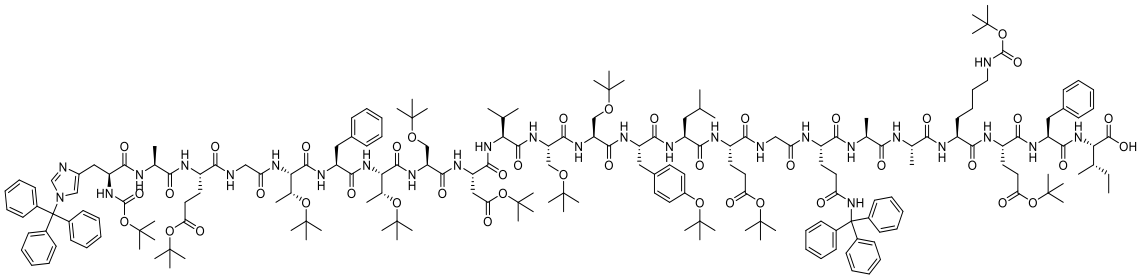


Attempted synthesis: To a solution of peptide **143** (50 mg, 0.05 mmol) in DMF (3 mL) under argon at -20 °C was added Trt-SH (418 mg, 1.51 mmol), PyBOP (131 mg, 0.25 mmol) and DIPEA (43 µL, 0.25 mmol). The reaction was stirred at -20 °C for 1.5 h. DMF was removed under reduced pressure and the solid was triturated with Et_2O (2 x 10 mL)

to yield a white solid. MS analysis revealed only a trace amount of the desired product which was not isolated.

HRMS (*m*/*z* ESI$^+$) calcd. for $C_{70}H_{90}N_8NaO_{11}S$ = 1273.6342 $(M + H)^+$. Found = 1273.6326.

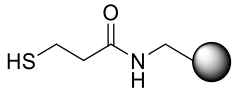
Boc-His(Trt)-Ala-Glu(O*t*Bu)-Gly-Thr(O*t*Bu)-Phe-Thr(O*t*Bu)-Ser(O*t*Bu)-Asp(O*t*Bu)-Val-Ser(O*t*Bu)-Ser(O*t*Bu)-Tyr(O*t*Bu)-Leu-Glu(O*t*Bu)-Gly-Gln(Trt)-Ala-Ala-Lys(Boc)-Glu(O*t*Bu)-Phe-Ile-COOH (146)



Loading of Fmoc-Ile-OH to 2-chlorotrityl chloride resin (186 mg, 0.26 mmol) was performed according to procedure G. The remaining 22 protected amino acid residues were coupled as per procedure I. Following the final coupling of Boc-His(Trt)-OH, the peptide was cleaved from the resin using a solution of TFA (0.05 mL), TES (0.25 mL) and CH_2Cl_2 (4.70 mL) which was added to the resin and agitated at rt for 5 min, then drained. The filtrate was concentrated under a flow of N_2, followed by precipitation of the peptide with Et_2O (10 mL) at 0 °C. The crude peptide suspension was centrifuged and the supernatant decanted. The peptide pellet was washed again with Et_2O (2 x 10 mL) at 0 °C, centrifuged and collected. The crude material was dried *in vacuo* to yield a white solid (201 mg, 20% crude yield) which was used without further purification.

HRMS (*m*/*z* MALDI$^+$) calcd. for $C_{198}H_{287}N_{27}O_{43}Na$ = 3754.0999 $(M + Na)^+$. Found = 3754.0952.

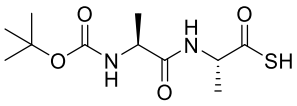
Resin intermediate (148)



A solution of 3-(tritylthio)propanoic acid **150** (139 mg, 0.40 mmol), PyBOP (208 mg, 0.40 mmol), NMM (88 µL, 0.80 mmol) and HOAt (54 mg, 0.4 mmol) in DMF (2 mL) and added to a syringe containing ChemMatrix AM resin **149** (100 mg, 0.1 mmol). The

reaction was agitated for 16 h and the syringe was then drained and washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL). The coupling of 3-(tritylthio)propanoic acid was then repeated under the same conditions. The unreacted amines on the resin were capped by the addition of a solution of Ac_2O (1 mL), DIPEA (2 mL) and DMF (7 mL) to the resin for 30 min. The resin was washed again with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL). Trityl deprotection was performed by the addition of 20% (v/v) TFA, 3% TES (v/v) in CH_2Cl_2 (5 mL) to the syringe for 3 x 5 min. The resin was washed with CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL) and used directly in the next step.

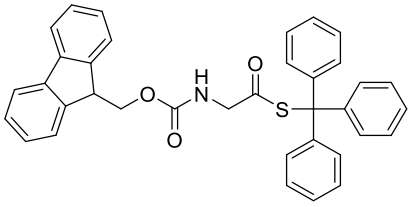
(*S*)-2-((*S*)-2-((*tert*-butoxycarbonyl)amino)propanamido)propanethioic *S*-acid (155)



Attempted synthesis: To resin intermediate **148** (0.36 mmol) was added a solution of Boc-Ala-OH (339 mg, 1.80 mmol), EDC·HCl (334 mg, 1.80 mmol) and DMAP (33 mg, 0.27 mmol) in CH_2Cl_2 (3 mL) for 1.5 h. The resin was washed with CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL). A solution of 50% (v/v) TFA, 3% (v/v) TES in CH_2Cl_2 (5 min) was added to the resin for 2 x 10 min. The resin was washed with CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL). A solution of Boc-Ala-OH (339 mg, 1.80 mmol), EDC·HCl (334 mg, 1.80 mmol) and DMAP (33 mg, 0.27 mmol) in CH_2Cl_2 (3 mL) for 45 min. The resin was washed with CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL) and dried *in vacuo*. H_2O (5 mL) was added to the resin for 20 min then drained. A solution of $(NH_4)_2S$ (aq. 21%; 3 mL) was added to phosphate buffer (pH 8; 10 mL) which was added to the resin for 45 min. The resin was drained, and the filtrate was acidified to pH 2 using aq. HCl (1 M) and degassed with argon. MS analysis revealed the presence of the desired thioacid **155**. The aq. solution was extracted with EtOAc (3 x 5 mL) and the organic layer was evaporated under reduced pressure to yield multiple products as a yellow oil.

HRMS (*m*/*z* ESI$^-$) calcd. for $C_{11}H_{19}N_2O_4S$ = 275.1071 $(M - H)^+$. Found = 275.1079.

_S_-Trityl 2-((((9_H_-fluoren-9-yl)methoxy)carbonyl)amino)ethanethioate (161)



To a solution of Fmoc-Gly-OH (3.00 g, 10.09 mmol), EDC·HCl (3.87 g, 20.18 mmol) and DMAP (123 mg, 1.01 mmol) in anhydrous CH_2Cl_2 (100 mL) under argon at rt was added Trt-SH (3.35 g, 12.10 mmol). The reaction was stirred at rt for 16 h and the desired product was then isolated by column chromatography (SiO_2, EtOAc:hex 3:7 v/v) to afford a white solid (5.04 g, 90%).

R_f = 0.49 (EtOAc:hex 3:7 v/v);

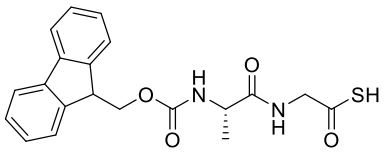
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.74 (d, *J* = 7.6 Hz, 2H, Ar-CH), 7.55 (d, *J* = 7.6 Hz, 2H, Ar-CH), 7.37 (t, *J* = 7.6 Hz, 2H, Ar-CH), 7.33-7.25 (m, 17H, Ar-CH), 5.30 (t, *J* = 5.8 Hz, 1H, NH), 4.42 (d, *J* = 6.8 Hz, 2H, Fmoc CH_2), 4.24 (t, *J* = 6.8 Hz, 1H, Fmoc CH), 4.15 (d, *J* = 5.6 Hz, 2H, Gly CH_2) ppm;

^{13}C NMR (101 MHz, $CDCl_3$) δ 194.6 (SC=O), 156.0 (C=O), 143.8 (Ar-qC), 143.4 (Ar-CH), 141.4 (Ar-qC), 129.8, 127.9 (Ar-CH), 127.7 (Ar-CH), 127.4 (Ar-CH), 127.1, 125.1, 120.0 (Ar-CH), 70.0 (qC), 67.3 (Fmoc CH_2), 50.5 (Gly CH_2), 47.1 (Fmoc-CH) ppm;

HRMS (*m/z* ESI$^+$) calcd. for $C_{36}H_{29}NNaO_3S$ = 578.1766 $(M + Na)^+$. Found 578.1760;

ν_{max} (thin film)/cm^{-1} 1696 (C=O), 1444 (Ar C-C), 1239 (CO).

(_S_)-2-(2-((((9_H_-Fluoren-9-yl)methoxy)carbonyl)amino)propanamido)ethanethioic _S_-acid (165)



Attempted synthesis: Fmoc-Gly-SH (**160**) was prepared from Fmoc-Gly-STrt (**161**) (100 mg, 0.18 mmol) as per procedure D. Anhydrous CH_2Cl_2 (5 mL) was added to a syringe containing 2-chlorotrityl chloride resin (30 mg, 0.045 mmol) which was agitated for 20 min and then drained. To the syringe was added a solution of the freshly prepared thioacid **160** in anhydrous CH_2Cl_2 (5 mL) under argon. The syringe was agitated for 16 h

and subsequently washed with CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL). A few resin beads were subjected to a test cleave using 1% (v/v) TFA, 5% (v/v) TES in CH_2Cl_2 (1 mL) for 5 min which, following evaporation of the solvents, yielded there desired product as a white residue: Fmoc-Gly-SH **160** (HRMS (*m/z* ESI^-) calcd. for $C_{17}H_{14}NO_3S$ = 312.0670 $(M - H)^-$. Found 312.0702). The remainder of the resin was treated with a solution of 6% (v/v) piperazine in DMF (4 mL) for 7 min and then washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL) and DMF (3 x 5 mL). A solution of Fmoc-Ala-OH (56 mg, 0.18 mmol), PyBOP (92 mg, 0.18 mmol) and NMM (20 µL, 0.18 mmol) in DMF (1 mL) was added to the resin which was then agitated for 45 min. The resin was drained and washed with DMF (3 x 5 mL), CH_2Cl_2 (3 x 5 mL), DMF (3 x 5 mL) and CH_2Cl_2 (3 x 5 mL) and dried under reduced pressure. CH_2Cl_2 (5 mL) was added to the resin for 20 min and then drained. The peptide was cleaved from the resin using 5% (v/v) TFA, 5% (v/v) TES in CH_2Cl_2 (1 mL) for 5 min and the solvents were evaporated under a stream of N_2 and washed with anhydrous Et_2O (2 x 2 mL) to yield a white solid. MS analysis revealed a minor trace of the desired compound.

HRMS (*m/z* ESI^-) calcd. for $C_{20}H_{19}N_2O_4S$ = 383.1144 $(M - H)^-$. Found 383.1129.

7.7 Materials and Methods for Enzyme Assays

7.7.1 LspA FRET Assay[77]

Performed by Dr. Jonathan Bailey and Dr. Samir Olatunji - The optimised LspPae assay was performed using FRET peptide **55** (30 µM) in 100 mM MES/NaOH pH 5.4, 150 mM NaCl and 0.05% (w/v) LMNG at 37 °C. A final assay volume of 50 µL was used with a final DMSO concentration of 10% (v/v). Reaction mixtures were pipetted into wells of a 384-well plate and preincubated for 10 min at 37 °C in a plate reader prior to starting the assay by addition of LspPae (100 nM). The optimised LspMrs assay was performed using FRET peptide **55** (80 µM) in 100 mM Tris/HCl pH 7.8, 150 mM NaCl and 0.09% (w/v) DDM at 37 °C. A final assay volume of 50 µL was used with a final DMSO concentration of 10% (v/v). Reaction mixtures were pipetted into wells of a 384-well plate and preincubated for 10 min at 37 °C in a plate reader prior to starting the assay by addition of LspMrs (300 nM). Reactions were monitored by fluorescence (excitation 320 nm, emission 420 nm) for 30 to 60 min at 37 °C.

7.7.2 Lit TLC Assay

Performed by Dr. Samir Olatunji and Dr. Dietmar Weichert – The activity of Lit was assayed using peptide **79** (150 µM) in 25 mM MES/NaOH pH 5.4, 200 mM NaCl, 0.02% (w/v) LMNG at 37 °C at a final reaction volume of 50 µL. The reaction was initiated by adding Lit (10 µM) and then stopped after the indicated time by flash freezing in liquid N_2. Aq. EtOH (70% v/v; 50 µL) was added to the frozen reaction mix samples which were vortexed for 10-20 s at 20 °C. $CHCl_3$ (30 µL) was added and the reaction was vortexed for 30 s to extract the substrate and product. The mixture was centrifuged for 2 min at 13,000 × *g* and 20 °C and the lower organic phase was collected. The organic phase was spotted on a silica gel 60 F_{254} TLC plate. The plate was dried *in vacuo* for 10 min. The TLC plate was developed using a mobile phase consisting of $CHCl_3$:MeOH:28% NH_4OH_{aq} (8:2:0.1 v/v/v.) and the substrate and the product were visualised on the TLC plate under UV at 254 nm. A digital image of the UV-illuminated TLC plate was taken, and the substrate and product band intensities were quantified by image analysis using ImageJ. The results were plotted using Prism 8 (GraphPad).

7.8 Methods for MD and QM/MM Simulations

Computational work was performed by Dr. Warispreet Singh and Dr. Irena Tikhonova, School of Pharmacy, QUB.

7.8.1 System Preparations for MD Simulations

The crystal structure of Lit at 1.95 Å resolution was used as an initial structure for MD simulations**.** The MD substrate, Cys-Ser-Ser-Lys, with the cysteine residue attached to a dipalmitoyl (*R*-stereoisomer) group was built using Maestro 2018-4 in the Schrodinger software suite. The structure of Lit with two monoolein molecules in the vicinity of catalytic His85 and His153 provided a good starting point for substrate docking. After removing the structured monooleins and water molecules from the Lit structure, the MD substrate was placed into the active site with the *sn*-2 chain toward the catalytic histidines. The substrate-Lit complex was then subjected to a short energy minimization using Maestro 2018-4. The product-Lit complex was generated by removing the *sn*-2 chain from the substrate DAG and linking it *via* an amide bond to the α-amino group of cysteine. The product complex was also energy optimized in Maestro 2018-4. Similarly, the Lit complex with lipopeptide monoacylated at the *sn*-1 position was created from the substrate-Lit complex. The structures of intermediate 1 (IM1) and intermediate 2 (IM2)

complexes were obtained from the QM/MM study. The force field parameters for the substrate, product, monoacylated lipopeptide, IM1, and IM2 were developed in the Antechamber program of AmberTools18 using the GAFF force field.

The protonation states of titratable residues at pH 7.0 were predicted using the PROPKA program available in Maestro 2018-4. Aspartate, glutamate, arginine, and lysine residues were assigned a negative or a positive charge, as appropriate. Non-catalytic histidine residues were modelled as neutral, with a hydrogen atom bound to either the delta or epsilon nitrogen depending on which tautomeric state optimized the local hydrogen-bonding network. Protonation states of catalytic His85 and His153 were assigned taking into consideration their local environment, QM/MM, and mutagenesis data. The side chain of His153 was in the neutral form with a proton at the delta nitrogen in the MD simulations of Lit bound to substrate and to monoacylated lipopeptide, and in the empty form. His153 was in the charged imidazolium form in MD simulations of Lit bound to the product and in the IM1 and IM2 states. Initial exploratory simulations were performed with His85 in positively charged and neutral forms. His85 in the charged form was chosen for the final MD simulations because the QM/MM analysis showed that the 8-membered heterocyclic intermediate (IM2) of the transacytlation reaction can only form in the presence of protonated His85.

Six Lit biosystems were used in this study. They include the empty form (where the two structured monoolein molecules have been removed from the active site pocket), substrate-, product- and monoacylated lipopeptide-bound forms, and IM1 and IM2 states in a hydrated lipid bilayer. These were constructed to study enzyme dynamics and the reaction mechanism. The lipid bilayer consisted of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (POPG) in a 3:1 mole ratio. The position of the Lit molecule across the lipid bilayer was established using the Orientation of Protein in Membranes (OPM) server. The CHARMM-GUI Lipid Builder was used to insert the Lit structure into the lipid bilayer consisting of 120 lipid molecules in each leaflet of the membrane. TIP3P water molecules were used to solvate the bilayer and counterions were added using Monte-Carlo simulations at a concentration of 0.15 M NaCl. The final systems comprised ~88,000 atoms with a box dimension of 95 x 95 x 100 $Å^3$.

7.8.2 MD Simulations Protocols

All MD simulations were performed using the Compute Unified Device Architecture (CUDA) version of particle-mesh Ewald molecular dynamics (PMEMD) in Amber18 on graphics processing units (GPUs). All systems were subjected to energy minimisation using the steepest descent (1,000 steps) followed by the conjugate gradient (1,000 steps) methods. Firstly, the protein and bilayer were restrained using a potential of 5 kcal mol^{-1} $Å^2$ and only solvent and ions were allowed to relax. This was followed by a full minimisation of the entire systems using steepest descent (5,000 steps) and conjugate gradient (5,000 steps) methods.

The systems were heated for 100 ps from 0 K to 100 K in the NVT ensemble using a Langevin thermostat with harmonic restraints of 5 kcal mol^{-1} $Å^{-2}$ on the non-hydrogen atoms of the lipids, protein, and lipopeptide. Initial velocities were sampled from a Boltzmann distribution. Next, the system was heated to 310 K over 300 ps in the NPT ensemble. Equilibration was performed at 310 K and 1 bar in an NPT ensemble for 100 ns in five steps: I) the membrane lipid head groups were allowed to move, while the whole system was restrained using 5 kcal mol^{-1} $Å^{-2}$ for 5 ns; II) the lipid bilayer was allowed to move, while the rest of the system was restrained using 5 kcal mol^{-1} $Å^{-2}$ for 10 ns; III) the lipid bilayer including water and ions were allowed to move, while the protein was restrained using 1 kcal mol^{-1} $Å^{-2}$ for 15 ns; and V) the whole system was allowed to equilibrate without any restrains for 70 ns. The final production step of 500 ns was run at 310 K and 1 bar in the NPT ensemble using the Langevin thermostat and Monte Carlo barostat. The simulations were performed using a time step of 2 fs. Non-bonded interactions were cut off at 10.0 Å and long-range electrostatic interactions were calculated using PMEMD. The SHAKE algorithm was used to constrain bond lengths. Five replica runs of 500 ns MD simulation each were performed for all systems. The FF14SB and Lipid14 force fields were used in all simulations. The simulations were performed on Kelvin2 and JADE clusters.

7.8.3 MD Simulations Trajectory Analysis

Trajectory reimaging, root mean square deviation, root mean square fluctuations, distances, hydrogen bonding, principal component analysis, and cluster analysis were conducted using the AmberTools18 CPPTRAJ package. Visual Molecular Dynamics (VMD) was used for visualization and movies. PyMOL and Maestro 2018-4 were used

to make images. The residue-ligand interaction energy was calculated using the 'namdenergy.tcl' script version 1.4 of NAMD 6.8.1.

7.8.4 QM/MM Adiabatic Potential Energy Calculations

Structures for the QM/MM calculations were obtained from the cluster analysis of production trajectories of Lit in complex with the lipopeptide substrate and product. The structures were first subjected to 2,000 steps of steepest descent energy minimisation followed by 2,000 steps of conjugate gradient minimisation using Amber18. The lipid bilayer was excluded in the QM/MM setup, but the water molecules that were present near the active site from the MD simulations were included. The QM/MM calculations were performed using the ChemShell suite 2019. All the quantum mechanical (QM) calculations were performed with the density functional theory (DFT) using the B3LYP functional with def2-SVP def2/J basis set. The D3 dispersion corrections with BJ damping were also included in the DFT calculations. The QM calculations was run using the ORCA 4.2.0 package and the molecular mechanics (MM) calculations were computed using the DL_POLY MD simulation package. The FF14SB force fields were used for the MM calculations. The RIJCOSX approximation with the TightSCF criteria, Grid4 and GridX4 were used in the QM calculations. The effect of the protein environment on the polarization of the QM wave function was described by the electronic embedding scheme. The QM region (74 atoms in total) consists of DAG-Cys, the α-amino group of the cysteine residue and His153 and His85 residues with an overall charge of +2. The residues and water molecules, which were within 8 Å of the acylated cysteine were allowed to move freely and the rest of the system was frozen during the geometry optimization. Hydrogen link atoms were used to saturate the dangling bonds at the QM region.

The transition sate structures associated with the reaction mechanism of Lit were obtained by running the minimum energy path calculations using the nudged elastic band method. The transition state structures were then subjected to full optimization using the dimer method. The transition state structures were validated by the presence of a single imaginary frequency at 310 K using the thermal keyword in the ChemShell suite. The distort keyword was then used to obtain the corresponding local minima structures connecting the transition states. Rate constants were calculated using the Arrhenius equation at T = 310 K using a pre-exponential factor derived from transition state theory,

$$k_{cat} = \kappa(K_b T/h)e^{-(Ea/RT)}$$

where κ is the reflection coefficient that was set equal to 1, E_a is the activation energy barrier, R is the gas constant, K_b is Boltzmann's constant, and h is Planck's constant.

References

1. P. W. Smith, K. Watkins and A. Hewlett, *Am. J. Infect. Control*, 2012, **40**, 35-42.
2. *Global Action Plan On Antimicrobial Resistance*, Geneva, World Health Organization, 2015.
3. K. Smith, *Front. Immunol.*, 2012, **3**, 369, 1-12
4. S. M. Blevins and M. S. Bronze, *Int. J. Infect. Dis.*, 2010, **14**, 744-751.
5. R. Emmerich and O. Löw, *Zeitschrift für Hygiene und Infektionskrankheiten*, 1899, **31**, 1-65.
6. S. Fetzner and Steffen L. Drees, *Chem. Biol.*, 2013, **20**, 1438-1440.
7. K. J. Williams, *J. R. Soc. Med.*, 2009, **102**, 343-348.
8. J. J. Abraham, *Br J Vener Dis*, 1948, **24**, 153-161.
9. G. Domagk, *Dtsch med Wochenschr*, 1935, **61**, 250-253.
10. B. Ribeiro da Cunha, L. P. Fonseca and C. R. C. Calado, *Antibiotics*, 2019, **8**, 45.
11. R. I. Aminov, *Front. Microbiol.*, 2010, **1**, 134-134.
12. A. Fleming, *Br. J. Exp. Pathol.*, 1929, **10**, 226-236.
13. A. Perciaccante, A. Coralli, D. Lippi, O. Appenzeller and R. Bianucci, *Intern Emerg Med*, 2019, **14**, 1363-1364.
14. *Chemistry of Penicillin*, Princeton University Press, New Jersey, 1949.
15. E. Chain, H. W. Florey, A. D. Gardner, N. G. Heatley, M. A. Jennings, J. Orr-Ewing, A. G. Sanders and L. F. Peltier, *Clin. Orthop. Relat. Res.*, 2005, **439**, 23-26.
16. A. Schatz, E. Bugle and S. A. Waksman, *Proc. Soc. Exp. Biol. Med.*, 1944, **55**, 66-69.
17. K. Lewis, *Nature*, 2012, **485**, 439-440.
18. K. Lewis, *Cell*, 2020, **181**, 29-45.
19. E. P. Abraham and E. Chain, *Nature*, 1940, **146**, 837-837.
20. *Antimicrobial Resistance: Global Report On Surveillance*, Geneva, World Health Organization, 2014.
21. J. M. Munita and C. A. Arias, *Microbiol. Spectr.*, 2016, **4**, 1-24.
22. C. Kirchhelle, *Palgrave Commun.*, 2018, **4**, 96.
23. *2019 Antibacterial Agents In Clinical Development: An Analysis Of The Antibacterial Clinical Development Pipeline*, Geneva, World Health Organization, 2019.
24. *Global Priority List Of Antibiotic-Resistant Bacteria To Guide Research, Discovery, And Development Of New Antibiotics*, Geneva, World Health Organization, 2017.
25. D. J. Payne, M. N. Gwynn, D. J. Holmes and D. L. Pompliano, *Nat. Rev. Drug Discov.*, 2007, **6**, 29-40.
26. H. I. Zgurskaya, C. A. Löpez and S. Gnanakaran, *ACS Infect. Dis.*, 2015, **1**, 512-522.
27. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Del. Rev.*, 2001, **46**, 3-26.

28. E. N. Parker, B. S. Drown, E. J. Geddes, H. Y. Lee, N. Ismail, G. W. Lau and P. J. Hergenrother, *Nat. Microbiol.*, 2020, **5**, 67-75.
29. M. F. Richter, B. S. Drown, A. P. Riley, A. Garcia, T. Shirai, R. L. Svec and P. J. Hergenrother, *Nature*, 2017, **545**, 299-304.
30. R. O'Shea and H. E. Moser, *J. Med. Chem.*, 2008, **51**, 2871-2878.
31. F. Yoshimura and H. Nikaido, *J. Bacteriol.*, 1982, **152**, 636.
32. J. C. Venter *et. al.*, *Science*, 2001, **291**, 1304-1351.
33. S. Narita and H. Tokuda, *Biochim Biophys Acta Mol Cell Biol Lipids*, 2017, **1862**, 1414-1423.
34. R. J. Havel, *Adv. Exp. Med. Biol.*, 1975, **63**, 37-59.
35. A. Kovacs-Simon, R. W. Titball and S. L. Michell, *Infect. Immun.*, 2011, **79**, 548-561.
36. M. J. Nadolski and M. E. Linder, *FEBS J.*, 2007, **274**, 5202-5210.
37. K. Hantke and V. Braun, *Eur. J. Biochem.*, 1973, **34**, 284-296.
38. K. M. Armbruster, G. Komazin and T. C. Meredith, *J. Bacteriol.*, 2019, **201**, e00195-00119.
39. H. Nakayama, K. Kurokawa and B. L. Lee, *The FEBS journal*, 2012, **279**, 4247-4268.
40. K. Kurokawa, K. H. Ryu, R. Ichikawa, A. Masuda, M. S. Kim, H. Lee, J. H. Chae, T. Shimizu, T. Saitoh, K. Kuwano, S. Akira, N. Dohmae, H. Nakayama and B. L. Lee, *J. Biol. Chem.*, 2012, **287**, 13170-13181.
41. A. M. Glauert and M. J. Thornley, *Annu. Rev. Microbiol.*, 1969, **23**, 159-198.
42. Y. Kamio and H. Nikaido, *Biochemistry*, 1976, **15**, 2561-2570.
43. A. T. Asmar and J.-F. Collet, *FEMS Microbiol. Lett.*, 2018, **365**, fny199
44. N. Buddelmeijer, *FEMS Microbiol. Rev.*, 2015, **39**, 246-261.
45. H. C. Wu, J. S. Lai, S. Hayashi and C. Z. Giam, *Biophys J*, 1982, **37**, 307-315.
46. V. Braun and H. C. Wu, in *New Comprehensive Biochemistry*, eds. J. M. Ghuysen and R. Hakenbeck, Elsevier, 1994, vol. 27, pp. 319-341.
47. H. Shruthi, P. Anand, V. Murugan and K. Sankaran, *Mol. Biosyst.*, 2010, **6**, 999-1007.
48. J. M. Crane and L. L. Randall, *EcoSal Plus*, 2017, **7**, 2.
49. S. Inouye, S. Wang, J. Sekizawa, S. Halegoua and M. Inouye, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, **74**, 1004-1008.
50. M. M. Babu, M. L. Priya, A. T. Selvan, M. Madera, J. Gough, L. Aravind and K. Sankaran, *J. Bacteriol.*, 2006, **188**, 2761-2773.
51. K. Sankaran and H. C. Wu, *J. Biol. Chem.*, 1994, **269**, 19701-19706.
52. M. Tokunaga, H. Tokunaga and H. C. Wu, *Proc. Natl. Acad. Sci. U. S. A.*, 1982, **79**, 2255-2259.
53. K. Yamaguchi, F. Yu and M. Inouye, *Cell*, 1988, **53**, 423-432.
54. A. Tschumi, C. Nai, Y. Auchli, P. Hunziker, P. Gehrig, P. Keller, T. Grau and P. Sander, *J. Biol. Chem.*, 2009, **284**, 27146-27156.
55. D. Vidal-Ingigliardi, S. Lewenza and N. Buddelmeijer, *J. Bacteriol. Res.*, 2007, **189**, 4456.

56. D. A. Widdick, M. G. Hicks, B. J. Thompson, A. Tschumi, G. Chandra, I. C. Sutcliffe, J. K. Brülle, P. Sander, T. Palmer and M. I. Hutchings, *Mol. Microbiol.*, 2011, **80**, 1395-1412.
57. K. Kurokawa, H. Lee, K.-B. Roh, M. Asanuma, Y. S. Kim, H. Nakayama, A. Shiratsuchi, Y. Choi, O. Takeuchi, H. J. Kang, N. Dohmae, Y. Nakanishi, S. Akira, K. Sekimizu and B. L. Lee, *J. Biol. Chem.*, 2009, **284**, 8406-8411.
58. M. V. Serebryakova, I. A. Demina, M. A. Galyamina, I. G. Kondratov, V. G. Ladygina and V. M. Govorun, *J. Biol. Chem.*, 2011, **286**, 22769-22776.
59. K. M. Armbruster and T. C. Meredith, *J. Bacteriol.*, 2017, **199**, e00099-00017.
60. J. H. Gardiner, G. Komazin, M. Matsuo, K. Cole, F. Götz and T. C. Meredith, *mBio*, 2020, **11**, e01619-01620.
61. E. Mifsud, A. Tan and D. Jackson, *Front. Immunol.*, 2014, **5**.
62. H. D. Brightbill, D. H. Libraty, S. R. Krutzik, R. B. Yang, J. T. Belisle, J. R. Bleharski, M. Maitland, M. V. Norgard, S. E. Plevy, S. T. Smale, P. J. Brennan, B. R. Bloom, P. J. Godowski and R. L. Modlin, *Science*, 1999, **285**, 732-736.
63. A. P. Basto and A. Leitão, *J. Immunol.*, 2014, 619410.
64. K. Takeda, O. Takeuchi and S. Akira, *J. Endotoxin Res.*, 2002, **8**, 459-463.
65. R. B. Johnson, S. Köhl, K. Wiesmüller, G. Jung and W. G. Bessler, *Immunobiology*, 1983, **165**, 27-35.
66. K. Garland, H. Pantua, M.-G. Braun, D. J. Burdick, G. M. Castanedo, Y.-C. Chen, Y.-X. Cheng, J. Cheong, B. Daniels, G. Deshmukh, Y. Fu, P. Gibbons, S. L. Gloor, R. Hua, S. Labadie, X. Liu, R. Pastor, C. Stivala, M. Xu, Y. Xu, H. Zheng, S. B. Kapadia and E. J. Hanan, *Biorg. Med. Chem. Lett.*, 2020, **30**, 127419.
67. M. Morr, O. Takeuchi, S. Akira, M. M. Simon and P. F. Muhlradt, *Eur. J. Immunol.*, 2002, **32**, 3337-3347.
68. J. Y. Kang, X. Nan, M. S. Jin, S.-J. Youn, Y. H. Ryu, S. Mah, S. H. Han, H. Lee, S.-G. Paik and J.-O. Lee, *Immunity*, 2009, **31**, 873-884.
69. M. S. Jin, S. E. Kim, J. Y. Heo, M. E. Lee, H. M. Kim, S.-G. Paik, H. Lee and J.-O. Lee, *Cell*, 2007, **130**, 1071-1082.
70. M.-T. Nguyen, J. Uebele, N. Kumari, H. Nakayama, L. Peter, O. Ticha, A.-K. Woischnig, M. Schmaler, N. Khanna, N. Dohmae, B. L. Lee, I. Bekeredjian-Ding and F. Götz, *Nat. Commun.*, 2017, **8**, 2246.
71. B. M. Staehlin, J. G. Gibbons, A. Rokas, T. V. O'Halloran and J. C. Slot, *Genome Biol. Evol.*, 2016, **8**, 811-826.
72. S. D. Rogers, M. R. Bhave, J. F. Mercer, J. Camakaris and B. T. Lee, *J. Bacteriol.*, 1991, **173**, 6742.
73. K. Gan, S. D. Gupta, K. Sankaran, M. B. Schmid and H. C. Wu, *J. Biol. Chem.*, 1993, **268**, 16544-16550.
74. G. Mao, Y. Zhao, X. Kang, Z. Li, Y. Zhang, X. Wang, F. Sun, K. Sankaran and X. C. Zhang, *Nat. Commun.*, 2016, **7**, 10198.
75. C. Robichon, D. Vidal-Ingigliardi and A. P. Pugsley, *J. Biol. Chem.*, 2005, **280**, 974-983.
76. S. D. Gupta, K. Gan, M. B. Schmid and H. C. Wu, *J. Biol. Chem.*, 1993, **268**, 16551-16556.

77. S. Olatunji, X. Yu, J. Bailey, C.-Y. Huang, M. Zapotoczna, K. Bowen, M. Remškar, R. Müller, E. M. Scanlan, J. A. Geoghegan, V. Olieric and M. Caffrey, *Nat. Commun.*, 2020, **11**, 140.
78. M. T. Nguyen and F. Götz, *Microbiol. Mol. Biol. Rev.*, 2016, **80**, 891.
79. K. M. Lehman and M. Grabowicz, *Antibiotics*, 2019, **8**, 163.
80. J. Bubeck Wardenburg, W. A. Williams and D. Missiakas, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 13831-13836.
81. H. Stoll, J. Dengjel, C. Nerz and F. Götz, *Infect. Immun.*, 2005, **73**, 2411-2423.
82. B. A. Bray, I. C. Sutcliffe and D. J. Harrington, *Microbiology*, 2009, **155**, 1451-1458.
83. A. Hamilton, C. Robinson, I. C. Sutcliffe, J. Slater, D. J. Maskell, N. Davis-Poynter, K. Smith, A. Waller and D. J. Harrington, *Infect. Immun.*, 2006, **74**, 6907.
84. S. Okugawa, M. Moayeri, A. P. Pomerantsev, I. Sastalla, D. Crown, P. K. Gupta and S. H. Leppla, *Mol. Microbiol.*, 2012, **83**, 96-109.
85. C. M. Petit, J. R. Brown, K. Ingraham, A. P. Bryant and D. J. Holmes, *FEMS Microbiol. Lett*, 2001, **200**, 229-233.
86. L. Vogeley, T. El Arnaout, J. Bailey, P. J. Stansfeld, C. Boland and M. Caffrey, *Science*, 2016, **351**, 876-880.
87. E. Rosenberg, B. Vaks and A. Zuckerberg, *Antimicrob. Agents Chemother.*, 1973, **4**, 507-513.
88. M. Inukai, M. Nakajima, M. Osawa, T. Haneishi and M. Arai, *J. Antibiot. Res.*, 1978, **31**, 421-425.
89. M. Inukai, M. Takeuchi, K. Shimizu and M. Arai, *J. Antibiot. Res.*, 1978, **31**, 1203-1205.
90. Y. Xiao, K. Gerth, R. Muller and D. Wall, *Antimicrob. Agents Chemother.*, 2012, **56**, 2014-2021.
91. V. Simunovic, J. Zapp, S. Rachid, D. Krug, P. Meiser and R. Müller, *Chembiochem*, 2006, **7**, 1206-1220.
92. M. Inukai, R. Enokita, A. Torikata, M. Nakahara, S. Iwado and M. Arai, *J. Antibiot. Res*, 1978, **31**, 410-420.
93. S. Kitamura, A. Owensby, D. Wall and D. W. Wolan, *Cell Chem Biol*, 2018, **3**, 301-308
94. E. Rosenberg, J. M. Porter, P. N. Nathan, A. Manor and M. Varon, *Bio/Technology*, 1984, **2**, 796-799.
95. T. Kiho, M. Nakayama, K. Yasuda, S. Miyakoshi, M. Inukai and H. Kogen, *Bioorg. Med. Chem.*, 2004, **12**, 337-361.
96. F. Graef, S. Gordon and C.-M. Lehr, in *How to Overcome the Antibiotic Crisis : Facts, Challenges, Technologies and Future Perspectives*, eds. M. Stadler and P. Dersch, Springer International Publishing, Cham, 2016, pp. 475-496.
97. A. M. Pappenheimer, in *Bacterial Vaccines*, ed. R. Germanier, Academic Press, Boston, 1984, pp. 1-36.
98. B. Bizzini, in *Bacterial Vaccines*, ed. R. Germanier, Academic Press, Boston, 1984, pp. 37-68.

99. R. Germanier, in *Bacterial Vaccines*, ed. R. Germanier, Academic Press, Boston, 1984, pp. 137-165.
100. F. M. Collins, in *Bacterial Vaccines*, ed. R. Germanier, Academic Press, Boston, 1984, pp. 373-418.
101. P. W. Kantoff, C. S. Higano, N. D. Shore, E. R. Berger, E. J. Small, D. F. Penson, C. H. Redfern, A. C. Ferrari, R. Dreicer, R. B. Sims, Y. Xu, M. W. Frohlich and P. F. Schellhammer, *N. Engl. J. Med.*, 2010, **363**, 411-422.
102. I. Mellman, G. Coukos and G. Dranoff, *Nature*, 2011, **480**, 480-489.
103. D. D. Chaplin, *J. Allergy Clin. Immunol.*, 2010, **125**, S3-S23.
104. A. Di Pasquale, S. Preiss, F. Tavares Da Silva and N. Garçon, *Vaccine*, 2015, **3**, 320-343.
105. S. Calabro, M. Tortoli, B. C. Baudner, A. Pacitto, M. Cortese, D. T. O'Hagan, E. De Gregorio, A. Seubert and A. Wack, *Vaccine*, 2011, **29**, 1812-1823.
106. O. Takeuchi and S. Akira, *Cell*, 2010, **140**, 805-820.
107. L. D. Fletcher, L. Bernfield, V. Barniak, J. E. Farley, A. Howell, M. Knauf, P. Ooi, R. P. Smith, P. Weise, M. Wetherell, X. Xie, R. Zagursky, Y. Zhang and G. W. Zlotnick, *Infect. Immun.*, 2004, **72**, 2088-2100.
108. E. Fikrig, S. W. Barthold, F. S. Kantor and R. A. Flavell, *Science*, 1990, **250**, 553-556.
109. A. R. Wijayadikusumah, L. C. Sullivan, D. C. Jackson and B. Y. Chua, *Amino acids*, 2017, **49**, 1691-1704.
110. U. Buwitt-Beckmann, H. Heine, K. H. Wiesmüller, G. Jung, R. Brock and A. J. Ulmer, *FEBS J.*, 2005, **272**, 6354-6364.
111. G. Agnihotri, B. M. Crall, T. C. Lewis, T. P. Day, R. Balakrishna, H. J. Warshakoon, S. S. Malladi and S. A. David, *J. Med. Chem.*, 2011, **54**, 8148-8160.
112. T. Seyberth, S. Voss, R. Brock, K. H. Wiesmüller and G. Jung, *J. Med. Chem*, 2006, **49**, 1754-1765.
113. A. B. Schromm, J. Howe, A. J. Ulmer, K. H. Wiesmuller, T. Seyberth, G. Jung, M. Rossle, M. H. Koch, T. Gutsmann and K. Brandenburg, *J. Biol. Chem.*, 2007, **282**, 11030-11037.
114. O. Takeuchi, A. Kaufmann, K. Grote, T. Kawai, K. Hoshino, M. Morr, P. F. Mühlradt and S. Akira, *J. Immunol. Res.* 2000, **164**, 554-557.
115. W. Wu, R. Li, S. S. Malladi, H. J. Warshakoon, M. R. Kimbrell, M. W. Amolins, R. Ukani, A. Datta and S. A. David, *J. Med. Chem.*, 2010, **53**, 3198-3213.
116. J. Metzger, G. Jung, W. G. Bessler, P. Hoffmann, M. Strecker, A. Lieberknecht and U. Schmidt, *J. Med. Chem.*, 1991, **34**, 1969-1974.
117. B. L. Lu, G. M. Williams, D. J. Verdon, P. R. Dunbar and M. A. Brimble, *J. Med. Chem.*, 2020, **63**, 2282-2291.
118. J. W. Metzger, K. H. Wiesmuller and G. Jung, *Int. J. Pept. Protein Res.*, 1991, **38**, 545-554.
119. C.-H. Leng, S.-J. Liu, H.-W. Chen and P. Chong, *Expert Rev. Vaccines*, 2015, **14**, 1623-1632.
120. F. Shen, Y.-C. Huang, S. Tang, Y.-X. Chen and L. Liu, *Isr. J. Chem.*, 2011, **51**, 940-952.

121. D. Ghaderi, R. E. Taylor, V. Padler-Karavani, S. Diaz and A. Varki, *Nat. Biotechnol.*, 2010, **28**, 863-867.
122. J. D. Valderrama-Rincon, A. C. Fisher, J. H. Merritt, Y.-Y. Fan, C. A. Reading, K. Chhiba, C. Heiss, P. Azadi, M. Aebi and M. P. DeLisa, *Nat. Chem. Biol.*, 2012, **8**, 434-436.
123. R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149-2154.
124. R. Behrendt, P. White and J. Offer, *J. Pept. Sci.*, 2016, **22**, 4-27.
125. B. Henkel, L. Zhang and E. Bayer, *Liebigs Annalen*, 1997, **10**, 2161-2168.
126. N. Thieriet, F. Guibé and F. Albericio, *Org. Lett.*, 2000, **2**, 1815-1817.
127. S. Sakakibara, Y. Shimonishi, Y. Kishida, M. Okada and H. Sugihara, *Bull. Chem. Soc. Jpn.*, 1967, **40**, 2164-2167.
128. L. A. Carpino and G. Y. Han, *J. Am. Chem. Soc.*, 1970, **92**, 5748-5749.
129. R. H. Angeletti, L. F. Bonewald and G. B. Fields, *Methods Enzymol.*, 1997, **289**, 697-717.
130. C. D. Chang and J. Meienhofer, *Int. J. Pept. Protein Res.*, 1978, **11**, 246-249.
131. A. Dryland and R. C. Sheppard, *J. Chem. Soc., Perkin Trans.*, 1986, **1**, 125-137.
132. S.-S. Wang, *J. Am. Chem. Soc.*, 1973, **95**, 1328-1333.
133. H. Rink, *Tetrahedron Lett.*, 1987, **28**, 3787-3790.
134. K. Barlos, D. Gatos, J. Kallitsis, G. Papaphotiu, P. Sotiriu, Y. Wenqing and W. Schäfer, *Tetrahedron Lett.*, 1989, **30**, 3943-3946.
135. J. Ryttersgaard, B. Due Larsen, A. Holm, D. H. Christensen and O. Faurskov Nielsen, *Spectrochimica Acta Part A*, 1997, **53**, 91-98.
136. B. D. Larsen, D. H. Christensen, A. Holm, R. Zillmer and O. F. Nielsen, *J. Am. Chem. Soc.*, 1993, **115**, 6247-6253.
137. S. Wang and Y. Ishii, *J. Am. Chem. Soc.*, 2012, **134**, 2848-2851.
138. L. K. Mueller, A. C. Baumruck, H. Zhdanova and A. A. Tietze, *Front. Bioeng. Biotechnol*, 2020, **8**, 162, 1-17.
139. F. Weygand, W. Steglich, J. Bjarnason, R. Akhtar and N. M. Khan, *Tetrahedron Lett.*, 1966, **7**, 3483-3487.
140. C. Hyde, T. Johnson, D. Owen, M. Quibell and R. C. Sheppard, *Int. J. Pept. Protein Res.*, 1994, **43**, 431-440.
141. T. Wöhr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. Sun and M. Mutter, *J. Am. Chem. Soc.*, 1996, **118**, 9218-9227.
142. T. Wöhr and M. Mutter, *Tetrahedron Lett.*, 1995, **36**, 3847-3848.
143. M. Quibell, W. G. Turnell and T. Johnson, *Tetrahedron Lett.*, 1994, **35**, 2237-2238.
144. S. L. Pedersen, A. P. Tofteng, L. Malik and K. J. Jensen, *Chem. Soc. Rev.*, 2012, **41**, 1826-1844.
145. B. Bacsa, K. Horváti, S. Bõsze, F. Andreae and C. O. Kappe, *J. Org. Chem.*, 2008, **73**, 7532-7542.
146. C. Loffredo, N. A. Assunção, J. Gerhardt and M. T. M. Miranda, *J. Pept. Sci.*, 2009, **15**, 808-817.
147. S. A. Palasek, Z. J. Cox and J. M. Collins, *J. Pept. Sci.*, 2007, **13**, 143-148.

148. C. Echalier, S. Al-Halifa, A. Kreiter, C. Enjalbal, P. Sanchez, L. Ronga, K. Puget, P. Verdié, M. Amblard, J. Martinez and G. Subra, *Amino Acids*, 2013, **45**, 1395-1403.
149. A. K. Tickler and J. D. Wade, *Curr Protoc Protein Sci*, 2007, **50**, 18.18.11-18.18.16.
150. S. A. Kates, B. F. McGuinness, C. Blackburn, G. W. Griffin, N. A. Solé, G. Barany and F. Albericio, *J. Pept. Sci.*, 1998, **47**, 365-380.
151. F. García-Martín, M. Quintanar-Audelo, Y. García-Ramos, L. J. Cruz, C. Gravel, R. Furic, S. Côté, J. Tulla-Puche and F. Albericio, *J. Comb. Chem.*, 2006, **8**, 213-220.
152. J. K. Magtaan, M. Devocelle and F. Kelleher, *J. Pept. Sci.*, 2020, **26**, e3250.
153. T. J. Lukas, M. B. Prystowsky and B. W. Erickson, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 2791.
154. E. Bayer, G. Jung, I. Halász and I. Sebestian, *Tetrahedron Lett.*, 1970, **11**, 4503-4505.
155. 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, X. Ye, A. J. Mijalis, M. D. Simon, A. Loas, S. Liu, C. Jessen, T. E. Nielsen and B. L. Pentelute, *Science*, 2020, **368**, 980.
156. P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776.
157. V. Agouridas, O. El Mahdi, V. Diemer, M. Cargoët, J.-C. M. Monbaliu and O. Melnyk, *Chem. Rev.*, 2019, **119**, 7328-7443.
158. S. Ingale, T. Buskas and G.-J. Boons, *Org. Lett.*, 2006, **8**, 5785-5788.
159. E. C. B. Johnson and S. B. H. Kent, *J. Am. Chem. Soc.*, 2006, **128**, 6640-6646.
160. L. Z. Yan and P. E. Dawson, *J. Am. Chem. Soc.*, 2001, **123**, 526-533.
161. C. Kan and S. J. Danishefsky, *Tetrahcdron*, 2009, **65**, 9047-9065.
162. H. Rohde and O. Seitz, *Peptide Science*, 2010, **94**, 551-559.
163. B. Kwon, D. Tietze, P. B. White, S. Y. Liao and M. Hong, *Protein Sci.*, 2015, **24**, 1087-1099.
164. M. Dittmann, J. Sauermann, R. Seidel, W. Zimmermann and M. Engelhard, *J. Pept. Sci.*, 2010, **16**, 558-562.
165. A. Otaka, S. Ueda, K. Tomita, Y. Yano, H. Tamamura, K. Matsuzaki and N. Fujii, *Chem. Commun.*, 2004, 1722-1723.
166. C. L. Hunter and G. G. Kochendoerfer, *Bioconj. Chem.*, 2004, **15**, 437-440.
167. E. Bianchi, R. Ingenito, R. J. Simon and A. Pessi, *J. Am. Chem. Soc.*, 1999, **121**, 7698-7699.
168. S. Jin, R. J. Brea, A. K. Rudd, S. P. Moon, M. R. Pratt and N. K. Devaraj, *Nat. Commun.*, 2020, **11**, 2793.
169. W. Trowitzsch, V. Wray, K. Gerth and G. Höfle, *Chem. Commun.*, 1982, 1340-1342.
170. D. Murray and M. Wigglesworth, in *High Throughput Screening Methods: Evolution and Refinement*, The Royal Society of Chemistry, London, 2017, pp. 1-15.

171. G. A. Jones and D. S. Bradshaw, *Front. Phys.*, 2019, **7**, 100.
172. T. Förster, *Annalen der Physik*, 1948, **437**, 55-75.
173. G. Zhang, in *Assay Guidance Manual*, eds. G. S. Sittampalam, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. P. Austin, J. Baell, B. Bejcek, J. M. M. Caaveiro, T. D. Y. Chung, N. P. Coussens, J. L. Dahlin, V. Devanaryan, T. L. Foley, M. Glicksman, M. D. Hall, J. V. Haas, S. R. J. Hoare, J. Inglese, P. W. Iversen, S. D. Kahl, S. C. Kales, S. Kirshner, M. Lal-Nag, Z. Li, J. McGee, O. McManus, T. Riss, P. Saradjian, O. J. Trask, Jr., J. R. Weidner, M. J. Wildey, M. Xia and X. Xu, Bethesda (MD), 2004.
174. M. Singh, M. Watkinson, E. M. Scanlan and G. J. Miller, *RSC Chem. Biol.*, 2020, **1**, 352-368.
175. J. C. Davies, *Paediatr Respir Rev*, 2002, **3**, 128-134.
176. P. Farrell, S. Joffe, L. Foley, G. J. Canny, P. Mayne and M. Rosenberg, *Ir Med J*, 2007, **100**, 557-560.
177. E. B. Breidenstein, C. de la Fuente-Nunez and R. E. Hancock, *Trends Microbiol.*, 2011, **19**, 419-426.
178. V. D. Rosenthal, H. Bijie, D. G. Maki, Y. Mehta, A. Apisarnthanarak, E. A. Medeiros, H. Leblebicioglu, D. Fisher, C. Álvarez-Moreno, I. A. Khader, M. Del Rocío González Martínez, L. E. Cuellar, J. A. Navoa-Ng, R. Abouqal, H. Guanche Garcell, Z. Mitrev, M. C. Pirez García, A. Hamdi, L. Dueñas, E. Cancel, V. Gurskis, O. Rasslan, A. Ahmed, S. S. Kanj, O. C. Ugalde, T. Mapp, L. Raka, C. Yuet Meng, T. A. Thu le, S. Ghazal, A. Gikas, L. P. Narváez, N. Mejía, N. Hadjieva, M. O. Gamar Elanbya, M. E. Guzmán Siritt and K. Jayatilleke, *Am J Infect Control*, 2012, **40**, 396-407.
179. R. M. Klevens, M. A. Morrison, J. Nadle, S. Petit, K. Gershman, S. Ray, L. H. Harrison, R. Lynfield, G. Dumyati, J. M. Townes, A. S. Craig, E. R. Zell, G. E. Fosheim, L. K. McDougal, R. B. Carey and S. K. Fridkin, *JAMA*, 2007, **298**, 1763-1771.
180. A. Das, S. Mahale, V. Prashar, S. Bihani, J. L. Ferrer and M. V. Hosur, *J. Am. Chem. Soc.*, 2010, **132**, 6366-6373.
181. M. Wiktor, D. Weichert, N. Howe, C. Y. Huang, V. Olieric, C. Boland, J. Bailey, L. Vogeley, P. J. Stansfeld, N. Buddelmeijer, M. Wang and M. Caffrey, *Nat. Commun.*, 2017, **8**, 15952.
182. M. Duchene, C. Barron, A. Schweizer, B. U. von Specht and H. Domdey, *J. Bacteriol.*, 1989, **171**, 4130-4137.
183. J. F. Riordan, M. Sokolovsky and B. L. Vallee, *Biochemistry*, 1967, **6**, 358-361.
184. V. De Filippis, R. Frasson and A. Fontana, *Protein Sci.*, 2006, **15**, 976-986.
185. J. Ø. Duus, M. Meldal and J. R. Winkler, *J. Phys. Chem. B*, 1998, **102**, 6413-6418.
186. Y. Liu, W. Zhang, Q. He, F. Yu, T. Song, T. Liu, Z. Zhang, J. Zhou, P. G. Wang and W. Zhao, *Chem. Commun.*, 2016, **52**, 10886-10889.
187. T. Maegawa, Y. Fujiwara, T. Ikawa, H. Hisashi, Y. Monguchi and H. Sajiki, *Amino acids*, 2009, **36**, 493-499.
188. B. Neises and W. Steglich, *Angew. Chem.*, 1978, **17**, 522-524.

189. A. Kumar, A. Sharma, E. Haimov, A. El-Faham, B. G. de la Torre and F. Albericio, *Org. Process. Res. Dev.*, 2017, **21**, 1533-1541.
190. L. S. Lin, T. Lanza, S. E. de Laszlo, Q. Truong, T. Kamenecka and W. K. Hagmann, *Tetrahedron Lett.*, 2000, **41**, 7013-7016.
191. R. Santini, M. C. Griffith and M. Qi, *Tetrahedron Lett.*, 1998, **39**, 8951-8954.
192. C. A. Chantell, M. A. Onaiyekan and M. Menakuru, *J. Pept. Sci.*, 2012, **18**, 88-91.
193. J. Coste, D. Le-Nguyen and B. Castro, *Tetrahedron Lett.*, 1990, **31**, 205-208.
194. G. B. Fields, in *Peptide Synthesis Protocols*, eds. M. W. Pennington and B. M. Dunn, Humana Press, Totowa, NJ, 1995, pp. 17-27.
195. V. Krchnak, J. Vagner and M. Lebl, *Int. J. Pept. Protein Res.*, 1988, **32**, 415-416.
196. E. Kaiser, R. L. Colescott, C. D. Bossinger and P. I. Cook, *Anal. Biochem*, 1970, **34**, 595-598.
197. K. M. Bromfield, J. Cianci and P. J. Duggan, *Molecules*, 2004, **9**, 427-439.
198. S. Singh, I. Khaytin, S. Botsko, G. Crossley, D. K. Plank, Y. Lefievre, F. Giuffrida and M. W. Pennington, *Lett. Pept. Sci.*, 2002, **9**, 221-225.
199. T. Kaiser, G. J. Nicholson, H. J. Kohlbau and W. Voelter, *Tetrahedron Lett.*, 1996, **37**, 1187-1190.
200. S. Kitamura and D. W. Wolan, *FEBS Lett.*, 2018, **592**, 2289-2296.
201. M. G. Acker and D. S. Auld, *Perspect. Sci.*, 2014, **1**, 56-73.
202. J. W. Williams and J. F. Morrison, in *Methods Enzymol.*, Academic Press, 1979, vol. 63, pp. 437-467.
203. E. N. Hudson and G. Weber, *Biochemistry*, 1973, **12**, 4154-4161.
204. W. H. Melhuish, *J. Phys. Chem.*, 1961, **65**, 229-235.
205. G. T. Wang, E. Matayoshi, H. Jan Huffaker and G. A. Krafft, *Tetrahedron Lett.*, 1990, **31**, 6493-6496.
206. B. Valeur and M. B.-Santos, in *Effects of Intermolecular Photophysical Processes on Fluorescence Emission*, 2nd ed., Wiley, Germany, 2012, pp. 141-179
207. R. van Slageren, G. den Boef and W. E. van der Linden, *Talanta*, 1973, **20**, 501-512.
208. S. Kumar Panigrahi and A. Kumar Mishra, *J. Photochem. Photobiol. C*, 2019, **41**, 100318.
209. L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox and R. Hilgenfeld, *Science*, 2020, **368**, 409.
210. G. Lu, Y. Xu, K. Zhang, Y. Xiong, H. Li, L. Cui, X. Wang, J. Lou, Y. Zhai, F. Sun and X. C. Zhang, *Nat. Commun.*, 2017, **8**, 15948.
211. C. L. Noland, M. D. Kattke, J. Diao, S. L. Gloor, H. Pantua, M. Reichelt, A. K. Katakam, D. Yan, J. Kang, I. Zilberleyb, M. Xu, S. B. Kapadia and J. M. Murray, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 6044-6053.
212. H. C. Pace and C. Brenner, *Genome Biol.*, 2001, **2**, 1-9.
213. J. Y. Kang, X. Nan, M. S. Jin, S. J. Youn, Y. H. Ryu, S. Mah, S. H. Han, H. Lee, S. G. Paik and J. O. Lee, *Immunity*, 2009, **31**, 873-884.
214. Y.-K. Li, H.-J. Yao and I. H. Pan, *J. Biochem.*, 2000, **127**, 315-320.

215. C. E. Weller, A. Dhall, F. Ding, E. Linares, S. D. Whedon, N. A. Senger, E. L. Tyson, J. D. Bagert, X. Li, O. Augusto and C. Chatterjee, *Nature Commun.*, 2016, **7**, 12979.
216. V. Pintado, C. Cabellos, S. Moreno, M. A. Meseguer, J. Ayats and P. F. Viladrich, *Medicine*, 2003, **82**, 346-364.
217. G. Illuminati and L. Mandolini, *Acc. Chem. Res.*, 1981, **14**, 95-102.
218. D. S. Kemp and R. I. Carey, *J. Org. Chem.*, 1993, **58**, 2216-2222.
219. J. Offer, *J. Pept. Sci.*, 2010, **94**, 530-541.
220. G. L. Thomas and R. J. Payne, *Chem Commun*, 2009, 4260-4262.
221. H. Y. Chow, Y. Zhang, E. Matheson and X. Li, *Chem. Rev.*, 2019, **119**, 9971-10001.
222. D. S. Kemp, R. I. Carey, J. C. Dewan, N. G. Galakatos, D. Kerkman and S. L. Leung, *J. Org. Chem.*, 1989, **54**, 1589-1603.
223. D. S. Kemp, D. J. Kerkman, S.-L. Leung and G. Hanson, , *J. Org. Chem.*, 1981, **46**, 490-498.
224. V. Popov, S. S. Panda and A. R. Katritzky, *J. Org. Chem.*, 2013, **78**, 7455-7461.
225. S. S. Panda, A. A. Oliferenko, H. M. Marwani and A. R. Katritzky, *Mendeleev Commun.*, 2014, **24**, 75-77.
226. S. Sohel and T. Opatz, *Arkivoc*, **2014**, 92-108.
227. R. Oliyai, T. J. Siahaan and V. J. Stella, *Pharm. Res.*, 1995, **12**, 323-328.
228. R. Oliyai and V. J. Stella, *Biorg. Med. Chem. Lett.*, 1995, **5**, 2735-2740.
229. C. M. Rojas and J. Rebek, *J. Am. Chem. Soc.*, 1998, **120**, 5120-5121.
230. M. Skwarczynski and Y. Kiso, *Curr Med Chem*, 2007, **14**, 2813-2823.
231. L. P. Miranda, W. D. F. Meutermans, M. L. Smythe and P. F. Alewood, *J. Org. Chem.*, 2000, **65**, 5460-5468.
232. M. Vila-Perello, Y. Hori, M. Ribo and T. W. Muir, *Angew. Chem. Int. Ed. Engl.*, 2008, **47**, 7764-7767.
233. D. S. Kemp, *Biopolymers*, 1981, **20**, 1793-1804.
234. D. S. Kemp, S.-L. Leung and D. J. Kerkman, *Tetrahedron Lett.*, 1981, **22**, 181-184.
235. M. El Khatib, M. Elagawany, F. Jabeen, E. Todadze, O. Bol'shakov, A. Oliferenko, L. Khelashvili, S. A. el-Feky, A. Asiri and A. R. Katritzky, *Org. Biomol. Chem.*, 2012, **10**, 4836-4838.
236. A. R. Katritzky, S. R. Tala, N. E. Abo-Dya, T. S. Ibrahim, S. A. El-Feky, K. Gyanda and K. M. Pandya, *J. Org. Chem.*, 2011, **76**, 85-96.
237. A. A. Oliferenko and A. R. Katritzky, *Org. Biomol. Chem.*, 2011, **9**, 4756-4759.
238. G. E. Schujman and D. de Mendoza, *Nat. Chem. Biol.*, 2006, **2**, 573-574.
239. C. O. Rock, S. E. Goelz and J. E. Cronan, Jr., *J. Biol. Chem.*, 1981, **256**, 736-742.
240. F. Hillmann, M. Argentini and N. Buddelmeijer, *J. Biol. Chem.*, 2011, **286**, 27936-27946.
241. S. Jackowski and C. O. Rock, *J. Biol. Chem.*, 1986, **261**, 11328-11333.
242. Y. Arai, S. Inuki and Y. Fujimoto, *Bioorganic Med. Chem. Lett.*, 2018, **28**, 1638-1641.
243. H. E. Gottlieb, V. Kotlyar and A. Nudelman, *J. Org. Chem.*, 1997, **62**, 7512-7515.

244. A. Gutteridge and J. M. Thornton, *Trends Biochem. Sci.*, 2005, **30**, 622-629.
245. A. Röttig and A. Steinbüchel, *Microbiol. Mol. Biol. Rev.*, 2013, **77**, 277.
246. M.-T. Nguyen, M. Matsuo, S. Niemann, M. Herrmann and F. Götz, *Front. Microbiol.*, 2020, **11**.
247. I. Bertini, C. Luchinat and G. Parigi, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2002, **40**, 249-273.
248. I. Solomon, *Phys. Rev.*, 1955, **99**, 559-565.
249. C. A. Softley, M. J. Bostock, G. M. Popowicz and M. Sattler, *J. Biomol. NMR*, 2020, **74**, 287-309.
250. M. A. Wells, G. S. Jackson, S. Jones, L. L. Hosszu, C. J. Craven, A. R. Clarke, J. Collinge and J. P. Waltho, *Biochem. J*, 2006, **399**, 435-444.
251. K. M. Armbruster, G. Komazin and T. C. Meredith, *J. Biol. Chem.*, 2020, **295**, 10195-10211.
252. J. Dadová, S. R. G. Galan and B. G. Davis, *Curr. Opin. Chem. Biol.*, 2018, **46**, 71-81.
253. T. H. Wright, B. J. Bower, J. M. Chalker, G. J. L. Bernardes, R. Wiewiora, W.-L. Ng, R. Raj, S. Faulkner, M. R. J. Vallée, A. Phanumartwiwath, O. D. Coleman, M.-L. Thézénas, M. Khan, S. R. G. Galan, L. Lercher, M. W. Schombs, S. Gerstberger, M. E. Palm-Espling, A. J. Baldwin, B. M. Kessler, T. D. W. Claridge, S. Mohammed and B. G. Davis, *Science*, 2016, **354**, 1465.
254. B. Josephson, C. Fehl, P. G. Isenegger, S. Nadal, T. H. Wright, A. W. J. Poh, B. J. Bower, A. M. Giltrap, L. Chen, C. Batchelor-McAuley, G. Roper, O. Arisa, J. B. I. Sap, A. Kawamura, A. J. Baldwin, S. Mohammed, R. G. Compton, V. Gouverneur and B. G. Davis, *Nature*, 2020, **585**, 530–537
255. D. A. Contreras-Cruz, M. A. Sánchez-Carmona, F. A. Vengoechea-Gómez, D. Peña-Ortíz and L. D. Miranda, *Tetrahedron*, 2017, **73**, 6146-6156.
256. H. M. Key and S. J. Miller, *J. Am. Chem. Soc.*, 2017, **139**, 15460-15466.
257. B. Amulic, S. L. Knackstedt, U. Abu Abed, N. Deigendesch, C. J. Harbort, B. E. Caffrey, V. Brinkmann, F. L. Heppner, P. W. Hinds and A. Zychlinsky, *Dev. Cell*, 2017, **43**, 449-462 e445.
258. R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683-720.
259. F. L. Zhang and P. J. Casey, *Annu. Rev. Biochem.*, 1996, **65**, 241-269.
260. J. Wang, S. M. Schiller and P. G. Schultz, *Angew. Chem. Int. Ed.*, 2007, **46**, 6849-6851.
261. Y. O. You, M. R. Levengood, L. A. F. Ihnken, A. K. Knowlton and W. A. van der Donk, *ACS Chem. Biol.*, 2009, **4**, 379-385.
262. J. M. Chalker, S. B. Gunnoo, O. Boutureira, S. C. Gerstberger, M. Fernández-González, G. J. L. Bernardes, L. Griffin, H. Hailu, C. J. Schofield and B. G. Davis, *Chem. Sci.*, 2011, **2**, 1666.
263. R. E. Stafford, T. Fanni and E. A. Dennis, *Biochemistry*, 1989, **28**, 5113-5120.
264. W. Lo, S. Huang, S.-L. Zheng and R. H. Holm, *lnorg. Chem.*, 2011, **50**, 11082-11090.
265. J. M. Chalker, G. J. Bernardes, Y. A. Lin and B. G. Davis, *Chem. Asian J.*, 2009, **4**, 630-640.

266. D. de la Fuente-Herreruela, A. K. Monnappa, M. Muñoz-Úbeda, A. Morallón-Piña, E. Enciso, L. Sánchez, F. Giusti, P. Natale and I. López-Montero, *J. Nanobiotechnology*, 2019, **17**, 77-77.
267. R. Petracca, K. A. Bowen, L. McSweeney, S. O'Flaherty, V. Genna, B. Twamley, M. Devocelle and E. M. Scanlan, *Org. Lett.*, 2019, **21**, 3281-3285.
268. C. Orskov, L. Rabenhøj, A. Wettergren, H. Kofod and J. J. Holst, *Diabetes*, 1994, **43**, 535-539.
269. D. Hinnen, *Diabetes Spectrum*, 2017, **30**, 202.
270. U. Ragnarsson and L. Grehn, *RSC Advances*, 2013, **3**, 18691-18697.
271. S. L. Cohen, C. Price and J. Vlasak, *J. Am. Chem. Soc.*, 2007, **129**, 6976-6977.
272. K. Barlos, O. Chatzi, D. Gatos and G. Stavropoulos, *Int. J. Pept. Protein Res.*, 1991, **37**, 513-520.
273. M. P. Doyle and C. T. West, *J. Org. Chem.*, 1975, **40**, 3835-3838.
274. T. S. Chisholm, D. Clayton, L. J. Dowman, J. Sayers and R. J. Payne, *J. Am. Chem. Soc.*, 2018, **140**, 9020-9024.
275. Y. Kajihara, A. Yoshihara, K. Hirano and N. Yamamoto, *Carbohydr. Res.*, 2006, **341**, 1333-1340.
276. B. L. Wilkinson, C. K. Chun and R. J. Payne, *Biopolymers*, 2011, **96**, 137-146.
277. M. C. Munson, C. Garcia-Echeverria, F. Albericio and G. Barany, *J. Org. Chem.*, 1992, **57**, 3013-3018.
278. N. N, V. M. Thimmalapura, B. Hosamani, G. Prabhu, L. R. Kumar and V. V. Sureshbabu, *Org. Biomol. Chem.*, 2018, **16**, 3524-3552.
279. X. Zhang, X.-W. Lu and C.-F. Liu, *Tetrahedron Lett.*, 2008, **49**, 6122-6125.
280. X. Li, T. Kawakami and S. Aimoto, *Tetrahedron Lett.*, 1998, **39**, 8669-8672.
281. M. J. Schöwe, O. Keiper, C. Unverzagt and V. Wittmann, *Chemistry (Weinheim an der Bergstrasse, Germany)*, 2019, **25**, 15759-15764.
282. A. E. Owens, I. de Paola, W. A. Hansen, Y. W. Liu, S. D. Khare and R. Fasan, *J. Am. Chem. Soc.*, 2017, **139**, 12559-12568.
283. T. Yamamura, S. Suzuki, T. Taguchi, A. Onoda, T. Kamachi and I. Okura, *J. Am. Chem. Soc.*, 2009, **131**, 11719-11726.
284. F. Shen, Z. P. Zhang, J. B. Li, Y. Lin and L. Liu, *Org. Lett.*, 2011, **13**, 568-571.
285. Y. Tsuda, Y. Okada, M. Tanaka, N. Shigematsu, Y. Hori, T. Goto and M. Hashimoto, *Chem. Pharm. Bull. (Tokyo)*, 1991, **39**, 607-611.
286. M. Kurimura, A. Ochiai and K. Achiwa, *Chem. Pharm. Bull.*, 1993, **41**, 1965-1970.
287. S. Heuking, A. Iannitelli, A. Di Stefano and G. Borchard, *Int. J. Pharm.*, 2009, **381**, 97-105.
288. G. P. Gential, N. I. Ho, F. Chiodo, N. Meeuwenoord, F. Ossendorp, H. S. Overkleeft, G. A. van der Marel and D. V. Filippov, *Bioorg. Med. Chem. Lett.*, 2016, **26**, 3641-3645.
289. P. Niederhafner, M. Safarik, E. Brichtova and J. Sebestik, *Amino Acids*, 2016, **48**, 1087-1098.
290. F. Reichel, A. M. Roelofsen, H. P. M. Geurts, T. I. Hämäläinen, M. C. Feiters and G.-J. Boons, *J. Am. Chem. Soc.*, 1999, **121**, 7989-7997.

291. O. Busnel, F. Carreaux, B. Carboni, S. Pethe, S. V. Goff, D. Mansuy and J. L. Boucher, *Bioorg. Med. Chem.*, 2005, **13**, 2373-2379.
292. G. Jung, C. Carrera, H. Brückner and W. G. Bessler, *Liebigs Ann. Chem.*, 1983, 1608-1622.
293. Y. A. Lin, O. Boutureira, L. Lercher, B. Bhushan, R. S. Paton and B. G. Davis, *J. Am. Chem. Soc.*, 2013, **135**, 12156-12159.
294. G. L. Weber, R. C. Steenwyk, S. D. Nelson and P. G. Pearson, *Chem. Res. Toxicol.*, 1995, **8**, 560-573.
295. L. Moroder, H.-J. Musiol and G. Siglmüller, *Synthesis*, 1990, 889-892.
296. L. A. Adams, V. K. Aggarwal, R. V. Bonnert, B. Bressel, R. J. Cox, J. Shepherd, J. de Vicente, M. Walter, W. G. Whittingham and C. L. Winn, *J. Org. Chem.*, 2003, **68**, 9433-9440.