# Development of N-terminal targeting ligands for protein-material conjugation

**Charis Miriam Warnock**

**MSc by Research**

**University of York**

**Chemistry**

**May 2022**

# Abstract.

Research within the Spicer group focuses on the synthesis of biomaterials for use in tissue engineering – the research I will present in this thesis focuses on the site-selective immobilization of extracellular matrix (ECM) signalling proteins to a synthetic tissue scaffold via the N-terminus. There are many different types of tissue scaffold that have been developed over the past few decades; synthetic tissue scaffolds are of particular interest due to their highly tuneable properties and biocompatibility. However, the main drawback of synthetic tissue scaffolds is their lack of cell adhesive motifs and bioactivity – this can be introduced in several ways. In this project we have focused on introducing bioactivity through immobilisation of ECM proteins to the scaffold via the N-terminus. The N-terminus of proteins has a very unique chemistry and can be targeted site-selectively without disruption to the protein's structure or activity. Herein, we present the synthesis of poly (ethylene glycol) (PEG) tissue scaffolds functionalised with novel 2-pyridine carbaldehyde (2PCA) derivatives capable of specifically and reversibly binding to the N-terminus of proteins.

# Contents

## List of Figures.

# List of schemes.

# Abbreviations

| | |
|---|---|
| 2-EBA | 2-Ethynylbenzaldehyde |
| 2PCA | 2-Pyridine carbaldehyde |
| ABM | 2-(Aminooxy) ethyl-2-bromo-2-methylpropanoate |
| AIBN | Azobisisobutyronitrile |
| app | Apparent |
| Ar | Aromatic |
| ATRP | Atom transfer radical polymerisation |
| Boc | *Tert*-butoxycarbonyl |
| br | Broad |
| $CDCl_3$ | Deuterated chloroform |
| COSY | Correlated Spectroscopy NMR |
| d | Doublet |
| $D_2O$ | Dueterated water |
| DCM | Dichloromethane |
| dd | Doublet of doublets |
| dd | Doublet of doublets |
| DIC | *N,N'*-diisopropylcarbodiimide |
| DMF | Dimethylformamide |
| DMSO-*d$_6$* | Deuterated dimethyl sulfoxide |
| DTT | 1,4-Dithiothreitol |
| DVB | Divinylbenzene |
| ECM | Extracellular matrix |
| EDC·HCl | N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride |
| EGF | Epidermal growth factor |
| eq. | Equivalent |
| ESI | Electrospray ionisation |
| Et | Ethyl |
| Et | Ethyl |
| EtOAc | Ethylacetate |
| GFP | Green flourescent protein |
| h | Hours |
| HMBC | Heteronuclear Multiple Bond Correlation NMR |
| HPLC | High performance liquid chromatography |
| HRMS | High resolution electrospray ionisation mass spectra |
| HSQC | Heteronuclear Single Quantum Coherence NMR |
| Hz | Hertz |
| IKVAV | Peptide sequence Ile-Lys-Val-Ala-Val |
| IR | Infrared |
| *J* | Coupling constants reported in Hz |
| LC-MS | Liquid chromatography-mass spectometry |
| m | Multiplet |
| M | Molar |
| M | Molar |
| m.p. | Melting point |

| | |
|---|---|
| *m/z* | Mass to charge ratio |
| *m/z* | Mass to charge ratio |
| MBHA | Methylbenzhydryl amine |
| Me | Methyl |
| MeOH | Methanol |
| mmol | Millimolar |
| MSCs | Mesenchymal stem cells |
| $NH_2$ | Amine |
| NHS | N-hydroxy succinimide |
| NMR | Nuclear magnetic resonance |
| NPC | 4-Nitrophenyl chloroformate |
| O/N | Overnight |
| OH | Hydroxyl |
| OH | Hydroxyl |
| PAM | Poly(acrylamide) |
| PCL | Poly(caprolactone) |
| PDA | Pyridine-2,6-dicarboxaldehyde |
| PDGF | Platelet-derived growth factor |
| PEG | Poly(ethylene glycol) |
| PEGA | Poly(ethylene glycol) acylamide |
| Ph | Phenyl |
| Ph | Phenyl |
| PHEMA | Poly(hydroxyethyl methacrylate) |
| POEGMA | Poly(oligo(ethylene glycol) methyl ether methacrylate |
| p-TsOH·$H_2O$ | p-Toluenesulfonic acid monohydrate |
| p-TsOH·$H_2O$ | p-Toluenesulfonic acid monohydrate |
| PVA | Poly(vinyl acetate) |
| q | Quartet |
| RGD | Tripeptide sequence arginine-glycine-aspartic acid |
| RT | Room temperature |
| s | Singlet |
| Sept | Septet |
| SH | Thiol |
| SPPS | Solid-phase peptide synthesis |
| t | Triplet |
| TA | Tyramine |
| TFA | Trifluoroacetic acid |
| THF | Tetrahydrofuran |
| TIPS | Triisopropylsilane |
| TLC | Thin layer chromatography |
| Ts | Tosyl |
| Ts | Tosyl |
| UV | Ultraviolet |
| VEGF | Vascular endothelial growth factor |
| VS | Vinyl sulfone |
| $U_{max}$ | Absorption maxima |

# Acknowledgements.

First and foremost, I would like to thank my brilliant supervisor, Dr Chris Spicer, for his unwavering support and encouragement over the past few years. Thank you for allowing me to be the first student of the Spicer Group way back in 2018 when I was just an undergrad looking for a supervisor for my BSc project. Despite my best efforts to make you question letting me stay on in the group for my MSc, I have thoroughly enjoyed my time as your supervisee, and I couldn't have wished for a better supervisor.

To the other members of the Spicer Group over the past few years, AKA. The Spicer Girls, thank you for your friendship and support, both inside the lab and out. Nick, Laetitia, Lydia, Natalia and Taylor, I couldn't have done it without you.

I would also like to thank the members of the E214 lab – you've all be great and incredibly helpful when I've needed it. I would like to give a special mention to Iman, without whom the lab would most definitely fall apart.

To my friends outside the lab; Tom, Erin, Sophie, Mandy, Mischa, Ruth, Justin, Hannah, Kate, Remy and Cameron – thank you for your friendship and entertainment when the degree started to take its toll and thank you all for sticking with me all these years.

And last, but definitely not least, I would like to thank my family who have supported me through my many ups and downs and who have kept pushing me to work to my greatest potential. Mum, Rosalie, Bethany, Benedict, Lola, Star, Gemma – thank you for the encouraging words, the shoulders to cry on, and the reasons to laugh and keep on going. A special mention, goes to my Granda, without whom I would not have had the inspiration nor the means to pursue this degree – thank you for keeping me on my toes and never failing to ask, "so when is the dissertation due then?".

## Author's Declaration.

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

*Charis Warnock*

# 1. Introduction.

The research I will present in this thesis focuses on the development of novel 2-pyridine carboxaldehyde (2PCA) molecules used in the site-selective immobilization of extracellular matrix (ECM) signalling proteins to a synthetic poly (ethylene glycol) (PEG) tissue scaffold via the N-terminus of the protein. In this introduction, I will introduce the importance of biomaterials, and the influence of the (ECM) in cell/tissue development. I will briefly explore the different materials used to manufacture tissue scaffolds and will examine how we can mimic the ECM through careful polymer choice and the introduction of bioactivity. Finally, I will focus on the current methods of protein immobilisation, giving special attention to the chemistries developed by the Francis Group for the site-specific immobilisation of proteins via the N-terminus using 2PCA.

## 1.1.  Tissue Scaffolds.

### 1.1.1.  Brief introduction to scaffolds for tissue regeneration.

Tissue engineering involves the regeneration/replacement of diseased or damaged tissue; its need often arises due to age-related illness, genetic diseases, cancer, or trauma. As the body ages its ability to repair itself decreases (due to a loss of progenitor cells with age) and the speed at which repair takes place slows down.[1,2] Furthermore, disease and traumatic injury can compromise the body's ability to regenerate lost tissue and so the need for tissue engineering arises.[1,3] Rheumatoid arthritis, for example, is caused by the body attacking the tissue surrounding joints, which leads to inflammation and damage to the bones, cartilage, tendons and ligaments.[4] Current treatments can involve suppression of the immune system to prevent further damage, however in cases where damage has already been inflicted upon the joints, regeneration of tissues, such as cartilage could provide a suitable alternative/ combination treatment.

Currently clinical treatment for tissue replacement (i.e. skin grafts, organ transplants) focuses on autografts and allografts with autografts being dubbed the "gold standard".[3] Autografts use tissue from the patients own body, such as skin grafts, whereas allografts use tissue from a donor to replace the lost/damaged tissue of another individual. This approach

predominantly focuses on organ transplants, such as liver, kidney, and heart transplants. However, there are significant issues associated with both; autografts prove challenging in that it is difficult and expensive to obtain tissue (especially from the elderly) and requires additional surgery for the patient which also poses risks of donor-site morbidity and infection.[3] Furthermore, the ability to use autografts relies on the fact that the patient does not carry a disease that would prevent their own tissue from being used. Alternatively, allografts suffer from challenges accessing enough tissue for every patient, not to mention the risks of infection and disease from the donated tissue and the requirement that the patient takes immunosuppressants for the remainder of their life to prevent graft rejection.[1,5]

Due to the problems described above, biomaterial scaffolds provide an attractive alternative. Tissue scaffolds are porous, degradable 3D matrices that mimic the physical, mechanical, and bio-instructive properties of the ECM of native tissue.[6-8] They provide a structure to support cell adhesion, migration, proliferation, and differentiation into new tissue for the purpose of tissue regeneration and replacement.[7,9] Biomaterial scaffolds offer the ability to overcome the issues posed by autografts/allografts; tissues do not need to be painfully extracted from the patient/donor and by using stem cells from the patient the generated tissue shouldn't elicit an immune response, thereby removing the requirement for immunosuppressants. It also overcomes the supply issue of tissues that are often not available for transplant unless taken from a recently deceased donor, such as heart and lung tissue.

Tissue scaffolds should meet a range of requirements in order to fulfil their purpose: they must be cell adhesive, biocompatible, non-immunogenic, viscoelastic, cost-effective to produce, reproducible and scalable.[7-9] They must also possess good mechanical strength, both to support cell growth and the potential need to be handled in a surgical environment.[1] Scaffolds should ideally have a porous structure to facilitate the migration of cells within the matrix and allow the transport of nutrients and the removal of waste products.[7] They should be biodegradable and generate non-toxic degradation products able to be absorbed by the body or excreted. Furthermore, the rate of scaffold degradation should match the rate of tissue production and ECM formation by the cells; if the scaffold doesn't degrade fast enough tissue formation will be stunted, however if the rate of degradation is too fast then the scaffold may lose the mechanical strength required to support the growing tissue.[1,6]

There are different types of biomaterial technology. Examples include scaffolds that are used to grow a tissue in vitro that can then be implanted, or those which are seeded with cells, implanted, and then tissue forms in vivo, using the body as a bioreactor to direct cell growth.[2] In the case of the former, scaffolds are typically seeded with stem cells, such as mesenchymal stem cells (MSCs) which have the ability to differentiate into fat, muscle, cartilage and bone cells, before subjecting them to biological stimuli that promote and guide cell differentiation and proliferation.[1,2] The type of stem cells used for seeding are important in determining the type of tissue formed: in the case of regenerating cartilage for patients suffering with rheumatoid arthritis, mesenchymal stems cells would be ideal. In vitro, biological stimuli are provided either in the form of a bioreactor, which is a device capable of mimicking the physiological environment in vivo, or by functionalising the tissue scaffold with biomolecules typically found in the ECM, such as proteins, peptides, growth factors and polysaccharides (this will be explored more in section 1.2).[6,10] In this research we focus specifically on modifying the tissue scaffold with bioactive motifs to generate tissue in vitro.

### 1.1.2. Structure and function of the native extracellular matrix.

The ECM is a highly hydrated, highly organised, biological scaffold that makes up the interstitial space between cells; it provides not only physical scaffolding to cells but also a combination of biochemical and biomechanical signals to direct cell behaviours, such as, adhesion, proliferation, differentiation, and migration as well as maintenance of the physiological microenvironment.[11,12] The ECM binds cell signalling molecules such as growth factors, proteins, polysaccharides etc and interacts with cell-surface receptors (integrins) to stimulate signal transduction pathways which regulate gene expression and direct cell behaviour.[3,13] The structure and composition of the ECM is unique to each specific tissue but is typically made up of collagen, elastin, glycoproteins, and proteoglycans as well as carbohydrates.[6,9,14] The ECM is a dynamic environment: it undergoes frequent remodelling via proteolytic degradation, triggered by signals sent out by the cells - this is an essential property for tissue repair and cell migration.[13,14] Disruption to the structure and composition of the ECM can lead to tissue degeneration - as we age the ECM structure begins to change, collagen fibres are "inappropriately crosslinked" which leads to a more rigid tissue with weaker mechanical properties and lower elasticity.[14,15] The ECM provides mechanical cues to

influence cell behaviour, and so if the structure of the ECM is compromised it can lead to abnormal cell behaviour, disruption of the tissue structure and ultimately tumour growth.[14,15]

There are many different biomolecules in the ECM that have structural and/or functional roles. Growth factors are extracellular proteins that bind cell-surface receptors; they transmit signals to regulate both cellular activities and gene expression, which promotes cellular differentiation and tissue growth.[6,13] They are often non-covalently bound to the ECM, via glycoproteins and proteoglycans, for stability and bind to specific cell surface receptors to initiate their function.[3,9] There are hundreds of different types of growth factor including vascular endothelial growth factor (VEGF), which plays a major role in the development and growth of new blood vessels, as well as epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) to name a few.[2,13] They can be both anti- and pro-proliferative, responsible for encouraging and inhibiting growth; cell signalling is a very complex process and many growth factors are active at once, simultaneously giving cells many different instructions.[13,16]

In addition to growth factors, other ECM molecules play important roles; glycosaminoglycans and proteoglycans provide a reservoir for growth factors and cytokines, binding them to the ECM and regulating their activity.[9,13] However, they are also responsible for water retention and occupy a large majority of the interstitial space, giving rise to the gel-like properties of the ECM.[14,16] Fibronectin is a glycoprotein and has both structural and operational properties; in addition to growth factor binding, fibronectin promotes cellular adhesion as well as cell migration, differentiation, and wound healing.[3,14,16] Structural fibrous proteins, on the other hand, such as collagen, elastin, and fibrin provide tensile strength to the ECM and support and direct cell adhesion, migration, and tissue development.[9,11] Elastin also provides the ECM with elasticity - which is important in tissues, such as muscle, that undergo repeated stretching.[9,14]

Cells bind to the ECM though transmembrane receptors, known as integrins; integrins recognise the tripeptide sequence arginine-glycine-aspartic acid (RGD) present in ECM proteins, such as collagen, fibronectin and laminin. Upon binding of an ECM protein to the integrins of the cell membrane, biochemical signals are transduced to the nucleus to mediate gene expression.[3,13] Cadherins are also cellular transmembrane receptors, however they regulate cell-cell interactions within the tissue.[2]

### 1.1.3. How scaffold structure can mimic the native ECM to guide cell fate.

The main aim of biomaterial scaffolds is to mimic the functions of the native ECM, however not all extracellular matrices are alike - as mentioned earlier the structure and composition of the ECM is unique to the specific tissue in question.[9,14] For example, cartilage tissue does not have blood vessels; it is unlikely that VEGF, which induces angiogenesis, will be present in the ECM of cartilage tissue. Consequently, VEGF won't need to be incorporated into scaffolds intended for cartilage regeneration.[12] Although a major function of an artificial tissue scaffold is to provide the right biochemical cues to stem cells, the biomechanical properties of the scaffold are just as important. The mechanical stiffness of the scaffold plays a significant role in stem cell differentiation on a transcriptional level as cells can sense biomechanical forces – a stiffer matrix will drive stem cells towards an osteogenic (bone) lineage whereas softer matrices will drive them towards adipose (fat) tissue.[12,17,18] Matrix stiffness is also reported to affect cell spreading, proliferation and migration – the stiffness of the scaffold can be tuned precisely by altering crosslinking density and by the choice of precursor materials used.[19,20]

Cell fate also relies heavily on the biomolecules presented to them by the artificial cell scaffold – one of the most basic functions required is the presentation of cell adhesive motifs that bind cells to the scaffold. When materials are non-adherent, cells will adopt a rounded morphology but will spread and elongate in an ideal conformation on an interactive surface.[21] The incorporation of ECM biomolecules into artificial tissue scaffolds will be explored more in section 1.3.

## 1.2. Hydrogel Tissue Scaffolds.

### 1.2.1. Hydrogel properties.

Hydrogels are highly water-swollen 3D networks of crosslinked hydrophilic polymers and are a popular choice for tissue scaffolds due to their structural properties which mimic the highly hydrated native ECM.[9,12,20] They are primarily used in soft tissue regeneration, due to their mechanical similarities; hydrogels are well suited to cartilage and muscle tissues, making them ideal scaffolds on which to regenerate tissue for rheumatoid arthiritis.[6] Hydrogels can be natural, synthetic, or biosynthetic; natural hydrogels, such as collagen and fibrin, are

biocompatible, exhibit good porosity, permeability, biodegradability and structural biocompatibility towards tissues but often lack mechanical strength which results in unstable hydrogels over extended periods.[9,12] Synthetic hydrogels, such as poly (acrylamide) (PAM) and poly (ethylene glycol) (PEG), although inherently non-bioactive, can be easily modified with pre-selected biomolecules to tailor specific function, are highly tuneable and possess good mechanical stability and permeability that can be altered through crosslinking.[9,21]

The choice of hydrogel used for a biomaterial scaffold ultimately comes down to the target tissue they will be used for and should be designed with a desired outcome; hydrogel structure and physiochemical properties can be tuned by the choice of biomaterial used, the crosslinking methods and fabrication strategies employed.[20,21] Tissues, such as cartilage require bio-inductive effects to induce chondrogenic differentiation towards cartilage formation and low porosity due to the absence of blood vessels in cartilage tissues.[21] Whereas muscle tissue, which contains many blood vessels to meet the demand for oxygen, requires high porosity to enable adequate vascularisation. The porosity of the scaffold can be tailored through crosslinking. Stiffer hydrogels with higher crosslinking exhibit lower swelling and smaller pore size – porosity not only effects vascularisation, but also diffusion of nutrients and oxygen, removal of waste products, and cell migration (pores <10 μm in diameter inhibit cell movement).[19,21] On the other hand, crosslinking provides hydrogels with mechanical stability and imparts control over degradation rates – covalent crosslinking provides scaffolds with structures with higher resistance to mechanical forces, whereas physical crosslinking (i.e. hydrogen bonding, ionic interactions or van der Waal's interactions) provide scaffolds with viscoelastic behaviour, allowing the structure to recover from strain and enable remodelling.[12]

### 1.2.1.1.    Natural hydrogel scaffolds.

Natural polymers include polysaccharides (chitosan, cellulose, alginate) and proteins (collagen, gelatin, fibrin, elastin).[6,9] Natural polymers are often inherently bioactive and are therefore able to promote cell adhesion and proliferation. They also more closely mimic the native ECM in comparison to their synthetic counterparts.[9,13,20] On the other hand, they often have poor mechanical properties, and suffer from batch-to-batch variability, difficulties with homogeneity, unregulated cell-protein interactions, and rapid degradation under

physiological conditions.[6,13] Furthermore, these polymers are commonly sourced from animals or microorganisms, leading to risks of contamination, diseases, and immunogenicity.[11,13] The biodegradability of natural polymer scaffolds can be an advantage. As mentioned above, host cells need to degrade the scaffold over time to both replace it with their own ECM and continue tissue growth uninhibited, however controlling the degradation rate can still prove difficult.[1,8]

### 1.2.1.2. Synthetic hydrogel scaffolds.

In comparison to natural polymer scaffolds, synthetic polymers possess many advantages; they often have excellent mechanical properties that can be easily tuned, the scaffold architecture, chemical composition and degradation rate are all controllable, they lack immunogenicity, and they can be homogeneously and reproducibly fabricated.[9,13] Although they lack bioactivity and cell adhesive ligands, they can undergo facile modification with bioactive molecules which allows them to selectively provide bioactive groups and be tailored precisely to function in a specific manner, as opposed to natural (bioactive) polymers which may provide unwanted signalling to cells.[6,9,13] Common synthetic polymers used for hydrogels include, but are not limited to, poly(caprolactone) (PCL) poly (vinyl acetate) (PVA), poly(acrylamide) (PAM), poly(hydroxyethyl methacrylate) (PHEMA) and poly(ethylene glycol) (PEG) (as shown in figure 1.).[13,20] Some synthetic polymers aren't biodegradable (e.g. PEG), however, by incorporating degradable crosslinks into the scaffold architecture, the materials can be broken down into molecules small enough to be excreted by the renal system.[8] Polyesters contain biodegradable ester linkages, however they are hydrophobic and consequently don't form very good tissue scaffolds on their own. Fortunately, polyesters can be combined as amphiphilic block co-polymers to impart biodegradability to a hydrophilic polymer scaffold, such as PEG or PVA.[21]
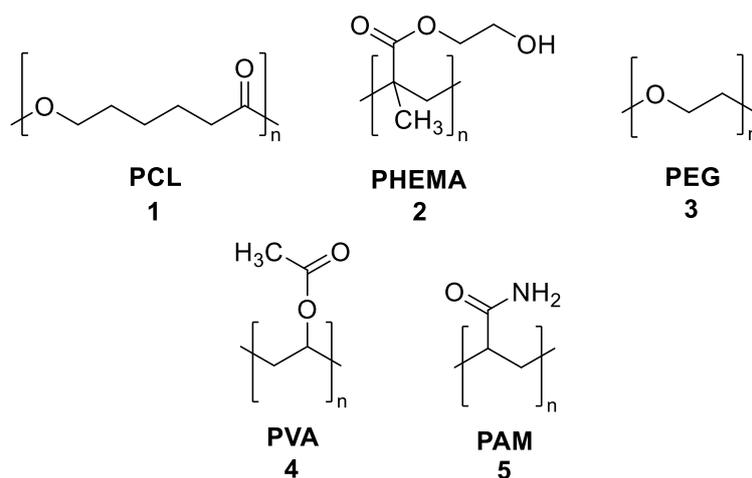
**Fig. 1.** Macromers of common synthetic polymers.

## 1.2.2. PEG hydrogel scaffolds.

PEG is the most frequently employed synthetic polymer in tissue engineering; PEG is bioinert, non-immunogenic, non-toxic, hydrophilic, displays cytocompatibility, and is very easily modified to install degradable linkers.[12,19] The most valuable property of PEG is its chemically and biologically inert nature, which provides a "blank slate" for modification and acts to minimise inflammatory response and non-specific protein binding.[19,21] The fundamental PEG structure is PEG diol with two hydroxyl end groups which can be easily modified and transformed into other functional groups.[13] The hydrophilic nature of PEG makes it a great base material to form hydrogel tissue scaffolds, however it also prevents natural cellular adhesion – this is overcome simply by the incorporation of cell adhesive motifs, most commonly peptides containing RGD (Arg-Gly-Asp) or IKVAV (Ile-Lys-Val-Ala-Val) sequences.[8,20,22] PEG can take the form of a linear chain, branched multiarmed-, or star-structures – hyperbranched and star macromers have increased interactions between polymer branches as opposed to linear chains increasing mechanical stability.[13] Branched structures have the added benefit of increased modification capability – the hydroxyl end groups on PEG chains are used for crosslinking and with star-macromers, for example, "sacrificial modification" of hydroxyl groups to introduce bioactivity has a reduced impact on the bulk properties of the material than would be observed for linear chain structures.[21,23] In its pure form, PEG is non-biodegradable however, degradable segments (e.g. Ester and disulphide linkages) can be incorporated into the structure.[13]

These advantageous properties of PEG make it an ideal polymer scaffold for our research. Specifically, its biologically inert properties make it particularly useful. In order to accurately test the effectiveness of our N-terminus binding molecule we need to make sure that the unmodified base polymer won't interact with and immobilise proteins and thereby make our results inconclusive. Additionally, we will be using an 8-arm PEG macromer which will allow us to use the end group of each arm either for crosslinking into a hydrogel or for protein immobilisation without sacrificing scaffold integrity.

## 1.3.  Introducing bioactivity.

One of the main ways to impart bioactivity on a material is through functionalisation with proteins and peptides. No one singular biomolecule can mimic the signalling complexity of the native ECM and often several different conjugation strategies need to be employed in order to successfully immobilise the range of biomolecules required for stem cell immobilisation, differentiation, and proliferation.[24] The suitability of a particular conjugation technique for protein immobilisation should be considered, which is determined by: site-selectivity of conjugation, accessibility of reactive handles on the protein to be attached, the ability to maintain bioactivity and proper function following conjugation, and the ability to control release profiles.[25,26] Ideally protein conjugation should take place in a chemo-selective manner, under aqueous conditions, whilst maintaining the native conformation of the protein.[27,28] Very often in the biomaterials community, protein immobilisation strategies will indiscriminately target abundant amino acids on the protein surface resulting in heterogeneous mixtures of protein orientations leading to a reduction in bioactivity.[18,29-31] In this section, we will discuss the most frequently used methods for protein immobilisation and highlight the specific motifs/amino acid residues commonly targeted as sites for conjugation.

### 1.3.1.  Conventional methods of protein immobilisation.

Protein immobilisation on a material surface can be carried out via physical interactions (hydrophobic, polar, ionic and van der Waal's interactions) or via covalent bonding. Typically, physical interactions tend to be weak and poorly controlled.[26] Covalent protein immobilisation generally results in better site-specificity by targeting a particular functional group, such as amines, alcohols, or carbonyls within amino acid side chains. However, proteins often contain many copies of the same amino acid residue on their surface which

can, as stated above, lead to heterogeneous binding profiles with the potential to block active sites.[25,27,31] The most common sites on proteins for site-selective conjugation are the cysteine and lysine residues, although conjugation to other functional motifs is possible.

### 1.3.1.1. Cysteine.

Targeting cysteine residues can be advantageous because of their low abundance on protein surfaces. However, the rarity of cysteine implies that they may play an important role in protein activity in which case, targeting the cysteine residue could be actively detrimental.[27,32] Of the 20 amino acids in the human body, cysteine residues are the only amino acid to contain a thiol group – this makes cysteine a particularly appealing target. Thiolate anions can undergo Michael addition reactions with $\alpha,\beta$-unsaturated carbonyls on the surface of a scaffold, in a highly specific reaction that proceeds under mild aqueous conditions.[24] Alternatively, thiols can react with an activated alkene attached to a scaffold, to form an alkyl sulphide via a nucleophilic catalysed reaction or a free-radical initiated mechanism (as shown in scheme 1.).[27]
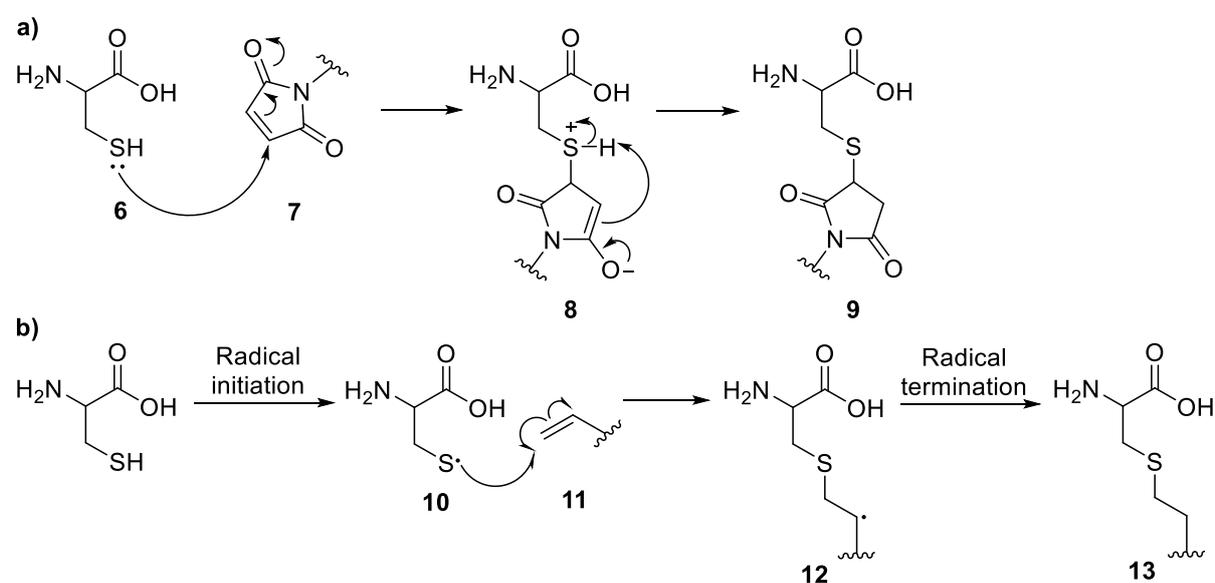


**Scheme. 1.** a) Michael-addition of a cysteine thiol to an $\alpha,\beta$-unsaturated carbonyl, b) Radical initiated thiolene reaction.

### 1.3.1.2. Lysine.

Amines are present on lysine side chains and the N-terminus of polypeptides/proteins. The high abundance of lysine residues on peptide and protein surfaces poses a major drawback due to the heterogeneity of conjugation products produced, which in turn leads to drops in bioactivity or function.[24] Additionally, most amine reactive functional groups will

preferentially react with the strongest nucleophile, which in some cases is not the amine.[33,34] Amines can undergo conjugation via amide formation which usually proceeds by activating the carboxyl-coupling partner (the tissue scaffold) with an NHS ester before conjugation of the amine-containing protein (as shown in scheme 2.).[33] Additionally NHS esters are hydrolysed in aqueous conditions which poses limitations for their use in hydrogels.[33] Alternatively, amines are able to react reversibly with aldehydes to form imines with a loss of water, however, in an aqueous environment the reverse reaction is strongly favoured – a second step can be implemented in which sodium cyanoborohydride irreversibly reacts with the imine to form the corresponding amine as a more stable product (as illustrated in scheme 2.).[35] Each copy of an amino acid residue experiences a unique environment dictated by their neighbouring residues. Consequently, a single type of amino acid can experience a range of different pKa values within the same protein – this can be exploited by altering the pH of a reaction to favour a specific residue.[28]
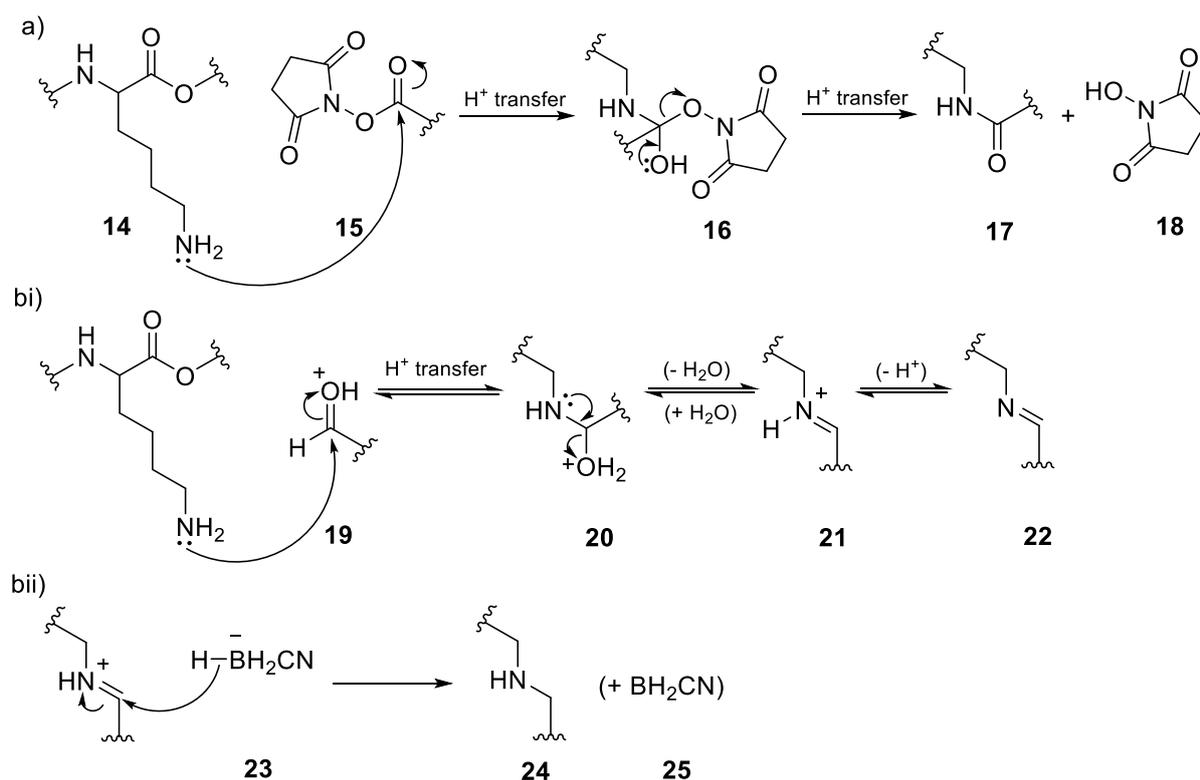


**Scheme. 2.** a) Amide bond formation between a lysine side chain and an NHS-activated carbonyl, bi) Reversible imine bond formation between a lysine side chain and an aldehyde, bii) Irreversible amine formation by reaction between an iminium ion and a cyanoborohydride anion.

## 1.4. N-terminus reactivity.

When it comes to targeting an amine motif on a protein, lysine side chains and the N-terminus of the protein (as shown in figure 2.) compete with each other. The N-terminal amine has a slightly lowered $pK_a$ ($N^\alpha$-$NH_2$, $pK_a$ 8) when compared with lysine side chains ($N^\varepsilon$-$NH_2$, $pK_a$ 10) due to the inductive effects of the nearby carbonyl; this means that the α-amine is less likely to be protonated at physiological pH and therefore experiences an increased reactivity towards electrophiles.[28,36] Additionally, the adjacent amide group on the peptide backbone, gives the N-terminal amine unique chemistry.[37] Controlling the pH of the medium enables site-selectivity of conjugation by increasing the reactivity of the N-terminal amine relative to the lysine side chains. At physiological pH the N-terminal amine is more reactive than the ε-$NH_2$ of lysine as lysine exhibits a higher degree of protonation making it a weaker nucleophile.[24,32] Lowering the pH favours the reactivity of the N-terminus, as α-amines can remain unprotonated at a lower pH than lysine amines, however this comes at the expense of reduced conjugation yields, as lowering the pH will increase protonation.[36] Single chain proteins possess just one N-terminus which is typically solvent exposed ($\sim$ 80%) making this site an excellent candidate for site-specific protein conjugation.[28,36] In addition N-terminus modification has minimal effect on the bulk protein structure, due to its accessible nature and typical lack of contribution to intramolecular protein binding.[36] Targeting the N-terminus of proteins for site-specific immobilisation is a very attractive technique; every protein and peptide, unless chemically modified, possesses an N-terminal amine and so mechanisms that specifically target the N-terminus can be applied to a wide variety of proteins.



**26**
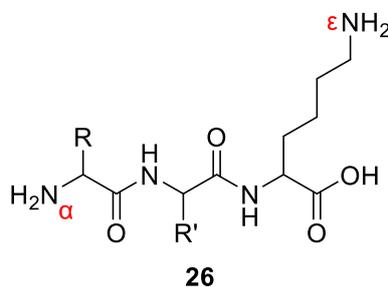
**Fig. 2.** Simplified peptide structure showing the α-$NH_2$ of the N-terminus and the ε-$NH_2$ of the lysine side chain (R and R' represent amino acid side chains, they can be the same or different).

### 1.4.1. N-terminus-specific reactions.

Exploiting N-terminus specific chemistry is not a new practice. Gaertner and Offord reported a method as early as 1996 to generate a reactive carbonyl motif (in the form of a glyoxylyl group) at the N-terminus of proteins comprising N-terminal serine or threonine via periodate oxidation - the carbonyl-functionalised N-terminus is then able to conjugate to aminooxy-functionalised PEG molecules via an oxime linkage (as shown in figure 3.).[38]
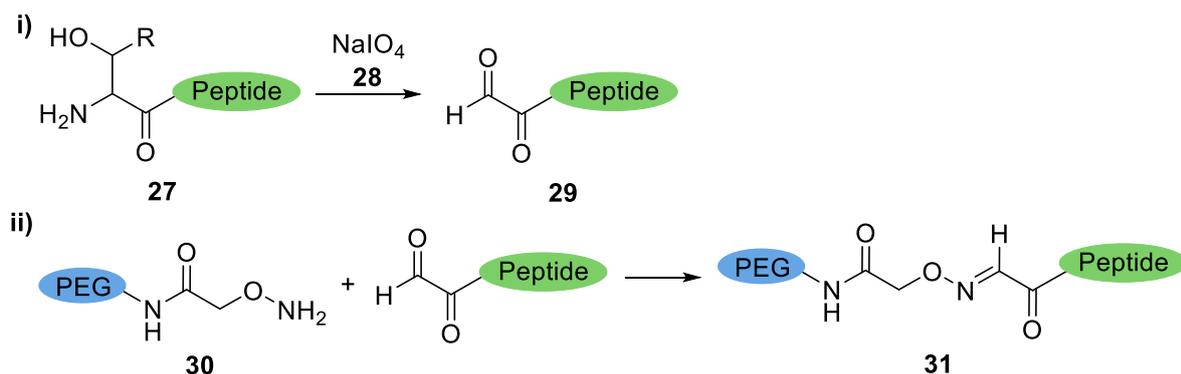


**Fig. 3.** i) Periodate oxidation of a serine (R = H) or threonine (R = CH$_3$) to form an N-terminal aldehyde, ii) immobilisation of a peptide containing an N-terminal aldehyde on an aminooxy-functionalised PEG molecule.
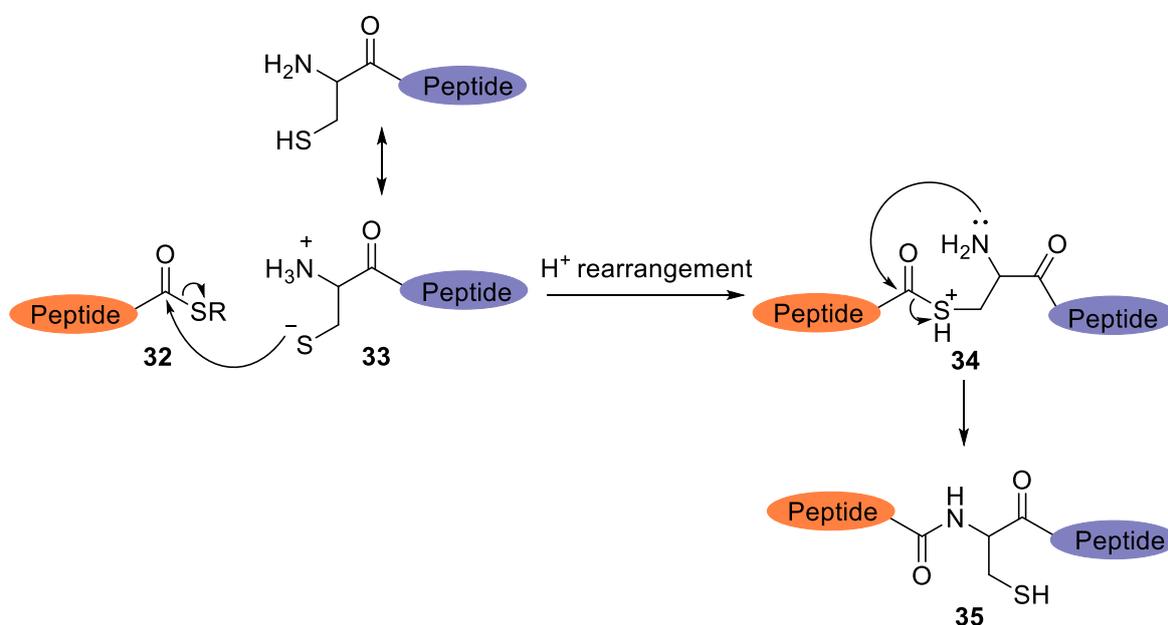


**Scheme. 3.** N-terminal cysteine-specific reaction with a modified-protein C-terminal thioester.

Specific N-terminal amino acid residues (cysteine, serine, and threonine) can be selectively targeted for modification, but conjugation relies on the protein chain having the specific N-terminal amino acid for each of the respective conjugation reactions to be made.[32] In 1994, the Kent group developed a reactive mechanism in which the thiol of an N-terminal cysteine attacks a thioester at the C-terminus of a separate peptide in a nucleophilic substitution reaction (as shown in scheme 3.).[39] Tam et al., reported a method of reacting N-terminal cysteine, serine and threonine peptides with aldehydes via thiazolidine (Cys) and oxazolidine (Ser and Thr) ring formation (as shown in scheme 4).[40] Targeting a specific N-terminal amino acid residue limits the type of bioconjugation technique that can be used. N-terminal amino acid residues vary depending on the protein and selective targeting often requires several steps in the bioconjugation reaction and long reaction times.[29,36]
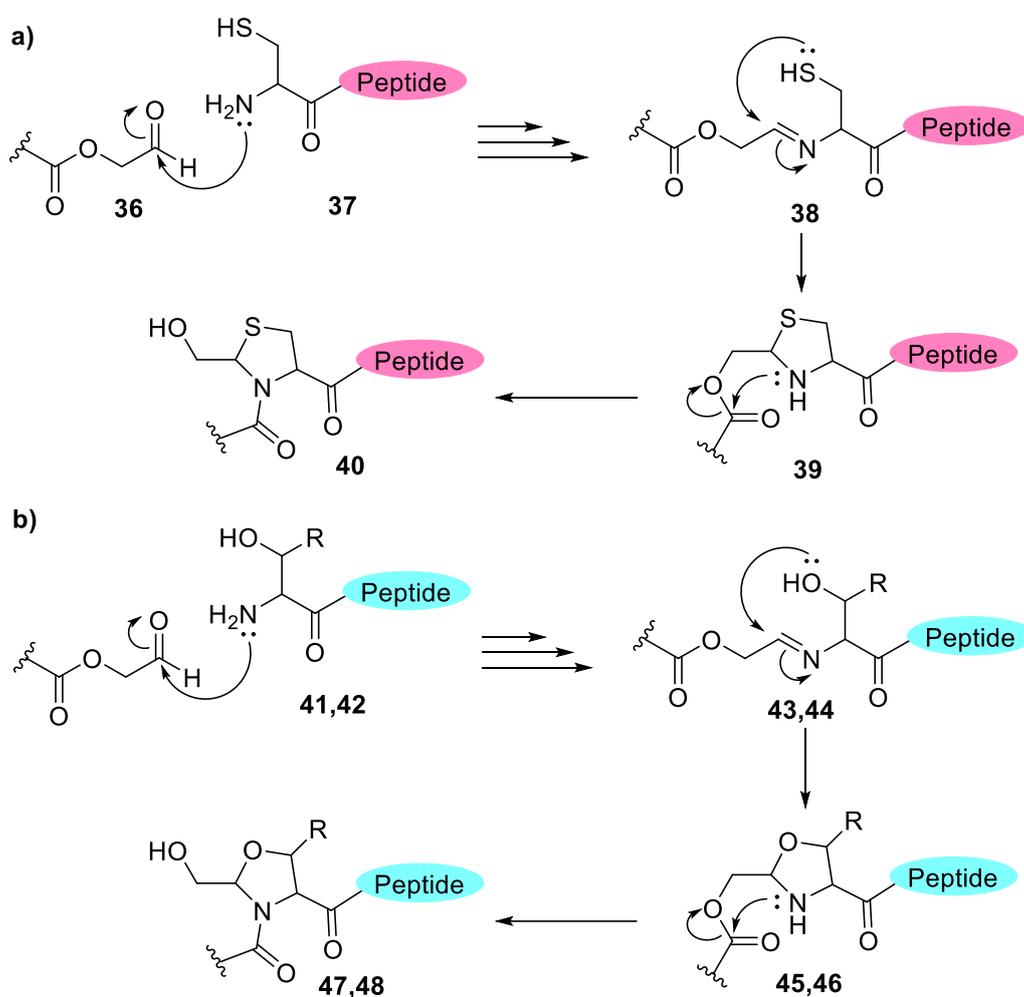


**Scheme. 4.** a) N-terminal cysteine-specific reaction with an aldehyde to form a cyclic thiazolidine conjugation product, b) N-terminal serine- and threonine-specific reaction with an aldehyde to form a cyclic oxazolidine conjugation product. For serine R = H (**41, 43, 47, 45**) and threonine R = CH$_3$ (**42, 44, 46, 48**).

Deng et al., have recently developed a one-step N-terminus modification of peptides and proteins using 2-ethynylbenzaldehydes (2-EBA) under mild reaction conditions by controlling pH of the medium.[41] Cyclic isoquiloliniums are formed from this reaction as a result of nucleophilic attack from the initially formed N-terminal imine to the alkyne group on 2-EBA (as shown in scheme 5).[41] Increasing the pH of the medium reduced site selectivity and lysine side chains began to react, and the same trend was observed at low temperatures.[41] The reaction relies on pH control to direct conjugation to the N-terminus, the isoquinolinium product is able to form at the lysine side chains and so the method is not entirely specific to the N-terminus. N-terminus reactions that rely on pH are difficult to control and can result in heterogeneous conjugation products. Additionally, the mechanism developed by Deng et al isn't site specific due to the ability of the lysine side chain to participate in the reaction.
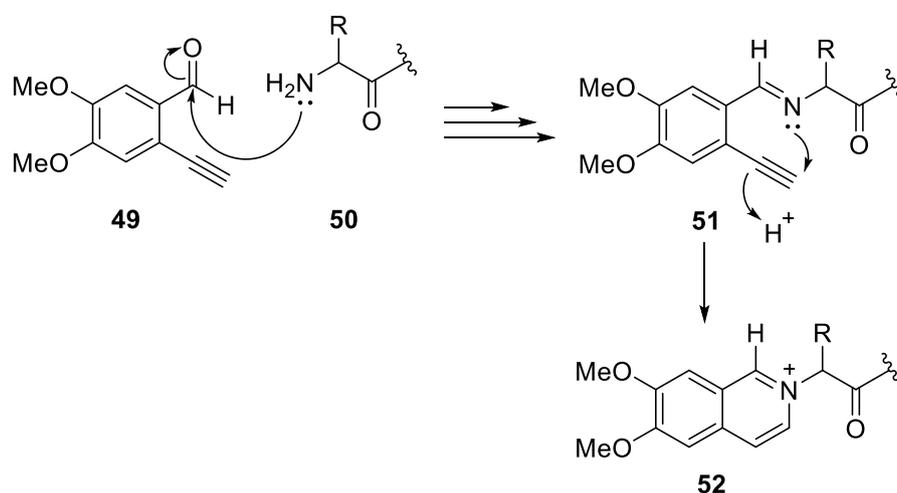


**Scheme. 5.** Reaction between an N-terminal amine and an aldehyde on a 2-EBA molecule, **49**, resulting in the formation of a cyclic isoquinolium product (R = amino acid side chain).

### 1.4.2. Francis Group development of N-terminus-specific 2PCA chemistry.

In 2015, the Francis Group, reported a novel procedure to selectively modify the N-terminus of proteins, in an aqueous environment, under biocompatible temperature and pH, which was able to modify proteins whilst maintaining their structural integrity and function.[29] The method developed is applicable to a wide range of peptides and proteins as it is not specific to one type of N-terminal amino acid residue. Additionally, it does not rely on pH control for N-terminus selectivity over lysine side chains. The procedure introduced 2-pyridine carboxaldehydes (2PCA) that initially form an imine with the $\alpha$-NH$_2$ of the N-terminus in a reversible reaction. This imine can then undergo subsequent cyclisation to form a stable cyclic

imidazolidinone product.[29] The cyclic imidazolidinone forms as the nucleophilic amide nitrogen on the peptide backbone attacks the electrophilic carbon of the N-terminal imine (as shown in scheme 6.). The ε-amino groups on the lysine side chain are unable to undergo cyclisation in this way due to the lack of a neighbouring amide group (as shown in scheme 7.).[18,29] The next closest amide on the resulting imine (58) could cyclise to form a 9-membered ring, however this is disfavoured due to the high ring strain of the resultant cyclic imidazolidinone product.
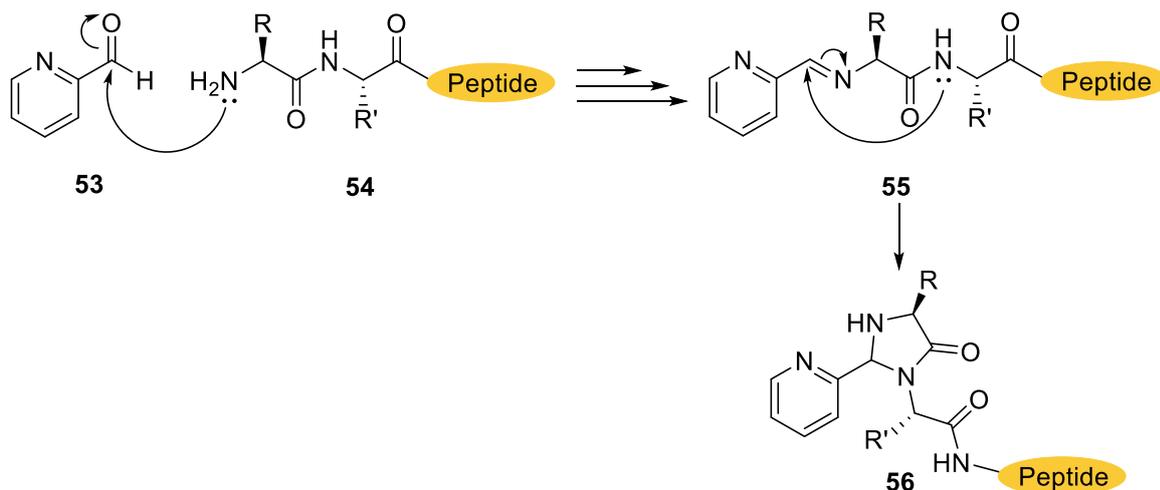


**Scheme. 6.** Selective modification of the N-terminal amine by 2PCA to form a stable cyclic imidazolidinone product (R and R' are amino acid side chains – R' ≠ proline side chain).

The reaction was shown to selectively modify the N-terminus of the protein despite the presence of several lysine side chains. This reaction requires no prior protein modification with the only requirement, that the N-terminus is solvent-exposed.[29,31] The two limitations are that: (1) imidazolidinone formation cannot occur when proline is the second amino residue in the peptide backbone as it lacks a primary amide group for cyclisation, and (2) for protein immobilisation purposes the protein must be single chain. A double chain protein, for example, could be immobilised at both N-termini in two separate directions which could cause denaturation if the immobilisation locations on the scaffold were far enough apart.[28,29]
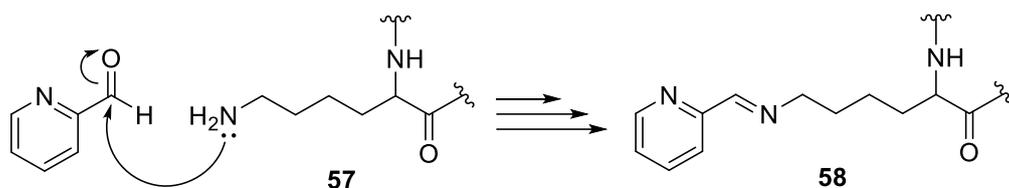
**Scheme. 7.** Modification of ε-NH$_2$ on a lysine side chain by 2PCA – due to the lack of neighbouring amide, the reaction stops at imine formation.

The Francis Group have subsequently developed this chemistry further. Lee et al., reported that a 2PCA-acrylamide conjugate could be polymerised into an acrylamide/bisacrylamide gel.[18] The 2PCA-acrylamide subunits were incorporated into the backbone of the poly (acrylamide) chain (as illustrated in figure 4.). Full length ECM proteins were then combined with the gel via selective reaction between the N-termini of the protein and the aldehyde of the 2PCA moiety to form 2PCA-acylamide-protein conjugates in which protein function was retained.[18] Using 0-0.1% stoichiometry of 2PCA-acrylamide relative to acrylamide monomer content, they functionalized the surface of the materials with full-length collagen, fibronectin and laminin. They then proceeded to seed the materials with cells and observed a direct increase in cell spreading in correlation with an increase in 2PCA-acrylamide content – from this they concluded that higher 2PCA content confers a higher concentration of immobilized ECM proteins.[18] This research introduces the ability to selectively control both immobilization and concentration of ECM proteins in a poly(acrylamide) gel scaffold through controlled modification by a 2PCA-containing moiety.
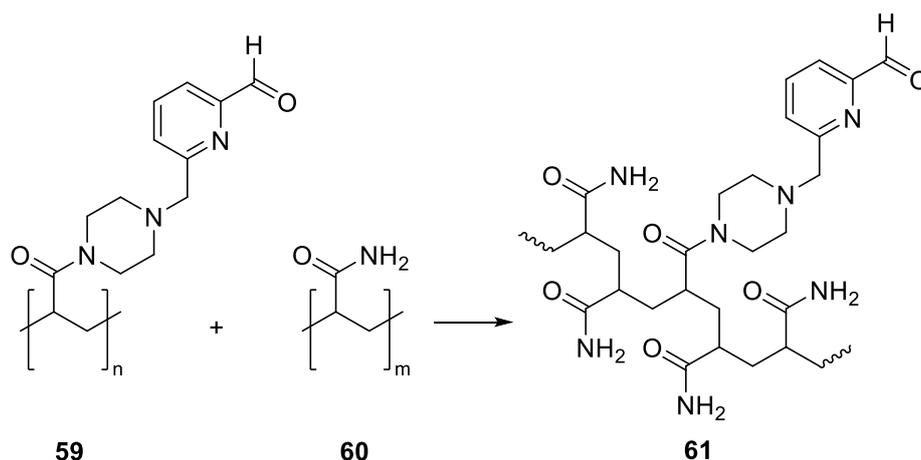


**Fig. 4.** Polymerisation of 2PCA-acrylamide with acrylamide to form a 2PCA-functionalised PAM network (n ≥ 1 and m > n).

Koo et al., later reported that proteins could be immobilised on a poly (ethylene glycol) acrylamide (PEGA) resin (as shown in scheme 8.) using a derivative of 2PCA.[26] Interestingly,

Imidazolidinone formation was found to be reversible, and upon removal of the excess free 2PCA reagent from the mixture the equilibrium favoured the unconjugated starting materials.[26] This allows the reversible immobilisation of proteins to scaffolds and is particularly useful for tissue engineering as different proteins are required at different stages in cell development – this mechanism imparts the ability to selectively attach and remove proteins when required. Hydroxylamine was used to competitively react with the aldehyde moiety on 2PCA, acting to "cap" the reactive handle, thereby releasing the previously immobilised proteins.[26] When the 2PCA reagent is immobilised on the PEGA scaffold the concentration of free 2PCA is fixed and therefore, given an excess of 2PCA motifs to proteins, there will be free unbound 2PCA on the scaffold and the forward reaction should be heavily favoured (as illustrated in figure 5.).[26]
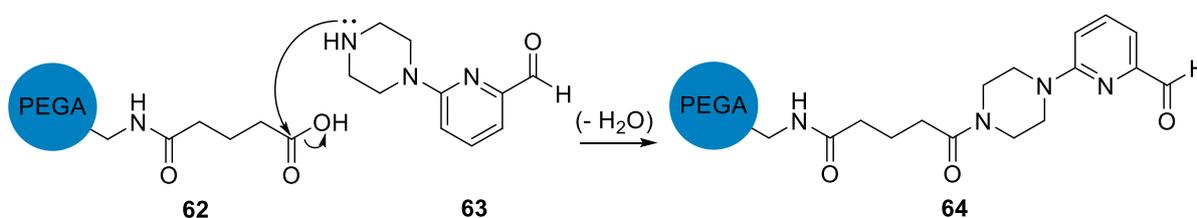


**Scheme. 8.** Modification of PEGA resin with a 2PCA derivative.
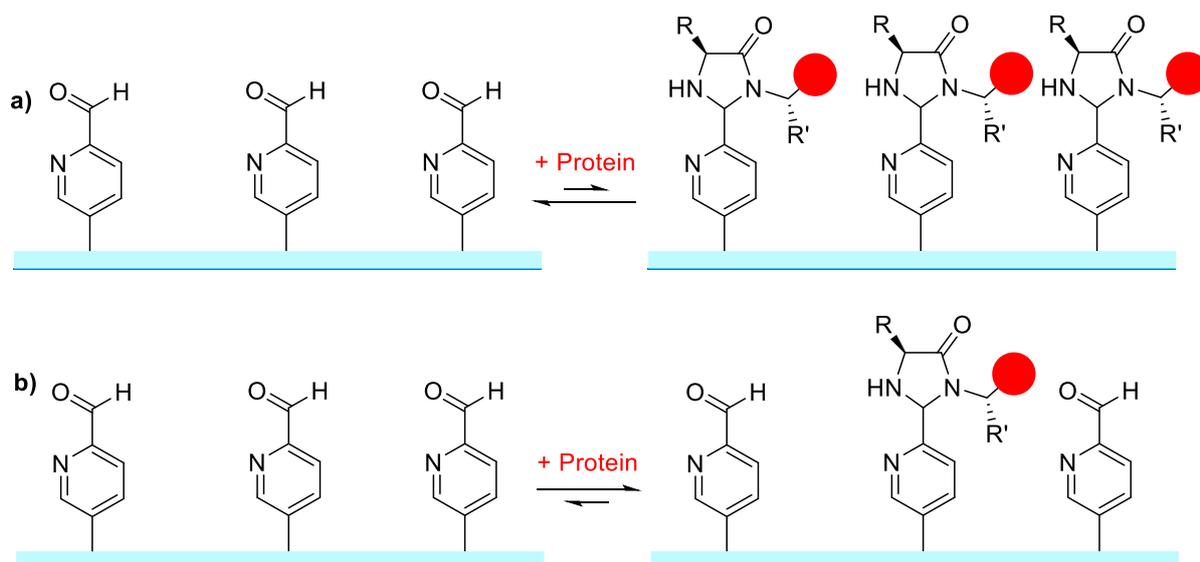


**Fig. 5.** A simplified illustration of how stoichiometry of peptide: 2PCA will favour the forward/reverse reaction. a) A 1:1 ratio will favour the reverse reaction, b) a lower ratio of peptide: 2PCA will favour the forward reaction (R and R' are amino acid side chains – R' ≠ proline side chain).

Additionally, Sangsuwan, Tachachartvanich and Francis, utilised 2PCA chemistry to conjugate amphiphilic polymers to proteins at their N-terminus via imidazolidinone formation. Green fluorescent protein (GFP) was used as the conjugated protein which underwent cytosolic delivery into the cell cytoplasm on account of the conjugation to the amphiphilic polymer, thereby providing evidence of the biocompatibility of 2PCA moieties with cells.[31]
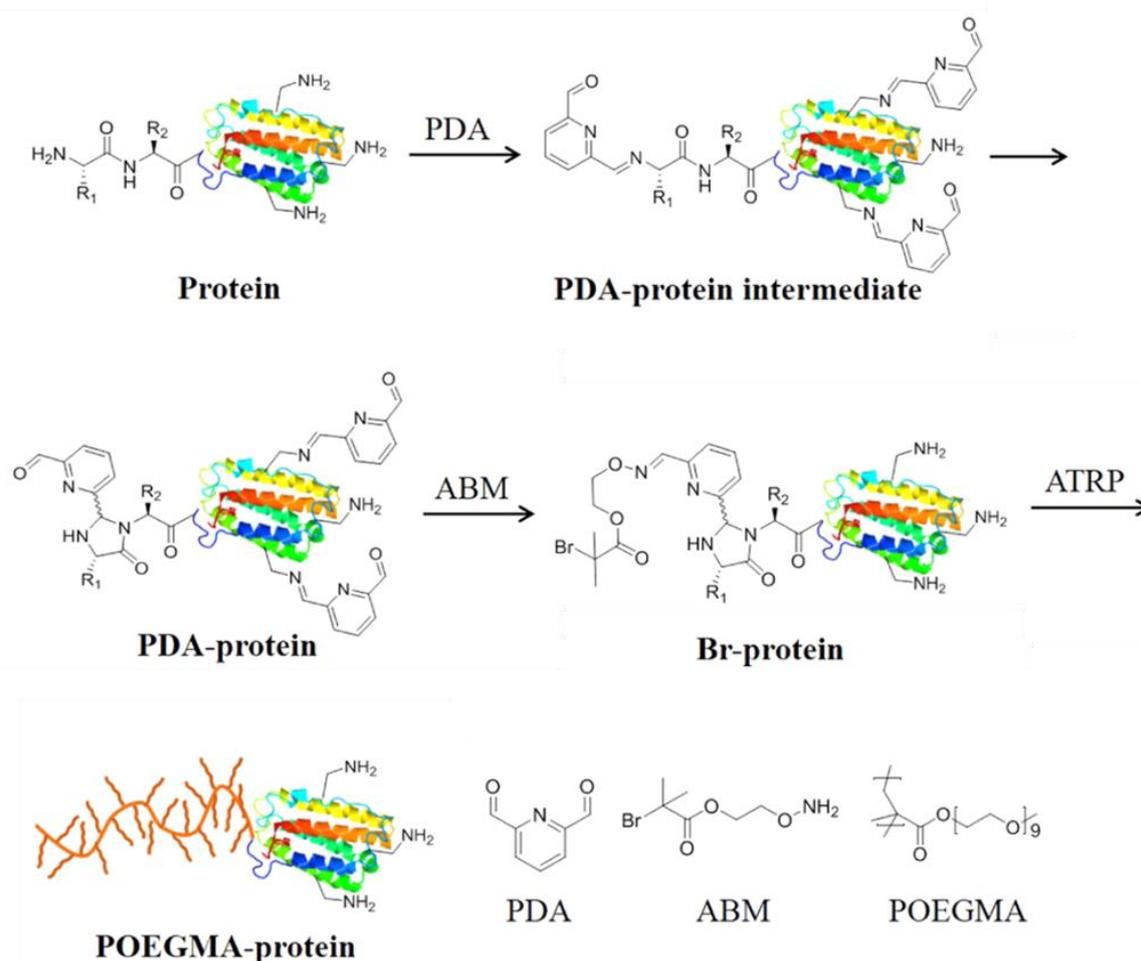


**Fig. 6.** Reproduced from Sun et al., to illustrate the use of PDA in modifying proteins with peptides in a grafting-from approach.[42]

This cutting edge 2PCA chemistry has been modified and utilised by other groups in the selective modification of proteins at the N-terminus; Sun et al., developed a method, using the chemistries established by the Francis group, to functionalise the N-terminus of a protein with pyridine-2,6-dicarboxaldehyde (PDA) (as shown in figure 6.).[42] One of the aldehydes of the PDA conjugates to the N-terminus whilst the other is subsequently functionalised with 2-(aminooxy) ethyl-2-bromo-2-methylpropanoate (ABM) which then undergoes atom transfer radical polymerisation (ATRP) with oligo(ethylene glycol) methyl ether methacrylate

(OEGMA).[42] As the ABM is added to the reaction, it outcompetes the lysine side chains to react with the PDA molecules, resulting in some molecules of PDA which contain an oxime linkage to ABM on both sides of the molecule.[42] This method enables protein conjugation using the unique reactivity of the protein N-terminus with PDA to form a protein-polymer complex in a grafting-from approach.[42]

In summary, synthetic polymer scaffolds are highly advantageous for use as scaffolds in tissue engineering, however their lack of bioactivity and cell adhesive properties is a major drawback. Protein immobilisation on the scaffold surface is a great means of imparting this bioactivity. The research developed by the Francis group provides a technique capable of immobilising proteins site-specifically in a manner which preserves their activity. The N-terminus specificity of the reaction allows it to be applied to a broad range of peptides and proteins, an ability which other immobilisation methods lack often due to their need for a particular amino acid. Several groups have modified this chemistry for protein tagging, immobilisation, and to increase the solubility of certain proteins. However, to the best of our knowledge, there are currently no reports of using 2PCA derivatives to both immobilise proteins on the scaffold and participate in crosslinking of the scaffold at the same time.

## 1.5.   Project aims.

The aim of this project was to adapt the 2PCA chemistry developed by the Francis group, to create novel 2PCA derivatives capable of binding to a range of different synthetic scaffolds. This project focuses on the synthesis of 2PCA-SH **65** and 2PCA-NH$_2$ **66** (as shown in figure 7.), as well as their incorporation onto the ends of 8-arm PEG macromers to be used in hydrogel scaffolds. The 2PCA molecules would have two functions: (1) to act as a bridge between the biologically inert PEG and the N-terminus of ECM cell signalling proteins, and (2) to crosslink modified-PEG macromers via a double ended N-terminal peptide crosslinker molecule. Additionally, we would synthesise 8-arm PEG-peptide macromers that would undergo subsequent crosslinking using molecules with a 2PCA moiety (**65** or **66**) on either end. By incorporating a larger proportion of 2PCA into the scaffold than N-terminal amines we would expect to observe a crosslinked hydrogel network with the presence of some free 2PCA motifs which could then be used to immobilise ECM signalling proteins. We aimed to observe conjugation between the free 2PCA and the N-terminus of peptides using a novel coumarin-based fluorophore **67** (as shown in figure 7.) that mimics the N-terminus of a peptide; upon

conjugation to the PEG-2PCA scaffold we would expect to be able to observe fluorescence of the material.
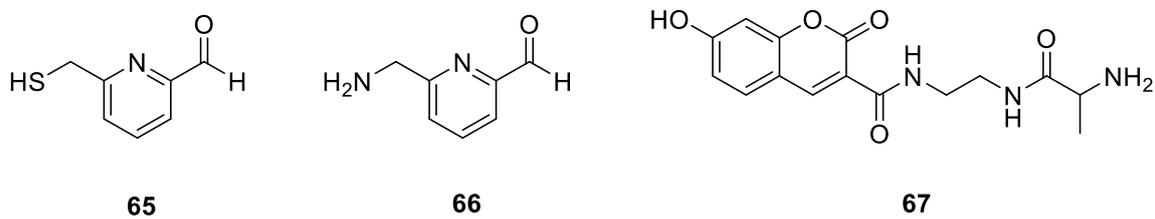


**Fig. 7.** Target molecules.

# 2.    Results and Discussion.

## 2.1. PCA Synthesis.

### 2.1.1. Summary.

The goal for PCA synthesis was to create two different novel derivatives of 2PCA, each with a functional group in the 6-position that was capable of binding the PCA molecule to a synthetic scaffold. The functional groups we chose to place in this position were an amine and a thiol respectively. As explored in section 1, there are many reactions that amines and thiols can take part in due to their nucleophilic natures, and it is for this reason that we chose these groups. Although, the aim of this project was to incorporate our 2PCA molecules to a PEG scaffold, the diversity of the reactions that amine and thiol groups can undergo makes 2PCA-$NH_2$ and 2PCA-SH applicable to different synthetic polymer scaffolds and therefore the scope for their use in future work is broad.

### 2.1.2. Installation of a methyl amine to the 6-positon of 2-pyridine carbaldehyde.

Initially, we began our path to synthesise 2PCA-$NH_2$ **66** using 2,6-pyridinedimethanol **68**; we first attempted to substitute one of the hydroxymethyl motifs for an amine following a procedure from a patent.[43] **68** in DCM was incubated with 0.5 M ammonia in THF at 0 °C, however NMR analysis showed no reaction had taken place. We proceeded by selectively oxidising one of the alcohol groups to form 2-pyridine carboxaldehyde (2PCA) **69** (Figure 8). We hypothesised that the hydroxymethyl group of **69** may be too unreactive to undergo nucleophilic substitution by ammonia, therefore, we activated the alcohol as a mesylate **70** to produce a better leaving group. Compound **70** was synthesised by reacting **69** with a slight excess of methanesulfonyl chloride in the presence of triethylamine and DCM - product formation was confirmed through both NMR and mass spectrometry. We subsequently attempted to react **70** with a 0.5 M solution of ammonia (in THF) but observed no reaction between the mesylate group and ammonia via ^{1}H NMR. We believe that, despite the reported procedure in the patent,[43] ammonia is too weak a nucleophile to attack the α-carbon even when activated as a mesylate leaving group. Retrospectively, we believe that the low temperatures used may have hindered the reaction, and perhaps heating the ammonia in solution may have improved reactivity.
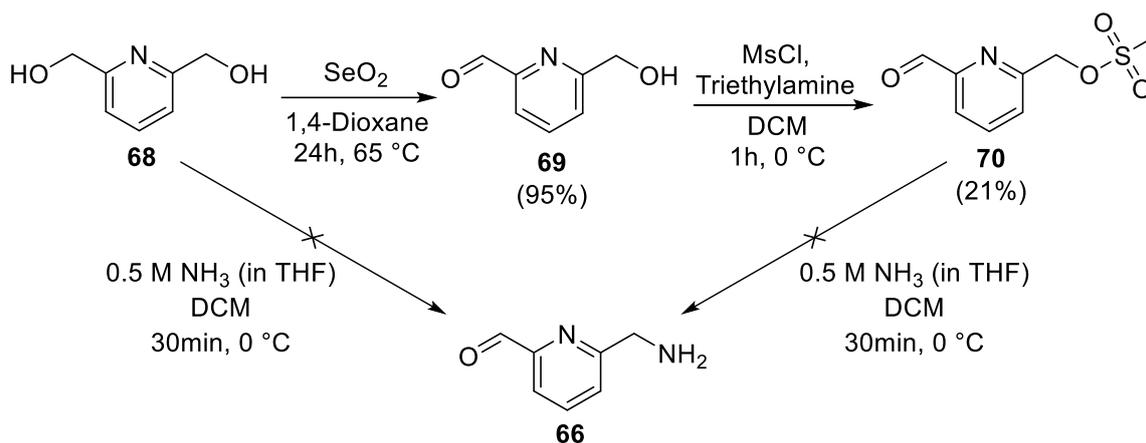
**Fig. 8.** Reaction pathway showing the initial attempts to functionalise 2PCA **(1a)** with an amine using NH₃ (in THF).
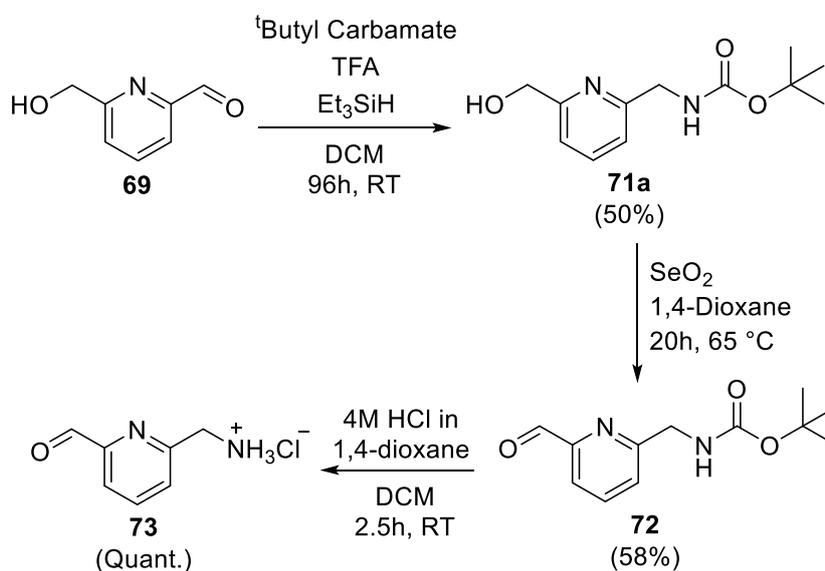


**Fig. 9.** Synthetic pathway to synthesise **73.**

As an alternative we developed a novel synthetic pathway to form 2PCA-NH₂ **73** using a Boc-protected amine. Modifying a reaction procedure from Dubé and Scholte,[44] we reacted **69** with *tert*-butyl carbamate in the presence of TFA and triethyl silane in DCM to generate **71a** (50%) via reductive amination. Although the *tert*-butyl carbamate has weak nucleophilic properties due to the electron withdrawing effects of the Boc-group, we suspect, with the presence of TFA in the reaction mixture, that the aldehyde will be protonated making the carbonyl carbon a much stronger electrophile enabling the reaction to take place (as shown in scheme 9). As expected, the *tert*-butyl carbamate reacted at the aldehyde, evidenced by

the loss of the aldehyde peak in the $^1$H NMR spectrum of **71a**. However, we also formed a major side-product (by mass) in the reaction, **71b** (23%) (as shown in scheme 10). Initially we believed this to be the imine of compound **71a** (**71c** – as shown in figure 10); we observed a reduction in the integration value of the α-CH from 2H to 1H in the NMR spectra of **71b**, compared to the spectra of **71a**. We suspected that the reduction in the number of hydrogens present on the α-carbon was due to the sp$_2$ hybridisation of the imine. We attempted to reduce the proposed imine (**71c**) using NaBH$_4$ in methanol, however this did not form compound **71a** as we were expecting, and no reaction was observed via $^1$H NMR. Retrospectively looking at the spectra we obtained of **71b** we now think that what may have taken place was double addition of *tert*-butyl carbamate (scheme 10). Initially we thought this would be unlikely due to steric hinderance, yet the failed reduction of what we suspected to be an imine (**71c**) and the $^1$H NMR spectra supports this hypothesis. The $^1$H NMR spectra of **71b** compared to **71a** (figures 11 and 12) showed in increase in the integration of the -NH peak from 1H to 2H accompanied by a shift downfield (from δ 5.65 to 6.18). We also see a decrease in the integration of the αH peak from 2H to 1H. Furthermore, we see the integration of the Boc signal double from 9H to 18H – initially we thought that this may be residual *tert*-butyl carbamate in the sample but have since concluded that this is due to the presence of a second -NHBoc group on the α-Carbon.
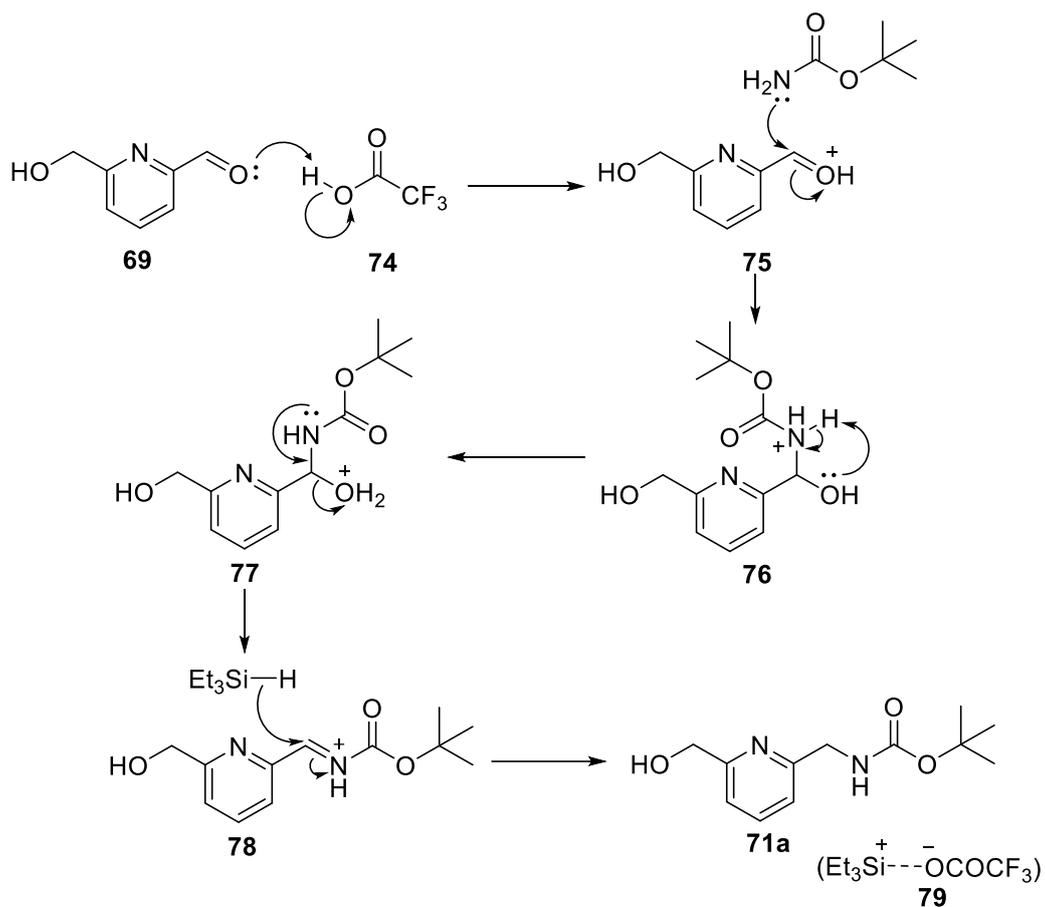
**Scheme. 9.** Reductive amination mechanism for the formation of **71a** from **69**.
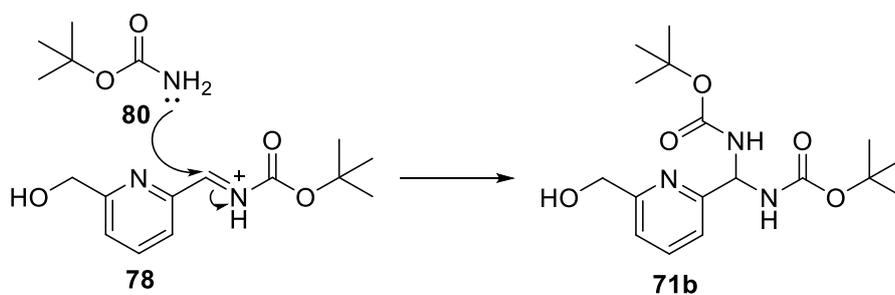


**Scheme. 10.** Hypothesised mechanism for the formation of **71b**.
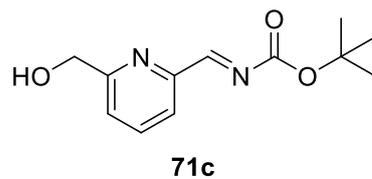


**Fig. 10.** Proposed structure of **71c**
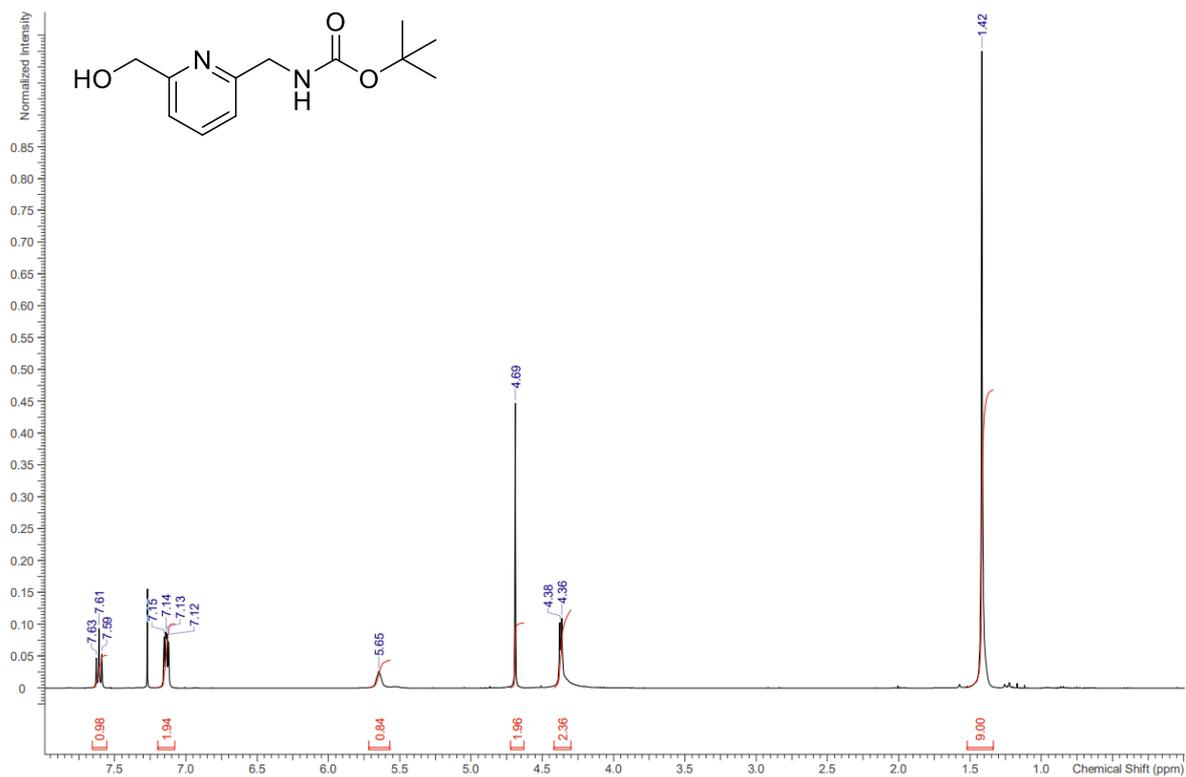
**Fig. 11.** 1H NMR spectra of compound **71a** in CDCl₃ (signal at 5.65 ppm represents the -NH proton and the signal at 4.37 represents -αCH adjacent to -NH).
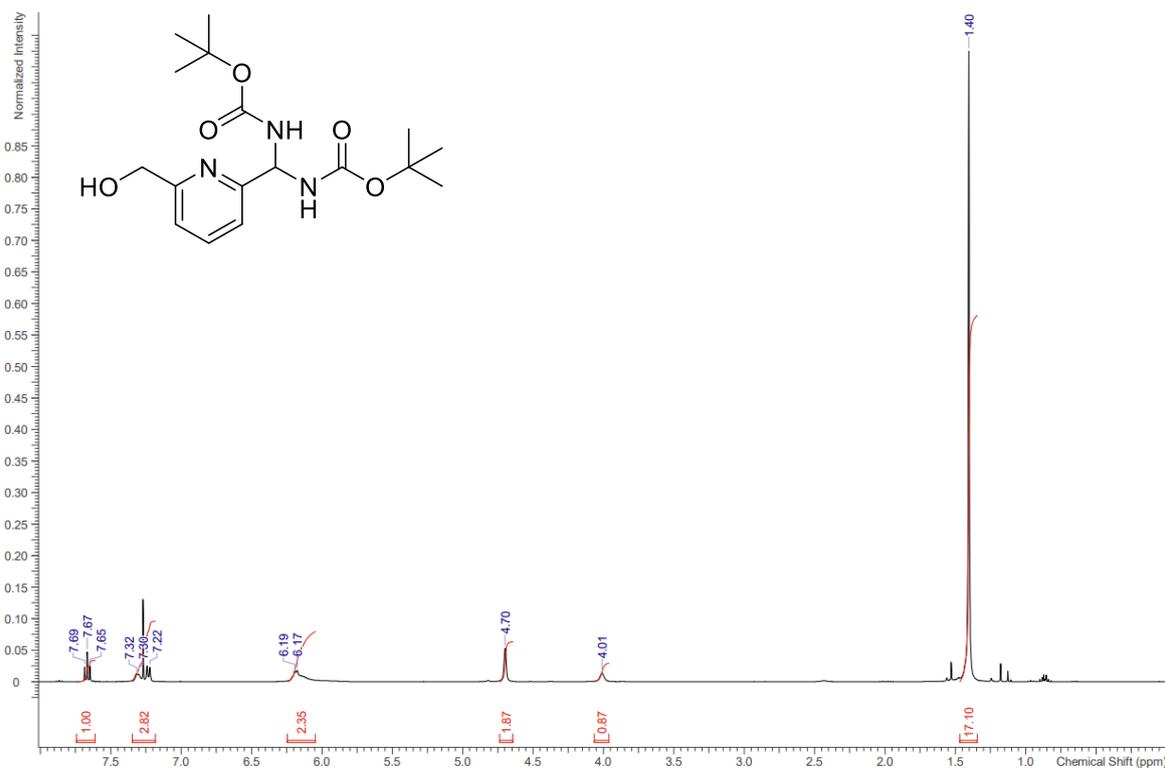


**Fig. 12.** 1H NMR spectra of compound **71b** in CDCl₃ (signal at 6.18 ppm represents the -NH protons and the signal at 4.01 represents -αCH adjacent to -NH).

Subsequently, we took compound **71a** and oxidised the hydroxymethyl group using $SeO_2$ to form the carboxaldehyde **72** in 58% yield. The final step, to remove the Boc protecting group from the amine, was first attempted using a 2:5 ratio of TFA in DCM however this was unsuccessful – though a reaction did take place, we were unable to successfully identify any products either by ^{1}H NMR spectra nor by mass spectrometry and we saw no evidence of any starting material remaining. The 1H NMR spectra was convoluted, and we were unable to accurately decipher any of the peaks. We successfully managed to deprotect the amine using a 4 M solution of HCl in Dioxane to produce the HCl salt **73** in quantitative yield. Disappointingly, we discovered that compound **73** was unstable and started to degrade at room temperature. After two weeks of the dry sample of **73** sat at room temperature we observed a colour change from a pale yellow to a dark orange/ brown, and a second ^{1}H NMR was taken. We observed a noticeable increase in both the number and integration value of aromatic peaks in the spectra of the degraded sample, in comparison to the initial spectra of **73**. Although the signals were weak, we were still able to observe an aldehyde peak - by altering the integration value of this peak to 1H, we were able to discern a considerable increase in the integration value of the peaks we suspected to be from the aromatic protons. From this we hypothesised that oligomerisation of **73** may be taking place, reacting together to form a chain of 2+ pyridine molecules connected via the $-NH_2$ and –CO groups as dimers and trimers. We attributed the increase in the aromatic proton signals visible in ^{1}H NMR to this (see scheme 11 for our hypothesised mechanism). Interestingly, no such reactions were observed within an NMR sample stored in $D_2O$ for at least 1 month. We concluded that any future samples of **73** would either need to be stored in aqueous solution or deprotected from **72** immediately before use. From scheme 11 we can see that a molecule of water is lost in the proposed mechanism, we suspect that storing **73** in an aqueous environment prevents the mechanism from occurring, in addition, imines are highly unstable in aqueous environments and any bonds formed would likely be hydrolysed.
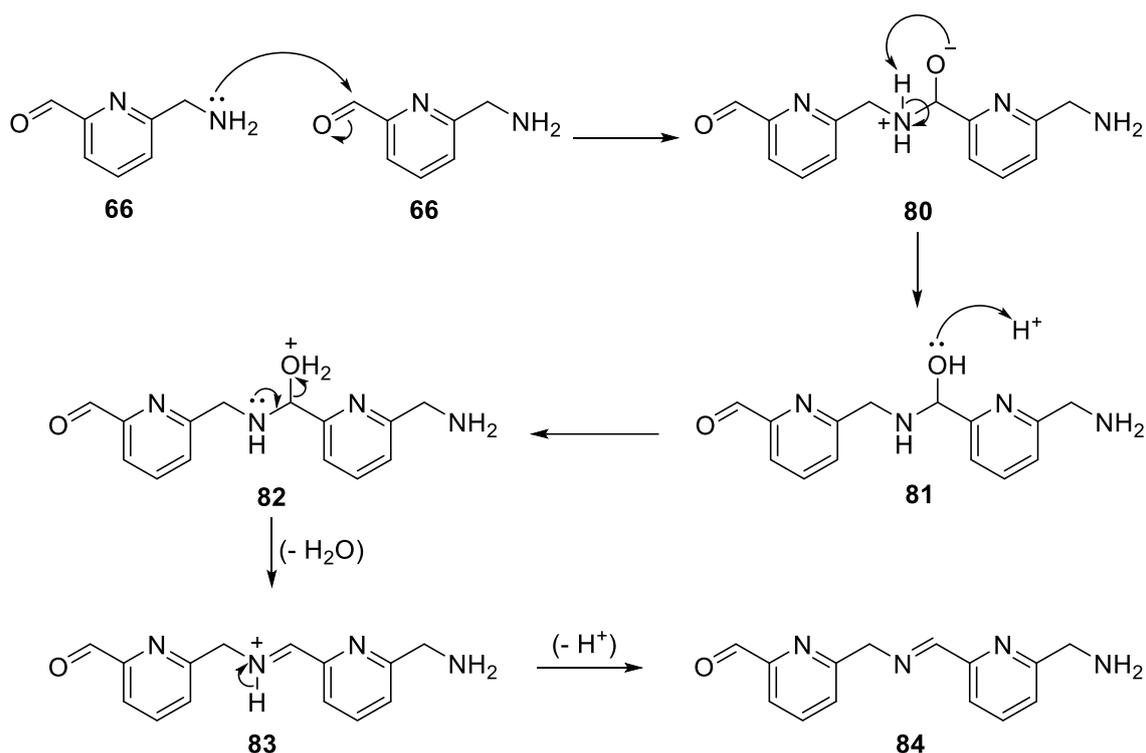
**Scheme. 11.** Proposed mechanism for oligomerisation between individual 2PCA-NH$_2$

After observing how unstable compound **73** was due to the intermolecular reactions occurring between the carbaldehyde and amine groups we decided to take a different route; using an acetal to protect the carbaldehyde group we planned to counteract the instability caused by reactions at this site. We planned to attempt the same reductive amination using *tert*-butyl carbamate, followed by a Boc-deprotection, as we conducted with **69** and **72** respectively. We then planned to conjugate these acetal-protected molecules to our scaffold via the amine motif before deprotecting the acetal to reinstate the aldehyde motif.
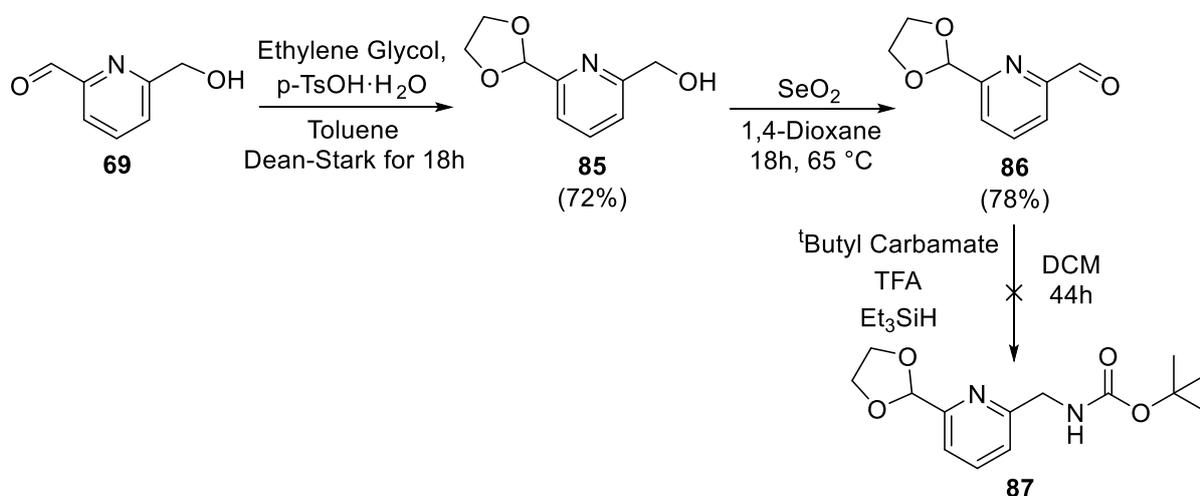


**Fig. 13.** Reaction pathway showing the attempted formation of **87** from **69**.

As shown in figure 13, we started by protecting the aldehyde on **69** using an excess of ethylene glycol and catalytic $p$TsOH·H$_2$O in toluene and heating to reflux for 19 hours under Dean-Stark conditions to form **85**. We needed to install an unprotected aldehyde motif to provide an electrophilic target for the subsequent reaction with tert-butyl carbamate. We therefore treated **85** with SeO$_2$ in 1,4-dioxane to oxidise the hydroxyl group and provide aldehyde, **86**. We then proceeded to reductively aminate **86** with *tert*-butyl carbamate, TFA and Et$_3$SiH in DCM, using the same stoichiometry used in the reaction to form **71a**. The solvents were removed under reduced pressure to give a yellow oil, however this was shown by $^1$H NMR to be **85**. We believe that the aldehyde was reduced in the reaction by triethyl silane to reform the alcohol. Due to time constraints, we ran out of time to pursue this reaction further.
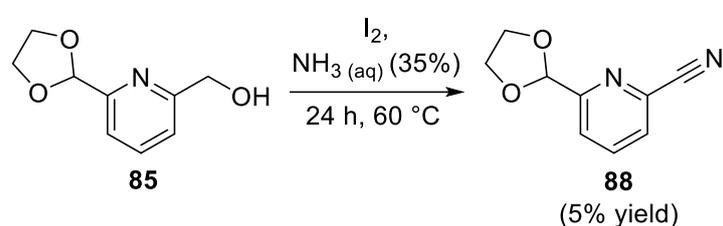


**Fig. 14.** Nucleophilic substitution of an alcohol by ammonia in the presence of iodine to form a nitrile, **88**.

Alternatively, we attempted a different route (as shown in figure 14) to the formation of the amine group by reacting **85** with an aqueous solution of ammonia (35%) in the presence of iodine (3.0 eq) to form a nitrile (**88**) in 5% yield. We would have then proceeded by converting the nitrile into an amine however due to the incredibly low yield of this reaction and time limitations we decided not to pursue it as an option.

### 2.1.3. Installation of a methyl thiol to the 6-positon of 2-pyridine carbaldehyde.

To functionalise 2PCA with a thiol motif, we began by reacting **69** with thionyl chloride in an S$_N$2 reaction to convert the alcohol into an alkyl chloride, as a better leaving group for the next step of the synthesis. Conversion to **89** was only 20%, and so larger scale reaction was required to provide enough material for subsequent steps. We suspect that the thionyl chloride was reacting with the aldehyde in competition with the alcohol group which caused the observed low yields of the desired product. The next step in the synthesis was to introduce a thiol at the site of the alkyl halide via a substitution reaction; initially we attempted to react **89** with thiourea in an ethanolic solution in the presence of 5M NaOH, under an argon

atmosphere. The crude mixture was shown to contain some of the desired product via mass spectrometry, however we were unable to isolate the product. We reattempted this procedure in the absence of NaOH; we took a preliminary $^1$H NMR of the crude mixture which indicated the presence of diethoxy acetal **91** (as shown in figure 15), which was also confirmed by mass spectrometry. Initially, we didn't want to introduce an acetal as it would add an extra step to the reaction process, so we reattempted the reaction with **89** using thiourea in 1,4-dioxane. Unfortunately, we did not observe any indication of the desired product in either NMR or mass spectrometry - the crude product was not particularly soluble in any of the commonly used solvents and so the NMR signals were barely visible, if visible at all, which meant we were unable to identify any possible products.
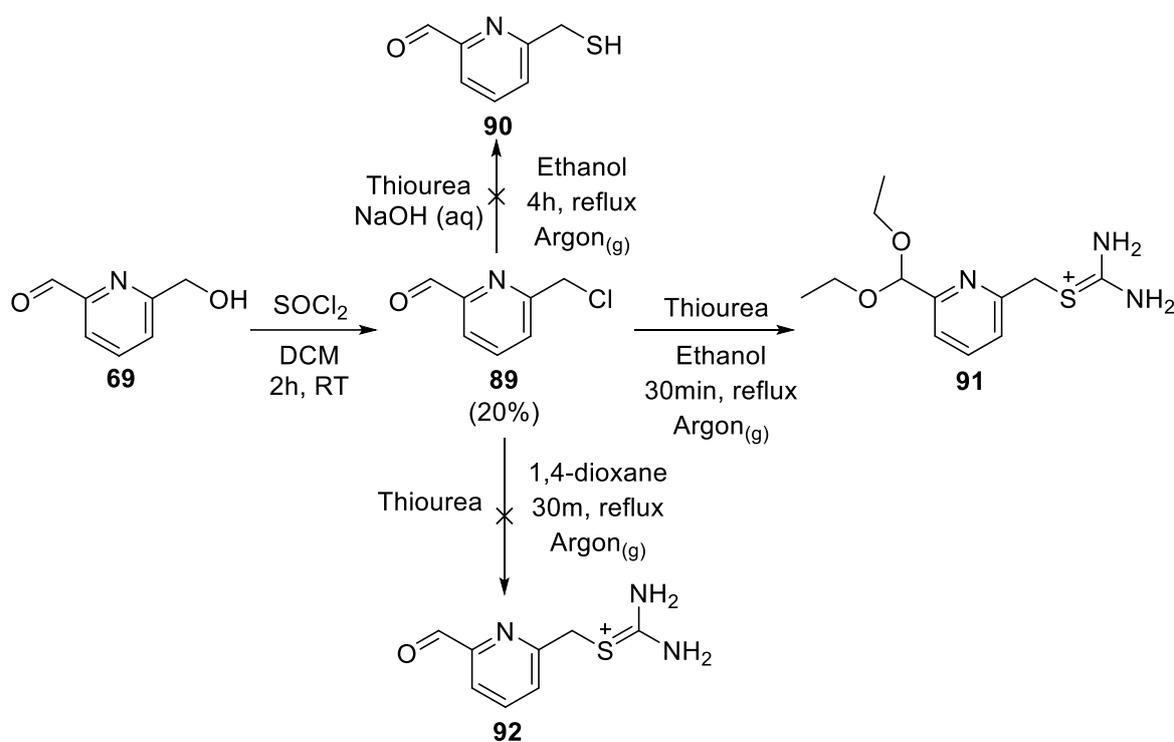


**Fig. 15.** Attempts to substitute alkyl halide with a thiol.
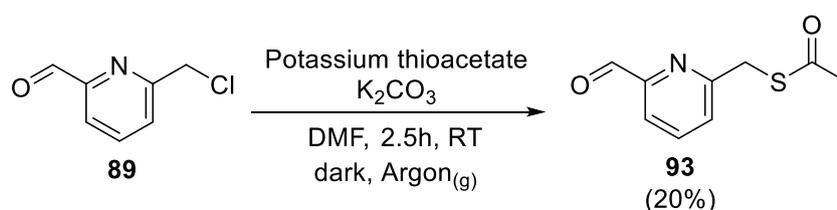


**Fig. 16.** Substitution of a chloride by thioacetate.

As an alternative approach, we introduced a thiol via an intermediate thioacetate by reacting **89** with potassium thioacetate in the presence of potassium carbonate and DMF to give **93**. This reaction, like the synthesis of **89**, was also very low yielding at only 20% yield (figure 16). As with the SOCl₂, we suspect that the thioacetate was interacting with the aldehyde, therefore we decided to protect the aldehyde with an acetal and reattempt these two reactions (figure 17). We began by substituting the hydroxyl group of **85** for an alkyl chloride using thionyl chloride in DCM which gave **94** in 77% yield. Next, we introduced the thiol in the form of a thioacetate by reacting **94** with potassium thioacetate and potassium carbonate in DMF which gave us **95** in 80% yield. We can see from both reactions that by protecting the aldehyde functionality, product yield increased by over 50% in both cases (figure 18). This backs up our suspicions that the aldehyde was interacting with both the thionyl chloride and potassium thioacetate, preventing the desired reactions from occurring effectively.
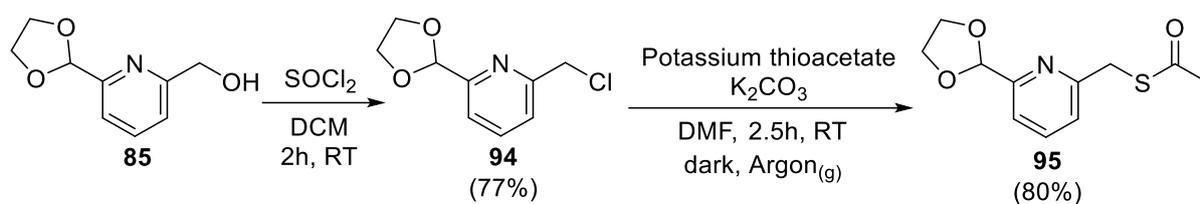


**Fig. 17.** Acetal protected synthesis of thioacetate functionalised 2PCA.

| PRODUCT YIELD (%) | | Functional Group | |
|---|---|---|---|
| | | Aldehyde | Acetal |
| **Reactants** | SOCl₂ in DCM | 20 | 77 |
| | CH₃COS⁻K⁺/K2CO₃ in DMF | 28 | 80 |

**Fig. 18.** Comparison of reaction yields between the aldehyde and acetal-protected derivatives of 2PCA.

Despite the favourability of the acetal-protected synthetic route, we were unable to convert the thioacetate group to a thiol. We attempted deprotection of **95** using potassium carbonate in an ethanolic solution for 22 hours; the mass spectrum indicated that there was some desired product in the reaction however the NMR of the residue was weak and convoluted and showed no indication of the product.

Initially we hadn't wanted to add in the extra step involved in deprotecting the acetal that was formed when we reacted **89** with thiourea in an ethanolic solution. Repeating the reaction with 1,4-dioxane instead of ethanol to avoid the generation of the acetal proved unsuccessful. As we had observed much higher yields by protecting the aldehyde with an acetal intentionally, we decided to reattempt a reaction between the alkyl chloride and thiourea, as we had previously unsuccessfully attempted with **89**. By TLC the residue was shown to contain a mixture of products however by $^1$H NMR it was shown to be a pure sample of the thiol product, **96**. Conversely, the procedure was poorly reproducible and on scale up we found that the produced residue contained many more impurities/side products. Despite this it was shown by mass spectrometry that the desired product was amongst those in the mixture. Attempts to isolate the desired product were unsuccessful, with mixtures being obtained. Due to time constraints, we weren't able to attempt to improve the scalability of the reaction.
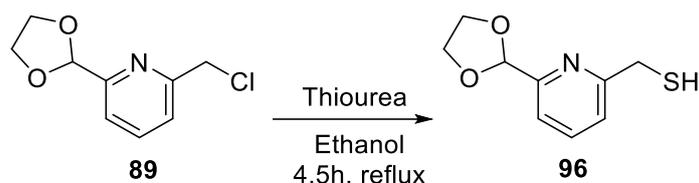


**Fig. 19.** Acetal protected synthesis of thiol functionalised 2PCA using thiourea.

## 2.2. PEG synthesis.

### 2.2.1. Summary.

The purpose of our research was to develop a derivative(s) of 2PCA capable of binding to a range of different synthetic polymer-based tissue scaffolds. We chose to use PEG hydrogels to test the viability of our 2PCA N-terminus linker molecules due to their excellent biocompatibility and highly tuneable properties (as described in more detail in section 1.2.2.). PEG is bioinert so any peptide/protein conjugation to the material would be directly relatable to the presence of the 2PCA molecule used. Our initial aim was to functionalise 8-arm PEG with our molecules of 2PCA-NH$_2$ and 2PCA-SH before hydrogelation into a tissue scaffold. As stated in section 1.2.2., using a multibranched/star macromer increases the capability for modification with a reduced effect on scaffold stability compared to linear chains. We planned to subsequently functionalise the scaffold with a fluorescently labelled N-terminus mimic

and/or green-fluorescent protein (GFP) to observe conjugation efficiency through fluorescence.

We began by using commercially available 8-arm PEG-OH ($M_w$ = 20,000 Da) as our base material. In order to couple our PCA-derivatives to the end of the PEG chains at the -OH terminus we needed to first modify the -OH group to create a functional group capable of reacting with the -NH$_2$ and -SH moieties of the PCA-derivatives **73** and **96** synthesised in the previous section (3.1.2. and 3.1.3.). There are many different functionalised PEG molecules that can be used for this purpose; to conjugate the thiol containing 2PCA we chose to functionalise PEG with a vinyl sulfone (VS) (other functional groups can be used such as PEG-maleimide or PEG-halide, for example, as shown in fig 20.).[46] We chose PEG-vinyl sulfone due to the stability of vinyl sulfones in aqueous conditions - maleimides on the other hand are prone to ring-opening hydrolysis in aqueous conditions, producing reactive functional groups that may interact unfavourably with the proteins we are attempting to conjugate to the scaffold. As hydrogels are, by design, aqueous environments, functional groups such as maleimide are  less suitable for this use. Similarly, there are several functionalised PEG molecules designed for conjugation to amines (as shown in figure 20); PEG-NHS (N-hydroxy succinimide), PEG-NPC (4-Nitrophenyl chloroformate) and PEG-acrylate, to name a few, would all be capable to binding to -NH$_2$ functionalised 2PCA.[46] However, amine conjugation to PEG-acrylate is slow, and *N*-hydroxy succinimide esters (akin to maleimides) are hydrolysed at the ester bond in aqueous conditions. In comparison, PEG-NPC reacts readily with amines and is more stable than NHS esters.[47]

Once functionalised with the desired functional group (VS or NPC), we planned that each PEG molecule would then be functionalised with 2PCA at the end of each arm via the -NH$_2$ and -SH moieties respectively. We hypothesised that by using a branched-peptide (with two N-termini) we could subsequently be able to crosslink the PCA functionalised PEG molecules to form hydrogels (as shown in figure 21). By ensuring we used an appropriate ratio of 2PCA groups to di-peptide (approximately 1:0.45 to account for the two N-termini) we would leave a proportion of the 2PCA aldehyde groups exposed and free to bind the N-termini of signalling proteins.
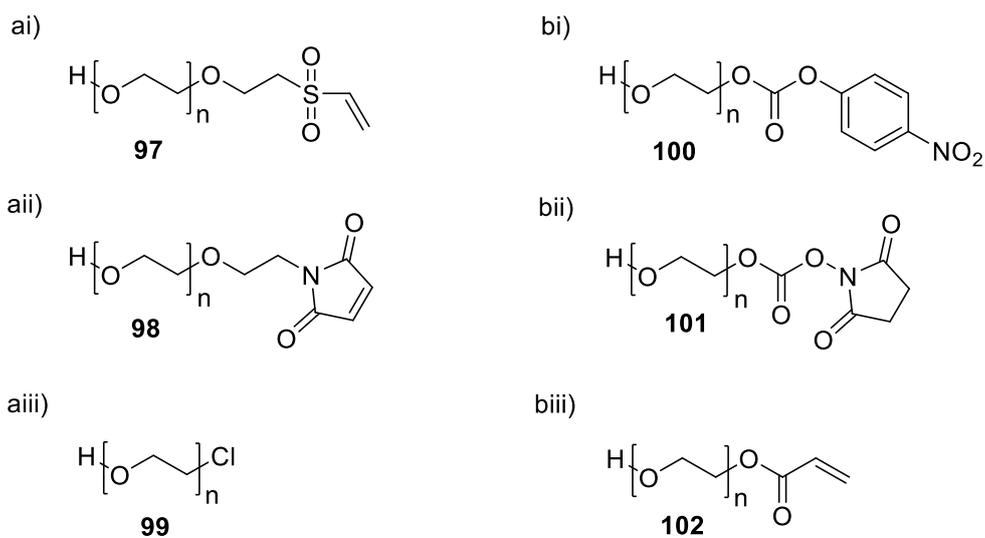
**Fig. 20.** A) Thiol reactive functionalised PEG ai) PEG-vinyl sulfone aii) PEG-maleimide aiii) PEG-halide. B) Amine reactive functionalised PEG bi) PEG-NPC bii) PEG-NHS biii) PEG-acrylate.



= Peptide crosslinker sequence AAKAA

= immobilised 2PCA motif
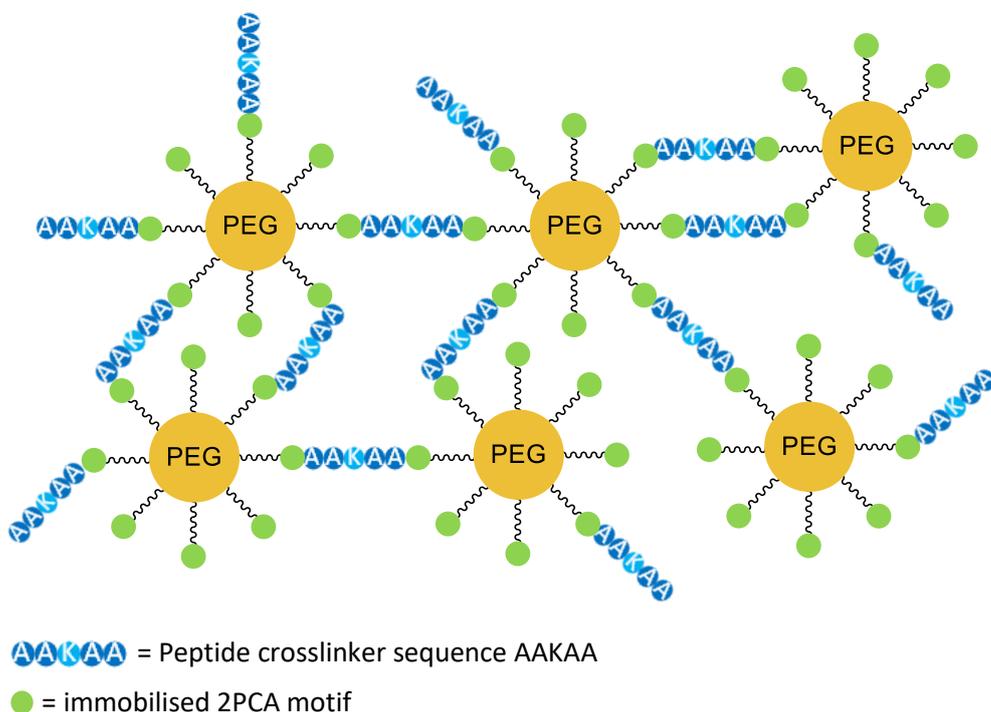
**Fig. 21.** Simplified example of crosslinking between 8-arm PEG-2PCA and branched peptide, K(AA)$_2$.

## 2.2.2. PEG synthesis.

### 2.2.2.1. PEG-NH-PCA.

We devised a PEG molecule functionalised using 4-nitrophenyl chloroformate (NPC) as a suitable pair for conjugation with PCA-NH$_2$ (**73**). Using a literature procedure,[48] 8-arm PEG-OH **103** (figure 22) was combined with 4-nitrophenyl chloroformate in the presence of triethylamine in DCM to afford PEG-NPC, **104**. As we had observed that PCA-NH$_2$ **73** was so unstable we decided to test the conjugation of PEG-NPC and an amine containing molecule prior to the attempt with PCA-NH$_2$ **73** – this was to ensure that this reaction would proceed in the way we had predicted it to before deprotecting the Boc group of **72**, as this needed to be carried out immediately before use of the resulting PCA-NH$_2$ **73**. For the amine containing molecule, we used tyramine; tyramine is an inexpensive commercially available compound of a similar weight and size to PCA-NH$_2$ **73** and contains a primary amine that we expected to react in the same way with PEG-NPC as the primary amine on **73**. From this preliminary reaction between PEG-NPC and tyramine we were able to determine suitable stoichiometry for the reactants and reaction conditions required for the reaction between PEG-NPC and **73** – we concluded that a large excess of PCA-NH$_2$ **73** would be required to modify all PEG arms, therefore we used a ratio of 1:40 of PEG-NPC: PCA-NH$_2$ **73** under an Argon atmosphere and a reaction time of between 18-22 hours.
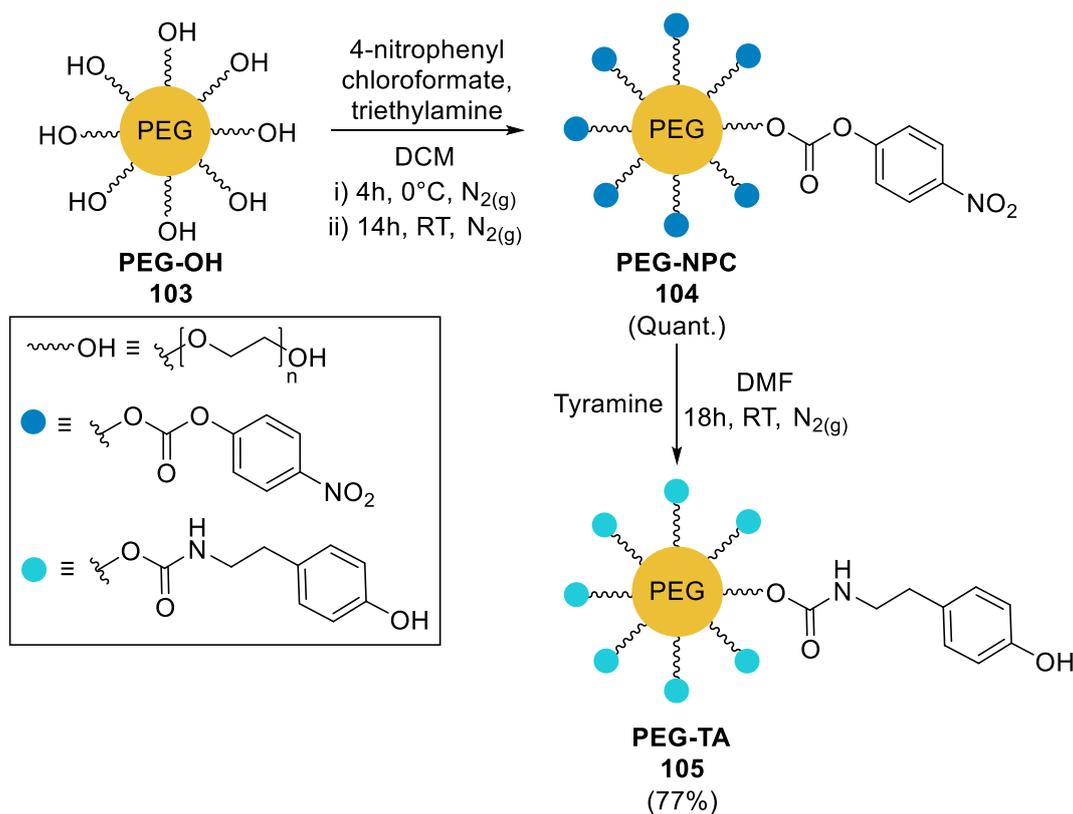
**Fig. 22.** Synthesis of PEG-tyramine (PEG-TA) **105**.

Once we had successfully conjugated tyramine to PEG-NPC we proceeded to attempt the reaction using **73** in place of tyramine. We began by removing the Boc protecting group from **72** using HCl in dioxane, as described above, to afford **73** which was reacted immediately with PEG-NPC (as shown in figure 23). As the hydrochloride salt was formed in the deprotection of the Boc group from **6** we introduced triethylamine into the reaction mixture which hadn't been required in the reaction between PEG-NPC and tyramine. We proceeded by attempting to precipitate the product in ice cold diethyl ether, as previously described for the purification of PEG derivatives,[48-50] however upon addition to the diethyl ether we observed that a brown gel had formed over the course of the reaction. We attempted to use multiple organic and aqueous solvents to solubilise the gel formed to little avail. When placed in deionised water the gel leached out a yellow colour into the solution but remained as a gel; we believe that the PCA-NH$_2$ molecules formed a crosslinked network by conjugating to both the PEG-NPC and other individual PCA-NH$_2$ molecules. When we attempted to analyse the solution which the gel leached colour into by $^1$H NMR, however no signals relating to the PCA nor the PEG were observed. We hypothesise that the PCA-NH$_2$ was able to crosslink the PEG into a gel through reacting with other PCA-NH$_2$ molecules (in a similar way to the mechanism shown in

46

figure 22, section 2.1.2.) – we suspect that by protecting the aldehyde motif (e.g., with an acetal), crosslinking would be inhibited, and the reaction would have proceeded as expected, in a similar way to tyramine. Due to time limitations, we were not able to synthesise an acetal protected form of PCA-NHBoc **87** (in figure 13) and so we were not able to investigate the hypothesised reaction between the acetal-protected form of **73** and PEG-NPC.
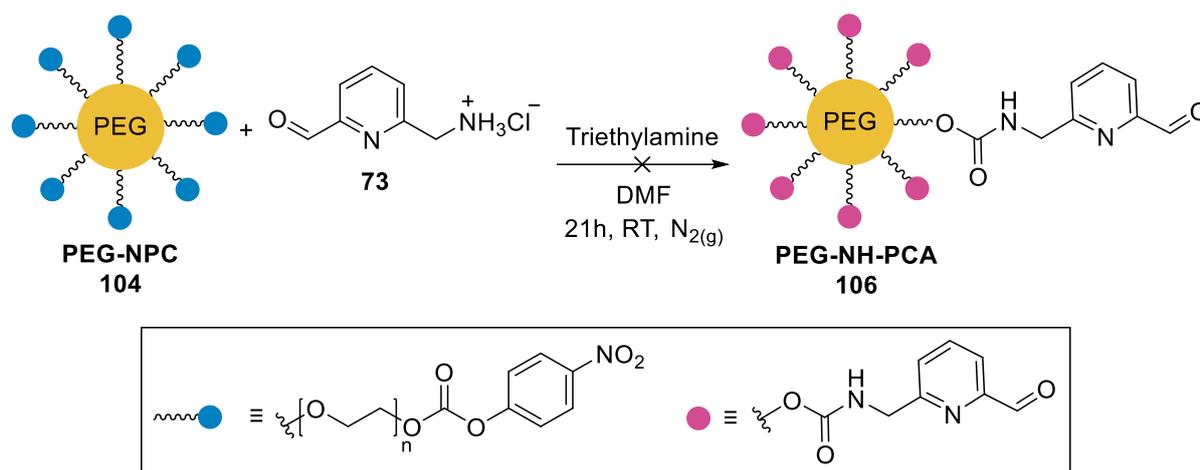


**Fig. 23.** Failed synthesis of PEG-NH-PCA **106**.

## 2.2.2.2. PEG-SH-PCA.

Alternatively, we attempted to conjugate the acetal-protected 2PCA-SH **96** to PEG; we hypothesised that by conjugating the acetal-protected 2PCA to the scaffold we would prevent any unwanted side-reactions from occurring at the aldehyde, and as we observed in the synthesis of **94** and **95**, yield had been higher when the aldehyde was protected. Once we had installed the acetal-protected 2PCA-SH **96** onto PEG, we planned to subsequently deprotect the acetal to reinstate the aldehyde in situ. We initially chose to facilitate conjugation of **96** to PEG using a vinyl sulfone(VS); vinyl sulfone is highly selective towards thiols compared to amines and does not react with any amino acid residues besides cysteine and lysine, with the former being much more favourable.[49-51]Vinyl sulfone presents pseudo-inert properties; if any of the vinyl sulfone moieties remain unmodified after reacting with **96** there is only small possibility for unwanted protein immobilisation at the site of the vinyl sulfone. This is due to the rare and often non-solvent accessible characteristics of cysteine and the steric shielding of lysine groups that make conjugation to such a bulky motif unfavourable.[51] A thiol will react readily to an activated alkene moiety, presented on vinyl sulfone and also on several other functionalised PEG molecules, such as PEG-maleimide, PEG-acrylate, and PEG-allyl. Divinyl

sulfone was added to a solution of PEG-OH and NaH (60% in oil) in DCM and reacted over a period of 48 hours under an argon atmosphere.[49,50] We proceeded to precipitate the product in ice cold diethyl ether, before decanting off the solvent and drying the product under vacuum to give a white solid. Upon analysis by $^1$H NMR it was observed that no conjugation had taken place and the divinyl sulfone was free in solution. This reaction was repeated, and the same outcome observed. This was surprising as the reaction we followed had been previously reported in the literature – we suspect that there may have been an issue with the batch of PEG used for this reaction. Several containers of PEG were in use at the time, and it is possible that we used a different batch to carry out this reaction to the one we used to synthesise PEG-NPC.



**Fig. 24.** Failed synthesis of acetal-protected PEG-SH-PCA **108**.

As an alternative to PEG-VS we proposed that PEG-allyl would serve as a suitable substitute (as shown in figure 24) as it would provide the necessary alkene group to which the -SH moiety on PCA-SH could be conjugated via a radical thiol-ene reaction.[49] We added allyl bromide to a solution of PEG-OH and NaH (60% in oil) in DCM. The product was precipitated in ice cold hexane, before decanting off the solvent and drying the product under vacuum to give PEG-allyl.

Using a literature procedure,[52] we combined PEG-allyl and **96** in the presence of radical initiator, AIBN, in dry DMF under a nitrogen atmosphere. The solution was heated to 65 °C to generate radicals of isobutyronitrile from AIBN. The residue was collected and put on for $^1$H NMR analysis which showed that **96** was free in solution and no conjugation had taken place. We suspect that the reaction environment wasn't sufficiently inert and oxygen in the environment may have quenched the radicals. Unfortunately, due to time constraints we were not able to repeat this procedure and as we were unable to successfully conjugate PEG-allyl and **96** we did not attempt to deprotect the acetal to reform the aldehyde.

## 2.3. Crosslinker Synthesis.

### 2.3.1. Summary.

Our aim was to crosslink our PEG hydrogels in the same way as we would immobilise ECM proteins - via imidazolidinone formation (as shown in figure 25). Our initial plan was to modify PEG using our 2PCA molecules and then crosslink the macromers into hydrogels through interaction between the 2PCA moieties and a di-N-termini peptide crosslinker (see figure 25). We would take advantage of the stability of the cyclic imidazolidinone to maintain scaffold integrity and we would be able to use any non-crosslinked PCA motifs for ECM protein immobilisation.
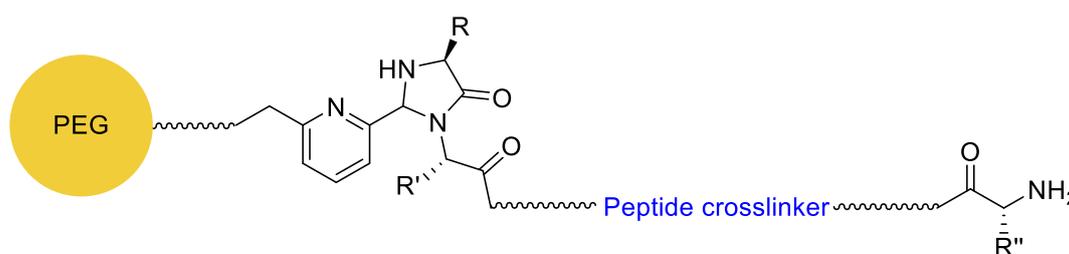


**Fig. 25.** Simplified diagram showing conjugation between a 2PCA-modified PEG and a di-N-termini peptide crosslinker via formation of a cyclic imidazolidinone ring.

### 2.3.2. Di-N-termini peptide-crosslinker synthesis.

If we had managed to succeed in functionalising PEG with either 2PCA-SH or 2PCA-NH$_2$, we planned to crosslink the PEG-PCA molecules into a hydrogel using a short chain branched peptide with a two N-termini. For this we synthesised a short-chain peptide K(AA)$_2$ **109** on a CEM Liberty Lite Automated Microwave Peptide Synthesiser, according to the manufacturer's

standard protocols. We used a Di-Fmoc Lysine residue (Fmoc-Lys(Fmoc)-OH) as the first amino acid so that addition of subsequent amino acids would lead to a branched structure. We then synthesised Ala-Ala coming off each arm - as the alanine residues bound to the lysine through their carboxyl groups we ended up with an N-terminal amino group at either end of the peptide (as shown in figure 26). We needed to attach two amino residues to either arm of the lysine residue in order to replicate the environment of the N-terminal α-amide present in a native protein. Our only limitation in choosing an amino acid to use was that the amino residue directly neighbouring the terminal residue must not be proline, as the imidazolidinone bond we aimed to generate between the PEG-2PCA and the peptide relies on the presence of an α-amide. Subsequently, we would have proceeded to crosslink the PEG-PCA molecules into a hydrogel, using a lower equivalent of N-terminus motifs: PCA-motifs to ensure the presence of exposed PCA motifs in the hydrogel to bind signalling peptides.
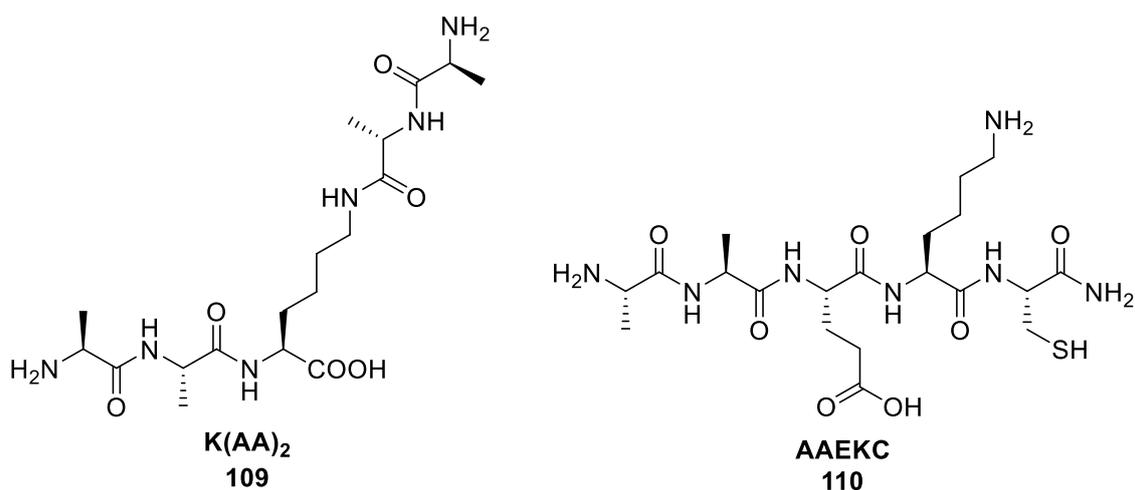


**Fig. 26.** Di-N-termini peptide K(AA)$_2$ **109** and peptide AAEKC with a C-terminal amide **110**.

## 2.3.3. Di-PCA crosslinker Synthesis.

Due to the difficulties we experienced attempting to conjugate 8-arm PEG with 2PCA-NH$_2$ and acetal-protected 2PCA-SH we decided to explore the idea of functionalising PEG with a short chain peptide (AAEKC). We hypothesised that the PEG-peptide macromers could then be crosslinked using di-2PCA molecules in a hydrogel network (this idea is expanded in section 4.2. Future work). For this we synthesised a short-chain peptide with a C-terminal amide; peptide H$_2$N-AAEKC-CONH$_2$ **110** (figure 26) was synthesised on a CEM Liberty Lite Automated Microwave Peptide Synthesiser, according to the manufacturer's standard protocols. We proposed to conjugate the peptide to PEG-VS via the thiol of the C-terminal cysteine, leaving

the N-terminus free to cross-link the PCA. However, due to the difficulties we faced in synthesising PEG-VS we weren't able to synthesise the PEG-peptide molecule, and given the time limitations we faced we were unable to explore using a different PEG molecule.

We hypothesised that by functionalising PEG with a short-chain peptide (AAEKC) we could use a di-PCA molecule (**114**) to crosslink the PEG into a hydrogel. Using a 2PCA motif ratio slightly below 1:1 we would be able to crosslink the hydrogel whilst leaving some of the 2PCA motifs unreacted and able to conjugate to signalling peptides.

We started by reacting triethylene glycol **111** with *p*-toluenesulfonyl chloride and triethylamine in DCM to form **112**. We chose triethylene glycol to provide a flexible and bioinert linker between the two PCA motifs. Due to the increased yields we observed using the acetal-protected 2PCA molecules (as demonstrated in section 3.1.3.), we decided to functionalise **112** with acetal-protected 2PCA **85** before deprotecting the acetal to form the aldehyde. Once we had activated the hydroxyl groups of triethylene glycol with tosyl chloride we proceeded to react **112** with acetal-protected 2PCA **85** in the presence of NaH (60% in oil) in anhydrous THF to form **113** in 22% yield. We suspect that yield may have been low due to complications with purification via flash column chromatography. It may also be possible that a larger excess of **85** was required to ensure efficient substitution took place. Subsequently, we would have proceeded by deprotecting the acetal groups to re-instate the aldehydes and generate the di-PCA crosslinker **114** (as shown in figure 27), however due to time constraints we were unable to take the synthesis further.
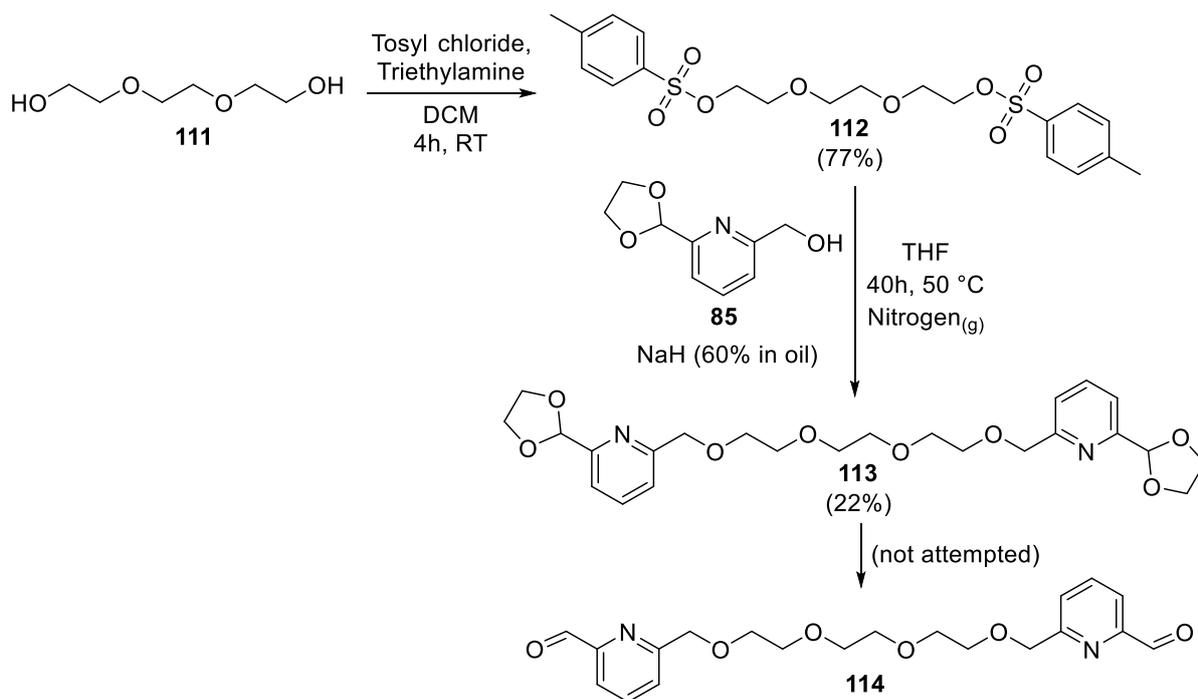
**Fig. 27.** Synthesis of di-2PCA crosslinker **114**.

## 2.4. Fluorophore-N-terminus-mimic Synthesis.

### 2.4.1. Summary.

To test the efficiency of binding between the aldehyde of the 2PCA functionalised hydrogels and the N-terminus of a protein we planned to first test the binding using fluorophore functionalised with an N-terminus mimicking alpha-amino amide. We introduced N-terminus functionality to the fluorophore through integration of alanine via an amide bond to mimic the α-amide of the native N-terminus. We then hoped to observe the conjugation between the PCA presented on the PEG-hydrogels and the N-terminus of our fluorophore. Upon successful conjugation we would observe fluorescence under excitation of the hydrogels and from here we would be able to quantify formation of the imidazolidinone bond and measure conjugate stability. We would also be able to measure physical properties of the hydrogel, such as strength and elasticity, by observing fluorescence of the hydrogel when tested using rheology.

Coumarins tend lack toxicity in mammalian systems and have been used in many drugs, the biocompatibility of coumarin and its fluorescent properties make it an ideal candidate to develop an N-terminus-mimic fluorophore.[53] The synthesis of this novel fluorophore involved

combining two subunits, the coumarin fluorophore and an N-terminal alanine, via a linker molecule, ethylenediamine – the alanine needed to be bound to the fluorophore via an amide bond that mimicked the α-amide bond present between amino residues in a native protein, the ethylenediamine introduces the amide bond at the alanine end of the compound whilst securing alanine to the coumarin fluorophore.

## 2.4.2. Synthesis.

The synthesis of the coumarin-derivative N-terminus-mimic (as shown in figure 29) began by synthesising 7-hydroxycoumarin-3-carboxylic acid **122** which we would then combine with an N-terminus mimic. Coumarin isn't naturally fluorescent - electron withdrawing groups at the 3-position, in addition to electron donating groups at the 7-position, create a transfer of electrons within the molecule, which imparts highly fluorescent properties.[54] As we will be utilising the fluorescent capabilities of coumarin, we are able to install electron withdrawing groups, such as the amide bond in **124**, which help to increase the fluorescence.

Commercially available 2,4-dihydroxybenzaldehyde **119** was reacted with diethyl malonate **120** in the presence of piperidine in EtOH to afford **121**. This was then hydrolysed using NaOH (2 M) to afford the carboxylic acid **122** which forms the coumarin portion of the N-terminal fluorophore. In order to combine the coumarin **122** and the N-terminus moiety **118**, both of which present an exposed carboxyl group, we utilised a diamine linker - forming an amide linkage to both the coumarin and N-terminus group respectively. For this we used ethylenediamine **115**; since we only wanted the coumarin to attach to one end of the diamine we protected one end using a Boc-protecting group. Ethylenediamine was combined with Boc anhydride in DCM to afford **116** (figure 28). We proceeded by activating the carboxylic acid of **122** using N-hydroxy succinimide (NHS) in the presence of EDC·HCl in DMF to afford the NHS ester **123**. The NHS ester moiety provided a superior leaving group to facilitate the next step in which an amide linkage was formed between **116** and **123** in the presence of triethylamine in MeOH to afford **124** in 38% yield. The low yield may be due to unwanted side-reactions between the secondary amine and the NHS-activated carbonyl group. Once we had successfully combined the coumarin portion of the fluorophore with the diamine linker we proceeded to deprotect the Boc group using TFA in MeOH to afford TFA salt **125**.
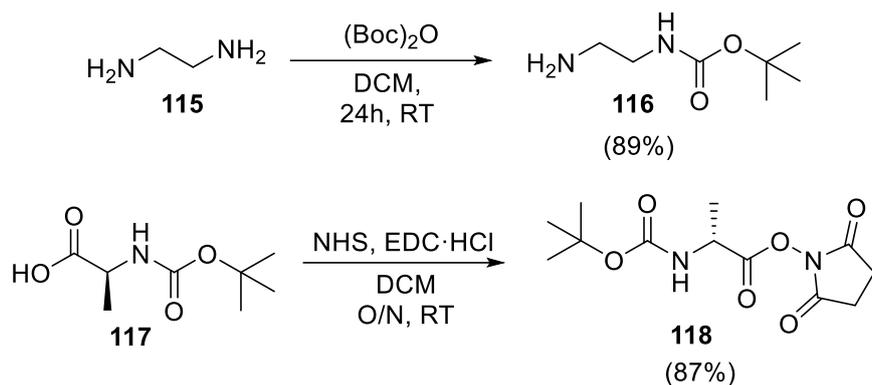
**Fig. 28.** Synthesis of **116** and **118**.
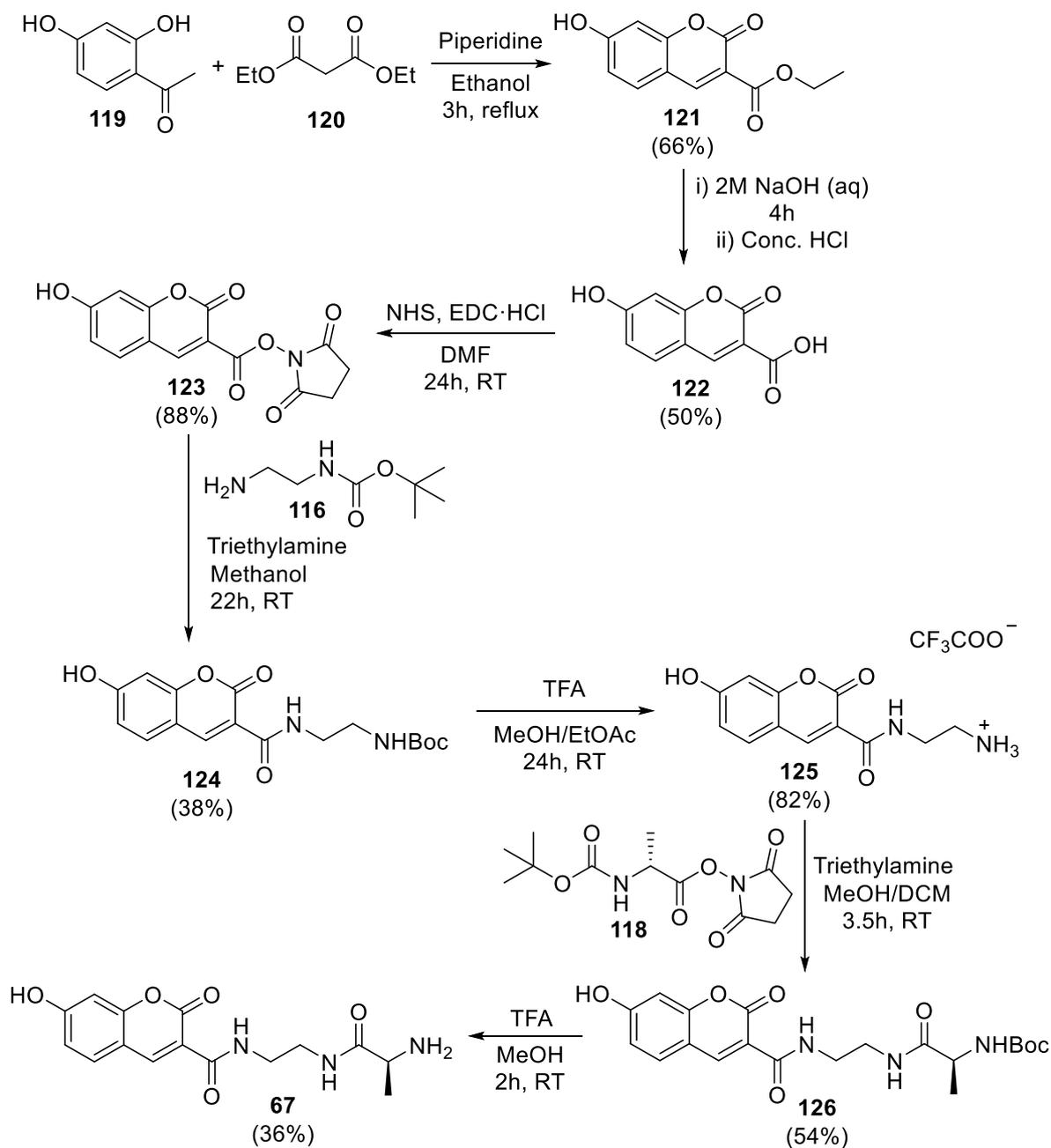


**Fig. 29.** Synthesis of coumarin fluorophore N-terminus mimic **67**.

To form the N-terminus moiety of our fluorophore we used the commercially available amino acid Boc-Ala-OH **117** which we activated as the NHS ester in the presence of EDC·HCl in DCM to afford **118**. The subsequent reaction between **118** and **125** in which the coumarin and N-terminus units of the fluorophore were combined via the diamine linker moiety, afforded **126** in 54% yield. We then proceeded to deprotect the Boc group of **126** using TFA in DCM to afford the TFA salt, **67** in 36% yield. We observed that removal of the Boc group required additional TFA to be added, therefore we suspect that the low yield may be due to the reaction time being too short.

# 3. Conclusion and future work.

## 3.1. Conclusion.

In this project we have devised synthetic routes to form novel 2PCA-NH$_2$, acetal-protected PCA-SH and a fluorescent N-terminus mimicking coumarin fluorophore.

We have shown that the aldehyde on 2PCA derivatives is incredibly reactive and often interferes with reactions taking place at other sites on the molecule, however this can be overcome by protecting the aldehyde group with a cyclic acetal, increasing reaction yield by over 50% in two applicable reactions.

We have also developed a novel precursor to a di-2PCA ethylene glycol-based crosslinker, which should be able to undergo facile acetal deprotection to form the reactive di-aldehyde crosslinker.

## 3.2. Future work.

As stated earlier, the Covid-19 outbreak led to limited time in the lab and time wasted resynthesizing compounds made before the first lockdown. This presented issues towards the end of the project as we ran out of time to complete the synthesis of some of our target molecules.

In the short term, it would be desirable to improve the scale-up reaction between **89** and thiourea to form **96**. Additionally, once this is achieved, we would be able to proceed with deprotecting the acetal on **96** to form the last of our target molecules 2PCA-SH **65** (figure 30).
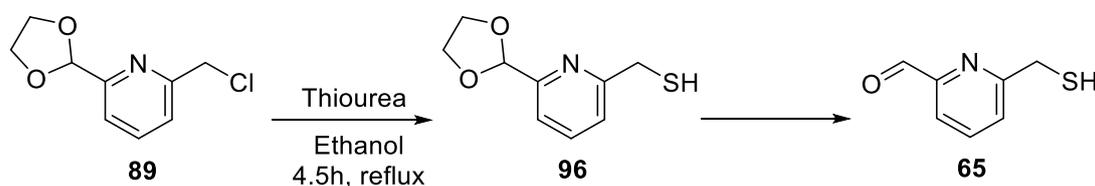


**Fig. 30.** Short-term 2PCA-SH reactions.

Additionally, it would be interesting to explore the reasons behind why **86** failed to react with $^t$Butyl carbamate in a reductive amination. There is a possibility that the reaction wasn't left for long enough, as **69** reacted over a period of 96 hours to afford **71a**, whereas, due to time limitations, **86** was only reacted over a period of 44 hours (figure 31).
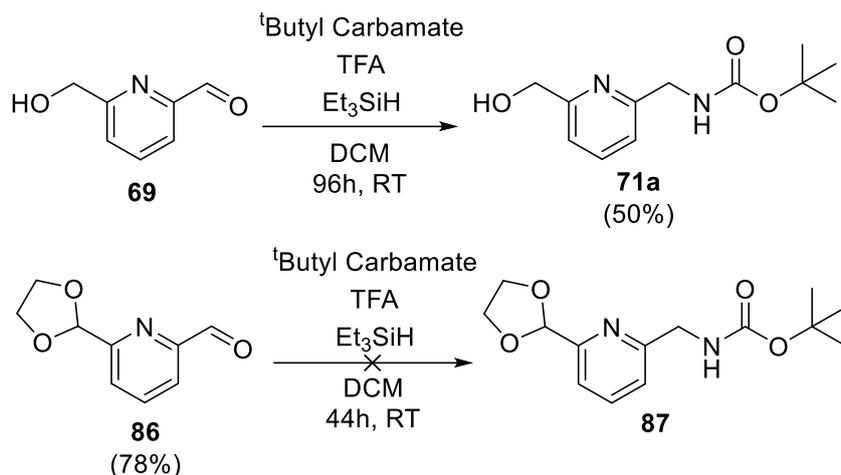
**Fig. 31.** Comparison between the success of the reductive amination reactions of **69** and **86**.

Another short-term objective includes the preparation of PEG-vinyl sulfone, the method we used is widely reported in the literature and we were, therefore, surprised when it didn't work. Using, a fresh batch of PEG-OH we would be able to retry the reaction and hopefully achieve successful modification. This would then allow us to functionalise PEG-VS with our preformed AAEKC peptide. It would then involve a simple procedure to deprotect the acetal groups on our di-2PCA crosslinker molecules, allowing us to combine our PEG-peptide molecules and 2PCA crosslinkers in a hydrogel scaffold.

Longer term, we would like to have carried out testing of the 2PCA crosslinked PEG-hydrogels and how we might have altered the components to impart different properties to the gel. It would be interesting to examine how varying the number of crosslinks affects both hydrogel stability and protein conjugation, which we would test through rheological experiments and fluorescence using our fluorescent coumarin N-terminus mimic.

Additionally, it would be interesting to explore how multi-chain proteins/peptides, with more than one N-terminus, interact with the scaffold – would they bind twice via two or more N-termini or would it be possible to "cap" one N-terminus and use the other for single-site protein immobilisation.

# 4. Experimental.

## 4.1. General considerations.

All solvents were used as supplied (Analytical or HPLC grade), without prior purification. Reagents were purchased from Sigma-Aldrich and used as supplied, unless otherwise indicated. All aqueous solutions were prepared using deionised water. Brine refers to a saturated solution of sodium chloride. Petrol refers to the fraction of petroleum ether boiling in the range 40 – 60 °C.

Proton nuclear magnetic resonance ($^1$H NMR) spectra were recorded on a Jeol ECX-400 (400 MHz) or Bruker AVIIIHD (500 MHz) spectrometer. Carbon nuclear magnetic resonance ($^{13}$C NMR) spectra were recorded on a Jeol ECX-400 (100 MHz) spectrometer. NMR shifts were assigned using COSY, HSQC and HMBC spectra. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard ($^1$H NMR: $CDCl_3$ = 7.26; MeOD = 3.31; $D_2O$ = 4.69; DMSO-$d_6$ = 2.50 and $^{13}$C NMR: $CDCl_3$ = 77.16, MeOD = 49.00, DMSO-$d_6$ = 39.52). Coupling constants ($J$) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, sept = septet, m = multiplet, app = apparent, br = broad. Melting points (m.p.) were recorded on a Gallenkamp melting point apparatus. Infrared (IR) spectra were recorded on a Perkin Elmer UATR Two FT-IR spectrometer. Absorption maxima ($\upsilon_{max}$) are reported in wavenumbers ($cm^{-1}$). UV-Vis spectra were recorded on a Perkin Elmer Lambda 25 spectrophotometer in a quartz cuvette or glass with a pathlength of 1 cm. High resolution electrospray ionisation (ESI) mass spectra (HRMS) were recorded on a Bruker Campact TOF-MS or a Jeol AccuTOF GCx-plus spectrometer. m/z values are reported in Daltons.

Thin layer chromatography (TLC) was carried out using aluminium backed sheets coated with 60 $F_{254}$ silica gel (Merck). Visualization of the silica plates was achieved using a UV lamp ($\lambda_{max}$ = 254, 302, or 366 nm), and/or stained with the appropriate staining solution; potassium permanganate (5% $KMnO_4$ in 1M NaOH with 5 % potassium carbonate), or ninhydrin (1.5% ninhydrin, 3% AcOH in $n$-butanol).  Flash column chromatography was carried out using Geduran Si 60 (40-63 μm) (Merck). Mobile phases are reported as % volume of more polar solvent in less polar solvent.

Liquid chromatography-mass spectrometry (LC-MS) was performed on a HCTultra ETD II ion trap spectrometer, coupled to an Ultimate300 HPLC using an Accucore C18 column (150 × 2.1 mm, 2.6 μm particle size). Water (solvent A) and acetonitrile (solvent B), both containing 0.1% formic acid, were used as the mobile phase at a flow rate 0.3 mL min$^{-1}$. LC traces were measured via UV absorption at 220, 270, 280, and 220-400 nm. The solvent gradient was programmed as shown in figure 41 and spectra were analysed using the Bruker Data Analysis 4.1 software.
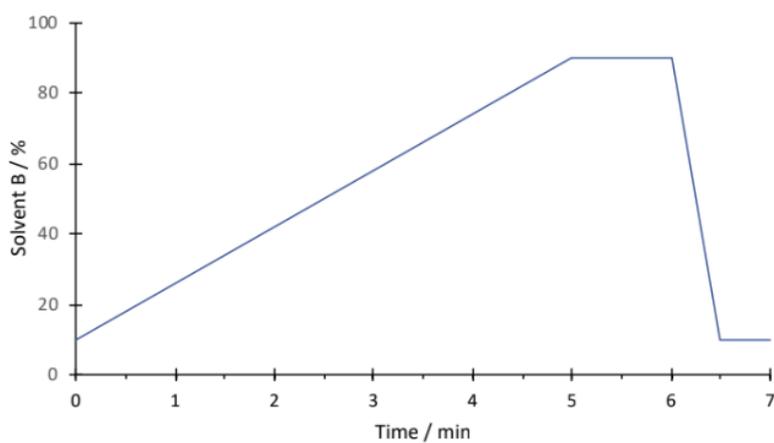


**Fig. 32.** Percentage of solvent B used during LC-MS over the duration of analysis.

## 4.2. Experimental Procedures and Characterisation.

### 4.2.1. PCA synthesis.

**6-(Hydroxymethyl)-2-pyridinecarbaldehyde (69)**
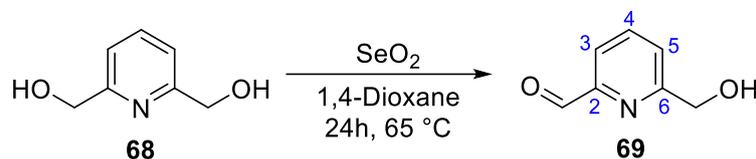


A mixture of 2,6-pyridinedimethanol **68** (1.00 g, 7.20 mmol, 1.0 eq.) and selenium dioxide (400 mg, 3.6 mmol, 0.50 eq.) in dioxane (20ml) was stirred at 65 °C for 24 hours. After cooling to RT, the crude mixture was filtered under vacuum, and the filtrate concentrated under vacuum. The residue was purified by flash column chromatography eluting with 2.5% MeOH: DCM. The pure fractions were combined, and the solvent removed under reduced pressure to give **69** (0.94 g, 6.9 mmol, 63%) as an off-white solid.

**¹H NMR** (400 MHz, CDCl$_3$), δ 10.09 (s, 1 H, C<u>H</u>O), 7.92-7.88 (m, 2 H, ArH3 and ArH5), 7.55-7.51 (m, 1 H, Ar<u>H</u>4), 4.89 (s, 2 H, C<u>H</u>$_2$). The spectroscopic data were consistent with that reported in the literature.[55]

**(6-Formyl-2-pyridinyl)methyl methanesulfonate (70)**
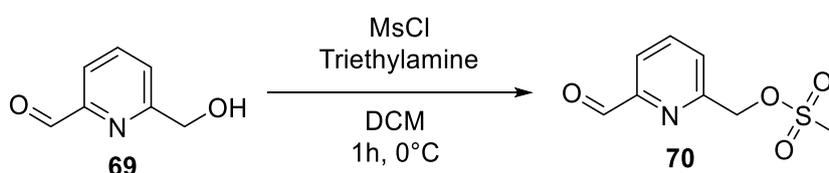


Triethylamine (0.45 mL, 3.1 mmol, 3.0 eq.) was added to a solution of **69** (0.15 g, 1.1 mmol, 1.0 eq.) in DCM (5 mL) and stirred at 0 °C for 5 minutes. Methanesulfonyl chloride (0.10 mL, 1.3 mmol, 1.2 eq.) was then added dropwise to the reaction mixture and left to stir for 1 hour at 0 °C. The reaction was quenched with saturated NaHCO$_3$ (aq) (10 mL), and the aqueous layer was extracted using DCM (3 × 10 mL). The combined organics were dried over MgSO$_4$ and the solvent removed under vacuum to give **70** (50 mg, 0.23 mmol, 21%) as a brown oil.

**¹H NMR** (400 MHz, CDCl$_3$), δ 10.04 (s, 1 H, C<u>H</u>O), 7.99-7.91 (m, 2 H, ArH3 and ArH5), 7.74-7.71 (m, 1 H, Ar<u>H</u>4), 5.42 (s, 2 H, C<u>H</u>$_2$), 3.15 (s, 3 H, C<u>H</u>$_3$). The spectroscopic data were consistent with that reported in the literature.[55]

## Tert-butyl N-[(6-hydroxymethyl-2-pyridyl)methyl]carbamate (71a)
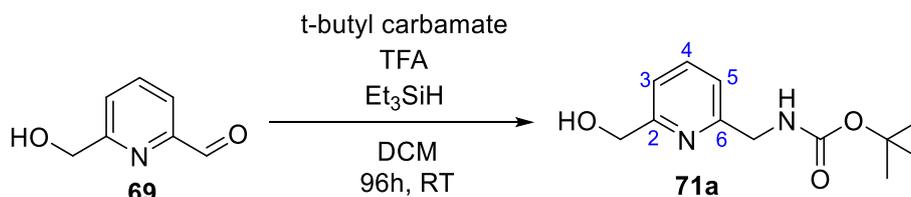


Triethyl silane (24 mL, 150 mmol, 10.0 eq.) was added dropwise to a solution **69** (2.0 g, 15 mmol, 1.0 eq.), TFA (3.4 mL, 45 mmol, 3.0 eq.) and *tert*-butyl carbamate (3.5 g, 30 mmol, 2.0 eq.) in DCM (100 mL) and stirred for 96 hours. The solvent was evaporated under reduced pressure and the residue dissolved in ethyl acetate (100 ml). The organics were washed with saturated $NaHCO_3$ (aq) (3 × 100 ml) and brine (1 × 100 ml), dried over $MgSO_4$ and the solvent removed under reduced pressure. The residue was purified by flash column chromatography eluting with 50-100% EtOAc: Petroleum ether (40-60 °C). The pure fractions were combined, and the solvent removed under reduced pressure to give **71a** (1.8 g, 7.4 mmol, 50%) as a pale yellow oil.

**$^1$H NMR** (400 MHz, CDCl$_3$), δ 7.63-7.59 (m, 1 H, ArH4), 7.15-7.12 (m, 2 H, ArH3 and ArH5), 5.65 (br s, 1 H, NH), 4.69 (s, 2 H, CH$_2$OH), 4.37 (d, *J* = 5.3 Hz, 2 H, CH$_2$NHBoc), 1.42 (s, 9H, 3 x CH$_3$); **$^{13}$C NMR** (101 MHz, CDCl$_3$), δ 158.85 (ArC2), 156.72 (ArC6), 155.97 (CO), 137.35 (ArC4), 120.00 (ArC3), 118.95 (ArC5), 79.48 (C(CH$_3$)$_3$), 64.05 (CH$_2$OH), 45.50 (CH$_2$NHBoc), 28.28 (CH$_3$); ν$_{max}$: (FT-ATR)/cm$^{-1}$: 3338, 2978, 2932, 1689, 1597, 1516, 1455, 1366, 1249, 1163, 1048, 946, 862, 759, 610; **HRMS**: *m/z* (ESI$^+$) calc. for $C_{12}H_{18}N_2O_3$ [M+H]$^+$ requires 261.1210, found 261.1213.

## Tert-butyl N-[(6-formyl-2-pyridyl)methyl]carbamate (72)
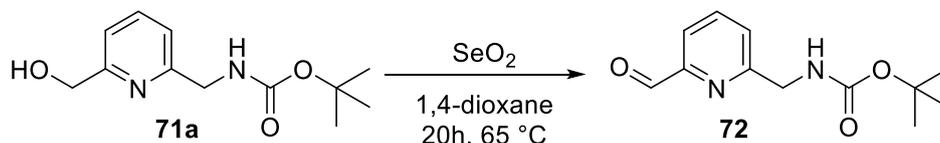


Selenium Dioxide (0.49 g, 4.40 mmol, 1.0 eq.) was added to a solution of **71a** (1.05 g, 4.40 mmol, 1.0 eq.) in 1,4-dioxane (100 mL) and stirred at 65 °C for 20 hours. After cooling to RT, the reaction mixture was filtered, and the filtrate concentrated under vacuum to give **72** (0.60 g, 2.54 mmol, 58%) as a dark yellow oil.

**$^1$H NMR** (400 MHz, CDCl$_3$), δ 10.06 (s, 1H, CHO), 7.88-7.83 (m, 2 H, ArH3 and ArH5), 7.54-7.50 (m, 1 H, ArH4), 5.61 (br s, 1 H, NH), 4.53 (d, *J* = 5.50 Hz, 2 H, CH$_2$NHBoc), 1.47 (s, 9H, 3 x CH$_3$);

**¹³C NMR** (101 MHz, CDCl₃), δ 193.30 (CHO), 158.51 (ArC6), 155.97 (COO), 152.10 (ArC2), 137.68 (ArC4), 125.93 (ArC5), 120.31 (ArC3), 79.81 (C(CH₃)₃), 45.49 (CH₂NHBoc), 28.36 (CH₃); **νmax**: (FT-ATR)/cm⁻¹: 3356, 2977, 2933, 1705, 1513, 1367, 1249, 1167, 1050, 951, 863, 783, 645 ; **HRMS**: *m/z* (ESI⁺) calc. for C₁₂H₁₆N₂O₃ [M+Na]⁺ 259.1053, found 259.1054.

### 6-(Aminomethyl)pyridine-2-carbaldehyde hydrochloride (73)
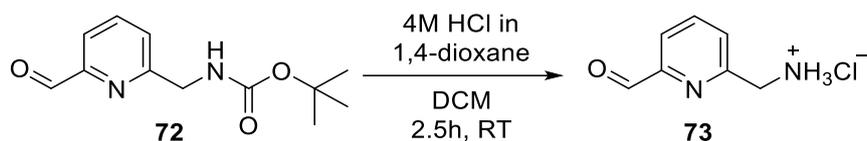


HCl in Dioxane (4.0 M, 3.2 mL, 13 mmol, 11.0 eq.) was added to a solution of **72** (0.28 g, 1.2 mmol, 1.0 eq.) in DCM (30 mL). After 2 hours an extra 2.0 eq of 4M HCl in Dioxane (0.58 mL) was added and the reaction left to stir for an additional 0.5 hours. The solvent was then removed under vacuum to give **73** (230 mg, 1.1 mmol, 95%) as a yellow solid.

**¹H NMR** (400 MHz, D₂O), δ 9.92 (s, 1H, ArC2CH), 7.97-7.95 (m, 1 H, ArC4H), 7.66-7.64 (m, 1 H, ArC3H/ArC5H), 7.50-7.49 (m, 1 H, ArC3H/ArC5H), 4.39 (s, 2H, CH₂NHBoc); **HRMS**: *m/z* (ESI⁺) calc. for C₇H₉N₂O [M+H]⁺ 137.0709, found 137.0709 Due to the instability of the compound, it was used without further analysis.

### [6-(1,3-Dioxolan-2-yl)-2-pyridinyl]methanol (85)
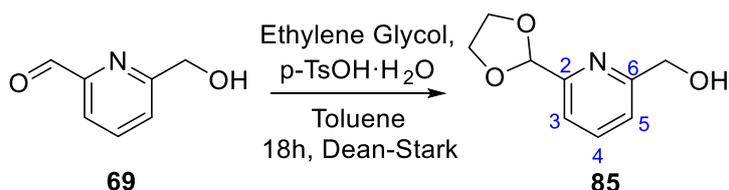


**69** (3.6 g, 26 mmol, 1.0 eq) was added to a mixture of ethylene glycol (10 mL, 180 mmol, 7.0 eq) and *p*-toluenesulfonic acid monohydrate (490 mg, 2.6 mmol, 0.1 eq) in toluene (150 mL) and refluxed under Dean-Stark conditions for 18 hours. After cooling to RT, the solvent was removed under reduced pressure. The residue was purified by flash column chromatography, eluting with 2% MeOH:DCM. The product containing fractions were combined and the solvent removed under reduced pressure. The residue was then purified by a second flash column eluting with 50% EtOAc: Petrol. The pure fractions were combined, and the solvent removed under reduced pressure to give **85** (3.4 g, 19 mmol, 72%) as a yellow oil.

**¹H NMR** (400 MHz, CDCl₃), δ 7.74-7.70 (m, 1H, Ar$\underline{H}$4), 7.44 (d, *J* = 7.3 Hz, 1H, Ar$\underline{H}$3), 7.26 (d, *J* = 7.8 Hz, 1H, Ar$\underline{H}$5), 5.84 (s, 1H, Ar2C$\underline{H}$), 4.76 (s, 1H, Ar6C$\underline{H}_2$), 4.19-4.03 (m, 4H, C$\underline{H}_2$C$\underline{H}_2$), 3.73 (br s, 1H, O$\underline{H}$); **¹³C NMR** (101 MHz, CDCl₃), δ 159.41 (Ar$\underline{C}$6), 156.19 (Ar$\underline{C}$2), 137.77 (Ar$\underline{C}$4), 121.09 (Ar$\underline{C}$5), 119.31 (Ar$\underline{C}$3), 103.68 (Ar2$\underline{C}$H), 65.80 ($\underline{C}$H₂$\underline{C}$H₂), 64.42 (Ar6$\underline{C}$H₂); **ν**$_{max}$: (FT-ATR)/cm⁻¹: 3350, 2890, 1598, 1455, 1368, 1212, 1104, 1027, 991, 945, 781, 649; **HRMS**: *m/z* (ESI⁺) calc. for C₉H₁₁NO₃ [M+Na]⁺ 204.0631, found 204.0632.

### 6-(1,3-dioxolan-2-yl)picolinaldehyde (86)
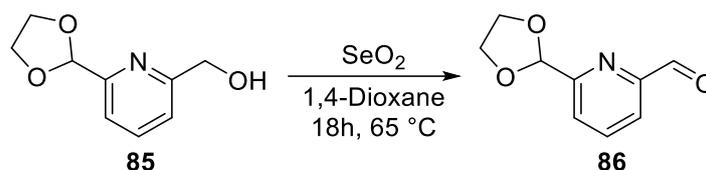


SeO₂ (120 mg, 1.1 mmol, 0.5 eq) was added to **85** (400 mg, 2.2 mmol, 1.0 eq) in 1,4-dioxane (20 mL) and heated to 65 °C for 18 hours. After cooling to RT, the crude mixture was filtered under vacuum, and the filtrate concentrated under vacuum to give **86** (310 mg, 1.7 mmol, 78%) as a yellow oil.

**¹H NMR** (400 MHz, CDCl₃), δ 10.11 (s, 1H, Ar6C$\underline{H}$O), 7.98-7.91 (m, 2H, Ar$\underline{H}$3 and Ar$\underline{H}$5), 7.80-7.78 (m, 1H, Ar$\underline{H}$4), 5.94 (s, 1H, Ar2C$\underline{H}$), 4.25-4.10 (m, 4H, C$\underline{H}_2$C$\underline{H}_2$); **HRMS**: *m/z* (ESI⁺) calc. for C₉H₉NO₃ [M+Na]⁺ 202.0475, found 202.0477. The spectroscopic data were consistent with that reported in the literature.[56]

### 2-(1,3-dioxolan-2-yl)-6-cyanopyridine (88)
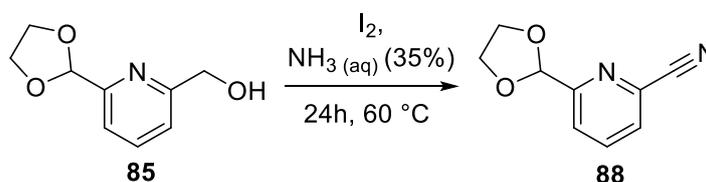


Iodine (420 mg, 1.7 mmol, 3.0 eq) was added to a solution of **85** (100 mg, 0.55 mmol, 1.0 eq) in aqueous ammonia (35%) (2.5 mL, 25 mmol, 45.0 eq) and heated to 60 °C for 24 hrs. After cooling to RT, the reaction was quenched with water (20 mL) and saturated Na₂SO₃ (aq) (3.0 mL). The aqueous layer was extracted with DCM (4 x 40 mL). The organics were combined, dried over MgSO₄ and the solvent removed under reduced pressure. The residue was then purified by flash column chromatography, eluting with 0-1% MeOH:DCM. The pure fractions

were combined, and the solvent removed under reduced pressure to give **88** (5 mg, 28 μmol, 5%) as a yellow/green oil.

**¹H NMR** (400 MHz, CDCl₃), δ 7.93-7.89 (m, 1H, Ar$\underline{H}$3), 7.79-7.77 (m, 1H, Ar$\underline{H}$4), 7.72-7.70 (m, 1H, Ar$\underline{H}$5) 5.86 (s, 1H, Ar2C$\underline{H}$), 4.23-4.08 (m, 4H, C$\underline{H_2}$C$\underline{H_2}$); **¹³C NMR** (101 MHz, CDCl₃), δ 159.27 (Ar$\underline{C}$2), 137.91 (Ar$\underline{C}$3), 133.32 (Ar$\underline{C}$6), 128.66 (Ar$\underline{C}$5), 124.10 (Ar$\underline{C}$4), 116.89 ($\underline{C}$N), 102.72 (Ar2$\underline{C}$H), 65.79 ($\underline{C}$H₂CH₂); **ν$_{max}$**: (FT-ATR)/cm$^{-1}$: 3090, 2894, 2242, 1589, 1434, 1360, 1109, 1026, 991, 948, 819; **HRMS**: *m/z* (ESI⁺) calc. for C₉H₈N₂O₂ [M+Na]⁺ 199.0478, found 199.0480.

### 6-(chloromethyl)-pyridine-2-carboxaldehyde (89)
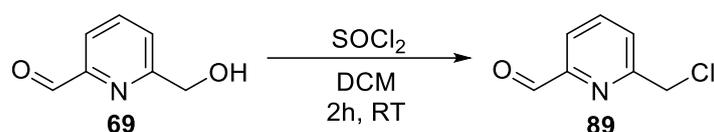


A solution of thionyl chloride (2.3 mL, 31 mmol, 1.5 eq.) in DCM (30 ml) was added dropwise to a pre-cooled solution of **69** (2.8 g, 21 mmol, 1 eq.) in DCM (150 ml) at 10-15 °C. The reaction was then allowed to return to RT and left to stir for 2 hours. The reaction mixture was then quenched dropwise with cooled saturated NaHCO₃ (aq) (150 ml) until effervescence stopped. The organic layer was separated and then further washed with saturated NaHCO₃ (aq) (2 × 150 mL) and brine (1 × 150 mL). The organic layer was then dried over MgSO₄ and the solvent removed under reduced pressure to give **89** (0.65 g, 4.2 mmol, 20%) as a red/brown oil.

**¹H NMR** (400 MHz, CDCl₃), δ 10.07 (s, 1 H, C$\underline{H}$O), 7.96-7.91 (m, 2 H, Ar$\underline{H}$3 and Ar$\underline{H}$5), 7.77-7.72 (m, 1 H, Ar$\underline{H}$4), 4.77 (s, 2 H, C$\underline{H_2}$). The spectroscopic data were consistent with that reported in the literature.[57]

### 6-[(Acetylsulfanyl)methyl]pyridine-2-carbaldehyde (93)
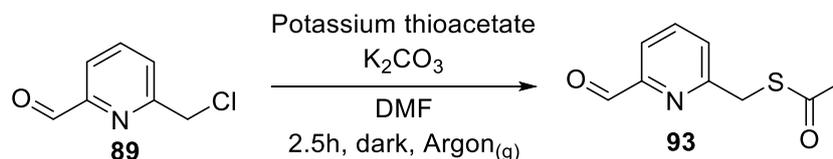


**89** (50 mg, 0.32 mmol, 1.0 eq.) was reacted with potassium thioacetate (40 mg, 0.35 mmol, 1.1 eq.) and potassium carbonate (90 mg, 0.64 mmol, 2.0 eq.) in degassed DMF (2 ml) under argon for 2.5 hours in the dark. The reaction was then filtered under vacuum and the filtrate diluted with ethyl acetate (100 ml). The organics were washed with water (2 × 100 ml) and

brine (1 × 100 ml), dried over MgSO$_4$ and the solvent removed under vacuum. The residue was then purified by flash column chromatography eluting with 20% EtOAc: Petroleum Ether. The pure fractions were combined and the solvent removed under reduced pressure to give **93** (18 mg, 9 µmol, 28%) as a pale orange solid.

**$^1$H NMR** (400 MHz, CDCl$_3$), δ 10.00 (s, 1 H, C<u>H</u>O), 7.82-7.78 (m, 2 H, ArH3 and ArH5), 7.59-7.55 (m, 1 H, ArH4), 4.31 (s, 2 H, C<u>H</u>$_2$), 2.36 (s, 3 H, C<u>H</u>$_3$); **$^{13}$C NMR** (101 MHz, CDCl3), δ 195.01 (<u>C</u>OCH$_3$), 193.45 (<u>C</u>HO), 158.69 (Ar<u>C</u>6), 152.65 (Ar<u>C</u>2), 138.07 (Ar<u>C</u>5), 127.87 (Ar<u>C</u>4), 120.48 (Ar<u>C</u>3), 35.19 (<u>C</u>H$_2$), 30.52 (<u>C</u>H$_3$); $\nu_{max}$: (FT-ATR)/cm$^{-1}$: 3387, 2929, 2832, 1693, 1589, 1455, 1354, 1133, 960, 783, 749, 627; **HRMS**: *m/z* (ESI$^+$) calc. for C$_9$H$_{10}$NO$_2$S [M+H]$^+$ 196.0429, found 196.0427; **m.p.** 44.5-49.8 °C.

**2-(chloromethyl)-6-(1,3-dioxolan-2-yl)pyridine (94)**
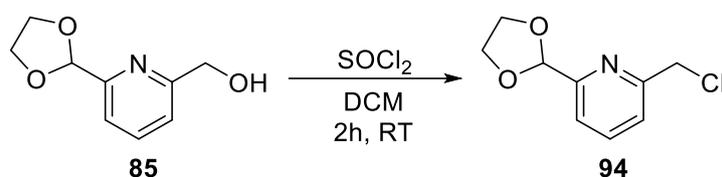


A solution of thionyl chloride (0.37 mL, 5.1 mmol, 2.0 eq) in DCM (20 mL) was added dropwise to a pre-cooled solution of **85** (460 mg, 2.5 mmol, 1.0 eq) in DCM (20 mL) at 10-15 °C. After 2 hours at room temperature the reaction was quenched dropwise with saturated NaHCO$_3$ (aq) (40 mL) until effervescence stopped. The aqueous layer was then extracted with DCM (2 x 40 mL), and the combined organics dried over MgSO$_4$ and the solvent removed under reduced pressure to give **94** (390 mg, 2.0 mmol, 77%) as a pale-yellow oil.

**$^1$H NMR** (400 MHz, CDCl$_3$), δ 7.83-7.79 (m, 1H, ArH4), 7.53 (d, *J* = 4.1 Hz, 1H, ArH5), 7.52 (d, *J* = 4.1 Hz, 1H, ArH3), 5.87 (s, 1H, Ar2C<u>H</u>), 4.73 (s, 2H, Ar6C<u>H</u>$_2$), 4.22-4.06 (m, 4H, C<u>H</u>$_2$C<u>H</u>$_2$); **$^{13}$C NMR** (101 MHz, CDCl$_3$), δ 156.34 (Ar<u>C</u>2 and Ar<u>C</u>6), 137.81 (Ar<u>C</u>4), 122.98 (Ar<u>C</u>5), 119.71 (Ar<u>C</u>3), 103.20 (Ar2<u>C</u>H), 65.57 (<u>C</u>H$_2$<u>C</u>H$_2$), 46.45 (Ar6<u>C</u>H$_2$); $\nu_{max}$: (FT-ATR)/cm$^{-1}$: 2888, 1595, 1460, 1365, 1248, 1103, 947, 819, 749, 714, 644; **HRMS**: *m/z* (ESI$^+$) calc. for C$_9$H$_{10}$ClNO$_2$ [M+Na]$^+$ 222.0292, found 222.0290.

## 1-({[6-(1,3-dioxolan-2-yl)pyridine-2-yl]methyl}sulfanyl)ethan-1-one (95)
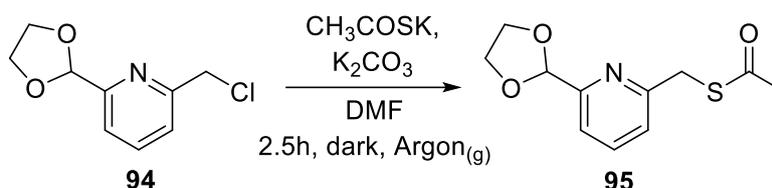


**94** (100 mg, 0.50 mmol, 1.0 eq) was added to a solution of potassium thioacetate (70 mg, 0.61 mmol, 1.2 eq) and potassium carbonate (150 mg, 1.1 mmol, 2.2 eq) in degassed anhdrous DMF (3.0 mL) under argon and left for 2.5 hours in the dark. The reaction mixture was diluted to 50 mL with EtOAc and filtered under vacuum. The organics were washed with water (2 x 50 mL) and brine (1 x 50 mL), dried over MgSO$_4$ and the solvent removed under reduced pressure to give **95** (96 mg, 0.4 mmol, 80%) as a brown oil.

**$^1$H NMR** (400 MHz, CDCl$_3$), δ 7.71-7.67 (m, 1H, ArH3), 7.45-7.43 (m, 1H, ArH4), 7.37-7.35 (m, 1H, ArH5), 5.83 (s, 1H, CHS), 4.29 (s, 2H, Ar6CH$_2$), 4.18-4.05 (m, 4H, CH$_2$CH$_2$), 2.36 (s, 3H CH$_3$); **$^{13}$C NMR** (101 MHz, CDCl$_3$), δ 195.08 (COCH$_3$), 157.12 (ArC6), 156.57 (ArC2), 137.77 (ArC3), 123.65 (ArC5), 119.00 (ArC4), 103.26 (Ar2CH), 65.55 (CH$_2$CH$_2$), 35.07 (CH$_2$S), 30.25 (CH$_3$); **ν$_{max}$**: (FT-ATR)/cm$^{-1}$: 2917, 1689, 1594, 1458, 1358, 1227, 1103, 957, 818, 752, 628; **HRMS**: *m/z* (ESI$^+$) calc. for C$_{11}$H$_{13}$NO$_3$S [M+Na]$^+$ 262.0508, found 262.0506.

## [6-(1,3-Dioxolan-2-yl)pyridine-2-yl]methanethiol (96)
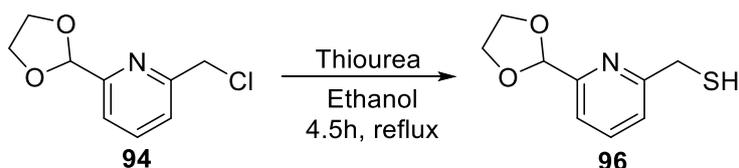


**94** (110 mg, 0.53 mmol, 1.0 eq) was combined with thiourea (81 mg, 1.1 mmol, 2.0 eq) in ethanol (5 mL) and heated to reflux for 4.5 hours. After cooling to RT, the solvent was removed under reduced pressure to give **96** (150 mg, 0.78 mmol, Quant.) as a pale brown solid.

**$^1$H NMR** (400 MHz, MeOD), δ 7.95 (m, 1H, ArH4), 7.59 (dd, *J* = 0.9, 7.8 Hz, 1H, ArH3), 7.51 (dd, *J* = 0.9, 7.8 Hz, 1H, ArH5), 5.83 (s, 1H, Ar2CH), 4.51 (s, 2H, Ar6CH$_2$), 4.14-4.05 (m, 4H, CH$_2$CH$_2$); **$^{13}$C NMR** (101 MHz, MeOD), δ 158.56 (ArC2), 157.23 (ArC6), 140.86 (ArC4), 125.11 (ArC5), 122.19 (ArC3), 103.94 (Ar2CH), 66.82 (CH$_2$CH$_2$), 37.56 (Ar6CH$_2$); **ν$_{max}$**: (FT-ATR)/cm$^{-1}$: 3290,

3173, 1610, 1467, 1395, 1106, 1085, 732, 587, 478; **HRMS**: *m/z* (ESI[+]) calc. for $C_9H_{11}NO_2S$ [M+H][+] 198.0583, found 198.0578.

### 4.2.2. PEG synthesis.

**PEG-NPC (104)**
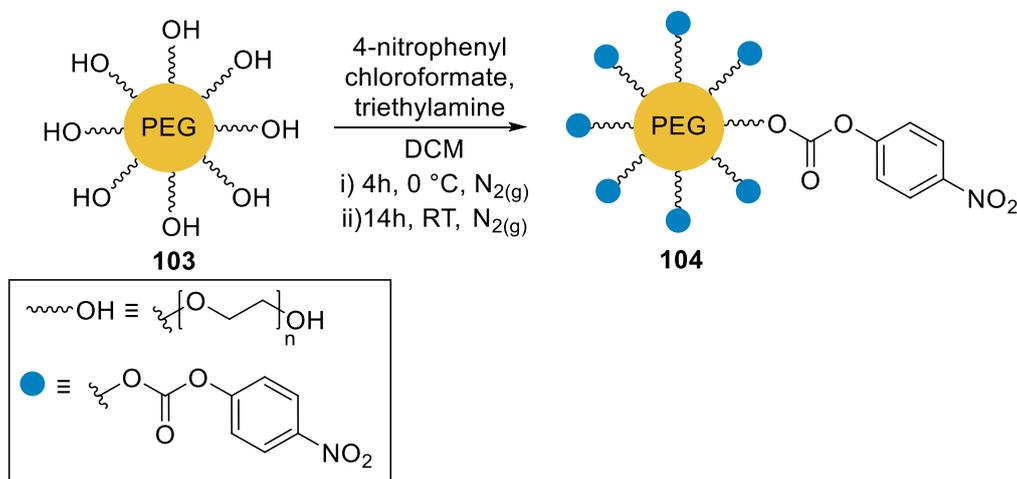


8-arm PEG-OH **103** (20 kDa $M_w$, 0.5 g, 25 µmol, 1.0 eq) was added to a solution of triethylamine (0.07 mL, 0.5 mmol, 20 eq) in DCM (5 mL) and cooled to 0 °C. 4-Nitrophenyl chloroformate (100 mg, 0.50 mmol, 20 eq) in DCM (5 mL) was added dropwise to the reaction mixture and then left for 4 hours under nitrogen at 0 °C. After 4 hours the reaction mixture was allowed to return to RT and left for a further 14 hours. The mixture was concentrated under reduced pressure to ~ 3.0 mL and subsequently, the PEG product was precipitated in diethyl ether (100 mL) and the precipitate isolated via centrifuge (3000 rpm, 6 min). The isolated precipitate was redissolved in DCM (10 mL) and the solvent was then removed under vacuum to give PEG-NPC **104** (0.58 mg, 27 µmol, Quant) as a yellow solid.

**^{1}H NMR** (400 MHz, CDCl$_3$), *δ* 8.29 (d, *J* = 9.2 Hz, 16 H, ArC<u>H</u>), 7.40 (d, *J* = 9.2 Hz, 16H ArC<u>H</u>). PEG peaks are not assigned for clarity. The spectroscopic data were consistent with that shown for 4-nitrophenyl chloroformate in the literature.[58]

**PEG-TA (105)**
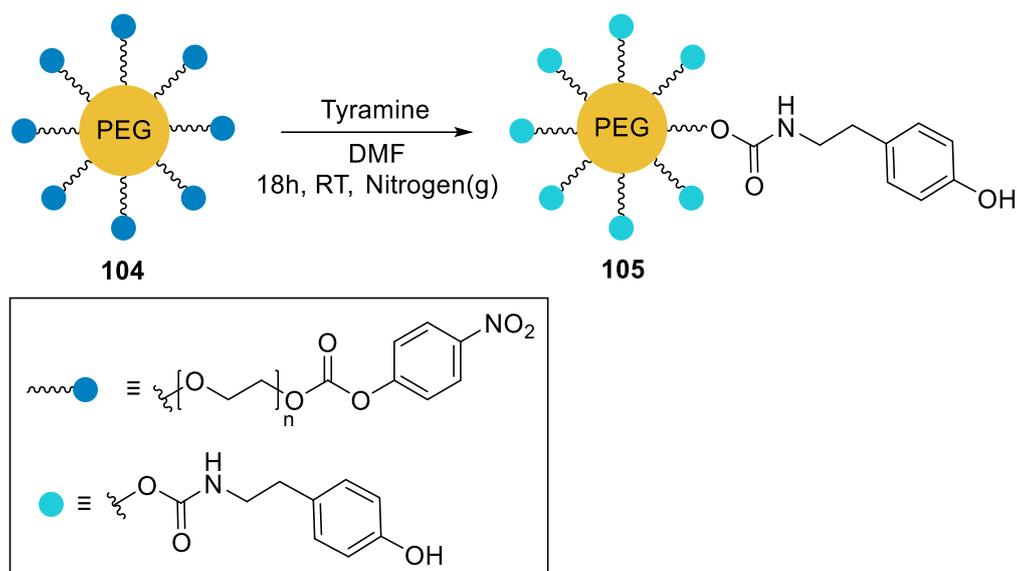


Tryramine (150 mg, 1.1 mmol, 40 eq) in DMF (5 mL) was added dropwise to a solution PEG-NPC **104** (0.56 g, 26 μmol, 1.0 eq) in DCM (5 mL). After complete addition, the reaction was left for 18 hours under a nitrogen atmosphere. The mixture was concentrated under reduced pressure to ~ 3.0 mL and subsequently, the PEG product was precipitated in excess diethyl ether (100 mL) to remove *p*-nitrophenol. The precipitate was then dissolved in DCM and filtered over sinter to remove any unreacted tyramine. The filtrate was then dried under vacuum to give PEG-TA **105** (430 mg, 20 μmol, 77%) as an off white solid solid.

**$^1$H NMR** (400 MHz, CDCl$_3$), δ 7.02 (d, *J* = 8.2 Hz, 16 H, ArC<u>H</u>), 6.80 (d, *J* = 8.2 Hz, 16H ArC<u>H</u>), 3.47 (t, *J* = 5.0 Hz, 16 H, C<u>H</u>$_2$), 2.73 (t, *J* = 5.0 Hz, 16 H, C<u>H</u>$_2$). PEG peaks are not assigned for clarity. The spectroscopic data were consistent with that reported for tyramine in the literature.[59]

**PEG-NH-PCA (106)**
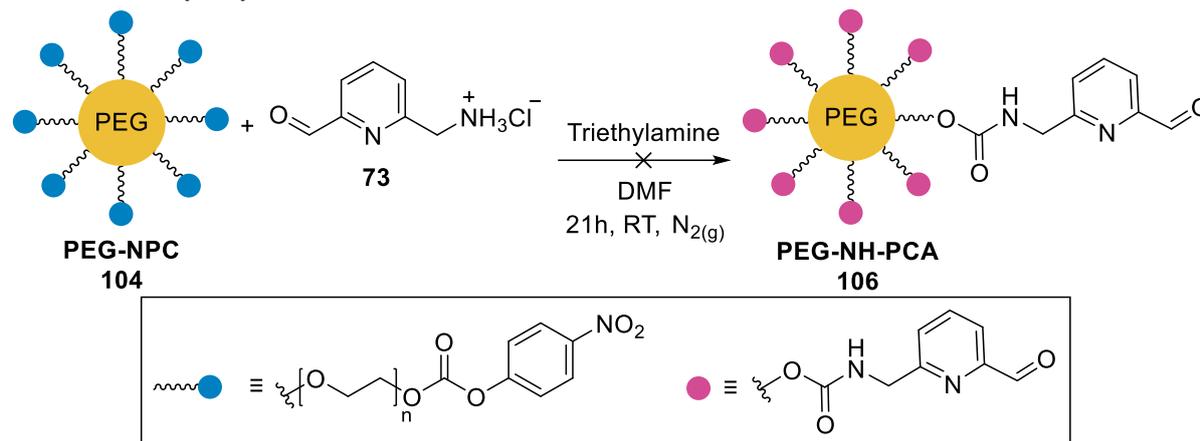


Triethylamine (0.33 mL, 2.4 mmol, 100.0 eq) and **73** (170 mg, 0.96 mmol, 40.0 eq) were combined in DMF (5 mL). PEG-NPC **104** (490 mg, 24 µmol, 1.0 eq) in DMF (5 mL) was then added dropwise to the reaction mixture. Once all the PEG-NPC had been added the reaction mixture was degassed for 30 mins and then left for a further 21 hours under nitrogen. We attempted to precipitate the product in ice cold diethyl ether (100 mL) only to find that the reaction had formed a brown gel. We were unable to dissolve any of the gel for analysis.

**PEG-Allyl (107)**
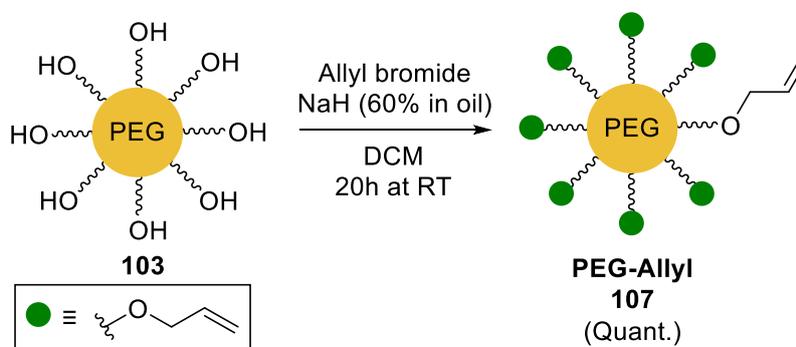


NaH (60% in oil, 30 mg, 0.75 mmol, 30 eq) was added slowly to 8-arm PEG-OH **103** (20 kDa $M_w$, 0.5 g, 25 µmol, 1.0 eq) in DCM (10 mL) and stirred for 15 minutes. Allyl bromide (0.06 mL, 0.75 mmol, 30 eq) was then added and the reaction left for 20 hours. The reaction mixture was filtered through cotton wool and the solvent removed under reduced pressure. The residue was dissolved in DCM (50 mL) and the organics washed with brine (3 x 50 mL), dried over $MgSO_4$ and concentrated to ~5 mL under reduced pressure. The remaining sample was then precipitated in hexane (100 mL) and the precipitate repeated via centrifugation. The hexane was decanted off and the precipitate collected in DCM and combined before removing

the solvent under reduced pressure to give PEG-Allyl **107** (0.55 g, 27 µmol, Quant) as a white solid.

**¹H NMR** (400 MHz, CDCl₃), δ 5.97-5.87 (m, 8H, CH₂CHCH), 5.26-5.17 (m, 8H, CH₂CHCH), 4.04-4.02 (m, 16H, CH₂CHCH). PEG peaks are not assigned for clarity. The spectroscopic data were consistent with that reported for allyl alcohol in the literature.[60]

**PEG-SH-PCA(acetal) (108)**
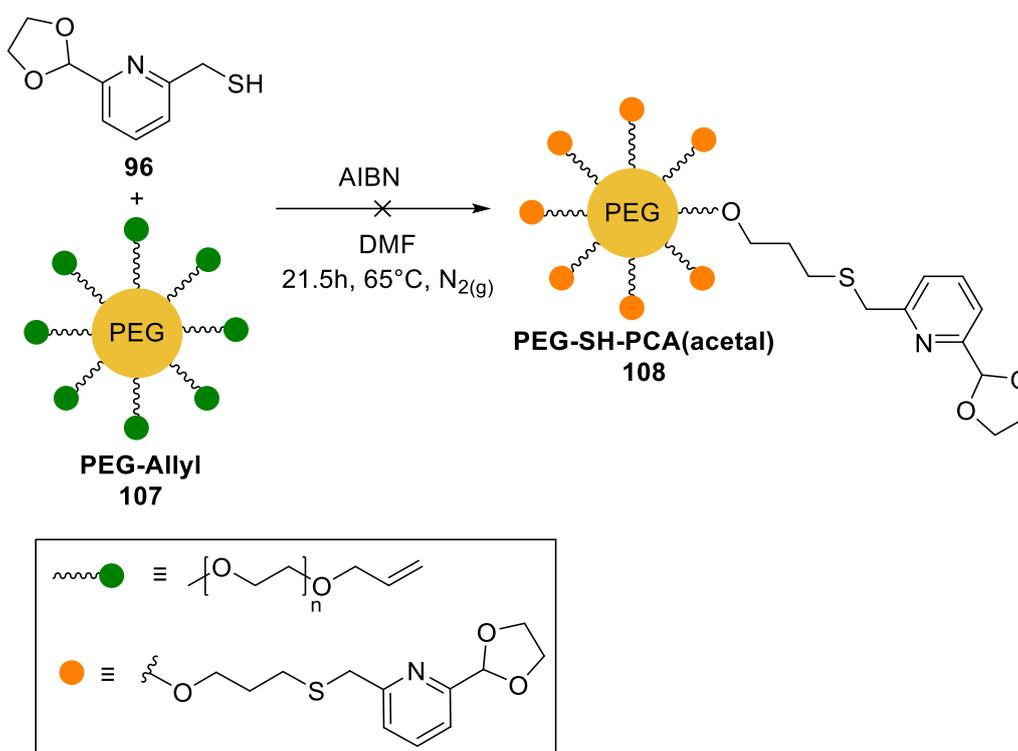


PEG-Allyl **107** (41 mg, 2 µmol, 1.0 eq) in a solution of DMF (5 mL) was degassed under nitrogen for 20 mins before adding AIBN (25 mg, 0.15 mmol, 100 eq) and **96** (30 mg, 0.15 mmol, 100 eq) and degassing under nitrogen for a further 20 mins. The reaction mixture was then left for 21.5 hours under nitrogen at 65 °C. Attempted to precipitate product in ice cold diethyl ether (100 mL). We were unable to achieve a good separation via centrifugation and so removed all solvent under reduced pressure. Redissolved residue in ~5 mL methanol and reattempted to precipitate product in ice cold diethyl ether (100 mL). Removed solvent under

reduced pressure. Added ice cold diethyl ether (100 mL) to residue and stirred for 10 mins before filtering through sinter. Residue collected and put on for NMR – NMR showed that **96** hadn't attached to PEG-Allyl.

### 4.2.3. Crosslinker Synthesis.

**Triethylene glycol ditosylate (112)**
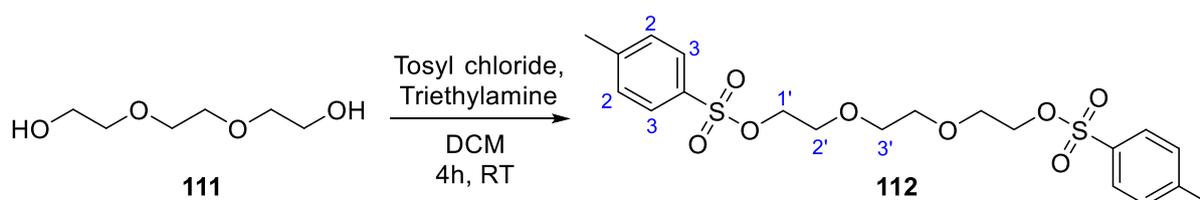


Triethylene glycol **111** (1.0 mL, 7.5 mmol, 1.0 eq) was added to a solution of *p*-toluenesulfonyl chloride (3.1 g, 16 mmol, 2.2 eq) and triethylamine (3.1 mL, 22 mmol, 3.0 eq) in DCM (40 mL) and stirred for 4 hours. The reaction mixture was then diluted with DCM (60 mL) and the organics washed with water (2 x 100 mL) and brine (1 x 100 mL), dried over $MgSO_4$ and the solvent removed under reduced pressure. The residue was purified by flash column chromatography, eluting with 50% EtOAc:petroleum ether (40-60 °C). The pure fractions were combined, and the solvent removed under reduced pressure to give **112** (2.7 g, 5.8 mmol, 77%) as a white solid.

**$^1$H NMR** (400 MHz, CDCl$_3$), δ 7.79 (d, *J* = 8.2 Hz, 4H, ArH2/ArH3), 7.34 (d, *J* = 8.2 Hz, 4H, ArH2/ArH3), 4.14 (t, *J* = 4.6 Hz, 4H, CH$_2$-1'/CH$_2$-2'), 3.65 (t, *J* = 4.6 Hz, 4H, CH$_2$-1'/CH$_2$-2'), 3.53 (s, 4H, CH$_2$-3'), 2.45 (s, 6H, CH$_3$); **HRMS**: *m/z* (ESI$^+$) calc. for C$_{20}$H$_{26}$O$_8$S$_2$ [M+Na]$^+$ 481.0961, found 481.0953. The spectroscopic data were consistent with that shown reported in the literature.[61]

**2-(1,3-dioxolan-2-yl)-6-{12-[6-(1,3-dioxolan-2-yl)pyridine-2-yl]-2,5,8,11-tetraoxadodecan-1-yl}pyridine (113)**
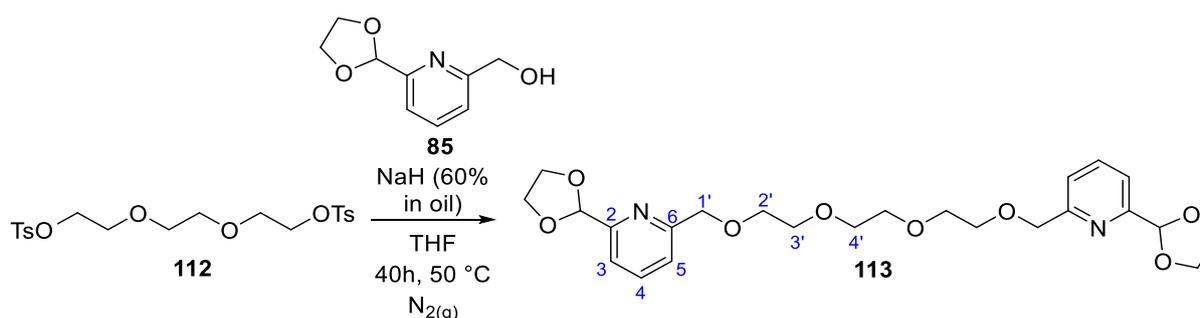
NaH (60% in oil, 72 mg, 1.8 mmol, 2.5 eq) was added slowly to a solution **85** (130 mg, 0.72 mmol, 1.0 eq) in anhdrous THF (10 mL). Once effervescence had stopped, **112** (170 mg, 0.36 mmol, 0.5 eq) was added and the reaction left for 40 hours at 50 °C under nitrogen. After cooling to RT, the reaction was quenched with water (15 mL) and aqueous citric acid (10% w/v, 1.0 mL), and the aqueous layer extracted with DCM (5 x 50 mL). The organics were combined, dried over MgSO$_4$ and the solvent removed under reduced pressure. The residue was purified by flash column chromatography, eluting with 0-3.5% MeOH:DCM. The pure fractions were combined, and the solvent removed under reduced pressure to give **113** (39 mg, 80 µmol, 22%) as a colourless oil.

**$^1$H NMR** (400 MHz, CDCl$_3$), δ 7.75-7.71 (m, 2H, Ar$\underline{H}$4), 7.49 (d, $J$ = 7.8 Hz, 2H, Ar$\underline{H}$5), 7.43 (d, $J$ = 7.8 Hz, 2H, Ar$\underline{H}$3), 5.81 (s, 2H, Ar2C$\underline{H}$), 4.71 (s, 4H, C$\underline{H}_2$-1'), 4.19-4.03 (m, 8H, C$\underline{H}_2$C$\underline{H}_2$), 3.70 (m, 12H, C$\underline{H}_2$C$\underline{H}_2$); **$^{13}$C NMR** (101 MHz, CDCl$_3$), δ 158.49 (Ar$\underline{C}$6), 156.15 (Ar$\underline{C}$2), 137.41 (Ar$\underline{C}$4), 121.45 (Ar$\underline{C}$5), 118.99 (Ar$\underline{C}$3), 103.52 (Ar2$\underline{C}$H), 73.85 (C$\underline{H}_2$-1'), 70.69/70.53/70.23 (C$\underline{H}_2$-2',C$\underline{H}_2$-3' and C$\underline{H}_2$-4'), 65.51 (C$\underline{H}_2$C$\underline{H}_2$); $\nu_{max}$: (FT-ATR)/cm$^{-1}$: 3446, 2895, 1719, 1596, 1460, 1354, 1105, 946, 800, 645; **HRMS**: $m/z$ (ESI$^+$) calc. for C$_{24}$H$_{32}$N$_2$O$_8$ [M+Na]$^+$ 499.2051, found 499.2056.

4.2.4. Fluorophore synthesis.

**N-Boc-Ethylenediamine (116)**
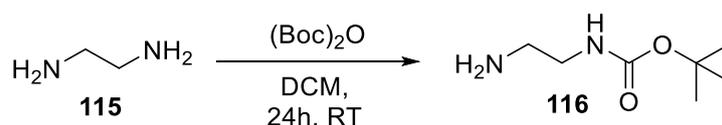


A solution of di-*tert*-butyl dicarbonate (3.1 g, 14 mmol, 1.0 eq) in DCM (150 mL) was added dropwise to a solution of ethylene diamine **115** (5.6 mL, 84 mmol, 6.0 eq) in DCM (50 mL). The reaction mixture was then left to stir for 24 hours. The solvent was then removed under vacuum and the residue dissolved in 2 M Na$_2$CO$_3$ (aq) (100 mL). The aqueous was then extracted with DCM (3 × 100 mL). The organic layers were combined and dried over MgSO$_4$ before the solvent was removed under vacuum to give **116** (2.0 g, 12 mmol, 89%) as a pale-yellow oil.

**$^1$H NMR** (400 MHz, CDCl$_3$), δH 5.00 (br s, 1 H, N$\underline{H}$), 3.17-3.13 (m, 2 H, C$\underline{H}_2$), 2.79-2.76 (m, 2 H, C$\underline{H}_2$), 1.42 (s, 9 H, $^t$Bu). The spectroscopic data were consistent with those reported in the literature.[62]

## Boc-Ala-O-N-hydroxysuccinimide (118)
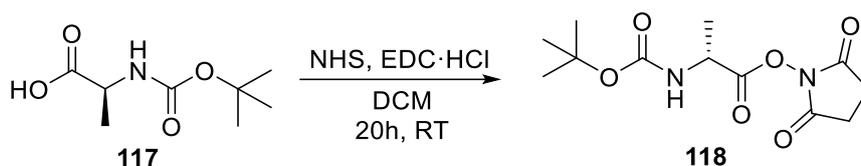


EDC·HCl (2.0 g, 11 mmol, 2.0 eq) was added to a solution of Boc-L-Ala-OH **117** (1.0 g, 5.3 mmol, 1.0 eq) and NHS (0.90 g, 7.9 mmol, 1.5 eq) in DCM (60 mL) and stirred at RT for 20 hours. The solution was diluted to 100 mL with DCM and the organics washed with water (2 x 100 mL) and brine (1 x 100 mL), dried over MgSO$_4$ and the solvent removed under reduced pressure to give **118** (1.3 g, 4.6 mmol, 87%) as a white solid.

**$^1$H NMR** (400 MHz, CDCl$_3$), δH 5.00 (d, $J$ = 4.0 Hz, 1 H, NH), 4.73-4.70 (m, 1 H, Hα), 2.85 (s, 4 H, CH$_2$CH$_2$), 1.58 (d, $J$ = 8.0 Hz, 3 H, Hβ), 1.46 (s, 9 H, $_t$Bu). The spectroscopic data were consistent with those reported in the literature.[63]

## 7-Hydroxycoumarin-3-carboxylic acid ethyl ester (121)
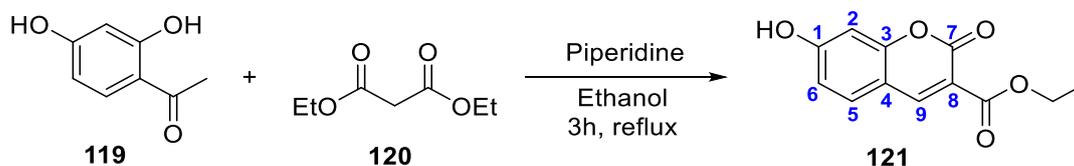


2, 4-Dihydroxybenzaldehyde **119** (2.0 g, 15 mmol, 1.0 eq), diethyl malonate **120** (2.7 mL, 17 mmol, 1.2 eq) and piperidine (0.07 mL, 0.73 mmol, 0.05 eq) were combined in ethanol (30 mL) and refluxed for 3 hours. After cooling to RT, the solvent was removed under vacuum and the residue was concentrated onto silica and purified by flash column chromatography eluting with 20% EtOAc: Petroleum Ether (40-60 °C). The pure fractions were combined, and the solvent removed under vacuum to give **121** (2.2 g, 9.2 mmol, 66%) as a yellow powder.

**$^1$H NMR** (400 MHz, CDCl$_3$), δH 8.56 (s, 1H, ArH9), 7.52 (d, $J$ = 8.7 Hz, 1 H, ArH5), 7.00 (d, $J$ = 2.3 Hz, 1H, ArH2), 6.91 (dd, $J$ = 8.7, 2.3 Hz, 1H, ArH6), 4.42 (q, $J$ = 7.3, 2H, CH$_2$), 1.41 (t, $J$ = 7.3, 3H, CH$_3$). The spectroscopic data were consistent with those reported in the literature.[64]

## 7-Hydroxycoumarin-3-carboxylic acid (122)
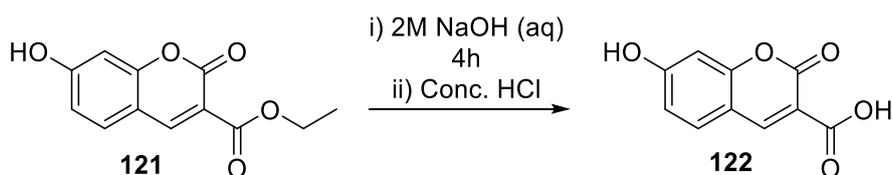
**121** (2.0 g, 8.5 mmol, 1.0 eq) was dissolved in 2 M NaOH (aq) (25 mL) and stirred for 4 hours. Concentrated HCl was then added dropwise to the reaction mixture until a yellow precipitate formed. The precipitate was collected by filtration and washed with HCl (2 × 10 mL). The solid was then dissolved in methanol (20 mL) which was then removed under vacuum to give **122** (0.88 g, 4.3 mmol, 50%) as a dark yellow/ orange powder.

**$^1$H NMR** (400 MHz, Acetone-$d_6$), δ 8.86 (s, 1H, ArH9), 7.89 (d, $J$ = 8.7 Hz, 1 H, ArH5), 7.04 (dd, $J$ = 8.7, 2.3 Hz, 1H, ArH6), 6.91 (d, $J$ = 2.3 Hz, 1H, ArH2). The spectroscopic data were consistent with those reported in the literature.[65]

**7-Hydroxycoumarin-3-carboxylic acid N-succinimidyl ester (123)**



EDC·HCl (120 mg, 0.60 mmol, 1.0 eq) and NHS (70 mg, 0.60 mmol, 1.0 eq) were added to a solution of **122** (120 mg, 0.60 mmol, 1.0 eq) in DMF (2.5 mL) and left to stir at RT for 24 hours. The reaction was then quenched with aqueous citric acid (10% w/v, 10 mL). The aqueous layer was extracted using EtOAc (3 × 30 mL). The combined organic layers were washed with brine (2 × 80 mL), dried over MgSO$_4$ and the solvent removed under vacuum to give **123** (160 mg, 0.53 mmol, 88%) as a yellow solid.

**$^1$H NMR** (400 MHz, MeOH), δH 8.95 (s, 1 H, ArH9), 7.71 (d, $J$ = 8.4 Hz, 1 H, ArH5), 6.90 (dd, $J$ = 8.4, 2.3 Hz, 1 H, ArH6), 6.76 (d, 2.3 Hz, 1 H, ArH2), 2.90 (s, 4 H, CH$_2$CH$_2$). The spectroscopic data were consistent with those reported in the literature.[66]

**Tert-butyl N-{2-[(7-hydroxy-2-oxo-2H-chromen-3-yl)formamido]ethyl}carbamate (124)**
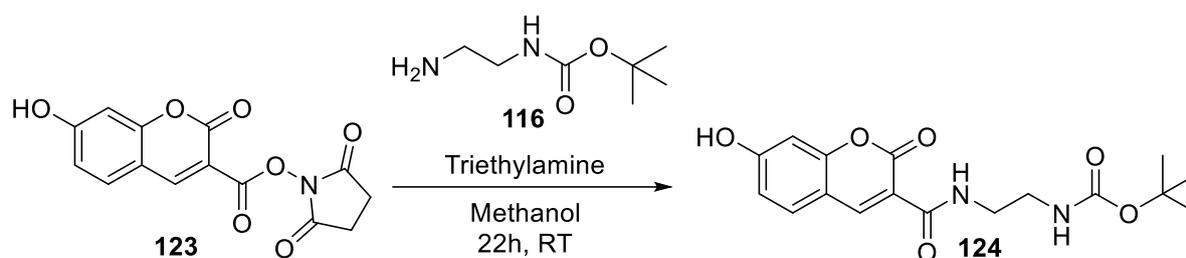
**116** (0.22 g, 1.4 mmol, 1.0 eq) was added to a solution of **123** (0.42 g, 1.4 mmol, 1.0 eq) and triethylamine (0.38 mL, 2.8 mmol, 2.0 eq) in methanol (50 mL) and left to stir for 22 hours. The solvent was removed under reduced pressure and the residue re-dissolved in EtOAc (60 mL). The organic layer was washed with water (2 × 50 mL) and brine (1 × 50 mL) and then dried over MgSO₄ and filtered. The Solvent was removed under reduced pressure to give **124** (180 mg, 0.52 mmol, 38%) as a yellow powder.

**¹H NMR** (400 MHz, DMSO-$d_6$), δ 8.79 (s, 1H, Ar$\underline{H}$9), 8.71 (t, $J$ = 6.0 Hz, 1 H, ArCON$\underline{H}$), 7.82 (d, $J$ = 8.7 Hz 1 H, Ar$\underline{H}$5), 6.95 (t, $J$ = 5.5 Hz, 1 H, N$\underline{H}$Boc), 6.88 (dd, $J$ = 8.7, 2.3 Hz, 1 H, Ar$\underline{H}$6), 6.80 (d, $J$ =2.3 Hz, 1H, Ar$\underline{H}$2), 3.36 (m, 2H, NHC$\underline{H}_2$), 3.10 (m, 2H, C$\underline{H}_2$NHBoc), 1.36 (s, 9H, $^t$Bu) ; **¹³C NMR** (101 MHz, DMSO-$d_6$), δ 163.88 (Ar$\underline{C}$3), 161.83 (Ar$\underline{C}$O), 160.91 (Ar$\underline{C}$8), 156.33 (Ar$\underline{C}$4), 155.74 (NH$\underline{C}$OO$^t$Bu), 148.04 (Ar$\underline{C}$9), 132.02 (Ar$\underline{C}$5), 114.43 (Ar$\underline{C}$6), 113.47 (Ar$\underline{C}$7), 111.01 (Ar$\underline{C}$1), 101.80 (Ar$\underline{C}$2), 77.69 ($\underline{C}$(CH₃)₃), 40.13 ($\underline{C}$H₂NHBoc), 39.63 (NH$\underline{C}$H₂), 28.22 (C($\underline{C}$H₃)₃); **ν$_{max}$**: (FT-ATR)/cm⁻¹: 3329, 2983, 2931, 1707, 1617, 1539, 1453, 1368, 1223, 1166, 851, 796; **HRMS**: $m/z$ (ESI⁺) calc. for $C_{17}H_{20}N_2O_6$ [M+Na]⁺ 371.1214, found 371.1212; **m.p.** 221.1-224.4 °C.

**N-(2-aminoethyl)-7-hydroxy-2-oxo-2H-chromene-3-carboxamide; trifluoroacetic acid (125)**
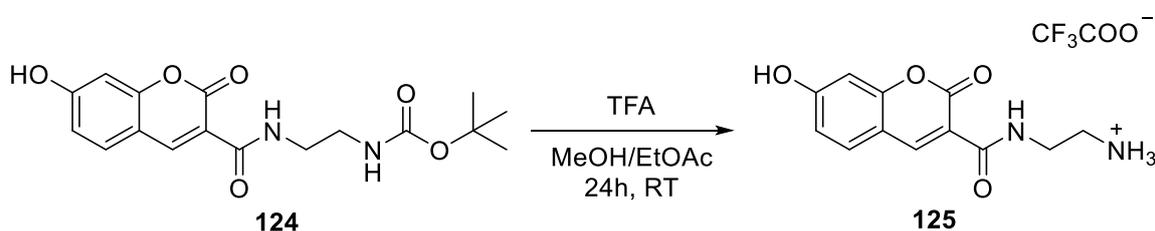


**124** (0.012 g, 0.034 mmol, 1.0 eq) was added to a mixture of TFA (2 mL) and methanol (1.5 mL) and left to stir. After 2 hours another 1 mL of TFA was added to the reaction mixture and left to stir for a further 22 hours. The solution was diluted to 15 mL with methanol and the solvent was removed under reduced pressure to give **125** (10 mg, 28 μmol, 82%) as a brown solid.

**¹H NMR** (400 MHz, MeOD), δ 8.79 (s, 1 H, Ar$\underline{H}$9), 7.67 (d, *J* = 8.2 Hz, 1 H, Ar$\underline{H}$5), 6.90 (dd, *J* = 8.2, 2.3 Hz, 1 H, Ar$\underline{H}$6), 6.78 (d, *J* = 2.3 Hz, 1 H, Ar$\underline{H}$2), 3.70 (t, *J* = 5.5 Hz, 2 H, NHC$\underline{H}_2$), 3.18 (t, *J* = 5.5 Hz, 2 H, C$\underline{H}_2$NH$_2$); **¹³C NMR** (101 MHz, MeOD), δ 165.90 (Ar$\underline{C}$3/Ar$\underline{C}$O), 165.87 (Ar$\underline{C}$3/Ar$\underline{C}$O), 162.94 (Ar$\underline{C}$8), 158.38 (Ar$\underline{C}$4), 150.01 (Ar$\underline{C}$9), 132.99 (Ar$\underline{C}$5), 115.72 (Ar$\underline{C}$6), 114.08 (Ar$\underline{C}$7), 112.65 (Ar$\underline{C}$1), 103.05 (Ar$\underline{C}$2), 40.94 ($\underline{C}$H$_2$NH$_3^+$), 38.44 (NH$\underline{C}$H$_2$); **ν$_{max}$**: (FT-ATR)/cm$^{-1}$: 3073, 1678, 1616, 1461, 1372, 1327, 1201, 1133, 837, 796, 723; **HRMS**: *m/z* (ESI$^+$) calc. for C$_{12}$H$_{12}$N$_2$O$_4$ [M+H]$^+$ 249.0870, found 249.0869; **m.p.** 210.7-216.0 °C.

## Tert-butyl N-[(1S)-1-({2-[7-hydroxy-2-oxo-2H-chromen-3-yl)formamido]ethyl}carbamoyl)ethyl]carbamate (126)
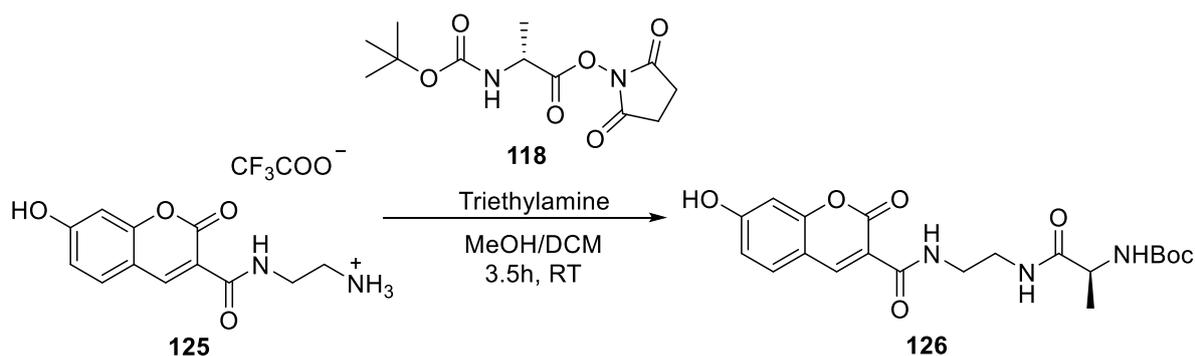


A solution of **118** (70 mg, 0.24 mmol, 1.0 eq) in DCM (10 mL) was added to a solution of **125** (60 mg, 0.24 mmol, 1.0 eq) and triethylamine (0.1 mL, 0.72 mmol, 3.0 eq) in methanol (10 mL) and stirred for 3.5 hours at RT. The solvents were removed under reduced pressure and the residue re-dissolved in EtOAc (20 mL). The organics were washed with water (2 × 20 mL) and brine (1 × 20 mL) and then dried over MgSO$_4$. The residue was then purified by flash column using 30% EtOAc: petroleum ether (40-60 °C) and the product containing eluent was combined. The solvent was removed under reduced pressure to give **126** (53 mg, 0.13 mmol, 54 %) as a yellow solid.

**¹H NMR** (400 MHz, MeOD), δ 8.74 (s, 1H, Ar$\underline{H}$9), 7.65 (d, *J* = 8.2 Hz, 1H, Ar$\underline{H}$5), 6.87 (dd, *J* = 8.2, 2.3 Hz, 1H, Ar$\underline{H}$6), 6.75 (d, *J* = 2.3, 1H, Ar$\underline{H}$2), 4.03 (q, *J* = 6.9 Hz, 1H, Hα), 3.58-3.52 (m, 2H, C$\underline{H}_2$), 3.46-3.38 (m, 2H, C$\underline{H}_2$), 1.41 (s, 9H, $^t$Bu), 1.29 (d, *J* = 6.9 Hz, 3H, Hβ); **¹³C NMR** (101 MHz,

MeOD), δ 176.34 (COCαH), 165.97 (ArC3), 164.93 (ArCO), 162.92 (ArC8), 158.30 (ArC4), 157.54 (COOᵗBu), 149.60 (ArC9), 132.85 (ArC5), 115.74 (ArC6), 114.28 (ArC7), 112.59 (ArC1), 103.04 (ArC2), 80.47 (C(CH₃)₃), 51.70 (Cα), 40.08 (CH₂NH), 39.94 (NHCH₂), 28.61 (C(CH₃)₃), 18.51 (Cβ); $\nu_{max}$: (FT-ATR)/cm$^{-1}$: 3317, 2979, 2934, 2486, 1706, 1536, 1455, 1368, 1226, 1166, 795; **HRMS**: *m/z* (ESI⁺) calc. for $C_{20}H_{25}N_3O_7$ [M+Na]⁺ 442.1585, found 442.1592; **m.p.** 185.8-195.6 °C.

**(2S)-2-amino-N-{2-[(7-hydroxy-2-oxo-2H-chromen-3-yl)formamido]ethyl}propenamide; trifluoroacetic acid (127)**
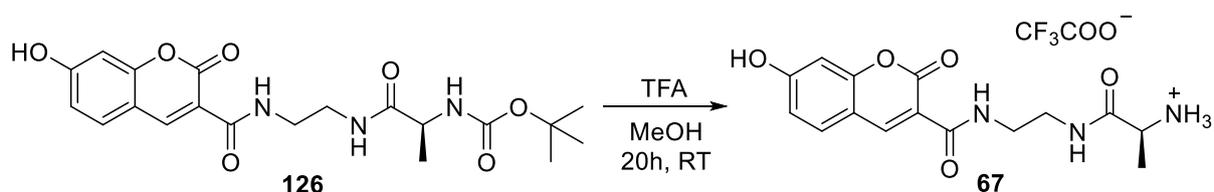


**126** (53 mg, 0.13 mmol, 1.0 eq) was added to a solution of TFA (5 mL) in MeOH (10 mL) and left to stir for 2 hours. The solvent was removed under reduced pressure and analysed by NMR. The NMR showed that not all the Boc group was removed and so the crude product was redissolved in TFA (10 mL) and MeOH (10 mL) and left to stir for 20 hours. The solvent was removed under reduced pressure to give **67** (15 mg, 47 μmol, 36%) as a brown oil.

**¹H NMR** (400 MHz, MeOD), δ 8.76 (s, 1H, ArH9), 7.66 (d, *J* = 8.4 Hz, 1H, ArH5), 6.89 (dd, *J* = 2.3, 8.4 Hz, 1H, ArH6), 6.77 (d, *J* = 2.3 Hz, 1H, ArH2), 3.89 (q, *J* = 6.9 Hz, 1H, Hα), 3.51 (m, 4H, CH₂CH₂), 1.50 (d, *J* = 6.9 Hz, 3H, Hβ); **¹³C NMR** (101 MHz, MeOD), δ 171.39 (COCαH), 166.01 (ArC3), 165.25 (ArCO), 163.12 (ArC8), 158.52 (ArC4), 150.03 (ArC9), 133.13 (ArC5), 115.91 (ArC6), 114.51 (ArC7), 112.88 (ArC1), 103.23 (ArC2), 50.45 (Cα), 40.96 (CH₂NH), 39.46 (NHCH₂), 17.72 (Cβ); $\nu_{max}$: (FT-ATR)/cm$^{-1}$: 3095, 2925, 2854, 1677, 1543, 1453, 1374, 1203, 1135, 798, 723, 518; **HRMS**: *m/z* (ESI⁺) calc. for $C_{15}H_{17}N_3O_5$ [M+H]⁺ 320.1241, found 320.1246.

4.2.5. Peptide synthesis.

Solid-phase peptide synthesis (SPPS) was performed on a CEM Liberty Lite Automated Microwave Peptide Synthesiser, according to the manufacturers standard protocols. Briefly,

Fmoc-protected amino acids (5 equiv., 0.2 M in DMF) were coupled in the presence of *N,N'*-diisopropylcarbodiimide (DIC, 15 equiv.) and Oxyma Pure (5 equiv.), as coupling agent and base respectively, under microwave irradiation at a temperature of 90 $^{\circ}$C for 2 minutes. Cysteine amino acids were coupled at 50 °C for 10 minutes. Fmoc deprotection was performed using 20% piperidine in DMF at 90 $^{\circ}$C for 60 seconds. Syntheses were performed on a 0.1 mmol scale, using Rink Amide MBHA resin (*C*-terminal amide, 0.5 mmol/g, 1% DVB, 100-200 mesh, Fluorochem). Prior to cleavage, the resin was washed sequentially with DCM (3 x 15 mL) and methanol (3 x 15 mL). Peptides were cleaved from the resin in 20 mL of cleavage cocktail (90% TFA, 5% $H_2O$, 3% TIPS, 2% DTT for Cys-containing sequences) for 3 hrs (unless stated otherwise). After filtration, the resin was washed extensively with DCM (3 x 10 mL) and the filtrate concentrated *in vacuo* to ~2 mL volume. The residue was dropped into ice cold diethyl ether (~100 mL), and the resultant precipitate collected by centrifugation (3000 rpm, 5 min), resuspended in diethyl ether, and centrifuged again. The residual solid was allowed to air dry for 10 min, then dissolved in deionized water (10 mL) and dried by lyophilisation. Peptides were typically pure by LC-MS analysis and used directly.

$$H_2N-Ala-Ala-Lys-Ala-Ala-NH_2$$
$$|$$
$$COOH$$

**109**
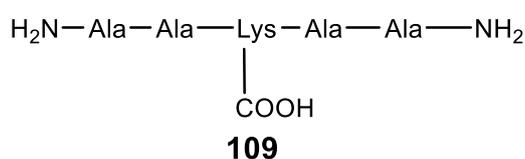
Peptide K(AA)$_2$ **109** was synthesised according to the general SPPS procedure to yield an off-white solid (26 mg, 60 µmol, 11%).

$$H_2N-Ala-Ala-Glu-Lys-Cys-CONH_2$$

**110**
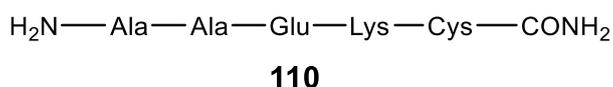
Peptide AAEKC **110** was synthesised according to the general SPPS procedure with a resin cleavage time of 20 hours to yield an off-white solid (110 mg, 0.19 mmol, 34%).
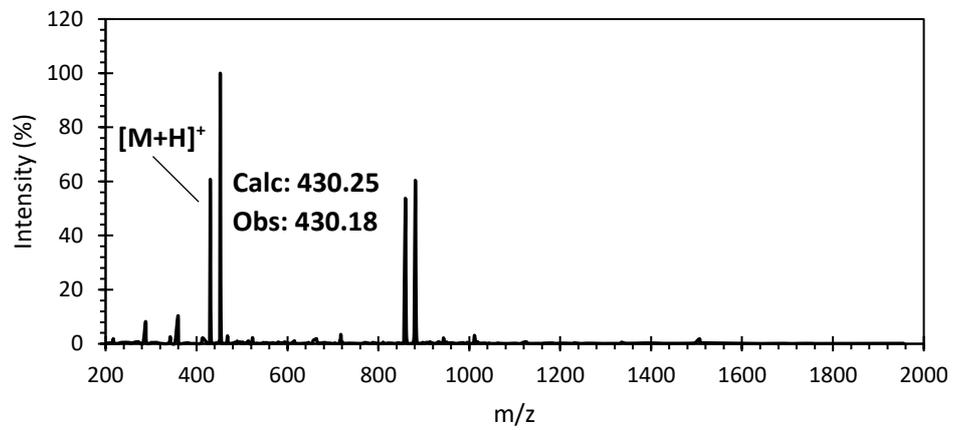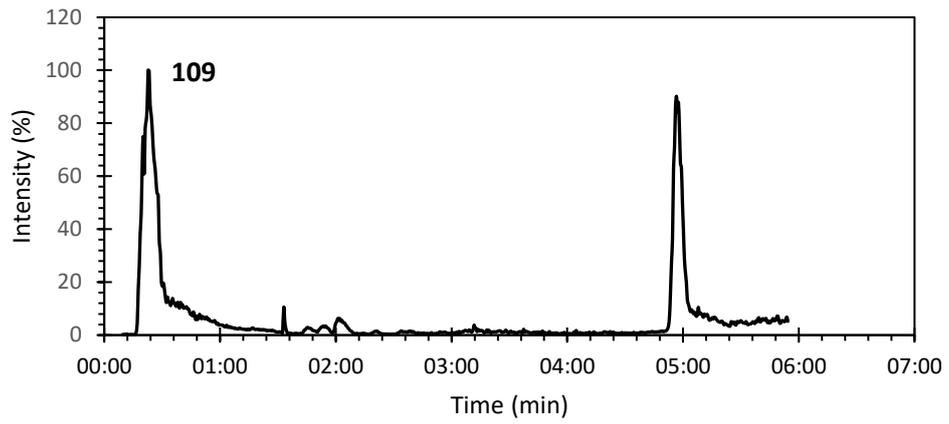
**Fig. 33.** 220 nm LC-MS UV absorption trace of peptide **109** with the corresponding MS spectrum.
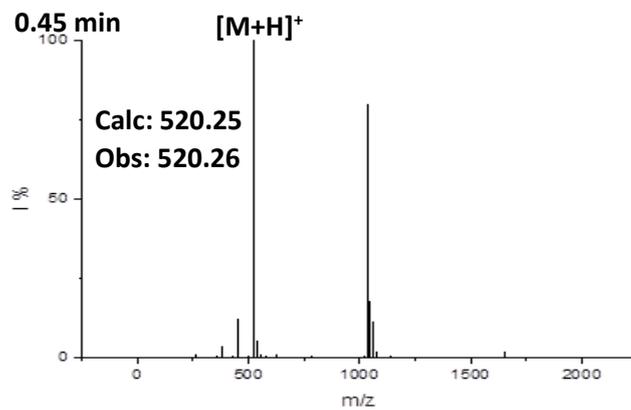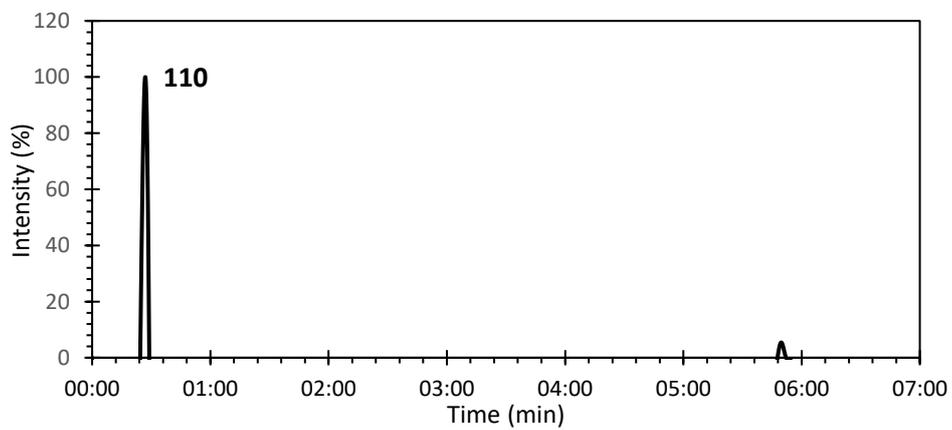


**Fig. 34.** 220 nm LC-MS UV absorption trace of peptide **110** with the corresponding MS spectrum.

# 5. Bibliography.

**[1]** O'Brien, F.J. (2011). Biomaterials & scaffolds for tissue engineering. *Materials Today*, 14(3), pp.88–95.

**[2]** Place, E.S., Evans, N.D. and Stevens, M.M. (2009). Complexity in biomaterials for tissue engineering. *Nature materials*, 8(6), pp.457–470.

**[3]** Oliver-Cervelló, L., Martin-Gómez, H. and Mas-Moruno, C. (2021). New trends in the development of multifunctional peptides to functionalize biomaterials. *Journal of Peptide Science*, 28(1).

**[4]** NHS Choices (2019). *Treatment - Rheumatoid arthritis*. [online] NHS. Available at: https://www.nhs.uk/conditions/rheumatoid-arthritis/treatment/ [Accessed 2 Apr. 2022].

**[5]** Kong, F., Mehwish, N., Niu, X., Lin, M., Rong, X., Hu, F. and Lee, B.H. (2021). Personalized hydrogels for individual health care: preparation, features, and applications in tissue engineering. *Materials Today Chemistry*, 22.

**[6]** Todros, S., Todesco, M. and Bagno, A. (2021). Biomaterials and Their Biomedical Applications: From Replacement to Regeneration. *Processes*, 9(11), p.1949.

**[7]** Liu, H., Lu, J., Jiang, Q., Haapasalo, M., Qian, J., Tay, F.R. and Shen, Y. (2021). Biomaterial scaffolds for clinical procedures in endodontic regeneration. *Bioactive Materials*, 12.

**[8]** Affar, A., Xu, F. and Hoare, T. (2018). *Poly(ethylene glycol) (PEG) and Synthetic PEG Derivatives for Tissue Engineering and Cell Delivery*. [online] sigmaaldrich.com. Available at: https://www.sigmaaldrich.com/GB/en/technical-documents/technical-article/materials-science-and-engineering/tissue-engineering/polyethylene-glycol-and-synthetic-peg-derivatives [Accessed 22 Mar. 2022].

**[9]** Litowczenko, J., Woźniak-Budych, M.J., Staszak, K., Wieszczycka, K., Jurga, S. and Tylkowski, B. (2021). Milestones and current achievements in development of multifunctional bioscaffolds for medical application. *Bioactive Materials*, 6(8), pp.2412–2438.

**[10]** Selden, C. and Fuller, B. (2018). Role of Bioreactor Technology in Tissue Engineering for Clinical Use and Therapeutic Target Design. *Bioengineering*, 5(2), p.32.

**[11]** Firipis, K., Nisbet, D.R., Franks, S.J., Kapsa, R.M.I., Pirogova, E., Williams, R.J. and Quigley, A. (2021). Enhancing Peptide Biomaterials for Biofabrication. *Polymers*, 13(16), p.2590.

**[12]** Ngadimin, K.D., Stokes, A., Gentile, P. and Ferreira, A.M. (2021). Biomimetic hydrogels designed for cartilage tissue engineering. *Biomaterials Science*, 9(12), pp.4246–4259.

**[13]** Zhu, J. (2010). Bioactive modification of poly(ethylene glycol) hydrogels for tissue engineering. *Biomaterials*, 31(17), pp.4639–4656.

**[14]** Frantz, C., Stewart, K.M. and Weaver, V.M. (2010). The extracellular matrix at a glance. *Journal of Cell Science*, 123(24), pp.4195–4200.

**[15]** Lambert, C.R., Nijsure, D., Huynh, V. and Wylie, R.G. (2018). Hydrogels with reversible chemical environments for *in vitro* cell culture. *Biomedical Materials*, 13(4).

**[16]** Badylak, S.F. (2005). Regenerative medicine and developmental biology: The role of the extracellular matrix. *The Anatomical Record Part B: The New Anatomist*, 287B(1), pp.36–41.

**[17]** Wen, J.H., Vincent, L.G., Fuhrmann, A., Choi, Y.S., Hribar, K.C., Taylor-Weiner, H., Chen, S. and Engler, A.J. (2014). Interplay of matrix stiffness and protein tethering in stem cell differentiation. *Nature Materials*, 13(10), pp.979–987.

**[18]** Lee, J.P., Kassianidou, E., MacDonald, J.I., Francis, M.B. and Kumar, S. (2016). N-terminal specific conjugation of extracellular matrix proteins to 2-pyridinecarboxaldehyde functionalized polyacrylamide hydrogels. *Biomaterials*, 102, pp.268–276.

**[19]** Caliari, S.R. and Burdick, J.A. (2016). A practical guide to hydrogels for cell culture. *Nature Methods*, 13(5), pp.405–414.

**[20]** Li, X., Sun, Q., Li, Q., Kawazoe, N. and Chen, G. (2018). Functional Hydrogels With Tunable Structures and Properties for Tissue Engineering Applications. *Frontiers in Chemistry*, 6.

**[21]** Spicer, C.D. (2019b). Hydrogel scaffolds for tissue engineering: the importance of polymer choice. *Polymer Chemistry*, 11, pp.184–219.

**[22]** Madl, C.M., LeSavage, B.L., Dewi, R.E., Dinh, C.B., Stowers, R.S., Khariton, M., Lampe, K.J., Nguyen, D., Chaudhuri, O., Enejder, A. and Heilshorn, S.C. (2017). Maintenance of neural progenitor cell stemness in 3D hydrogels requires matrix remodelling. *Nature Materials*, 16(12), pp.1233–1242.

**[23]** Kim, J., Kong, Y.P., Niedzielski, S.M., Singh, R.K., Putnam, A.J. and Shikanov, A. (2016). Characterization of the crosslinking kinetics of multi-arm poly(ethylene glycol) hydrogels formed via Michael-type addition. *Soft Matter*, 12(7), pp.2076–2085.

**[24]** Spicer, C.D., Pashuck, E.T. and Stevens, M.M. (2018). Achieving Controlled Biomolecule–Biomaterial Conjugation. *Chemical Reviews*, 118(16), pp.7702–7743.

**[25]** Chen, Y.-X., Triola, G. and Waldmann, H. (2011). Bioorthogonal Chemistry for Site-Specific Labeling and Surface Immobilization of Proteins. *Accounts of Chemical Research*, 44(9), pp.762–773.

**[26]** Koo, B., Dolan, N.S., Wucherer, K., Munch, H.K. and Francis, M.B. (2019). Site-Selective Protein Immobilization on Polymeric Supports through N-Terminal Imidazolidinone Formation. *Biomacromolecules*, 20(10), pp.3933–3939.

**[27]** Fisher, S.A., Baker, A.E.G. and Shoichet, M.S. (2017). Designing Peptide and Protein Modified Hydrogels: Selecting the Optimal Conjugation Strategy. *Journal of the American Chemical Society*, 139(22), pp.7416–7427.

**[28]** Reddy, N.C., Kumar, M., Molla, R. and Rai, V. (2020). Chemical methods for modification of proteins. *Organic & Biomolecular Chemistry*, 18(25), pp.4669–4691.

**[29]** MacDonald, J.I., Munch, H.K., Moore, T. and Francis, M.B. (2015). One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. *Nature Chemical Biology*, 11(5), pp.326–331.

**[30]** Maza, J.C., Ramsey, A.V., Mehare, M., Krska, S.W., Parish, C.A. and Francis, M.B. (2020). Secondary modification of oxidatively-modified proline N-termini for the construction of complex bioconjugates. *Organic & Biomolecular Chemistry*, 18(10), pp.1881–1885.

**[31]** Sangsuwan, R., Tachachartvanich, P. and Francis, M.B. (2019). Cytosolic Delivery of Proteins Using Amphiphilic Polymers with 2-Pyridinecarboxaldehyde Groups for Site-Selective Attachment. *Journal of the American Chemical Society*, 141(6), pp.2376–2383.

**[32]** Gilmore, J.M., Scheck, R.A., Esser-Kahn, A.P., Joshi, N.S. and Francis, M.B. (2006). N-Terminal Protein Modification through a Biomimetic Transamination Reaction. *Angewandte Chemie International Edition*, 45(32), pp.5307–5311.

**[33]** Koniev, O. and Wagner, A. (2015). Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem. Soc. Rev.*, 44(15), pp.5495–5551.

**[34]** Smith, G.P. (2006). Kinetics of Amine Modification of Proteins. *Bioconjugate Chemistry*, 17(2), pp.501–506.

**[35]** Borch, R.F., Bernstein, M.D. and Durst, H.D. (1971). Cyanohydridoborate anion as a selective reducing agent. *Journal of the American Chemical Society*, 93(12), pp.2897–2904.

**[36]** Rosen, C.B. and Francis, M.B. (2017). Targeting the N terminus for site-selective protein modification. *Nature Chemical Biology*, 13(7), pp.697–705.

**[37]** Jiang, H., Chen, W., Wang, J. and Zhang, R. (2022). Selective N-terminal modification of peptides and proteins: Recent progresses and applications. *Chinese Chemical Letters*, 33(1), pp.80–88.

**[38]** Gaertner, H.F. and Offord, R.E. (1996). Site-Specific Attachment of Functionalized Poly(ethylene glycol) to the Amino Terminus of Proteins. *Bioconjugate Chemistry*, 7(1), pp.38–44.

**[39]** Dawson, P., Muir, T., Clark-Lewis, I. and Kent, S. (1994). Synthesis of proteins by native chemical ligation. *Science*, 266(5186), pp.776–779.

**[40]** Tam, J.P., Yu, Q. and Miao, Z. (1999). Orthogonal ligation strategies for peptide and protein. *Biopolymers*, 51(5), pp.311–332.

**[41]** Deng, J.-R., Lai, N.C.-H., Kung, K.K.-Y., Yang, B., Chung, S.-F., Leung, A.S.-L., Choi, M.-C., Leung, Y.-C. and Wong, M.-K. (2020). N-Terminal selective modification of peptides and proteins using 2-ethynylbenzaldehydes. *Communications Chemistry*, 3(1).

**[42]** Sun, J., Liu, X., Guo, J., Zhao, W. and Gao, W. (2020). Pyridine-2,6-dicarboxaldehyde-Enabled N-Terminal In Situ Growth of Polymer–Interferon α Conjugates with Significantly Improved Pharmacokinetics and In Vivo Bioactivity. *ACS Applied Materials & Interfaces*, 13(1), pp.88–96.

**[43]** Wei, D. and Li, B. (Apalene Technology Co., LTD.) (2019). *Linear alpha-olefin catalysts and preparation and use thereof*. [online] Available at:

https://worldwide.espacenet.com/publicationDetails/originalDocument?FT=D&date=20191 015&DB=&locale=en_EP&CC=US&NR=10441946B2&KC=B2&ND=1 [Accessed 1 Nov. 2019].

**[44]** Dubé, D. and Scholte, A.A. (1999). Reductive N-alkylation of amides, carbamates and ureas. *Tetrahedron Letters*, 40(12), pp.2295–2298.

**[46]** Creative PEGWorks | PEG Products Leader. (n.d.). *PEGylation Chemistry*. [online] Available at: https://creativepegworks.com/knowledge-center/pegylation-chemistry [Accessed 28 Mar. 2022].

**[47]** Ghosh, A.K. and Brindisi, M. (2015). Organic Carbamates in Drug Design and Medicinal Chemistry. *Journal of Medicinal Chemistry*, 58(7), pp.2895–2940.

**[48]** Moriyama, K., Minamihata, K., Wakabayashi, R., Goto, M. and Kamiya, N. (2013). Enzymatic preparation of streptavidin-immobilized hydrogel using a phenolated linear poly(ethylene glycol). *Biochemical Engineering Journal*, 76, pp.37–42.

**[49]** Lutolf, M.P. and Hubbell, J.A. (2003). Synthesis and Physicochemical Characterization of End-Linked Poly(ethylene glycol)-co-peptide Hydrogels Formed by Michael-Type Addition. *Biomacromolecules*, 4(3), pp.713–722.

**[50]** Zustiak, S.P. and Leach, J.B. (2010). Hydrolytically Degradable Poly(Ethylene Glycol) Hydrogel Scaffolds with Tunable Degradation and Mechanical Properties. *Biomacromolecules*, 11(5), pp.1348–1357.

**[51]** Morpurgo, M., Veronese, F.M., Kachensky, D. and Harris, J.M. (1996). Preparation and Characterization of Poly(ethylene glycol) Vinyl Sulfone. *Bioconjugate Chemistry*, 7(3), pp.363–368.

**[52]** Buwalda, S.J., Dijkstra, P.J. and Feijen, J. (2012). In Situ Forming Poly(ethylene glycol)-Poly(L-lactide) Hydrogels via Michael Addition: Mechanical Properties, Degradation, and Protein Release. *Macromolecular Chemistry and Physics*, 213(7), pp.766–775.

**[53]** Hoult, J.R.S. and Payá, M. (1996). Pharmacological and biochemical actions of simple coumarins: Natural products with therapeutic potential. *General Pharmacology: The Vascular System*, 27(4), pp.713–722.

**[54]** Warrier, S. and Kharkar, P.S. (2018). Highly selective on-off fluorescence recognition of Fe3+ based on a coumarin derivative and its application in live-cell imaging. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 188, pp.659–665.

**[55]** Griswold, A.R., Cifani, P., Rao, S.D., Axelrod, A.J., Miele, M.M., Hendrickson, R.C., Kentsis, A. and Bachovchin, D.A. (2019). A Chemical Strategy for Protease Substrate Profiling. *Cell Chemical Biology*, 26(6), pp.901-907.e6.

**[56]** Younes, A.H., Zhang, L., Clark, R.J., Davidson, M.W. and Zhu, L. (2010). Electronic structural dependence of the photophysical properties of fluorescent heteroditopic ligands – implications in designing molecular fluorescent indicators. *Organic & Biomolecular Chemistry*, 8(23), p.5431.

**[57]** Paolucci, G., Zanella, A., Bortoluzzi, M., Sostero, S., Longo, P. and Napoli, M. (2007). New constrained geometry catalysts-type yttrium, samarium and neodymium derivatives in olefin polymerization. *Journal of Molecular Catalysis A: Chemical*, 272(1-2), pp.258–264.

**[58]** Sedlák, M., Pejchal, V., Imramovský, A. and Hanusek, J. (2013). A Safe and Efficient Route for the Preparation of Nitro-, Dinitro-, and Trinitrophenyl Chloroformates. *Synthetic Communications*, 43(17), pp.2365–2369.

**[59]** McAllister, M.I., Boulho, C., McMillan, L., Gilpin, L.F., Wiedbrauk, S., Brennan, C. and Lennon, D. (2018). The production of tyramine *via* the selective hydrogenation of 4-hydroxybenzyl cyanide over a carbon-supported palladium catalyst. *RSC Advances*, 8(51), pp.29392–29399.

**[60]** Wang, Z.J., Jackson, W.R. and Robinson, A.J. (2015). A simple and practical preparation of an efficient water soluble olefin metathesis catalyst. *Green Chemistry*, 17(6), pp.3407–3414.

**[61]** Tran, F., Odell, A., Ward, G. and Westwood, N. (2013). A Modular Approach to Triazole-Containing Chemical Inducers of Dimerisation for Yeast Three-Hybrid Screening. *Molecules*, 18(9), pp.11639–11657.

**[62]** Aubineau, T. and Cossy, J. (2013). Chemoselective alkynylation of N-sulfonylamides versus amides and carbamates – Synthesis of tetrahydropyrazines. *Chemical Communications*, 49(32), p.3303.

**[63]** Pastuszak, J., Gardner, J.H., Singh, J. and Rich, D.H. (1982). Improved synthesis of [4-alanine]chlamydocin: cyclization studies of tetrapeptides containing five .alpha. substituents. *The Journal of Organic Chemistry*, 47(15), pp.2982–2987.

**[64]** Khan, D., Mukhtar, S., Alsharif, M.A., Alahmdi, M.I. and Ahmed, N. (2017). PhI(OAc) 2 mediated an efficient Knoevenagel reaction and their synthetic application for coumarin derivatives. *Tetrahedron Letters*, 58(32), pp.3183–3187.

**[65]** Tian, M., Liu, C., Dong, B., Zuo, Y. and Lin, W. (2019). A dual-site controlled ratiometric probe revealing the simultaneous down-regulation of pH in lysosomes and cytoplasm during autophagy. *Chemical Communications*, 55(70), pp.10440–10443.

**[66]** Amara, N., Foe, I.T., Onguka, O., Garland, M. and Bogyo, M. (2019). Synthetic Fluorogenic Peptides Reveal Dynamic Substrate Specificity of Depalmitoylases. *Cell Chemical Biology*, 26(1), pp.35-47.e7.