Synthesis of Chemical Probes Targeting Membrane Proteins

By Gerard Reid
12305431

Research project presented to Trinity College Dublin, The University of Dublin, for the degree of MSc. By research

Under the supervision of Prof. Martin Caffrey

2019

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.

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.

_____

Gerard Reid

Contents

Acknowledgements

Firstly, I would like to thank Prof. Martin Caffrey for allowing me to do such an exciting project in the Membrane Structural & Functional Biology group. Secondly, I would also like to thank the entire MS&FB group, in particular Dr. Dietmar Weichert, for their help throughout the project. I would also like to thank everyone in the School of Chemistry who helped me to obtain my results: Dr. Manuel Ruether, Dr. Martin Feeney, Dr. Gary Hessman and Dr. John O'Brien.

Summary

Membrane proteins carry out important physiological functions, including roles as enzymes and receptors, and are the targets for 50% of current drugs on the market,[1] indicating their pharmacological importance. Knowledge of the structure and function of membrane proteins is essential in the development of effective therapeutics by structure-based drug design. This project involves the synthesis and characterisation of chemical probes to facilitate structure and function studies on membrane proteins in the Membrane Structural & Functional Biology research group. This project can be divided in three parts; firstly, due to the emerging threat of antibiotic resistance, replenishing the antibiotic development pipeline is of paramount importance. The Membrane Structural & Functional Biology research group has identified the bacterial enzymes lipoprotein diacylglyceryl transferase (Lgt) and apolipoprotein N-acyl transferase (Lnt), as possible targets for new antibiotics. In this work, a number of chromophore labelled phospholipids were synthesised to support the development of a continuous Förster Resonance Energy Transfer assay for both lipoprotein diacylglyceryl transferase and apolipoprotein N-acyl transferase. Such a high-throughput assay could be used to identify inhibitors. Secondly, a monoacylglycerol that is used in the *in meso* method for membrane protein crystallisation was synthesised. In addition, the isomerisation of monoacylglycerols that are used in the *in meso* method and the effect of monoacylglycerol isomerisation on membrane protein crystallisation was studied. Thirdly, synthetic efforts towards unnatural sulfur and selenium containing tryptophan derivatives were carried out in order to facilitate the determination of membrane protein structures.

Abbreviations

1-MAG	1-Monacylglycerol
2-MAG	2-Monacylglycerol
9-BBN	9-Boravicyclo[3.3.1]nonane
Abz	Aminobenzoic acid
Ac_2O	Acetone
AcOH	Acetic acid
AIDS	Acquired immune deficiency syndrome
$CDCl_3$	Deuterated chloroform
$CHCl_3$	Chloroform
CH_2Cl_2	Dichloromethane
D_2O	Deuterated water
DAG	Diacylglyceryl
DCC	*N*,*N*'-Dicyclohexylcarbdodiimide
DMAP	4-Dimethylaminopyridine
DMF	*N*,*N*-Dimethylformamide
DMS	Dimethylsulfide
DMSO	Dimethylsulfoxide
DNP	Dinitrophenyl
EtOAc	Ethyl acetate
FRET	Förster Resonance Energy Transfer
GPCR	G protein-coupled receptor
HCl	Hydrochloric acid
HGT	Horizontal gene transfer
HIV	Human immunodeficiency virus
IR	Infrared
KOH	Potassium hydroxide
Lgt	Lipoprotein diacylglyceryl transferase
Lnt	Apolipoprotein N-acyl transferase
Lol	Lipoprotein outer membrane localisation
LspA	Lipoprotein signal peptide II
MAD	Multi-wavelength anomalous dispersion
MAG	Monoacylglycerol

MDR	Multi drug resistant
MeCN	Acetonitrile
MeOH	Methanol
MRSA	Methicillin-resistant *Staphylococcus aureus*
NMR	Nuclear magnetic resonance
PC	Phosphatidylcholine
Pdb	Protein data bank
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PLP	Pyridoxal 5'-phosphate
RM IM	*Rhizomucor miehei* immobilised
S-SAD	Sulfur-single wavelength anomalous dispersion
Se-Met	Selenomethionine
TBAB	Tetra-*n*-butylammonium bromide
TBAF	Tetra-*n*-butylammonium fluoride
TBDMS	*tert*-Butyldimethylsilyl
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin-layer chromatography
WHO	World health organisation

Chapter 1: Introduction

1.1 Antibiotic resistance

Antimicrobial resistance is predicted to become the greatest threat to human life, causing 10 million deaths per year by 2050 and projected to incur costs of €100 trillion.[2] Antimicrobial resistance encompasses resistance achieved by fungi, viruses, parasites and bacteria. Antibiotic resistance can be defined as the ability of bacteria to change their genetic material in response to antibiotic treatment, that renders the antibiotic ineffective.[2–4] As many antibiotics were originally found in nature (many were discovered as secondary metabolites in actinomycetes bacteria)[5] and not a man-made invention, small-scale innate resistance in bacteria has been shown to have taken place thousands of years before humans existed.[6] Secondary metabolites in bacteria are organic compounds that are not essential for growth/development, unlike primary metabolites such as proteins and carbohydrates.[5]

There are many mechanisms bacteria can take advantage of in order to gain resistance. For example, *Staphylococcus aureus* can replicate through 20 generations in under 12 hours, giving over 1 million progeny.[4,7] This is possible due to *S. aureus'* short doubling time of less than 30 minutes, allowing the bacteria to produce hundreds of mutations.[7]

In addition, bacteria do not only rely on replication for producing favourable mutations, but can acquire genetic material from neighbouring bacteria in a process known as horizontal gene transfer (HGT).[6,8] In HGT, there is a physical movement of DNA from one bacterium to another, followed by incorporation into the latter's genome.[6] In certain situations this can be a competitive process, where one bacterium kills another for their DNA, as seen in the case of *Streptococcus pneumonia*, where toxic, small peptides (bacteriocins) are secreted to lyse neighbouring bacteria.[9] HGT has been shown to be clinically relevant in the spread of antibiotic resistance.[10]

Humans have had a profound impact on the spread of antibiotic resistance, due to a number of factors. Firstly, antibiotic use alone will cause a selective pressure on the bacteria to develop resistance in order to survive. Combining antibiotic misuse by the

public (where antibiotic courses are not finished because individuals begin to feel better) and over-prescription by practitioners has had a detrimental effect on antibiotic performance. Secondly, the over-use of antibiotics in agriculture and in veterinary medicine and its knock-on effect cannot be overstated. Antibiotics are used in agriculture mostly as prophylactics and growth-promoters[4] at sub-therapeutic levels.[11] The use at sub-therapeutic levels favours the emergence of a large number of resistance genes.[11] This is because the sub-therapeutic levels do not kill the bacteria, but exposes them to antibiotics, to which they can begin to develop resistance. Astoundingly, in the United States, over 80% of all antibiotic consumption can be ascribed to agriculture,[4,12] with 60% of those antibiotics being of critical, medical importance for humans.[4]

1.2 ESKAPE pathogens

Just over 10 years ago, a special group of bacteria were highlighted due to their clinical importance and grouped together to form the 'ESKAPE' Pathogens. The group is made up by: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumonia*, *Acinetobacter baumanii*, *Pseudomonas aeruginosa*, and *Enterobacter* species.[13] Of the six ESKAPE pathogens, four represent Gram-negative bacteria: *K. pneumonia*, *A. baumanii*, *P. aeruginosa*, and *Enterobacter* species. The complex membrane structure of Gram-negative bacteria makes them difficult to treat (Fig. 1.1 a, b).[14]

The ESKAPE pathogens' clinical importance stems from their ability to 'escape' current antibiotic treatment by showing high levels of antibiotic resistance, representing a new wave of resistant bacterial strains that has not been observed before.[13–15] These levels of resistance convey the full extent of the impressive genetic flexibility possessed by bacteria.

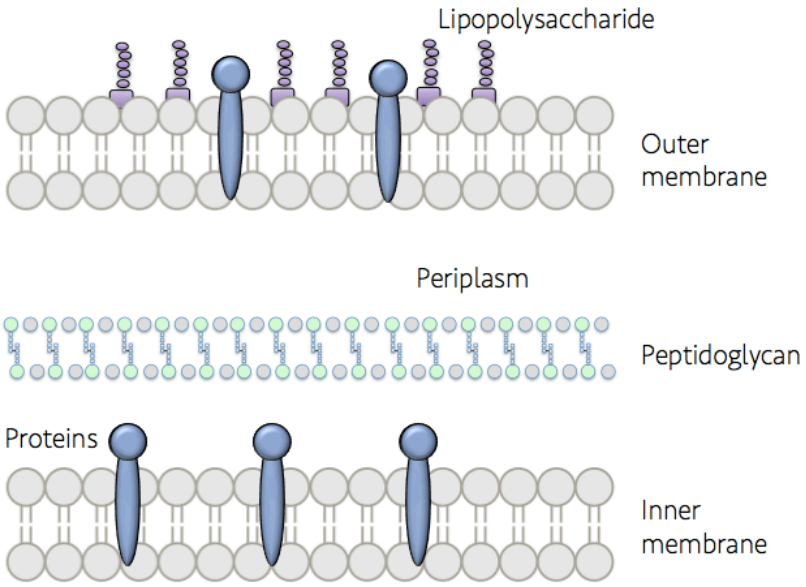


Figure 1.1a The membrane structure in Gram-negative bacteria. Gram-negative bacteria have a more complex membrane structure than their Gram-positive counterparts. Gram-negative bacteria contain an inner and an outer membrane with a thin layer of peptidoglycan in their periplasm (the area between the inner membrane and the outer membrane). The peptidoglycan has a layered structure that is strengthened by cross-links. The outer membrane is functionalised with lipopolysaccharide on the outer leaflet and with lipoproteins on the inner leaflet. The presence of a second membrane barrier makes antibiotic development for Gram-negative bacteria difficult.

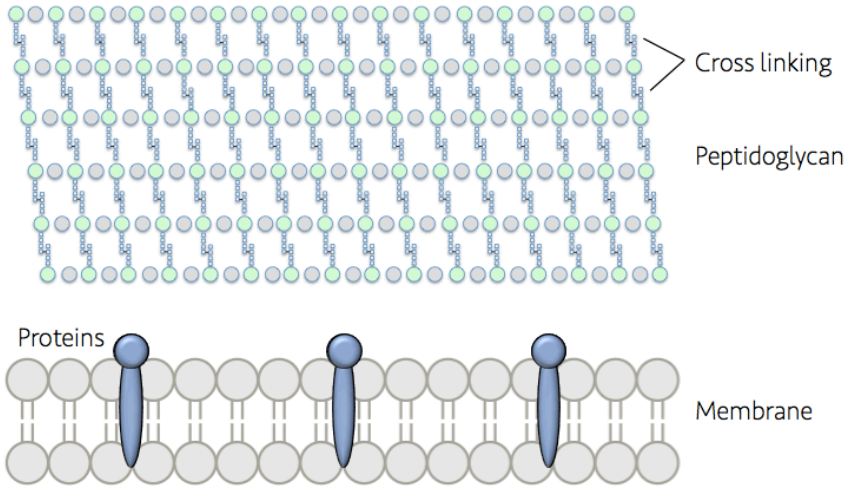


Figure 1.1b The structure of the Gram-positive bacteria membrane. In contrast to Gram-negative bacteria, the envelope of Gram-positive bacteria is characterised by a lack of an outer membrane. A single cytoplasmic membrane is protected by a thick layer of peptidoglycan, a far thicker layer than is present in Gram-negative bacteria.

Enzymes that form cross-links in peptidoglycan can be inhibited by antibiotics, such as penicillins, cephalosporins and glycopeptides.

ESKAPE pathogens cause the majority of nosocomial (hospital acquired) infections.[13] They tend to be opportunistic pathogens that infect immuno-compromised individuals, such as cystic fibrosis patients, cancer patients and burn victims.[14] The consequences of these nosocomial infections by ESKAPE pathogens include increased mortality and prolonged stays in hospitals. In the U.S., every year more people die from methicillin-resistant *Staphylococcus aureus* (MRSA) nosocomial infections alone than from HIV (AIDS) and tuberculosis combined.[15]

There is currently no new therapy for multi-drug resistant (MDR) *P. aeruginosa* or carbapenem resistant *K. pneumonia*, causing clinicians to revert to older drugs such as colistin (polymixin E).[15] Colistin was developed in the 1950s, but was then discarded due to concern regarding its neuro- and nephrotoxicity.[14–16] Due to its reduced use, the drug has not been invalidated by resistance. It is therefore being kept as an 'antibiotic of last resort' for MDR Gram-negative infections. As the use of colistin now begins to increase, as will associated resistance in Gram-negative bacteria, highlighting the need for novel antibiotics with novel drug targets and mechanisms of action.

1.3 Current antibiotics & need for new drugs

Antibiotics currently available for use include penicillins and cephalosporins (alone and in combination with beta-lactamase inhibitors), glycopeptides, macrolides, polymixins, carbapenems, lipopeptides and tetracyclines.

These current classes of antibiotics were nearly all discovered during the 'golden period' of antibiotic discovery between the 1940s and the 1960s. Antibiotics have facilitated advances in modern medicine, as without them common procedures such as dental surgery or cesarean section could not take place safely. The development of antibiotics was one of the greatest human feats, but in the years since 1960, the majority of new antibiotics have been 'me too' drugs.[14,17] Companies have identified current, commercially successful antibiotics on the market and introduced minor modifications to the core structure of established compounds to create so-called 'me too' analogues. The best performing analogue then made its way through clinical

trials and onto the market. These 'me too' antibiotics do not have a novel mode of action, and will therefore be susceptible to the same resistance mechanisms. The 'me too' antibiotics are financially less risky to develop, as the likelihood of toxicity and secondary interactions are reduced when compared with the development of a brand new antibiotic classes against new targets.

The current antibiotic pipeline is not sufficient to combat the alarming rates of resistance, particularly among ESKAPE pathogens.[14] The pipeline has been slowly decimated because antibiotic development is not an economically attractive option for large pharmaceutical companies. The fact that many pharmaceutical companies are pulling out of research and development of new antibiotics is disappointing but not surprising.[4] From a purely financial point of view, there are a number of factors that will undermine antibacterial investment. Firstly, the limited duration of antibiotic treatment is the main disadvantage.[4] Antibiotics will typically be prescribed for a period of only 5-10 days, unlike other medicines that may have to be taken everyday for a lifetime. This will naturally result in lower sales and revenue. Secondly, antibiotics usually have a low price per dose, so companies cannot recoup the losses from short duration of treatments by inflating price.[4] Thirdly, once the product makes it to market, there is no guarantee of market longevity due to potential antibiotic resistance.[4,14] Lastly, if a successful antibiotic is developed, *i.e.*, an antibiotic that causes low-level resistance in bacteria, it will most likely be kept as a drug of last resort, again reducing sales and revenue. The lack of interest from many pharmaceutical companies has depleted the antibiotic pipeline and passed the burden of antibiotic discovery to academic institutions.

1.4 Lipoprotein processing pathway – A novel drug target

Lipidation is a common post-translational modification to proteins in bacteria. Both Gram-positive and Gram-negative bacteria contain a large amount of lipoproteins (*P. aeruginosa* and *S. aureus* have 175 and 67 lipoproteins, respectively).[18,19] The importance of protein lipidation in bacteria is illustrated by the number of vital physiological roles carried out by lipoproteins, including roles in transport, signal transduction, virulence and in penicillinase activity.[20–23]

Lipoproteins in bacteria are processed by a three-enzyme pathway termed the lipoprotein processing pathway (Fig. 1.2). This pathway is exclusive to bacteria, *i.e.*, not possessed by humans.[21,24] It also has been shown that these enzymes are essential in many organisms for bacteria survival and virulence.[21]

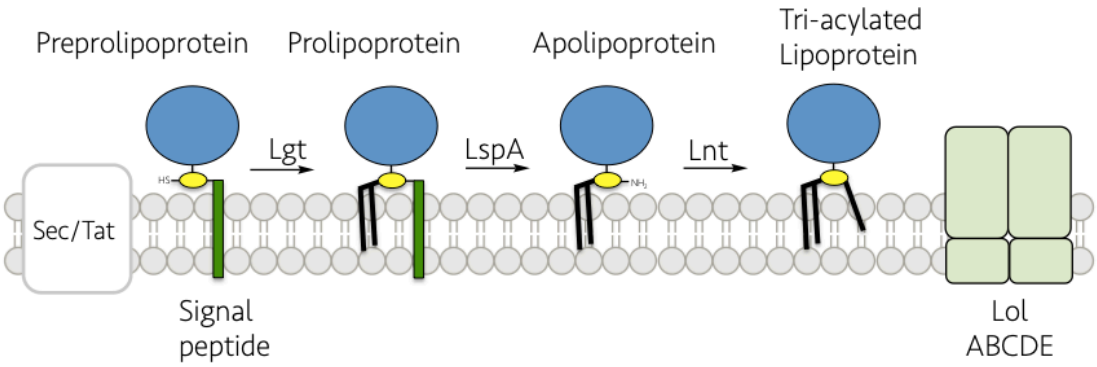


Figure 1.2 The post-translational lipoprotein processing pathway. The preprolipoproteins arrive via the Sec or Tat pathways and are anchored in the inner membrane by a single transmembrane helix called the signal peptide (green). The first step of the scheme involves the addition of a diacylglyceryl group to the free thiol of an invariant cysteine residue (yellow) by lipoprotein diacylglyceryl transferase (Lgt). Next, lipoprotein signal peptidase II (LspA) cleaves off the signal peptide and reveals a primary amine of a conserved cysteine. Finally, apolipoprotein *N*-acyl transferase (Lnt) acylates the apolipoprotein at the *N*-terminus of the lipobox cysteine to form the tri-acylated lipoprotein. The lipoprotein may then go on to be localised to the inner leaflet of the outer membrane by the Lol machinery. A detailed description of the reactions catalysed by Lgt, LspA and Lnt is provided below.

In order to begin structure-based drug design, it is necessary to have a high-resolution structure of the target protein. The structure of lipoprotein diacylglycerl transferase (Lgt) was solved by Zhang *et al.*[25] and the structures of lipoprotein signal peptidase II (LspA) and apolipoprotein *N*-acyl transferase (Lnt) were determined by Vogeley *et al.*[18] and Wiktor *et al.*[26] from this research group. As lipoproteins are predominantly membrane-associated proteins, lipidation is a crucial modification in facilitating their anchorage into the membrane. Lipidation can also play a role in increasing the stability and flexibility of the protein.[21]

All bacterial preprolipoproteins contain a four-residue lipobox motif that allows recognition by Lgt. The residues that make up the lipobox motif consist of $L(VI)^{-3}$-$A(S)^{-2}$-$G(A)^{-1}$ C^{+1}, with the cysteine at the +1 position being absolutely conserved and

a limited amount of variation allowed with the first three residues. The numbers in superscript (-3 to +1) indicate the position of proteolytic cleavage by LspA. Cleavage occurs between residues -1 and + 1 and the cysteine at the +1 position will become the new N-terminal amino acid.

After lipoprotein processing, the Lol (Lipoprotein Outer Membrane Localisation) machinery then sorts the tri-acylated lipoproteins. The Lol system in Gram-negative bacteria is influenced by the amino acid occupying the +2 position. The presence of an Asp residue at the +2 position causes localisation to the inner membrane, whereas any other amino acid in the +2 position results in localisation to the outer membrane.[27] Of course, in Gram-positive bacteria, no such differentiation needs to take place.

1.4.1 Lgt (lipoprotein diacylglyceryl transferase)

Lgt is an integral membrane protein present in both Gram-negative and Gram-positive bacteria and has been shown to be essential in all Gram-negative bacteria, while its absence in Gram-positive bacteria results in attenuated virulence.[24] Inhibition or depletion of Lgt will cause accumulation of preprolipoprotein and will prevent the growth of bacteria.[28]

The preprolipoprotein is inserted into the inner membrane from the cytoplasm by the Sec or Tat secretion machinery. Lgt catalyses the transfer of a diacylglyceryl group from a phosphatidylglycerol (PG) phospholipid to the free thiol group of the conserved cysteine in the lipobox of the preprolipoprotein. The transfer is achieved through a thioether linkage and the products are a di-acylated ‘prolipoprotein’ and the PG headgroup (Fig. 1.3).

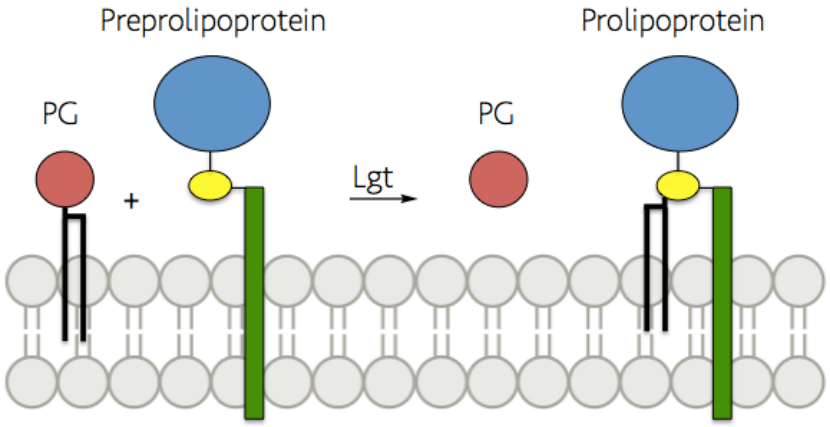


Figure 1.3 Lgt catalyses the transfer of a diacylglyceryl group from a phosphatidylglycerol (PG) to the free thiol of the lipobox cysteine (yellow).

Kinetics experiments on the specificity of the reaction between the preprolipoprotein and Lgt have been carried out.[21] Lgt shows specificity for PG phospholipids over other phospholipid headgroups. A high level of specificity is exhibited, indicating that the headgroup plays a significant role in binding to the active site. From the crystal structure, the glycerol headgroup of the PG forms favourable electrostatic interactions with a conserved Arg residue.[25] The Lgt/phospholipid reaction will not take place if another phospholipid headgroup is employed.[21]

1.4.2 LspA (lipoprotein signal peptidase II):

The second enzyme in the pathway is LspA. The enzyme is a member of the aspartyl protease family, as its active site consists of two aspartyl residues.[18] At physiological pH, these aspartic acid residues will both be deprotonated and participate in the protease activity by deprotonating water molecules and stabilising intermediates. As indicated previously, LspA cleaves the prolipoprotein – the product of Lgt – at the lipobox cysteine, freeing the cysteine for further processing (Fig. 1.4).

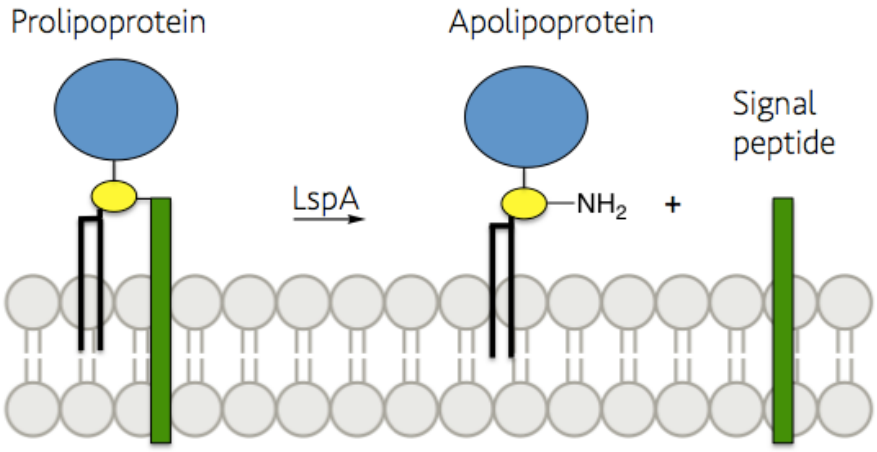


Figure 1.4 LspA catalyses the proteolytic cleavage of the signal peptide (green) at the N-terminus of the lipobox cysteine (yellow). The resulting apolipoprotein is still anchored to the membrane by the diacylglyceryl group.

Unlike Lgt and Lnt, the LspA reaction does not require phospholipids.[21] The product of the LspA reaction is simply termed 'apolipoprotein' and now contains a primary amine at the N-terminal lipobox cysteine. This primary amine will now act as a substrate for the following enzyme in the pathway; Lnt. LspA can be inhibited by the antibiotics globomycin and myxovirescin. Inhibition causes accumulation of the diacylglyceryl-modified prolipoproteins and subsequent cell death.[29]

1.4.3 Lnt (Apoliprotein *N*-acyl transferase)

Lnt is the final enzyme in the pathway and catalyses the formation the mature tri-acylated lipoprotein that may go on to be sorted by the Lol pathway. Lnt is not as widely conserved as Lgt and LspA, and is only present in proterobacteria and actinomycetes.[21]

Lnt contains a membrane domain and an extra-cellular nitrilase-like domain. The latter domain harbours an active site cleft containing a catalytic triad comprised of C, E and K.[26] Lnt catalyses the transfer of a fatty acid from a phospholipid substrate to the N-terminus of the lipobox cysteine (Fig. 1.5). The transfer is thought to occur via a two-step mechanism. The fatty acid is first transferred from the *sn*-1 position (Fig. 1.6) of the phospholipid to the cysteine of the catalytic triad of Lnt, generating a thioester intermediate (Fig. 1.7). Then, the di-acylated lipoprotein (LspA product)

binds to this intermediate resulting in the transfer of the fatty acid to the free N-terminal group of the lipoprotein's lipobox cysteine.

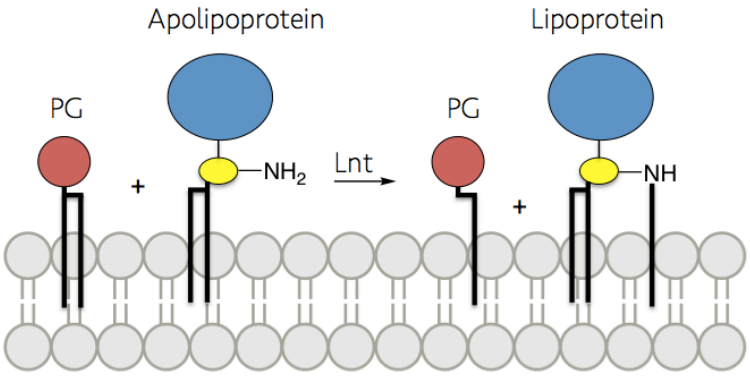


Figure 1.5 Lnt catalyses the transfer of a fatty acid from the *sn*-1 position of a phospholipid to the N-terminus of the lipobox cysteine (yellow).

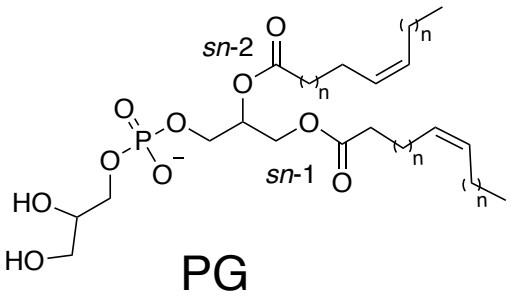


Figure 1.6 Structure of a PG phospholipid showing the difference between the *sn*-1 and *sn*-2 positions. The *sn*-1 alcohol is a primary alcohol, whereas the *sn*-2 alcohol is secondary.

Kinetic studies of the specificity of phosphate headgroup have also been carried out for Lnt.[30] It was discovered that Lnt has a preference for phosphatidylethanolamine (PE) phospholipids, with a saturated fatty acid at the *sn*-1 position and an unsaturated fatty acid at the *sn*-2 position. While there is a preference for PE, PG also supports Lnt activity.[30]

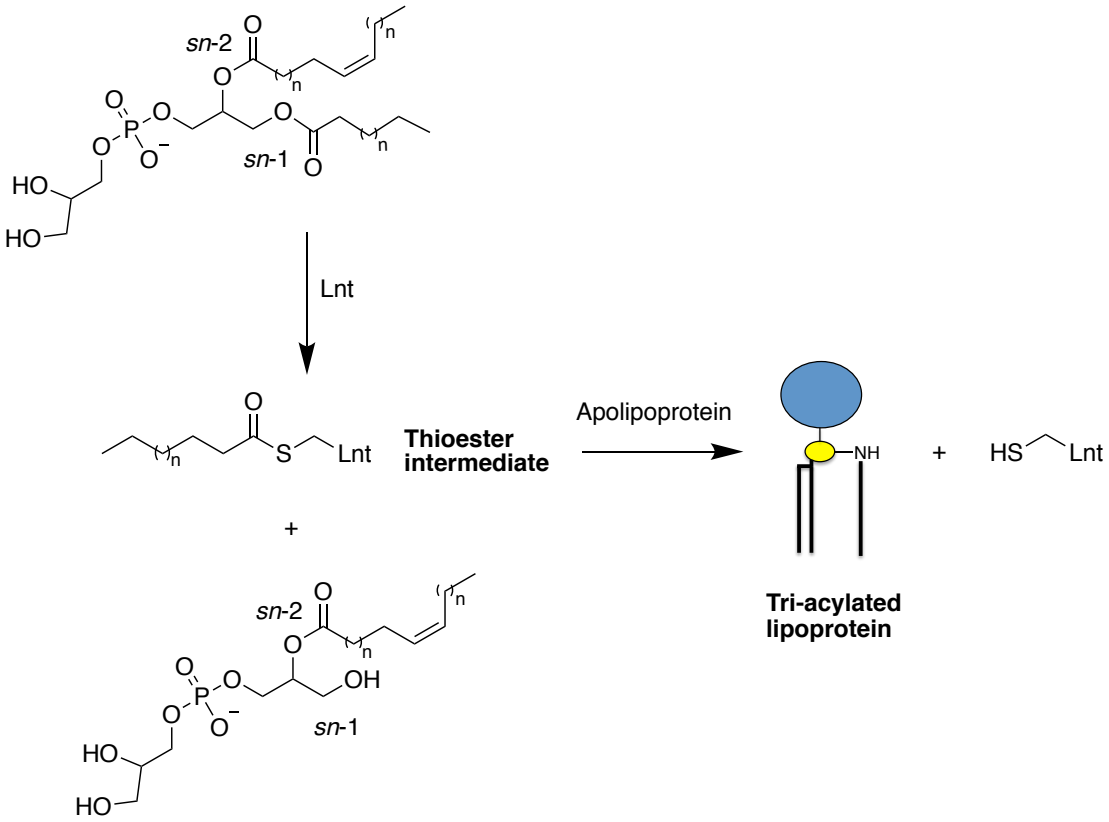


Figure 1.7 Illustrating the two-step mechanism of Lnt. Firstly, the *sn*-1 fatty acid of a phospholipid is transferred to the catalytic cysteine of Lnt (a 1-lyso phospholipid is formed). In the second step, the fatty acid of the thioester intermediate is transferred to the apolipoprotein to form the tri-acylated lipoprotein and the free thiol of Lnt is restored.

Lnt is thought to be an essential enzyme in many Gram-negative bacteria.[31,32] The lack of the enzyme, or the inhibition of the enzyme will result in cell death.[31]

1.4.4 Lipoprotein protein processing pathway and drug development

All three enzymes in the lipoprotein processing pathway would be good candidates for structure-based drug design. Their potential as future druggable targets lies in the fact that these new antibiotics would have a novel mode of action. Very few new classes of antibiotics have been brought to the market in the last 40 years, since the end of the 'golden era' of antibiotic development.[33]

The enzymes in the lipoprotein processing pathway are essential to the survival and virulence of many organisms.[21] In addition, there already exists two known inhibitors of LspA; globomycin and myxovirescin (Fig 1.8). The existence of these inhibitors

gives optimism for future structure-based drug design of LspA and the entire pathway, as knowledge of their structures provides medicinal chemists a scaffold on which to build a library of potential new antibiotics.

Importantly, Lgt, LspA and Lnt are not present in humans.[18] Although Lnt contains a nitrilase-like domain, mammalian nitrilases act on different substrates.[26] The active site cleft faces outwards in Lnt, to provide an opening for substrates and products,[26] giving an attractive feature for drug design.

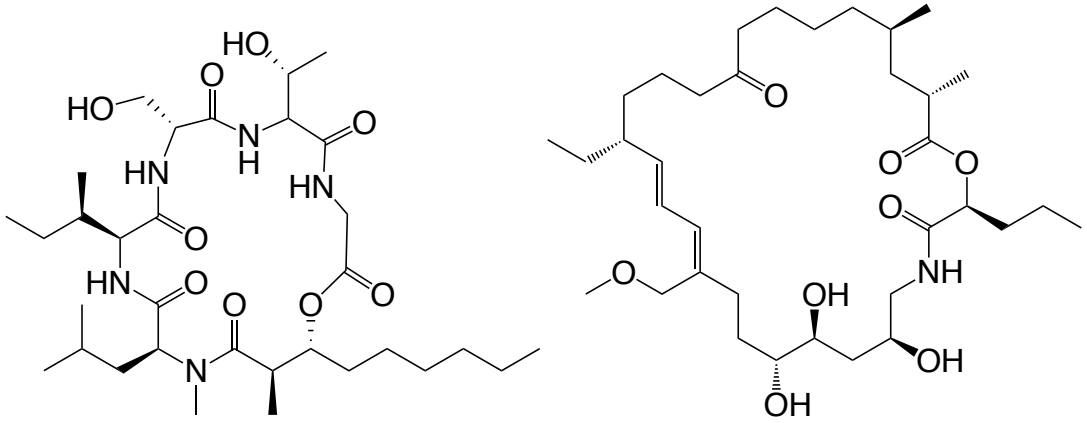
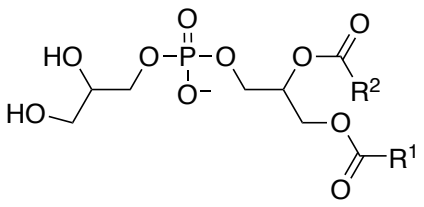


Figure 1.8 The structures of the depsi-peptide, globomycin (left), and myxovirescin (right). Both are inhibitors of the second enzyme in the pathway, LspA.

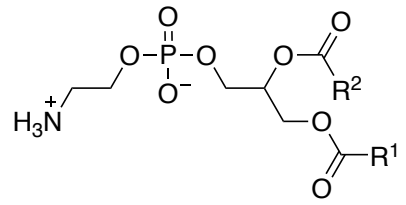
1.5 Glycerophospholipids

Glycerophospholipids are a diverse group of amphiphilic molecules that are biologically important and perform a variety of roles; as second messengers, as energy and heat sources, in drug delivery systems and most markedly, acting as a barrier in membrane lipids in both humans and bacteria.[34–36] For example, in *E. coli*, the membrane composition is 75% PE, 20% PG and cardiolipin.[36]

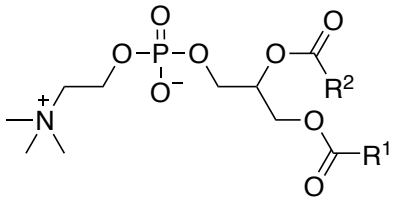
Glycerophospholipids can be divided into a number of families based on their headgroup structure (Fig. 1.9).



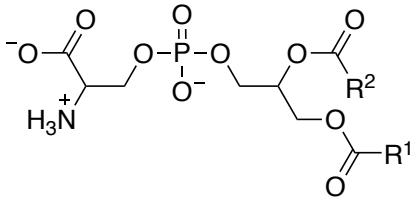
Phosphatidylglycerol



Phosphatidylethanolamine



Phosphatidylcholine



Phosphatidylserine

Figure 1.9 The structures of the most common families of glycerophospholipids. PG, PE, phosphatidylcholine (PC) and phosphatidylserine (PS). R^1 and R^2 represent fatty acids.

Glycerophospholipids contain a diacylated glycerol spine (shown with R^1 and R^2 fatty acids). The *sn*-1 position (R^1) is most often a long, saturated fatty acid chain, while the *sn*-2 position is predominantly occupied by a long, unsaturated fatty acid. At the centre is a negatively charged phosphate.

The discrimination between the glycerophospholipid families is achieved by the variety shown in their headgroups. PE contains an ethylene linker to a primary amine cation that neutralises the negatively charged phosphate centre, becoming zwitterionic. Phosphatidylcholine (PC) contains an ethylene linker to a choline group – a triply methylated amine cation, also reaching an overall neutral charge. PGs contain a second glycerol backbone, giving an overall negative charge to the molecule and phosphatidylserine (PS) contains an ethylene linker to an amino-carboxylic acid.

Glycerophospholipids are the source of the fatty acids that post-translationally modify the preprolipoprotein and the apolipoprotein in the lipoprotein processing pathway.[21] In this project, a number of glycerophospholipids were synthesised to support the development of a Förster Resonance Energy Transfer (FRET) assay (Chapter 2).

1.6 Structure determination – X-ray crystallography:

A main focus of this research group is structure elucidation by X-ray crystallography of membrane proteins using the *in meso* method. The three enzymes in the lipoprotein processing pathway are membrane proteins and two of the three structures were determined in this group by the *in meso* method (LspA and Lnt).[18,26] Currently, the most powerful tool for elucidating the structure of the described enzymes is macromolecular X-ray crystallography. The resulting crystal structure provides detailed information that is invaluable for structure-based drug design.

The *in meso* method was introduced in the late 1990s by Landau and Rosenbusch.[37] Its uptake into the field was slow at first, but the method has experienced an explosive growth since 2012, with the *in meso* method being responsible for over 500 structures in the pdb at the time of writing (400 since 2012). The *in meso* method involves reconstituting the membrane protein into a specific monoacylglycerol (MAG)/water mixture that forms a bicontinuous cubic phase. MAGs consist of glycerol with one fatty acid attached as an ester. The cubic phase can be made with knowledge of the phase diagram of the particular MAG used. For example, the phase diagram of monoolein (Fig. 1.10), the most commonly used lipid in the *in meso* method, shows the cubic phase will form at 40% hydration with water at 20 ^{o}C. The cubic phase mimics a biological membrane, and for this reason, it provides an environment in which membrane proteins should be stable, remain folded and have an opportunity to crystallise.

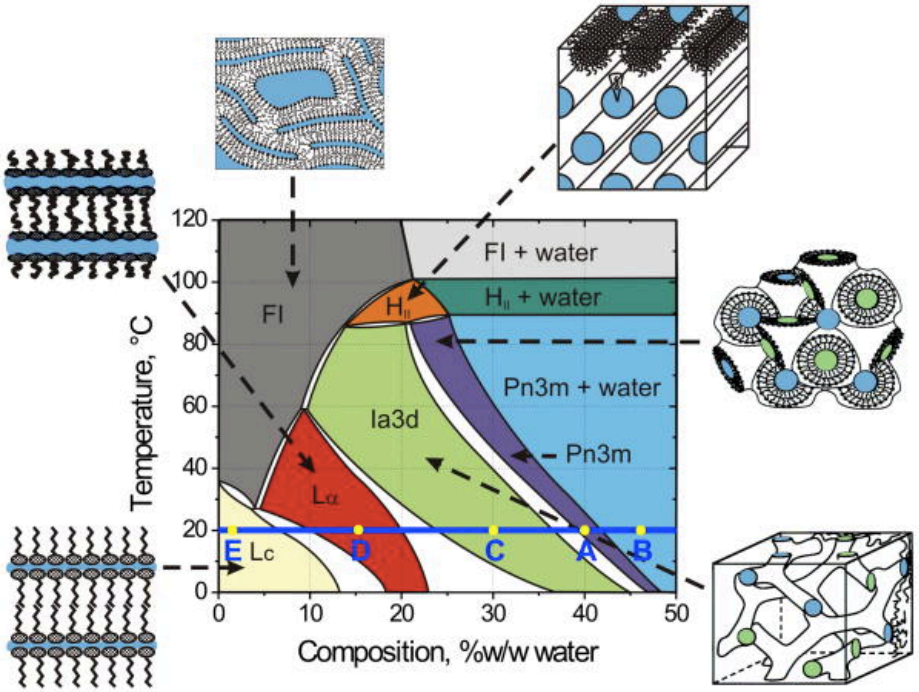


Figure 1.10 The phase diagram of monoolein, with temperature on the y-axis and the % of water on the x-axis. The Pn3m and Ia3d cubic phases can be reached at 20 ^{o}C and at 40% water composition.[38]

This protein containing cubic phase mixture is then spotted on a glass plate and covered with an excess of 'precipitant' solution. The exact mechanism of protein crystal formation in the cubic phase is not known, but a rational hypothesis has been proposed.[39,40] In this hypothesis, the protein molecules are free to diffuse through the cubic phase, but upon addition of the appropriate precipitant solution, a phase separation is triggered.[39] A local lamellar phase is formed, becomes saturated with protein and facilitates nucleation and crystal growth (Fig 1.11).

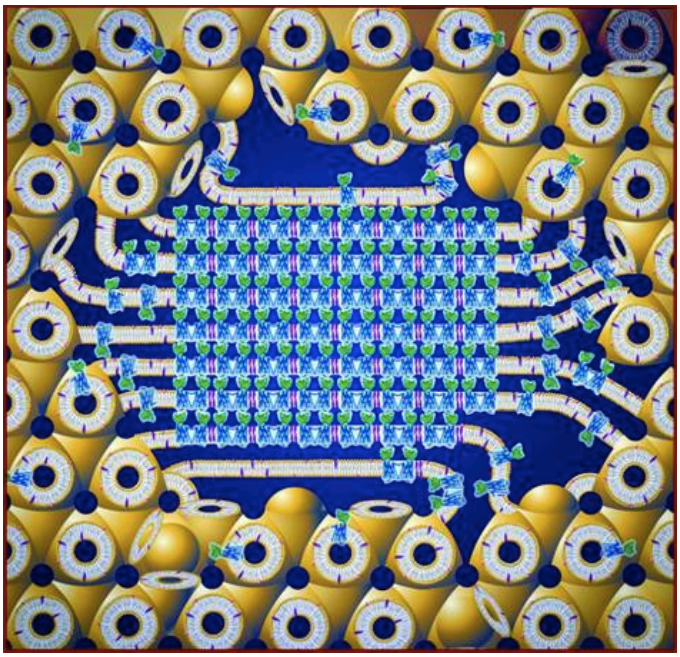


Figure 1.11 Schematic of the protein-laden cubic phase after addition of the precipitant solution. Phase separation occurs: the protein diffuses from the bicontinuous cubic phase (exterior) to a lamellar region (centre) to cause nucleation and crystal growth.[40]

The choice of the MAG that forms the cubic phase can be varied and is a crucial variable in the formation of structure grade crystals. Historically, since the inception of the *in meso* method, the cubic phase has been made using monoolein (9.9 MAG, Fig. 1.12).

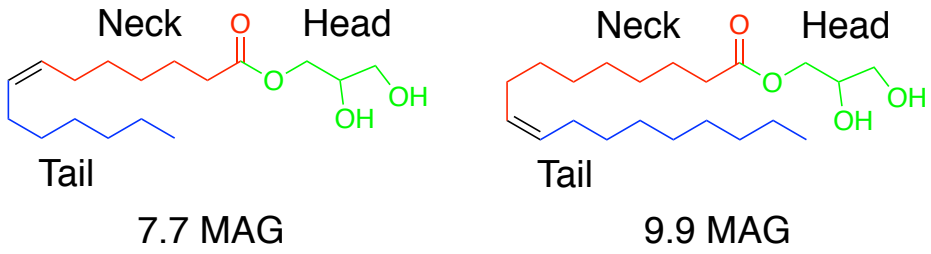


Figure 1.12 Structures of 7.7 MAG and 9.9 MAG (monoolein), commonly used MAGs in the *in meso* method.

The MAGs used in the *in meso* method have a similar structure - the polar, glycerol head group (green, Fig 1.13), the neck (red, carbons before the *cis* double bond) and the tail (blue, carbons after the *cis* double bond). The naming of MAGs relies upon the number of carbons either side of the double bond. The neck (N) of the compound

is defined as the part of the molecule between the ester and the double bond, while the portion after the double bond is the tail (T): N.T MAG, so the below compound (Fig. 1.13) is named 7.8 MAG, as it contains 7 carbons before the *cis* double bond (red) and 8 carbons after the *cis* double bond (blue).

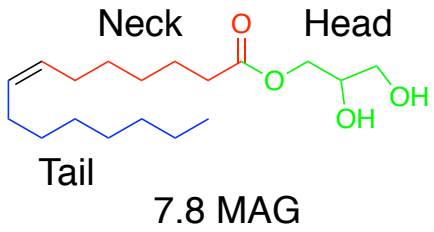


Figure 1.13 Structure showing the N.T naming system of MAGs.

In this project, 7.7 MAG (Fig 1.12, left) was synthesised according to a procedure developed in this research group, with some modifications.[41] 7.7 MAG is so named as it has 7 carbon atoms either side of the *cis* double bond, *i.e.*, a 14 carbon fatty acid chain.

The 7.7 MAG that was synthesised in this project was then used in the *in meso* method in this research group to grow protein crystals. 7.7 MAG is an alternative to 9.9 MAG and often leads to crystallisation of proteins that do not form crystals in 9.9 MAG.

7.7 MAG has recently become the lipid of choice to elucidate the structures of G-Protein Coupled Receptor (GPCR)-G protein complex,[42] as the shorter chain length (compared to 9.9 MAG) gives its cubic phase a narrower bilayer thickness and larger aqueous channels.[43] These larger aqueous channels are preferable for the crystallisation of GPCR-G protein complexes, as they contain large soluble domains that can diffuse freely in the enlargened aqueous channels formed by 7.7 MAG, but not 9.9 MAG.[43,44]

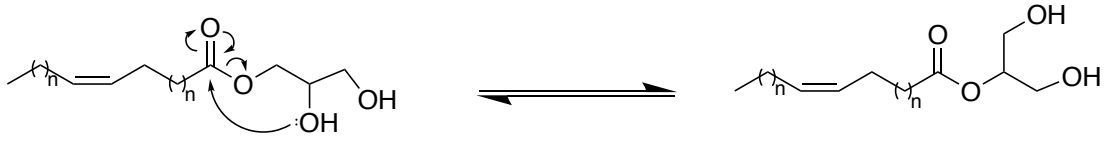
While 7.7 MAG can be obtained from commercial vendors, the cost is exorbitant and the quality of the material is inconsistent. This inconsistency has frustrated the progression of the field and led to the formation of a lipid synthesis program in the MS&FB group.

1.7 The role of 2-monoacylglycerols (2-MAGs) in crystallogenesis

The *in meso* method, as described in Section 1.6, involves a complex system that is made up by the membrane protein solubilised in aqueous solution using detergent, the MAG and the precipitant solution. The MAG is often presumed pure, but this is never truly the case, as 2-MAGs are formed by isomerisation/transesterification of 1-MAGs (such as 7.7 MAG and 9.9 MAG, Scheme 1.1).

In-house synthesised and commercial MAGs contain a small amount of their 2-MAG isomer, usually in the region of 1-10%. This small percentage originates from spontaneous isomerisation during preparation and purification. Therefore, the 2-MAG isomer of the MAG will be present in the cubic phase system of the *in meso* method in which protein crystals are grown, adding another variable to the method. In addition, further isomerisation may even take place in the cubic phase during crystallisation.

Scheme 1.1 Isomerisation/transesterification of 1-MAGs to 2-MAGs



The role of 2-MAGs in crystallogenesis using the *in meso* method, if any, is often overlooked and is completely unknown. In this project, 2-monoolein (Fig 1.14) was synthesised by a newly developed synthetic scheme (Chapter 3). This synthetic 2-monoolein could then be doped into its 1-isomer, monoolein, and would then allow the role of 2-MAG in the *in meso* to be studied.

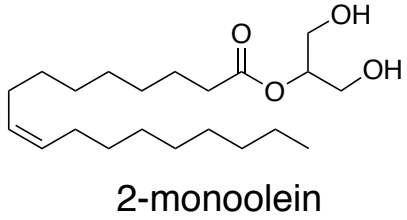


Figure 1.14 Structure of 2-monoolein

1.8 Solving the crystallographic phase problem

Once a crystallisation procedure has been optimised and crystals of the membrane protein of interest have been formed, the next step in macromolecular X-ray crystallisation is the use of X-rays to elucidate the structure. The beam of light from the synchrotron is then diffracted by electrons in the atoms of the protein to produce a unique diffraction pattern (Fig 1.15). In the process of recording diffraction information, intensities are registered, but the phase data of the crystal is lost. Solving the phase problem goes hand in hand with the crystallisation of proteins, as without the phase data the position of the atom in space is unknown, and therefore, the structure cannot be solved.

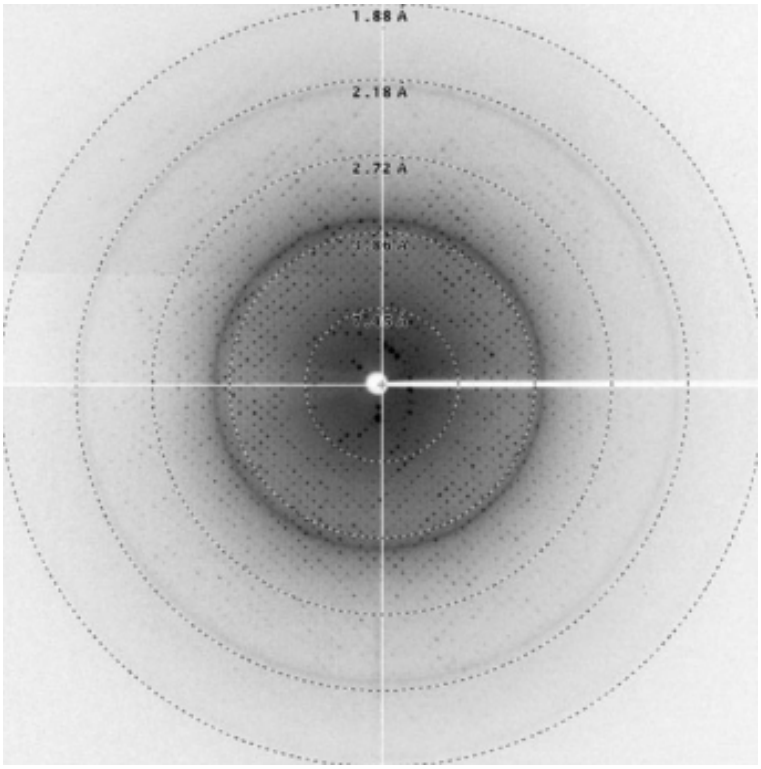


Figure 1.15 A sample diffraction pattern.[45]

Phasing represents a significant barrier in macromolecular crystallography but fortunately, many techniques for solving the phase problem have been developed, including the addition of heavy atoms to the structure and identifying their location in the unit cell and molecular replacement by using the phase information of a similar structure that has been solved before. Heavy atom-based phasing can include the substitution of methionine for seleno-methionine (Se-Met). Se-Met incorporation

allows the phase problem to be solved by single-wavelength anomalous dispersion (SAD) or multi-wavelength anomalous dispersion (MAD) methods, but Se-Met phasing can often be unreliable, as expression can often lead to lower yields, poorer diffracting crystals and a low signal to noise ratio, caused by the methionine side chain that is quite flexible and disordered.[46]

Recently, this research group explored the idea that sulfur or selenium containing derivatives of tryptophan could be used for phasing (Fig. 1.16). Tryptophan has the advantage that it is relatively abundant in membrane proteins and tryptophan residues are also usually well-defined in a membrane protein structure as they are important for anchoring the protein at the membrane interface.[47]

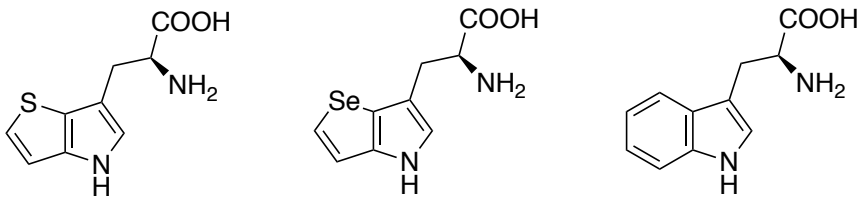


Figure 1.16 Structure of sulfur- and selenium-tryptophan (left and middle, respectively) and natural tryptophan (right). The indole ring of tryptophan is replaced with thienopyrrole and selenopyrrole moieties.

The substituted tryptophan would contain the sulfur/selenium incorporated into the ring moiety and should give a stronger signal during phasing, due to its reduced flexibility when compared to their S/Se-Met counterparts. In this project, the sulfur and selenium tryptophan derivatives were synthesised according to literature procedures[48] with minor modifications. The unnatural amino acids could then be incorporated into proteins and used to solve the crystallographic phase problem. If successful, the method could then go on to replace phasing using Se-Met, that is currently the most common method used, if molecular replacement cannot be carried out.

The research performed for this MSc has been organised in three different sub-projects and these will be discussed in the following chapters.

Chapter 2: Glycerophospholipid Synthesis for use in a FRET assay for Lgt and Lnt

2.1 Introduction

This chapter will deal with the synthesis of a number of chromophore labelled phosphatidylglycerol derivatives for possible use in Förster Resonance Energy Transfer (FRET) assays for Lgt and Lnt. The synthetic glycerophospholipids were to contain a chromophore at their *sn*-1 position.

2.1.1 Rationale for continuous Lnt FRET assays

A FRET assay consists of a donor and an acceptor, where the donor is fluorescent and the acceptor is a chromophore. The donor and acceptor would work in tandem; the fluorophore would transfer energy at a wavelength where the chromophore could absorb when they are in close proximity. This transfer of energy can be monitored by the increase/decrease of fluorescence and it shows that the fluorescent donor and the chromophore are now in close proximity to each other.

At the outset of this project, a number of assay ideas were considered and could have been included in this project (Fig. 2.1 – 2.3). The many variations of the FRET assay illustrate the complexity of the assay systems.

Assay System A:

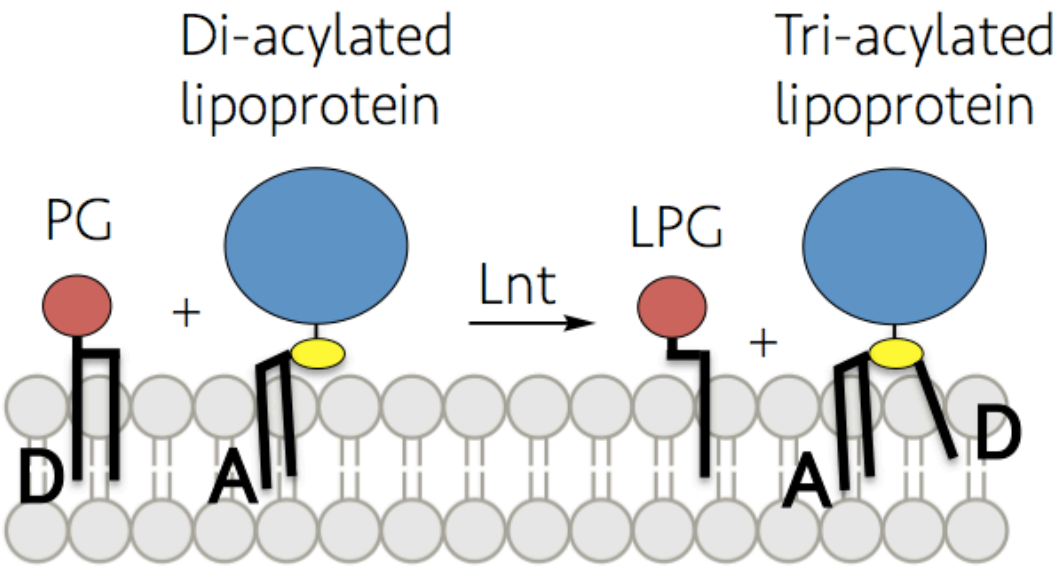


Figure 2.1 Assay idea involving a PG phospholipid with a fluorescent donor on the *sn*-1 chain of the phospholipid and acceptor on the diacylglyceryl moiety of the di-acylated lipoprotein (D = donor, A = acceptor).

In this assay, lyso-PG and a tri-acylated lipoprotein are the products of the assay. Fluorescence from the donor (D) on the PG will be observed at t_o. Upon transfer of the *sn*-1 chain containing the donor by Lnt to the di-acylated lipoprotein, FRET will take place, *i.e.*, quenching of fluorescence.

B:

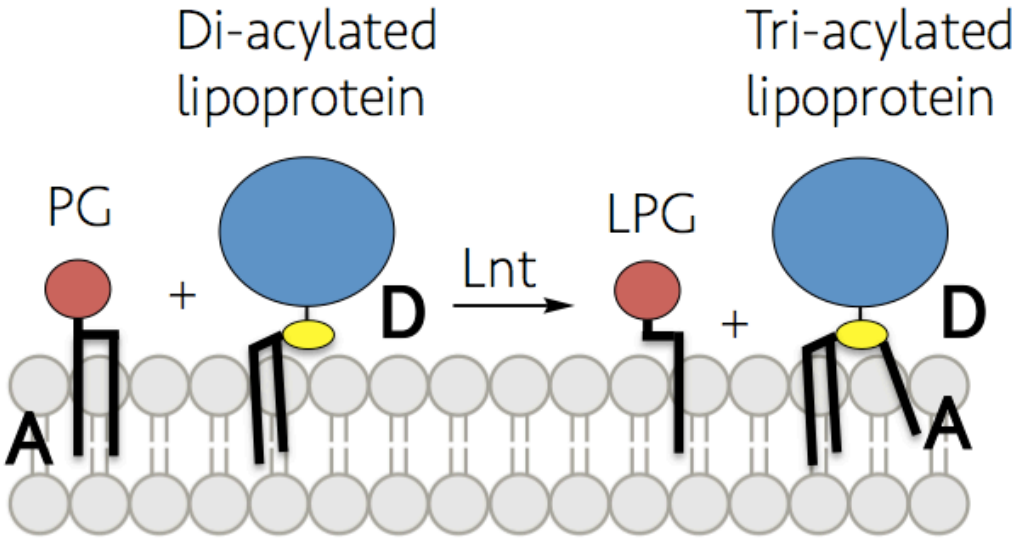


Figure 2.2 Assay idea involving a PG phospholipid with a fluorescence quencher on the *sn*-1 chain of the phospholipid and a donor incorporated into the amino acid sequence of the di-acylated lipoprotein (D = donor, A = acceptor).

Here, the fluorophore has been covalently added into the amino acid sequence of the di-acylated lipoprotein. The di-acylated lipoprotein will be fluorescent and this fluorescence will be quenched when FRET occurs as the acceptor containing *sn*-1 chain is transferred to the di-acylated lipopeptide by Lnt. Here, the fluorescent donor and the chromophore would be slightly farther away than in assay system A. FRET assays are sensitive to the distance between the FRET pair, with higher efficiency resulting from a smaller distance between them. The efficiency of a FRET assay is defined by the following equation: $E = R_o^6/(R_o^6 + r^6)$ where R_o is the Förster distance where half the energy is transferred and r is the actual distance between the FRET pair. FRET assays work best when the distance between the FRET pair is 20-90 Å.[49]

C:

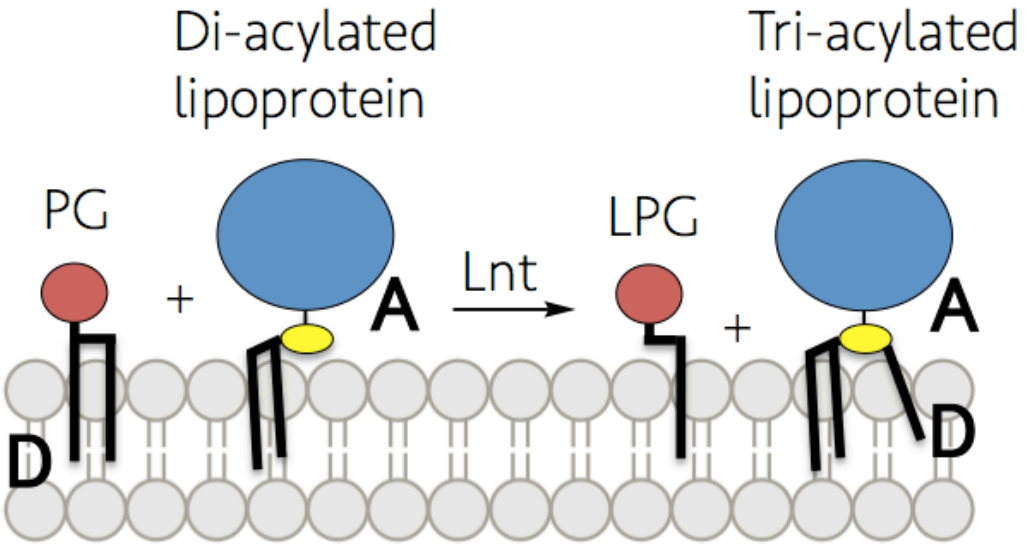


Figure 2.3 Assay idea involving a PG phospholipid with a fluorescent donor on the *sn*-1 chain of the phospholipid and acceptor incorporated in the amino acid sequence in the di-acylated lipoprotein (D = donor, A = acceptor).

This scenario is the converse of B. Fluorescence from the fluorophore will be quenched upon transfer of the *sn*-1 chain to the N-terminus of the diacylated lipoprotein that contains the chromophore incorporated into the amino acid sequence.

The assay system that was preferred was B, because the chromophore that was chosen for this FRET assay (dinitrophenyl, DNP, Fig. 2.4) could be added to a fatty acid of a PG phospholipid easily through chemical synthesis and the fluorescence donor

(aminobenzoic acid, Abz, Fig. 2.4) contains a carboxylic acid, allowing it to be easily incorporated into an amino acid sequence through peptide synthesis.

aminobenzoic acid (Abz) (fluorophore)

dinitrophenyl (DNP) (chromophore)

Figure 2.4 The DNP chromophore is in blue and forms part of a FRET pair with the aminobenzoic acid (Abz, shown as an aminobenzamide after incorporation via peptide synthesis) fluorophore (red).

These two compounds were chosen to comprise the FRET pair, as they are both small and neutral and, therefore, should not be too foreign to the active site of the enzymes. The use of Abz and DNP as fluorescent donor and acceptor has been reported in the literature.[50] In addition, the Abz fluorophore and the DNP chromophore have a good spectral overlap, *i.e.*, the region where DNP absorbs light overlaps well with the region where Abz fluoresces.

This proposed assay would be continuous and could potentially be used in a high-throughput system to identify new inhibitors of Lnt and, therefore, to identify new antibiotics. The assay will have three components: Lnt, the synthetic chromophore labelled phospholipid and a fluorophore labelled synthetic peptide. The fluorophore labelled synthetic peptide would be synthesised by Katie Bowen in the lab of Eoin Scanlan, School of Chemistry, TCD. The synthetic peptide is shown below in Fig. 2.5.

The short peptide (CSSK) would mimic a lipoprotein by containing a diacylglyceryl (DAG) moiety on its cysteine residue that is crucial in being recognised by Lnt as the lipobox cysteine. The Abz fluorophore (red) and is coupled to the side chain of a lysine residue. In the assay system B, Lnt would transfer the *sn*-1 fatty acid containing the chromophore label to the N-terminus of the DAGylated cysteine in the fluorescent peptide in Fig 2.5.

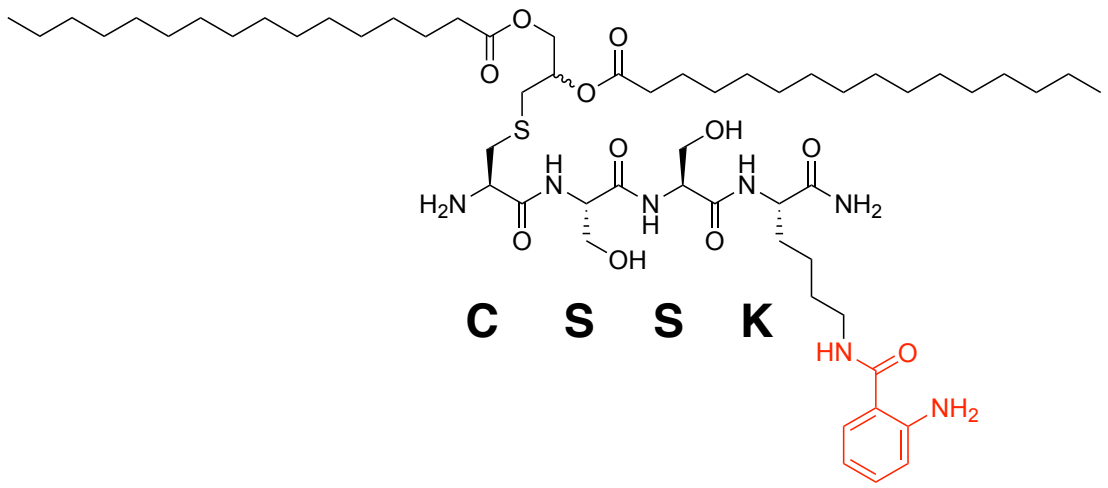


Figure 2.5 Structure of the fluorescent peptide to be used in the assay for Lnt. The Abz fluorophore is coloured red.

The synthetic chromophore labelled phospholipids in the asssay would be compounds **1-3**, below (Fig. 2.6), and were synthesised in this project.

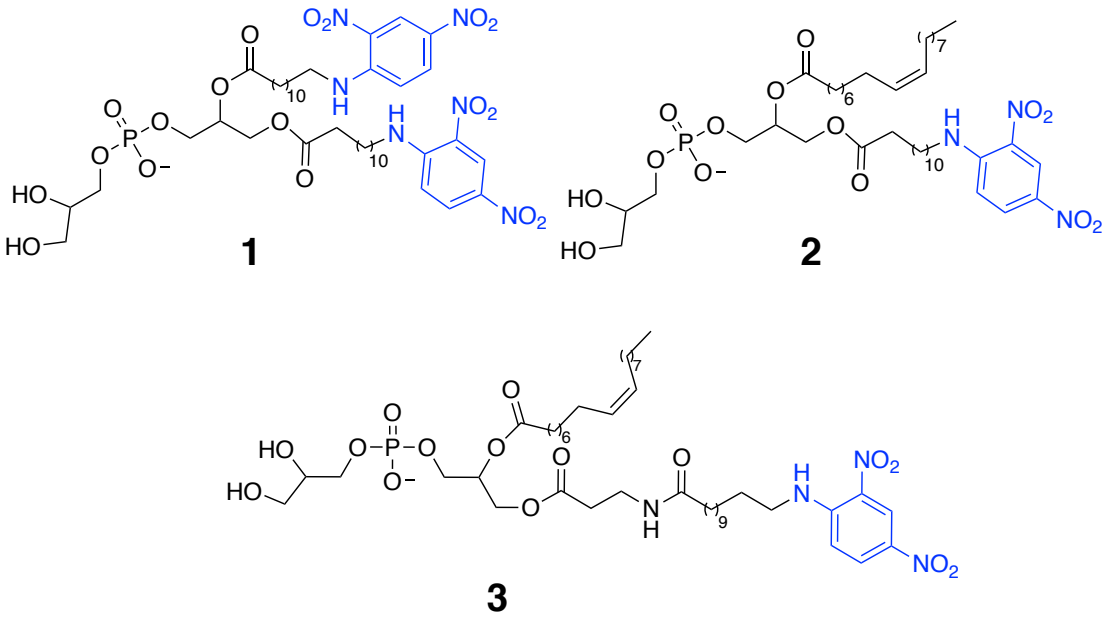


Figure 2.6 The structures of compounds **1-3**, DNP-chromophore labelled phosphatidylglycerol derviatives.

Glycerophospholipids (**1-3**) were synthesised in this project, as they are the source of the fatty acids that post-translationally modify lipoproteins in both Lgt and Lnt. As Lnt transfers only the *sn*-1 chain of the phospholipid, each of the synthetic phospholipids must contain the chromophore label at the *sn*-1 position in order to be substrates for Lnt. As Lgt does not discriminate between the *sn*-1 and *sn*-2 positions

and transfers the entire DAG moiety, compounds **1-3** can also be used in a functional assay for Lgt (below in Section 2.1.2). For synthetic ease, compound **1** contains the chromophore label at both the *sn*-1 and *sn*-2 positions.

2.1.2 Assay Ideas for Lgt

Lgt transfers the entire DAG moiety from a PG phospholipid to the preprolipoprotein, so the assay design is a little more straightforward than Lnt. The assay for Lgt that was considered is characterised in Fig. 2.7.

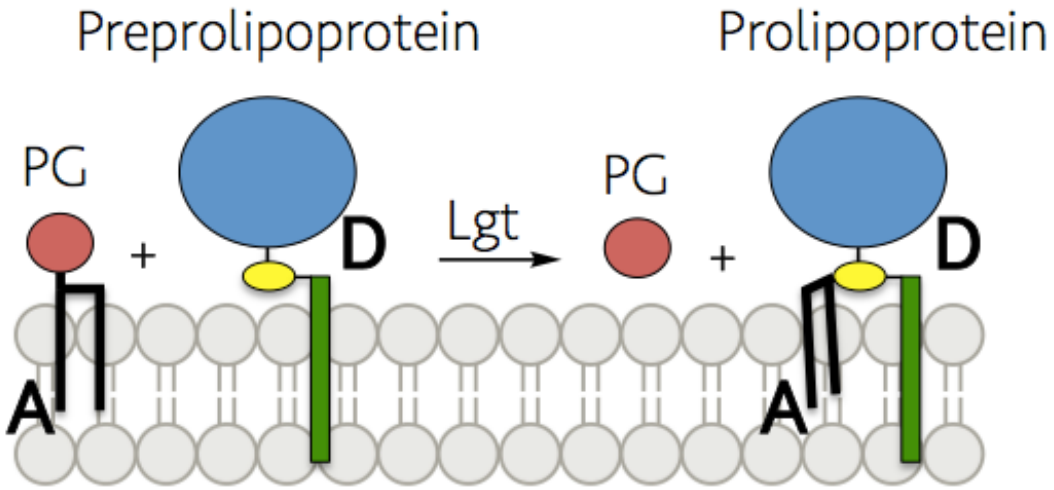


Figure 2.7 An idea for a continuous assay for Lgt that has been considered for this project. The PG *sn*-1 chain is a fatty acid containing the DNP chromophore and the Abz fluorophore is incorporated into the amino acid sequence (blue) (D = donor, A = acceptor).

Here, any specificity in relation to the *sn*-1 and *sn*-2 positions does not have to be considered. The entire diacylglyceryl moiety containing the quencher will be transferred to the lipobox cysteine (yellow), resulting in FRET and loss of fluorescence from the donor, Abz.

Lgt, a DNP labelled phospholipid (such as compounds **1-3**) and a fluorescent peptide, such as in Fig. 2.8, would make up the assay system.

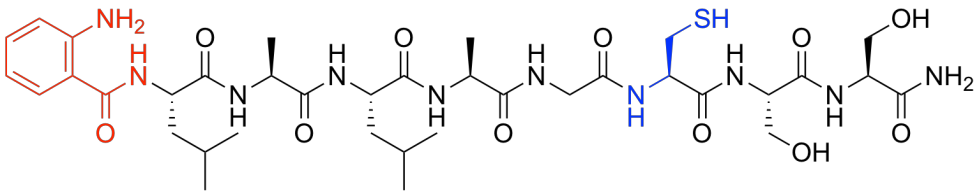


Figure 2.8 The structure of an Abz labelled peptide that could act as a substrate for Lgt. The synthetic peptide would contain an Abz fluorophore (red).

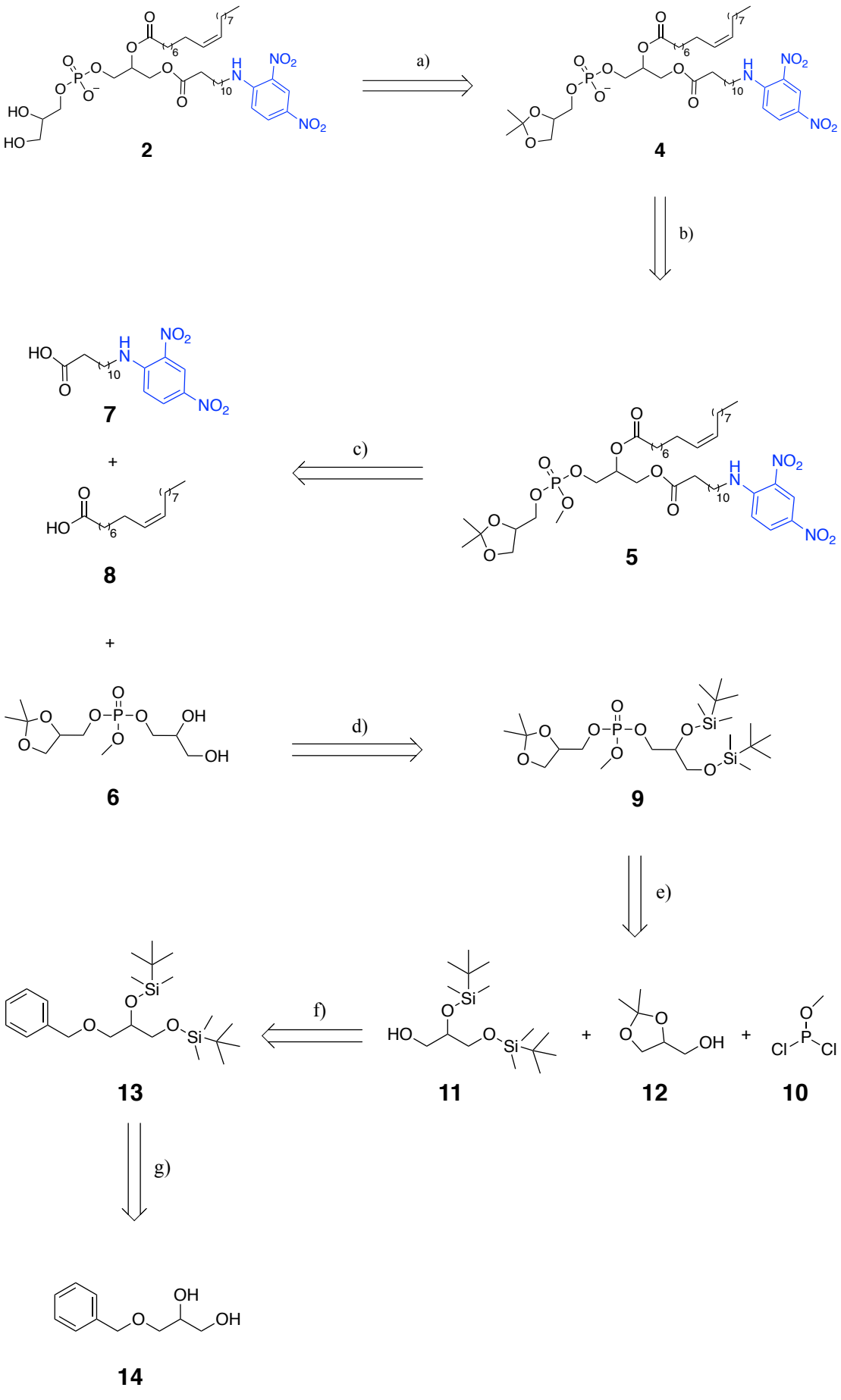
The Abz labelled peptide in Fig. 2.8 contains a LAGC sequence that mimics the lipobox of a preprolipoprotein and will be recognised by Lgt. Lgt will transfer to the diacylglyeryl moiety from the DNP labelled phospholipid (**1**-**3**) to the free thiol of the lipobox cysteine (blue).

For both the Lgt and Lnt FRET assays, the experiment will rely on observing the loss of fluorescence. At t_0, the Abz fluorophore will be fluorescent and only upon transfer of the DNP chromophore by Lgt/Lnt; a quenching of fluorescence will take place. This will make interpretation of the assay difficult due to a high signal to noise ratio. The assay would be more straightforward if at t_0 there was a baseline of no fluorescence, and upon the action of Lgt/Lnt fluorescence were to occur, but the Lnt and Lgt systems are quite complex.

2.1.3 Rationale for the synthesis of compounds 1-3

As there exists a number of families of glycerophospholipids, a choice had to be made in relation to what headgroup to install on the synthetic phospholipid. The choice was made with respect to evidence in the literature that showed Lgt has a preference for PG as substrate and while Lnt preferred PE, crucially, it still showed good activity for PG.[21] Therefore, a synthetic PG would have the ability to act as a substrate for both Lgt and Lnt. Retrosynthetic analyses for compounds **2**, **1** and **3** are presented (Schemes 2.1, 2.2 and 2.3).

Scheme 2.1 Retrosynthetic analysis of **2**


a)
b)
c)
d)
e)
f)
g)
2
4
5
6
7
8
9
10
11
12
13
14

The retrosynthetic analysis in Scheme 2.1 shows how the synthetic route to **2** was devised. The synthetic routes to **1** and **3** share many of the same steps. The final product, **2**, contains two free hydroxyl groups and these must be protected throughout the synthesis, as both are nucleophilic and will cause side product formation. The diol was protected as an acetal, as the acetal containing solketal, **12**, is commercial and this particular isopropylidene acetal has been successfully deprotected in the literature in the presence of a glycerophospholipid using trifluoracetic acid (TFA) and MeOH in CH_2Cl_2.[51]

The next step, b), in the retrosynthesis is the O-demethylation at the phosphate head group to give **5**. The first idea considered for the protection of this phosphate anion was a benzyl ether-protecting group, as hydrogenolysis under mild conditions would yield the anion. But due to the presence of the double bond in the oleyl fatty acid at the *sn*-2 position of **5**, hydrogenation of this alkene would have certainly taken place to give the saturated fatty acid. In addition, reduction of the nitro groups on the DNP chromophore would have occurred. To avoid the unwanted hydrogenation of the alkene and nitro group reductions, methyl deprotection of the oxygen was chosen. O-demethylation of a phospholipid was demonstrated in the literature using NaI in 2-butanone.[51,52]

The retrosynthesic step c) involves the removal of the fatty acid containing the DNP chromophore, **7**, from the *sn*-1 position and the oleic acid chain, **8**, from the *sn*-2 position to give **6** containing a diol. The aim of the synthesis was to have the fatty acid containing the DNP chromophore at the *sn*-1 position and a natural fatty acid at the *sn*-2 position. The literature suggested that a diol such as **6** could be acylated selectively in a sequential manner.[53] This sequential acylation was useful for this project, as it would allow the synthesis of mixed chained phospholipids from the diol **6**.

The diol **6** could be synthesised from silyl protected **9** in step d). Due to the presence of the isopropylidene acetal group in **9**, acidic conditions could not be employed to deprotect the silyl groups.

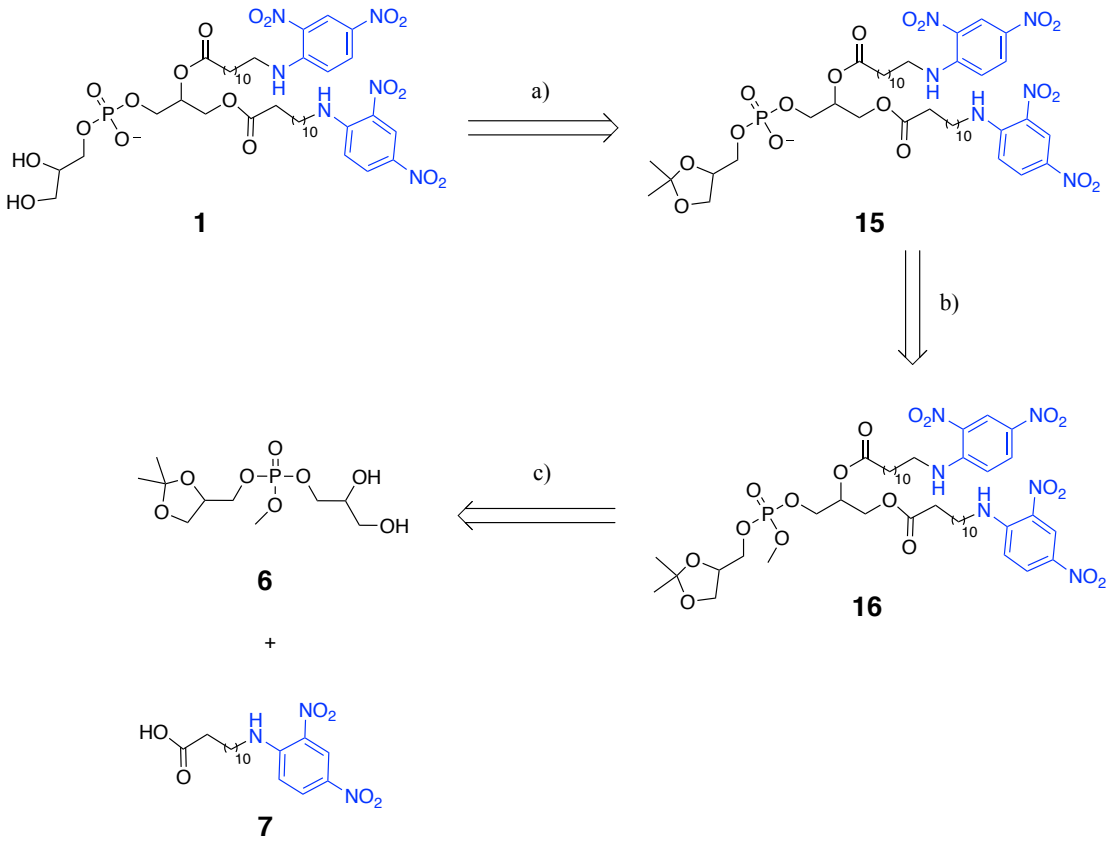
Compound **9** was synthesised by the 'phosphate installation' reaction, e), between

methyl dichlorophosphite, **10**, the silyl-protected glycerol, **11**, and solketal, **12**. The phosphite could be then oxidised to the phosphate using *meta*-chloroperoxybenzoic acid (mCPBA).

The next retrosynthetic step, f), is from **11** to **13**, where the benzyl ether of **13** is removed by hydrogenolysis. The benzyl-ether protecting group was chosen as it is orthogonal to silyl protecting groups, *i.e.*, the benzyl-ether protecting group can be removed under conditions that do not affect the *tert*-butyldimethylsilyl (TBDMS) groups.

The final retrosynthetic step is the silyl protection of the commercial **14** to form **13**.

Scheme 2.2 Retrosynthetic analysis of **1**

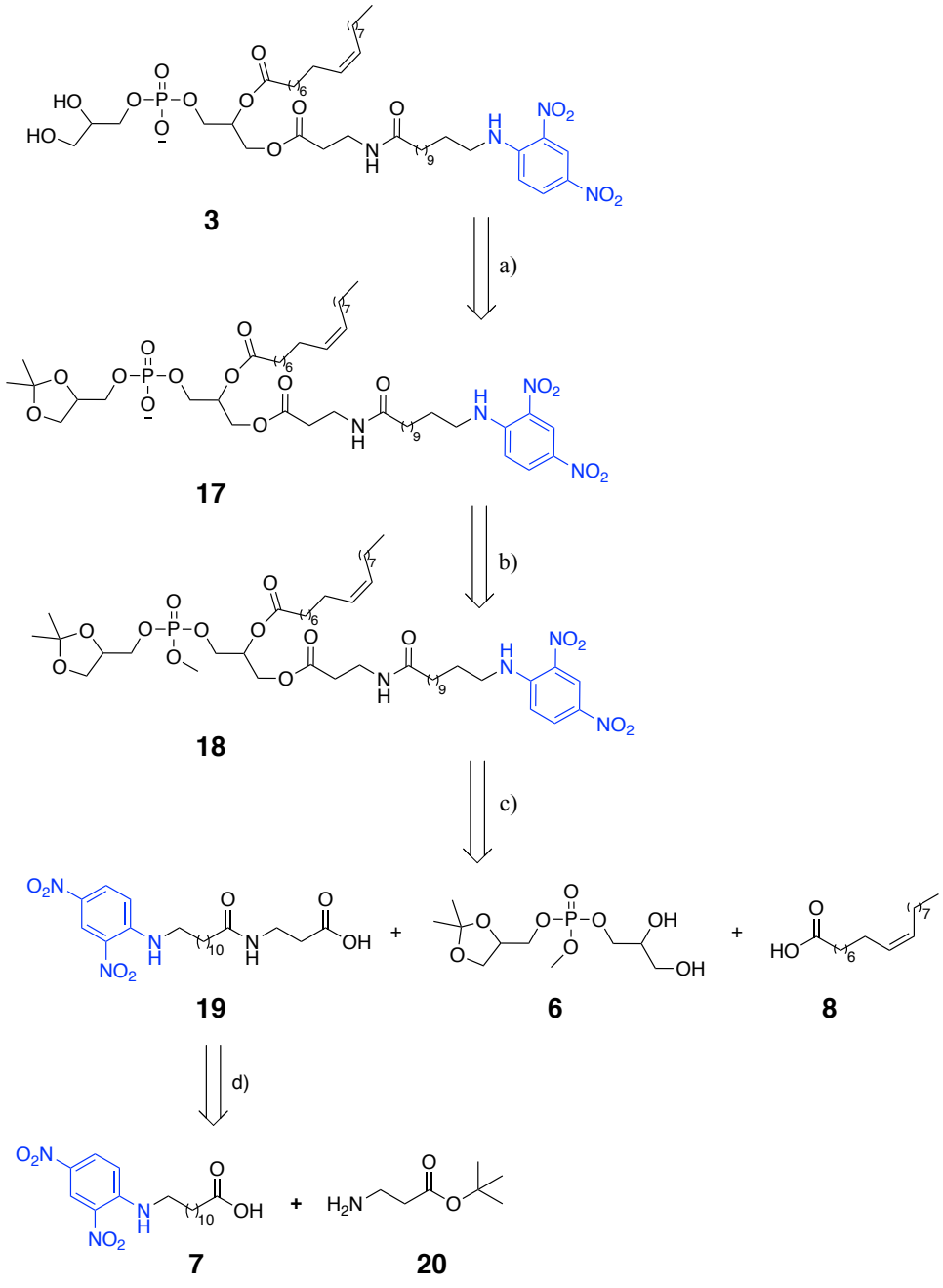


The first two retrosynthetic transformations for **1** are the same as for **2**, acetal deprotection, a), using TFA and MeOH in CH_2Cl_2[51] and O-demethylation, b), using NaI in 2-butanone.[51,52] Compound **1** contains the chromophore label at both its *sn*-1 and *sn*-2 positions. The reason for this is due to the synthetic ease of coupling the same fatty acid, **7**, to both positions rather than only the *sn*-1 position. Here, the

Steglich esterification does not require sequential acylation that is present in the retrosynthesis of **2**. The Steglich esterification is carried with diol **6** and 2 eq. of **7**.

However, while **1** was easier to synthesise than **2** or **3**, there were concerns over the suitability of **1** as a substrate for Lnt and Lgt. These concerns were due to the presence of two unnatural, bulky fatty acids that may negatively affect the enzymatic activity. For this reason, **2** and **3** contain a more natural fatty acid, oleic acid, **8**, at the *sn*-2 position.

Scheme 2.3 Retrosynthetic analysis of **3**

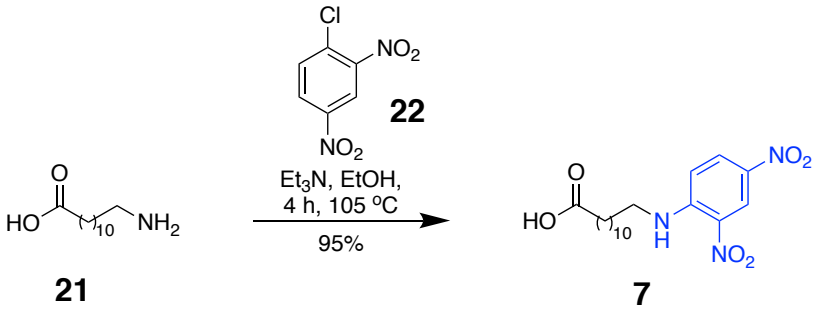


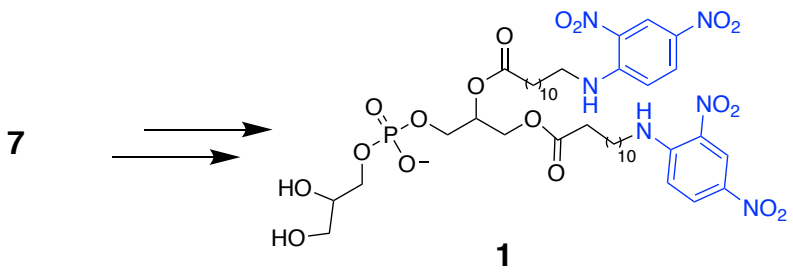
The first two retrosynthetic transformations for **3** are the same as for **1** and **2**, acetal deprotection, a), and O-demethylation, b). Compound **3** contains an additional β-alanine linker on the *sn*-1 chain, when compared to **2**. This linker was incorporated to extend the chromophore farther out of the active site cleft of Lnt, hoping to minimise detrimental interactions between the DNP chromophore and Lnt. Therefore, the Steglich esterification will be different to that employed in the synthetic route to **1** and **2**. A sequential acylation is carried out, but the carboxylic acid used to couple to the *sn*-1 position is **19** and oleic acid, **8**, is coupled to the *sn*-2 position. **19** is synthesised by an amide coupling reaction between **7** and **20**, using the Mukaiyama reagent as coupling agent.

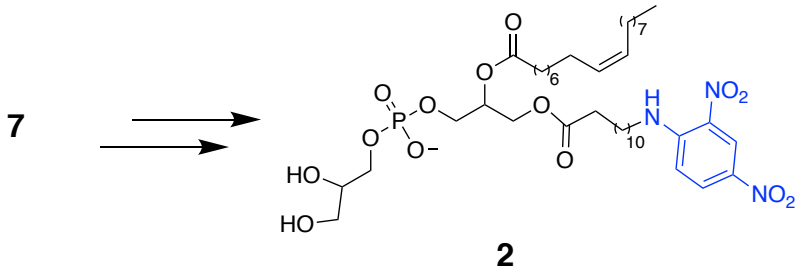
2.1.4 Objectives of this Project:

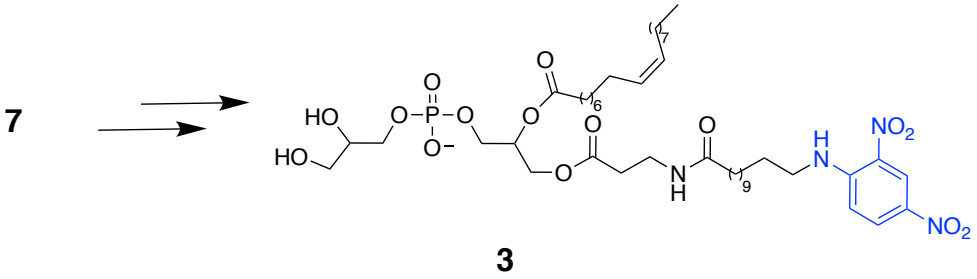
The objective of the phospholipid synthesis project is to synthesise a number of PG derivatives for use in the development of continuous FRET assays for the lipoprotein processing enzymes Lgt and Lnt. The PG derivatives contain a DNP chromophore that would act as a fluorescence acceptor. The DNP chromophore (blue) can be incorporated into a fatty acid by the reaction between 12-aminododecanoic acid, **21**, and 2,4-dinitrochlorobenzene, **22.**

Scheme 2.4 Incorporation of DNP chromophore









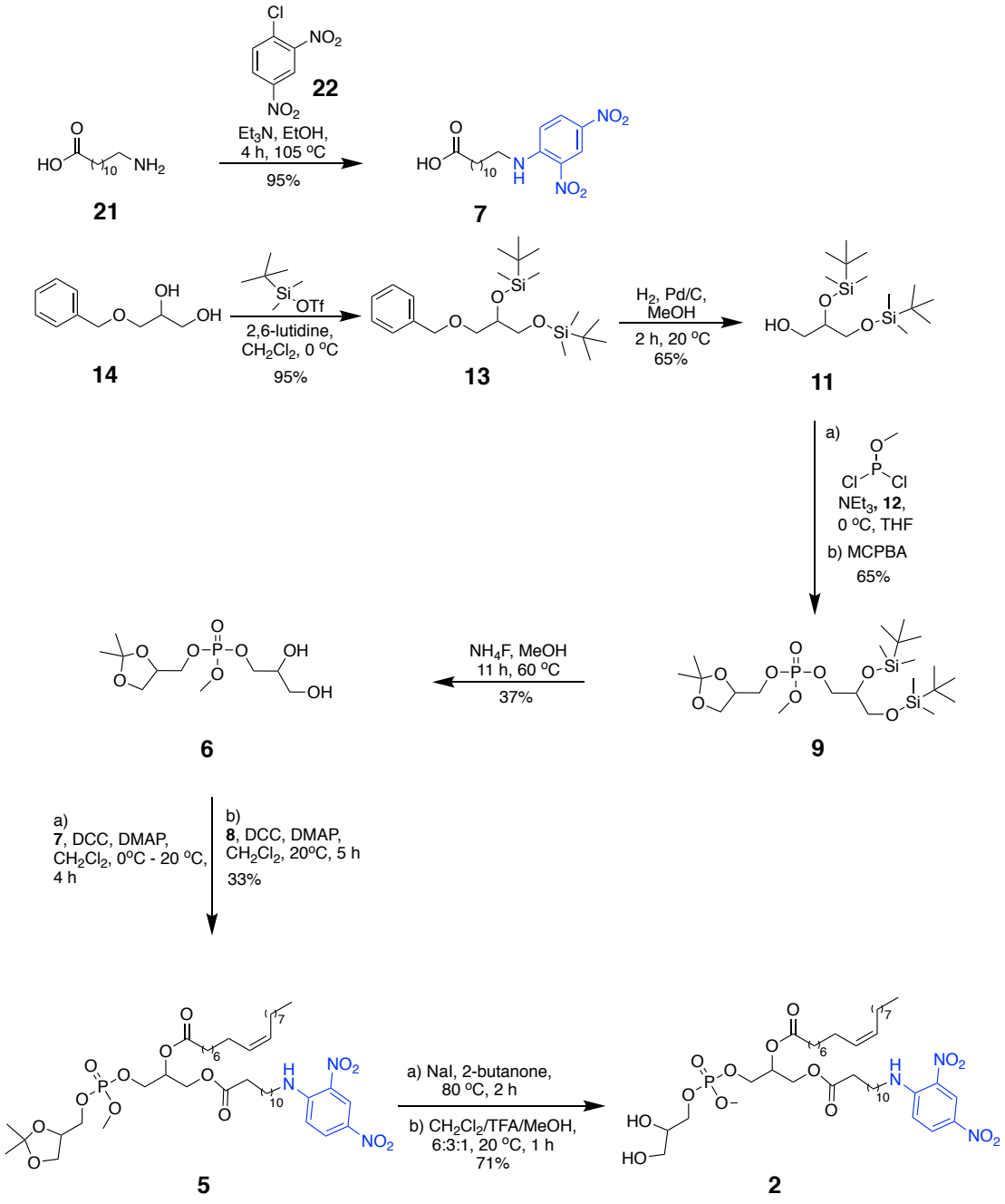
As mentioned previously, Lnt transfers only the *sn*-1 chain of a glycerophospholipid to its lipoprotein substrate. Therefore, this project will involve the synthesis of glycerophospholipids containing a chromophore label at its *sn*-1 position.

In addition, biochemical synthesis of chromophore labelled phospholipids using an immobilised lipozyme was to be explored. Enzymes offer a higher level of selectivity than traditional organic synthesis and usually require fewer steps. In order to take advantage of this selectivity, and to reduce the overall environmental impact of this synthesis, enzymatic synthesis of chromophore labelled phospholipids was explored using lipozyme RM IM.

2.2 Results and Discussion

An overview of the synthetic scheme for the preparation of **2** is shown in Scheme 2.5.

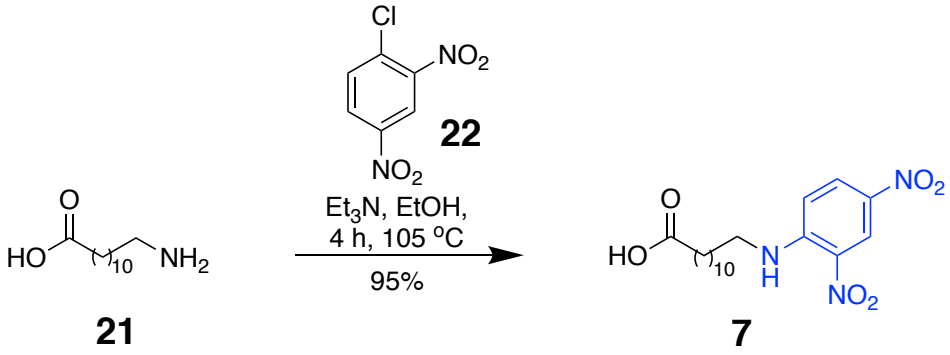
Scheme 2.5 Synthesis of chromophore labelled phospholipid **2**



In the first step of the scheme, the chromophore labelled fatty acid **7** is synthesised (Scheme 2.6). 12-aminododecanoic acid, **21**, and 1-chloro-2,4-dinitrobenzene, **22**, were heated to reflux in EtOH with triethylamine used as base for 4 h. The

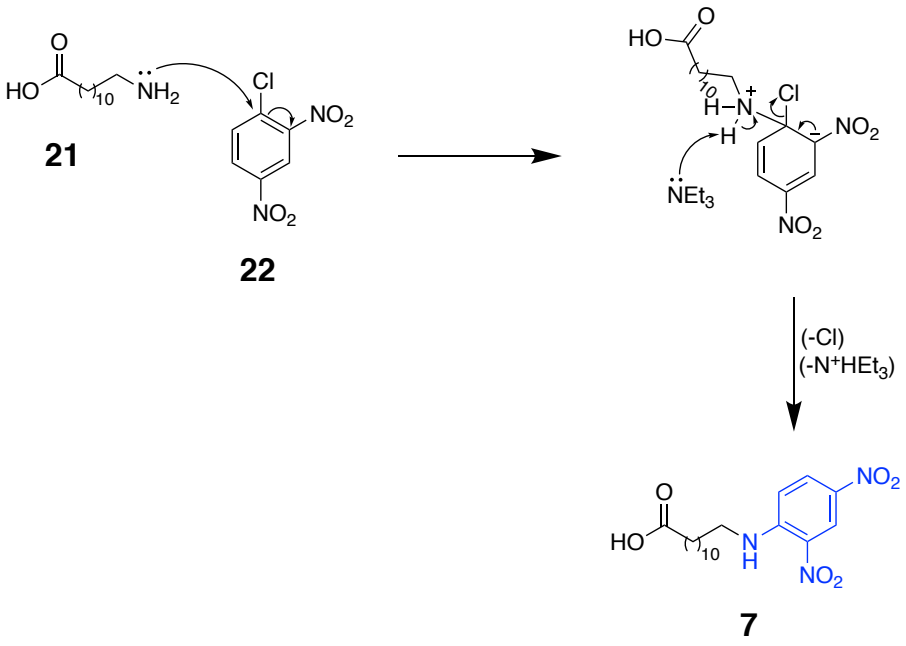
nucleophilic primary amine displaces the chlorine via an addition-elimination mechanism (Scheme 2.7). The deactivating nitro groups ortho and para to the chlorine promote this reaction by removing the electron density from the phenyl ring.

Scheme 2.6 Synthesis of fatty acid **7**



The chloride was easily displaced to give near quantitative conversion to the product. During the work-up of this reaction, it was necessary to dilute the reaction mixture with water and acidify the solution with 5 N HCl to pH 1-3, in order to re-protonate the carboxylic acid, otherwise, extraction into an organic solvent is ineffective.

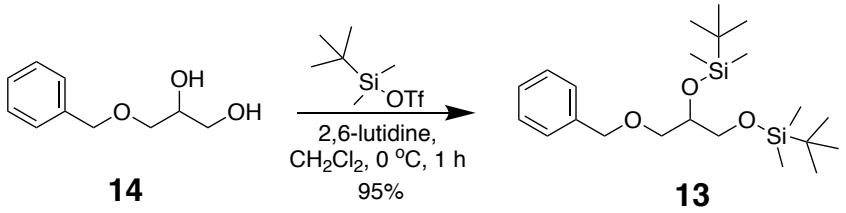
Scheme 2.7 Mechanism of the addition-elimination mechanism to synthesise **7**



In the protecting group strategy employed in this synthesis, a benzyl-ether protected

glycerol that is commercially available, **14**, is protected at both free hydroxyl groups with *tert*-butyldimethylsilyl trifluoromethanesulfonate (TBDMS-OTf, Scheme 2.8).

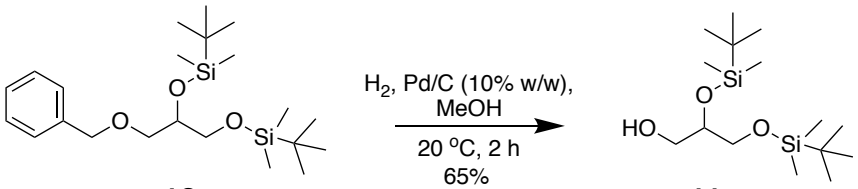
Scheme 2.8 Silyl protection of **14**



14 **13**

The silyl protection of hydroxyl groups is common in organic chemistry[54,55] and this reaction was carried out according to Bailey *et al.*,[55] using TBDMS-OTf and 2,6-lutidine as basic catalyst in anhydrous CH_2Cl_2 at 0 °C for 1 h. The reaction furnished the triply protected glycerol product **13** in a 95% yield.

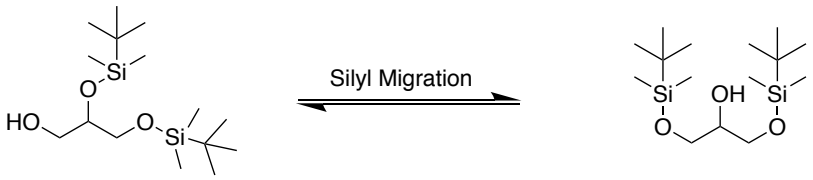
Scheme 2.9 Hydrogenolysis of **13**



13 **11**

The next step in the synthetic scheme is the removal of the benzyl ether protecting group using Pd/C mediated hydrogenolysis (Scheme 2.9). The reaction worked well in MeOH but showed no product formation in EtOAc after 2 h. This solvent specificity is not unusual for hydrogenation reactions, which tend to be quite sensitive to solvent effects. After 2 h in MeOH, complete conversion to **11** was observed. In removal of the benzyl-ether, a free hydroxyl has been formed and unfortunately, this hydroxyl group may now participate in silyl migration reactions to produce compounds **11** and **23** (Scheme 2.10). This research group has considerable interest in and experience with acyl migration in monoacylglycerol compounds and while this silyl migration was not anticipated, it did not come as a complete surprise.

Scheme 2.10 Silyl migration of **11**



11 **23**

Compound **23** was identified by its similarity to **11** on TLC (Fig. 2.9). Due to our experience dealing with migration in glycerol derivatives, efforts to reduce the migration were eventually enforced, such as reducing reaction time and work-up time, reducing the temperature of the water bath during solvent removal and storage of the product at -80 ^{o}C and in the dark.

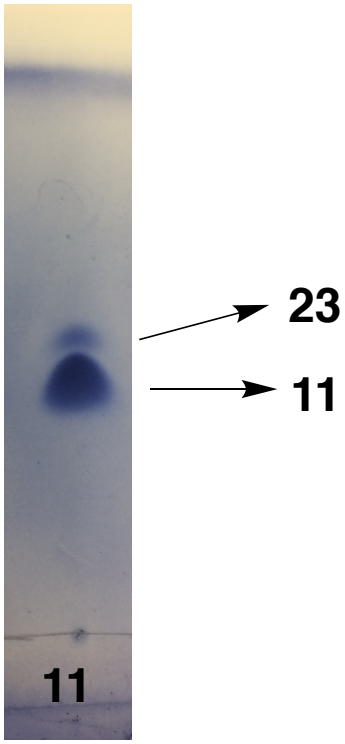
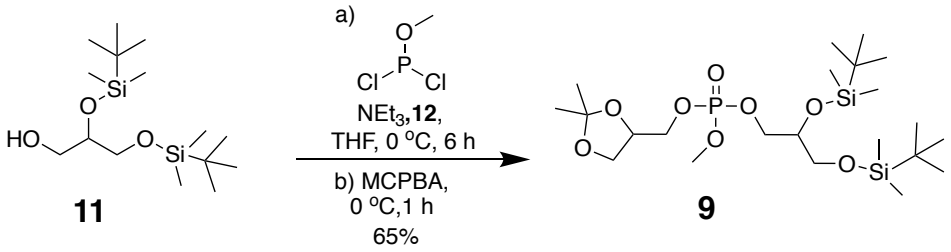


Figure 2.9 TLC of the reaction mixture of the benzyl-ether deprotection, showing a second sport that may be formed by silyl migration to **23**, run in hexanes/EtOAc, 19:1. The spot labelled **23** could be removed by flash chromatography over silica gel.

Next, isomerically pure compound **11** was coupled to the phosphate backbone and glycerol headgroup in a reaction termed the 'phosphate installation reaction' (Scheme

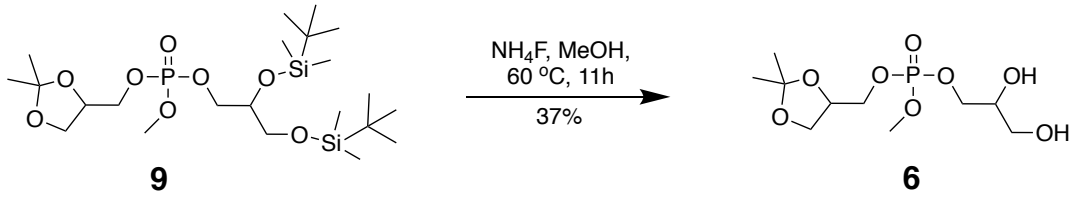
2.11). This reaction was carried out according to a literature protocol[56] and used methyl dichlorophosphite, **10**, as the phosphorylating reagent, triethylamine as basic catalyst and solketal, **12**, as a protected glycerol coupling partner. **12** contains an acetal group protecting the 1- and 2-positions of glycerol. It has often been employed in lipid and phospholipid synthesis as a protecting group for glycerol and its removal in the presence of a phosphate has been described.[57] The reaction worked well on the first occasion (65% yield), but could not be reproduced. Each time, the reaction showed a different TLC profile and a drastically reduced yield of just 10%.

Scheme 2.11 Phosphate installation reaction to form **9**



After several repeat reactions, it was realised that the methyl dichlorophosphite had been hydrolysed by water vapour in air. Once a fresh reagent was used for each phosphate installation reaction, it became reliable and gave consistent yields of **12** (60-65%). Meta-chloroperoxybenzoic acid (mCPBA) was the oxidising agent of choice to yield the phosphate from the phosphite and although literature suggested the presence of unwanted side products, mCPBA worked well in our hands.[57]

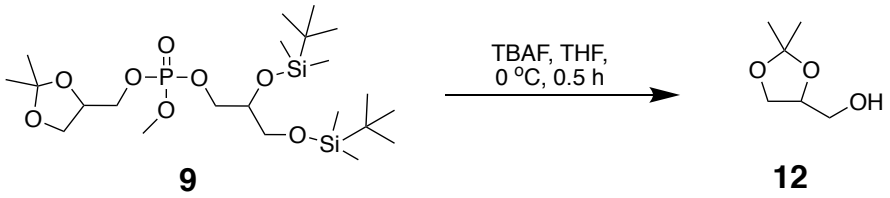
Scheme 2.12 Silyl deprotection of **9**



The next step of this scheme was the silyl-deprotection of **9** (Scheme 2.12). Due to the presence of the acetal protecting group in **9**, acidic conditions could not be used to selectively remove the TBDMS protecting groups. Fortunately, silicon has exceptional affinity for fluoride and this was exploited in our attempts to remove the two TBDMS groups. In the case of **9**, treatment with TBAF resulted in complete cleavage at the phosphate, yielding **12** as the main product of the reaction (Scheme

2.13).

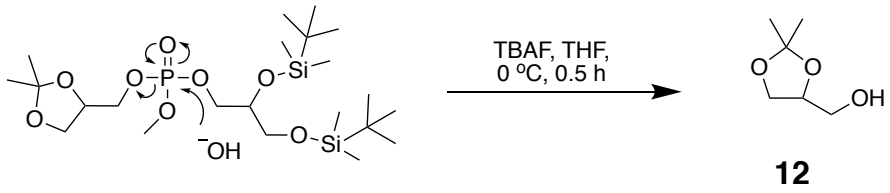
Scheme 2.13 Cleavage of **9** after treatment with TBAF



9 **12**

This cleavage is likely due to the basicity associated with the fluoride anion in TBAF[58] and the unavoidable presence of water that is associated with solutions of the hydrogroscopic TBAF in THF, causing the formation of hydroxide anion and complete cleavage of the phosphate, which is highly susceptible to base. The mechanism of the cleavage is shown in Scheme 2.14.

Scheme 2.14 Mechanism of the cleavage of **9**



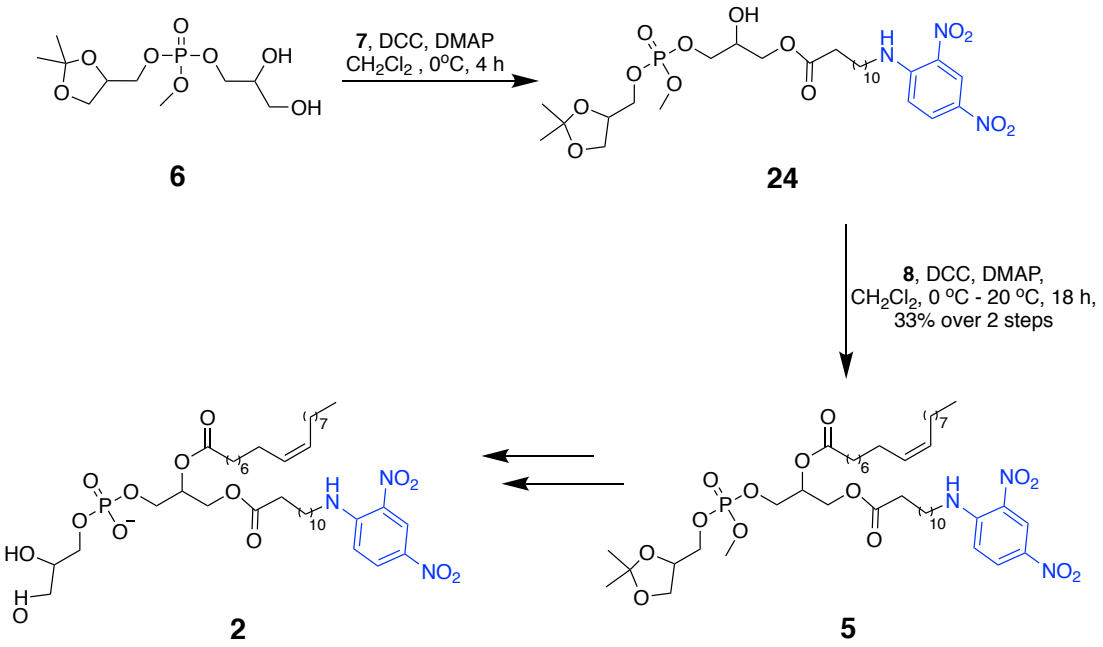
12

Due to the ineffectiveness of TBAF in this reaction, our attention turned to an inorganic source of fluoride, such as NH_4F. With NH_4F the basicity is reduced and the water content can be controlled. Small-scale test reactions (10 mg of **9**) with NH_4F in methanol at 60 °C were characterised by TLC Mass Spectrometry and showed good product formation with minor cleavage evident, resulting in the formation of **12**. The scale-up of the reaction showed the same profile and gave a yield of 37%.

The literature suggested that a primary and secondary diol could be acylated selectively in a sequential manner.[53] This selective acylation was interesting, as it would allow us to synthesise a phospholipid with different fatty acids at the *sn*-1 and *sn*-2 positions from diol **6**. The selectivity during the acylation arises from the control of temperature and the equivalents of carboxylic acids employed in the reaction. At 0 °C, only the primary alcohol would react with the coupling agent and 1 eq. of carboxylic acid. After primary alcohol acylation, the secondary alcohol can then be reacted with the coupling agent and another carboxylic acid at 20 °C, as illustrated in

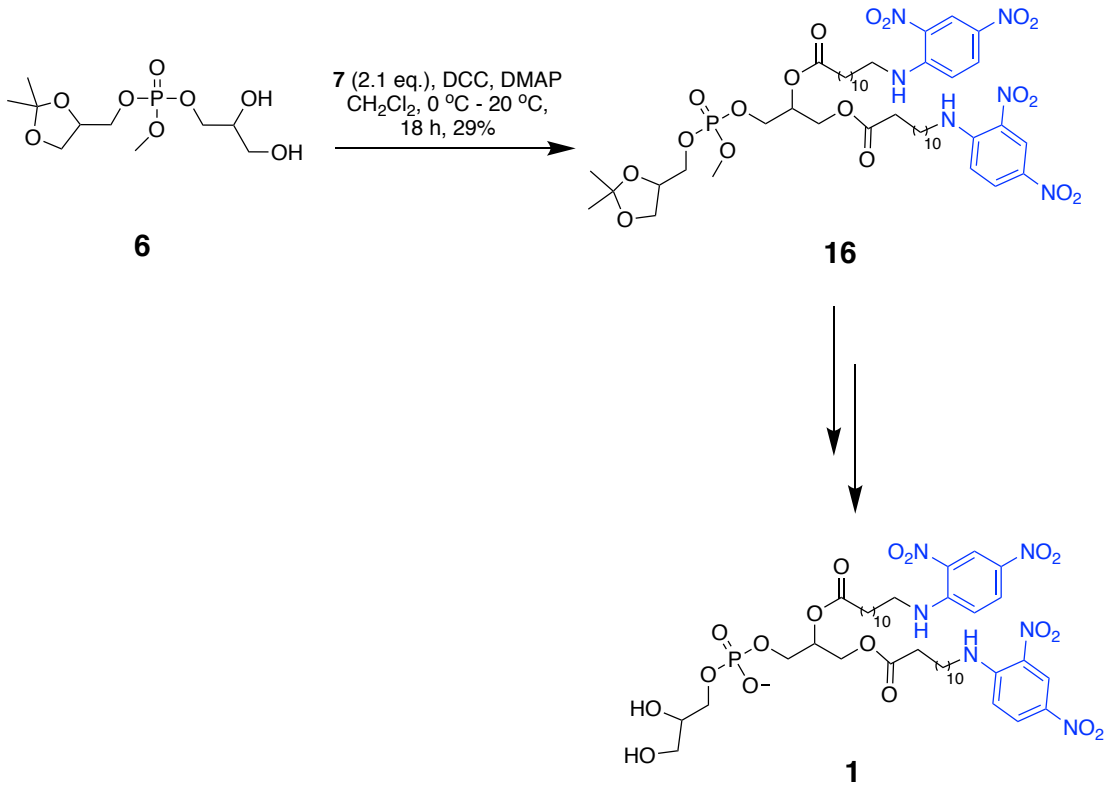
Scheme 2.15.

Scheme 2.15 Sequential acylation of **6**



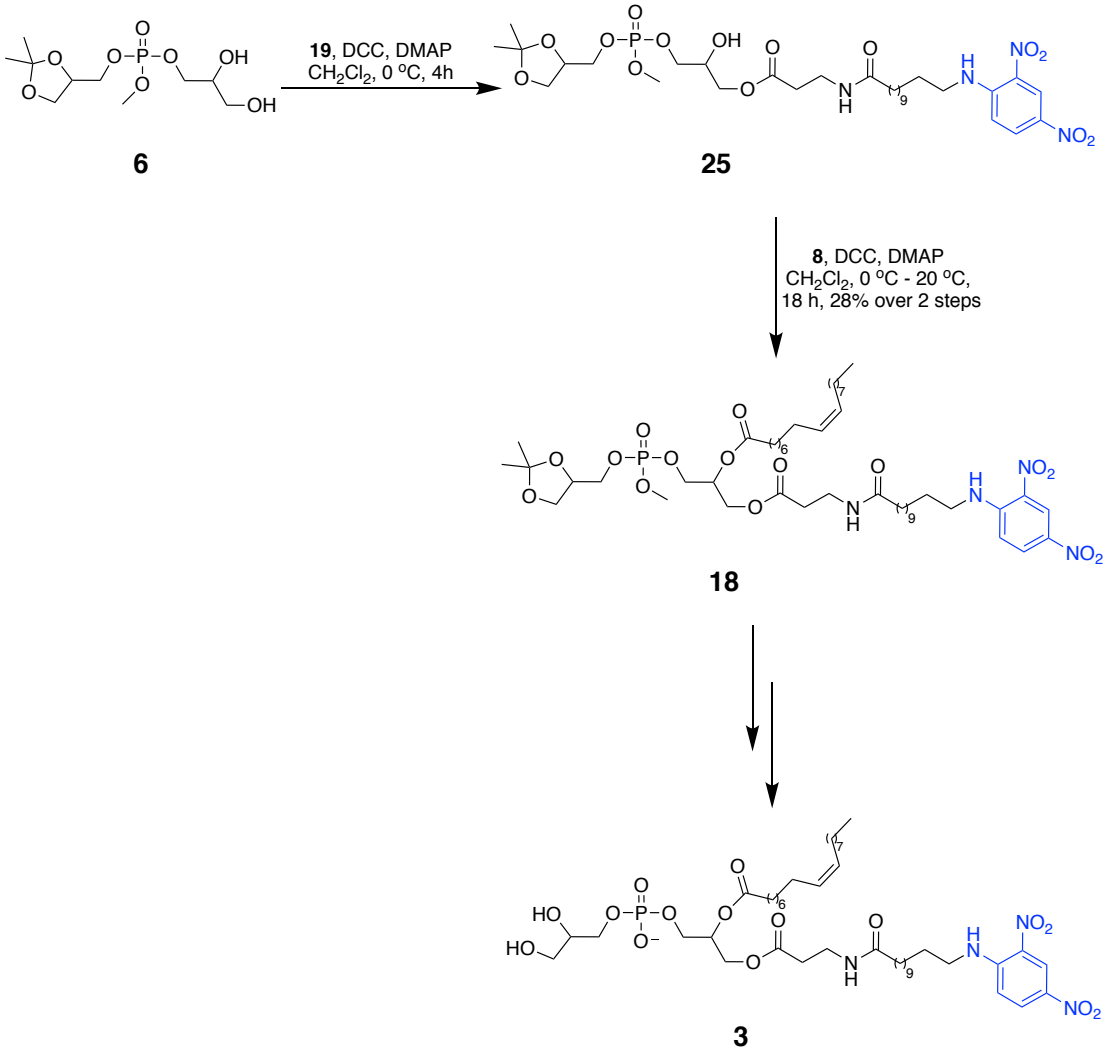
First, at 0 °C, **6**, the coupling reagent (DCC), the catalyst (DMAP) and the fatty acid chain we planned to couple to the primary alcohol position, **7** were added to anhydrous CH_2Cl_2. According to the literature, the primary alcohol should be acylated at 0 °C, then further coupling agent, catalyst, and a different fatty acid, **8**, could be added and would couple to the secondary alcohol at 20 °C. This sequential acylation worked well to give the desired phospholipid **5** (33%). A common side product of Steglich esterifications is dicyclohexylurea, formed from DCC. The dicyclohexylurea content in **5** could be reduced following precipitation out of CH_2Cl_2 at 0 °C and completely removed following flash chromatographic purification after the next reaction has been carried out.

Scheme 2.16 Steglich esterification of **6** with **7**



Compound **1** contains the DNP chromophore label at both the *sn*-1 and *sn*-2 positions. The two DNP labels were incorporated by Steglich esterification (Scheme 2.16). Here, the sequential acylation was not required, 2.1 eq. of **7** was added to the diol **6**, DCC and DMAP in anhydrous CH_2Cl_2.

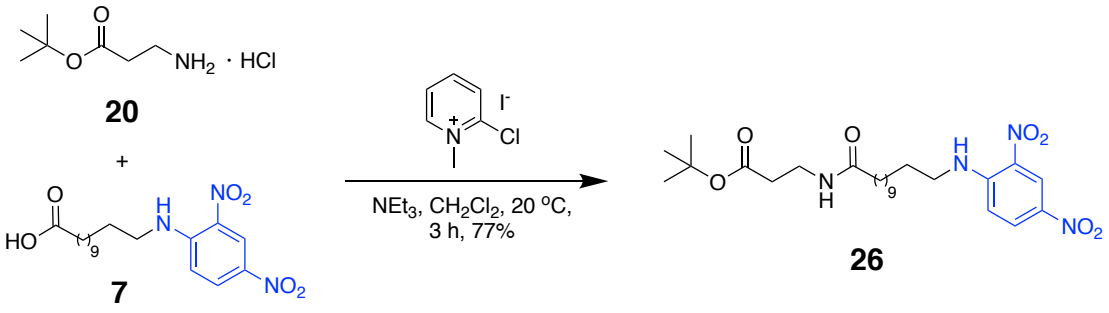
Scheme 2.17 Sequential acylation of **6**



As compound **3** contains an extended fatty acid at its *sn*-1 position, compound **6** is sequentially acylated with the fatty acid, **19**, and then oleic acid, **8**, to form **18** (Scheme 2.17).

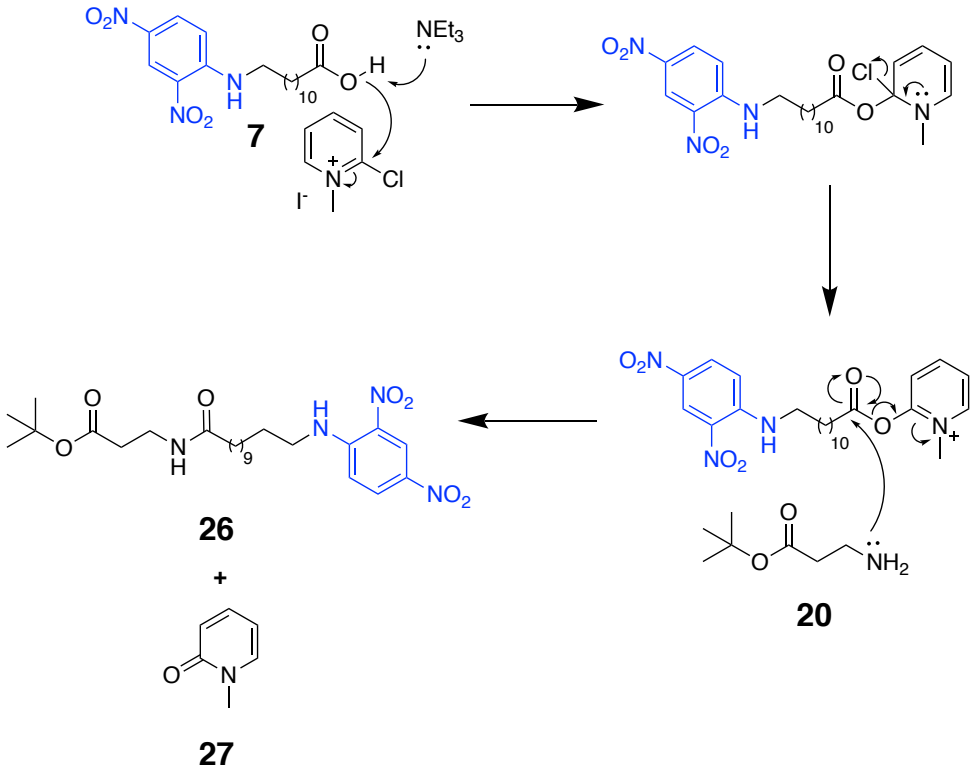
The extended fatty acid chain containing the DNP chromophore, **19**, is synthesised in two steps; first, an amide coupling reaction between **7** and **20** using the Mukaiyama reagent as coupling agent (Scheme 2.18) and second, the subsequent deprotection of the *tert*-butyl ester using TFA.

Scheme 2.18 Amide coupling reaction between **7** and **20**



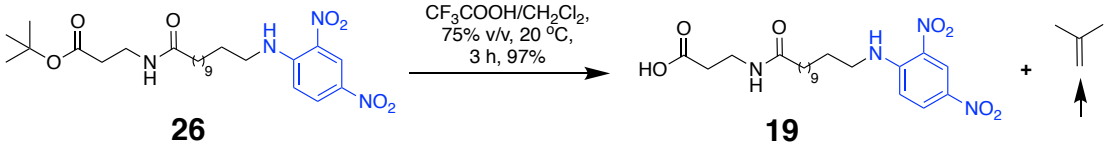
An amide coupling reaction between **20** and **7** allowed the formation of the amide, **26**. The amide coupling was achieved using the Mukaiyama reagent and triethylamine in anhydrous CH_2Cl_2 giving a yield of 77%. It was necessary to protect the carboxylic acid of **20** as an ester, to prevent selectivity issues if two unprotected carboxylic acids are present. The mechanism of the reaction is shown below (Scheme 2.19). The removal of the side product, **27**, is carried out during flash chromatography over silica gel, where the non-polar **27** elutes before the more polar, **26**.

Scheme 2.19 Mechanism of the amide coupling reaction between **7** and **20**



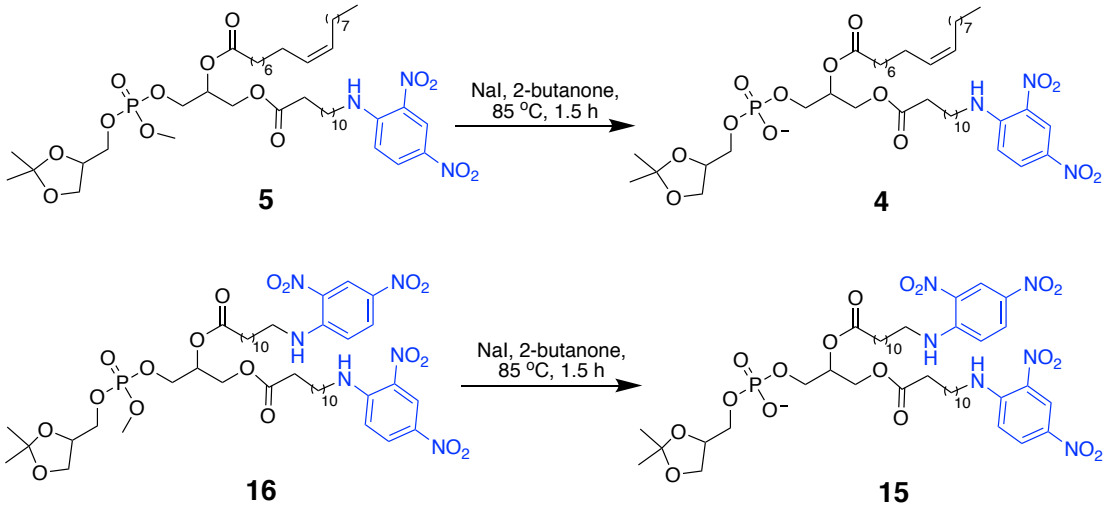
The deprotection of the *tert*-butyl group of **26** was achieved through exposure to TFA in anhydrous CH_2Cl_2 (Scheme 2.20). TFA has a pK_a of 0.5 and protonates the ester, leading to the loss of the stable *tert*-butyl cation that can then rearrange and leave the reaction mixture as isobutylene gas. The reaction gave the product, **19**, in 97% yield with chromatographic purification unnecessary. Remaining TFA in the reaction mixture proved troublesome, but was removed after several rounds of co-evaporation with toluene and overnight drying under vacuum.

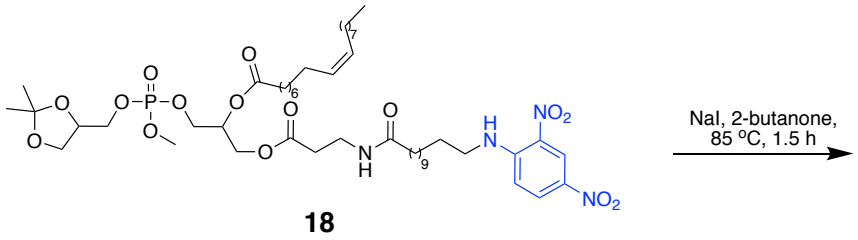
Scheme 2.20 Deprotection of the *tert*-butyl ester



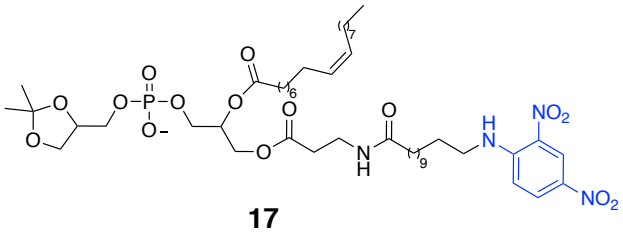
After the Steglich esterifications to synthesise **5**, **16** and **18** the next steps in each of the syntheses was to deprotect the phosphate anion that is protected as a methyl group and acetal deprotection. The O-demethylation of the methoxy groups is achieved by heating **5**, **16** and **18** in the presence of NaI in 2-butanone according to the literature (Scheme 2.21).[51] The reaction is fast (1.5 h) and there is no side product formation evident. The solvent, 2-butanone, is volatile and can be easily evaporated. Remaining NaI in the reaction mixture was removed by several rounds of washing with water and the acetal deprotection was carried out without further purification.

Scheme 2.21 O-demethylation of **5**, **16** and **18**





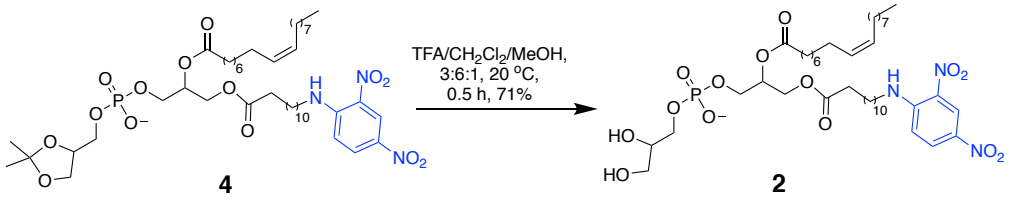
18



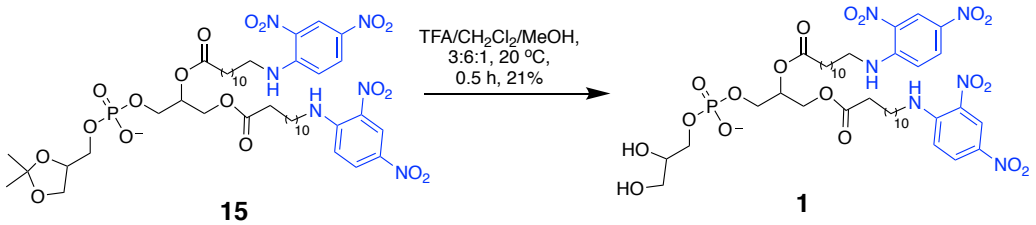
17

The final step in the syntheses of **1**, **2** and **3** is the removal of the acetal-protecting groups in **4**, **15** and **17** to yield the fully deprotected chromophore labelled-phospholipid (Scheme 2.22). Acetals are commonly deprotected under mild acidic conditions, such as AcOH in H_2O,[59] but the presence of water was not preferred for this particular reaction, due to the fear of cleavage of the phosphate.

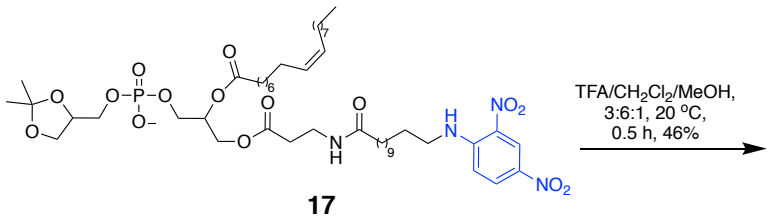
Scheme 2.22 Acetal deprotection of **4**, **15** and **17**



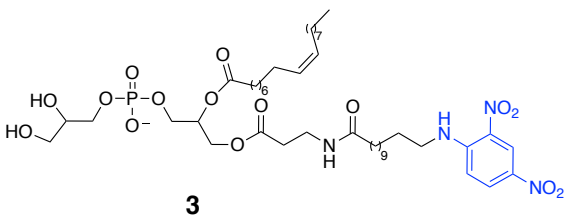
4 **2**



15 **1**



17



3

A procedure was found in the literature that utilised a CH_2Cl_2/TFA/MeOH (6:3:1) mixture for 0.5 h at 0 ^oC.[51] These conditions were applied to the crude products **4**, **15** and **17** and gave complete acetal deprotection after 0.5 h. The yields shown are over the two steps of O-demethylation and acetal deprotection. The lower yields seen for **1** and **3**, of 21% and 46% are most likely due to their high polarity and subsequent difficulty in elution off a normal phase silica column.

After flash chromatographic purification of **2** and **3** over silica gel was carried out, a minor spot on TLC co-eluted (Fig. 2.10 & Fig 2.11).

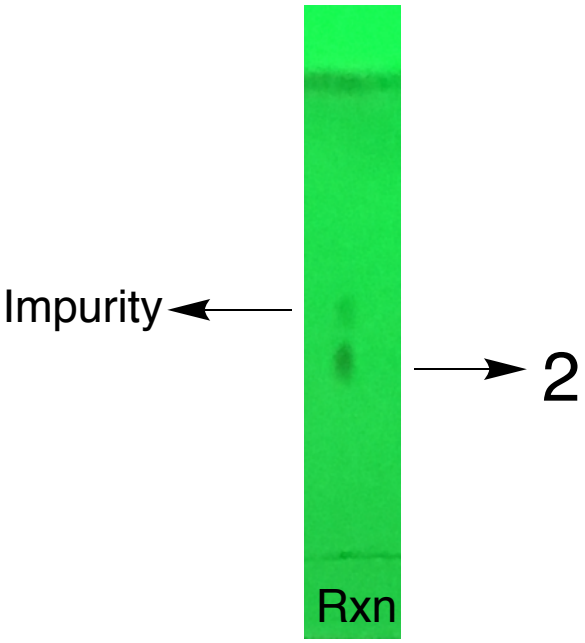


Figure 2.10 TLC after flash chromatography of **2** (Rxn), run in $CHCl_3$/MeOH/AcOH, 75:23:2.

A similar impurity could also be seen after flash chromatography of **3** (Fig 2.11).

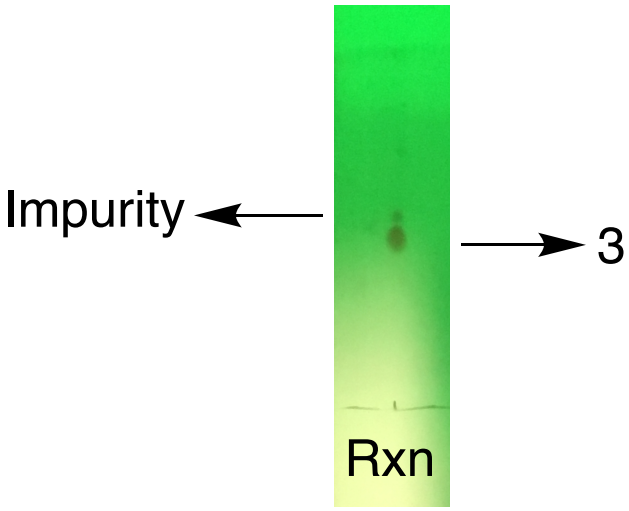


Figure 2.11 TLC after flash chromatography of **3** (Rxn), run in $CHCl_3/Ac_2O/MeOH/AcOH/H_2O$, 10:4:2:2:1.

Again, an impurity with a slightly higher R_f than the phospholipid product can be seen. The presence of an impurity is not visible in the ^1H NMR spectrum of **2** or **3**. A likely explanation for the presence of this spot is the presence of compounds **28** and **29** (Fig. 2.12).

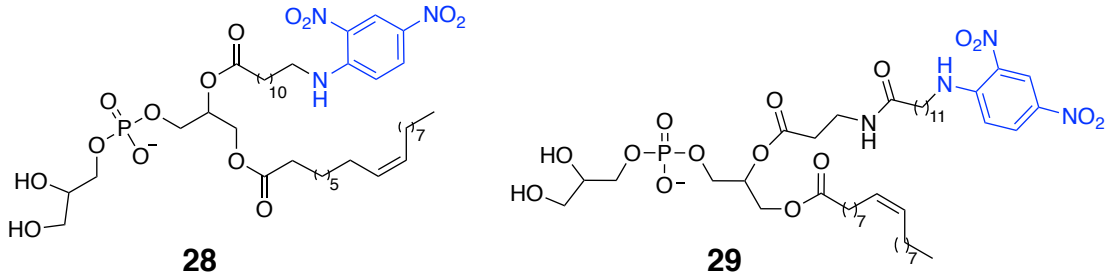


Figure 2.12 Structures of the possible impurities, **28** and **29**.

Compounds **28** and **29** are isomers of **2** and **3**, respectively. They both contain the DNP label at the *sn*-2 position instead of the intended *sn*-1 position and may have formed during the sequential acylations of **6**.

As flash chromatography was not successful in separating the two spots, preparative TLC was performed on **2** that contained two spots. The preparative TLC purification was successful in separating the two spots, but unfortunately introduced another impurity at the solvent front (Fig. 2.13)

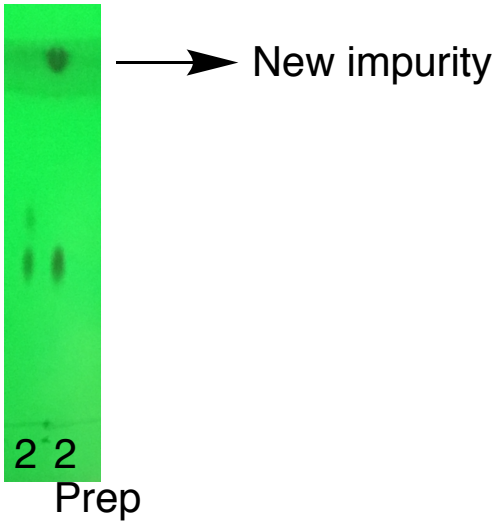


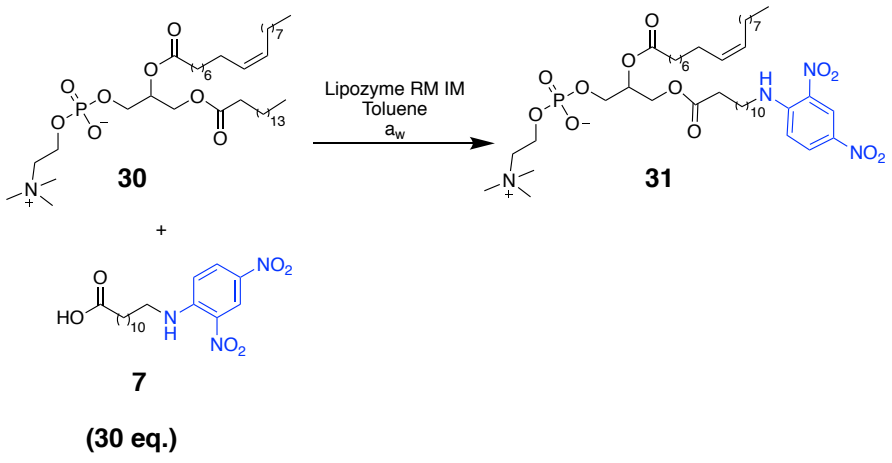
Figure 2.13 TLC of **2** after flash chromatography (left, **2**) and **2** after preparative TLC (2 Prep, right) run in $CHCl_3$/MeOH/AcOH, 75:23:2.

After preparative TLC purification of **2**, the minor impurity spot has been successfully removed, but a new impurity has been introduced at the solvent front. This new impurity is likely the organic binder that is present in the silica on the TLC plate. To avoid introducing this impurity, TLC plates with an inorganic binder were used to purify both **2** and **3** and the final, pure compounds were handed over for biological evaluation.

2.3 Enzymatic synthesis – preliminary work using immobilised lipozyme

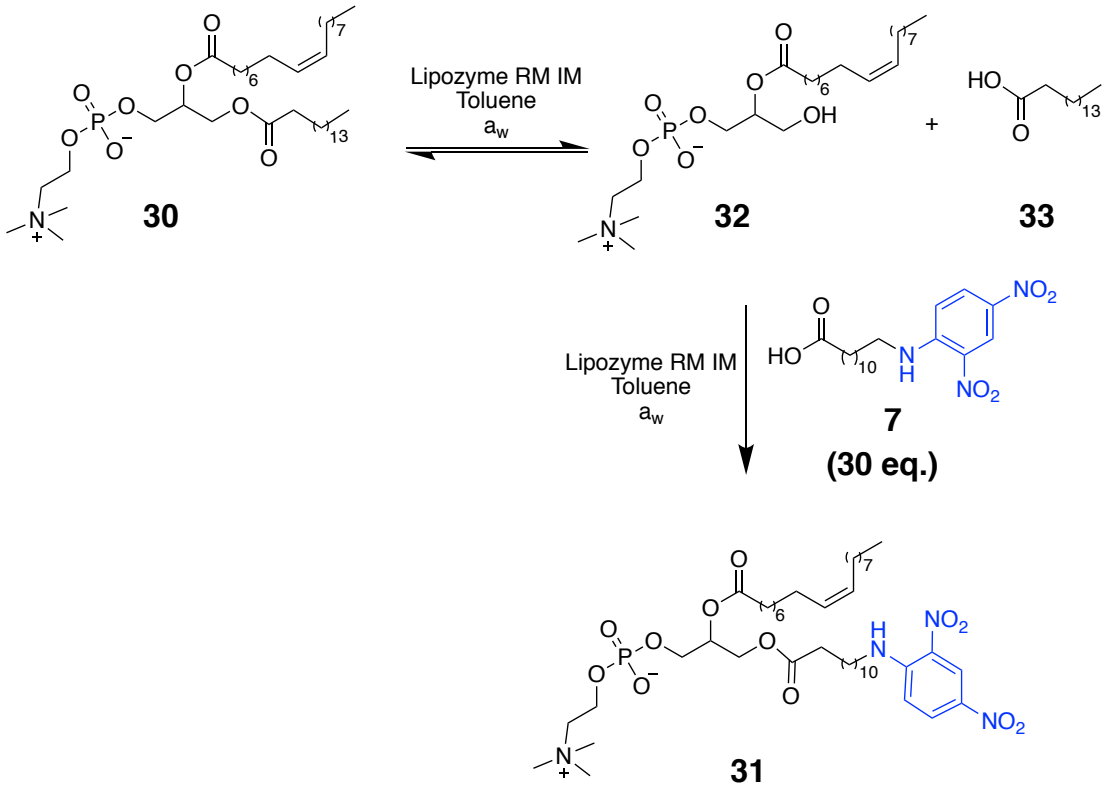
Enzymatic synthesis of DNP labelled phospholipids was explored and led to the synthesis of **31** below (Scheme 2.23).

Scheme 2.23 Enzymatic synthesis of **31**



The enzymatic reaction was carried out inside an open 20 mL vial inside a closed vessel containing a saturated solution of $MgCl_2$. The saturated $MgCl_2$ allowed the water activity (a_w) to be modified to a level of 0.33. As indicated above, the reaction mixture contained 30 eq. of **7**. This high excess was necessary to push the reaction equilibrium in the direction of the product, **31** (Scheme 2.24).

Scheme 2.24 Mechanism of the enzymatic reaction



The reaction proceeds through an intermediate, **32**, with a free hydroxyl group. At this stage of the reaction, the lipozyme RM IM may add the fatty acid, **7**, onto the free hydroxyl group of **32**, or the fatty acid, **33** (palmitic acid), that has just been removed. Due to the large excess of **7** that is present in the reaction mixture, the reaction proceeds in the direction of the product, **31**. The product, **31**, was found by HRMS (M^+ + Na, Fig. 2.14).

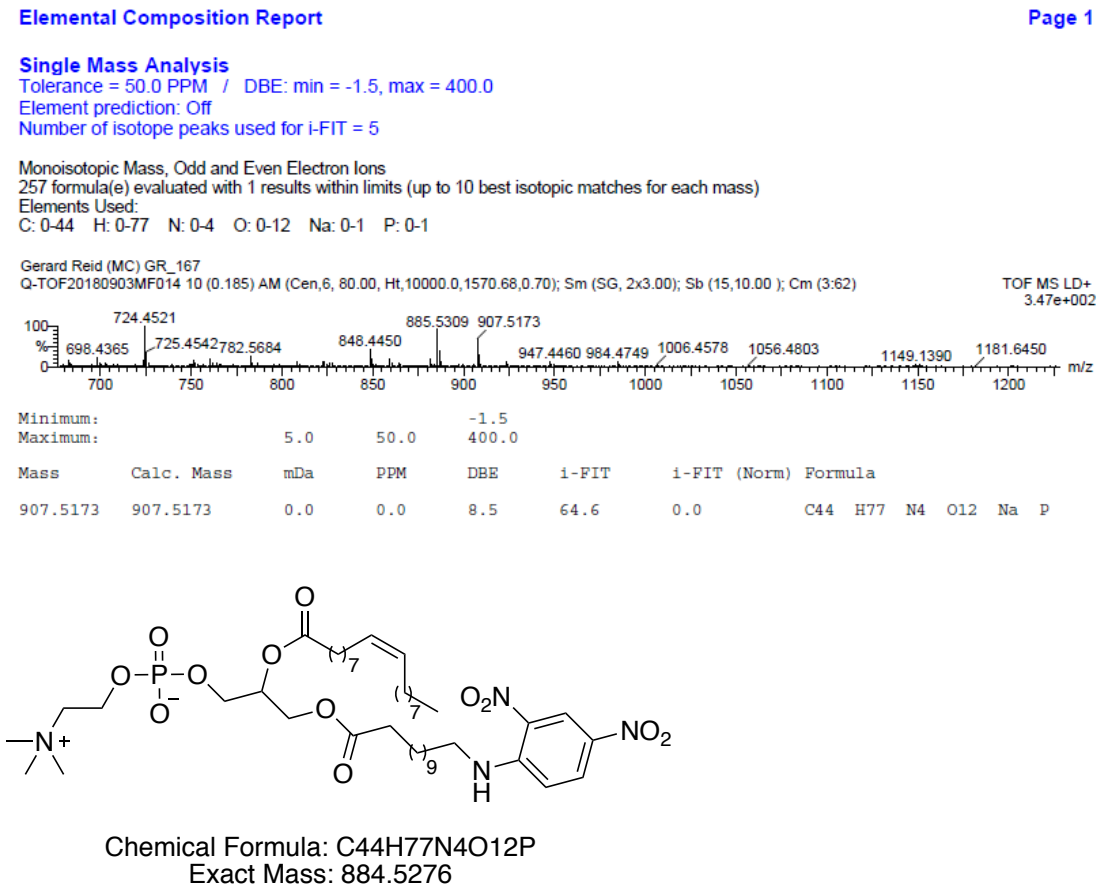


Figure 2.14 Result of the HRMS MALDI experiment of the product **31**.

Although the turnover to **31** was quite low, this result shows that the biochemical synthesis of chromophore labelled phospholipids for this project is possible and can be used in the future.

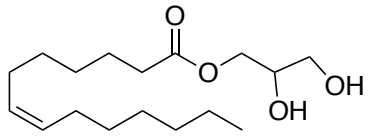
2.4 Conclusions and Future Work

The chromophore labelled phospholipids, **1-3** were synthesised using chemical synthesis by a newly developed synthetic scheme. Suspected silyl migration products during benzyl-ether deprotection of **13** could be removed by flash chromatography over silica gel. The cleavage of **9** during silyl deprotection using TBAF was avoided by employing NH_4F for silyl deprotection reactions. Additionally, a possible new route to chromophore labelled phospholipids via enzymatic synthesis has been developed. Compounds **1-3** were handed over for biological evaluation in facilitating structure and function studies on the bacterial enzymes Lgt and Lnt.

Chapter 3: *In meso* Method

3.1 Introduction

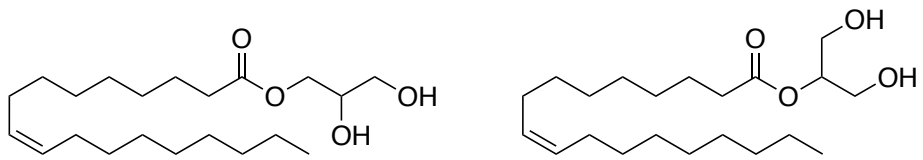
This chapter deals with the synthesis of monoacylglycerols (MAGs) for use in the *in meso* method of membrane protein crystallisation outlined in detail in Section 1.6. The synthesis of MAGs, such as 7.7 MAG (referred to as **34** from here, Fig. 3.1), was established in this research group.[41] The synthetic scheme developed was a modular synthesis that hinged on a Suzuki coupling reaction that crucially conserves the *cis* double bond of the MAG. The conservation of the *cis* double bond is essential, as MAGs can only enter the cubic phase, and therefore be used in the *in meso* method, if they contain a *cis* double bond. In the *in meso* method, the amphiphilic **34** is mixed with an aqueous solution containing a membrane protein of interest solubilised in detergent. If mixed in the correct quantities, a cubic phase is formed that acts as a conduit for the formation of membrane protein crystals.



34

Figure 3.1 The structure of **34**.

As MAGs such as **34** contain two free hydroxyl groups, **34** can spontaneously isomerise to its 2-MAG isomer. For this reason, MAGs that are employed in the *in meso* method can possibly contain some amount of its 2-isomer. Furthermore, once the MAG is mixed with the aqueous solution containing the membrane protein during the *in meso* method, it is stored (at 4 ^{o}C or 20 ^{o}C, depending on crystallisation experiement conditions) for a period of time (usually 1-30 days), it is likely that isomerism to its 2-isomer can occur during this period. The role of the 2-isomer, if any, in the *in meso* method is not known. For this reason, 2-monoolein was synthesised in this project (Fig. 3.2).



35 **36**

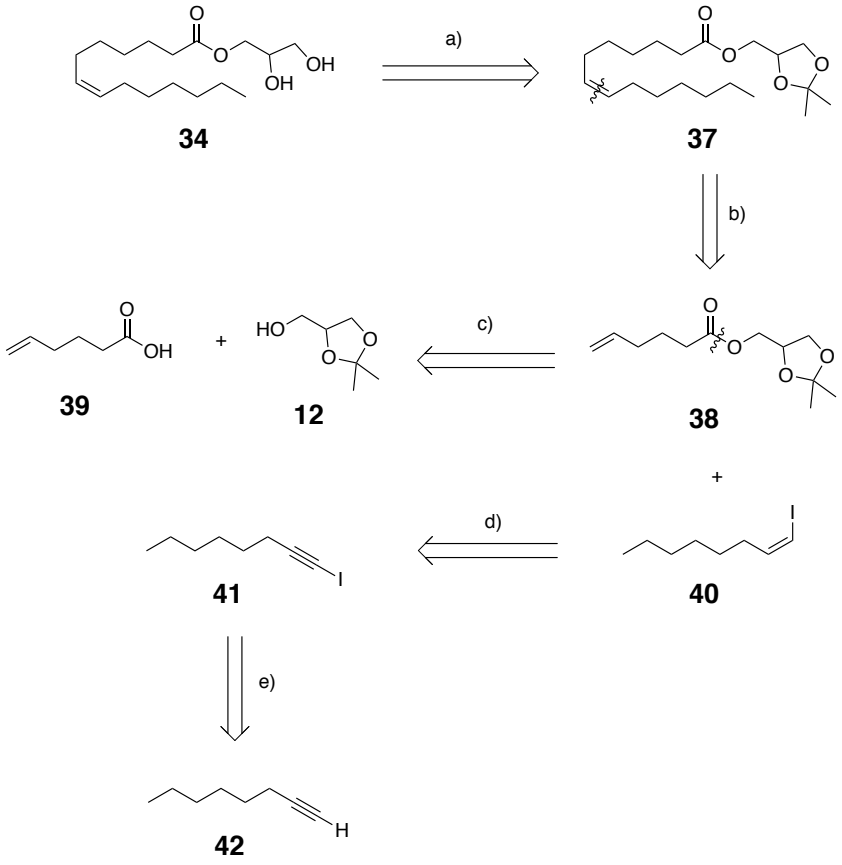
Figure 3.2 The structure of monoolein (left, **35)** and 2-monoolein (right, **36**) the 2-isomer of monoolein. Monoolein is a MAG commonly used in the *in meso* method.

3.2 Objectives of this project

The objectives of this project were: 1) to synthesise compound **34**, to be used as the host lipid in the *in meso* method for membrane protein crystallisation and 2) to synthesise the 2-MAG, 2-monoolein, **36**, that could go on to be used in the *in meso* method to understand the role, if any, of 2-MAGs in crystallogenesis.

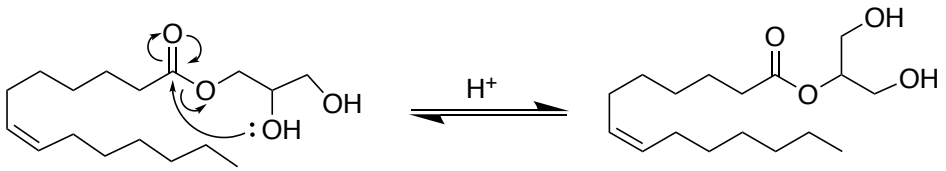
A retrosynthetic analysis is carried out for **34** and **36** in Scheme 3.1 and Scheme 3.3, respectively.

Scheme 3.1 Retrosynthetic analysis of **34**



The first retrosynthetic transformation, a), is an acetal deprotection. The protection of the diol in **34** is necessary in order to prevent side product formation during the synthesis. In this synthetic scheme, protection of the diol is carried out by an isopropylidene acetal group. The isopropylidene protection was originally used in the first reported synthesis of **34**.[41] Upon deprotection, the free hydroxyl groups can now take part in acyl migration to form its 2-isomer, **43** (Scheme 3.2).

Scheme 3.2 Mechanism of the isomerisation of **34**



34 **43**

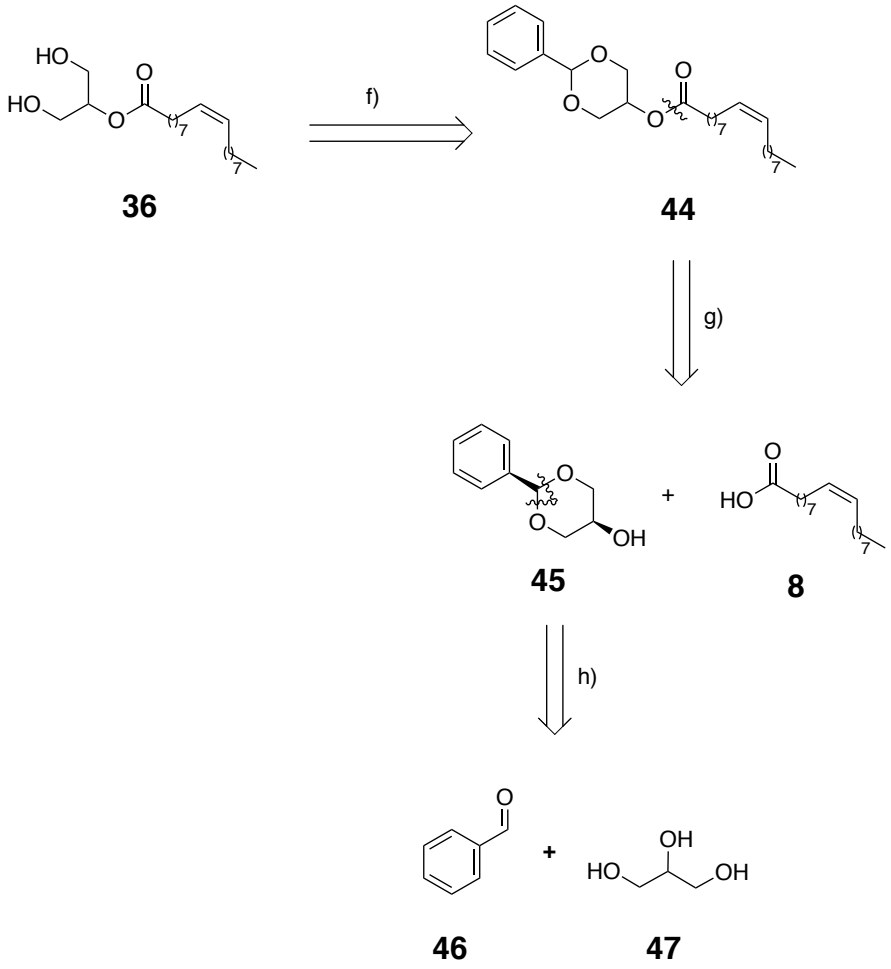
The next step in the retrosynthesis, b), is the Suzuki coupling of **38** and **40** that was carried out in this project as described by Hart *et al*.[41] The *cis* double bond of **34** is synthesised by $HBCy_2$ reduction, d), of the iodo-alkyne, **41**.[60] The presence of the bulky cyclohexane rings on the borane reducing agent allows the formation of the *cis* double bond, due to steric effects.

The iodine needs to be introduced into **42** to act as a leaving group for the subsequent Suzuki coupling reaction. The iodine is installed via a CuI/tetrabutylammonium bromide (TBAB) catalysed iodination under air in water.[61] This deviates from Hart *et al*.[60] that used n-BuLi and I_2. The CuI was preferred in this project as it offered a more environmentally benign alternative that avoided the use of flammable n-BuLi.

Compound **38** can be synthesised by a Steglich esterification between **12** and **39**. Both **12** and **39** are commercially available and are coupled using DCC and DMAP as described by Hart *et al.*[60]

A challenge throughout this synthesis is the large-scale reactions required, due to the need for gram quantities of the final product, **34**. For example, the synthesis begins with the iodination of 7.5 g of **42** and the coupling of ~ 5g of **39** and **12**.

Scheme 3.3 Retrosynthetic analysis of **36**



The first retrosynthetic step, f), from the target compound, **36**, is an acetal deprotection that gives **44**. Compound **44** contains a benzylidene acetal that protects the 1- and 3-positions of its glycerol moiety. Benzylidene acetal deprotection can be achieved by hydrogenolysis or exposure to acidic conditions. As **44** contains an alkene, hydrogenolysis can not be used to deprotect the acetal, as hydrogenation of the double bond in **44** would occur.

The second retrosynthetic transformation, g), is a Steglich esterification to give **45** and **8** from the ester, **44**. As the fatty acid chain of **36** is commercial (**8**, oleic acid), it can be coupled by a Steglich esterification to the free hydroxyl group of **45**. Therefore, the Suzuki reaction that was necessary to synthesise the MAG, **34**, is not necessary in this synthetic scheme. This greatly reduces the steps in the synthetic

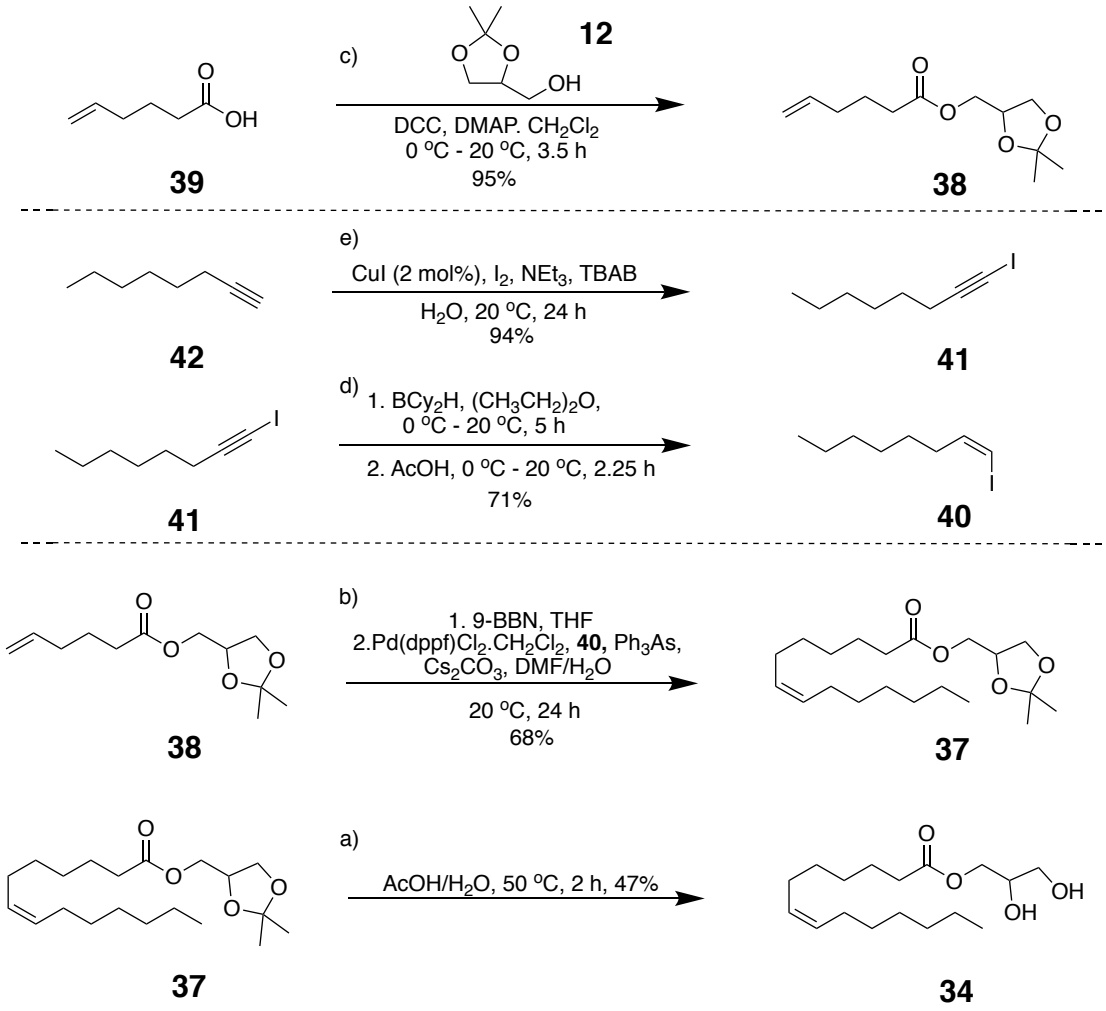
scheme.

The synthesis of **45**, is an acetal protection reaction, h), where benzaldehyde, **46**, and glycerol, **47**, and *p*-toluenesulfonic acid are heated to reflux in benzene. The water that is formed during the reaction is removed using a Dean-Stark apparatus.

The benzylidene acetal protecting group strategy was chosen for this project, as the literature suggested that the condensation reaction between **46** and **47** would yield 1,3-protected and 1,2-protected products, but that **45** could then be isolated following recrystallisation.[62–64]

3.3 Results and Discussion – Synthesis of 34

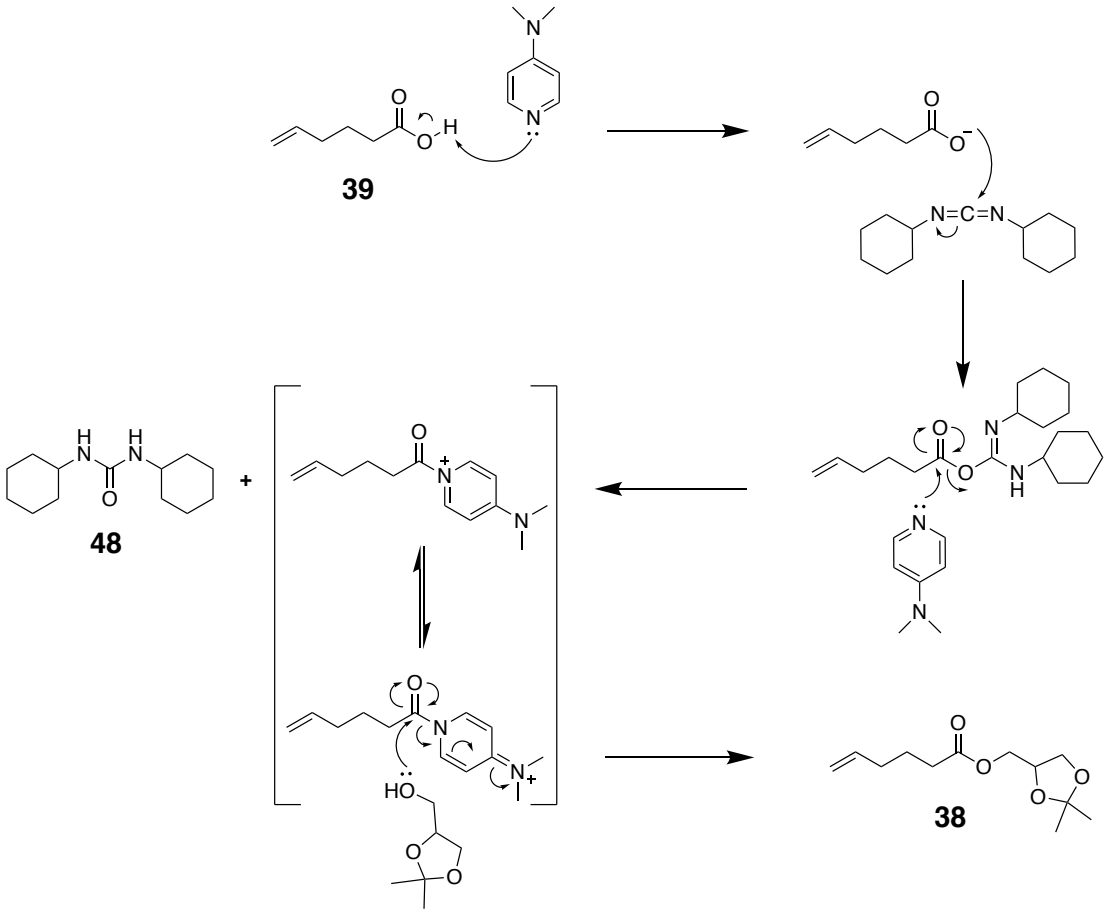
Scheme 3.4 The synthetic route to **34** in this project



The first step in the synthesis of **34** (Scheme 3.5, c) is a Steglich esterification

between **39** and **12** as described by Hart *et al.*[60] The reaction proceeded as expected to give yields of 90-95%. The proposed mechanism of the esterification is shown in Scheme 3.5.

Scheme 3.5 Proposed mechanism of the Steglich esterification

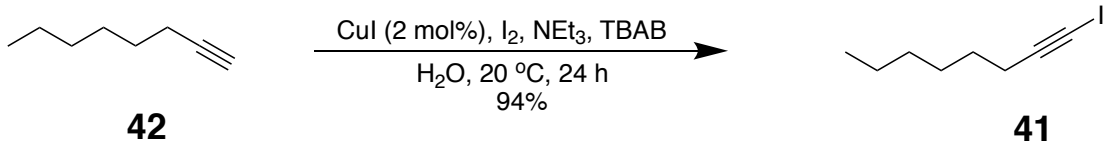


In the proposed mechanism. DMAP deprotonates **39** that then reacts with the electrophilic diimide of DCC. DMAP then acts as a nucleophile towards the **39**-DCC complex to form dicyclohexylurea (**48**) and a reactive intermediate that acylates the alcohol, **12**, to form **38**.

The purification was challenging, due to the large scale of the reaction and therefore, the large amount of the dicyclohexylurea side product, **48**, that is produced. Removal of **48** can be achieved by precipitation out of CH_2Cl_2 at 0 ^oC, subsequent filtration and flash chromatography over silica gel.

The iodination of the terminal alkyne (e) was necessary in order to incorporate a leaving group for the later Suzuki coupling reaction.

Scheme 3.6 CuI/TBAB catalysed iodination of **42**



In the original published procedure, **42** is iodinated using n-BuLi and I_2.[41] Due to the flammable nature of n-BuLi, this procedure was changed by a previous group member to a two-phase CuI/TBAB catalysed iodination (Scheme 3.6).[61] The phases in the reaction consist of water and an organic layer containing **42** and triethylamine. TBAB acts as a phase transfer catalyst to transport **42** to the interface of the two phases. It is here that the alkyne of **42** interacts with CuI, lengthening the C-H bond and allowing triethylamine to deprotonate. The copper acetylide will then act as a nucleophile towards the I_2 forming the iodo-alkyne and regenerating the copper catalyst.

The CuI/TBAB catalysis reaction offers a cheap, safe and more environmentally benign alternative to the previous n-BuLi reaction. Interestingly, the product, **41**, runs as two spots on TLC (Fig. 3.3). The reason for this is not known.

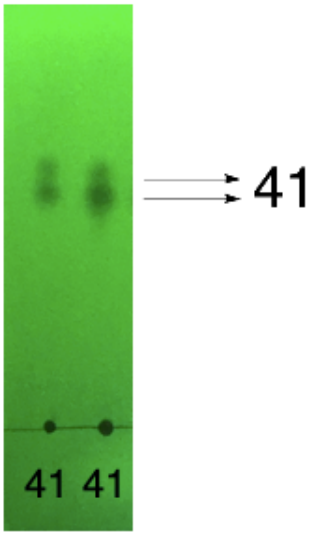
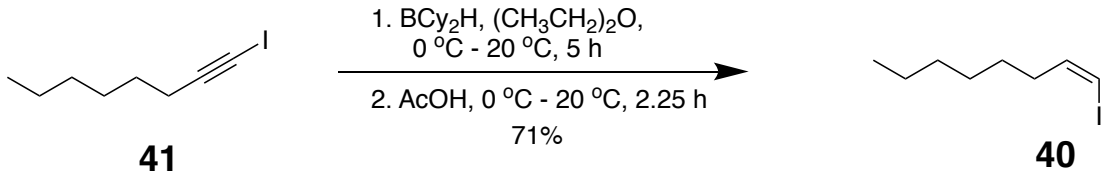


Figure 3.3 TLC of the CuI/TBAB catalysed iodination of **42**, run in hexanes and visualised under UV light at 254 nm.

The next step in the synthetic scheme is the reduction of **41** to **40** (d) that was carried out as described by Hart *et al.*[60] (Schemes 3.7 and 3.8). The reaction involves the *in situ* formation of the $HBCy_2$ (Cy = cyclohexyl) reducing agent by the coordination of

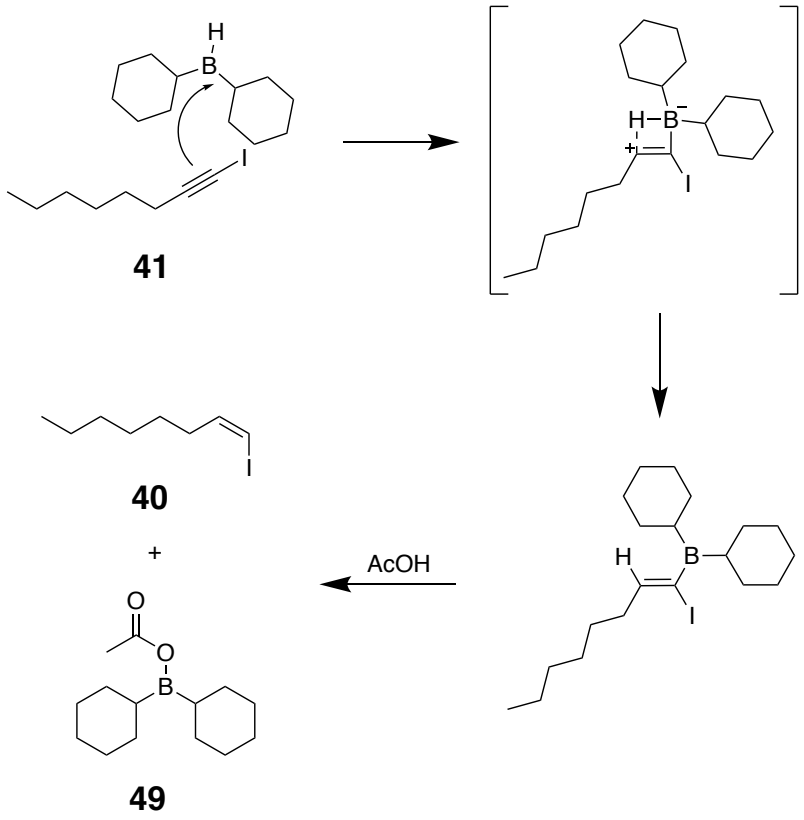
BH_3.DMS to 2 eq. of cyclohexene. The cyclohexyl groups now bound to the boron give the reducing agent a large steric presence. This steric presence is the source of the *cis* selectivity in the reaction, as it will strongly disfavour the formation of a *trans* transition state. Subsequent protonolysis using 16 eq. AcOH furnishes **40** in a yield of 71%.

Scheme 3.7 $HBCy_2$ reduction of **41**



41 **40**

Scheme 3.8 Mechanism of the reduction of **41**



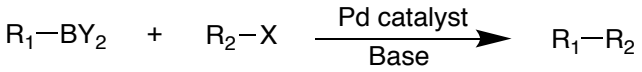
41 **40** + **49**

Following protonolysis, the *cis* alkene **40** is formed with the side product **49**. Addition of ethanolamine to the crude product during work-up results in the formation of an ethanolamine-BCy_2 complex that precipitates out and can be removed by filtration.

The Suzuki coupling reaction (b) is extremely valuable in organic chemistry, as it offers the opportunity to create carbon-carbon bonds. It is named after the scientist

Akira Suzuki and its discovery led to his reception of the 2010 Nobel Prize in Chemistry along with Heck and Negishi for the discovery and development of carbon-carbon bond formation by palladium catalysed cross-coupling reactions (Scheme 3.9).

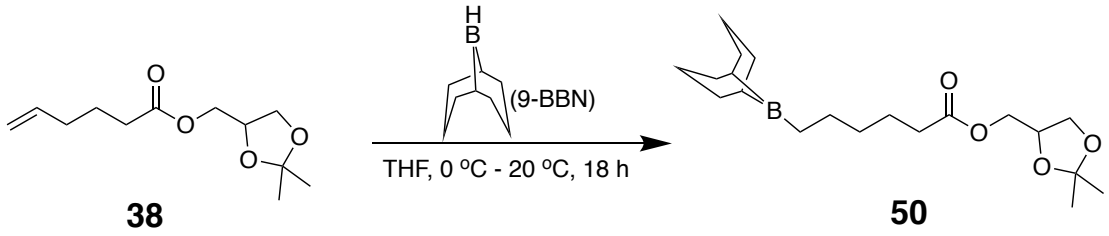
Scheme 3.9 General scheme of a Suzuki reaction



The Suzuki reaction allows the formation of carbon-carbon bonds from the reaction between an organic boronic acid (R_1-BY_2) and an organic halide (R_2-X) in the presence of a palladium catalyst and base.

In this project, the boronic acid necessary for a Suzuki coupling is synthesised *in situ* by adding 9-BBN to **38** in THF (Scheme 3.10).

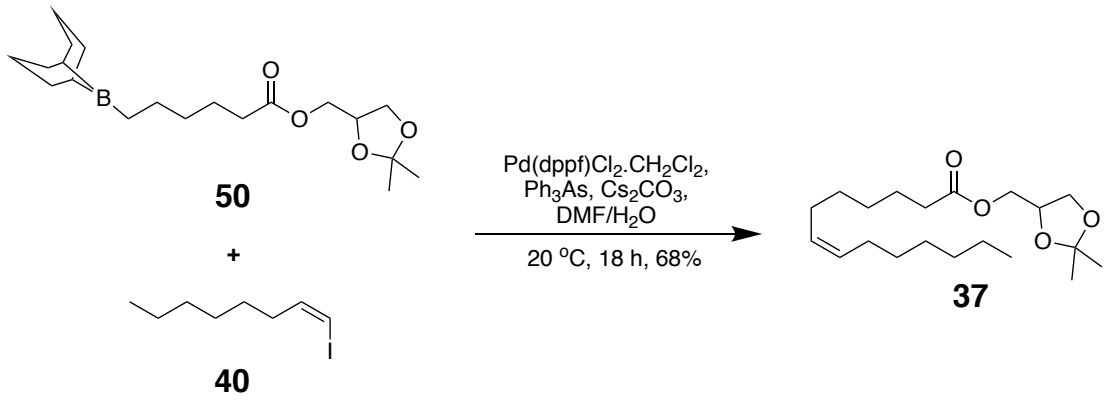
Scheme 3.10 Preparation of the boronic acid **50**



38 **50**

The reaction to prepare of the boronic acid from **50** was allowed to stir for 18 h at 20 ^{o}C to ensure complete transformation to the boronic acid. In this reaction, 9-BBN adds to the double bond in **38**. The boronic acid solution is then diluted with water and transferred to the Suzuki reaction flask containing the catalyst ($Pd(dppf)Cl_2.CH_2Cl_2$), ligand (Ph_3As), base ($CsCO_3$) and solvent (DMF/H_2O) (Scheme 3.11).

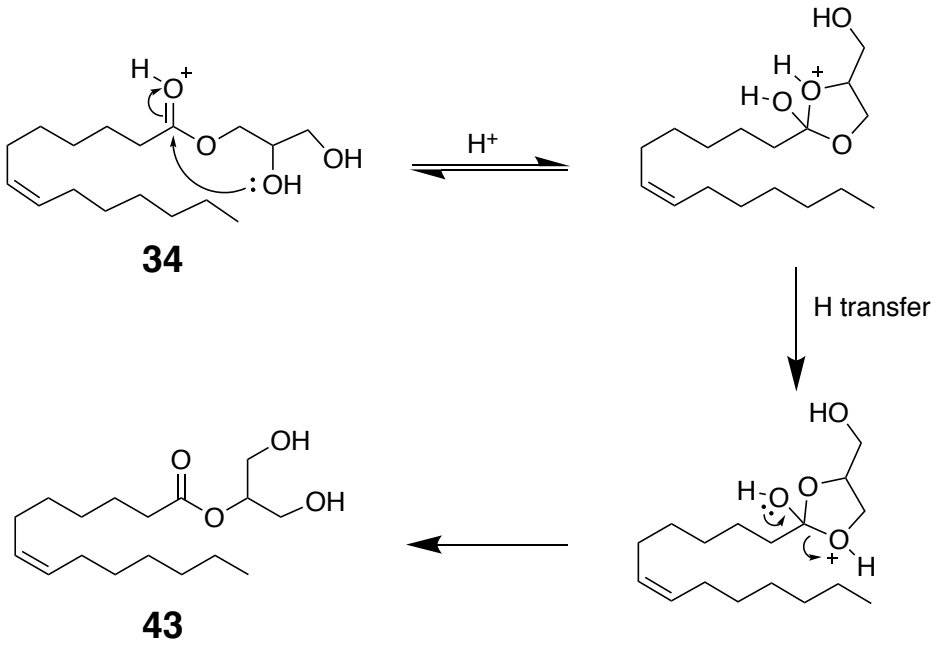
Scheme 3.11 Suzuki coupling to form **37**



The Ph_3As in the reaction mixture will coordinate to the $Pd(dppf)Cl_2$, to remove the two chlorines and form the active catalyst, $Pd(dppf)(Ph_3As)_2$. Next, oxidative addition of **40** will occur, followed by transmetallation of the boronic acid, **50**, and reductive elimination to give **37** and to regenerate the catalyst. Critically, the Suzuki reaction conserves the *cis* double bond of **40** in the product, **37**. Purification of **37** is challenging due to the presence of some unreacted **38** in the crude product that behaves similarly to **37** during flash chromatography over silica gel. Multiple flash chromatographic columns were necessary to fully remove **38**. In addition, the palladium catalyst is extracted into the organic layer during extraction with ether during work-up. The catalyst gives the crude product a black colour and was also removed during flash chromatography over silica gel, as it does not elute through the column.

While the Suzuki coupling step is undoubtedly the most critical in this synthetic scheme, the acetal deprotection (a) is also important as it can determine the level of the 2-isomer, **43**, formed. The presence of **43** is important for this work, as it affects the purity of the final product. Acidic conditions are necessary for acetal deprotection, but, unfortunately, acidic conditions also promote the transesterification reaction of **34** to **43** (Scheme 3.12).

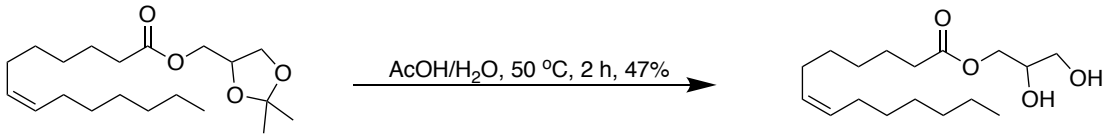
Scheme 3.12 Intramolecular transesterification of **34** to **43**



34 **43**

In the published protocol from this research group,[60] 2 M HCl in MeOH were the conditions used for the deprotection of **37**. In this project, deprotection conditions were changed to AcOH/H_2O (4:1, v/v)[65] (Scheme 3.13) with the idea that the weaker acid conditions would result in lower levels of isomerisation to **43**.

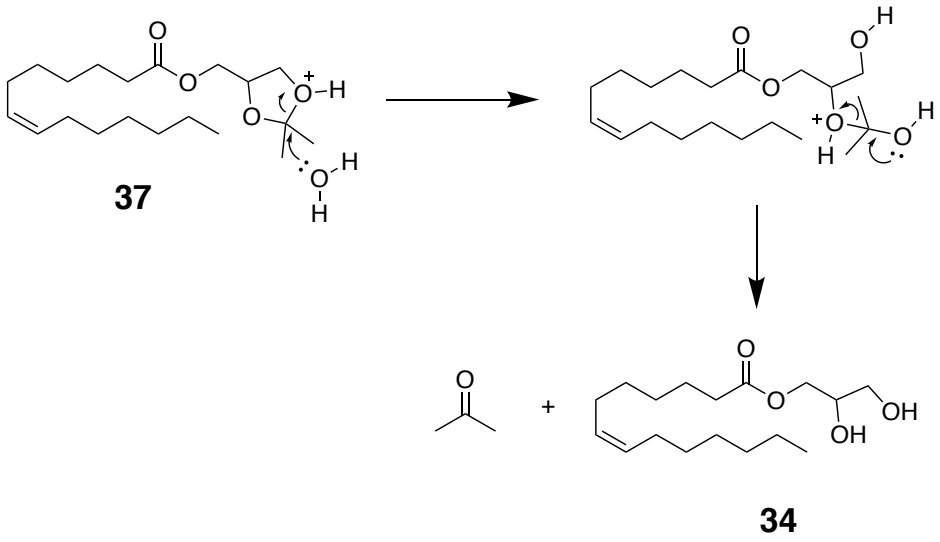
Scheme 3.13 Acetal deprotection of **37**



37 **34**

The mechanism of this acetal deprotection reaction is shown in Scheme 3.14, acidic conditions protonate the acetal followed by nucleophilic attack of water to form acetone and **34**.

Scheme 3.14 Mechanism of the acetal deprotection



After the AcOH/H_2O deprotection and subsequent flash chromatography over silica gel, the amount of 2-MAG, **43**, was quantified at 3.8% using ^{1}H NMR (Fig. 3.4).

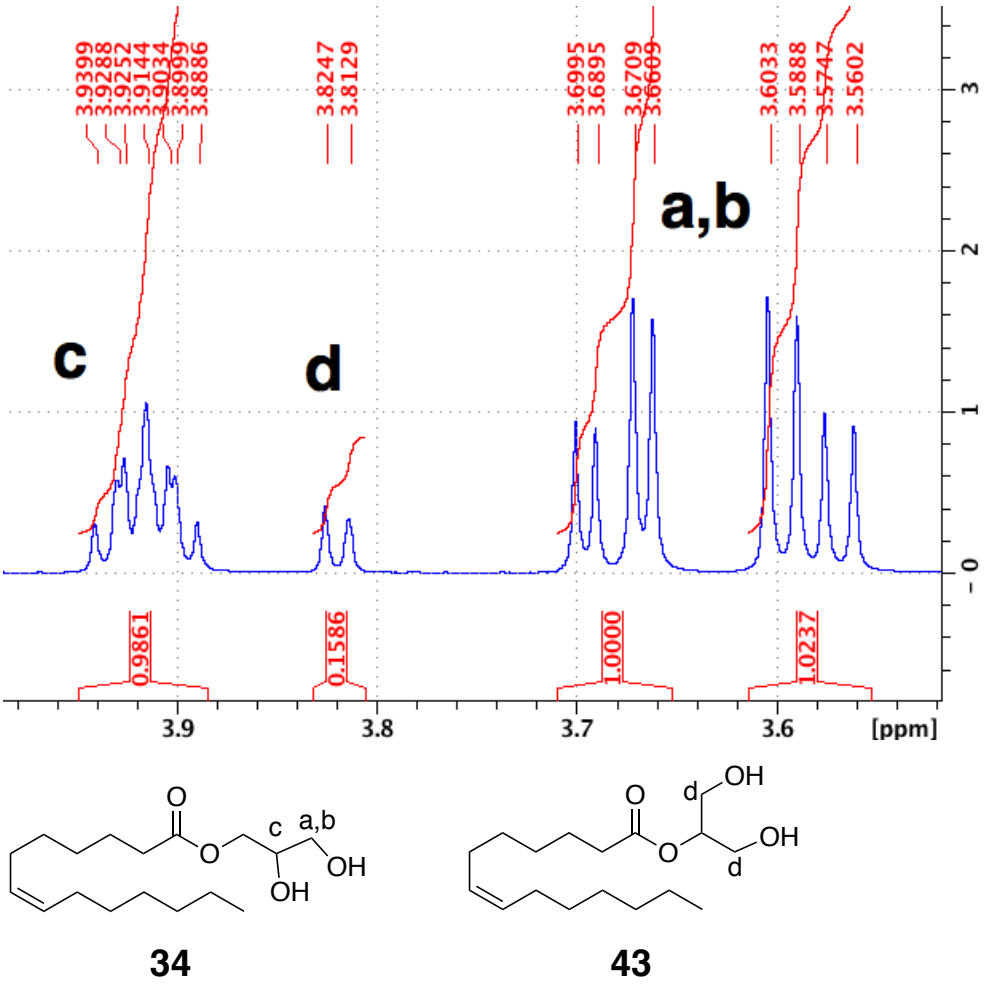


Figure 3.4 ^{1}H NMR of **34** in $CDCl_3$ showing the presence of small amounts of **43**.

The doublet at 3.82 ppm in Fig. 3.4 represents the 2 methylene groups at d in **43** and integrates for 0.1586. As this peak relates to 4 protons, the integral is divided by 4 to give 0.039 that gives a wt% value of 3.8% of **43** present. The goal of this synthesis was to achieve the lowest level of isomerisation to **43** as possible, *i.e.*, achieving the highest purity of **34**. This reduced level of 3.8% of **43** after acetal deprotection was an improvement on the previous reported literature value of 5%.[41]

Unfortunately, a minor impurity spot with a similar R_f to **34** could be seen by TLC (Fig. 3.5), even after the first flash chromatographic purification over silica gel. Achieving a high level of purity was essential if **34** was to be used in the *in meso* method. For this reason, further flash chromatographic purifications over silica gel were carried out. Following the first purification of the acetal deprotection, there was 5.0 g of **34** (containing 3.8% of **43** and an impurity visible on TLC, Fig. 3.5).

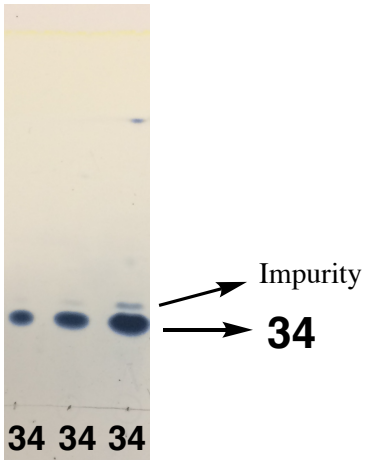
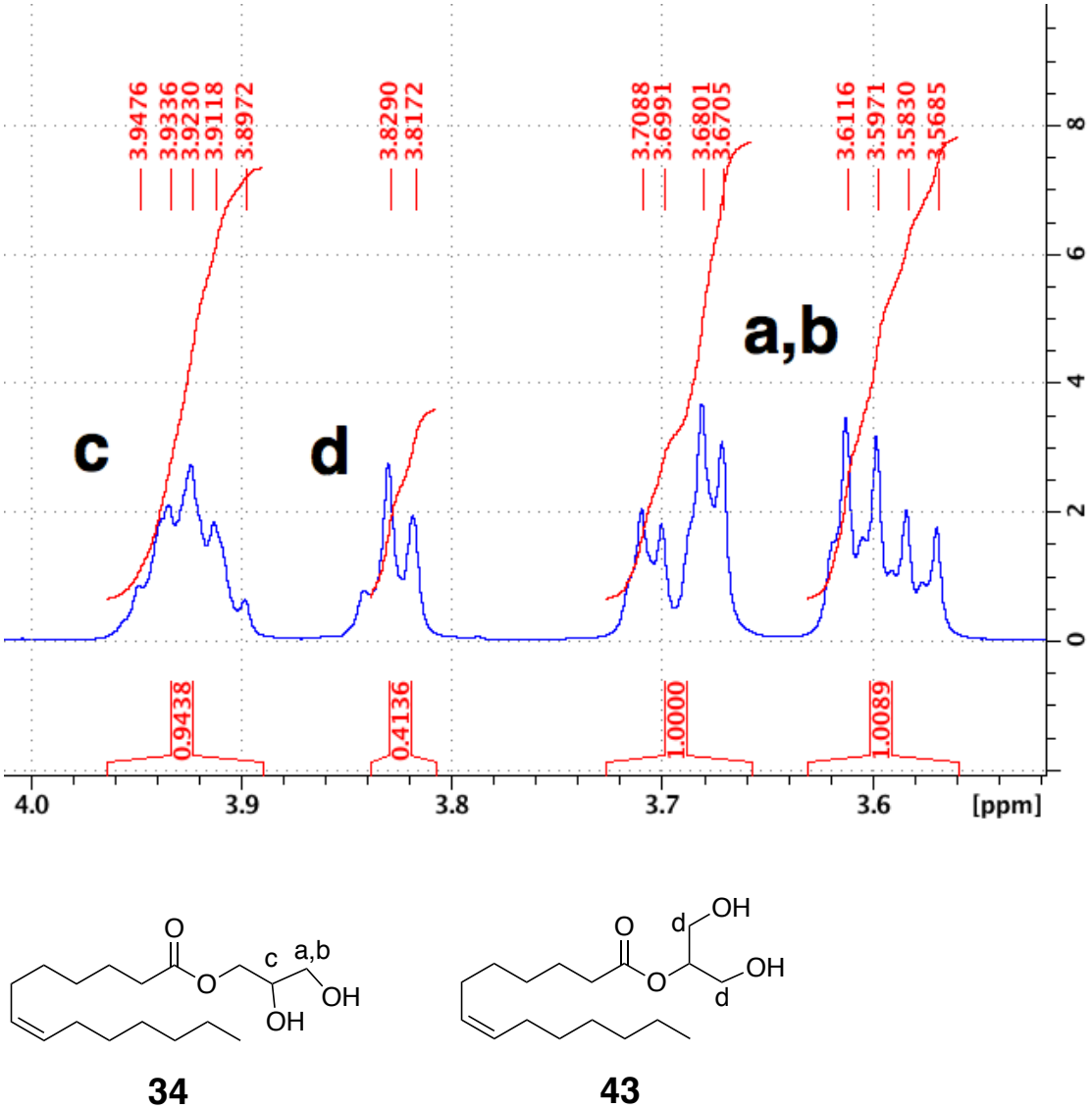


Figure 3.5. TLC of **34** after one flash chromatographic purification over silica gel, run in hexanes/Ac_2O/EtOAc, 75:24:1, stained in ammonium molybdate and heated to 180 ^oC using a heat-gun.

A second flash chromatography purification was performed on the 5.0 g of **34**, but separation between the spots for **34** and the impurity was not achieved. Early fractions during purification were yellow in colour and contained the impurity alone. These fractions were discarded. Later fractions were mixed, containing both **34** and the impurity and were pooled together. In order to achieve better separation, the 5.0 g of

34 was divided into two 2.5 g samples and purified separately. This allowed slightly better separation of the impurity spot to yield 3.1 g of **34**.

While the three further flash chromatographic purifications over silica gel were successful in removing the impurity, transesterification occurred throughout the purifications and the amount of **43** rose from 3.8 wt% to 9.3 wt% (by ^{1}H NMR, Fig. 3.6). The transesterification from **34** to **43** would have been promoted by the slightly acidic silica gel, and during rotary evaporation at 30 ^{o}C. In order to decrease isomerisation in the future, temperatures during rotary evaporation can be reduced. In addition, smaller-scale reactions can be carried out that should offer shorter purification times that will reduce isomerisation.



34 **43**

Figure 3.6 NMR of **34** after three further flash chromatography purifications in $CDCl_3$ still showing the presence of **43**.

The integral for the 2 methylene groups of **43** is 0.41, which gives 9.3 wt% of **43**.

Fortunately, a method of separating **43** from **34** was developed in this research group[41] and optimized in the literature by Zhang *et al.*[66] A low temperature recrystallization of a mixture of **34** and **43** in MeCN allows the opportunity for removal of **43**. At -20 °C, **34** will crystallize out of the bulk solution, whereas **43** will remain in solution. One round of recrystallization resulted in approximately a halving of the content of **43**. Two rounds of recrystallization were carried out on **34** in this project, reducing the content of **43** from 9.3 wt% to 2.0 wt% (Fig. 3.7).

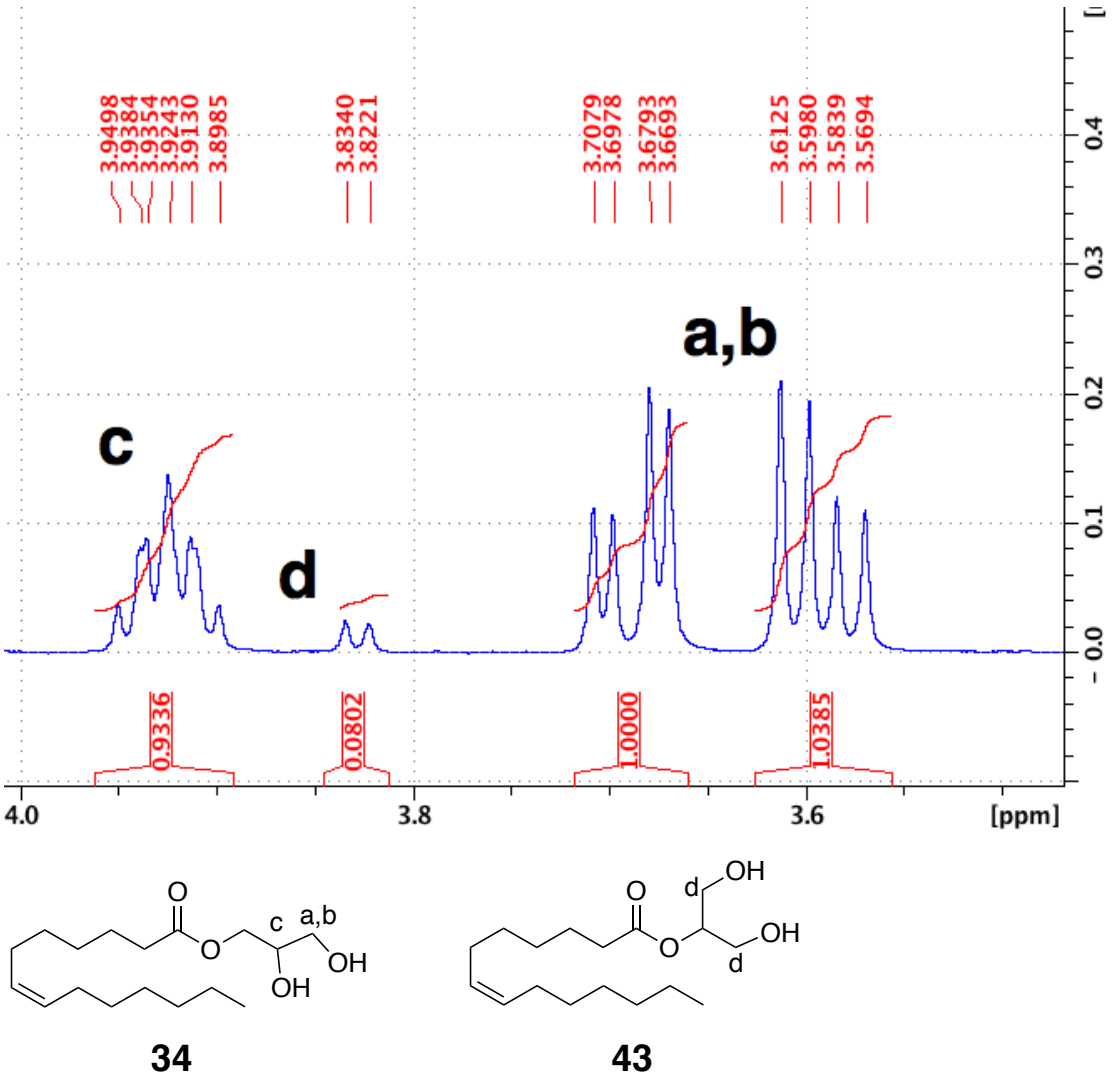
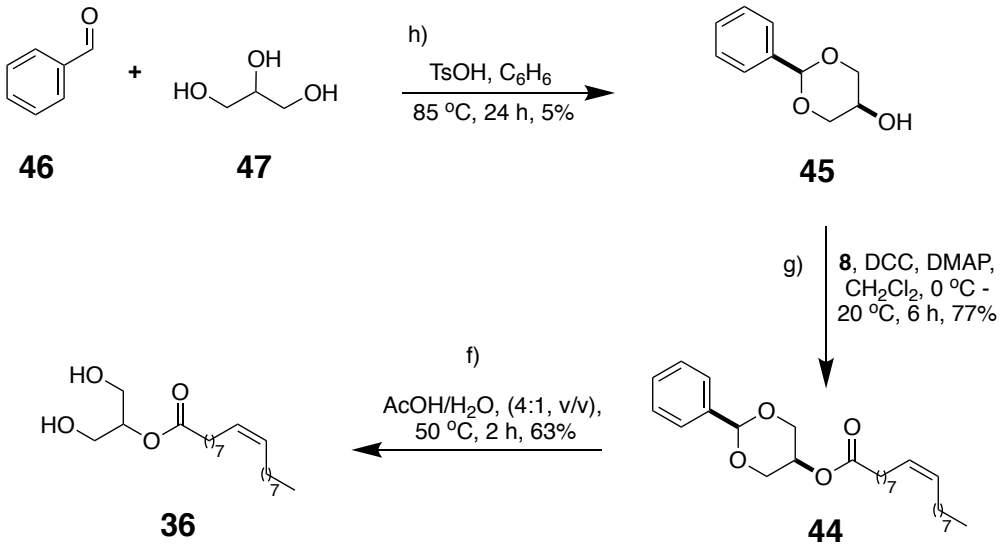


Figure 3.7 ^{1}H NMR of **34** following two rounds of recrystallization.

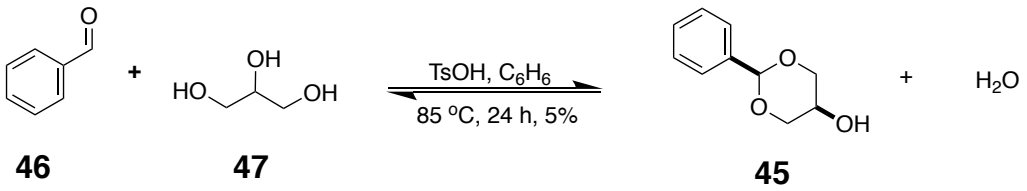
After two rounds of recrystallisation from MeCN, the integral for d is 0.08, giving a value of 2.0 wt% of **43**.

Scheme 3.15 Synthetic scheme to synthesise **36**



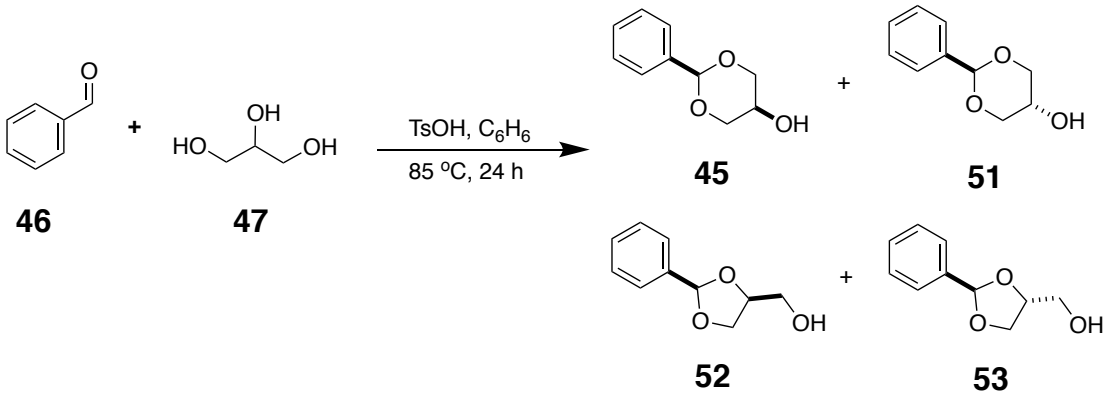
Next, the benzylidene acetal was introduced into **47** by condensation with **46** in benzene at 85 ^{o}C using *p*-toluenesulfonic acid (TsOH) as catalyst (h). The water produced in the condensation reaction was removed from the reaction mixture using Dean-Stark apparatus, as water would hydrolyse the acetal product, **45** (Scheme 3.16).

Scheme 3.16 Acetal protection of **47**



TsOH catalyses the reaction by protonating **46** and promoting nucleophilic addition. The three hydroxyl groups of **47** act as nucleophiles towards the electrophilic aldehyde of **46** to give four products, **48** and **51**-**53** (Scheme 3.17).

Scheme 3.17 Products of the condenation of **46** and **47**

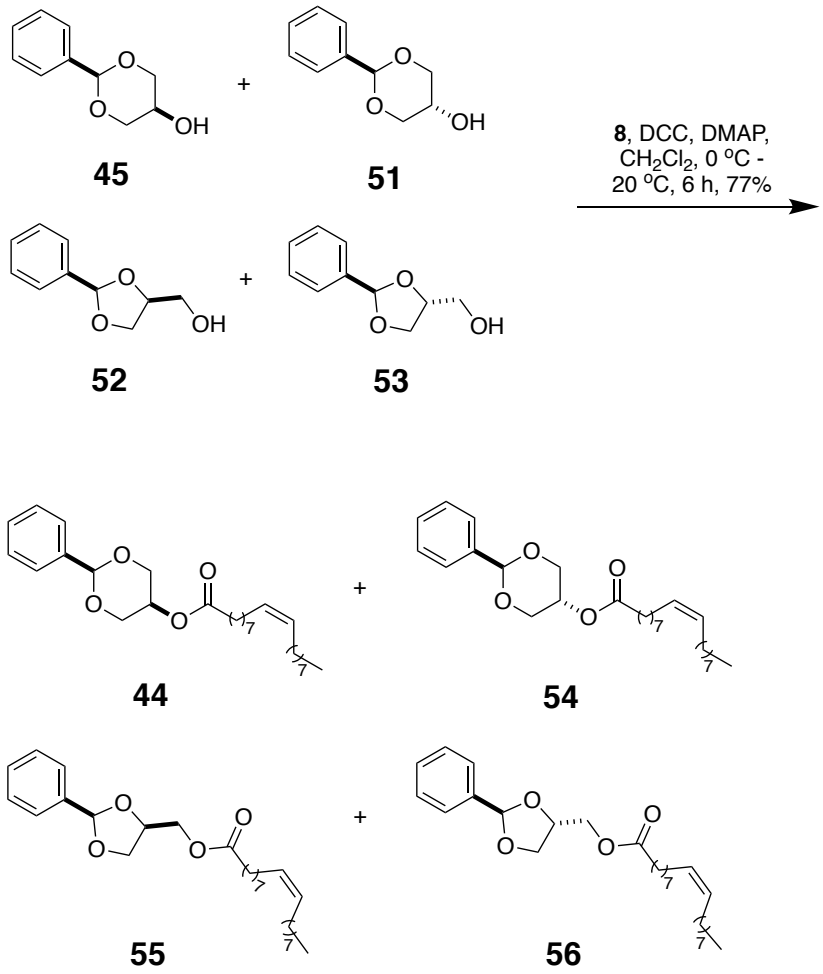


46 **47** **45** **51** **52** **53**

While the yield was 5%, the starting materials are cheap and common enabling a simple scale-up to afford large amounts of **45**. Recrystallisation from petroleum ether/toluene (1:1, v/v) at -10 °C according to Zheng *et al*.[67] afforded **45** with ~10 wt% of **51**-**53**. Flash chromatography over silica gel could not separate **45** and **51**-**53** but the resulting products of the Steglich esterification with **8** (Section 3.3.8) could be easily separated.

The 1,2- and 1,3-protected products of the benzylidene acetal protection: **45** and ~10 wt% of **51**-**53**, were coupled to oleic acid, **8**, using Steglich esterification conditions as illustrated in Scheme 3.18 (g).

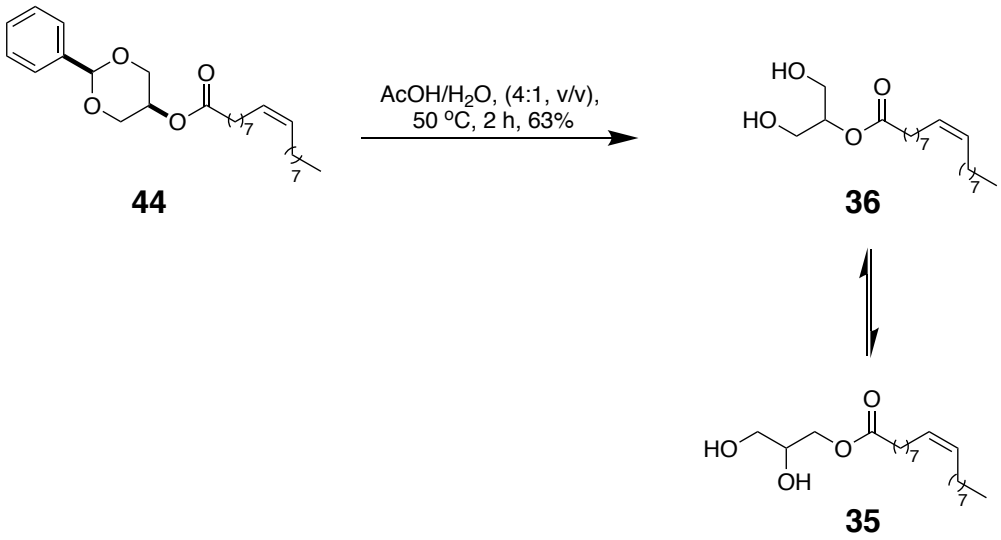
Scheme 3.18 Steglich esterification of **45** and **51**-**53**



The desired product, **44**, could be separated from **54**-**56** by flash chromatography over silica gel to give a final yield of 77%.

The final step in the synthesis is deprotection of the benzylidene acetal of **44** (Scheme 3.19, f).

Scheme 3.19 Benzylidene acetal deprotection of **44**



As the deprotection reaction furnishes **36** containing two free hydroxyl groups, transesterification of **36** to its 1-MAG isomer (monoolein), **35**, can take place. Deprotection reactions were carried out in AcOH/H_2O (4:1, v/v) at 50 ^oC with reaction times of 2 h and 4 h. Unsurprisingly, lower isomerisation to **35** was observed at a reaction time of 2 h (Table 3.1).

Table 3.1 Isomerisation of **36** during benzylidene acetal deprotection

Temperature	Reaction time (h)	Reaction conditions	Isomerisation to 35	Yield
50 ^0C	2	AcOH/H_2O (4:1, v/v)	20 wt%	63%
50 ^oC	4	AcOH/H_2O (4:1, v/v)	50 wt%	79%

While the 4 h reaction showed a higher yield of 79%, a balance between overall yield and extent of isomerisation must be struck for this reaction. The 2 h deprotection was preferred in this project, as after purification the product only contained 20 wt% of **35**, compared to 50 wt% for the deprotection reaction lasting 4 h.

The synthetic **36** (containing 20 wt% of **35**) was doped into commercially available

35 (containing 1 wt% of **36**) to create a number of samples of monoolein that contain a varying concentration of 2-monoolein (between 1 and 18 wt%).

This was done by dissolving **36** and **35**, separately in $CHCl_3$ and combining the appropriate amount of each solution, following by drying under a stream of N_2 and then under vacuum overnight. Doping was carried out in order to create a series of samples of the 1-isomer containing varying amounts of the 2-isomer (**36**). This series can then be used in the *in meso* method of membrane protein crystallisation of the membrane proteins in order to probe the effects of the presence of 2-MAG isomers on crystallogenesis.

3.4 Conclusions and Future Work

For this research group, the synthesis of MAGs to be used in the *in meso* method is essential, due to the inconsistent nature of commercially available MAGs. In this project, **34** was synthesised following a procedure by Hart *et al.*[60] with some modifications. Recrystallisation from MeCN was carried out on **34** in order to reduce its 2-MAG content. The final product can now be used in the *in meso* method in order to grow membrane protein crystals.

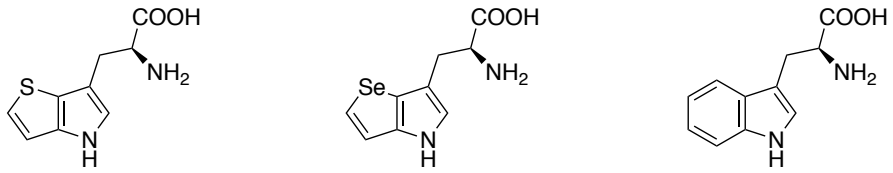
As a 2-MAG, **36**, has been synthesised and doped into **35** to create a series of MAGs containing between 1 and 18 wt% 2-MAG, the series can be used to determine the effect, if any, of 2-isomers in the *in meso* method. The results will be beneficial, as if 2-isomers have no effect, this research group will no longer carry out recrystallisations on MAGs that reduce overall yield. If 2-isomers are found to have a positive effect on the crystallisation of membrane proteins, then the *in meso* method can be used with a MAG with a high wt% of its 2-isomer.

Chapter 4: Solving the crystallographic phase problem

4.1 Introduction

The incorporation of non-canonical amino acids into peptides and proteins has found myriad applications in biological chemistry, such as in protein NMR, or, as in this project, solving the cystallographic phase problem.[46,68] This chapter will deal with the synthesis of sulfur and selenium substituted tryptophan derivatives, **57** and **58** (Fig. 4.1), following a synthetic procedure devised by Soth *et al.* and improved upon by Phillips *et al*.[48,69] While the synthesis of **58** was the original goal of this project, **57** was synthesised to optimize conditions for **58**. As the project progressed, advancements in synchrotron technology gave the application of S-SAD phasing a real boost.[46] These developments allowed access to significantly longer X-ray wavelengths (3 Å – 4 Å), thereby offering greater sensitivity towards the anomalous diffraction from sulfur atoms in proteins, which makes **57** more interesting. As previously stated, **57** was originally synthesised in order to optimize conditions for **58**, but now **57** can be used in phasing.

Following synthesis, **57** and **58** would be incorporated into membrane proteins and would then be used to try and solve the crystallographic phase problem, as described in Section 1.8.

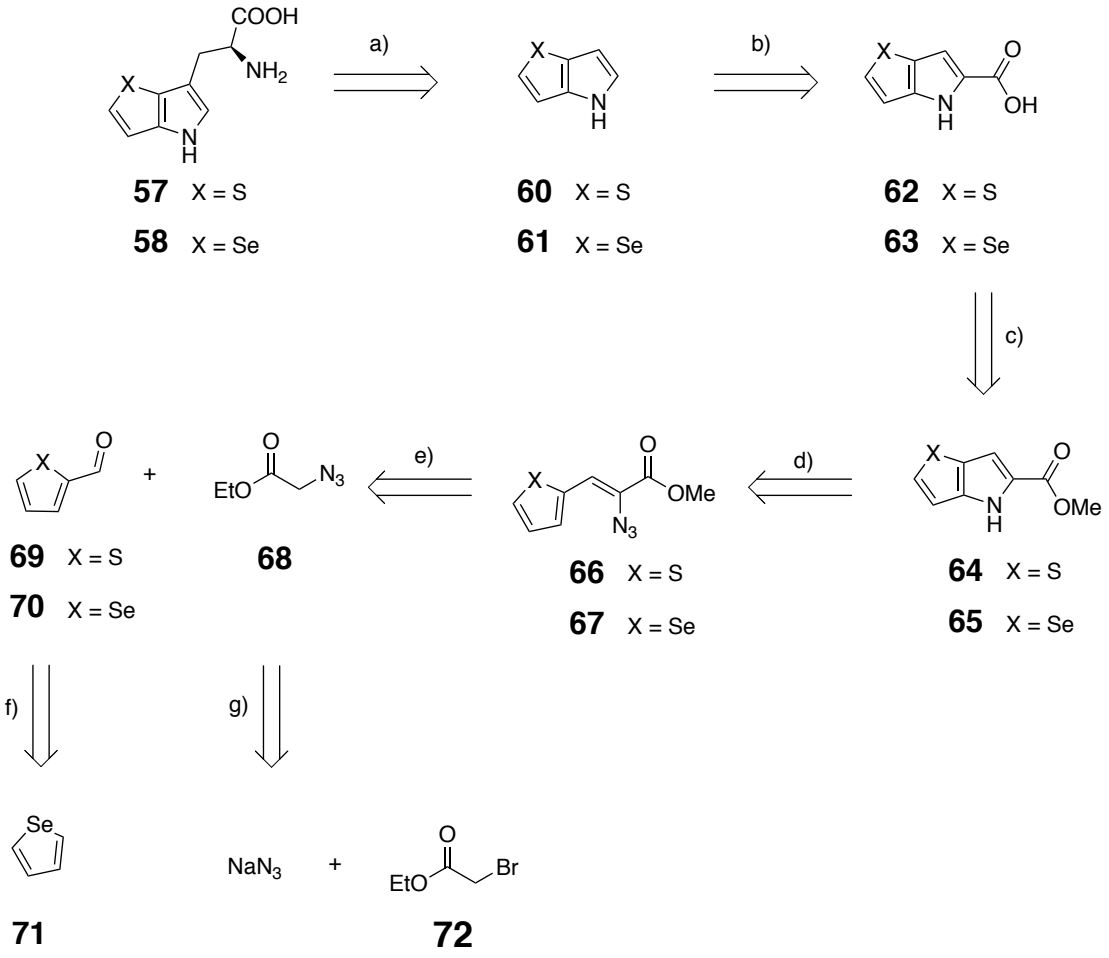


57 **58** **59**

Figure 4.1. The indole ring of tryptophan, **59** is replaced with thienopyrrole and selenopyrrole moieties in **57** and **58**, respectively.

A retrosynthetic analysis of **57** and **58** is shown in Scheme 4.1.

Scheme 4.1 Retrosynthetic analysis of **57** and **58**



The first retrosynthetic step, **a)**, is the removal of the amino-acid moiety from **57** and **58** give the heterocycles, **60** and **61**. The incorporation of the amino acid moiety can be carried out by enzymatic synthesis using tryptophan synthase according to Hoesl *et al.*[70]

Step **b)** is a decarboxylation that removes the carboxylic acid groups of **62** and **63** to to yield **60** and **61**. The decarboxylation of **62** and **63** has been described by Welch *et al.*[48]

In order for the decarboxylation step in **b)** to be possible, the esters of **64** and **65** have to be hydrolysed to carboxylic acids. Step **c)** is a base-catalysed ester hydrolysis that has been described by Welch *et al.*[48] using KOH.

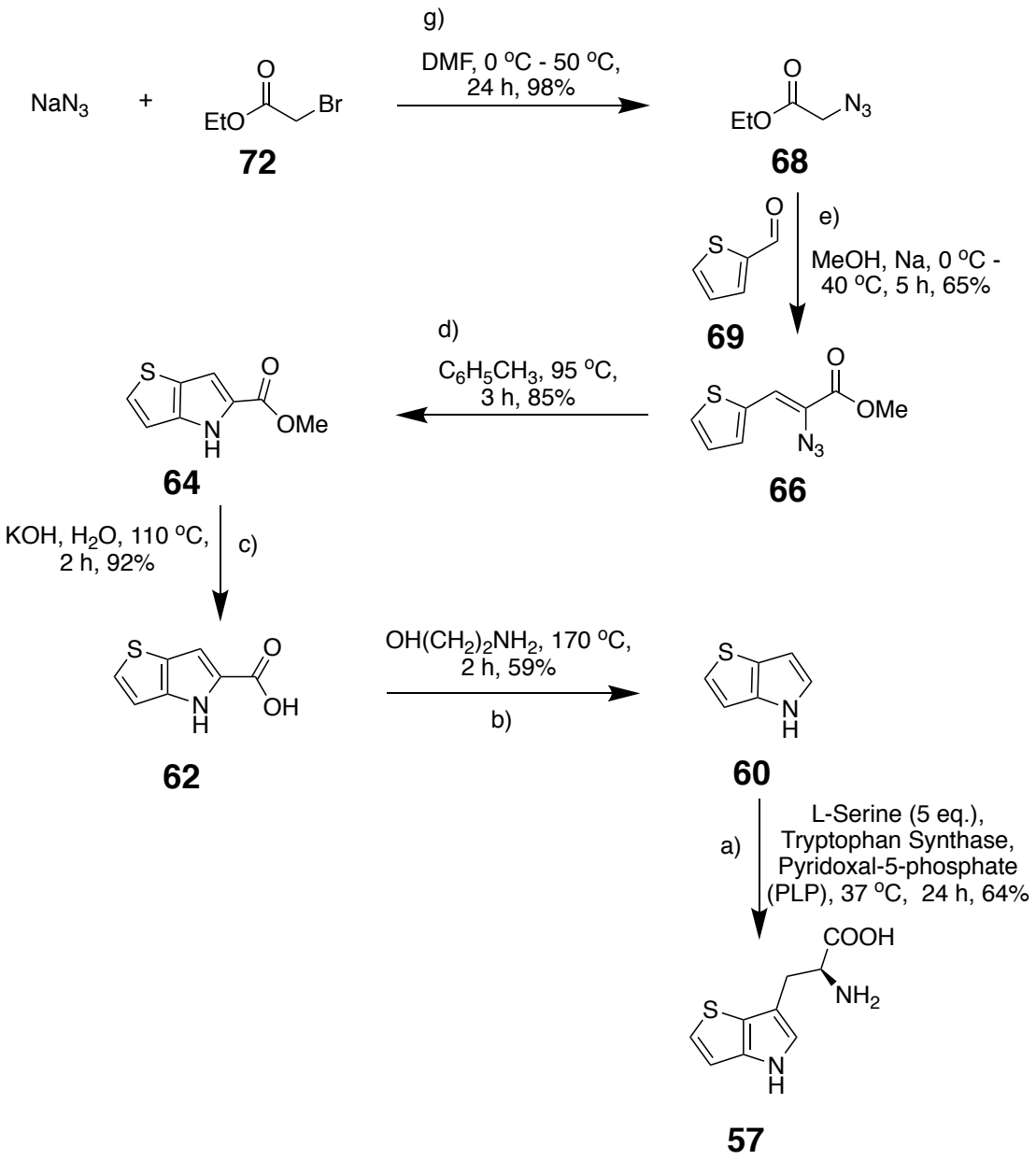
Step **d)** is a Hemetsberger reaction transformation. The Hemetsberger reaction allows the cyclisation of **66** and **67** to form **64** and **65**.

The azides, **66** and **67** are synthesised by a Knoevenagel condensation reaction, **e)**, between **68** and **69**/**70**. **69** is commercially available, whereas **70** is not, but can be synthesised by a Vilsmeier reaction, **f)** starting from the commercially available selenophene, **71**.

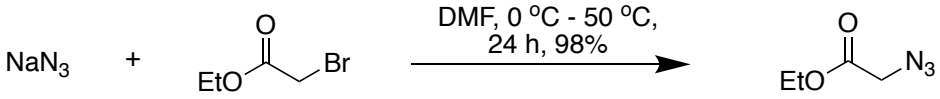
Ethyl azidoacetate, **68**, can be synthesised by an S_N2 reaction between sodium azide and ethyl bromoacetate, **72.**

4.2 Results & Discussion:

Scheme 4.2 Full synthetic scheme of **57**



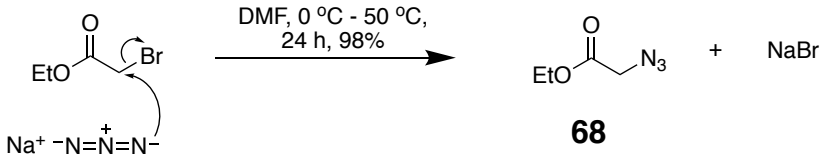
Scheme 4.3 Synthesis of **68**



72 **68**

The first step in the synthesis of both **57** and **58** is an S_N2 (substitution nucleophilic bimolecular) reaction between sodium azide and **72** (g). Heating to 50 °C for 24 h in DMF gives full conversion to the **68**. This reaction proceeds through an S_N2 mechanism, as the new bond between the azide and **72** is formed at the same time as the C-Br bond is cleaved (Scheme 4.4).

Scheme 4.4 S_N2 mechanism of the synthesis of **68**



68

The reaction could be reliably monitored by infrared spectroscopy (IR): the appearance of a strong absorption at 2103 cm^{-1} indicative of an azide functional group (Fig. 4.2).

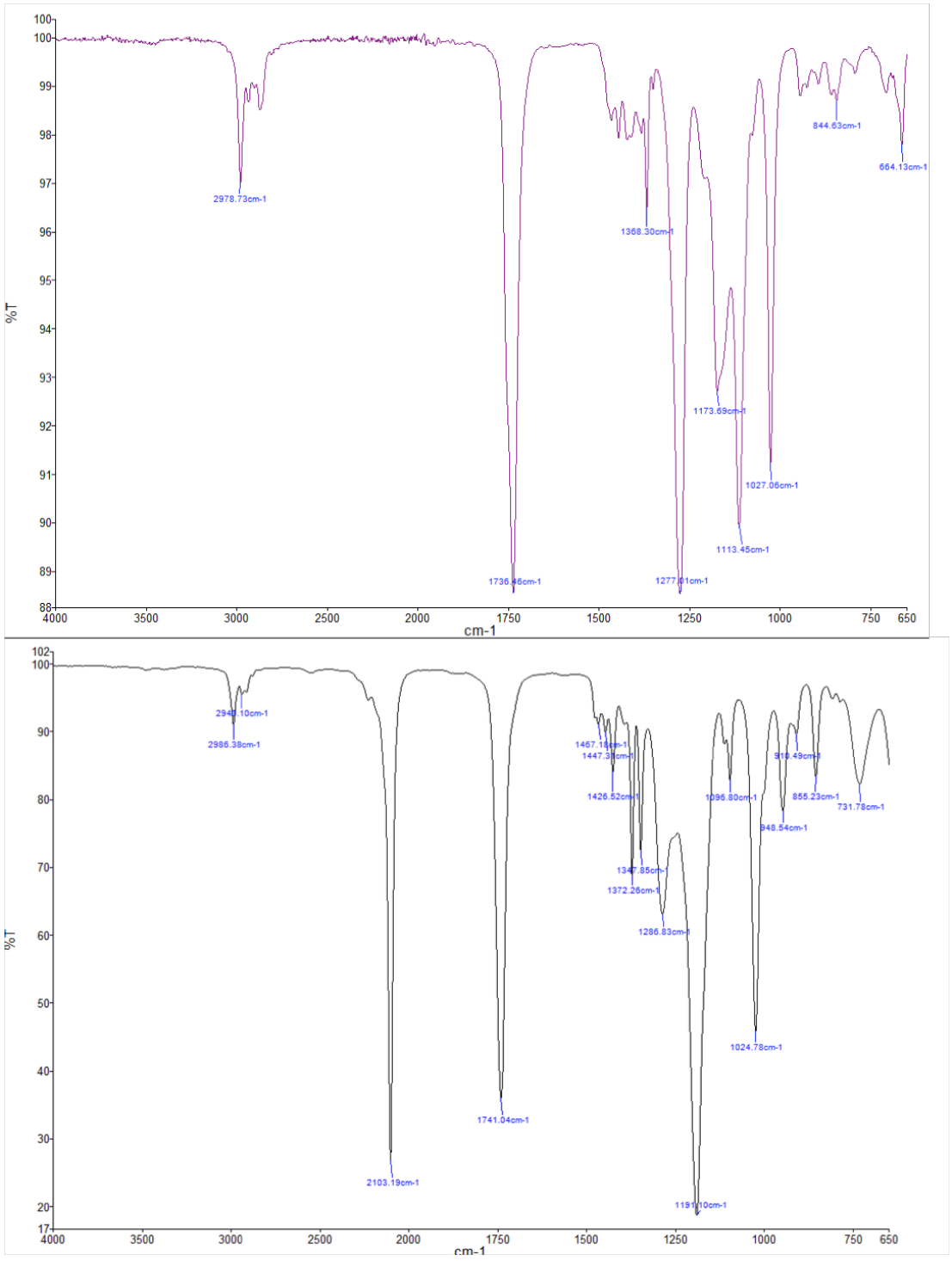
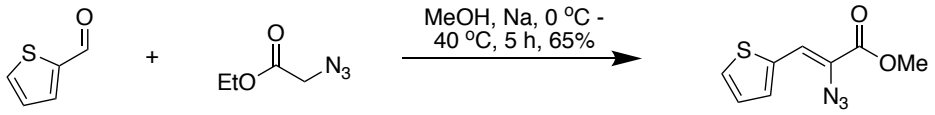


Figure 4.2 IR of the starting material, **72**, (top, pink) and the product, **68** (bottom, black). Azide peak at 2103 cm^{-1}.

During the work-up of the first synthesis of **68** carried out, losses occurred during rotary evaporation of solvent under reduced pressure due to the low boiling point of **68** (44-46 ^{o}C) to give a slightly reduced yield of 78%. In subsequent work-ups, lower pressures during rotary evaporation were avoided to ensure maximum yield.

Next, a Knoevenagel condensation reaction between **68** and **69** yielded the alkene **66** (Scheme 4.5, e). Addition of sodium metal to anhydrous methanol formed sodium methoxide *in situ* and deprotonated at the α-position to the ester in **68**. The reaction was sluggish if the solvent was not completely anhydrous, as this led to the formation of hydroxide base instead of methoxide. Hydroxide's lower pKa of 14, led to reduced product formation; otherwise it afforded **66** in reliable yields of 65%. The use of MeOH as solvent results in the inconsequential transesterification of **68** from the ethyl ester to the methyl ester.

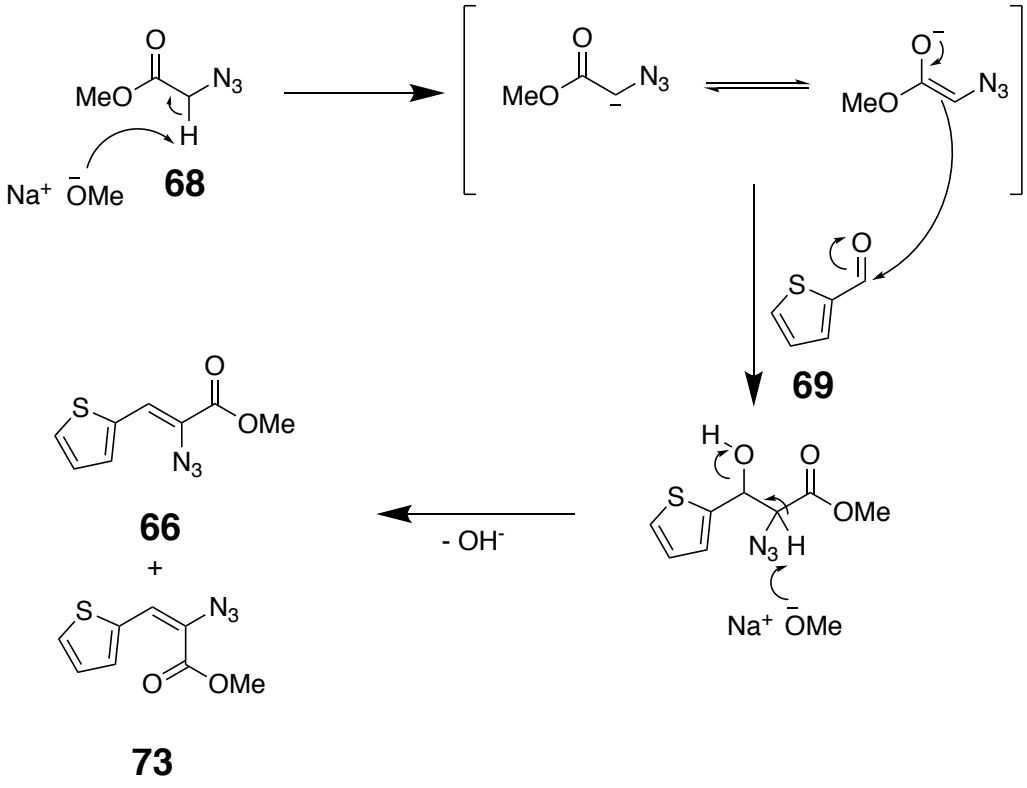
Scheme 4.5 Synthesis of **66**



69 **68** **66**

The mechanism of the condensation reaction is shown in Scheme 4.6 and leads to the formation of *cis* and *trans* products, which run as two spots on TLC.

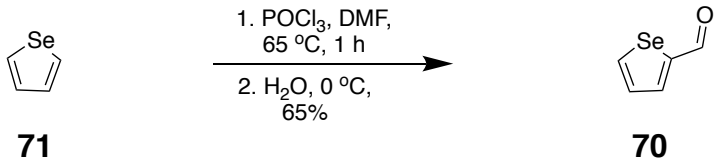
Scheme 4.6 Mechanism of the Knoevenagel condensation to form **66**



Sodium methoxide deprotonates **68** at the α-carbon to form an enolate that then acts as a nucleophile towards the aldehyde of **69**. A subsequent deprotonation and loss of water leads to the formation of the *cis* and *trans* products, **66** and **73**.

For the Knoevenagel reaction in the synthesis of **67**, selenophene-2-carbaldehyde, **70**, is not commercially available. Fortunately, selenophene, **71**, is available and a Vilsmeier reaction allows access to **70** (Scheme 4.7, f)

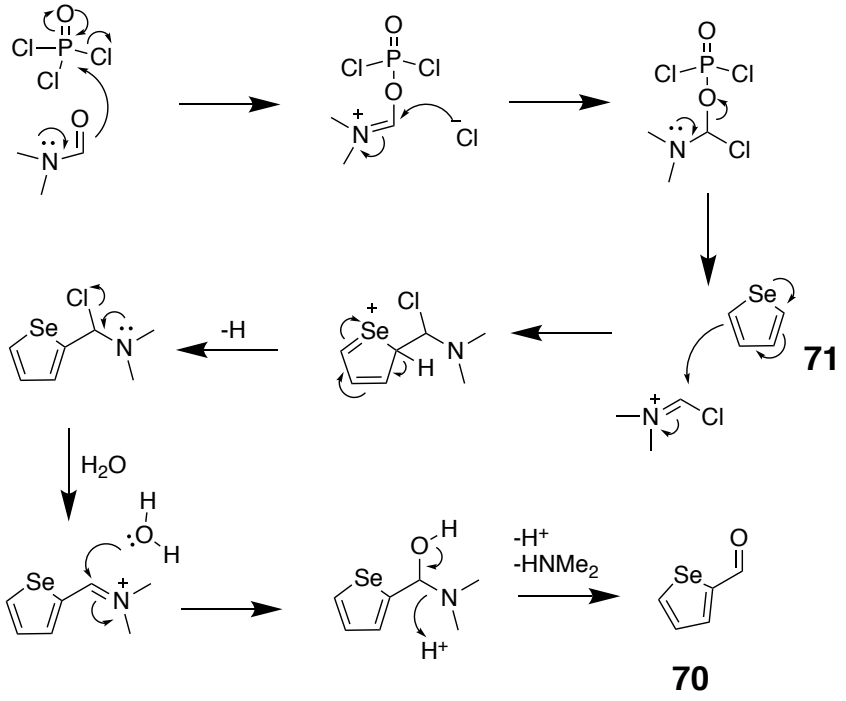
Scheme 4.7 Vilsmeier reaction to synthesise **70**



71 **70**

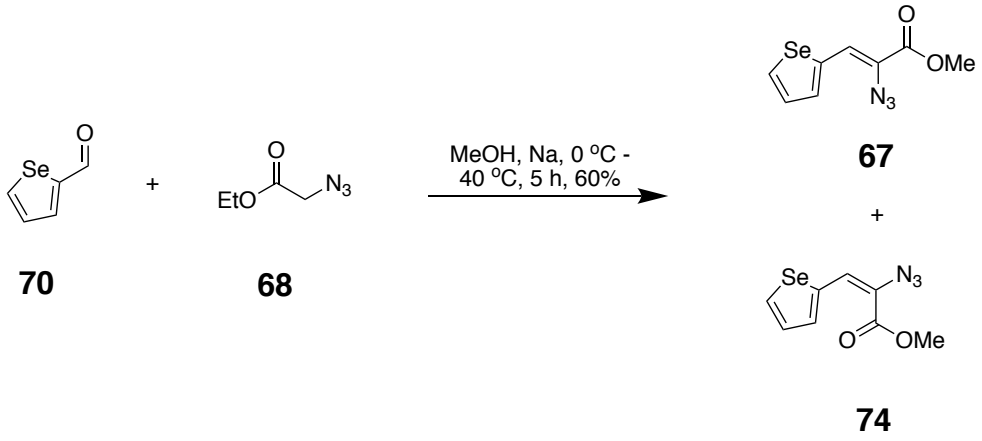
A Vilsmeier reaction involves the formation of an electrophilic intermediate by the reaction of DMF with phosphorus oxychloride, $POCl_3$. This intermediate is then electrophilic enough to react with selenophene (or other aromatic species such as thiophene, pyrrole or furan). The mechanism of the Vilsmeier reaction is shown in Scheme 4.8.

Scheme 4.8 Vilsmeier reaction mechanism



Once **70** has been formed, the Knoevenagel condensation with **68** can be carried out (Scheme 4.9).

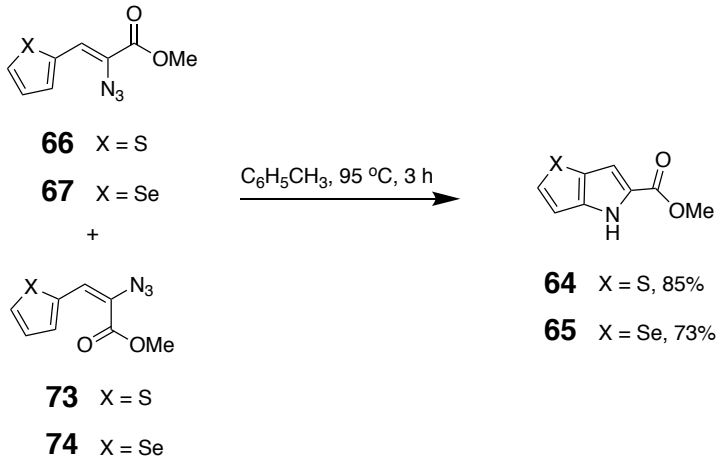
Scheme 4.9 Knoevenagel condensation to synthesise **67** and **74**



The conditions for this reaction were the same as those to synthesise **66**. The Knoevenagel condensation of **70** and **68** gave slightly lower yields (41-60%) than observed under the same reaction conditions using **69**. This decline in yield with a reaction in the scheme involving the seleno-derivative was observed throughout the project, and can most likely be ascribed to the tendency of selenium containing compounds to be less stable than their sulfur analogues,[71] resulting in the formation of side products during each reaction and loss of yield. Again, the reaction produced *cis* and *trans* products, **67** and **74**.

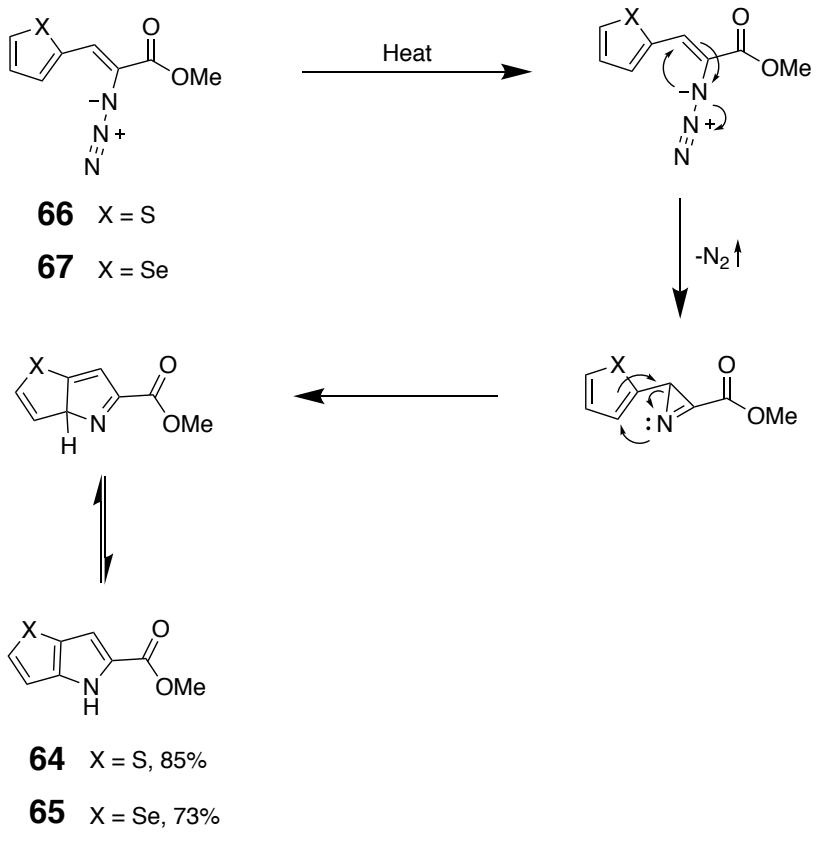
The next step in the synthetic scheme is heating of the products of the Knoevenagel condensations (**66** & **73**, **67** & **74**) to release N_2 gas and cause cyclisation via a Hemetsberger reaction (Scheme 4.10, d). Xylenes is used as solvent in the published protocol,[48] but was substituted for toluene in this project, due to the ease of removal of toluene over xylenes.

Scheme 4.10 Hemetsberger reaction to synthesise **64** and **65**



The Hemetsberger reaction takes advantage of thermal decomposition of azides and is often employed in indole synthesis.[72] An advantage of the Hemetsberger reaction is that it can be effected simply with heating, but care must be taken in preparation of large amounts of possibly explosive azide starting materials. The reaction is thought to proceed through a mechanism illustrated in Scheme 4.11.

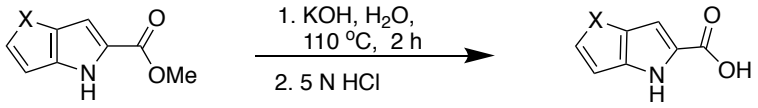
Scheme 4.11 Proposed mechanism of the Hemetsberger reaction



Upon heating, the loss of N_2 gas occurs and can be observed as it bubbles out of the reaction mixture. The proposed Hemetsberger mechanism proceeds through a nitrene intermediate that first adds to the double bond of **66**/**67** then rearranges to form the aromatic **64**/**65**.

In order to deprotect the ester, a suspension of **64** in water is refluxed in an excess of KOH (Scheme 4.12, c). After 2 h, near quantitative conversion takes place and the carboxylic acid product can be precipitated out of the reaction mixture using 5 N HCl to adjust the pH to 1-3 and protonate the carboxylate to give the product at a yield of 90-95%. Similarly, the Se containing ester, **65**, can be readily hydrolysed under the same conditions (66-73%).

Scheme 4.12 Base-catalysed ester hydrolysis

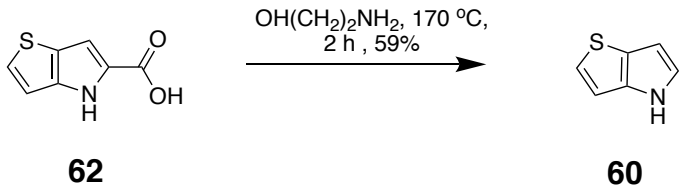


64 X = S
65 X = Se

62 X = S, 90-95%
63 X = Se, 66-73%

The decarboxylation reaction proved to be somewhat of a bottleneck in both syntheses, due to low yields being attained due to degradation of the very unstable heterocyclic products, **60** and **61** at the necessary temperatures for loss of CO_2. The published protocol reports decarboxylation of **62** in glycerol at 165 °C,[48] but a satisfactory yield was not achieved in our hands (<30%). However, modifying the procedure to decarboxylation at 170 °C in ethanolamine gave a much improved yield of 59% (Scheme 4.13). The increase in yield is most likely due to the acid starting material, **62**, being deprotonated by ethanolamine and becoming polar enough to go into solution. In glycerol, the starting material is present as a suspension. Furthermore, ethanolamine is known to absorb CO_2 as a carbamate,[48] which will promote the loss of CO_2. Also, the higher electron density of the deprotonated carboxylate may facilitate decarboxylation.

Scheme 4.13 Decarboxylation of **62**



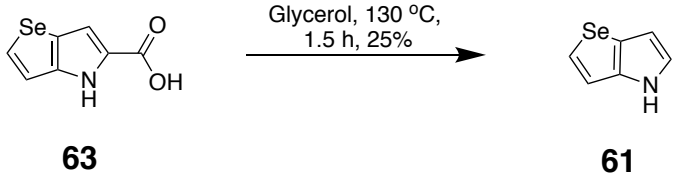
62 **60**

The very unstable S-heterocycle product, **60**, degrades readily[48] and one suspected method of degradation is via polymerisation. In order to minimise polymerisation, the decarboxylation reaction was carried out using large amounts of ethanolamine to reach a final concentration of **60** of 0.1 M.

Decarboxylation of the Se carboxylic acid, **63**, is reported to take place in glycerol at 165 °C for 0.75 h giving a yield of 75%,[48] but the yield could not be reproduced in this project. During the reaction, degradation of **61** could be seen during the reaction by the darkening of the reaction mixture resulting in a 27% yield.

Decarboxylation in glycerol of **63** at slightly lower temperatures of 120 °C and 130 °C was then carried out due to the instability of **61** at 165 °C. Decarboxylation of **63** at 120 °C in glycerol for 0.25 h delivered a yield of only 5%, but an increase in temperature to 130 °C and time to 1.5 h gave 25% yield of **61** (Scheme 4.14). Due to time constraints, decarboxylation of **63** has not yet been carried out in ethanolamine, which may offer an increased yield.

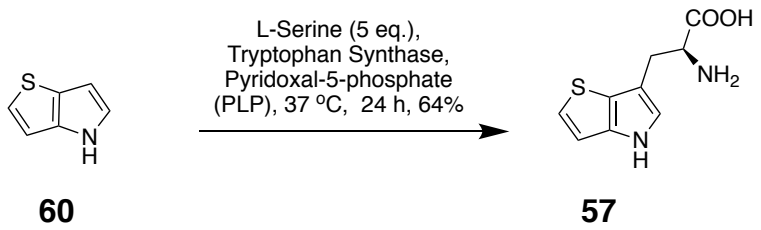
Scheme 4.14 Decarboxylation of **63** in glycerol



63 **61**

The final step of the synthetic scheme is an enzymatic synthesis using tryptophan synthase to convert **60**, in the presence of L-serine and pyridoxal-5-phosphate (PLP) to the final product, **57** (Scheme 4.15). The enzymatic reaction was not carried out for **61**. As this research group is a biochemistry group, we had the capability to produce the enzyme ourselves. The enzyme was produced by Dr. Coilín Boland, and the reaction consistently gave **57** in yields of 58-64%.

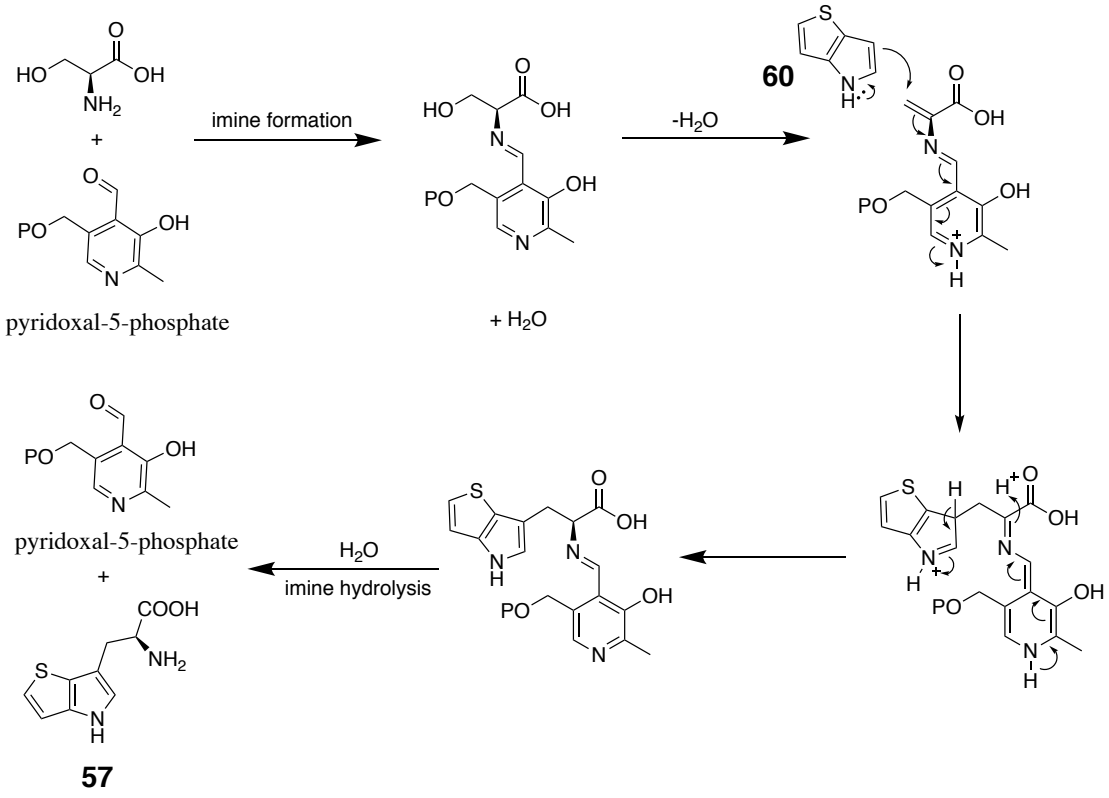
Scheme 4.15 Enzymatic synthesis of **57**



60 **57**

The mechanism of the reaction is shown in Scheme 4.16.

Scheme 4.16 Mechanism of the reaction between **60**, L-serine and PLP catalysed by tryptophan synthase



In the first step of the mechanism, the primary amine of L-serine and the PLP cofactor react to form an imine. Next, a dehydration occurs of the hydroxyl group on the L-serine side chain to form a Michael acceptor than becomes an electrophile for **60**. Subsequent rearrangements and an imine hydrolysis gives the final product **57** and regenerates PLP. Tryptophan synthase catalyses the reaction.

The reaction mixture after work-up consists of the 4 remaining equivalents of L-

serine and **57**. A TLC of the reaction mixture is shown below in Fig 4.3.

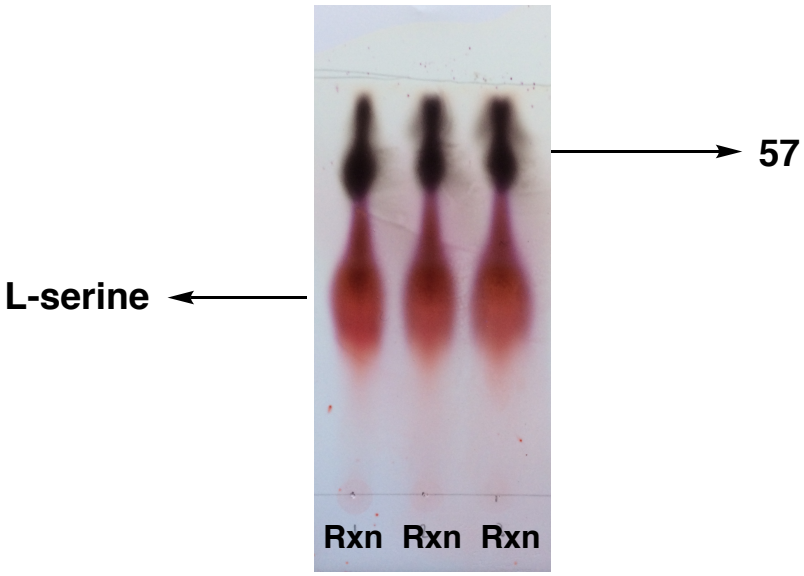


Figure 4.3 TLC of each of the reaction mixture run in Ac_2O/H_2O/AcOH, 12:3:3, after 24 h. TLC stained in ninhydrin and heated to 180 ^oC using a heat-gun.

Compound **57** is too polar to be purified by standard flash chromatography over normal phase silica gel. For that reason, it was decided to purify the crude over a Sephadex G10 resin (formed by the cross linking of dextran and epichlorohydrin specifically for small molecules) size exclusion column.

It was expected that the larger **57** would elute first, followed by the smaller L-serine. Eluting with water separated the two amino acids, surprisingly, L-serine eluted first, followed by **57**. The 4 eq. of L-serine remaining in the reaction mixture most likely form an oligomer that elutes faster through the size exclusion column than **57**. The fractions containing **57** were combined, frozen at -80 ^oC and lyophilised to give **57** as a brown powder.

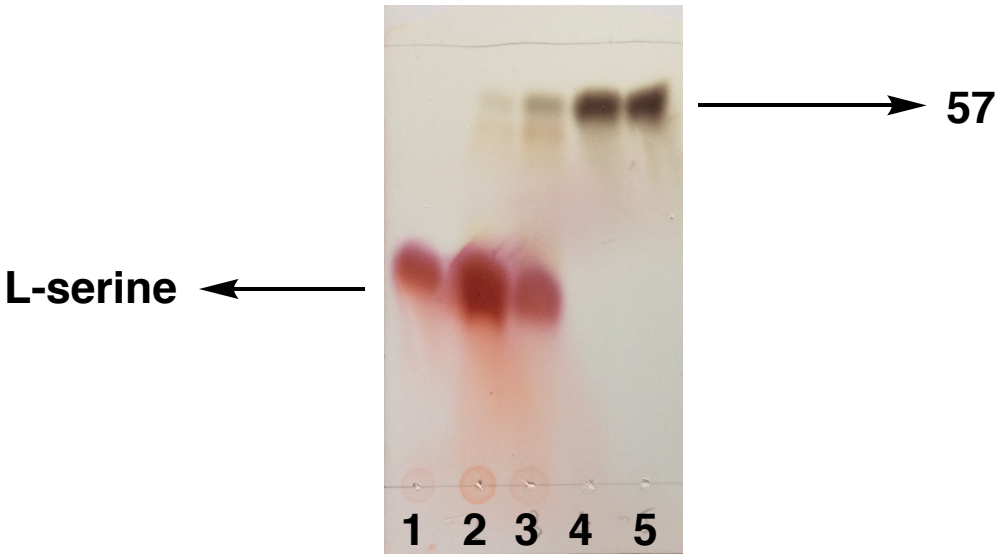


Figure 4.4 TLC showing the fractions of the column (1-5, 25-50 mL fractions). **57** elutes in fractions 4 and 5 during every run. TLC stained in ninhydrin and heated to 180 ^{o}C using a heat-gun. A ^{1}H NMR spectrum on the purified product is shown in Fig 4.5.

^{1}H NMR:

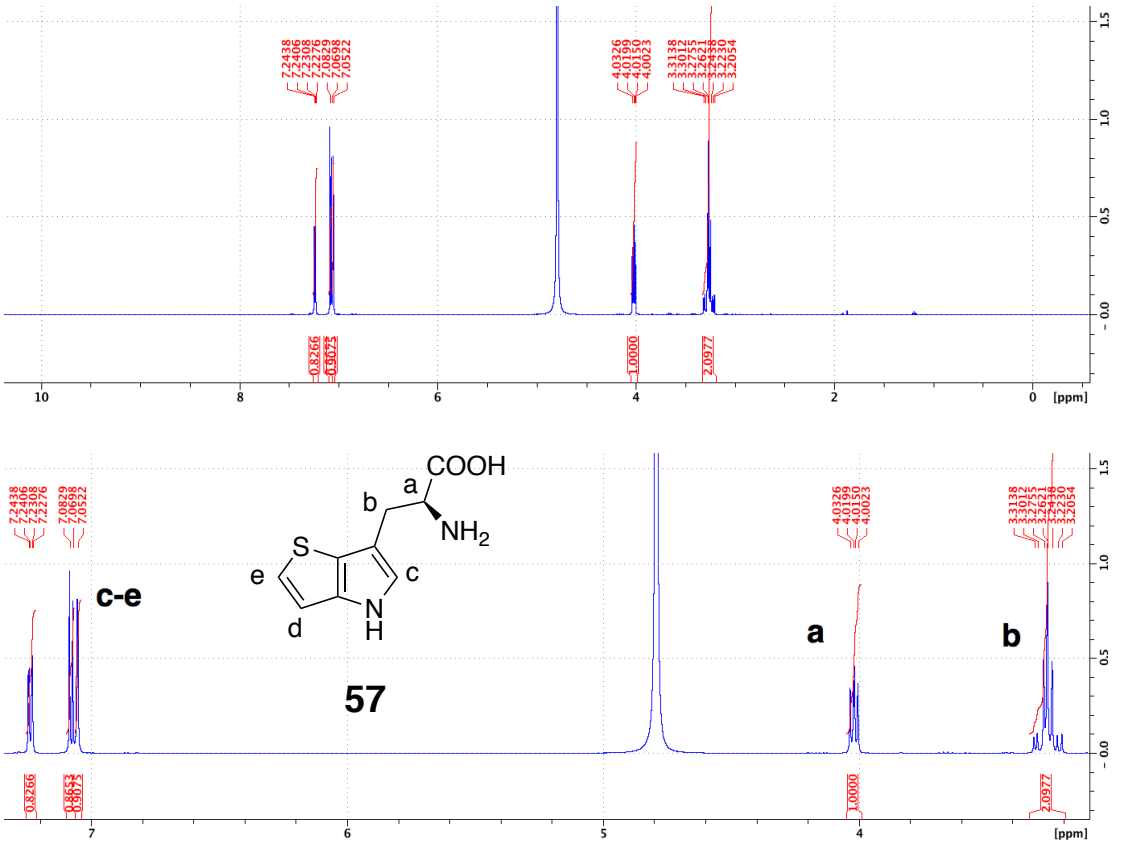


Figure 4.5 ^{1}H NMR of the purified product in D_2O (residual water peak at 4.79 ppm).

The peaks present in the spectrum are labelled and represent a-e as marked on compound **57**. The ^{1}H NMR values correspond to the literature.[70]

4.3 Conclusions and Future Work

The unnatural tryptophan derivative **57** was synthesised according to literature procedures, with minor modifications.[48,70] Low yields in the decarboxylation reaction of **62** that was a bottleneck in the synthesis of **57** were optimised by the change in solvent from glycerol to ethanolamine. This optimization allowed access to large quantities of **57**. Compound **57** can now be incorporated into proteins in order to solve the phase problem using S-SAD. In this technique, proteins are expressed in a tryptophan free medium, with an exess of **57** present. If phasing using **57** proves to be successful, the method may take over from phasing using Se-Met incorporation.

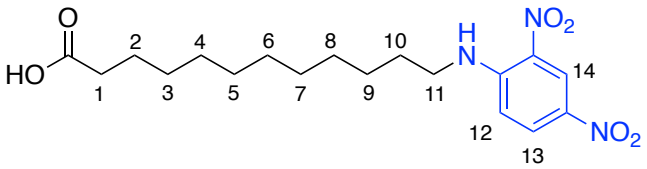
Chapter 5: Experimental Section

5. Materials and Methods

All commercial chemicals were obtained from Sigma-Aldrich, Fisher, Fluorochem or Fluka and used without further purification. Deuterated solvents for NMR use were purchased from Apollo. Analytical TLC was performed using Merck Kieselgel 60 F254 silica gel plates. Visualisation was by UV light (254 nm) using a UVGL-58 handheld UV lamp or by staining with ammonium molybdate (ammonium molybdate 5 g, ceric sulfate 0.2 g, 5% aq. H_2SO_4 100 mL) and heating using a heat-gun. Flash chromatography was carried out using silica gel, particle size 0.04 – 0.063 mm. NMR spectra were recorded on Agilent DD2/LH or Bruker Advance II/DPX spectrometers, operating at 400.13 MHz and 600.1 MHz for ^{1}H NMR; 100.6 MHz and 150.9 MHz for ^{13}C-NMR. Shifts are referenced to the internal solvent signals. NMR data were processed using TopSpin software. HRMS spectra were measured on a Micromass LCT electrospray TOF instrument with a WATERS 2690 autosampler and methanol/acetonitrile as carrier solvent. Infrared spectra were recorded on a Perkin Elmer Spectrum One FT-IR Spectrometer equipped with a Universal ATR sampling accessory.

5.1 Procedures and Data

*12-((2,4-dinitrophenyl)amino)dodecanoic acid (**7**)*



To 12-aminododecanoic acid (1.0 g; 4.6 mmol) suspended in EtOH (18 mL) was added triethylamine (1.40 g; 13.8 mmol) and 1-chloro-2,4-dinitrobenzene (0.94 g; 4.6 mmol). The reaction mixture was heated to 105 ^{o}C for 4 h. The EtOH was then removed under vacuum. The crude was re-dissolved in water (25 mL) and transferred to a separating funnel. The pH was adjusted to 1-3 using 5 N HCl. The product was extracted into DCM (10 x 10 mL), dried over $MgSO_4$ and concentrated under vacuum. The crude was purified over silica gel eluting with EtOAc/AcOH, 9:1, to yield the product as a yellow solid (1.66 g; 95%).

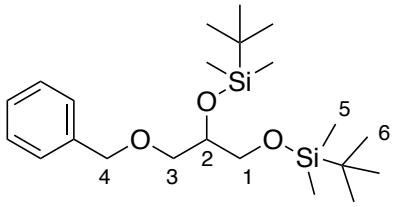
^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 1.28 (12H, m, H3-H8), 1.43 (2H, m, H9), 1.61 (2H, quint, J=7.4 Hz, H2), 1.77 (2H, quint, J=7.4 Hz, H10), 2.33 (2H, t, J=7.4 Hz, H1), 3.40 (2H, m, H11), 6.91 (1H, d, J=9.6 Hz, H12), 8.25 (1H, dd J=9.6, 2.6 Hz, H14), 8.55 (1H, s, NH), 9.11 (1H, d, J=2.6 Hz, H13), 10.18 (1H, s, OH).

^{13}C NMR (100 MHz, $CDCl_3$, *ppm*): 24.68, 26.92, 28.69, 29.02, 29.16, 29.19, 29.33, 29.39, 29.41, 34.14, 43.64, 113.94, 124.37, 130.23 (qC), 130.34, 135.90 (qC), 148.42 (qC), 180.14 (C=O).

HRMS (m/z ES^{+}): Found 404.1798 (M^{+} + Na. $C_{18}H_{27}N_3NaO_6$: Calculated 404.1792).

ν_{max} (cm^{-1}): 2928 (CH), 1707 (COOH), 1325 (C=C).

*5-((benzyloxy)methyl)-2,2,3,3,8,8,9,9-octamethyl-4,7-dioxa-3,8-disiladecane (**13**)*



To 3-(benzyloxy)propane-1,2-diol (1 g; 5.5 mmol) in anhydrous CH_2Cl_2 (8 mL) at 0^{o}C was added 2,6-lutidine (1.8 g; 16.5 mmol). *tert*-Butyldimethylsilyl trifluoromethanesulfonate (3.2 g; 12.1 mmol) was then added dropwise over 5 minutes. The reaction mixture was stirred for 1 h at 0 ^{o}C. After this time, the solvent was evaporated using a stream of N_2 gas and the crude was purified over silica gel eluting with hexanes/EtOAc, 49:1, to yield the product as a colourless oil (2.2 g; 97%).

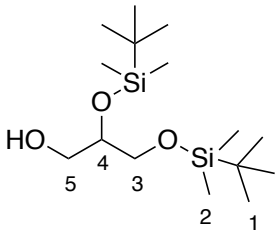
^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.10 (12H, m, H5), 0.93 (18H, s, H6), 3.47 (1H, dd, J=4.3, 5.5 Hz, H3), 3.62 (3H, m, H1 & H3), 3.90 (1H, m, H2), 4.58 (2H, s, H4), 7.37 (5H, m, Ar-H).

^{13}C NMR (101 MHz, $CDCl_3$, *ppm*): -5.38, -5.32, -4.65, -4.58, 18.23 (qC), 18.37 (qC), 25.90, 25.98, 65.11 (CH_2), 72.25 (CH_2), 72.79 (CH), 73.36 (CH_2), 127.42, 127.56, 128.28, 138.60.

HRMS (m/z ES$^+$): Found 411.2750 (M^+ + H. $C_{22}H_{43}O_3Si_2$: Calculated 411.2745)

ν_{max} (cm^{-1}): 2928 (CH), 1251, 829, 774

*2,3-bis((tert-butyldimethylsilyl)oxy)propan-1-ol (**11**)*

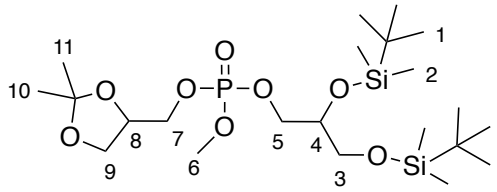


To **13** (500 mg; 1.2 mmol) was added MeOH (8 mL). The reaction mixture was placed under inert atmosphere (Ar gas) using a balloon. Palladium on carbon (10% loading, 140 mg) was added through a briefly opened septum and the reaction mixture was evacuated and refilled with H_2 gas (3x). After 2 h, the reaction mixture was filtered through celite and purified over silica gel eluting with hexanes:EtOAc, 19:1, to yield the product as a colourless oil (280 mg; 73%).

^{1}H NMR: (400 MHz, $CDCl_3$m *ppm*): 0.08 (12H, m, H2), 0.89 (18H, s, H1), 3.62 (4H, m, H3 & H5), 3.77 (1H, m, H4).

Analytical data corresponds to the literature.[73]

*2,3-bis((tert-butyldimethylsilyl)oxy)propyl ((2,2-dimethyl-1,3-dioxolan-4-yl)methyl) methyl phosphate (**9**)*



To methyl dichlorophosphite ($MeOPCl_2$; 1.2 g; 9.0 mmol) and triethylamine (NEt_3; 2.09 g; 20.7 mmol) in anhyd. THF (30 mL) was added **11** (0.78 g; 2.4 mmol) in anhyd. THF (30 mL) dropwise over 3 h at 0 ^oC. After this time, **12** (2.43 g; 19.0 mmol) was added over 5 minutes at 0 ^oC. After a further 3 h, mCPBA (meta chloro peroxybenzoic acid; 2.43 g; 14.1 mmol) was added through a briefly removed septum and the reaction mixture was stirred at 0 ^oC for 1 h. The reaction mixture was then quenched using NaS_2O_3 (10% aq. soln., 165 mL) and $NaHCO_3$ (sat. soln., 45 mL) and extracted into DCM (3 x 150 mL). The organic phases were dried over $MgSO_4$ and concentrated under vacuum to yield a yellow oil that was purified by flash chromatography over silica gel eluting with hexanes/EtOAc, 7:3, to yield the product as a colourless oil (0.78 g; 65%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.09 (12H, m, H2), 0.90 (18H, s, H1), 1.37 (3H, s, H10/H11), 1.44 (1H, s, H10/H11), 3.57 (2H, m, H3), 3.83 (5H, m, H6, H9 & H4), 4.06 (5H, m, H5, H7 & H9), 4.32 (1H, m, H8).

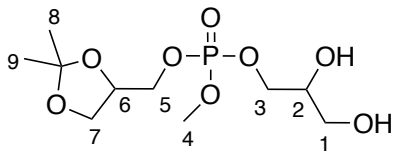
^{13}C NMR (100 MHz, $CDCl_3$, *ppm*): -5.47, -5.43, -4.78, -4.70, 18.09 (qC), 18.28 (qC), 25.27, 25.75, 25.87, 26.72, 54.42 (P-O-$\underline{CH_3}$), 64.03, 66.16, 67.47, 69.13, 71.97, 74.04, 109.88 (qC).

^{31}P NMR (100 MHz, $CDCl_3$, *ppm*): 0.16 (s)

HRMS (m/z ES$^+$): Found 529.2775 (M^+ + H. $C_{22}H_{50}O_8PSi_2$: Calculated 529.2776).

ν_{max} (cm^{-1}): 2954 (CH), 1022, 832, 775.

*2,3-dihydroxypropyl ((2,2-dimethyl-1,3-dioxolan-4-yl)methyl) methyl phosphate (**6**)*



To **9** (70 mg; 0.13 mmol) was added MeOH (5 mL) and ammonium fluoride (70 mg; 1.82 mmol). The reaction was heated to 60 ^oC for 11 h. After this time, the solvent

was removed under a stream of N_2 gas and the crude was purified over silica gel eluting with CH_2Cl_2/MeOH, 19:1, to give the product as a colourless oil (16 mg; 39%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 1.38 (3H, s, H8/H9), 1.46 (3H, s, H8/H9), 3.69 (2H, m, H1), 3.84 (4H, m, H4 & H7), 3.94 (1H, m, H2), 4.12 (5H, m, H3, H5 & H7), 4.35 (1H, m, H6).

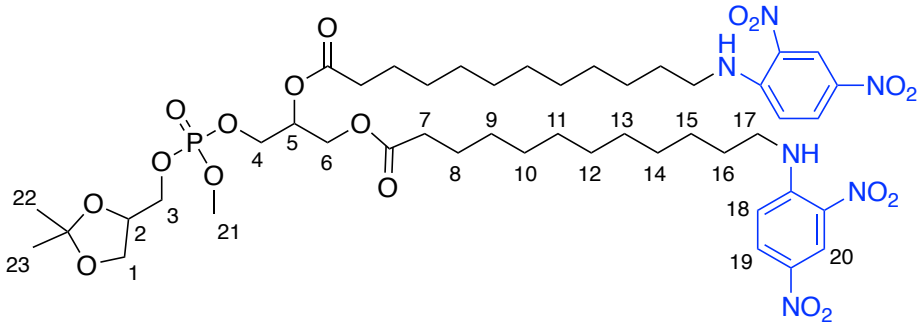
^{13}C NMR (100 MHz, $CDCl_3$, *ppm*): 25.20, 26.67, 54.75, 62.64, 65.83, 67.96, 68.90, 70.58, 74.05, 110.10 (qC).

^{31}P NMR (100 MHz, $CDCl_3$, *ppm*): 1.26 (s).

HRMS (m/z ES^{+}): Found 301.1038 (M^{+} + H. $C_{10}H_{22}O_8P$: Calculated 301.1046).

ν_{max} (cm^{-1}): 3391 (OH), 2955 (CH), 1012 (P=O).

*3-((((2,2-dimethyl-1,3-dioxolan-4-yl)methoxy)(methoxy)phosphoryl)oxy)propane-1,2-diyl bis(12-((2,4-dinitrophenyl)amino)dodecanoate) (**16**)*



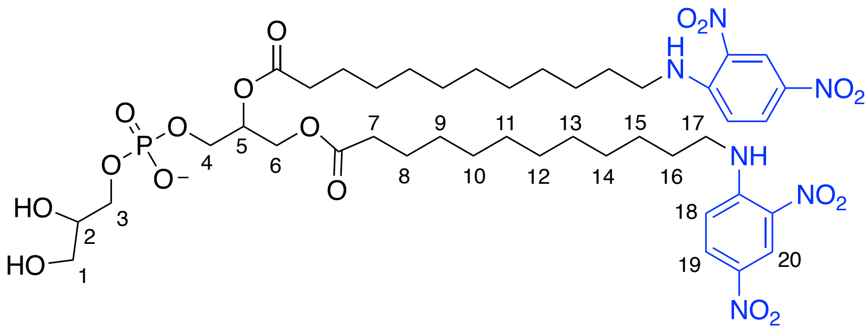
To **6** (20 mg; 0.07 mmol) was added anhydrous DCM (1 mL), DMAP (1 mg; 8 uM) and **7** (64 mg; 0.17 mmol) at 0 ^{o}C. DCC (52 mg; 0.3 mmol) in anhydrous DCM (1 mL) was added over 5 min. The reaction mixture was then removed from the ice bath and left stirring at room temp. (20 ^{o}C) for 18 h. The crude was purified by flash chromatography over silica gel eluting with hexanes/EtOAc/AcOH, 5:4.9:0.1 to give the product as a yellow solid (20 mg; 29%).

^{1}H NMR (600 MHz, $CDCl_3$, *ppm*): 1.28 (24H, s, H9-H14), 1.35 (3H, s, H22/H23), 1.42 (3H, s, H22/H23), 1.46 (4H, m, H15), 1.60 (4H, m, H8), 1.77 (4H, quint, J= 7.4 Hz, H16), 2.32 (4H, m, H7), 3.40 (4H, m, H17), 3.80 (4H, m, H21 & H1), 4.05 (3H, m, H3 & H1), 4.19 (3H, m, H6 & H4), 4.32 (2H, m, H2 & H4), 5.23 (1H, m, H5), 6.92 (2H, d, J=9.6 Hz, H18), 8.27 (2H, dd, J=2.6, 9.6 Hz, H19), 8.55 (2H, s, 2xNH), 9.14 (2H, d, J=2.6 Hz, H20).

^{31}P NMR (100 MHz, $CDCl_3$, *ppm*): 0.01 (s).

HRMS (m/z, ES$^+$): Found 1049.4439 (M^+ + Na. $C_{46}H_{71}N_6NaO_{18}P$: Calculated 1049.4454).

*2,3-bis((12-((2,4-dinitrophenyl)amino)dodecanoyl)oxy)propyl (2,3-dihydroxypropyl) phosphate (**1**)*



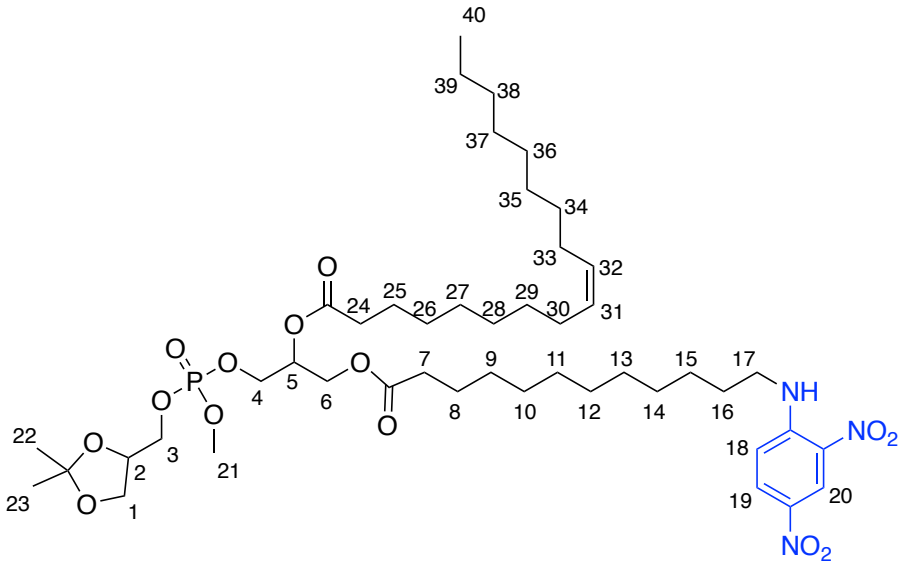
To **15** (20 mg; 0.020 mmol) was added 2-butanone (1.5 mL) and NaI (6 mg; 0.040 mmol). The flask was placed in a pre-heated oil bath at 80 ^{o}C under Ar. After 1.75 h, the reaction was removed from the oil bath and allowed to cool. The solvent was removed under a stream of N_2 gas and the crude product (yellow powder) was redissolved in $CHCl_3$/MeOH, 5:1 (12 mL) and washed with 2 N HCl (15 mL) and sat. aq. NaCl (15 mL). The combined aqueous layers were washed with $CHCl_3$/MeOH, 5:1 (3 x 12 mL). The combined organic layers were dried over $MgSO_4$ and concentrated under vacuum to yield a yellow solid. The yellow solid was re-dissolved in CH_2Cl_2/TFA/MeOH, 6:3:1 (1 mL). The mixture was stirred for 0.5 h. After this time, the reaction was stopped after removal of the CH_2Cl_2/TFA/MeOH, 6:3:1, under a stream of N_2 gas. The crude was purified over silica gel eluting with $CHCl_3$/MeOH, 1:1, to give the product as a yellow solid (4 mg; 21%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 1.24 (24H, s, H9-H14), 1.42 (4H, m, H15), 1.56 (4H, m, H8), 1.76 (4H, m, H16), 2.29 (4H, m, H7), 3.40 (4H, m, H17), 4.03 (9H, m, H1-H4 & H6), 5.22 (1H, s, H5), 6.92 (2H, d, J=9.6 Hz, H18), 8.25 (2H, dd, J=2.6, 9.6 Hz, H19), 8.55 (2H, s, 2xNH), 9.10 (2H, d, J=2.6 Hz, H20).

^{31}P NMR (100 MHz, $CDCl_3$, *ppm*): -1.12 (bs).

HRMS (MALDI): Found 971.4056 (M^-. $C_{42}H_{64}N_6O_{18}P$: Calculated 971.4015).

1-((((2,2-dimethyl-1,3-dioxolan-4-yl)methoxy)(methoxy)phosphoryl)oxy)-3-((12-((2,4-dinitrophenyl)amino)dodecanoyl)oxy)propan-2-yl oleate ***(5)***



To **6** (25 mg; 0.083 mmol) was added anhydrous CH_2Cl_2 (1 mL), DMAP (1 mg; 0.008 mmol) and **7** (34 mg; 0.087 mol) at 0 ^{o}C. DCC (33 mg; 0.16 mmol) in anhydrous CH_2Cl_2 (1 mL) was added over 5 min. The reaction was allowed to stir for 4 h at 0 ^{o}C. After this time **8** (31 mg; 0.11 mmol) and DCC (33 mg; 0.16 mmol) was added. The reaction mixture was left stirring for 18 h at 20 ^{o}C. The crude was purified by flash chromatography over silica gel eluting with hexanes/EtOAc/AcOH, 5:4.9:0.1 to give the product as a yellow solid (20 mg; 33%).

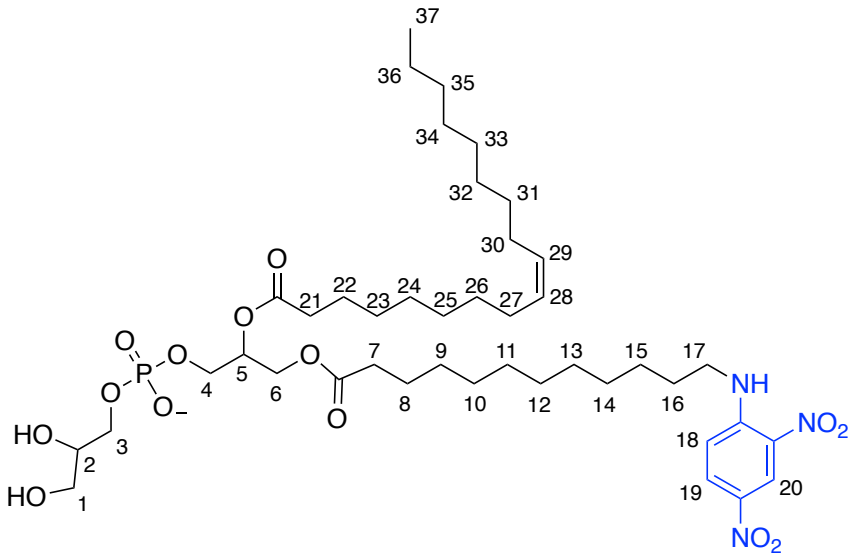
^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.88 (3H, t, J=6.7 Hz, H40), 1.28 (32H, m, H9-H14, H26-H29, H34-H39), 1.36 (3H, s, H22/H23), 1.43 (3H, s, H22/H23), 1.47 (2H, m, H15), 1.60 (4H, m, H8 & H25), 1.78 (2H, m, H16), 2.01 (4H, m, H30 & H33),

2.32 (4H, m, H7 & H24), 3.40 (2H, m, H17), 3.80 (4H, m, H21 & H1), 4.06 (3H, m, H3 & H1), 4.18 (3H, m, H6& H4), 4.33 (2H, m, H2 & H4), 5.23 (1H, m, H5), 5.34 (2H, m, H31 & H32), 6.91 (1H, d, J=9.6 Hz, H18), 8.27 (1H, dd, J=2.6, 9.6 Hz, H19), 8.56 (1H, s, NH), 9.15 (1H, d, J=2.6 Hz, H20).

^{31}P NMR (100 MHz, $CDCl_3$, *ppm*): 0.03 (s).

HRMS (m/z ES$^-$): Found 926.5122 (M^- - H. $C_{46}H_{77}N_3O_{14}P$: Calculated 926.5148).

*2,3-dihydroxypropyl (3-((12-((2,4-dinitrophenyl)amino)dodecanoyl)oxy)-2-(oleoyloxy)propyl) phosphate (**2**)*



To **5** (25 mg; 0.027 mol) was added NaI (6 mg; 0.054 mol) and 2-butanone (2.2 mL). The reaction mixture was heated to 80 ^oC under Ar for 2 h, after which time the solvent removed under a stream of N_2 gas. The residue was then dissolved in $CHCl_3$/MeOH, 5:1 (15 mL), washed with 2 N HCl (15 mL) and sat. aq. NaCl (15 mL). The aqueous layers were washed with $CHCl_3$/MeOH, 5:1 (2 x 15 mL). The organic layers were combined, dried over $MgSO_4$ and concentrated under vacuum to give a yellow oil. This yellow oil was then dissolved in CH_2Cl_2/CF_3COOH/MeOH, 6:3:1 (1.4 mL) and stirred for 0.5 h at 20 ^oC. The solvents were co-evaporated with toluene under a stream of N_2 gas and purified over silica gel eluting with $CHCl_3$/MeOH, 7:3, to yield the product as a yellow oil (19 mg; 71%) from which 12

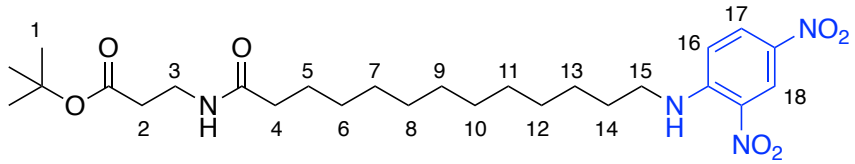
mg was purified by preparative TLC, eluting with $CHCl_3/Ac_2O/MeOH/H_2O$, 10:4:2:2:1, to yield **2** as a yellow oil (10.5 mg).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.86 (3H, t, J=5.9 Hz, H37), 1.26 (32H, m, H9-H14, H23-H26, H31-H36), 1.43 (2H, m, H15), 1.57 (4H, m, H8 & H22), 1.77 (2H, m, H16), 1.99 (4H, m, H27 & H30), 2.29 (4H, m, H7 & H21), 3.40 (2H, m, H17), 4.05 (9H, m, H1-H4 & H6), 5.21 (1H, s, H5), 5.32 (2H, m, H28 & H29), 6.92 (1H, d, J=9.6 Hz, H18), 8.26 (1H, dd, J=2.6, 9.6 Hz, H19), 8.55 (1H, s, NH), 9.12 (1H, d, J=2.6 Hz, H20).

^{31}P NMR (100 MHz, $CDCl_3$, *ppm*): -3.32 (bs).

HRMS (m/z ES$^-$): Found 872.4672 (M^- - H. $C_{42}H_{64}N_6O_{18}P$: Calculated 872.4679).

*tert-butyl 3-(12-((2,4-dinitrophenyl)amino)dodecanamido)propanoate (**26**)*



To **7** (200 mg; 0.52 mmol) in anhydrous CH_2Cl_2 (8 mL) was added Mukaiyama's reagent (2-chloro-1-methylpyridinium iodide, 402 mg; 1.57 mmol) at 20 ^oC. After 5 minutes, NEt_3 (320 mg; 3.00 mmol) and **20** (95 mg; 0.52 mmol) were added. The reaction mixture was stirred at 20 ^oC for 3 h. After this time, the solvent was removed under a stream of N_2 gas and the residue was purified by flash chromatography over silica gel eluting with hexanes/EtOAc/AcOH, 5:4.9:0.1 to yield the product as a yellow solid (205 mg; 77%).

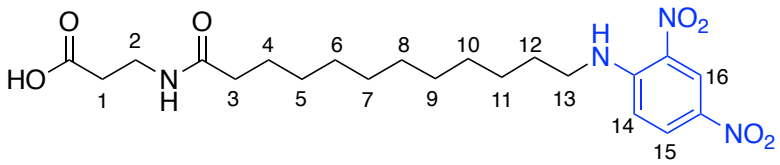
^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 1.28 (12H, s, H6-H11), 1.45 (11H, s, H1 & H12), 1.60 (2H, m, H5), 1.77 (2H, quint, J=7.3 Hz, H13), 2.14 (2H, t, J=7.3 Hz, H4), 2.44 (2H, t, J=5.9 Hz, H2), 3.40 (2H, m, H14), 3.47 (2H, m, H3), 6.05 (NH, s, HN-C=O), 6.91 (1H, d, J=9.6 Hz, H16), 8.28 (1H, dd, J=2.6, 9.6 Hz, H17), 8.55 (1H, s, NH), 9.14 (1H, d, J=2.6 Hz, H18).

^{13}C NMR (100 MHz, $CDCl_3$, *ppm*): 25.70, 26.91, 28.12, 28.69, 29.15, 29.22, 29.30, 29.36, 29.38, 29.42, 34.90, 35.14, 36.84, 43.63, 81.09 (qC), 113.90, 124.42, 130.28, 130.36, 135.95 (qC), 148.41 (qC), 172.21 (C=O), 173.04 (C=O).

HRMS (APCI): Found 509.2974 (M^+ + H. $C_{21}H_{31}N_4O_7$ Requires 598.2969).

ν_{max} (cm^{-1}): 3350 (NH), 2925, 1619 (N-C=O), 1522, 1424, 1287.

*3-(12-((2,4-dinitrophenyl)amino)dodecanamido)propanoic acid (**19**)*



To **26** (200 mg; 0.39 mmol) in CH_2Cl_2 (1.25 mL) at 20 °C was added CF_3COOH (3.75 mL) dropwise over 10 minutes. The reaction mixture was stirred for 3 h. The solvents were removed by co-evaporation with toluene under a stream of N_2 and subsequent drying overnight under vacuum to yield the product as a yellow solid (170 mg; 97%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 1.29 (12H, s, H5-H10), 1.47 (2H, m, H11), 1.61 (2H, m, H4), 1.79 (2H, quint, J=7.3 Hz, H12), 2.20 (2H, t, J=7.3 Hz, H3), 2.61 (2H, t, J=5.3 Hz, H1), 3.43 (2H, m, H13), 3.55 (2H, m, H2), 6.31 (NH, s), 6.94 (1H, d, J=9.6 Hz, H14), 8.28 (1H, dd, J=2.6, 9.6 Hz, H15), 8.58 (NH, s), 9.14 (1H, d, J=2.6 Hz, H16), 9.36 (OH, s).

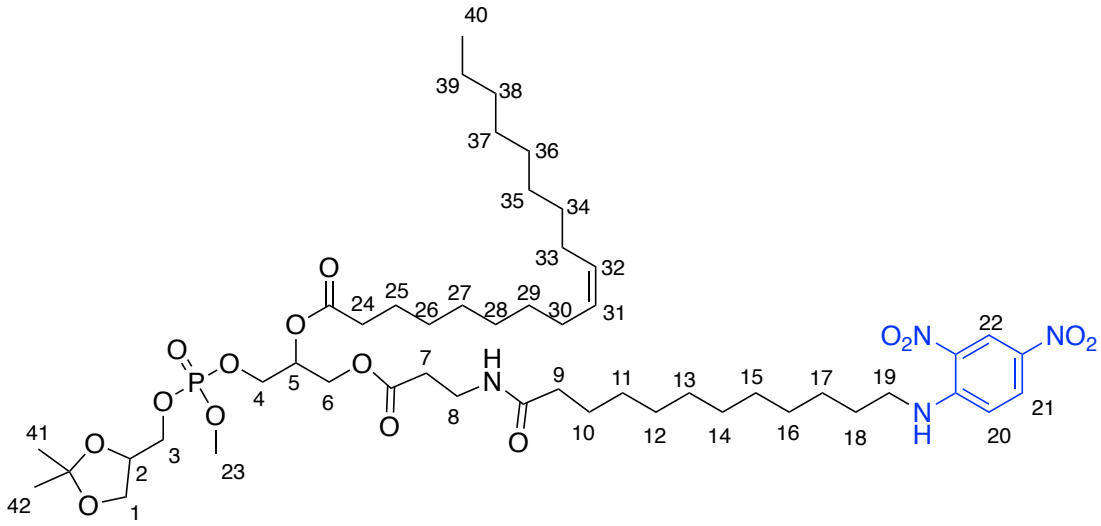
^{13}C NMR (100 MHz, $CDCl_3$, *ppm*): 25.65, 26.88, 28.66, 29.12, 29.17, 29.24, 29.34, 29.35, 29.38, 33.97, 34.89, 36.65, 43.62, 113.98, 124.39, 130.22 (qC), 130.36, 135.90 (qC), 148.44 (qC), 174.19 (C=O), 176.36 (C=O).

HRMS (APCI): Found 451.2195 (M^+ + H. $C_{21}H_{31}N_4O_7$: Calculated 451.2198).

ν_{max} (cm^{-1}): 3348 (OH), 2927, 1707 (C=O), 1618 (N-C=O), 1325.

1-((((2,2-dimethyl-1,3-dioxolan-4-yl)methoxy)(methoxy)phosphoryl)oxy)-3-((3-(12-

*((2,4-dinitrophenyl)amino)dodecanamido)propanoyl)oxy)propan-2-yl oleate (**18**)*



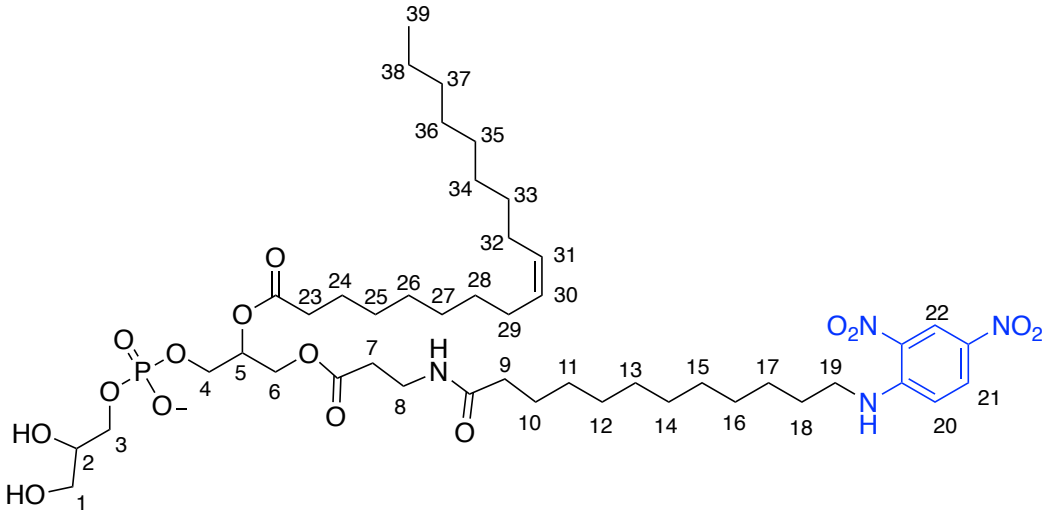
To **6** (25 mg; 0.083 mmol) was added anhydrous CH_2Cl_2 (1 mL), DMAP (1 mg; 0.008 mmol) and **7** (34 mg; 0.087 umol) at 0 ^oC. DCC (33 mg; 0.16 mmol) in anhydrous CH_2Cl_2 (1 mL) was added over 5 min. The reaction was allowed to stir for 4 h at 0 ^oC. After this time **8** (31 mg; 0.108 mmol) and DCC (33 mg; 0.16 mmol) was added. The ice bath was removed and the reaction mixture was left stirring for 18 h at 20 ^oC. The crude was purified by flash chromatography over silica gel eluting with hexanes/EtOAc/AcOH, 5:4.9:0.1 to give the product as a yellow solid (23 mg; 28%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.87 (3H, t, J=6.7 Hz, H40), 1.28 (32H, m, H11-H16, H26-H29, H34-H39), 1.35 (3H, s, H41/42), 1.42 (5H, s, H41/42 & H17), 1.62 (4H, m, H10 & H25), 1.77 (2H, m, H18), 2.00 (4H, m, H30 & H33), 2.15 (2H, t, J=7.6 Hz, H9), 2.31 (2H, m, H24), 2.43 (2H, t, J=7.4 Hz, H7), 3.40 (2H, m, H19), 3.50 (2H, m, H8), 3.80 (4H, m, H23 & H1), 4.05 (3H, m, H3 & H1), 4.17 (3H, m, H6 & H4), 4.33 (2H, m, H2 & H4), 5.23 (1H, m, H5), 5.33 (2H, m, H31 & H32), 6.38 (NH, s), 6.91 (1H, d, J=9.6 Hz, H20), 8.27 (1H, dd, J=2.6, 9.6 Hz, H21), 8.55 (NH, s), 9.14 (1H, d, J=2.6 Hz, H22).

^{31}P NMR: (100 MHz, $CDCl_3$, *ppm*): 0.02 (s).

HRMS (APCI): Found 999.5673 (M^+ + H. $C_{49}H_{84}N_4O_{15}P$: Calculated 999.5663).

2,3-dihydroxypropyl (3-((3-(12-((2,4-dinitrophenyl)amino)dodecanamido)propanoyl)oxy)-2-(oleoyloxy)propyl) phosphate ***(3)***



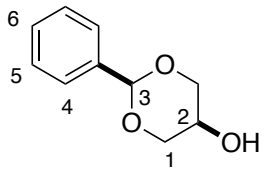
To **18** (23 mg; 0.023 mmol) was added NaI (7 mg; 0.046 mol) and 2-butanone (2.2 mL). The reaction mixture was heated to 80 ^{o}C for 2 h, after which time the solvent removed under a stream of N_2 gas. The residue was then dissolved in $CHCl_3$/MeOH, 5:1 (15 mL), washed with 2 N HCl (15 mL) and sat. aq. NaCl (15 mL). The aqueous layers were washed with $CHCl_3$/MeOH, 5:1 (2 x 15 mL). The organic layers were combined, dried over $MgSO_4$ and concentrated under vacuum to give a yellow oil. This yellow oil was then dissolved in CH_2Cl_2/CF_3COOH/MeOH, 6:3:1 (1.4 mL) and stirred for 1 h at 20 ^{o}C. The solvents were co-evaporated with toluene under a stream of N_2 gas and purified over silica gel eluting with $CHCl_3$/MeOH, 7:3, to yield the product as a yellow oil (9.5 mg; 46%), from which 3.0 mg was purified by preparative TLC, eluting with $CHCl_3$/Ac_2O/MeOH/H_2O, 10:4:2:2:1, to yield **3** as a yellow oil (2.5 mg).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.89 (3H, t, J=6.7 Hz, H39), 1.28 (32H, s, H11-H16, H25-H28, H33-H38), 1.46 (2H, m, H17), 1.60 (4H, m, H24&H10), 1.77 (2H, m, H18), 2.01 (4H, m, H29 & H32), 2.16 (2H, m, H9), 2.32 (2H, m, H23), 2.56 (2H, m, H7), 3.43 (2H, m, H19), 3.51 (2H, m, H8), 3.69-4.40 (9H, m, H1-H4 & H6), 5.17 (1H, m, H5), 5.35 (2H, m, H30&H31), 6.57 (NH, s), 6.96 (1H, d, J=9.6 Hz, H20), 8.29 (1H, dd, J=2.6, 9.6 Hz, H21), 8.59 (NH, s), 9.14 (1H, d, J=2.6 Hz, H22).

^{31}P NMR (100 MHz, $CDCl_3$, *ppm*): -0.5 (bs)

HRMS (MALDI): Found 967.5045 ([M^+ + H] + Na. $C_{45}H_{77}N_4O_{15}NaP$: Calculated 967.5021).

*2-phenyl-1,3-dioxan-5-ol (**45**)*

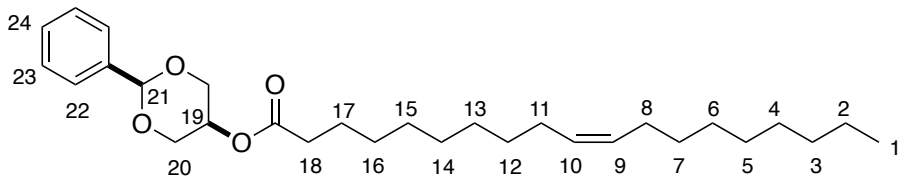


To **47** (20 g; 217 mmol) in pet. ether/benzene (1:1, 30 mL) was added **46** (24 g; 226 mmol) and *p*-toluenesulfonicacid (0.2 g). The reaction mixture was heated to 85 ^oC for 24 h. The reaction mixture was neutralised using 0.5 M KOH and the organic layer was cooled to -10 ^oC using a liq. N_2/MeCN bath. The solid was removed using filtration, then re-suspended in toluene/pet.ether. The mixture was cooled to -10 ^oC and the white solid removed by filtration to give the product as a white solid (2.8 g; 5%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 3.04 (1H, d, J=11.0 Hz, OH), 3.64 (1H, d, J=11.0 Hz, H2), 4.16 (4H, m, H1), 5.56 (1H, s, H3), 7.39 (3H, m, H5&H6), 7.50 (2H, m, H4).

Analytica data corresponds to the literature.[74]

*2-phenyl-1,3-dioxan-5-yl (Z)-nonadec-10-enoate (**44**)*



To **45** (300 mg, 1.7 mmol) was added DMAP (23 mg, 0.2 mmol), anhydrous CH_2Cl_2 (5 mL) and **8** (452 mg, 1.6 mmol). The reaction mixture was placed in an ice bath for 10 min. In a separate flask, DCC (340 mg, 1.7 mmol) was dissolved in anhydrous

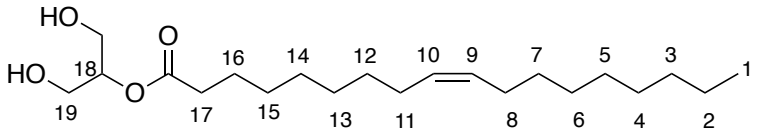
CH_2Cl_2 (4 mL). The DCC solution was added to the reaction flask by syringe over 10 min. The reaction mixture was stirred for 4.5 h at 20 ^oC. The mixture was filtered to remove dicyclohexylurea and concentrated under vacuum. The crude was purified by flash chromatography over silica gel eluting with hexanes/EtOAc, 19:1, to yield the product as an oil (570 mg; 77%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.88 (3H, t, J=6.8 Hz, H1), 1.30 (22H, s, H2-H7, H12-H16), 1.67 (2H, m, H17), 2.00 (4H, m, H8&H11), 2.43 (2H, t, J=7.6 Hz, H18), 4.17 (2H, dd, J=1.5, 12.9 Hz, H20), 4.29 (2H, dd, J=1.1, 12.9 Hz, H20), 4.72 (1H, s, H19), 5.34 (2H, m, H9&H10), 5.56 (1H, s, H21), 7.37 (3H, m, H23&H24), 7.51 (2H, m, H22).

^{13}C NMR (100 MHz, $CDCl_3$, *ppm*): 14.12, 22.69, 24.93, 27.17, 27.21, 29.09, 29.12, 29.20, 29.32, 29.52, 29.71, 29.77, 31.90, 34.38, 65.70, 69.12, 101.21, 126.02, 128.28, 129.06, 129.75 (C=C), 129.97 (C=C), 137.84 (qC), 173.84 (C=O).

ν_{max} (cm^{-1}): 2923, 1731 (C=O), 1084.

*1,3-dihydroxypropan-2-yl oleate (**36**)*



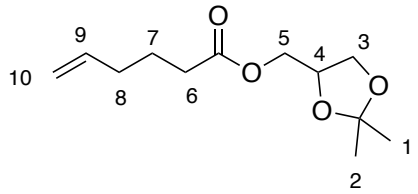
To **44** (600 mg; 1.35 mmol) was added AcOH/H_2O (4:1, v/v, 15 mL) and the reaction mixture was heated to 50 ^oC for 2 h. After this time, the reaction mixture was diluted with ether (5 mL) and neutralised with sat. $NaHCO_3$ (10 mL). The organic layer was dried over $MgSO_4$, concentrated under vacuum and purified over silica gel eluting with hexanes/EtOAc, 1:1, to yield the product as a colourless oil (300 mg; 63%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.88 (3H, t, J=6.8 Hz, H1), 1.29 (20H, m, H2-H7 & H12-H15), 1.65 (2H, m, H16), 2.01 (4H, m, H8&H11), 2.38 (2H, t, J=7.6 Hz, H17), 3.84 (4H, d, J=5.0 Hz, H19), 4.93 (1H, quint, J=5.0 Hz, H19), 5.35 (2H, m, H9&H10).

^{13}C NMR (100 MHz, $CDCl_3$, *ppm*): 14.10, 22.67, 24.94, 27.14, 27.21, 29.06, 29.08, 29.14, 29.30, 29.31, 29.51, 29.67, 29.75, 31.89, 34.32, 62.60, 75.06, 129.70 (C=C), 130.03(C=C), 174.02 (C=O).

HRMS (m/z ES$^+$): Found 357.3001 (M^+ + H. $C_{21}H_{41}O_4$: Calculated 357.2999.

*(2,2-dimethyl-1,3-dioxolan-4-yl)methyl hex-5-enoate (**38**)*

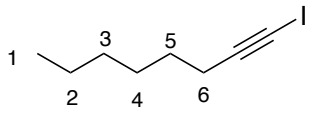


To DMAP (1.17 g; 9.6 mmol) and anhydrous CH_2Cl_2 (DCM, 80 mL) under argon in a 500 mL three-necked RBF was added **39** (10.00 g; 87.6 mmol) and **12** (12.04 g; 90.2 mmol). The reaction mixture was placed at 0 ^oC. After 10 minutes a DCC (18.62 g; 91.1 mmol) solution in CH_2Cl_2 (50 mL) was added dropwise over 1 h. The ice bath was then removed and the reaction mixture was allowed to stir at room temperature for 3.5 h. After this time, the reaction mixture was cooled to 0 ^oC, stirred for 1 h then filtered. The crude was purified by flash chromatography over silica gel, eluting with hexanes/EtOAc, 8.5:1.5 to give the product as a colourless oil (16.4 g; 82%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 1.35 (3H, s, H1/H2), 1.42 (3H, s, H1/H2), 1.73 (2H, quint, J=7.5 Hz, H7), 2.08 (2H, m, H8), 2.34 (2H, t, J=7.5 Hz, H6), 3.72 (1H, dd, J=6.2, 8.5 Hz, H3), 4.10 (3H, m, H3&H5), 4.30 (1H, m, H4), 5.00 (2H, m, H9), 5.75 (1H, m, H10).

Analytical data corresponds to the literature.[60]

*1-iodooct-1-yne (**41**)*



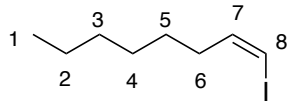
An oven-dried 1 L round bottomed flask (RBF) was charged with H_2O (200 mL), TBAB (21.85 g; 67.8 mmol) and CuI (1.5 mol%; 0.19 g; 0.7 mmol) producing a colourless solution. This was allowed to stir vigorously for 10 minutes. To this 1 L

RBF was added I_2 (34.40 g; 135.6 mmol) and NEt_3 (20.58 g; 28 mL; 203.4 mmol). This was again allowed to stir for 10 minutes, after which **42** (7.47 g; 10 mL; 67.8 mmol) was added. The 1 L RBF was stirred vigorously for 24 h at room temperature. The reaction mixture was transferred to a 1 L separating funnel and extracted with EtOAc (3 x 150 mL). The organic layers were washed with saturated aqueous sodium thiosulfate (3 x 100 mL). The washed organic layers were then dried over $MgSO_4$ and concentrated under vacuum. The crude was purified by flash chromatography over silica gel, eluting with hexane to give the product as an orange oil (15 g; 94%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.87 (3H, t, J=6.9 Hz, H1), 1.31 (6H, m, H2-H4), 1.49 (2H, m, J=7.24 Hz, H5), 2.34 (2H, t, J=6.9 Hz, H6).

Analytical data corresponds to the literature.[61]

*(Z)-1-iodooct-1-ene (**40**)*

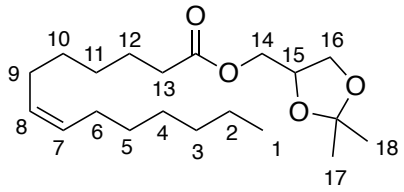


To a three-necked 250 mL round bottomed flask under argon was added anhydrous ether (60 mL) and BH_3.DMS complex (4.6 g; 60 mmol). The reaction mixture was then placed in an ice bath. After 10 min, cyclohexene (10.8 g; 132 mmol) was added dropwise over 10 min via plastic syringe. The reaction mixture was allowed to stir in ice bath for a further 15 min, then at 20 ^oC for 1 h. The reaction mixture was then placed in the ice bath for 5 min, before addition of **41** (13 g; 53 mmol), slowly over 10 min. The mixture was stirred at 0 ^oC for 0.5 h and 20 ^oC for 2 h. Next, acetic acid (16 eq.) was added slowly at 0 ^oC over 15 min, the reaction mixture is then left at 20 ^oC for 2.5 h. The reaction mixture was then diluted with diethyl ether (30 mL) and water (50 mL). The aqueous layer is discarded and the organic layer washed with water (3 x 60 mL). The organic layer was dried over $MgSO_4$ and swirled with ethanolamine (7 mL) to precipitate out the residual borane in the mixture. The white borane-ethanolamine precipitate was removed by filtration and the organic layers concentrated under vacuum to give the product as a yellow oil (9.1 g; 71%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.88 (3H, m, H1), 1.33 (6H, m, H2-H4), 2.12 (2H, m, H6), 6.15 (2H, m, H7&H8).

Analytical data corresponds to the literature.[60]

*(2,2-dimethyl-1,3-dioxolan-4-yl)methyl (Z)-tetradec-7-enoate (**37**)*

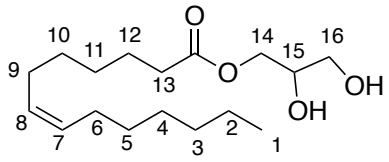


To an oven-dried 500 mL round-bottomed-flask under argon was added **38** (6.80 g; 29.8 mmol) and anhydrous THF (80 mL). The flask was placed in an ice-bath and 9-BBN from a 0.5 M soln. in THF (65.5 mL; 8.00 g; 32.8 mmol) was added. The reaction mixture was then stirred at 20 ^{o}C for 18 h. After this time, deionised water (90 mL) was added and the reaction mixture was stirred for 2 h then sparged with argon for 15 minutes. To a separate 1 L round-bottomed-flask under argon was added **40** (8.5 g; 35.7 mmol), cesium carbonate (9.71g; 29.8 mmol), triphenylarsine (0.91 g; 2.9 mmol) and DMF (140 mL). This mixture was sparged with argon for 15 minutes, then Pd(dppf)Cl_2.CH_2Cl_2 (1.23 g; 1.5 mmol) was added. The resultant red solution was again sparged with argon for 15 minutes, after which time the solution turned brown. The hydroboration mixture was then added via plastic syringe over 35 minutes and left stirring at 20 ^{o}C for 18 h. The reaction mixture was then transferred to a 2 L separating funnel containing 600 mL of brine. The aqueous was washed with ether (3 x 300 mL). The combined organic layers were dried over $MgSO_4$ and concentrated under vacuum to give a dark brown/black oil (6.50 g, 68%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.86 (3H, t, J=6.8 Hz, H1), 1.25 (12H, s, H2-H5, H10-H11), 1.35 (3H, s, H17/H18), 1.42 (3H, s, H17/H18), 1.61 (2H, m, H12), 1.99 (4H, m, H6&H9), 2.33 (2H, t, J=7.6 Hz, H13), 3.72 (1H, dd, J=6.2, 8.5 Hz, H16), 4.15 (3H, m, H16&H14), 4.30 (1H, m, H15), 5.33 (2H, m, H7&H8).

Analytical data corresponds to the literature.[60]

*2,3-dihydroxypropyl (Z)-tetradec-7-enoate (**34**)*



To **37** (6.10 g; 17.9 mmol) was added acetic acid/water (4:1, v/v; 200 mL). The reaction mixture was heated to 50 ^{o}C for 2 h. The reaction mixture was then added to sat. $NaHCO_3$ mixture (2 L) and ether (300 mL). The organic layer was concentrated under vacuum to give a yellow oil. The crude was purified via chromatography over silica gel, eluting with hexanes/EtOAc, 1:1, to give the product as a colourless oil (2.5 g; 47%). The oil was then recrystallised from MeCN at -20 ^{o}C to give **34** with a lower 2-MAG content (1.0 g).

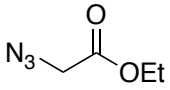
^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 0.88 (3H, t, J=6.7 Hz, H1), 1.32 (12H, m, H2-H5, H10-H11), 1.64 (2H, quint, J=7.2 Hz, H12), 2.01 (4H, m, H6&H9), 2.35 (2H, t, J=7.2 Hz, H13), 3.60 (1H, dd, J=5.8, 11.4 Hz, H16), 3.70 (1H, dd, J=4.0, 11.4 Hz, H16), 3.93 (1H, m, H15), 4.18 (2H, m, H14), 5.34 (2H, m, H7&H8).

^{13}C NMR (100 MHz, DMSO-d_6, *ppm*): 13.91, 22.05, 24.30, 26.45, 26.57, 28.06, 28.24, 28.79, 29.05, 31.10, 33.42, 62.61, 65.47, 69.26, 129.47 (C=C), 129.71(C=C), 172.89 (C=O).

HRMS (APCI): Found 301.2383 (M^+ + H. $C_{17}H_{33}O_4$: Calculated 301.2373).

Analytical data corresponds to the literature.[60]

*Ethyl azidoacetate (**68**)*



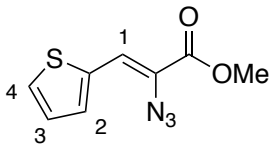
To **72** (25.0 g; 149.7 mmol) in DMF (80 mL) was added sodium azide (48.15 g, 740.7 mmol) at 0 ^{o}C. The reaction mixture was then heated to 50 ^{o}C for 18 h. After this time, water was added (100 mL). The product was extracted into ether (3 x 100 mL). The organic layers were combined, washed with water (2 x 100 mL) and brine (1 x 100

mL), dried over $MgSO_4$ and concentrated under vacuum to yield the product as a yellow liquid (15 g, 78%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 1.30 (3H, t, J=7.1 Hz, O-CH_2-$\underline{CH_3}$), 3.85 (2H, s, $\underline{CH_2}$), 4.25 (2H, q, J=7.1 Hz, O-$\underline{CH_2}$).

Analytical data corresponds to the literature.[48]

*methyl-2-azido-3-(thiophen-2-yl)acrylate (**66**)*

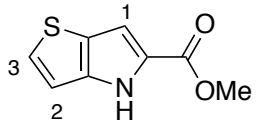


Sodium metal (0.40 g; 17.40 mmol; 3.9 eq) was added to anhydrous methanol (15 mL) at 0 °C. To this was added **69** (0.50 g; 0.42 mL; 4.46 mmol) and **68** (1.34 g; 10.37 mmol; 2.3 eq). The reaction was heated to 40 ^oC for 5 h. After this time, EtOAc (15 mL) and saturated ammonium chloride solution (15 mL) were added to the reaction mixture. The product was extracted using EtOAc (3 x 30 mL). The organic layers were dried over $MgSO_4$ and concentrated under vacuum. The crude product was purified by flash chromatography, eluting with hexanes/EtOAc, 95:5 to yield the product as a yellow oil (0.65; 65%).

^{1}H NMR (400 MHz; $CDCl_3$, *ppm*): 3.89 (3H, s, O$\underline{Me}$), 7.06 (1H, dd, J=3.8, 5.0 Hz, H3), 7.17 (1H, s, H1), 7.32 (1H, d, J=3.8 Hz, H2), 7.49 (1H, d, J=5.0 Hz, H4).

Analytical data corresponds to the literature.[70]

*methyl 4H-thieno[3,2-b]pyrrole-5-carboxylate (**64**)*

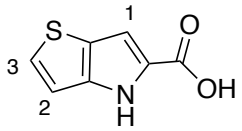


66 (0.61 g; 2.73 mmol) was dissolved in toluene (8 mL) and heated to reflux. After 1 h, the reaction mixture was allowed to cool to room temperature. The crude product was purified by flash chromatography eluting at hexanes/EtOAc, 9:1 to yield the product as a white solid (1.33 g; 85%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 3.89 (1H, s, OMe), 6.94 (1H, d, J=5.3 Hz, H2), 7.12 (1H, d, J=1.16 Hz, H1), 7.32 (1H, d, J=5.3 Hz, H3), 9.03 (NH, s).

Analytical data corresponds to the literature.[70]

*4H-thieno[3,2-b]pyrrole-5-carboxylic acid (**62**)*

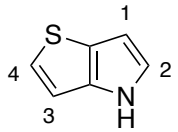


To a suspension of **64** (1.17 g; 6.4 mmol) in water (50 mL) was added potassium hydroxide (1.07 g; 19.2 mmol). The reaction was heated to reflux (110 ^oC) for 2 h. The reaction mixture was extracted with ether (2 x 50 mL) and the aqueous layer containing the product as the salt was acidified to pH 1-3 using 5 N HCl. The aqueous layer was then placed in the fridge for 2 h to aid precipitation as the protonated acid. The product was extracted using ether (3 x 50 mL). The organic layers were washed with water (2 x 50 mL) and brine (1 x 50 mL), dried over $MgSO_4$ and concentrated under vacuum to give the product as a white solid (0.81 g; 92%).

^{1}H NMR (DMSO-d_6, *ppm*): 6.94 (1H, d, J=5.3 Hz, H2), 7.00 (1H, s, H1), 7.45 (1H, d, J=5.3 Hz, H3), 11.86 (NH, s), 12.49 (OH, s).

Analytical data corresponds to the literature.[70]

*4H-thieno[3,2-b]pyrrole (**61**)*

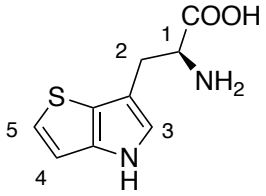


62 (430 mg; 2.6 mmol) was dissolved in ethanolamine (20 mL) and heated to 170 ^oC for 2 h. The reaction mixture was allowed to cool to room temperature. Water (20 mL) was added and it was extracted with 15% EtOAc in hexanes (3 x 50 mL). The organic layers were washed with water (2 x 100 mL) and brine, dried over $MgSO_4$ and concentrated under vacuum to give the product as a brown oil (190 mg, 59%).

^{1}H NMR: (DMSO-d_6, *ppm*): 6.30 (1H, m, H1), 6.96 (1H, d, J=5.2 Hz, H3), 7.01 (1H, m, H2), 7.13 (1H, d, J=5.2Hz, H4), 11.06 (NH, s).

Analytical data corresponds to the literature.[70]

*(S)-2-amino-3-(4H-thieno[3,2-b]pyrrol-6-yl)propanoic acid (**57**)*



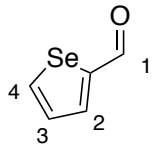
60 (360 mg; 2.92 mmol) was dissolved in DMSO (1.5 mL) and 0.5 mL was placed in three separate falcon tubes each containing L-serine (513 mg; 4.9 mmol) and 0.1 M ammonium bicarbonate buffer (pH = 8.0, 35 mL). To these falcon tubes was added pyridoxal-5-phosphate (PLP) co-enzyme (1.55 mg; 0.0058 mmol. Added from 4 mg/mL stock of PLP in ammonium bicarbonate buffer; 129 μL) and the Tryptophan Synthase enzyme (7.2 mg from a 27 mg/mL stock; 88 μL). The falcon tubes were placed at 37 °C with shaking (30 rpm). Every 1.5 h, a further addition of 1 eq. tryptophan synthase enzyme and 1 eq. PLP buffer were made, a total of 5 further additions were made. The reaction mixtures were combined and extracted with ether (2 x 150 mL), to remove any unreacted **60**. The aqueous layer was put down a sephadex G10 size exclusion column for purification eluting with water. The eluent was lyophilised to give the product as a brown solid (385 mg; 63%).

^{1}H NMR (400 MHz, D_2O, *ppm*): 3.26 (2H, m, H2), 4.02 (1H, dd, J=5.1, 7.0 Hz, H1), 7.05 (1H, d, J=1.3 Hz, H3), 7.08 (1H, d, J=5.2 Hz, H5), 7.24 (1H, d, J=5.2, 1.3 Hz, H4).

HRMS (*m/z* ES^+): Found 233.0348 (M^+ + Na. $C_9H_{10}N_2NaO_2S$: Calculated 233.0355).

Analytical data corresponds to the literature.[70]

*Selenophene-2-carbaldehyde (**70**)*

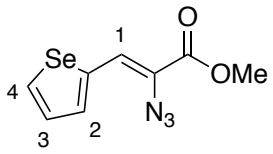


To **71** (3.5 g; 26.7 mmol) in DMF (3.12 g; 42.7 mmol) at 0 ^{o}C was added $POCl_3$ (6.55 g; 42.7 mmol). The reaction mixture was heated to 65 ^{o}C for 1 h. After this time, ice (20 g) was added. The organic layers were separated and dried over $MgSO_4$ and concentrated under vacuum to give the product as a brown oil (3.01 g; 65%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 7.46 (1H, dd, J=3.9, 5.4 Hz, H3), 8.01 (1H, dd, J=1.1, 3.9 Hz, H2), 8.48 (1H, m, H4), 9.80 (1H, s, H1).

Analytical data corresponds to the literature.[75]

*methyl (Z)-2-azido-3-(selenophen-2-yl)acrylate (**67**)*

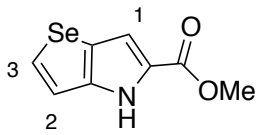


To sodium metal (1.80 g; 78.3 mmol) in a 100 mL round bottomed flask at 0 ^{o}C was added MeOH (40 mL). When the sodium had fully reacted, **70** (3.00 g; 18.9 mmol) and **68** (7.35 g; 78.3 mmol) was added. The reaction mixture was then allowed to heat to room temperature and left for 3 h. Ammonium chloride solution (30 mL) and EtOAc (30 mL) were added to the reaction mixture. The product was extracted into EtOAc (3 x 70 mL). The organic layers were washed with water (3 x 100 mL) and brine (1 x 100 mL), dried over $MgSO_4$ and concentrated under vacuum. The crude was purified over silica gel eluting with hexanes to give the product as a yellow oil (1.40 g; 35%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 3.89 (1H, s, OMe), 7.23 (1H, s, H1), 7.30 (1H, dd, J=3.9, 5.7 Hz, H3), 7.47 (1H, d, J=3.9 Hz, H2), 8.26 (1H, d, J=5.7 Hz, H4).

Analytical data corresponds to the literature.[48]

*methyl 4H-selenopheno[3,2-b]pyrrole-5-carboxylate (**65**)*

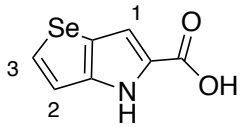


67 (1.4 g; 5.5 mmol) was dissolved in toluene (15 mL). The reaction mixture was then heated to reflux. After 1.5 h gas evolution ceased. The reaction mixture was allowed to cool to 20 ^{o}C then concentrated under vacuum to give a brown solid. The crude was purified by flash column chromatography over silica gel eluting with hexanes/EtOAc, 4:1, to give the product as a yellow solid (0.85 g; 68%).

^{1}H NMR (400 MHz, $CDCl_3$, *ppm*): 3.88 (3H, s, OMe), 7.17 (1H, s, H1), 7.23 (1H, d, J=5.8 Hz, H2), 7.87 (1H, d, J=5.8 Hz, H3), 9.13 (NH, s).

Analytical data corresponds to the literature.[48]

*4H-selenopheno[3,2-b]pyrrole-5-carboxylic acid (**63**)*

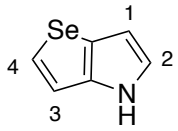


To **65** (850 mg; 3.7 mmol) was added water (40 mL) and potassium hydroxide (840 mg; 15.0 mmol). The reaction was heated to reflux (110 ^{o}C). After 2 h, the reaction mixture was allowed to cool to 20 ^{o}C. The organic material was then removed by extraction into ether (2 x 50 mL), as the charged product remains in the aqueous layer as the potassium salt. The aqueous layer was then acidified with 5 N HCl, (pH = 1-3) and stored at 4 ^{o}C to aid precipitation (1 h). The product was then extracted into ether (3 x 70 mL). The organic layers were combined, washed with water (2 x 70 mL) and brine (1 x 70 mL), dried over $MgSO_4$ and concentrated under vacuum to yield the product as a yellow solid (520 mg; 65%).

^{1}H NMR (400 MHz, DMSO-d_6, *ppm*): 7.03 (1H, s, H1), 7.19 (1H, d, J=5.8 Hz, H2), 7.92 (1H, d, J=5.8 Hz, H3), 11.94 (NH, s), 12.42 (OH, s).

Analytical data corresponds to the literature.[48]

*4H-selenopheno[3,2-b]pyrrole (**61**)*



63 (260 mg; 1.2 mmol) was suspended in glycerol (25 mL). The flask was added to a pre heated oil bath at 130 ^{o}C. After 1.5 h, the reaction mixture was poured onto ice. The mixture was then transferred to a separating funnel and extracted with hexane (3 x 70 mL). The organic layers were washed with water (2 x 100 mL) and brine (1 x 100 mL), dried over $MgSO_4$ and concentrated under vacuum to yield the product as a black oil (50 mg; 25%).

^{1}H NMR (400 MHz, DMSO-d$_6$, *ppm*): 6.35 (1H, m, H2), 6.96 (1H, m, H2), 7.22 (1H, d, J=5.6 Hz, H4), 7.62 (1H, dd, J=1.2, 5.6 Hz, H3), 11.17 (NH, s).

Analytical data corresponds to the literature.[48]

Bibliography:

(1) Boland, C.; Olatunji, S.; Bailey, J.; Howe, N.; Weichert, D.; Fetics, S. K.; Yu, X.; Merino-Gracia, J.; Delsaut, C.; Caffrey, M. Membrane (and Soluble) Protein Stability and Binding Measurements in the Lipid Cubic Phase Using Label-Free Differential Scanning Fluorimetry. *Anal. Chem.* **2018**, *90* (20), 12152–12160. https://doi.org/10.1021/acs.analchem.8b03176.

(2) Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf (accessed Jul 17, 2018).

(3) World Health Organisation. Antimicrobial Resistance http://www.who.int/en/news-room/fact-sheets/detail/antimicrobial-resistance (accessed Jul 17, 2018).

(4) Marston, H. D.; Dixon, D. M.; Knisely, J. M.; Palmore, T. N.; Fauci, A. S. Antimicrobial Resistance. *JAMA* **2016**, *316* (11), 1193. https://doi.org/10.1001/jama.2016.11764.

(5) Demain, A. L. Small Bugs, Big Business: The Economic Power of the Microbe. *Biotechnol. Adv.* **2000**, 16.

(6) Stokes, H. W.; Gillings, M. R. Gene Flow, Mobile Genetic Elements and the Recruitment of Antibiotic Resistance Genes into Gram-Negative Pathogens. *FEMS Microbiol. Rev.* **2011**, *35* (5), 790–819. https://doi.org/10.1111/j.1574-6976.2011.00273.x.
(7) Pray, L. Antibiotic Resistance, Mutation Rates and MRSA. *Nat. Educ.* **2008**, *1* (1), 30.
(8) Kirkup, B. Learning from Losers. *eLife* **2017**, *6*. https://doi.org/10.7554/eLife.32703.
(9) Wholey, W.-Y.; Kochan, T. J.; Storck, D. N.; Dawid, S. Coordinated Bacteriocin Expression and Competence in Streptococcus Pneumoniae Contributes to Genetic Adaptation through Neighbor Predation. *PLOS Pathog.* **2016**, *12* (2), e1005413. https://doi.org/10.1371/journal.ppat.1005413.
(10) Croucher, N. J.; Mostowy, R.; Wymant, C.; Turner, P.; Bentley, S. D.; Fraser, C. Horizontal DNA Transfer Mechanisms of Bacteria as Weapons of Intragenomic Conflict. *PLOS Biol.* **2016**, *14* (3), e1002394. https://doi.org/10.1371/journal.pbio.1002394.
(11) Marshall, B. M.; Levy, S. B. Food Animals and Antimicrobials: Impacts on Human Health. *Clin. Microbiol. Rev.* **2011**, *24* (4), 718–733. https://doi.org/10.1128/CMR.00002-11.
(12) Van Boeckel, T. P.; Brower, C.; Gilbert, M.; Grenfell, B. T.; Levin, S. A.; Robinson, T. P.; Teillant, A.; Laxminarayan, R. Global Trends in Antimicrobial Use in Food Animals. *Proc. Natl. Acad. Sci.* **2015**, *112* (18), 5649–5654. https://doi.org/10.1073/pnas.1503141112.
(13) Rice, L. B. Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE. *J. Infect. Dis.* **2008**, *197* (8), 1079–1081. https://doi.org/10.1086/533452.
(14) Pendleton, J. N.; Gorman, S. P.; Gilmore, B. F. Clinical Relevance of the ESKAPE Pathogens. *Expert Rev. Anti Infect. Ther.* **2013**, *11* (3), 297–308. https://doi.org/10.1586/eri.13.12.
(15) Boucher, H. W.; Talbot, G. H.; Bradley, J. S.; Edwards, J. E.; Gilbert, D.; Rice, L. B.; Scheld, M.; Spellberg, B.; Bartlett, J. Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **2009**, *48* (1), 1–12. https://doi.org/10.1086/595011.
(16) Biswas, S.; Brunel, J.-M.; Dubus, J.-C.; Reynaud-Gaubert, M.; Rolain, J.-M. Colistin: An Update on the Antibiotic of the 21st Century. *Expert Rev. Anti Infect. Ther.* **2012**, *10* (8), 917–934. https://doi.org/10.1586/eri.12.78.
(17) Brown, E. D.; Wright, G. D. Antibacterial Drug Discovery in the Resistance Era. *Nature* **2016**, *529* (7586), 336–343. https://doi.org/10.1038/nature17042.
(18) Vogeley, L.; El Arnaout, T.; Bailey, J.; Boland, C.; Stansfeld, P. J.; Caffrey, M. Structural Basis of Lipoprotein Signal Peptidase II Action and Inhibition by the Antibiotic Globomycin. *Science* **2016**, *351* (6275), 876–880.
(19) Bartual, S. G.; Alcorlo, M.; Martínez-Caballero, S.; Molina, R.; Hermoso, J. A. Three-Dimensional Structures of Lipoproteins from Streptococcus Pneumoniae and Staphylococcus Aureus. *Int. J. Med. Microbiol.* **2018**, *308* (6), 692–704. https://doi.org/10.1016/j.ijmm.2017.10.003.

(20) Kovacs-Simon, A.; Titball, R. W.; Michell, S. L. Lipoproteins of Bacterial Pathogens. *Infect. Immun.* **2011**, *79* (2), 548–561. https://doi.org/10.1128/IAI.00682-10.
(21) Buddelmeijer, N. The Molecular Mechanism of Bacterial Lipoprotein Modification—How, When and Why? *FEMS Microbiol. Rev.* **2015**, *39* (2), 246–261. https://doi.org/10.1093/femsre/fuu006.
(22) Hutchings, M. I.; Palmer, T.; Harrington, D. J.; Sutcliffe, I. C. Lipoprotein Biogenesis in Gram-Positive Bacteria: Knowing When to Hold 'em, Knowing When to Fold 'Em. *Trends Microbiol.* **2009**, *17* (1), 13–21. https://doi.org/10.1016/j.tim.2008.10.001.
(23) *Posttranslational Modifications. Pt. A: ...*; Wold, F., Moldave, K., Eds.; Methods in enzymology; Academic Press: Orlando, Fla., 1984.
(24) Sundaram, S.; Banerjee, S.; Sankaran, K. The First Nonradioactive Fluorescence Assay for Phosphatidylglycerol:Prolipoprotein Diacylglyceryl Transferase That Initiates Bacterial Lipoprotein Biosynthesis. *Anal. Biochem.* **2012**, *423* (1), 163–170. https://doi.org/10.1016/j.ab.2012.01.018.
(25) Mao, G.; Zhao, Y.; Kang, X.; Li, Z.; Zhang, Y.; Wang, X.; Sun, F.; Sankaran, K.; Zhang, X. C. Crystal Structure of E. Coli Lipoprotein Diacylglyceryl Transferase. *Nat. Commun.* **2016**, *7*, 10198. https://doi.org/10.1038/ncomms10198.
(26) Wiktor, M.; Weichert, D.; Howe, N.; Huang, C.-Y.; Olieric, V.; Boland, C.; Bailey, J.; Vogeley, L.; Stansfeld, P. J.; Buddelmeijer, N.; et al. Structural Insights into the Mechanism of the Membrane Integral N-Acyltransferase Step in Bacterial Lipoprotein Synthesis. *Nat. Commun.* **2017**, *8*, 15952. https://doi.org/10.1038/ncomms15952.
(27) Zückert, W. R. Secretion of Bacterial Lipoproteins: Through the Cytoplasmic Membrane, the Periplasm and Beyond. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2014**, *1843* (8), 1509–1516. https://doi.org/10.1016/j.bbamcr.2014.04.022.
(28) Pailler, J.; Aucher, W.; Pires, M.; Buddelmeijer, N. Phosphatidylglycerol::Prolipoprotein Diacylglyceryl Transferase (Lgt) of Escherichia Coli Has Seven Transmembrane Segments, and Its Essential Residues Are Embedded in the Membrane. *J. Bacteriol.* **2012**, *194* (9), 2142–2151. https://doi.org/10.1128/JB.06641-11.
(29) Inukai, M.; Nakajima, M.; Osawa, M.; Haneishi, T.; Arai, M. Globomycin, a New Peptide Antibiotic with Spheroplast-Forming Activity. *J. Antibiot. (Tokyo)* **1977**, *31* (5).
(30) Hillmann, F.; Argentini, M.; Buddelmeijer, N. Kinetics and Phospholipid Specificity of Apolipoprotein *N* -Acyltransferase. *J. Biol. Chem.* **2011**, *286* (32), 27936–27946. https://doi.org/10.1074/jbc.M111.243519.
(31) Gupta, K.; Donlan, J. A. C.; Hopper, J. T. S.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V. The Role of Interfacial Lipids in Stabilizing Membrane Protein Oligomers. *Nature* **2017**, *541* (7637), 421–424. https://doi.org/10.1038/nature20820.
(32) Buddelmeijer, N.; Young, R. The Essential *Escherichia Coli* Apolipoprotein *N* -Acyltransferase (Lnt) Exists as an Extracytoplasmic Thioester Acyl-

Enzyme Intermediate. *Biochemistry* **2010**, *49* (2), 341–346. https://doi.org/10.1021/bi9020346.
(33) Coates, A. R.; Halls, G.; Hu, Y. Novel Classes of Antibiotics or More of the Same?: New Antibiotic Classes Are Urgently Needed. *Br. J. Pharmacol.* **2011**, *163* (1), 184–194. https://doi.org/10.1111/j.1476-5381.2011.01250.x.
(34) Harayama, T.; Riezman, H. Understanding the Diversity of Membrane Lipid Composition. *Nat. Rev. Mol. Cell Biol.* **2018**, *19* (5), 281–296. https://doi.org/10.1038/nrm.2017.138.
(35) Li, J.; Wang, X.; Zhang, T.; Wang, C.; Huang, Z.; Luo, X.; Deng, Y. A Review on Phospholipids and Their Main Applications in Drug Delivery Systems. *Asian J. Pharm. Sci.* **2015**, *10* (2), 81–98. https://doi.org/10.1016/j.ajps.2014.09.004.
(36) Sohlenkamp, C.; Geiger, O. Bacterial Membrane Lipids: Diversity in Structures and Pathways. *FEMS Microbiol. Rev.* **2016**, *40* (1), 133–159. https://doi.org/10.1093/femsre/fuv008.
(37) Landau, E. M.; Rosenbusch, J. P. Lipidic Cubic Phases: A Novel Concept for the Crystallization of Membrane Proteins. *Proc. Natl. Acad. Sci.* **1996**, *93* (25), 14532–14535. https://doi.org/10.1073/pnas.93.25.14532.
(38) Caffrey, M.; Cherezov, V. Crystallizing Membrane Proteins Using Lipidic Mesophases. *Nat. Protoc.* **2009**, *4* (5), 706–731. https://doi.org/10.1038/nprot.2009.31.
(39) Caffrey, M. On the Mechanism of Membrane Protein Crystallization in Lipidic Mesophases †. *Cryst. Growth Des.* **2008**, *8* (12), 4244–4254. https://doi.org/10.1021/cg800693r.
(40) Caffrey, M. A Comprehensive Review of the Lipid Cubic Phase or *in Meso* Method for Crystallizing Membrane and Soluble Proteins and Complexes. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2015**, *71* (1), 3–18. https://doi.org/10.1107/S2053230X14026843.
(41) Yang, D.; Cwynar, V. A.; Hart, D. J.; Madanmohan, J.; Lee, J.; Lyons, J.; Caffrey, M. Preparation of 1-Monoacylglycerols via the Suzuki-Miyaura Reaction: 2,3-Dihydroxypropyl (*Z*)-Tetradec-7-Enoate. In *Organic Syntheses*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014; pp 183–201. https://doi.org/10.1002/0471264229.os089.19.
(42) Rasmussen, S. G. F.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; et al. Crystal Structure of the B2 Adrenergic Receptor–Gs Protein Complex. *Nature* **2011**, *477* (7366), 549–555. https://doi.org/10.1038/nature10361.
(43) Misquitta, L. V.; Misquitta, Y.; Cherezov, V.; Slattery, O.; Mohan, J. M.; Hart, D.; Zhalnina, M.; Cramer, W. A.; Caffrey, M. Membrane Protein Crystallization in Lipidic Mesophases with Tailored Bilayers. *Structure* **2004**, *12* (12), 2113–2124. https://doi.org/10.1016/j.str.2004.09.020.
(44) Li, D.; Shah, S. T. A.; Caffrey, M. Host Lipid and Temperature as Important Screening Variables for Crystallizing Integral Membrane Proteins in Lipidic Mesophases. Trials with Diacylglycerol Kinase. *Cryst. Growth Des.* **2013**, *13* (7), 2846–2857. https://doi.org/10.1021/cg400254v.
(45) Love, J.; Mancia, F.; Shapiro, L.; Punta, M.; Rost, B.; Girvin, M.; Wang, D.-N.; Zhou, M.; Hunt, J. F.; Szyperski, T.; et al. The New York Consortium on Membrane Protein Structure (NYCOMPS): A High-Throughput Platform for

Structural Genomics of Integral Membrane Proteins. *J. Struct. Funct. Genomics* **2010**, *11* (3), 191–199. https://doi.org/10.1007/s10969-010-9094-7.
(46) Bent, A. F.; Mann, G.; Houssen, W. E.; Mykhaylyk, V.; Duman, R.; Thomas, L.; Jaspars, M.; Wagner, A.; Naismith, J. H. Structure of the Cyanobactin Oxidase ThcOx from *Cyanothece* Sp. PCC 7425, the First Structure to Be Solved at Diamond Light Source Beamline I23 by Means of S-SAD. *Acta Crystallogr. Sect. Struct. Biol.* **2016**, *72* (11), 1174–1180. https://doi.org/10.1107/S2059798316015850.
(47) de Jesus, A. J.; Allen, T. W. The Role of Tryptophan Side Chains in Membrane Protein Anchoring and Hydrophobic Mismatch. *Biochim. Biophys. Acta BBA - Biomembr.* **2013**, *1828* (2), 864–876. https://doi.org/10.1016/j.bbamem.2012.09.009.
(48) Welch, M.; Phillips, R. S. Improved Syntheses of [3,2-b]- and [2,3-b]-Fused Selenolo- and Thienopyrroles, and of Furo[3,2-Fa]Pyrrole. *Heterocycl. Commun.* **1999**, *5* (4), 6.
(49) Maïo, I.; Barbier, P.; Allegro, D.; Brault, C.; Peyrot, V. Quantitative Analysis of Tau-Microtubule Interaction Using FRET. *Int. J. Mol. Sci.* **2014**, *15* (8), 14697–14714. https://doi.org/10.3390/ijms150814697.
(50) Oliveira, L. C. G.; Silva, V. O.; Okamoto, D. N.; Kondo, M. Y.; Santos, S. M. B.; Hirata, I. Y.; Vallim, M. A.; Pascon, R. C.; Gouvea, I. E.; Juliano, M. A.; et al. Internally Quenched Fluorescent Peptide Libraries with Randomized Sequences Designed to Detect Endopeptidases. *Anal. Biochem.* **2012**, *421* (1), 299–307. https://doi.org/10.1016/j.ab.2011.10.025.
(51) Murakami, K.; Molitor, E. J.; Liu, H. An Efficient Synthesis of Unsymmetrical Optically Active Phosphatidyl Glycerol. *J. Org. Chem.* **1999**, *64* (2), 648–651. https://doi.org/10.1021/jo981653p.
(52) McMillen, D. A.; Volwerk, J. J.; Ohishi, J.; Erion, M.; Keana, J. F. W.; Jost, P. C.; Griffith, O. H. Identifying Regions of Membrane Proteins in Contact with Phospholipid Head Groups: Covalent Attachment of a New Class of Aldehyde Lipid Labels to Cytochrome c Oxidase. *Biochemistry* **1986**, *25* (1), 182–193. https://doi.org/10.1021/bi00349a027.
(53) Martin, S. F.; Josey, J. A.; Wong, Y.-L.; Dean, D. W. General Method for the Synthesis of Phospholipid Derivatives of 1,2-O-Diacyl-Sn-Glycerols. *J. Org. Chem.* **1994**, *59* (17), 4805–4820. https://doi.org/10.1021/jo00096a023.
(54) Crouch, R. D. Recent Advances in Silyl Protection of Alcohols. *Synth. Commun.* **2013**, *43* (17), 2265–2279. https://doi.org/10.1080/00397911.2012.717241.
(55) Bailey, S.; Helliwell, M.; Teerawutgulrag, A.; Thomas, E. J. A Total Synthesis of Milbemycin G: Approaches to the C(1)–C(10)-Fragment and Completion of the Synthesis. *Org. Biomol. Chem.* **2005**, *3* (20), 3654. https://doi.org/10.1039/b508675b.
(56) Zumbuehl, A.; Zaffalon, P.-L. BODP - A Versatile Reagent for Phospholipid Synthesis. *Synthesis* **2011**, *5*, 778–782.
(57) Andresen, T. L.; Jensen, S. S.; Madsen, R.; Jørgensen, K. Synthesis and Biological Activity of Anticancer Ether Lipids That Are Specifically Released by Phospholipase A $_2$ in Tumor Tissue. *J. Med. Chem.* **2005**, *48* (23), 7305–7314. https://doi.org/10.1021/jm049006f.

(58) Pilcher, A. S.; Ammon, H. L.; DeShong, P. Utilization of Tetrabutylammonium Triphenylsilyldifluoride as a Fluoride Source for Nucleophilic Fluorination. *J. Am. Chem. Soc.* **1995**, *117* (18), 5166–5167. https://doi.org/10.1021/ja00123a025.
(59) Chen, C.-Y.; Han, W.-B.; Chen, H.-J.; Wu, Y.; Gao, P. Optically Active Monoacylglycerols: Synthesis and Assessment of Purity: Optically Active Monoacylglycerols. *Eur. J. Org. Chem.* **2013**, *2013* (20), 4311–4318. https://doi.org/10.1002/ejoc.201300247.
(60) Yang, D.; Cwynar, V. A.; Hart, D. J.; Madanmohan, J.; Lee, J.; Lyons, J.; Caffrey, M. Preparation of 1-Monoacylglycerols via the Suzuki-Miyaura Reaction: 2,3-Dihydroxypropyl (*Z*)-Tetradec-7-Enoate. In *Organic Syntheses*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014; pp 183–201. https://doi.org/10.1002/0471264229.os089.19.
(61) Chen, S.-N.; Hung, T.-T. Reusable and Efficient CuI/TBAB-Catalyzed Iodination of Terminal Alkynes in Water under Air. 4.
(62) Lyons, J. Towards the Rational Design of Lipids for Use in In Meso Crystallisation of Membrane Proteins: A Case Study with a Cytochrome Oxidase and a G-Protein Coupled Receptor, 2012.
(63) Wang, K.; Hawley, M. C.; DeAthos, S. J. Conversion of Glycerol to 1,3-Propanediol via Selective Dehydroxylation. *Ind. Eng. Chem. Res.* **2003**, *42* (13), 2913–2923. https://doi.org/10.1021/ie020754h.
(64) Hill, H. S.; Whelen, M. S.; Hibbert, H. Studies on the Reactions Relating to Carbohydrates and Polysaccharides. XV. The Isomeric Benzylidene Glycerols. *J. Am. Chem. Soc.* **1928**, *50* (8), 2235–2242. https://doi.org/10.1021/ja01395a025.
(65) Mori, K. Pheromone Synthesis. Part 253: Synthesis of the Racemates and Enantiomers of Triglycerides of Male Drosophila Fruit Flies with Special Emphasis on the Preparation of Enantiomerically Pure 1-Monoglycerides. *Tetrahedron* **2012**, *68* (40), 8441–8449. https://doi.org/10.1016/j.tet.2012.07.086.
(66) Fu, Y.; Weng, Y.; Hong, W.-X.; Zhang, Q. Efficient Synthesis of Unsaturated 1-Monoacyl Glycerols for in Meso Crystallization of Membrane Proteins. *Synlett* **2011**, *2011* (06), 809–812. https://doi.org/10.1055/s-0030-1259912.
(67) Zheng, Z.; Luo, M.; Yu, J.; Wang, J.; Ji, J. Novel Process for 1,3-Dihydroxyacetone Production from Glycerol. 1. Technological Feasibility Study and Process Design. *Ind. Eng. Chem. Res.* **2012**, *51* (9), 3715–3721. https://doi.org/10.1021/ie201710h.
(68) Pfefferkorn, C. M.; Lee, J. C. 5-Fluoro-d,l-Tryptophan as a Dual NMR and Fluorescent Probe of α-Synuclein. In *Intrinsically Disordered Protein Analysis*; Uversky, V. N., Dunker, A. K., Eds.; Humana Press: Totowa, NJ, 2012; Vol. 895, pp 197–209. https://doi.org/10.1007/978-1-61779-927-3_14.
(69) Soth, S.; Farnierz, M.; Paulmier, C. Recherches en série hétérocyclique. XXIX. Sur des voies d'accès à des thiéno, sélénolo, furo et pyrrolopyrroles. *Can. J. Chem.* **1978**, *56* (10), 1429–1434. https://doi.org/10.1139/v78-234.
(70) Hoesl, M. G.; Oehm, S.; Durkin, P.; Darmon, E.; Peil, L.; Aerni, H.-R.; Rappsilber, J.; Rinehart, J.; Leach, D.; Söll, D.; et al. Chemical Evolution of a

Bacterial Proteome. *Angew. Chem. Int. Ed.* **2015**, *54* (34), 10030–10034. https://doi.org/10.1002/anie.201502868.
(71) Bach, T. G. *Organoselenium Chemistry: A Practical Approach*; OUP Oxford, 1999.
(72) Inman, M.; Moody, C. J. Indole Synthesis – Something Old, Something New. *Chem Sci* **2013**, *4* (1), 29–41. https://doi.org/10.1039/C2SC21185H.
(73) Harried, S. S.; Croghan, M. D.; Kaller, M. R.; Lopez, P.; Zhong, W.; Hungate, R.; Reider, P. J. Stereoselective Synthesis of *Anti* - *N* -Protected 3-Amino-1,2-Epoxides by Nucleophilic Addition to *N* - *Tert* -Butanesulfinyl Imine of a Glyceraldehyde Synthon †. *J. Org. Chem.* **2009**, *74* (16), 5975–5982. https://doi.org/10.1021/jo900643b.
(74) Bartlett, N.; Gross, L.; Péron, F.; Asby, D. J.; Selby, M. D.; Tavassoli, A.; Linclau, B. Stereocontrol by Quaternary Centres: A Stereoselective Synthesis of (−)-Luminacin D. *Chem. - Eur. J.* **2014**, *20* (12), 3306–3310. https://doi.org/10.1002/chem.201304776.
(75) Kim, K.; Jo, C.; Easwaramoorthi, S.; Sung, J.; Kim, D. H.; Churchill, D. G. Crystallographic, Photophysical, NMR Spectroscopic and Reactivity Manifestations of the “8-Heteroaryl Effect” in 4,4-Difluoro-8-(C_4H_3X)-4-Bora-3*a*,4*a*-Diaza-*s*-Indacene (*X* = O, S, Se) (BODIPY) Systems. *Inorg. Chem.* **2010**, *49* (11), 4881–4894. https://doi.org/10.1021/ic902467h.