

MASTER OF SCIENCE BY RESEARCH

THE UNIVERSITY OF KENT

THE SCHOOL OF PHYSICAL SCIENCES

The Templatation of Amyloid by DNA and the Modulation of its Suprastructure

HANNAH CRAWFORD

SEPTEMBER 2019

ACKNOWLEDGEMENTS

There are many people I wish to acknowledge for their support during this research. I would firstly like to thank my supervisor, Dr Christopher Serpell, for his guidance during the whole project and for the opportunity to undertake this work. I am also appreciative to Dr Wei-Feng Xue, for his contribution and involvement throughout.

A special thank you to Liisa Lutter, from the Xue lab in the School of Biosciences, for expressing and purifying the protein and for our work together on the conjugations. I am grateful to Akiko Sato, working in the Serpell lab, for synthesising the Waltz peptide and for her collaboration on the conjugation work. Kevin Howland, from the School of Biosciences, kindly provided the mass spectrometry data, which has been much appreciated. The School of Physical Sciences at the University of Kent have provided the funding for this Masters, to which I am indebted.

In addition, I would also like to acknowledge and thank my friends, both within and out of the Serpell group, for their encouragement and assistance. Finally, I wish to thank my family for their support and reassurance throughout this project and for always believing in me.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	2
ABSTRACT	5
LIST OF ABBREVIATIONS	7
CHAPTER 1: INTRODUCTION	10
1.1 Amyloid	10
1.2 Biological Importance	10
1.3 Structure and Formation of Amyloid Fibrils	13
1.4 Mechanisms of Amyloid Growth	15
1.5 Analysis of Fibril Structure	17
1.6 Imaging Fibrils	18
1.7 Amyloid as a Functional Material	19
1.8 Biological Functions of Amyloid	21
1.9 Control of Amyloid Assembly	23
1.10 Control of Assembly with DNA	24
1.11 Control of Amyloid Assembly with DNA	26
1.12 Project Aims	27
CHAPTER 2: MALEIMIDE CHEMICAL LINKER SYNTHESIS	31
2.1 Introduction	31
2.2 General Considerations	32
2.3 Experimental	34
2.3.1 Synthesis of Linkers	34
2.3.2 Conjugation of Linker and Protein	39
2.4 Results and Discussion	40
2.4.1 Synthesis of Linkers	40

2.4.2 Conjugation of Linker and Protein	47
2.5 Conclusion	50
CHAPTER 3: CONJUGATION OF MALEIMIDE LINKER AND WALTZ PEPTIDE	51
3.1 Introduction	51
3.2 Experimental	52
3.3 Results and Discussion	52
3.4 Conclusion	55
CHAPTER 4: PHOSPHORAMIDITE MALEIMIDE CHEMICAL LINKER SYNTHESIS ...	56
4.1 Introduction	56
4.2 Experimental	58
4.2.1 Synthesis of DNA	58
4.2.2 Synthesis of Linkers	59
4.2.3 Conjugation of Linker and DNA.....	63
4.2.4 Conjugation of DNA-Linker and Protein	64
4.3 Results and Discussion	65
4.3.1 Synthesis of DNA	65
4.3.2 Synthesis of Linkers	66
4.3.3 Conjugation of Linker and DNA.....	70
4.3.4 Conjugation of DNA-Linker and Protein	76
4.4 Conclusion	79
CHAPTER 5: CONCLUSIONS, FURTHER WORK AND REFERENCES	81
5.1 Conclusion	81
5.2 Future Work	83
5.3 References	84

ABSTRACT

Amyloidogenic proteins are an important part of the constitution of our bodies; such examples, in humans alone, include their roles in peptide hormone storage, melanin biosynthesis and neural growth and repair, to name just a few.¹ A key step in the growth and formation of amyloid fibrils is elongation, the addition of protein monomers to the fibril ends. However, it is currently thought that this process can result in the formation of cytotoxic oligomers which, in vivo, can lead to neurodegenerative diseases. This research focused on synthesising two chemical linkers in order to conjugate a known amyloid-forming protein of interest with a strand of deoxyribose nucleic acid (DNA). As a result, it was hoped the effect of DNA on amyloid fibrils could be investigated, with the extent to which amyloid formation could be influenced and controlled explored, and thus the toxicity of the oligomers examined. Amyloid also have many material applications and hence the potential use of the materials in which these templated amyloid could be used could also be studied.

The first linkers to be made were maleimide chemical linkers, N-(6-acetamidopyridin-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)acetamide (**linker 1A**) and N-6-acetamidopyridin-yl-6-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)hexanamide (**linker 1B**). Synthesis of **linker 1A** proved challenging, but **linker 1B** led to more positive results. After synthesis and characterisation, the conjugation with the protein and DNA was carried out; the formation of the conjugates was validated using ultra-violet visible spectrometry, mass spectrometry, gel electrophoresis and amyloid formation assays using a fluorescent probe.

The second linkers produced were also maleimide linkers, 6-maleimidohexanoic acid N-hydroxysuccinimide ester (**linker 2A**) and 2-cyanoethyl 2-((3aR,4S,7aS)-4,7-dimethyl-1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-4,7-epoxyisoindol-2yl)ethyl)diospropylphosphoramidite (**linker 2B**). Conjugation to the DNA followed, with **linker 2A** conjugated successfully using a solution-based method, and thus the addition of the protein to the linker-DNA compound; likewise to **linker 1B**, the conjugation was monitored using ultra-violet-visible spectrometry,

mass spectrometry, gel electrophoresis and analysis of 96 well microplates. Conversely, the hand coupling of linker 2B to DNA attached to a solid-support bead was unavailing.

LIST OF ABBREVIATIONS

%	=	Percentage
°C	=	Degrees Celsius
µL	=	Microlitre
µM	=	Micromolar
Ac	=	Acetyl
AFM	=	Atomic force microscopy
C	=	Cysteine
CDCl ₃	=	Chloroform
Da	=	Daltons
DMAP	=	4-dimethylamino pyridine
DCM	=	Dichloromethane
DMF	=	Dimethylformamide
DMSO	=	Dimethyl sulfoxide
DNA	=	Deoxyribose nucleic acid
DTT	=	Dithiothreitol
EDC.HCl	=	N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride
ESI	=	Electrospray
Et ₃ N	=	Triethylamine
EtOAc	=	Ethyl acetate
F	=	Phenylalanine
G	=	Gravity
g	=	Grams
G	=	Glycine
GdnHCl	=	Guanidine hydrochloride
H	=	Histidine
HATU	=	Hexafluorophosphate azabenzotriazole tetramethyl uranium (1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazole[4,5- b]pyridinium-3-oxide-hexafluorophosphate)

HFIP	=	Hexafluoroisopropanol
HPLC	=	High performance liquid chromatography
I	=	Isoleucine
K	=	Lysine
LC-MS	=	Liquid chromatography mass spectrometry
LCys	=	Linker cysteine
M	=	Methionine
MALDI-TOF MS	=	Matrix-assisted laser desorption ionisation time-of-flight
mg	=	Milligrams
MHz	=	Megahertz
mL	=	Millilitre
mM	=	Millimolar
Mol equiv	=	Molar equivalent
MS	=	Mass spectrometry
MW	=	Molecular weight
m/z	=	Mass to charge ratio
N	=	Asparagine
NaCl	=	Sodium chloride
N-Boc Gly	=	N-(<i>tert</i> -butoxycarbonyl)glycine
ppm	=	Parts per million
R	=	Arginine
S	=	Serine
TCEP	=	Tris(2-carboxyethyl)phosphine
TFA	=	Trifluoroacetic acid
THF	=	Tetrahydrofuran
ThT	=	Thioflavin T
TIPS	=	Triisopropylsilane
TLC	=	Thin layer chromatography
Tris-HCl	=	Tris(hydroxymethyl)aminomethane hydrochloride
UV-Vis	=	Ultraviolet visible spectroscopy
V	=	Valine

WT = Wild type

Y = Tyrosine

CHAPTER 1: INTRODUCTION

1.1 Amyloid

Proteins are amyloidogenic if they are able to misfold and subsequently self-assemble, resulting in highly ordered forms of the proteins known as amyloid fibrils. The name 'amyloid' is derived from the Latin of starch, which arose from a reaction whereby iodine was used to stain brain tissue that had abnormal macroscopic appearances; the tissue stained blue on reaction with iodine and violet on addition of sulphuric acid, which was characteristic of starch. Later work, by Friedreich and Kekule in 1859, showed that amyloid actually did not contain carbohydrate as previously thought, due to its high nitrogen content, but was comprised of protein. Amyloid has since been studied as a protein and subsequently identified as a class of proteins that are able to undergo structural changes, resulting in fibril formation.²

1.2 Biological Importance

Since the discovery of amyloid formation, it has been associated with many medical conditions and diseases. It is now thought there are approximately 50 conditions which are linked to misfolding proteins into amyloid fibrils, including Alzheimer's disease, Huntington's disease and Type II Diabetes. The diseases associated with amyloid fibrils are suspected to be caused either by aggregation of fibrils in the brain, resulting in neurodegenerative conditions, localised aggregation in a single tissue type or aggregation in multiple tissues.³

For many years it was thought that the amyloid fibrils themselves were toxic and thus causing these conditions, but it has since been suggested that it is the smaller oligomers and small intermediate fibrils that are produced during the formation, shown in figure 1, that causes the toxic properties. This is a result of the loss of protein function once the fibril has formed and could also result from the gain of a toxic function.⁴

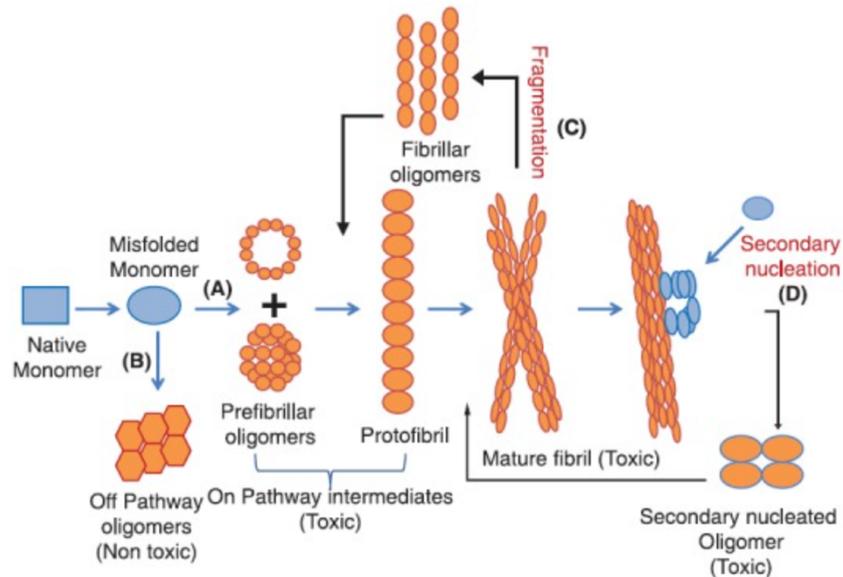


Figure 1: A diagram showing the formation of amyloid protein fibrils.⁵
 (a) Misfolded protein monomers form protofibrils and mature fibrils (b) Misfolded monomers can also form off-pathway oligomers (c) Fragmentation of mature fibrils produces toxic fibrillar oligomers, before reforming mature fibrils (d) Oligomerisation and secondary nucleation of mature fibrils yields more toxic oligomers from the monomers

Misfolded proteins are likely to be toxic as the unusual conformation will display some functional groups and hydrophobic surfaces that are usually contained within the folded protein. These groups could cause interactions that do not usually occur with other cell components, leading to cells malfunctioning. Smaller oligomers would also have a higher surface area, meaning more functional groups and side groups are exposed and free to interact, thus causing a higher toxicity relative to complete amyloid fibrils.⁶ The smaller oligomers formed in the process are also able to disrupt membranes, which can lead to disturbance of the transport of proteins and ion imbalances; a linear correlation has been reported between the level of Ca^{2+} across membranes when oligomers of the $\text{A}\beta_{42}$ protein bind to a cell membrane. This has resulted in cellular damage and ultimately could result in cellular death.⁷ Due to the instability and varying molecular weight and structure of toxic oligomers, their structure has not yet been determined in a high enough resolution to enable the molecular basis of the toxic properties to be resolved.

Although there are natural defences within cells that attempt to inhibit protein misfolding, these mechanisms can be overrun due to factors such as mutations within the cell, which could cause an increase in the rate of aggregation formation. The resulting accumulation of aggregates in affected tissues are predominant characteristics of these diseases; for example, extreme pain occurs in joints due to amyloid build up and memory impairment due to a build-up in the brain tissues.⁸

This area of research has been of much interest recently, but there are still many fundamentals that are yet to be fully understood. It is still unclear what causes the aggregation of amyloid to begin with and also the mechanism by which amyloid aggregates are recognised by other cells and their components.

Clinical trials have been carried out on Alzheimer's patients whereby drugs were used in order to stop amyloid formation. One issue to be faced is that drug companies need to determine that the aggregation of amyloid fibrils was reduced or stopped, whilst simultaneously the quality of life of the patients who took the drugs improved, particularly their memory and judgement. The testing of potential drugs for Alzheimer's is also a slow process as Alzheimer's is a progressive disease that causes the health and symptoms of a patient to deteriorate with time; the outcome of potential drugs that are started in the early onset of Alzheimer's would take a while to show, if at all.⁹

Semagacestat, shown in figure 2, is an example of one drug that has undergone clinical trials. The drug worked by inhibiting the enzyme that is responsible for the cleavage of amyloid precursor protein to pathogenic amyloid- β . However, the drug failed to show any cognitive improvement in patients treated with the drug compared to the placebo control and so the trial was stopped.¹⁰ Conversely, BAN2401 is a monoclonal antibody that is used to detect amyloid- β fibrils in the brains of Alzheimer sufferers. In over 800 patients treated with the antibody at five different doses, amyloid plaque levels were decreased compared to the control group. However, two separate trials of the cognitive function of these patients had contrasting results; one trial found a 30% reduction in slowing cognitive decline when

the highest dose of antibody was administered, whilst another trial failed to show a statistically significant difference between the patients and control group. Further work is thus needed into the effectiveness of the treatment.¹¹

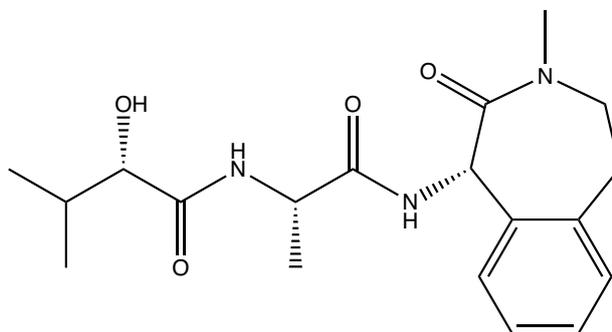


Figure 2: The structure of Semagacestat¹¹

These trials also highlight the issues faced with finding an effective therapeutic treatment. The most common explanation for trial failure is that the drugs are administered at the wrong time during the development of Alzheimer disease. It has been shown, by using positron emission tomography, that plaques form long before the symptoms of cognitive decline show and therefore amyloid- β levels may have plateaued by the time the drugs are being administered. Some trials are being carried out on patients who do not show any symptoms, whilst others are being carried out at different stages along the development of Alzheimer's in order to establish the most critical point at which drugs should be administered for the most effective results.¹²

1.3 Structure and Formation of Amyloid Fibrils

Although proteins have a typical native folded structure, when in a solution they can adopt many different structures that have varying energy levels. Recent research has begun on intrinsically disordered proteins that favour an overall higher energy structure and also on proteins which can vary their structure depending on the specific conditions in which they are found.

The energy levels of proteins can be described by the 'folding funnel' energy landscape, shown in figure 3, which suggests that protein structures have different energy levels depending on a number of conditions¹³. Influences, such as solvent interactions, enthalpic contributions from non-covalent bonds and shielding of apolar protein surfaces¹⁴, are all factors in determining the free energy and entropy of the protein. It is only when the high entropy of unfolded proteins is overcome by these interactions that the typical native folded structure forms¹⁵. It is possible though that whilst forming this lower energy structure the protein can misfold, resulting in an unfolded or half-folded state, represented on the funnel landscape as a local minimum¹³. Fibrils are consequently formed when many proteins interact with each other in a regular, periodic arrangement; the interaction of 'misfolded' proteins leads to disease related amyloid, whilst intrinsically disordered proteins can also form amyloid.¹⁶

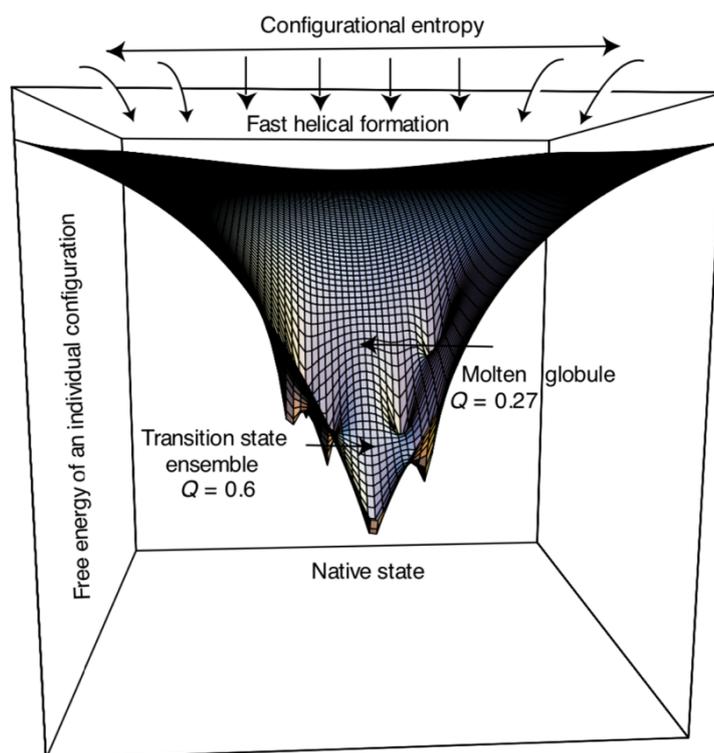


Figure 3: The folding funnel energy landscape of proteins¹³

The insoluble fibrils formed are straight, mainly unbranched structures that have a diameter between 6 - 12 nanometres and a length that ranges from a few nanometres to several micrometres, although this is variable depending on the protein monomer. The highly

ordered characteristic of amyloid originates from a core structure within the fibril that is composed of cross β -strands, a long extended chain of polypeptides. Hydrogen bonding between individual cross β -strands runs parallel to the fibril axis and creates a long β -sheet, which hydrogen bonds to other β -sheets to form the fibril core in protofilaments. The protofilaments are then able to twist around each other, forming fibrils.¹⁷ Figure 4 illustrates the structure of amyloid fibrils.

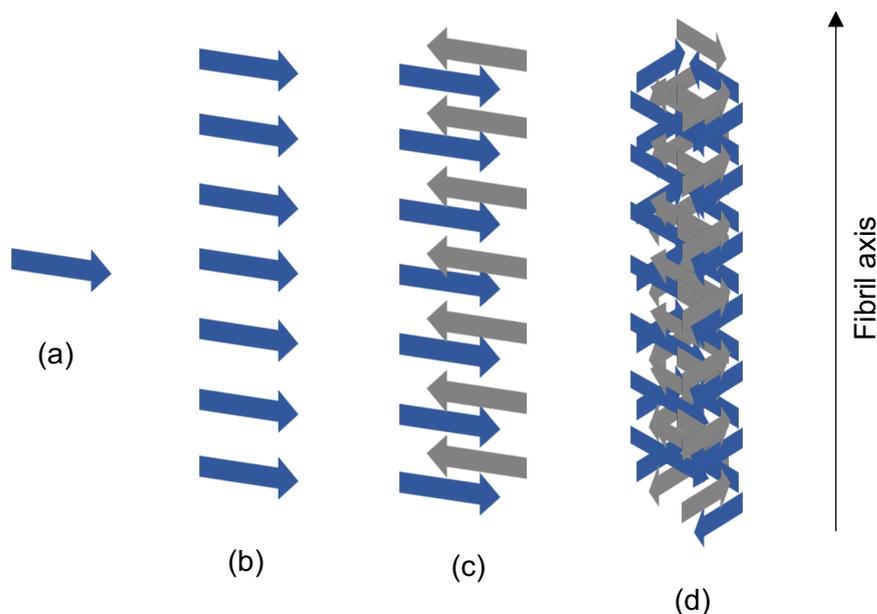


Figure 4: The structure of fibrils
(a) β -strand; (b) β -sheet; (c) protofilaments; (d) fibril core

1.4 Mechanisms of Amyloid Growth

Amyloid fibrils are generally considered to form by a nucleation growth mechanism.¹⁸ Either unfolded or partially folded protein monomers bind together in an ordered manner to form oligomers of varying structure and length. Oligomers can then further assemble to produce longer chains which form the precursors of amyloid fibrils, as shown in figure 5. During this process a high energy nucleus forms before polymerisation occurs rapidly.¹⁹ Although the conventional nucleation growth mechanism suggests that the growth should not increase exponentially, it is secondary processes, such as secondary nucleation and fibril fragmentation, that occur and provide exponential growth. The growth of amyloid fibrils

after the nucleation phase is exponential and there is a lag time in which the individual monomers aggregate.²⁰

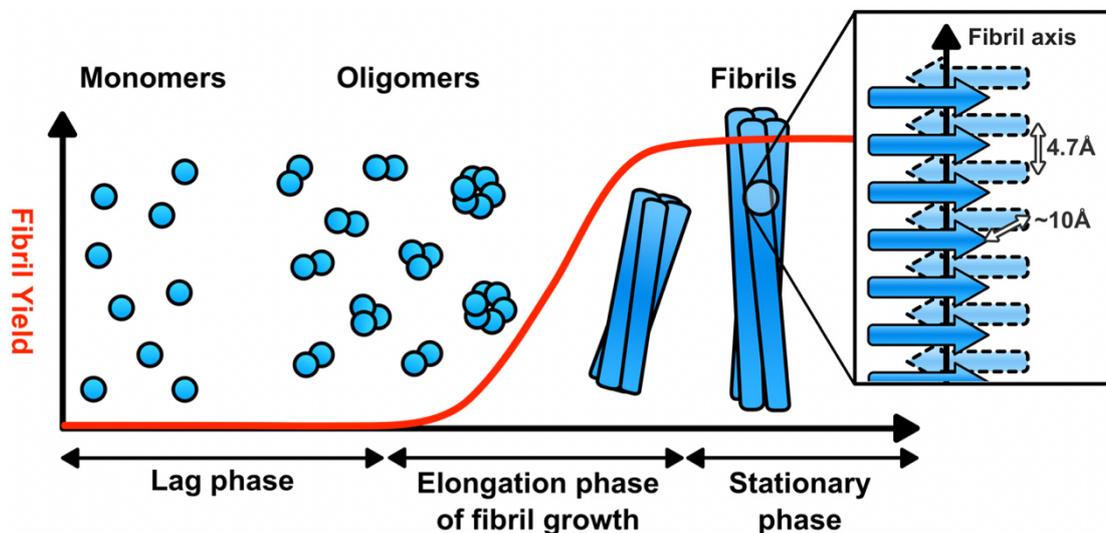


Figure 5: The growth of amyloid fibrils²⁰

During the assembly of the protein monomers, the monomer adopts a β -strand conformation; this is illustrated in figure 6.²¹ However, fibrils are able to fragment, producing new fibril ends to which monomers can bind; this further increases the growth rate.²²

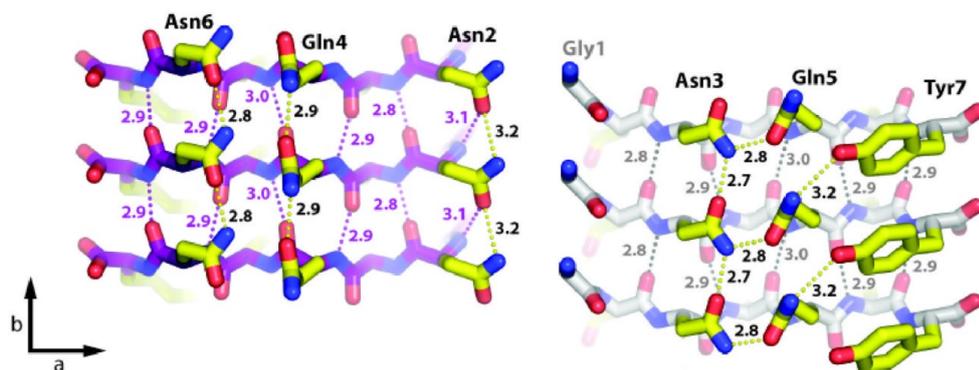


Figure 6: A side view of 3 β -strands of GNNQQNY protein, arranged to form a β -sheet arrangement. GNNQQNY is a 7-residue section of the fibril-forming N-terminus of the Sup35 protein²¹

Left: The β -strands viewed from the centre of the dry interface
 Right: The β -strands viewed from the middle of the wet interface

Secondary nucleation can also occur during fibril growth. This is a process where oligomers form on the surface of a fibril that has already been created, potentially leading to an autocatalytic amplification of the process. It is thought this secondary nucleation leads to formation of toxic oligomers and so could provide an area of research for medical advances.²³

1.5 Analysis of Fibril Structure

Amyloid fibrils are insoluble and thus there are limitations on the techniques that can be used in their analysis. Solution NMR does not yield any structure information, but more recently solid state NMR, has been used to further determine the atomic level structure.²⁴

X-ray fibre diffraction patterns of fibrils show the distances between the individual stacked β -strands and also between the β -sheets. The individual stacked β -strands are separated by a distance of approximately 0.47 nanometres, whilst the distance between the stacked β -sheets is between 0.6 and 1.1 nanometres²⁵. Fibrils however are heterogeneous polymers, formed by the nucleated growth mechanism in which growth is rapid along one dimension; the cross- β strands resemble a one dimensional crystal, with a limited capability to produce a three dimensional lattice. Resultantly x-ray diffraction is limited in the information it provides as this technique cannot fully resolve their structure.²⁶

Circular dichroism determines the secondary structure of the protein monomer in the amyloid structure by using polarized light in the UV region. It can also be used to quantify the amount of α -helixes, β -sheets and random coils found within the protein structure. The cross- β sheet structure of fibrils allows certain dye molecules to bind. Congo red can be used under cross-polarised light, producing a green birefringence.²⁷ Thioflavin T is a π -conjugated molecule that binds to the cross- β sheets; this changes the absorption and emission properties of the dye allowing the fibril structure to be analysed. As fibrils grow by the nucleated growth mechanism, there is a lag phase that is observed before the growth

phase. This lag phase allows the growth of the fibrils to also be monitored as the dye only binds once the cross- β sheets have been formed.²⁸

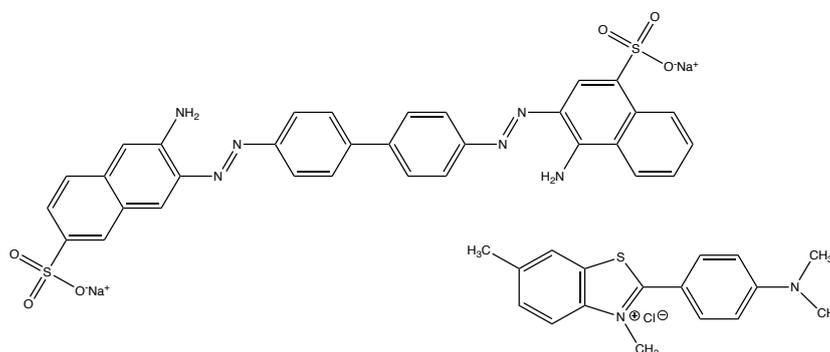


Figure 7: The structure of the dye molecules used
Left: Congo red dye
Right: Thioflavin T dye

Cryo-electron microscopy has recently shown that the structure of amyloid fibrils is comparable to steric zippers, observed in peptide assemblies. Steric zippers are short peptides that consist of self-complementary β -sheets and as a result there are some differences between the structure of steric zippers and the amyloid fibril core. Amyloid is usually composed of twisted fibrils with packing defects, whilst steric zippers are untwisted β -sheets that are packed together densely. Thus, the interactions within peptide microcrystals and amyloid fibrils are analogous.²⁹

1.6 Imaging Fibrils

Transmission electron microscopy enables individual fibrils to be studied. Negatively stained species can be visualised in a dried state under high vacuum, allowing the framework of fibrils and growing fibrils to be examined. It enables features as small as 1 – 2 nanometres to be seen, providing images of protein aggregates and fibrils and thus allowing the formation of amyloid fibrils from proteins to be followed.³⁰

Atomic force microscopy allows the study of the structural information of the amyloid fibrils in three-dimensions, including the size, helicity, dispersity and uniformity of the fibrils. However, more advanced atomic force microscopy measurements allow the study of the biophysical properties of amyloid on the nanoscale, particularly on the mechanism of amyloid formation and on the characteristic of the protein aggregated. This is an advantage over both circular dichroism and infrared spectroscopy as these techniques are unable to determine the properties of single amyloid fibril. Conversely, without these techniques, it would not be known how the fibrils interact and the structures that they form could not be determined and thus, the combination of all techniques is favourable. Recent developments on atomic force microscopy methods have also revealed the mechanical properties on the nanoscale.³¹

1.7 Amyloid as a Functional Material

As a result of the continuous hydrogen bonding in the core, amyloid fibrils are very strong, with high thermal and chemical stability. They also exhibit good mechanical and electrochemical properties along with biocompatibility. This can prove advantageous in applications of energy, optics and healthcare.

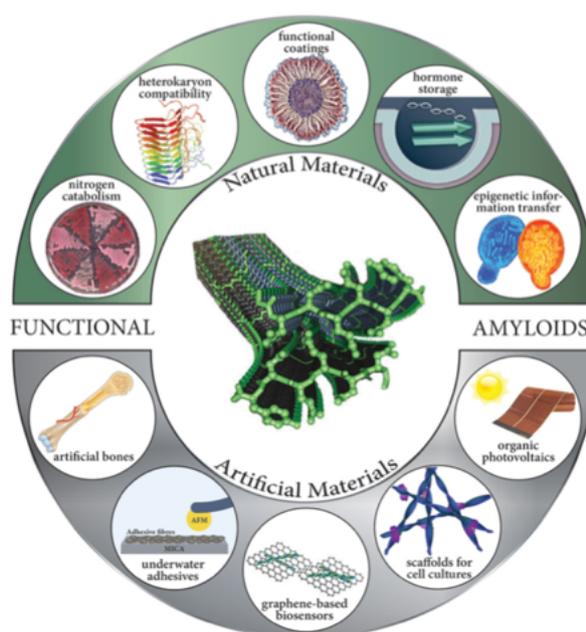


Figure 8: Potential uses of amyloid fibrils²⁸

Self-assembled peptide nanomaterials, that contain the cross- β sheet core, have recently been used to build functional materials with a range of applications. Photosynthetic nanostructures have been made using self-assembling peptides with photosensitive and photo catalysts in order to produce light harvesting peptide hydrogels, which could help advance the research in solar energy.³² Cathode materials can also be made using certain self-assembling peptides as a template. The reaction produced carbon-coated FePO_4 nanotubes that worked as a cathode in lithium-ion batteries.³³

Perylene bisimides are able to self-assemble into photoconductive one-dimensional structures; these structures have random orientation of the bisimides, which restricts their use in conductive wires. However, research has shown that by using magnetic fields, the self-assembly can be controlled to produce aligned, directionally dependent thin films. The strength of the magnetic field can be varied to change the degree of alignment; a stronger field results in more alignment whilst samples with a lower concentration of bisimides requires a weaker field for the same alignment effect to be produced. Studies showed that when the magnetic field was removed, the samples returned to their randomly orientated structures, but the alignment could be preserved if the sample was dried under the effect of the magnetic field. This highly ordered structure could potentially be used to create photoconductive films from an organic semi-conductor. However, it is not yet known the strength of magnetic field required in order to produce the directional dependence required within the films.³⁴

Self-assembling proteins can produce structures, with sizes ranging from the nanoscale to the macroscale. Self-assembling dipeptides can be modified to be incorporated into composite materials, forming supramolecular co-polymers. The resulting co-polymers have shown a potential to be used in nanostructures; they are able to form molecular motors with a metal-organic framework, deliver therapeutic agents and form piezoelectric-based sensors. Their potential in energy storage devices, with the energy storage of self-assembled dipeptides as part of a larger structure, is one research topic that is currently of much interest. Vertically aligned peptide nanotubes were used to produce electrostatic

ultracapacitors with a carbon nanotube structure. The peptide nanotube ultracapacitor showed a 30 times higher level of current response when compared to a standard carbon nanotube-coated electrode and other similar structures.³⁵

Amyloid have good electronic and optical properties, with some amyloid showing piezoelectric properties, allowing templates for biosensors and biochips to be produced. Nanotubes, made of self-assembled peptides, helped produced a composite electrode that had greater electrochemical activity; this could help improve the sensitivity of biosensors.³⁶

1.8 Biological Functions of Amyloid

One benefit to healthcare could be in producing hard or soft tissue regenerative materials. Cell growth is influenced by the mechanical properties of its surroundings and so recent research has been undertaken to modify the solidity of a gel template using peptides in order to influence the type of cell produced. Hydrogels are an effective template for cell growth due to their network structure and ability to incorporate and interact with cells and bioactive material. The self-assembly of peptides has been utilised to produce a nanofibrous, bioactive hydrogel that acts as a three-dimensional scaffold for cells that need fixing to a support. The hydrogel is a rigid network containing ligands on its surface, mimicking an extracellular matrix for which the cells can grow.³⁷ Peptide self-assembly offers structural flexibility and by using peptide amphiphiles as a template, the regeneration of bone and its growth can be stimulated. Similarly, a self-assembled peptide hydrogel scaffold was used to stimulate cartilage repair.³⁸

Despite the negative effects associated with the toxicity of amyloid, it is known that amyloid can provide many benefits both naturally and synthetically. One current area of research is the molecular level mechanism that distinguishes between toxic amyloid and beneficial amyloid. It is presently thought that bacteria ensure functional amyloid are produced by

controlling the nucleation step through a specific protein, eliminating uncontrolled polymerisation.³⁹

The stability and strength of amyloid fibrils, along with their high level of chemical and biological resistance, can provide an advantage for bacteria in their biofilms. Spherical *Escherichia coli* produced curli fibrils through a tightly controlled mechanism in which co-factors ensure the fibril production is not harmful in the intracellular environment. Comparable amyloid fibrils also constitute the biofilms of *Salmonella* and *Bacillus subtilis*.⁴⁰ Amyloid also aids the survival of *Escherichia coli* strains, by protecting against bactericidal activity. The stability of amyloid is also used by silk moths, who use an amyloid fold in eggshells; the eggshells contain chorion proteins which exist in predominantly amyloid form as it provides chemical separation and thus defence of the egg from the environment.⁴¹

Epigenetics refers to an inheritable change in the gene expression of a cell, without alterations to the DNA sequences. Advances in the field of epigenetics have since been made to modify particular proteins and DNA due to external stimuli or the environment; this has resulted in changes to the characteristics displayed.⁴² A 'switch' in gene expression can be caused by a change in the usual environment of a cell due to a certain factor. Amyloid fibrils are able to bring about a change in a cell by being constituted of proteins that act as a particular class of amyloid fibrils, known as prions, in yeast. When the protein is soluble, as a non-amyloid, they are able to act as normal but when the proteins aggregate into amyloid, the specific function is 'switched off'. This has been seen in Sup35P yeast proteins, which is a translation termination factor, causing the stop of protein synthesis. When Sup35P aggregates, the termination is 'switched off', which allows translation and protein synthesis to continue.⁴³

As a result, it is thought that more artificial materials could be produced using amyloid and their self-assembling mechanism, including underwater adhesives, artificial bones and organic photovoltaics.³² There is also potentially a more natural use of amyloid, such as that in nitrogen catabolism, functional coatings and hormone storage.⁴⁴

1.9 Control of Amyloid Assembly

Neurodegenerative diseases and other medical conditions all display amyloid aggregates in affected tissues. One study showed that amyloid formation uses an amyloid template in order to bind individual amyloid monomers. By creating a synthetic amyloid template with amyloid monomers attached to an activated surface, insulin monomers were able to deposit onto the template, forming fibrils. The fibrils formed at a much faster rate without a lag time and were found, by atomic force microscopy, to be longer and thinner than those usually produced in solution. Dyes were used to further study the fibrils formed, revealing that the amyloid formed by a structural change when the insulin monomers reacted with the template. As a result, this showed that a synthetic template can be used to control the assembly of insulin monomers.⁴⁵

The aggregation of amyloid monomers into fibrils is also thought to be influenced by environmental factors, such as metal ions and pH. This was investigated by using a synthetic template of β -amyloid monomers on an activated solid surface. Atomic force microscopy measurements showed that the addition of Cu^{2+} and Zn^{2+} ions resulted in the formation of amorphous aggregates but at an accelerated rate whereas Fe^{3+} caused amyloid fibrils to form at a neutral pH. However, when mildly acidic conditions were used none of the ions stimulated amyloid fibril formation.⁴⁶

Modification of the composition of peptides is another way that the self-assembly of the peptides can be altered. The C-terminus of D-tripeptide was marginally altered, which caused an increase in the self-assembling ability of the protein and an increase in the anti-cancer efficacy. This research could be used to increase the self-assembly formation of other cells and in turn the capability of the cell.⁴⁷

1.10 Control of Assembly with DNA

DNA can be folded in order to produce self-assembled intricate structures and shapes on the nanoscale. A single-stranded scaffold is produced for a desired shape, held in place by 'staple' strands which self-assemble to the scaffold strands. The first shape to be produced from this technique was a one-dimensional barcode array⁴⁸ before five helper strands and a long strand of DNA were combined in order to produce an octahedron.⁴⁹ Since this, many advances have been made in the self-assembly of DNA, with more intricate shapes such as a smiley face and larger structures, including extended periodic lattices and a hexamer of triangles being produced, shown in figure 9 below.⁵⁰

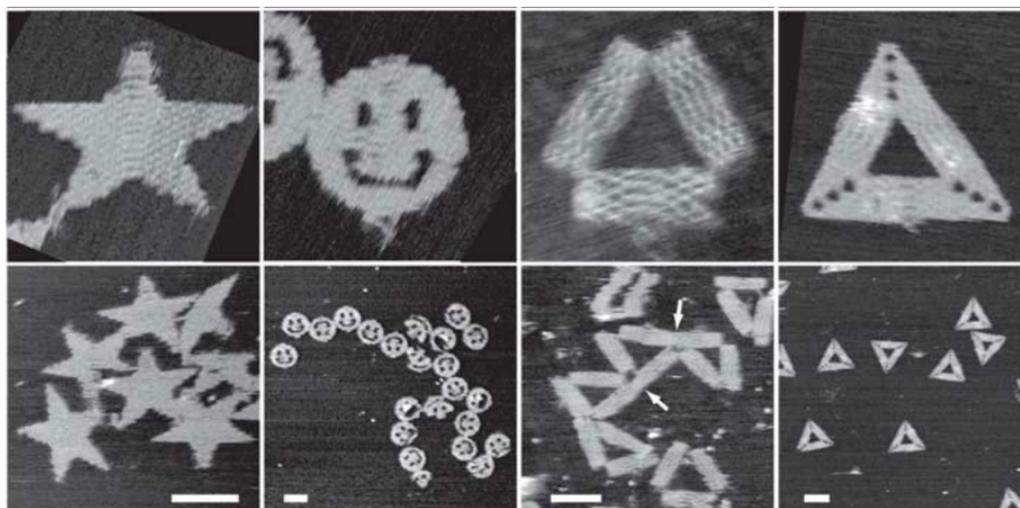


Figure 9: AFM images of four different DNA origami shapes. The scale bars in the bottom row of images represent 100 nm. From left to right: star; disk with holes; triangle produced by rectangular domains; triangle comprised of trapezoidal regions with bridges between.⁵⁰

As DNA assembly is both powerful and precise, it can be used to control the positioning of other nanomaterials. Research has shown that semiconductor and metal nanocrystals can be controllably bonded to a scaffold double-stranded DNA template in order to produce nanostructures with optical properties.⁵¹ Gold nanospheres and nanocubes were used to alter the structure and fibrillation kinetics of fibrils. Due to the higher affinity of gold nanospheres with amyloid- β peptide, thus leading to stronger interactions of the

nanospheres with the peptide, the increase in the rate of fibril formation was greater compared to gold nanocubes. As well as fibrils forming more rapidly with gold nanospheres than gold nanocubes, different morphologies were also seen in the fibrils formed. Gold nanocubes formed long mature fibrils that precipitated and intertwined with each other; in contrast, gold nanospheres formed fibrils that were shorter, wider and arranged themselves side-by-side.⁵² This contrast is clearly displayed in figure 10 below.

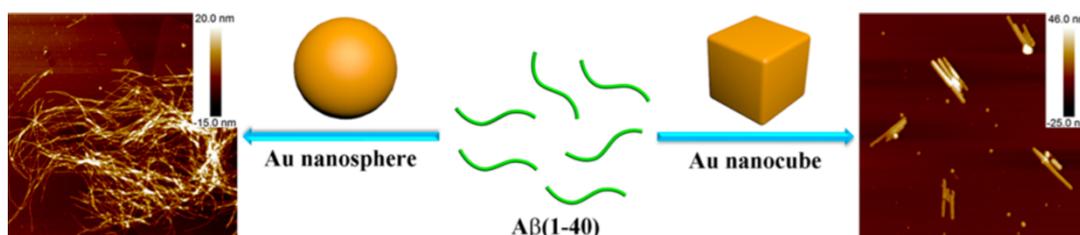


Figure 10: AFM images of:

Left: Aggregates of A β (1–40) and gold nanospheres ~20 nm in diameter after 72 hours incubation

Right: Aggregates of A β (1–40) and gold nanocubes ~20 nm in edge length after 72 hours incubation⁵²

Modified-DNA has been used in a confined template in order to mediate the assembly of these nanoparticles due to the sequence-specific binding of DNA. Research into the thermodynamic control of this assembly has enabled the templates to be specifically designed to yield nanoparticles on desired substrates with the correct size, shape and concentration. It is also thought that changes in temperature and time of the reaction will enable the growth of the substrate with nanoparticles to be most effective.⁵³

Following on from using modified-DNA as a template, nucleobase peptide amphiphiles, which are based on dipeptides, can exist as helical or curved substances in solution or as gels. The amphiphiles can self-assemble due to hydrophilic-hydrophobic interactions but can also interact with complementary nucleobases. This interaction allows the production of template structures with nucleobases that can be further varied due to their self-assembly formation.⁵⁴ This provides proof-of-principle for control of amyloid formation through interaction with DNA.

More recently, DNA has been used as a template in order to control the assembly of individual colloidal nanoparticles. A template of specific nucleic acids and polymer pores were able to control the assembly of nanoparticles of varying shapes and sizes, resulting in conformations that can be controlled and thus changed. It also enabled controlled assembly on both the nanometre and micrometre scale. Consequently, the light interactions of optical materials that use nanoparticles can be controlled and further studied.⁵⁵

1.11 Control of Amyloid Assembly with DNA

Research, by Gras' and Seeman's group in the USA, shows amyloid fibrils can be contained within DNA origami nanotubes. A DNA-peptide conjugate was annealed with a scaffold strand, which resulted in a DNA nanotube. The DNA region of the conjugate acted as a linker to form the nanotube, whilst the peptide was concealed inside the nanotube in order to nucleate the formation of amyloid fibrils. As a result, the amyloid fibrils were enclosed within the DNA nanotube, showing that the peptide successfully nucleated their formation and that a protein linked to a DNA molecule can be used to produce a bigger framework; this is shown below in figure 11.⁵⁶

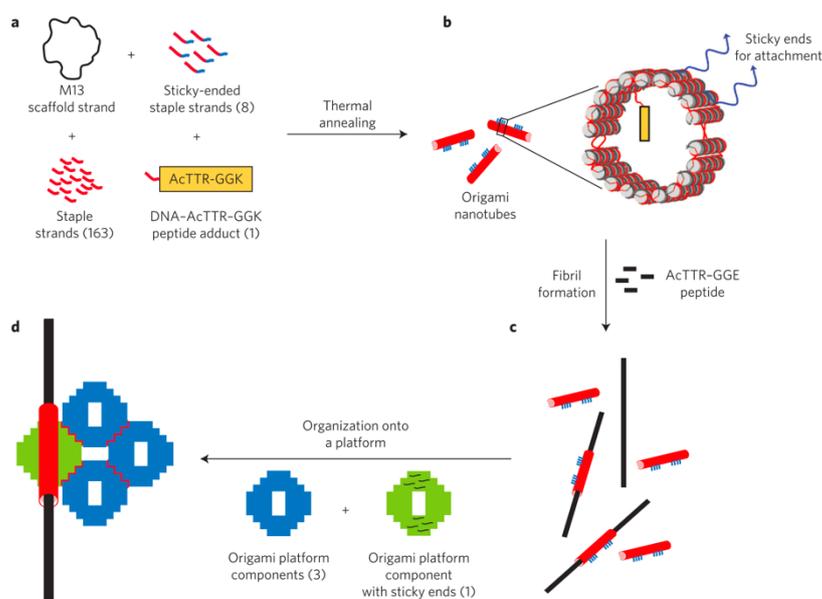


Figure 11: Process of organizing amyloid fibrils using DNA origami⁵⁶
 (a) The components required (b) A 20-helix DNA nanotube containing the peptide on the inside (yellow), and the sticky ends for attachment (blue) (c) The mixing of nanotubes with a peptide solution to form fibrils enclosed within the nanotubes. (d) The organised fibril on the platform.

A similar paper published showed that polymers were able to be templated using DNA bases. By using nucleobase-templated polymerisation with a DNA backbone of thymine with adenine monomers aligned complementarily, a polymer was able to be synthesised of a similar weight and length to the template polymer used.⁵⁷

The bacteria *Staphylococcus aureus* is known to cause persistent infections by accumulating and thus producing a biofilm, which is produced through a matrix containing polysaccharides, extracellular DNA and proteins, including amyloid fibrils. It has been shown that the formation of amyloid fibrils is stimulated by the extracellular DNA, which causes the formation of specific phenol soluble modulins peptides; in conditions where the extracellular DNA was not produced, or contained mutations, a lack of amyloid was seen. Phenol soluble modulins were found to be less cytotoxic than the soluble peptide equivalent, showing that extracellular DNA may help to reduce the toxic peptides by favouring peptide aggregation, thus producing amyloid fibrils. Despite this, it is not yet known how all of the matrix components interact in order to contribute to the formation of the biofilm.⁵⁸

Other research has shown that certain peptide-DNA conjugates are able to controllably self-assemble into amyloid-like fibrils. The fibrils form through a nucleation polymerisation process in which the driving force is the incompatibility of the hydrophobic peptides and hydrophilic DNA groups. π - π stacking and intermolecular interactions between the peptides and DNA groups leads to the formation of spherical structures, which act as nuclei for the nucleation step and thus can be combined into fibres with a β -sheet structure.⁵⁹

1.12 Project Aims

The overall aim of this project is to use DNA base pairing interactions in order to modify and the assembly of Sup35, a protein that is known to form amyloid. A chemical linker will be synthesised in order to link the DNA strands with the protein and the growth of fibrils monitored to assess whether the growth can be reduced or controlled. The DNA will serve as a 'template' to control the length of the fibrils produced; this template is hoped to control

the amount of protein monomers that can form fibrils and also potentially lead to less fragmentation by strengthening the fibril. Less fragmentation could help to reduce the amount of the cytotoxic fragments formed and associated with neurodegenerative disease.

Sup35 is an aggregate of eRF3, which is a yeast termination factor. The protein's amino acid sequence is comprised of 685 residues and can be separated into three domains: N, C and M. The N domain contains a succession of a nine-residue oligomer, PQGGYQQYN, with a total of 123 residues that are responsible for amyloid formation. There is a high quantity of polar residues in this domain, such as glutamine (Q), which is a common characteristic of amyloid-forming proteins.⁶⁰ Polar functional groups are also found in residues 124 to 245, the M domain, whilst the remaining residues, 246 to 683, form the C domain and are used in the termination activity of the protein.⁶¹ As this project will not be utilising the protein's termination activity, the smaller Sup35NM protein is to be used. Some modifications have been carried out, with a histidine tag added to the N domain and an additional short, flexible linker in the C domain; the flexible linker consists of serine (S) and glycine (G) in the arrangement of GSGGSGGSGC. An additional cysteine is also added in order to ensure the correct interactions take place, resulting in the Sup35NM LCys protein.⁶² This protein will be synthesised by Liisa Lutter, who works in the Xue lab in the Kent Fungal Group.

In order to link the protein with a DNA strand, maleimide chemical linkers will be synthesised, as shown in figure 12. This will covalently attach to the Sup35NM LCys protein, whilst also hydrogen bond to a single base of DNA, as drawn in figure 13. By creating a DNA backbone, the self-assembly of amyloid fibrils can be controlled. Fibril growth could be controlled using different length DNA backbones or different DNA sequences and the effect of DNA upon the assembly of Sup35NM LCys proteins studied. A 1:1:1 ratio of protein, chemical linker and DNA monomer will be used to begin with, but this will be varied in order to see whether the protein monomers still bind to the DNA backbone if not all of them have the chemical linker.

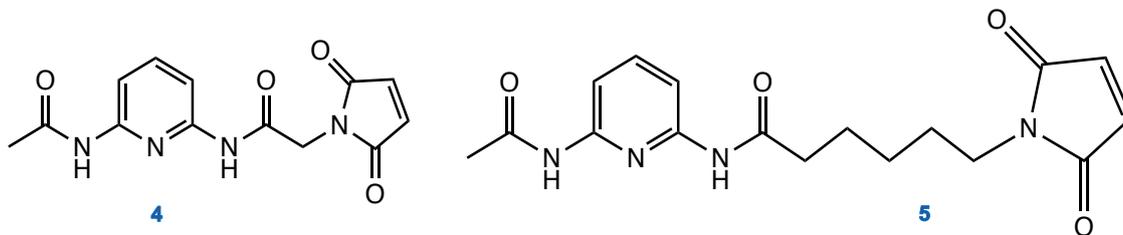


Figure 12: The two maleimide chemical linkers

Left: Compound 4

Right: Compound 5

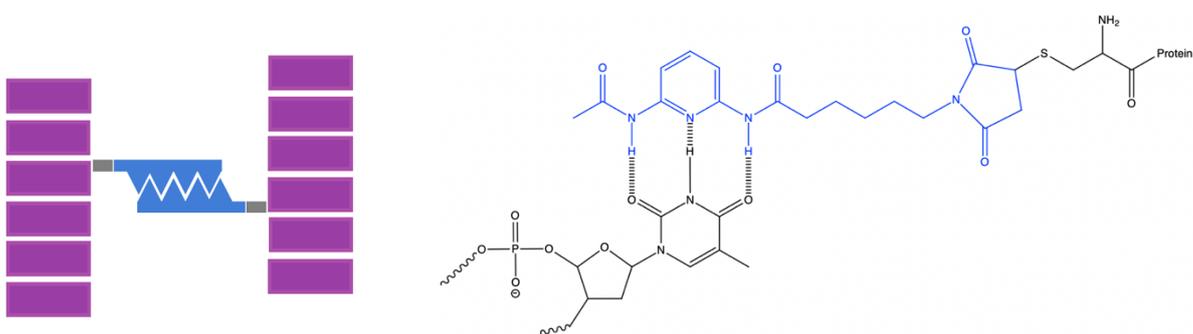


Figure 13:

Left: A cartoon diagram showing the cross-linking of two amyloid (purple) by complementary DNA strands (blue) covalently bonded to a phosphoramidite maleimide linker (grey)

Right: An example of the DNA-Linker-Protein conjugate using compound 5 as the linker

A second chemical linker, a maleimide phosphoramidite, will also be made in this project; the structures are shown in figure 14. This linker will be covalently bonded to a strand of DNA and also to the Sup35NM LCys protein. When the proteins assemble to form amyloid, one of the fibrils will be attached to the DNA strand; this DNA strand will be able to crosslink with a complementary DNA strand attached to another amyloid, as displayed in figure 15. The level of crosslinking can then be controlled by altering the number of fibrils in the amyloid that are covalently bonded to the linker and thus to a DNA strand.

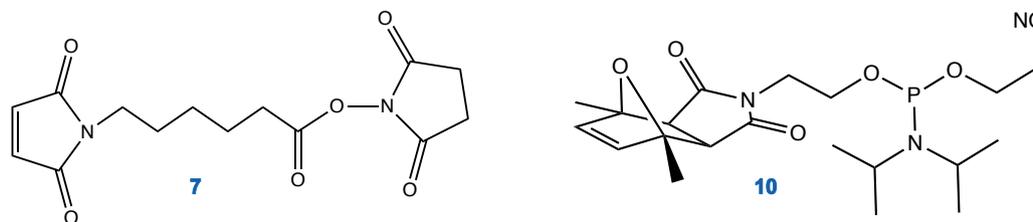


Figure 14: The two maleimide phosphoramidite chemical linkers
 Left: Compound **7**, with the maleimide ring
 Right: Compound **10**, which incorporates the phosphoramidite group and furan ring

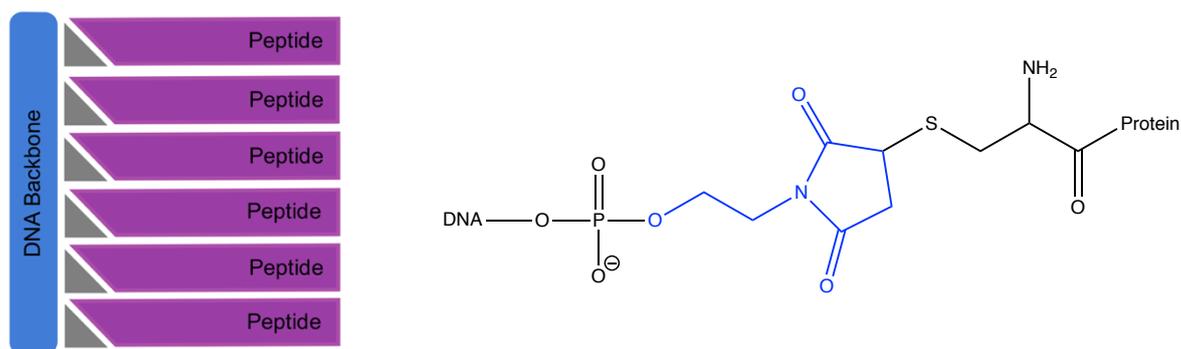


Figure 15:
 Left: A cartoon showing the controlled assembly of peptides (purple) by using a chemical linker (grey) and a DNA backbone (blue)
 Right: An example of the DNA-Linker-Protein conjugate, with compound **7** as the linker

CHAPTER 2: MALEIMIDE CHEMICAL LINKER SYNTHESIS

2.1 Introduction

Amyloid can broadly be defined as a protein that consists of polypeptide chains, generally with a β -sheet structure, that can aggregate into long fibres. This research focuses on synthesising a chemical linker whereby one end is able to covalently attach to the chosen protein, Sup35NM LCys, whilst the end hydrogen bonds to a strand of DNA, in order to prevent the self-assembly of the protein into amyloid fibrils.

The linkers, as shown again in figure 16, have two different functional groups. One end contains a maleimide ring, which is required to covalently attach to the sulphur containing linker cysteine on the protein; the double bond of the maleimide ring readily reacts with the thiol group of cysteine to form a stable carbon-sulphur bond. The opposite end of the linker contains three nitrogen groups, which are used to facilitate the hydrogen bonding with a base on the chosen DNA strand. Previous research has shown that thymine can be bonded to 2,6-Diaminopurine (DAP), a similar compound to adenine, by hydrogen bonding; as a consequence, the double helix formed predominantly had increased stability.⁶³ Two slightly different chemical linkers will be synthesised, with the length of the carbon chain between the nitrogen groups and the maleimide ring differing.

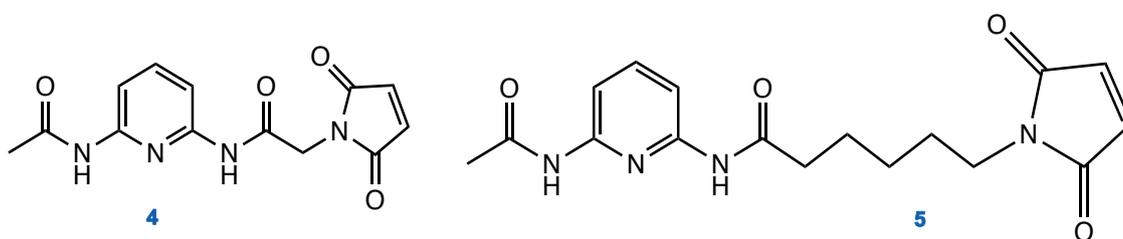


Figure 16: The two maleimide chemical linkers

Left: Compound 4

Right: Compound 5

The protein selected to be used is Sup35, a yeast termination factor, although a slightly modified version will be used. The smaller Sup35NM protein will be used, as the C domain is not used in the [PSI⁺] phenotype and only in the termination activity of the protein and hence not required. A linker cysteine group, LCys, will also be added resulting in the protein Sup35NM LCys.

Once conjugated, the DNA-linker-protein system, as shown in figure 17, will be studied to determine if the fibril growth can be controlled using the DNA backbone. Variations of DNA strands and sequences can be used to see if there is a changing effect on the assembly of fibrils.

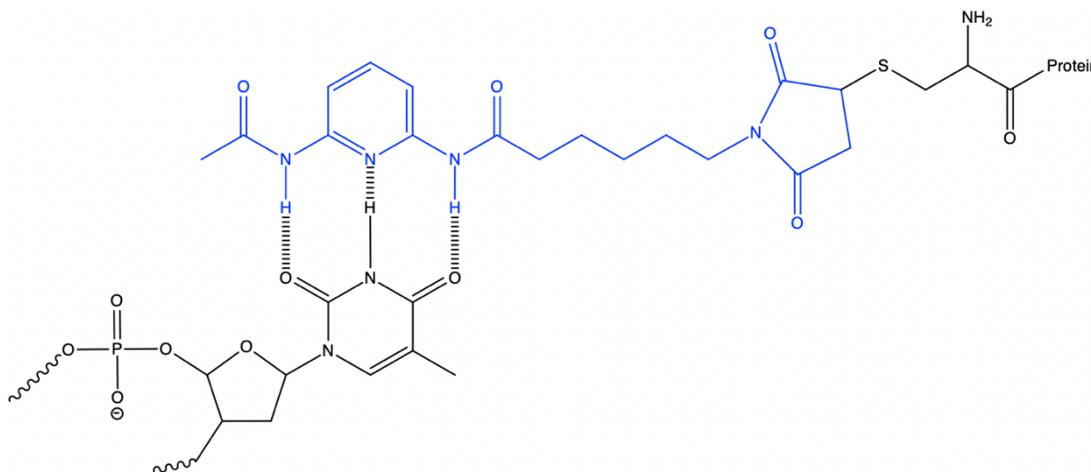


Figure 17: The DNA-Linker-Protein conjugate using the N-(6-Acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Hexanamide linker

2.2 General Considerations

All solvents, reagents and buffer components were purchased from Fisher or Sigma Aldrich, except those listed below.

NMR spectra were obtained on a Bruker AV2 400 MHz spectrometer; the spectra were calibrated to the centre of the set solvent peak and chemical shifts then reported in parts per million (ppm). Spectra were analysed using ACD labs.

UV-Vis Spectroscopy data was collected by a Nanodrop spectrophotometer. The pedestal was cleaned with ethanol and deionised water before a blank was run of deionised water (2 μ L). The sample was then loaded (2 μ L), a spectrum run in the wavelength of 220 – 350 nm and the process repeated 3 times. In order to determine the DNA concentration, the measured A260 value was multiplied by the A260 value calculated from the IDT Oligoanalyser tool⁶⁴.

MS data was attained by using one of two methods. For small organic molecules, a Thermo MSQ Plus LC-MS system was used. However, for larger samples, or those containing protein, peptide or DNA, a Bruker micrOTOF-Q Mass Spectrometer was used.

DNA strands were purchased from Integrated DNA Technologies or synthesised using an Expedite 8909 System (BioLytic). Phosphoramidite reagents (dA-CE, dT-CE, dG-CE, and dC-CE), cap A, cap B, deblock and oxidiser were purchased from Link Technologies; the phosphoramidite reagents were diluted as recommended using dry acetonitrile. The DNA strands were synthesised on a 1 μ M scale.

SDS-polyacrylamide gel electrophoresis (12 %) was also carried out. The gel was run for 60 minutes at 75 - 80 Volts, depending on the stack, before the voltage was increased and the gel run for a further 120 minutes.

Liisa Lutter from the School of Biosciences at the University of Kent expressed and purified the protein to be used. Elution G buffer was composed of 20 mM Tris-HCl, 0.5 M NaCl and 6 M GdnHCl, whilst the fibril formation buffer was a phosphate buffer.

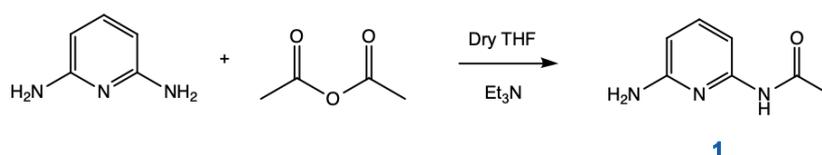
2.3 Experimental

2.3.1 Synthesis of Linkers

N-(6-Acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Acetamide

Linker

N-(6-Aminopyridin-2-yl)Acetamide (**1**)



Scheme 1: Acetylation of 2,6-diaminopyridine

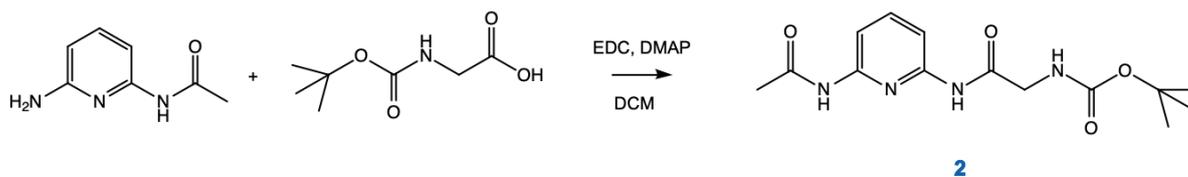
A solution of acetic anhydride (5.80 g, 0.057 mol, 1 mol equiv) in dry THF (15 mL) was added dropwise, over 2 hours, to a stirring solution of 2,6-diaminopyridine (7.52 g, 0.069 mol, 1.2 mol equiv) and triethylamine (8.33 g, 0.082 mol, 1.4 mol equiv) in dry THF (85 mL). Once added, the reaction solution was left to stir overnight. Water (40 mL) was then added and the solvent removed by rotary evaporation before the solution was left to recrystallise overnight. The solution was dried under vacuum in a Büchner funnel and washed with ice water and diethyl ether, producing dark green crystals (6.15 g, 0.0406 mol, 72 % yield).

¹H NMR (400 MHz, DMSO), ppm: 9.88 (1H, s, NH), 7.32 (1H, t, J = 7.9 Hz, ArH), 7.21 (1H, d, J = 7.7 Hz, ArH), 6.16 (1H, dd, J = 0.8 Hz, 8.0 Hz, ArH), 5.78 (2H, s, NH₂), 2.02 (3H, s, CH₃),

¹³C NMR (400 MHz, DMSO), ppm: 168.87, 158.83, 150.60, 138.91, 103.30, 100.79, 24.04

MS (ESI, positive mode): calculated [MW]H⁺ = 152.1, m/z found = 152.0

Tert-Butyl(2-((6-Acetamidopyridin-2-yl)Amino-2- Oxoethyl)Carbamate (2)



Scheme 2: Amide coupling of compound 1

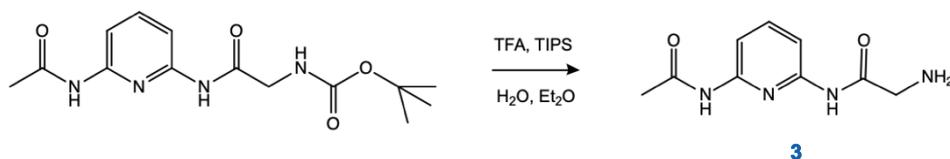
Compound 1 (3.71 g, 0.025 mol, 1 mol equiv) was added to DCM (100 mL) and stirred until dissolved. N-Boc Gly (4.49 g, 0.026 mol, 1 mol equiv), EDC-HCl (5.27 g, 0.028 mol, 1.1 mol equiv) and DMAP (a few grains) were added and the solution stirred overnight. The reaction solution was washed three times with 2M sodium hydroxide, once with water and then finally with brine. Magnesium sulphate was used to dry the organic layer before being filtered out. The solvent was removed by rotary evaporation, yielding a cream/light brown solid (3.66 g, 0.0118 mol, 48 % yield).

^1H NMR (400 MHz, CDCl_3), ppm: 8.32 (1H, s, NH), 7.97 (1H, s, ArH), 7.80 (1H, d, $J = 7.6$ Hz, ArH), 7.60 (1H, t, $J = 8.1$ Hz, ArH), 3.92 (2H, s, CH_2), 2.13 (3H, s, CH_3), 1.40 (9H, s, $\text{C}(\text{CH}_3)_3$)

^{13}C NMR (400 MHz, CDCl_3), ppm: 168.79, 168.25, 149.70, 148.86, 140.76, 109.83, 45.03, 28.33, 24.68

MS (ESI, positive mode): calculated $[\text{MW}]\text{H}^+ = 309.2$, m/z found = 332.1 (Na^+ adduct)

N-(6-Acetamidopyridin-2-yl)-2-Aminoacetamide (**3**)



Scheme 3: Deprotection of compound **2**

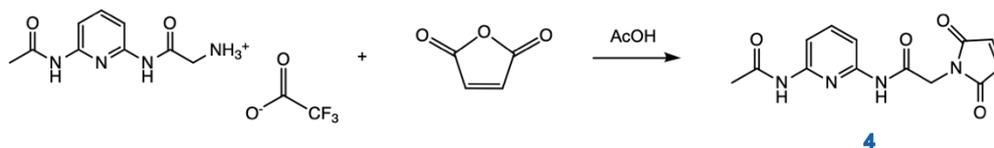
Compound **2** (0.90 g, 2.920 mmol) was dissolved in TFA (3.43 mL). TIPS (0.15 mL) and water (0.15 mL) were added and the solution stirred for 1 hour. Diethyl ether (47.00 mL) was then added and the solution stirred overnight. Overnight a precipitate had formed. The solution was centrifuged, the supernatant removed and diethyl ether added; this process was repeated three times. The product was dried under vacuum in a Büchner funnel and collected as a white solid (0.902 g, 4.33 mmol, 148 % apparent yield).

¹H NMR (400 MHz, DMSO), ppm: 10.56 (1H, s, NH), 10.12 (1H, s, NH), 8.13 (3H, s, NH₂), 7.78 (3H, t, J = 6.7 Hz, ArH), 3.37 (2H, s, CH₂), 2.12 (3H, s, CH₃)

¹³C NMR (400 MHz, DMSO), ppm: 169.84, 158.42, 149.84, 140.81, 115.27, 109.24, 24.51

MS (ESI, positive mode): calculated [MW]H⁺ = 209.1, m/z found = 209.2, 231.2 (Na⁺ adduct)

N-(6-Acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Acetamide (**4**)



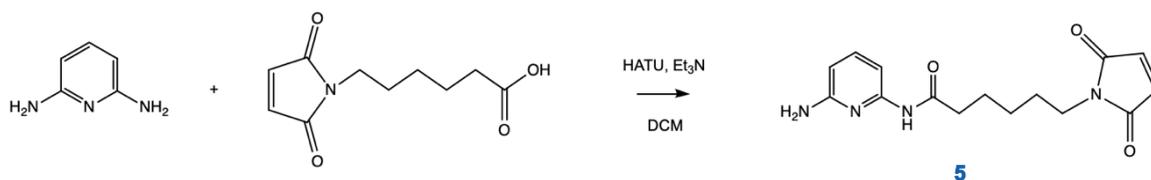
Scheme 4: Addition of maleic anhydride to compound **3**

Compound **3** (0.26 g, 7.948 x 10⁻⁴ mol, 1 mol equiv) was dissolved in acetic acid (8 mL) and maleic anhydride (0.09 g, 8.970 x 10⁻⁴ mol, 1.1 mol equiv) was added. The solution was

heated to 180 °C and refluxed overnight. The solvent was removed by rotary evaporation and the product collected as an orange residue (0.009 g, 3.122×10^{-5} mol, 4 % yield).

***N*-6-Acetamidopyridin-yl-6-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Hexanamide Linker**

N-(6-Amidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Hexanamide (**5**)



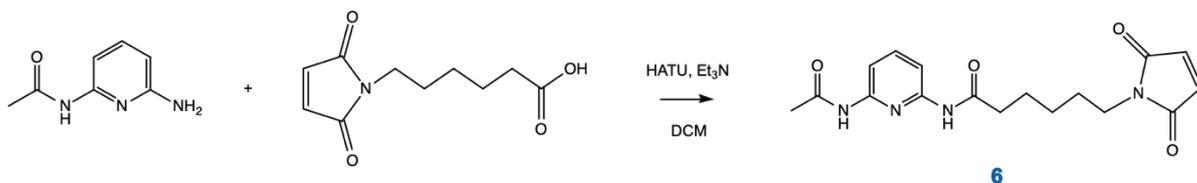
Scheme 5: Addition of 6-maleimido-hexanoic acid to 2,6-diaminopyridine

HATU (0.96 g, 2.522 mmol, 1.1 mol equiv) and 6-maleimido-hexanoic acid (0.54 g, 2.533 mmol, 1.1 mol equiv) were dissolved in a solution of triethylamine (0.35 mL, 1.1 mol equiv) and DCM (10.0 mL). The solution was stirred for 30 minutes, before 2,6-diaminopyridine (0.25 g, 2.300 mmol, 1 mol equiv) was added. The solution was left to stir overnight. The solution was washed with water, the organic layer dried with magnesium sulphate and the solvent removed by rotary evaporation, yielding the product as a light brown residue (0.59 g, 1.942 mmol, 84 % yield).

¹H NMR (400 MHz, CDCl₃), ppm: 9.01 (1H, s, NH), 7.77 (2H, d, J = 7.9 Hz, ArH), 7.64 (1H, t, J = 8.0 Hz, ArH), 7.45 (2H, d, J = 8.2 Hz, ArH), 6.62 (2H, s, CH), 3.44 (2H, t, J = 6.2 Hz, CH₂), 2.31 (2H, t, J = 8.0 Hz, CH₂), 2.13 (2H, s, NH₂), 1.64 (2H, m, J = 7.9 Hz, CH₂), 1.54 (2H, m, J = 7.6 Hz, CH₂), 1.30 (2H, m, J = 7.5 Hz, CH₂)

¹³C NMR (400 MHz, CDCl₃), ppm: 176.86, 175.79, 159.64, 157.53, 140.90, 134.88, 105.94, 41.54, 36.98, 29.89, 22.32, 18.52

N-(6-Acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Hexanamide (**6**)



Scheme 6: Addition of 6-maleimido-hexanoic acid to compound **1**

HATU (1.01 g, 2.649 mmol, 2 mol equiv) and 6-maleimido-hexanoic acid (0.56 g, 2.656 mmol, 2 mol equiv) were dissolved in a solution of triethylamine (0.37 mL, 2 mol equiv) and DCM (15 mL). The solution was stirred for 90 minutes, before compound **1** (0.25 g, 1.323 mmol, 1 mol equiv) was added and stirring continued overnight. The solution was washed with saturated potassium carbonate and brine, the organic layer dried with magnesium sulphate and the solvent removed by rotary evaporation, yielding the product as an orange oil (Crude product: 0.356 g, 1.034 mmol, 62 % yield). The product was purified by column chromatography, using a running solvent of 3 % methanol in DCM. The desired fractions were combined and the solvent removed by rotary evaporation. (Purified product: 0.015 g, 4.4×10^{-5} mol, 3 % yield).

¹H NMR (400 MHz, CDCl₃), ppm: 7.83 (5H, m, J = 7.7 Hz, 8.0 Hz, ArH and NH),
6.64 (2H, s, CH), 3.47 (2H, t, J = 7.2 Hz, CH₂), 2.32 (2H, t, J = 7.5 Hz, CH₂),
2.16 (2H, s, CH₃), 1.69 (2H, m, J = 7.6 Hz, CH₂), 1.56 (2H, m, J = 7.5 Hz, CH₂),
1.31 (2H, m, J = 7.7 Hz, CH₂)

¹³C NMR (400 MHz, CDCl₃), ppm: 176.30, 169.90, 147.87, 140.56, 133.00, 108.23, 37.90,
36.50, 27.80, 26.03, 23.53, 23.50

MS (ESI, positive mode): calculated [MW]H⁺ = 344.2, m/z found = 345.2

2.3.2 Conjugation of Linker and Protein

Sup35NM LCys in guanidine was thawed, DTT (0.1 mL) added and left for 1 hour. LCys in guanidine was then loaded into a PD10 column and eluted to a volume of 3.5 mL using Elution G buffer. The absorbance of the protein was measured and the concentration calculated to be 11.5 μ M in 3.5 mL.

Linker 1B (0.5 mg, 1.453 mmol) was dissolved in methanol (1 mL).

8 samples were then prepared using the protein and linker solutions, with 2, 5, 10 or 20 equivalents of linker to protein, in the following quantities;

2 equivalents of linker to protein:	0.39 μ L linker, 24.6 μ L protein
5 equivalents of linker to protein:	0.97 μ L linker, 24.03 μ L protein
10 equivalents of linker to protein:	1.94 μ L linker, 23.06 μ L protein
20 equivalents of linker to protein:	3.88 μ L linker, 21.12 μ L protein

4 of the samples, one of each of the different equivalents, were reacted overnight at room temperature, whilst the remaining 4 were conjugated overnight at 37 °C. The samples were then analysed using MALDI-TOF MS.

2.4 Results and Discussion

2.4.1 Synthesis of Linkers

N-(6-Acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Acetamide Linker

N-(6-Aminopyridin-2-yl)Acetamide

Compounds **1** – **4** were synthesised using a method from literature, that was adapted in order to produce the target molecule. The first reaction was an acetylation reaction for which the method used was taken directly from the paper.⁶⁵ All three attempts of this reaction have been carried out using the same protocol, as described above, with the products being successfully made with good yields. The literature procedure reported a yield of 47.3% and the reactions carried out produced yields between 55 % - 60 %; the slightly higher yield is most likely due to a by-product, in which there is the addition of the acetyl group onto both nitrogen atoms in the starting reagent (2,6-diaminopyridine). The ¹H and ¹³C NMR spectra produced are also very comparable to that in the literature, indicating the synthesis had been successful.⁶⁵

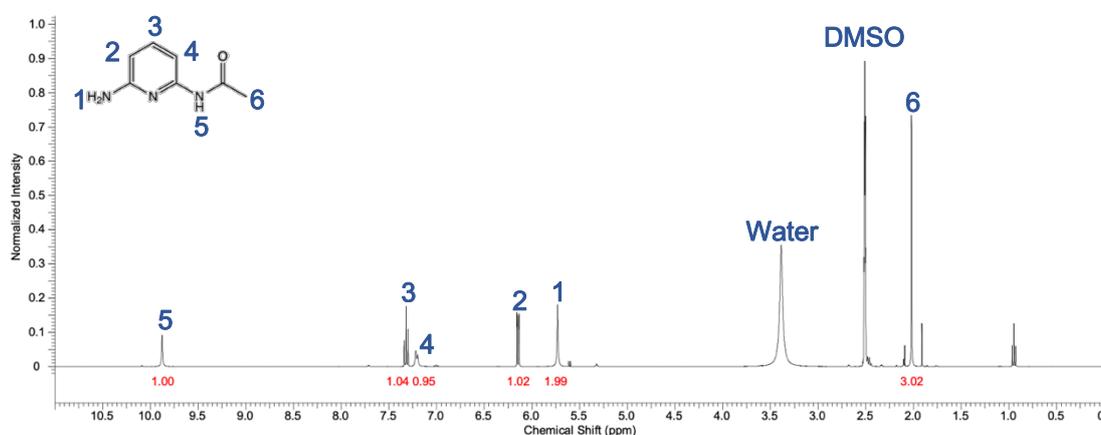


Figure 18: ¹H NMR of the acetylation of 2,6-diaminopyridine (compound **1**)

Tert-Butyl(2-((6-Acetamidopyridin-2-yl)Amino-2- Oxoethyl)Carbamate

This reaction was an amide coupling reaction of compound **1** with *N*-*boc* glycine, using EDC.HCl as the coupling reagent and DMAP as a catalyst, as widely published in literature.⁶⁶ The reaction worked well on both attempts, with the target product being synthesised. Despite this, on the second attempt of the reaction a black precipitate formed overnight; it is likely that the precipitate was undissolved starting reactant, *N*-(6-aminopyridin-2-yl)acetamide, as the reaction resulted in a lower yield compared to the previous attempt. The precipitate was removed by vacuum filtration on a Büchner funnel before the work-up was carried out as normal. The ¹H NMR spectra obtained for both of the reaction attempts, one of which is shown in figure 19, match well with the integration and peak shifts as expected.

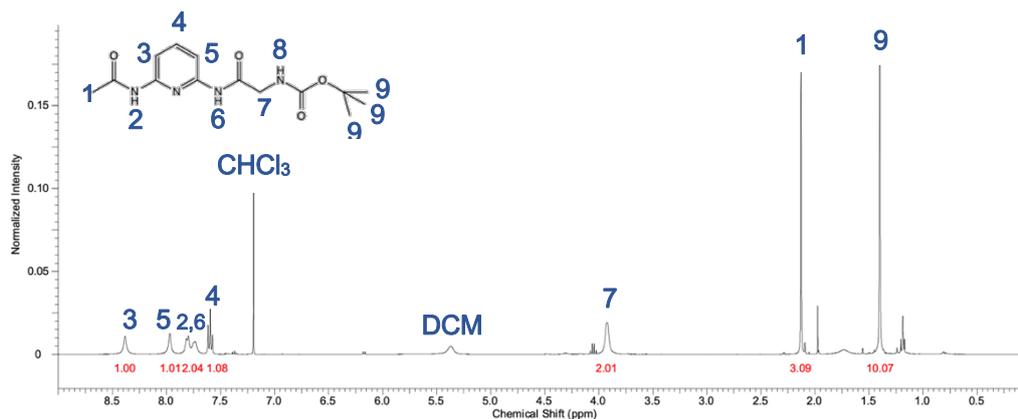


Figure 19: ¹H NMR of the amide coupling of *N*-(6-aminopyridin-2-yl)acetamide (compound **2**)

N-(6-Acetamidopyridin-2-yl)-2-Aminoacetamide

Reaction 3 of the synthesis was the deprotection of *tert*-butyl(2-((6-acetamidopyridin-2-yl)amino-2-oxoethyl)carbamate, which used TFA, TIPS and water. Deprotection reactions using TFA are common, whilst these particular reaction conditions are often used in the cleavage of peptides from the resin during synthesis.⁶⁷ All four attempts of this reaction have been successful, producing the desired product with good yields, although residue solvent has often remained. The main difference between these reactions attempts has

been the work-up method. Two of the attempts produced a fine precipitate that, when filtered on a Büchner funnel, flowed through and was not filtered out. As a result, the solution was centrifuged 3 times; the supernatant was removed and diethyl ether added after each spin before the solid was filtered out and dried. On the other two attempts of the reaction, a more crystalline and 'sticky' precipitate formed, which could be filtered out of the solution by vacuum filtration and dried.

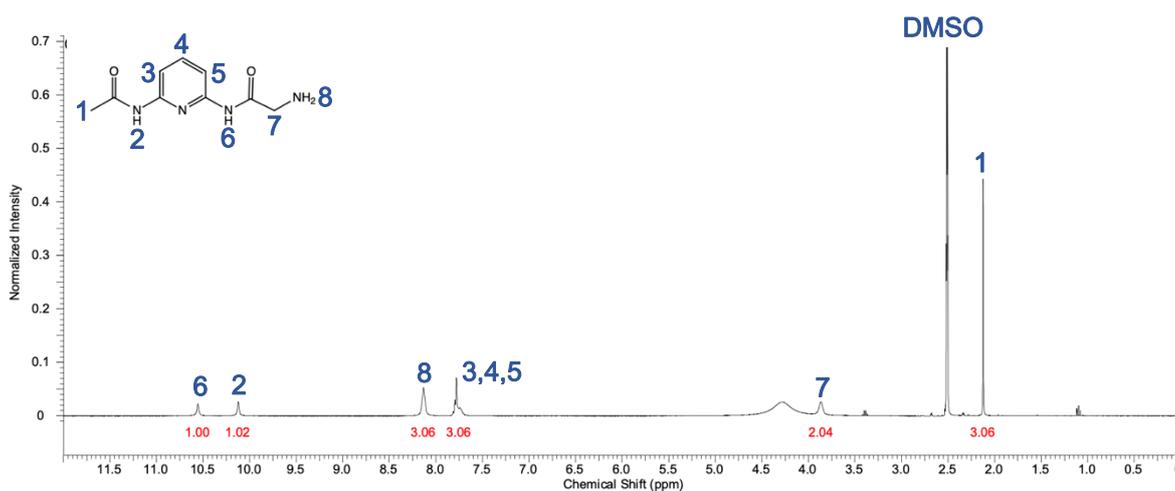


Figure 20: ^1H NMR of the deprotection of ^tBu (2-((6-acetamidopyridin-2-yl)amino-2-oxoethyl)carbamate (compound **3**)

N-(6-Acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Acetamide

The addition reaction of acid anhydrides to amines in order to synthesise amides are frequently undertaken.⁶⁸ However, this reaction has proved problematic and so the reaction conditions have been changed multiple times to try and obtain a successful product. Initially the protocol was adapted from literature but was adjusted to suit the desired target molecule.⁶⁹ As the preliminary reaction did not work, shown by the ^1H NMR spectrum, the first adjustment made was to decrease the reaction time from ~72 hours to just ~24 hours, in case the reaction equilibrium was favouring the reverse reaction, thus causing the maleimide ring to open.

Another change was made to the work-up; water was originally added to the reaction before washing with ethyl acetate, sodium hydroxide and brine. However, this was changed so instead, excess dichloromethane was added to increase the volume of the reaction solution to aid separation and the solution was then washed with hydrochloric acid, potassium carbonate and brine.

The final change made was to the ratio of the reagents and the reaction temperature. A large excess of maleic anhydride was originally used, but this was reduced to a 1 : 1.1 ratio of amide to anhydride. Alongside this reaction, the synthesis of maleimide alcohol was being carried out for the synthesis of the phosphoramidite maleimide linker; during this, the reaction temperature was increased to 180 °C as this increased the yield and so the reaction temperature of this reaction was likewise increased.

The ^1H NMR spectrum and MS data of the latest reaction attempt of the reaction showed that the desired product was not synthesised. The desired product would have an expected MW of 288.1 g mol $^{-1}$; ESI positive conditions were used and so the target peak should appear around 289.1 g mol $^{-1}$. However, in the MS spectrum obtained for the synthesised product, the target peak is not visible, with only peaks for fragments of the starting reactants present.

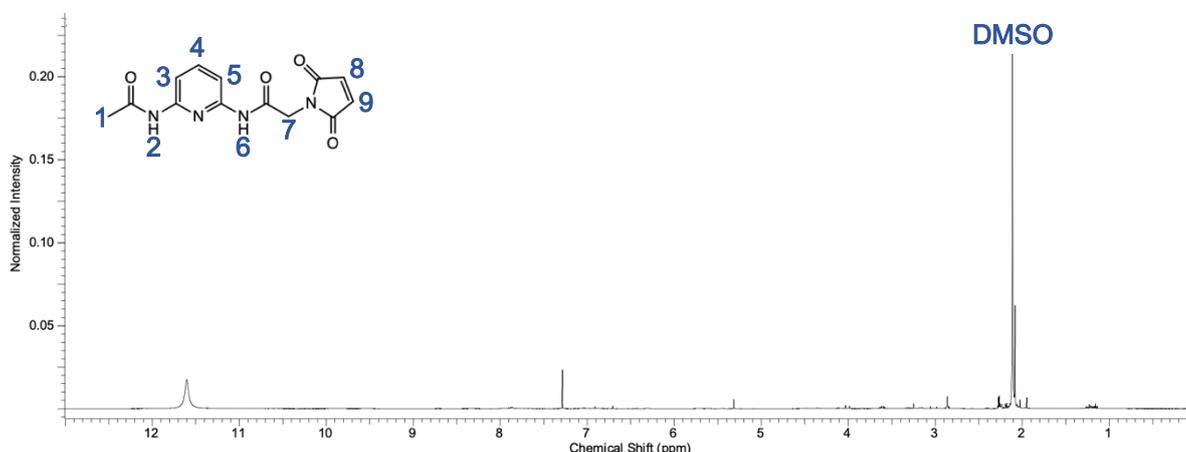


Figure 21: ^1H NMR of the addition of maleic anhydride to N-(6-acetamidopyridin-2-yl)-2-aminoacetamide (compound 4)

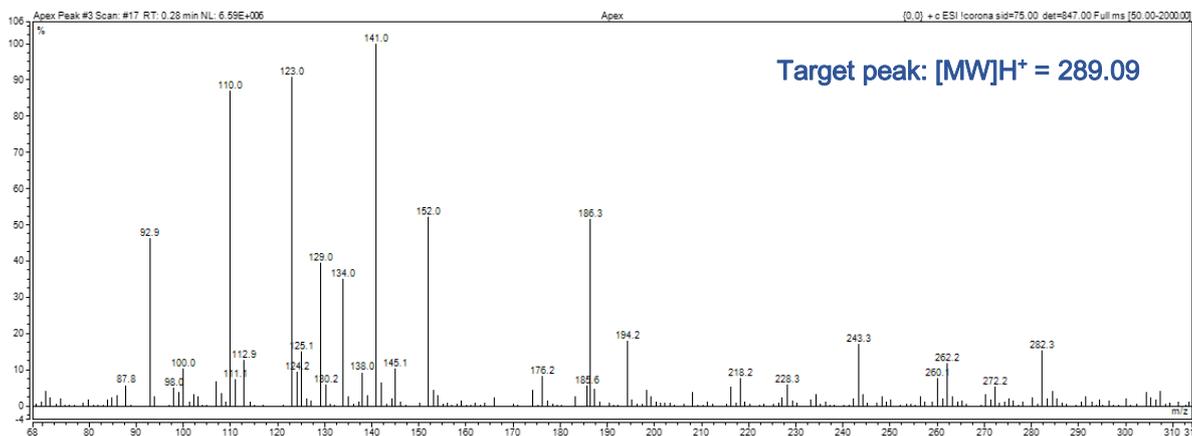


Figure 22: The mass spec data of the addition of maleic anhydride to N-(6-acetamidopyridin-2-yl)-2-aminoacetamide (compound **4**)

N-6-Acetamidopyridin-yl-6-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Hexanamide Linker

N-(6-Amidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Hexanamide

As the addition of maleic anhydride to N-(6-acetamidopyridin-2-yl)-2-aminoacetamide was proving unsuccessful and **linker 1A** was thus not being synthesized, a new method was devised which would yield a similar product, although with the longer alkyl chain.

This reaction was run as an experimental reaction to see if a new protocol, using HATU to generate an activated ester, would successfully yield the product. 2,6-Diaminopyridine was used as it is more nucleophilic than the N-(6-aminopyridin-2-yl)acetamide and so would hopefully produce the desired product.

The ¹H NMR spectrum, shown in figure 23, indicated that although only a crude product had been produced, as the solution contained some starting materials, the desired product had been formed. As a result, the same protocol was used for the addition of 6-maleimidohexanoic acid to N-(6-aminopyridin-2-yl)acetamide (compound **1**) with the addition of a purification step. Despite 2,6-diaminopyridine being more nucleophilic than N-(6-aminopyridin-2-yl)acetamide used in the next reaction, the acetamide would be better to use as the starting reagent as one of the nitrogen atoms is already acetylated. The

addition of 6-maleimidohexanoic acid to 2,6-diaminopyridine could lead to multiple side products being synthesised as the acid could react with either nitrogen atom, or to both.

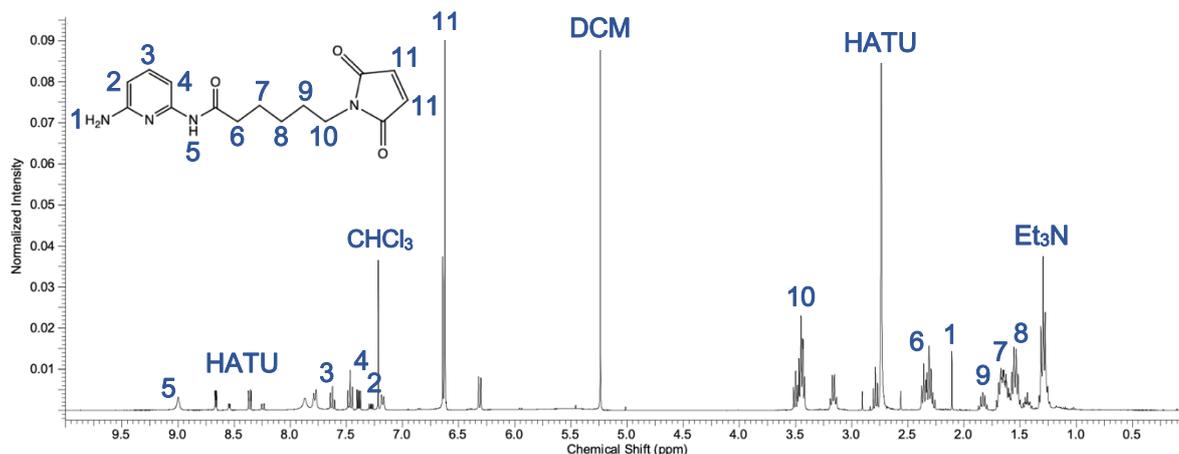


Figure 23: ¹H NMR of the addition of 6-maleimidohexanoic acid to 2,6-diaminopyridine (compound 5)

N-(6-Acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)Hexanamide

The first attempt of this reaction followed the same protocol as for the amide coupling of *N*-(6-aminopyridin-2-yl)acetamide and thus used EDC-HCl and DMAP in order to react the two reagents. This was unsuccessful as the reaction solution was swamped with DMAP, and so the reaction was repeated without the catalyst, but with no success, most probably due to the lack of catalyst. The next two attempts at the reaction used DCM as the solvent, but with varying ratios of the starting materials; 1 : 2 : 2 equivalents of amine : 6-maleimidohexanoic acid : EDC-HCl was then increased to 1 : 3 : 2. However, during the work-up of the reaction it was noted that not all the starting amine was dissolving in the solvent. Thus, the NMR spectra produced was mainly just the starting reagents, with traces of the desired product in very little quantity; little presence of the aromatic protons between 7 – 8 ppm expected in the product was seen in the spectra. The starting amine was also dissolved in THF but this also yielded no success.

HATU and triethylamine were used to form a HOAt ester, which was then reacted with the amine (compound 1). The NMR spectrum of the first attempt did not show the expected

peaks and so the ratios of the reagents were changed. The ratio of acetamide : 6-maleimidohexanoic acid : triethylamine : HATU was changed from 1 : 1.1 : 10 : 1.1 to 1 : 2 : 2 : 2. This attempt was more successful, with the product purified by column chromatography.

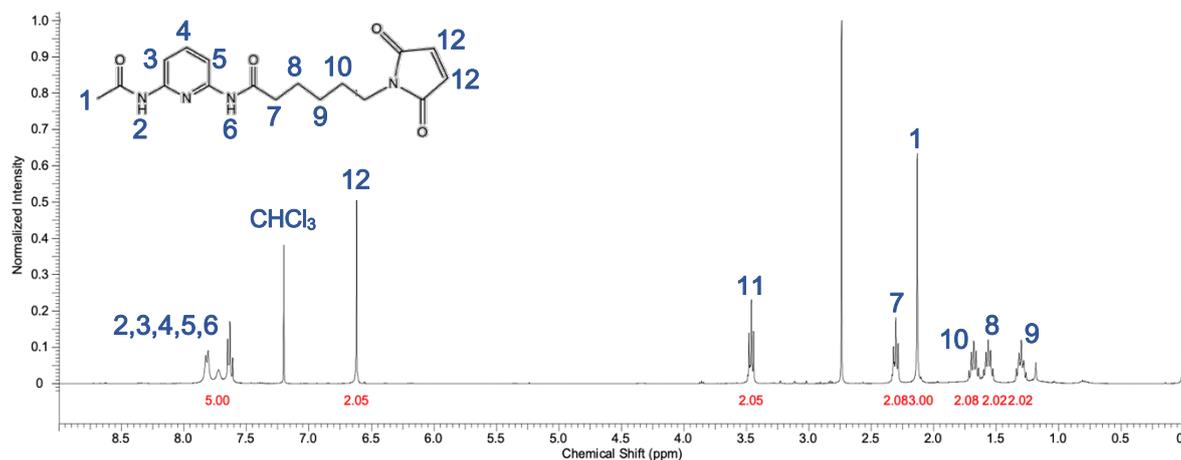


Figure 24: ¹H NMR of the addition of 6-maleimidohexanoic acid to N-(6-aminopyridin-2-yl)acetamide (compound 6)

This compound has an expected molecular weight of 344.15 g mol⁻¹, which can be seen in the MS data obtained. The negative ion mode shows the main peak at 343.14 g mol⁻¹ ([MW]H⁻) with an additional peak at 301.13 g mol⁻¹, which is due to the dissociation of the acetyl group. In the positive ion mode, there is the expected peak at 345.16 g mol⁻¹ ([MW]H⁺) but also a peak at 666.45 g mol⁻¹; this was shown to result from a contaminant in the HPLC system used to inject the sample when a blank run was undertaken. Therefore, the NMR spectrum and MS data showed that the desired compound had successfully been synthesised.

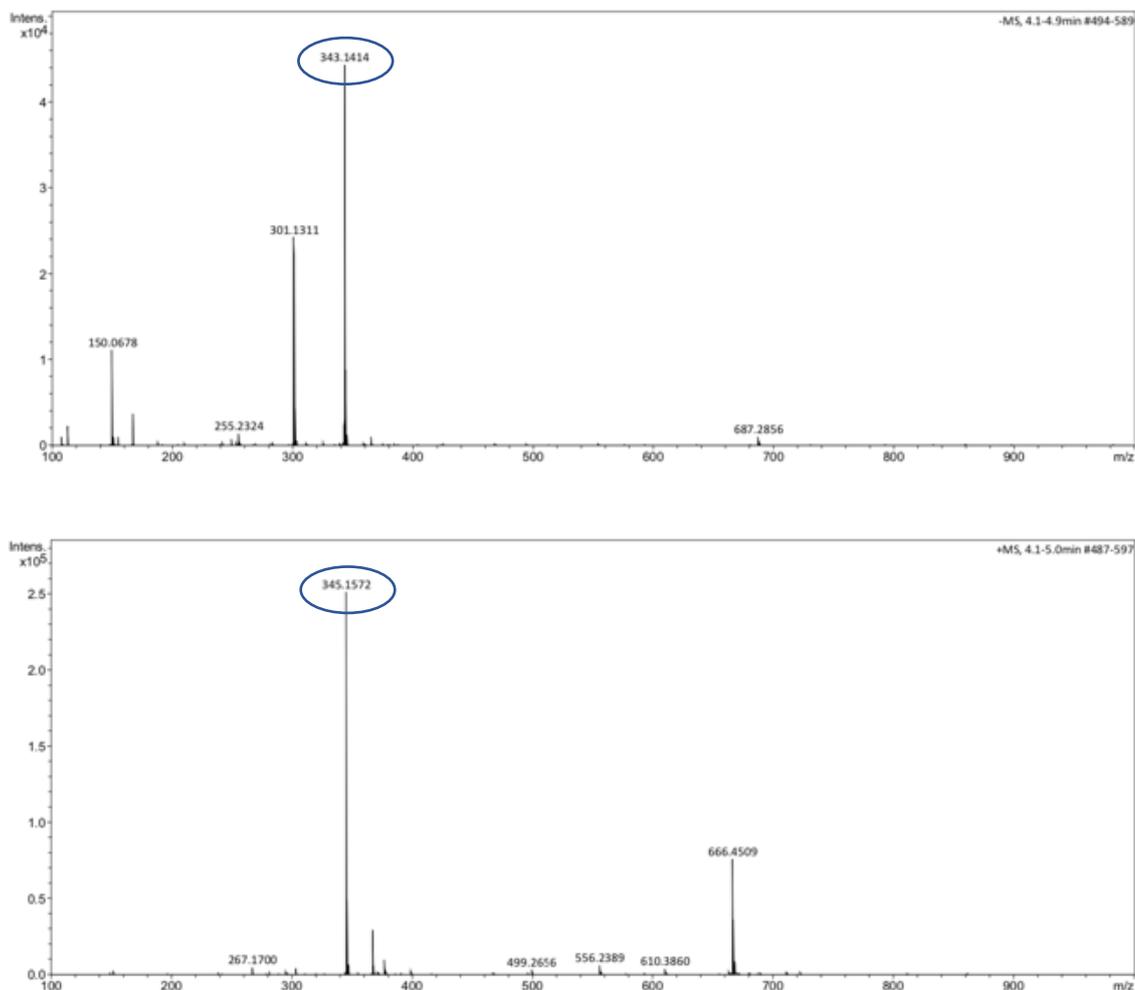


Figure 25: The mass spec data of the addition of 6-maleimido-hexanoic acid to N-(6-aminopyridin-2-yl)acetamide (compound **6**)
 Above: Negative ion mode
 Below: positive ion mode

2.4.2 Conjugation of Linker and Protein

The protein has a MW of 31,260 Da, which combined with the MW of the linker at 344 Da, would produce a conjugate with a MW of 31,604 Da. The samples with 2 and 5 equivalents of linker to protein showed no peak for the conjugate, whereas the samples containing 10 and 20 equivalents of linker to protein displayed a peak in the region of the MW of the conjugate at both room temperature and 37 °C. One sample for each temperature showed a difference of 15 Da, whilst another sample showed a difference of 13 Da. However, the MS data obtained showed that the samples conjugated at room temperature, displayed in figure 26, did not work as successfully as those left at 37 °C, shown in figure 27, as less conjugate was produced.

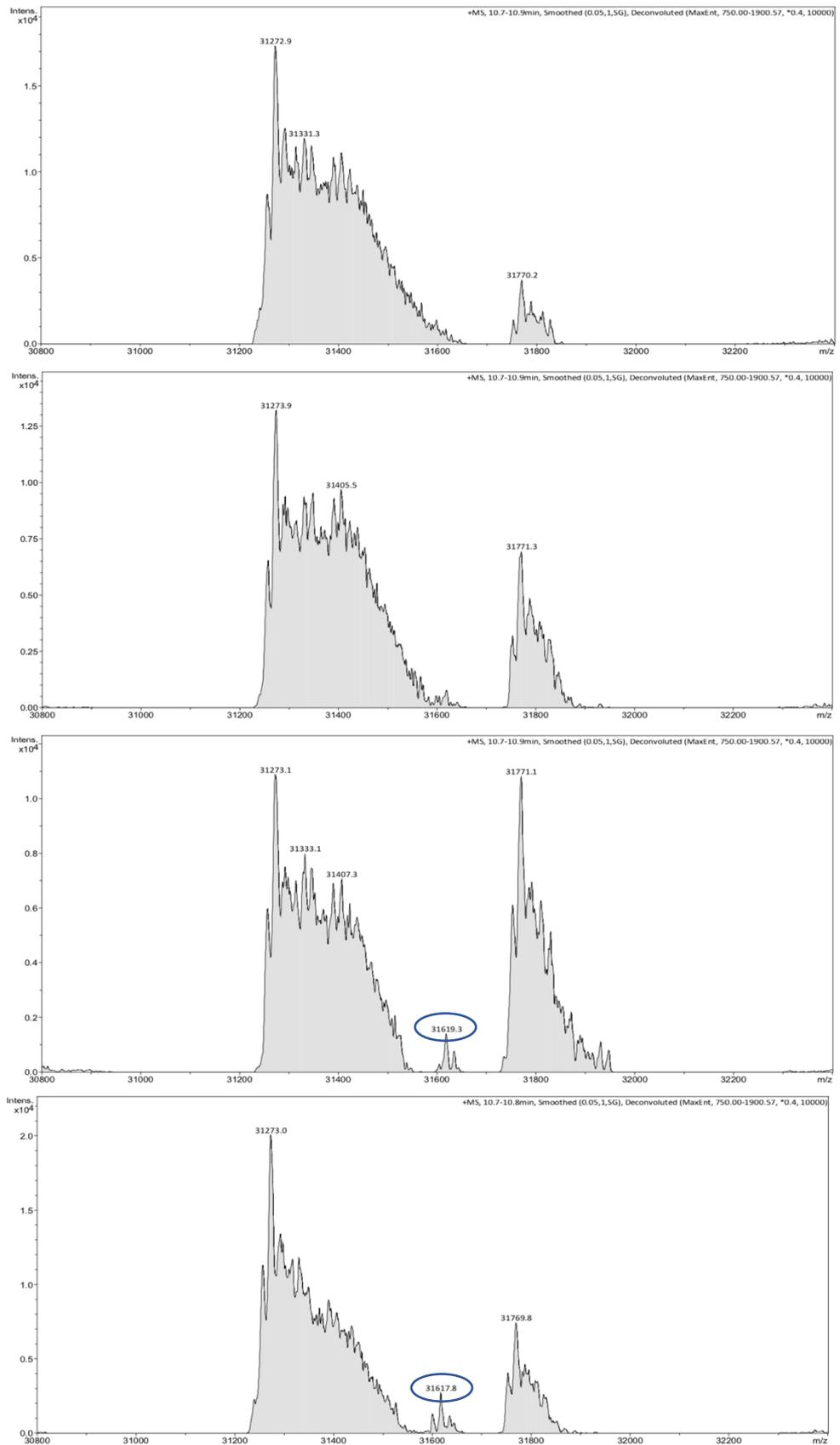


Figure 26: MALDI-TOF MS data of the conjugates incubated at room temperature
 From top to bottom: 2 equivalents; 5 equivalents; 10 equivalents; 20 equivalents

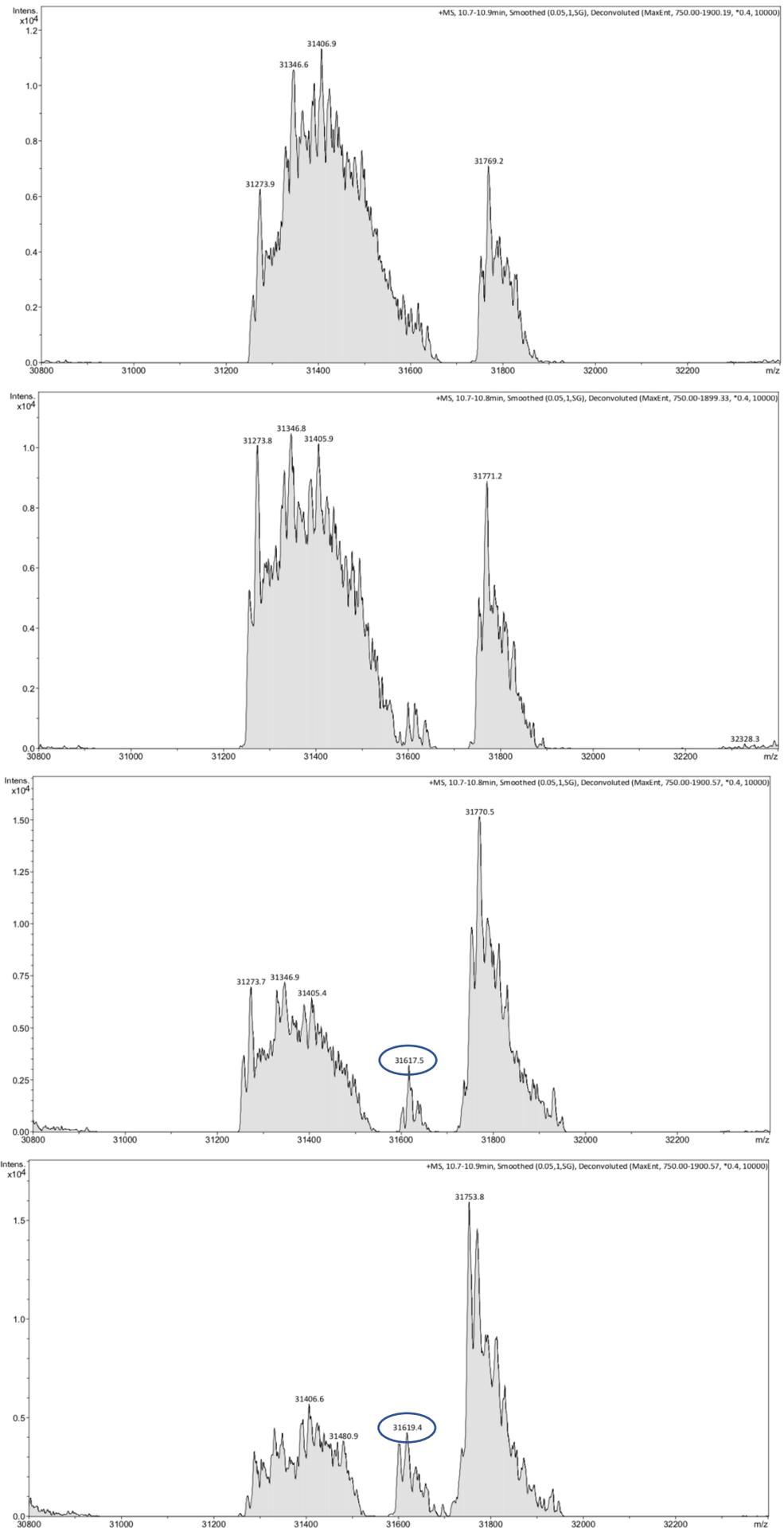


Figure 27: MALDI-TOF MS data of the conjugates incubated at 37°C. From bottom to top: 2 equivalents; 5 equivalents; 10 equivalents; 20 equivalents

2.5 Conclusion

Overall, the synthesis of **linker 1A**, the structure of which is shown in figure 28, was unsuccessful. The first three reactions in the synthesis of N-(6-acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)acetamide were carried out with positive results. However, the final reaction of the addition of maleic anhydride continuously proved problematic despite many attempts at changing the reaction conditions.

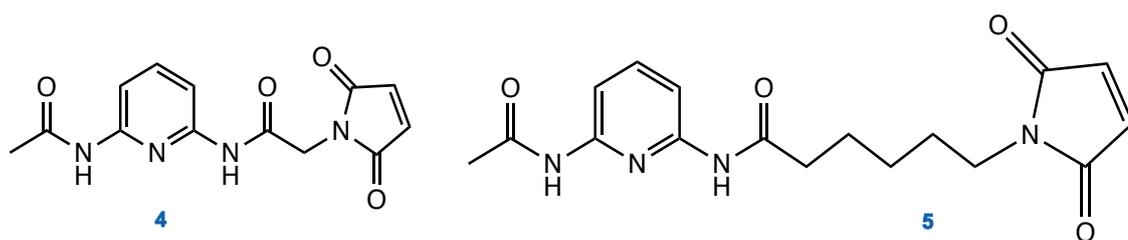


Figure 28: The two maleimide chemical linkers

Left: Compound **4**, which is **linker 1A**

Right: Compound **5**, which is **linker 1B**

Nevertheless, **linker 1B**, also shown in figure 28, was synthesised with more desirable outcomes. The target molecule, N-(6-acetamidopyridin-2-yl)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)hexanamide, was synthesised. The structure was confirmed by ^1H and ^{13}C NMR spectra and MS data. As a result, the conjugation with the protein was also carried out, which also yielded encouraging results, due to MS data confirmed that the conjugates had been produced.

3.1 Introduction

Peptides are molecules comprised of more than two amino acids, bonded through covalent peptide bonds between the carboxyl group of one amino acid and the amino group of the sequential amino acid. Some soluble peptide chains are able to self-assemble into insoluble fibrils, known as amyloid fibrils. The self-assembly causes the peptide to become resistance to degradation and thus is often associated with disease and neurodegenerative conditions.⁷⁰ From recent research has emerged the Waltz algorithm, which considers both the specific residue type and position of peptides in order to determine if the peptide contains any amyloidogenic regions and thus be able to form self-assembled fibrils. Three peptides which have been recognised as being able to form highly ordered amyloid fibrils have the structures of HYFNIF, RVFNIM and VIYKI.⁷¹

The main objective of this research was to conjugate a Waltz peptide with **linker 1B**, a maleimide chemical linker, previously synthesised in chapter 2; this could potentially allow DNA-templated amyloid as shown formerly in figure 13. The Waltz peptide, synthesised by Akiko Sato in the Serpell group, is comprised of 8 different amino acids with an acetyl group at one terminus and an amide group at the other; this gives the general structure of Ac-CGSGHYFNIF-NH₂, as shown in figure 29. Conjugation of the peptide and linker arises through thiol-maleimide chemistry, whereby the maleimide ring opens and reacts with the sulphur containing thiol group, present in the side chain of cysteine.

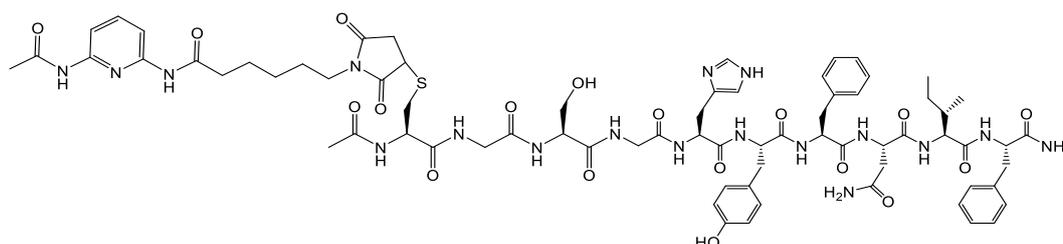


Figure 29: The conjugate of the Waltz peptide and maleimide chemical linker

3.2 Experimental

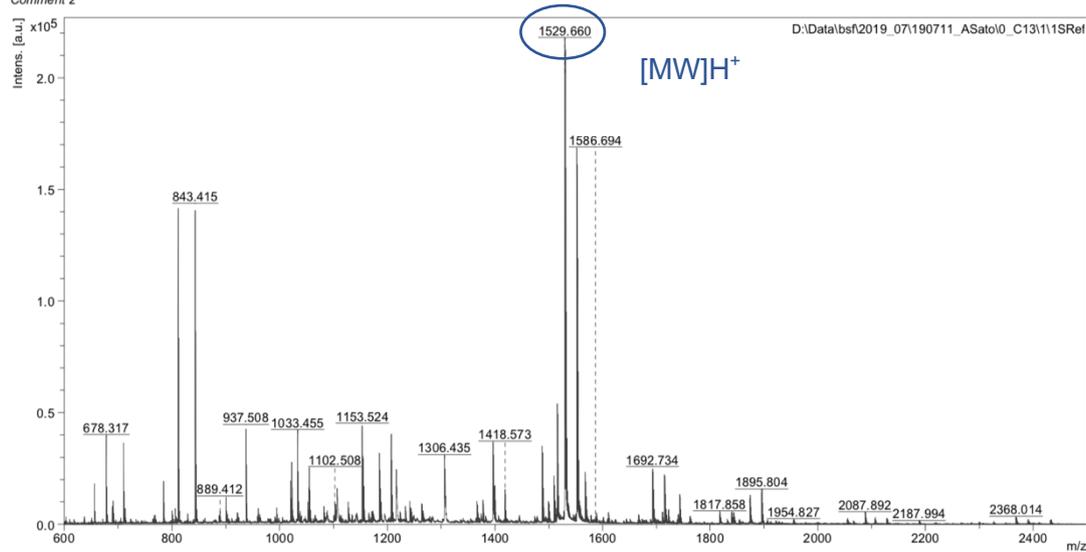
Two solutions (160 μL) were prepared, consisting of peptide (80 μL , 100 μM) and linker (80 μL , 200 μM). TCEP (4 mM) was also added to one of the solutions. TLC plates were run and MALDI-TOF MS data obtained.

3.3 Results and Discussion

A preliminary run of the conjugation was carried out, in which the linker and peptide solutions were combined in a 1:1 ratio (80 μL , 200 μM). The sample was heated at 37°C overnight, but analysis by TLC showed that the conjugation had not worked as no new bands could be seen. For the first attempt, the linker was dissolved in methanol, but the peptide solution was prepared in water. Waltz peptides form aggregates very quickly in water and so it was likely that the peptide had aggregate within itself, before being able to react with the linker.

As a result, the solvent system was changed to an organic solvent. The linker was again dissolved in methanol, but the peptide was prepared in a solution consisting of 1:3 parts HFIP and methanol; the presence of HFIP aids the solubility of the peptide, which is otherwise usually insoluble in methanol. Three samples were formed, each with the same amount of linker (200 μM), but with excess peptide (200 μM , 400 μM , 1000 μM respectively). TLC analysis showed the presence of new bands on the plate, which MALDI-TOF MS data confirmed to be the conjugates; the MW of the linker is 344.15 g mol^{-1} , the peptide has a MW of 1184.56 g mol^{-1} , which resulted in an expected MW of the conjugates at 1528.71 g mol^{-1} .

Comment 1 PL-1
Comment 2

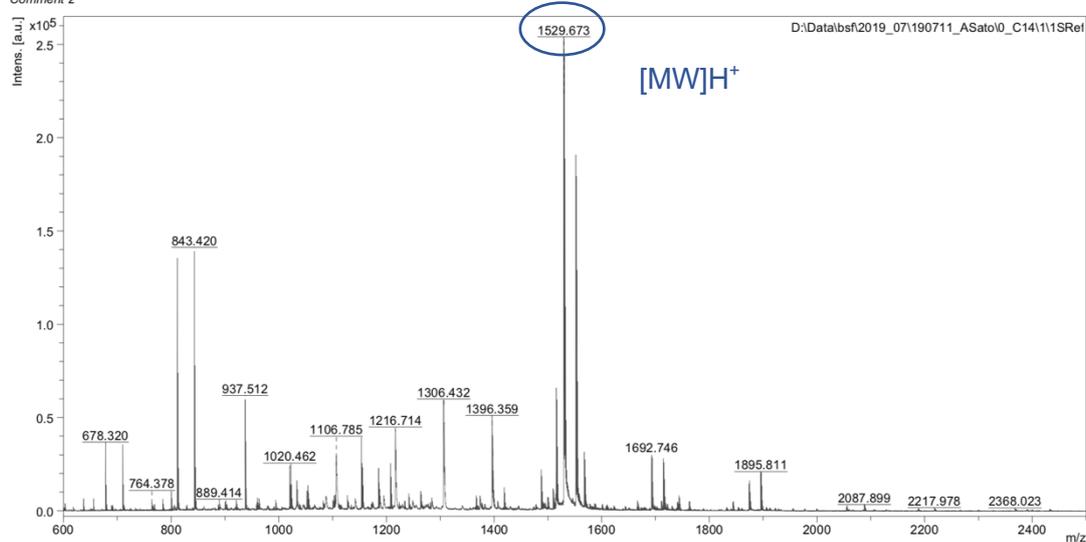


Bruker Daltonics flexAnalysis

printed: 7/11/2019 3:42:12 PM

Figure 30: MALDI-TOF MS data of 1:1 equivalent of peptide to linker

Comment 1 PL-2
Comment 2



Bruker Daltonics flexAnalysis

printed: 7/11/2019 5:00:48 PM

Figure 31: MALDI-TOF MS data of 2:1 equivalents of peptide to linker

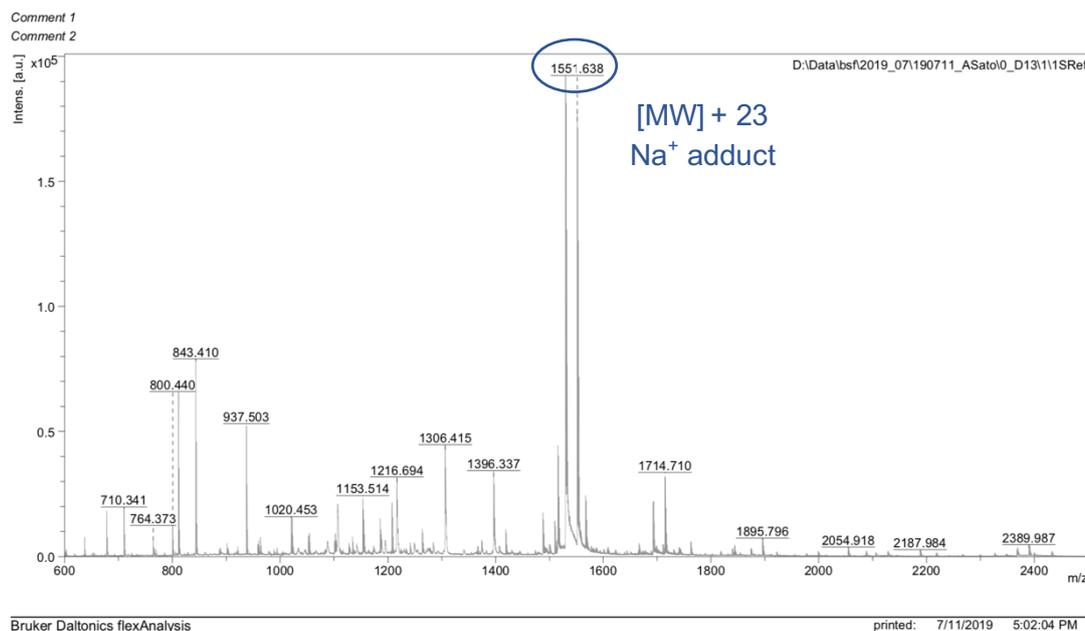


Figure 32: MALDI-TOF MS data of 5:1 equivalents of peptide to linker

From this, the conjugation was repeated, using the same solvent system before because it was shown to work. However, instead of using excess peptide, an excess of linker was used, as it is easier and more economical to produce. Two solutions (160 μL) were produced, with peptide (80 μL , 100 μM) and linker (80 μL , 200 μM); TCEP (4 mM) was also added to one of the solutions in order to reduce the disulphide bond. TLC plates were run again, which positively showed the appearance of new, potential conjugate bands. MS data for this sample displayed a peak in the region expected for a conjugate, shown below in figure 33; however, instead of the peak occurring at 1529.71 g mol^{-1} , which would represent $[\text{MW}]\text{H}^+$, the peak occurred at 1530.64 g mol^{-1} which corresponds to $[\text{MW}]\text{2H}^+$. Nevertheless, the sample that contained TCEP did not yield any visible conjugate bands.

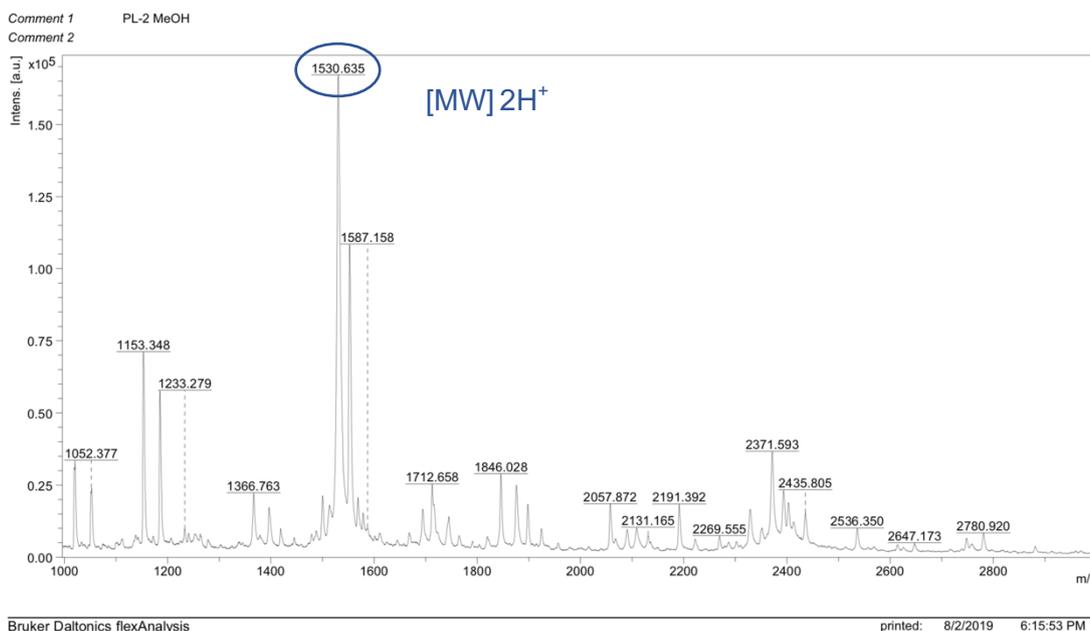


Figure 33: MALDI-TOF MS data of 1:2 equivalents of peptide to linker

In order to further analyse the conjugates, the reaction was scaled up using a mixture of the conditions used before. An organic solvent system remained, but without the presence of TCEP, and an excess of peptide was used as the purification of the conjugate solution would be easier. Three samples (160 μL) were prepared of the linker at a higher concentration (80 μL , 400 μM), combined with the peptide at 1 equivalent (80 μL , 400 μM) and at 2 equivalents (80 μL , 800 μM). After heating in the thermocycler, the TLC plates showed no new bands that could be characteristic of the conjugate.

3.4 Conclusion

On the whole, the conjugation of the Waltz peptide with **linker 1B**, N-6-acetamidopyridinyl-6-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)hexanamide, has produced positive results. Once the optimal conditions for the conjugation had been found, the reactions have been successful, with TLC and MS data confirming the expected conjugates had been made. Although further examination of the conjugates could not be carried out, it is known the conditions for which the conjugation works best and so there is potential for this work to provide interesting results in the future.

4.1 Introduction

This chapter of research focuses on synthesising two slightly different chemical linkers. The objective of these linkers is to provide a molecule that covalently connects a strand of DNA with a cysteine group of an amyloid-forming protein, Sup35NM LCys. When the protein forms amyloid fibrils, the DNA strand of one fibril will be able to crosslink with a DNA strand attached to another fibril by complementary hydrogen bonding. The level of crosslinking could potentially then be regulated by varying the amount of DNA strands bonded to the protein through the chemical linker.

The first linker, as shown on the left in figure 34, used a solution phase strategy to conjugate the linker to the DNA strand; it also contains a maleimide ring in order to covalently bond the linker with the protein through thiol-maleimide chemistry. The second linker, as shown on the right in figure 34, has a phosphoramidite group which allows the linker to be covalently bonded to the desired DNA strands by solid phase synthesis. The linker also contains a maleimide ring, protected by furan.

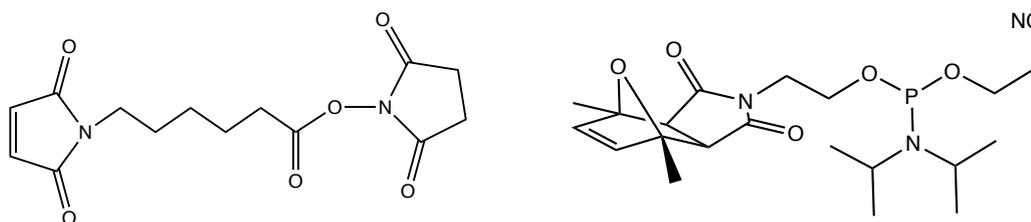


Figure 34: The two chemical linkers
Left: **Linker 2A** (compound **7**) with the maleimide ring
Right: **Linker 2B** (compound **10**) which incorporates the phosphoramidite group and furan ring

Sup35NM LCys continued to be the chosen protein for the conjugation reactions, synthesised by the Kent Fungal Group in the School of Biosciences. Once conjugated with the protein and DNA, the protein-linker-DNA conjugate will be able to assemble into amyloid. Two complementary DNA strands will be used so when assembled, the DNA strands will be able to crosslink to each other, allowing the templation of amyloid fibrils. Once achieved, the level of crosslinking can be varied by differing the number of fibrils in the amyloid that are covalently bonded to the linker and thus a DNA strand.

A similar technique has been used to create peptide-DNA conjugates that reversibly self-assemble into intertwined fibrils and bigger structures with fibrillar networks. Peptide-DNA structures were produced from nanofibers, formed from alkylated peptides, covalently bonded to an oligonucleotide. A peptide-DNA motif, with the complementary DNA sequence, was then reacted with this structure, which lead to the DNA strands cross-linking. This led to a hydrogel being created, whose formation could be reversed and broken down into the single-stranded DNA-peptide components; this was achieved either by the addition of an 'invader' oligonucleotide or by heating at an elevated temperature.⁷²

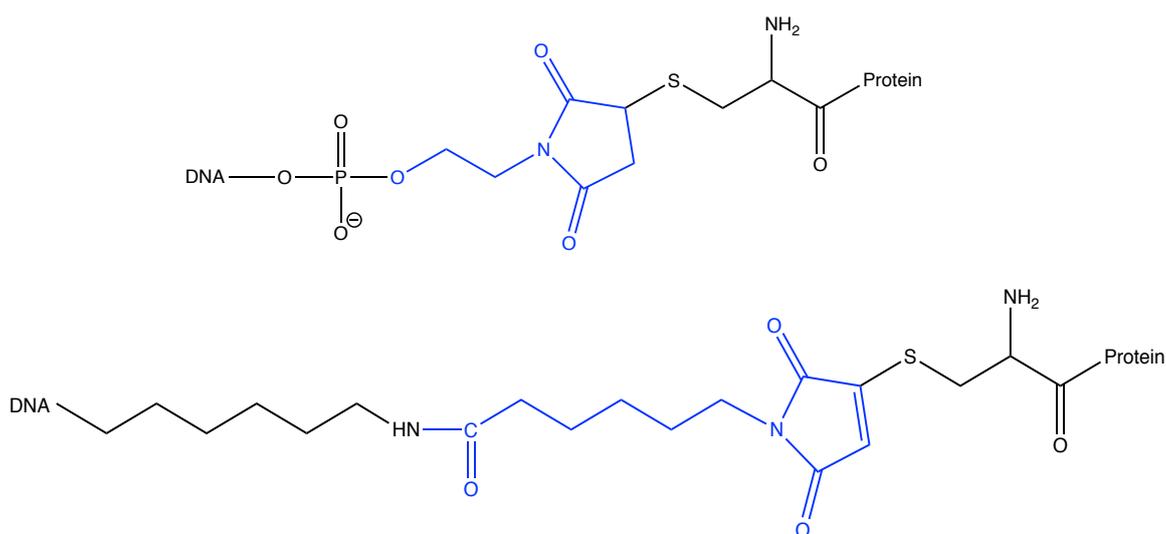


Figure 35: The DNA-Linker-Protein conjugate

Above: With compound **7** as the linker

Below: With compound **10** as the linker

4.2 Experimental

4.2.1 Synthesis of DNA

The DNA strands were synthesised on an Expedite 8909 DNA synthesiser by phosphoramidite chemistry that have been used previously in literature and within the group⁷³, as shown in figure 36.

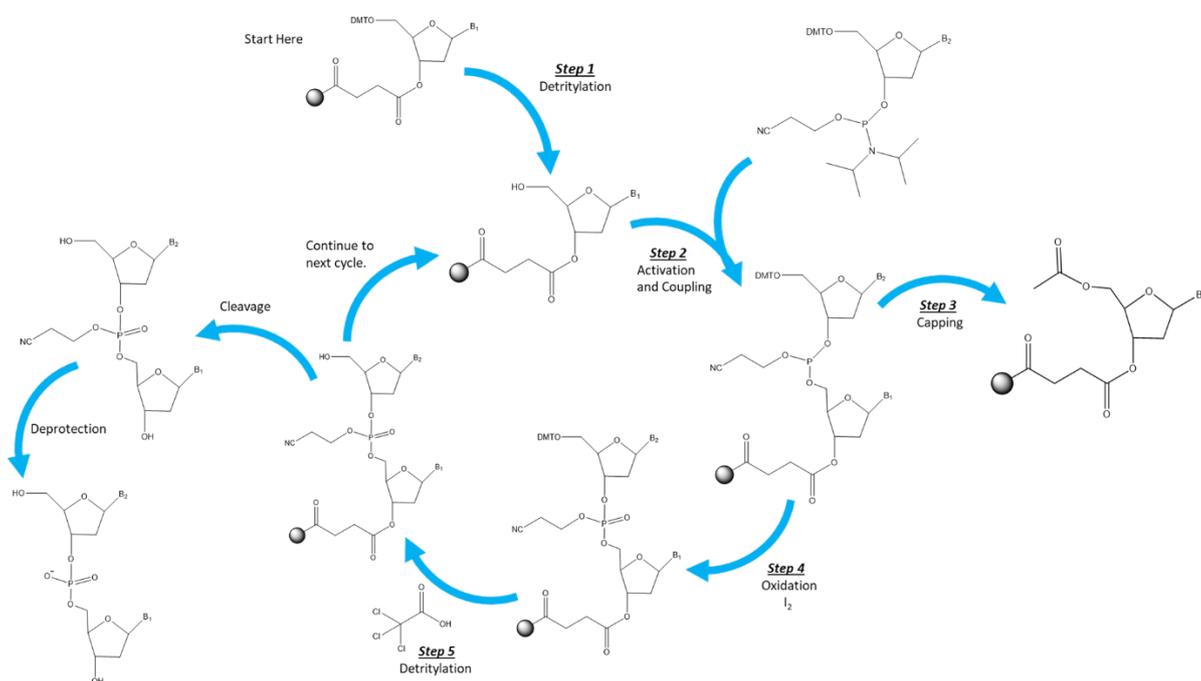


Figure 36: The synthesis of DNA by the phosphoramidite method

Two complementary DNA strands, with a spacer of 5 thymine bases, were made on the synthesiser. The strands required were,

DNA 1: TT TTT GTC AAC GGT TAT GTC

DNA 2: TT TTT GAC ATA CCA GTT GAC

A 5 base thymine spacer was used to add flexibility to the DNA strands as it is not involved in hybridisation. It eventually would increase the space between the DNA and protein, leading to less interference in the self-assembly of the DNA-linker-protein strands. These

DNA strands were used as they are complementary to each other, but do not self-dimerise, form any secondary structures, such as hairpins or loops, and show no secondary folding. DNA 1 has a melting temperature of 48.8 °C, whilst DNA 2 has a melting temperature of 48.2 °C.

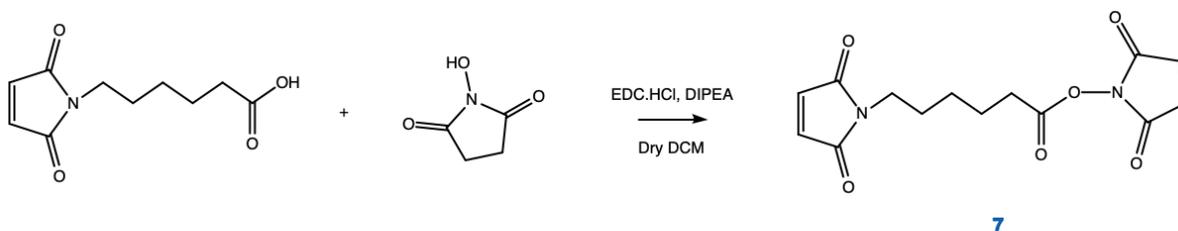
A gas leak test was carried out on the synthesiser and the lines were primed with the reagents before the desired sequence and settings were entered onto the programme. The column, containing the glass beads onto which the DNA was to be made, was placed into the synthesiser and the sequence started. The synthesis was monitored by measuring the optic absorption of the trityl cation at each detritylation step.

Once completed, the columns were removed from the synthesiser and placed in the freezer until needed.

4.2.2 Synthesis of Linkers

6-Maleimidohexanoic Acid N-Hydroxysuccinimide Ester Linker

6-Maleimidohexanoic Acid N-Hydroxysuccinimide Ester (7)



Scheme 7: NHS esterification of 6-maleimidohexanoic acid

Anhydrous DCM (100 mL) was added to a flask containing 6-maleimidohexanoic acid (0.45 g, 2.017 mmol, 1 mol equiv), *N*-hydroxysuccinimide (0.70 g, 6.108 mmol, 3 mol equiv) and EDC-HCl (1.83 g, 9.557 mmol, 4.8 mol equiv) and the reaction stirred under a nitrogen atmosphere overnight. DIPEA (0.68 g, 5.251 mmol, 2.6 mol equiv) was added and the

solution stirred for a further 2 hours. The solution was washed with hydrochloric acid (1 M), saturated potassium carbonate and brine. Magnesium sulphate was used to dry the organic layer and the solvent was removed by rotary evaporation. The product was collected as a pale yellow oil (0.609 g, 1.976 mmol, 98 % yield).

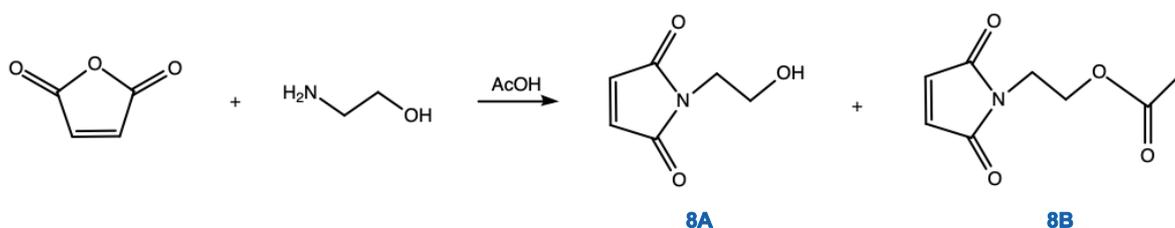
^1H NMR (400 MHz, CDCl_3), ppm: 6.63 (2H, s, ArH), 3.46 (2H, t, $J = 7.2$ Hz, CH_2), 2.78 (2H, s, CH_2), .54 (2H, t, $J = 7.1$ Hz, CH_2), 1.71 (2H, m, $J = 7.6$ Hz, CH_2), 1.56 (2H, m, $J = 7.4$ Hz, CH_2), 1.34 (2H, m, $J = 7.8$ Hz, CH_2)

^{13}C NMR (400 MHz, CDCl_3), ppm: 170.87, 169.17, 168.41, 134.09, 35.25, 30.78, 28.07, 25.82, 25.60, 24.07

MS (ESI, positive mode): calculated $[\text{MW}]\text{H}^+ = 309.1$, m/z found = 331.1 (Na^+ adduct)

2-Cyanoethyl 2-((3aR,4S,7aS)-4,7-Dimethyl-1,3-Dioxo-1,3,3a,4,7,7a-Hexahydro-2H-4,7-Epoxyisoindol-2-yl)ethyl) Diisopropylphosphoramidite Linker

1-(2-Hydroxyethyl)-1H-Pyrrole-2,5-Dione (**8**)⁷³



Scheme 8: Synthesis of compound **8**

Ethanolamine (0.85 g, 0.0139 mol, 1 mol equiv) and maleic anhydride (1.53 g, 0.016 mol, 1.1 mol equiv) were dissolved in acetic acid (85 mL). The solution was heated to 180 °C and refluxed overnight. Once cooled to room temperature the solvent was removed by rotary evaporation, producing an oil. DCM (20 mL) was then added and the solution washed with hydrochloric acid (1 M), saturated potassium carbonate and brine. Magnesium sulphate was used to dry the organic layer and the solvent was removed by rotary

evaporation. The product was collected as a light brown solid (1.349 g, 7.364 mmol, 53 % yield).

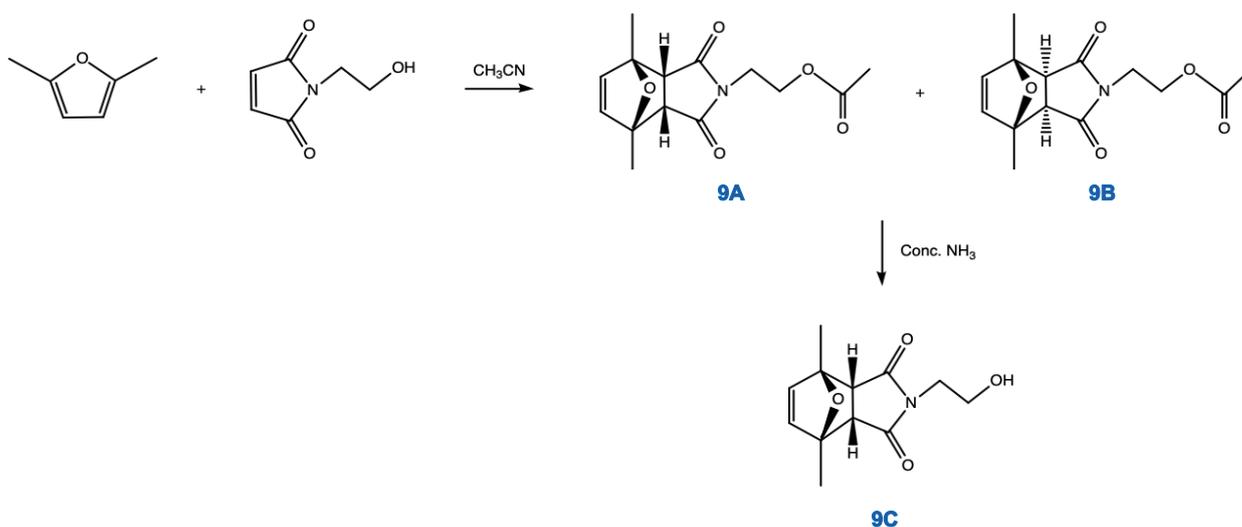
^1H NMR (400 MHz, CDCl_3), ppm: 6.66 (2H, s, ArH), 4.16 (2H, t, $J = 5.3$ Hz, CH_2),

3.73 (2H, t, $J = 5.3$ Hz, CH_2), 1.95 (3H, s, CH_3)

^{13}C NMR (400 MHz, CDCl_3), ppm: 170.44, 134.24, 61.43, 36.93, 20.74

MS (ESI, positive mode): calculated $[\text{MW}]\text{H}^+ = 142.0$, m/z found = 142.3

*2,5-Dimethylfuran-Maleimide Exo Adduct (9)*⁷³

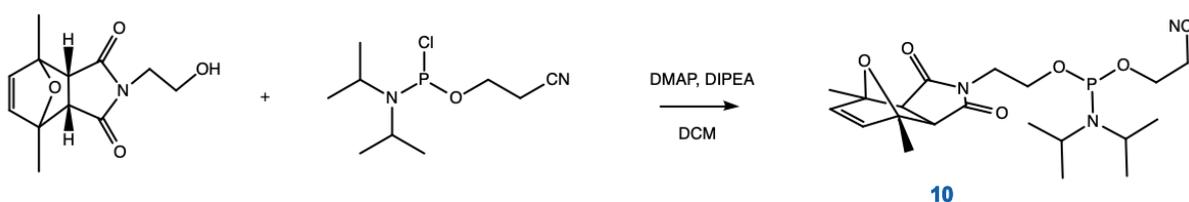


Scheme 9: Furan protection of compound **8B** by Diels-Alder chemistry

Compound **8B** (0.75 g, 4.099 mmol, 1 mol equiv) was dissolved in dry acetonitrile (11.5 mL) and 2,5-dimethylfuran (3.0 mL, 6.8 mol equiv) added. The solution was heated to 65 °C under a nitrogen atmosphere and left to stir for 6 hours. A ^1H NMR spectrum was taken to ensure the reaction had gone to completion. The solvent was then removed by rotary evaporation and concentrated ammonium hydroxide (11.5 mL) was then added directly to the product and the solution left to stir overnight. The solvent was removed by rotary evaporation and the resulting oil acidified with trifluoroacetic acid to pH 1-2. The product was extracted with DCM and magnesium sulphate used to dry the organic layer. The solvent was removed by rotary evaporation and the product collected as a brown liquid (1.247 g, 5.255 mmol, 128 % apparent yield).

^1H NMR (400 MHz, CDCl_3), ppm: 6.24 (2H, s, ArH), 3.68 (2H, m, $J = 1.8$ Hz, CH_2),
 3.64 (2H, m, $J = 1.8$ Hz, CH_2), 2.85 (2H, s, ArH), 2.69 (1H, s, OH), 1.62 (6H, s, CH_3)
 ^{13}C NMR (400 MHz, CDCl_3), ppm: 170.85, 134.08, 47.06, 38.62, 37.56, 26.22
 MS (ESI, positive mode): calculated $[\text{MW}]\text{H}^+ = 238.1$, m/z found = 238.2, 260.2 (Na^+
 adduct)

2-Cyanoethyl 2-((3aR,4S,7aS)-4,7-Dimethyl-1,3-Dioxo-1,3,3a,4,7,7a-Hexahydro-2H-4,7-Epoxyisoindol-2-yl)ethyl) Diisopropylphosphoramidite (**10**)⁷⁴



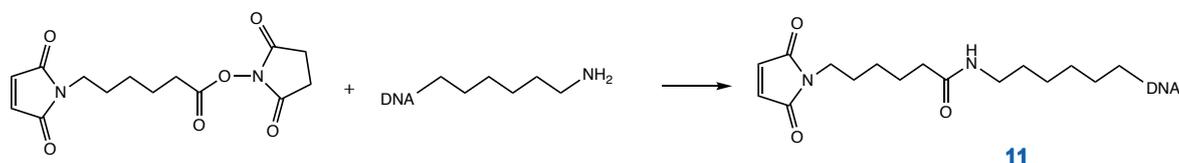
Scheme 10: Phosphoramidite synthesis of compound **9C**

Compound **9C** (0.06 g, 2.360×10^{-4} mol, 1 mol equiv) was dissolved in dry DCM (2 mL). 2-cyanoethyl phosphoramidite chloride (0.06 g, 2.360×10^{-4} mol, 1 mol equiv), *N,N*-diisopropylethylamine (0.15 g, 1.180 mmol, 5 mol equiv), DMAP (a few grains) and DCM (5 mL) were added and the reaction stirred under nitrogen for 150 minutes. The solution was washed with degassed saturated potassium carbonate and brine. A small silica column was prepared and the reaction solution loaded. Two fractions were collected using a solvent of DCM : triethylamine (95:5). The fractions were combined and the solvent was removed by rotary evaporation (0.097 g, 0.2375 mmol, 101 % apparent yield).

^{31}P NMR (400 MHz, CDCl_3), ppm: 147.28, 15.16, 11.40

4.2.3 Conjugation of Linker and DNA

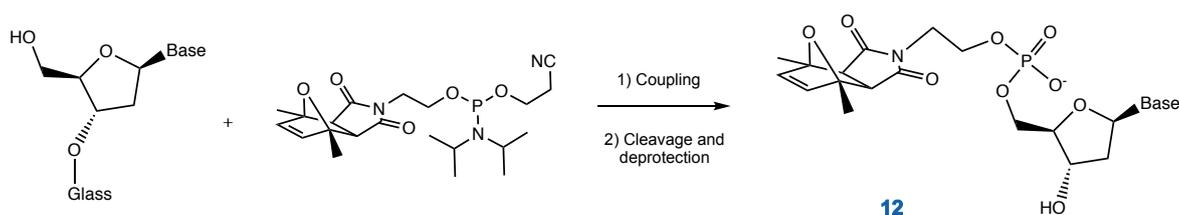
6-Maleimidohexanoic Acid N-Hydroxysuccinimide Ester Conjugation with DNA (**11**)



Scheme 11: Conjugation of compound **7** with DNA

Compound **7** (0.6 mg, 1.947 mmol) was dissolved in DMF (30 μ L). Stock solutions of DNA 1 and DNA 2 (100 μ M, 1 ml autoclaved water) were prepared and DMF-linker solution (15 μ L) was added to each of the prepared DNA stock solutions. The solutions were incubated at 37 $^{\circ}$ C for 90 minutes. The two solutions were then purified by size-exclusion chromatography ('desalting') using Zetadex resin (8 mL column) and eluted with autoclaved water. A total of 10 fractions were collected for each DNA strand and the UV-Vis absorbance measured on the NanoDrop spectrophotometer.

Phosphoramidite Maleimide Alcohol Conjugation with DNA (**12**)



Scheme 12: Conjugation of compound **10** with DNA

The DNA strands to be used were placed in the DNA synthesiser and washed with deblock, before the product (compound **10**) was hand-coupled to the two strands of pre-synthesised DNA. Dry acetonitrile (1 mL) and an activator, 5-ethylthio-1H-tetrazole in acetonitrile (1 mL), were added to the phosphoramidite product and the resulting solution split between the two DNA strands. After 30 minutes, the phosphoramidite solutions were removed from the two

DNA columns, which were then washed with acetonitrile and oxidiser. The solid support beads were removed from the two DNA columns and placed into corresponding vials. Concentrated ammonium hydroxide solution (1 mL) was added and the solution heated at 55 °C overnight, to remove the DNA-linker from the solid support beads. Once cooled to room temperature, the samples were placed into the centrifuge and spun for 30 minutes at 2000 x G. The supernatant was removed and purified by size-exclusion chromatography ('desalting') using a Zetadex column. The fractions produced were then dried by centrifugal vacuum concentration at 65 °C for 4 hours before being re-suspended into autoclaved water (1000 µL) and quantified by measuring optical absorbance at 260 nm.

4.2.4 Conjugation of DNA-Linker and Protein

6-Maleimidohexanoic Acid N-Hydroxysuccinimide Ester-DNA Conjugation with Protein

Sup35NM LCys protein in guanidine was thawed on ice, DTT (0.1 mL) added and left for 2 hours. A PD10 column was then run, using an Elution G buffer; the fluorescence of all the fractions was measured. Fractions 1 – 6 were combined, the absorbance re-measured and the protein concentration determined. The protein was split into two samples (1.2 mL each). Linker-DNA1 (93.0 µL) was added to one of the samples and linker-DNA2 (58.9 µL) was added to the other. The samples were left for 2 hours above ice. Separately, the protein (50 µL) was added to PEG-Maleimide (5 µL) and also to a fluorescent chloroacetamide (50 µL). LCys-DNA1 (100 µL), LCys-DNA2 (100 µL), a third sample of LCys-DNA1 (50 µL) and LCys-DNA2 (50 µL) combined, LCys-PEG-Maleimide (55 µL) and LCys-Chloroacetamide (100 µL) were all precipitated by ethanol precipitation. The samples were suspended in 9 volumes of cold, 100 % ethanol and reacted at -20 °C for 1 hour. After being centrifuged (10 minutes, 15000 x G), the pellet was dried and washed with cold ethanol. The ethanol was then aspirated, the sample dried with nitrogen and re-suspended in 8 M, fresh urea. All 5 samples were then loaded into a SDS-PAGE gel (12 % polyacrylamide) and run for 2 hours. The gel was imaged on the transilluminator to see the fluorescent chloroacetamide before being stained with SYBR safe (10,000 x dilution of

concentrate into 30 mL of SDS-PAGE running buffer) for 30 minutes. After imaging and washing in distilled water, Instant Blue was used to stain the gel for 20 minutes, before it was imaged. The samples were left at 2 °C for 48 hours and precipitated again via the above ethanol precipitation protocol. All 5 samples were then run on an SDS-PAGE gel (12 % polyacrylamide) in the same conditions and with the same staining procedure previously used.

4.3 Results and Discussion

4.3.1 Synthesis of DNA

The first attempt to synthesise the desired DNA strands failed twice due to a cytosine base failing to couple successfully; this could be seen in the trityl monitor graph.



Figure 37: Trityl of the first synthesis of DNA1



Figure 38: Trityl of the first synthesis of DNA2

As a result, the cytosine base was replaced with a fresh bottle of cytosine at a higher concentration (1 g dissolved in 10 mL of dry acetonitrile) and fresh bottles of activator and oxidant were also used. The synthesis was re-attempted, with no failures and a much better trityl growth; the concentration was also better, indicating the synthesis had been more successful.



Figure 39: Trityl of the second synthesis of DNA1



Figure 40: Trityl of the second synthesis of DNA2

4.3.2 Synthesis of Linkers

6-Maleimidohexanoic Acid *N*-Hydroxysuccinimide Ester Linker

6-Maleimidohexanoic Acid *N*-Hydroxysuccinimide Ester

This reaction has been carried out twice, using a method adapted from literature⁷⁵, with the product successfully made on both attempts. However, the first reaction produced a low yield (30 %) and so the subsequent reaction was slightly altered.

After stirring overnight under a nitrogen atmosphere, the reaction solution still contained some starting material. As a result, DIPEA was added to the reaction between 6-maleimidohexanoic acid and *N*-hydroxysuccinimide and stirring continued for a further hour; this pushed the reaction to completion and thus a much higher yield (98 %) was produced.

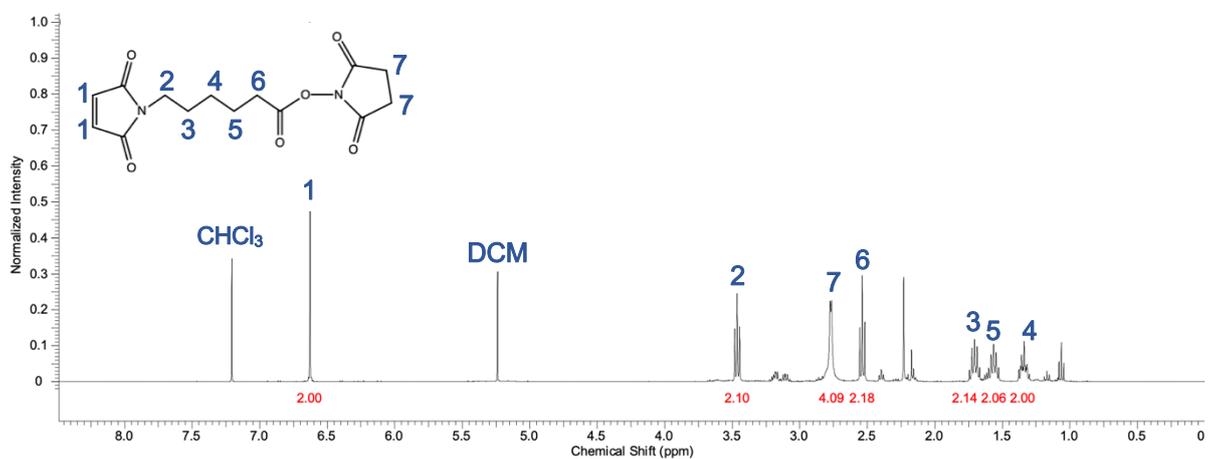


Figure 41: ¹H NMR of the NHS esterification of 6-maleimidohexanoic acid (compound 7)

2-Cyanoethyl 2-((3aR,4S,7aS)-4,7-Dimethyl-1,3-Dioxo-1,3,3a,4,7,7a-Hexahydro-2H-4,7-Epoxyisoindol-2-yl)ethyl) Diisopropylphosphoramidite Linker

1-(2-Hydroxyethyl)-1H-Pyrrole-2,5-Dione

The first attempt at synthesis followed a literature example⁷³ and was carried out at 170 °C and yielded a white, crystalline solid. However, the yield of this reaction was very low (18 %) and so subsequent reactions were carried out at 180 °C to increase the yield. The NMR spectrum of the product also contained a peak at around 2 ppm, which was determined to be part of the structure and not an impurity; it is the addition of an acetyl group onto the alcohol.

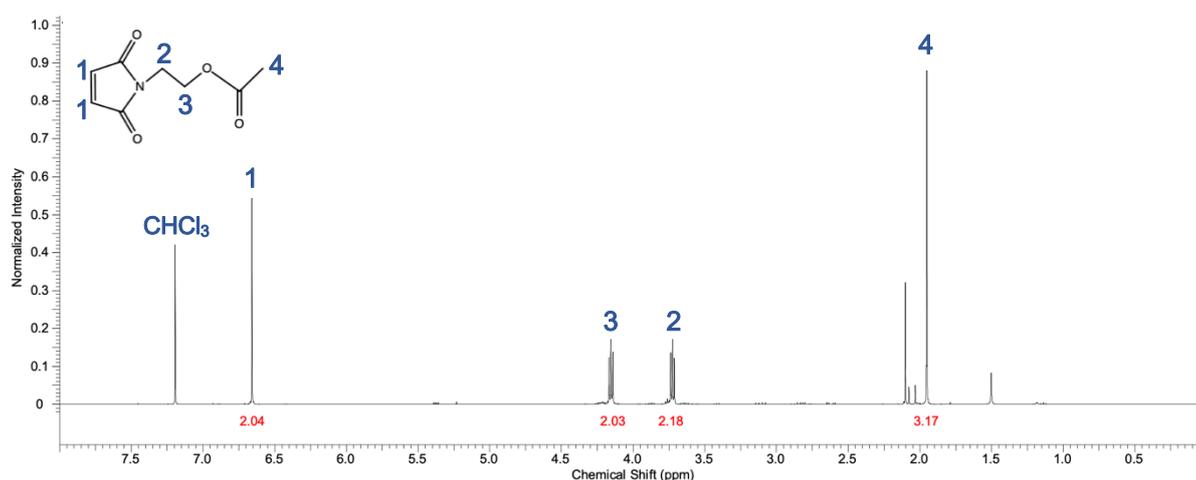


Figure 42: ¹H NMR of the synthesis of maleimide alcohol (compound **8**)

2,5-Dimethylfuran-Maleimide Exo Adduct

The first attempt of this reaction used a reaction temperature of 65 °C and stirred the reaction mixture under nitrogen overnight. However, the flask was submerged too far into the oil bath, which caused the solvent to evaporate and thus the reaction did not proceed to completion as the starting materials solidified; this was confirmed by the NMR spectrum obtained.

The reaction was then repeated and followed by TLC, which showed that the reaction had reached completion in 6 hours. The product was obtained as a yellow oil, containing both the *exo* and *endo* forms of the structure. Subsequent reaction conditions will use concentrated ammonium hydroxide and so only the product with the highest stability is desired. The *endo* form of the product has more steric hindrance than the *exo* form, caused by the alkyl chain of the adjoining ring and so in order to degrade the *endo* product, leaving the more stable *exo* product, the compound was stirred in concentrated ammonium hydroxide overnight. The solution was then washed with hydrochloric acid, potassium carbonate and brine, with the product collected as a clear oil residue, although in a low yield.

A third and fourth attempt at the reaction used a different work-up method, as found in literature.⁷³ Once the *exo* and *endo* forms of the product had been produced and stirred with concentrated ammonium hydroxide solution, the ammonium was removed by rotary evaporation. Trifluoroacetic acid was used to acidify the solution and the product then extracted with dichloromethane.

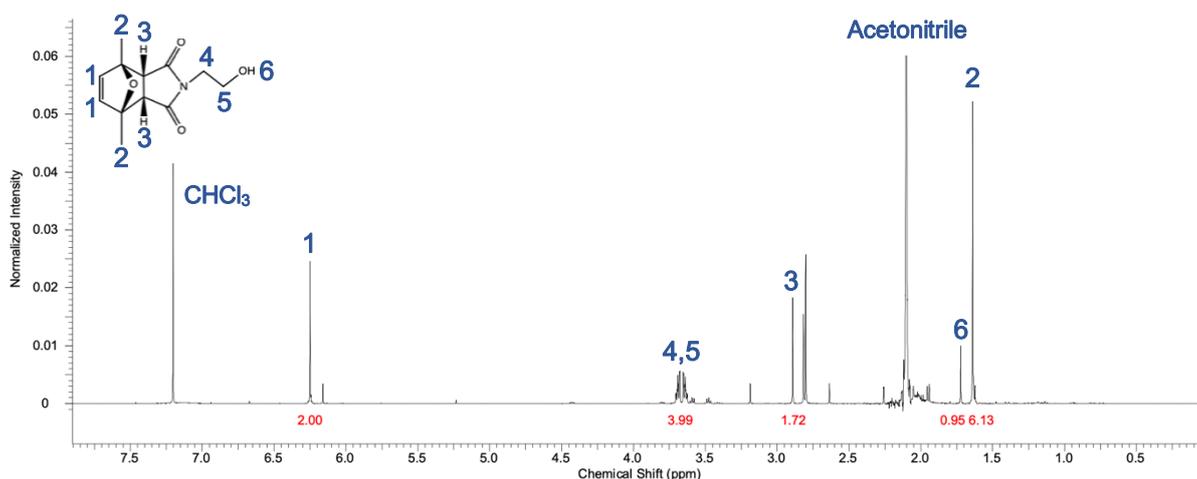


Figure 43: ¹H NMR of the furan protected maleimide alcohol Diels-Alder reaction (compound 9)

2-Cyanoethyl 2-((3aR,4S,7aS)-4,7-Dimethyl-1,3-Dioxo-1,3,3a,4,7,7a-Hexahydro-2H-4,7-Epoxyisoindol-2-yl)ethyl) Diisopropylphosphoramidite

This reaction used a method from literature⁷⁴, that had been adapted by another member of the group. It was found that the ratio of maleimide alcohol : 2-cyanoethyl phosphoramidite chloride : DIPEA as stated in the literature (1 : 1 : 5) worked well but a different work-up protocol was needed. The phosphorous is prone to oxidation so in order to reduce the chance of oxidation, only an aqueous wash was carried out in the hope that the product would be purified enough, but an adequate proportion of the phosphorous would remain unoxidised. Every solvent used in the washes were also de-gassed. A small silica column was then run, eluted with dichloromethane and triethylamine (95 : 5); triethylamine was used in a small quantity to help avoid oxidation of the phosphorous and to add basicity to the acidic silica used in the column.

The ³¹P NMR spectrum showed that although some of the phosphorus had oxidised, a large proportion of it remained unoxidised and so the product was used in subsequent steps; the linker would be in an excess when coupled to the DNA and so the coupling should still be effective.

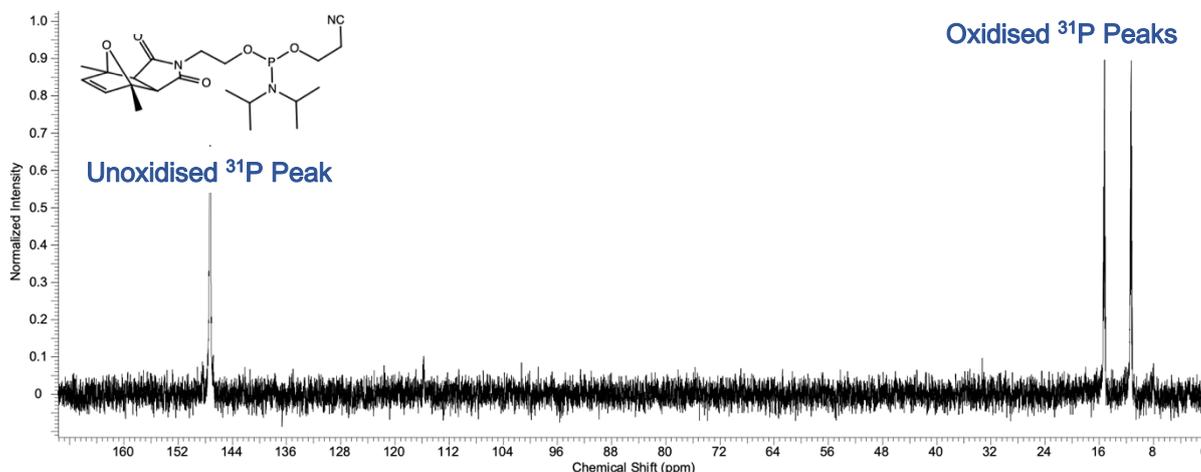


Figure 44: ³¹P NMR of the phosphoramidite maleimide alcohol synthesis (compound 10)

4.3.3 Conjugation of Linker and DNA

6-Maleimidohexanoic Acid N-Hydroxysuccinimide Ester-DNA Conjugation with DNA

The first three attempts at conjugating the NHS ester to the DNA strands resulted in samples that were very concentrated with DNA and linker, resulting in excess linker precipitating out of the fractions after they had been purified by size-exclusion chromatography ('desalting'); de-salting removes unwanted ions and small molecules. The fractions were centrifuged (2000 x G for 5 minutes), the supernatant removed and re-suspended into deionised water (1 mL). After storage in the freezer, more excess linker precipitated from the solution of some of the fraction, so these fractions were washed with ethyl acetate (500 μ L, three times each), before being purified using the previous protocol.

MS data showed that the conjugations had been unsuccessful as the only peaks visible in the spectra were those that corresponded to the DNA strand on its own and not the DNA-linker conjugate. DNA 1 has a MW of 6283.2 Da, DNA 2 has a MW of 6261.2 Da and the linker was a MW of 194.1; this should have produced DNA1-linker conjugates with MW 6477.3 Da and DNA2-linker conjugates of MW 6455.3 DA.

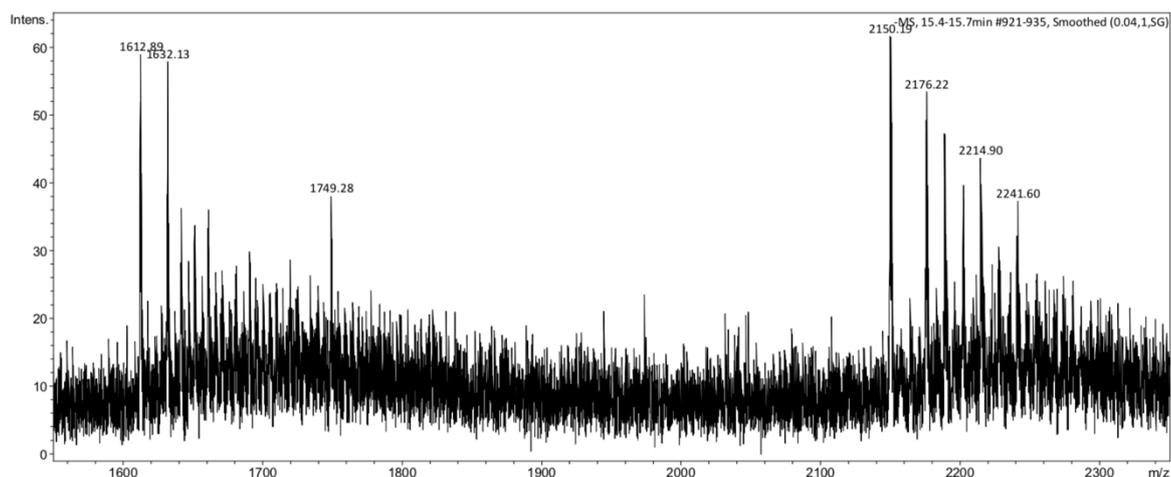
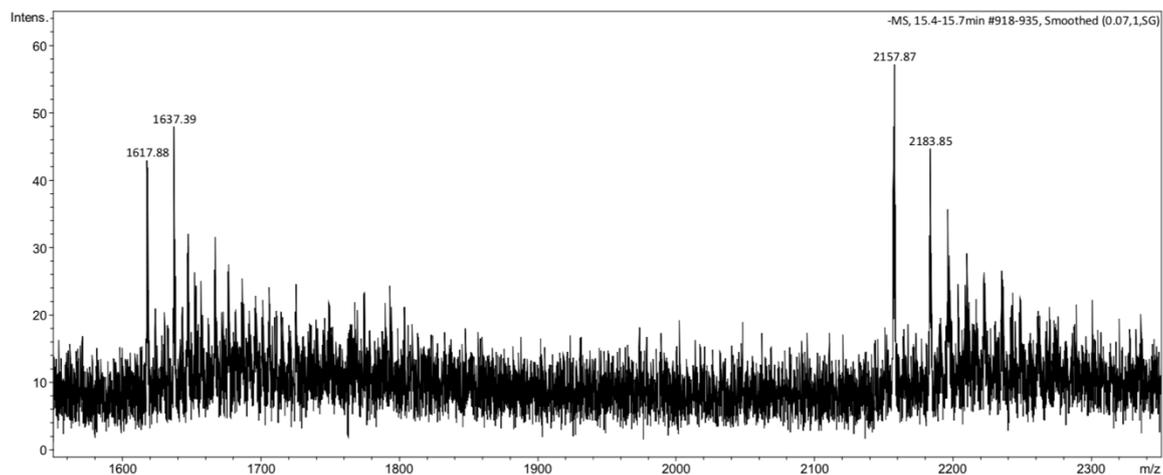


Figure 45: MALDI-TOF MS data of the linker-DNA conjugates
 Above: DNA1-Linker
 Below: DNA2-Linker

The UV-Vis spectrum of DNA1 showed fractions 3 and 4 did contain DNA as there was clear absorbance at A260, with fraction 4 containing the highest concentration of DNA. Fraction 5 did contain some DNA but was also of a lower purity as there were some contaminants around A235 and A300; the peak at A235 is likely to have been caused by other organic compounds and salts. The spectrum of DNA2 displayed the same results as the spectrum of DNA1, as shown in figure 46 below.

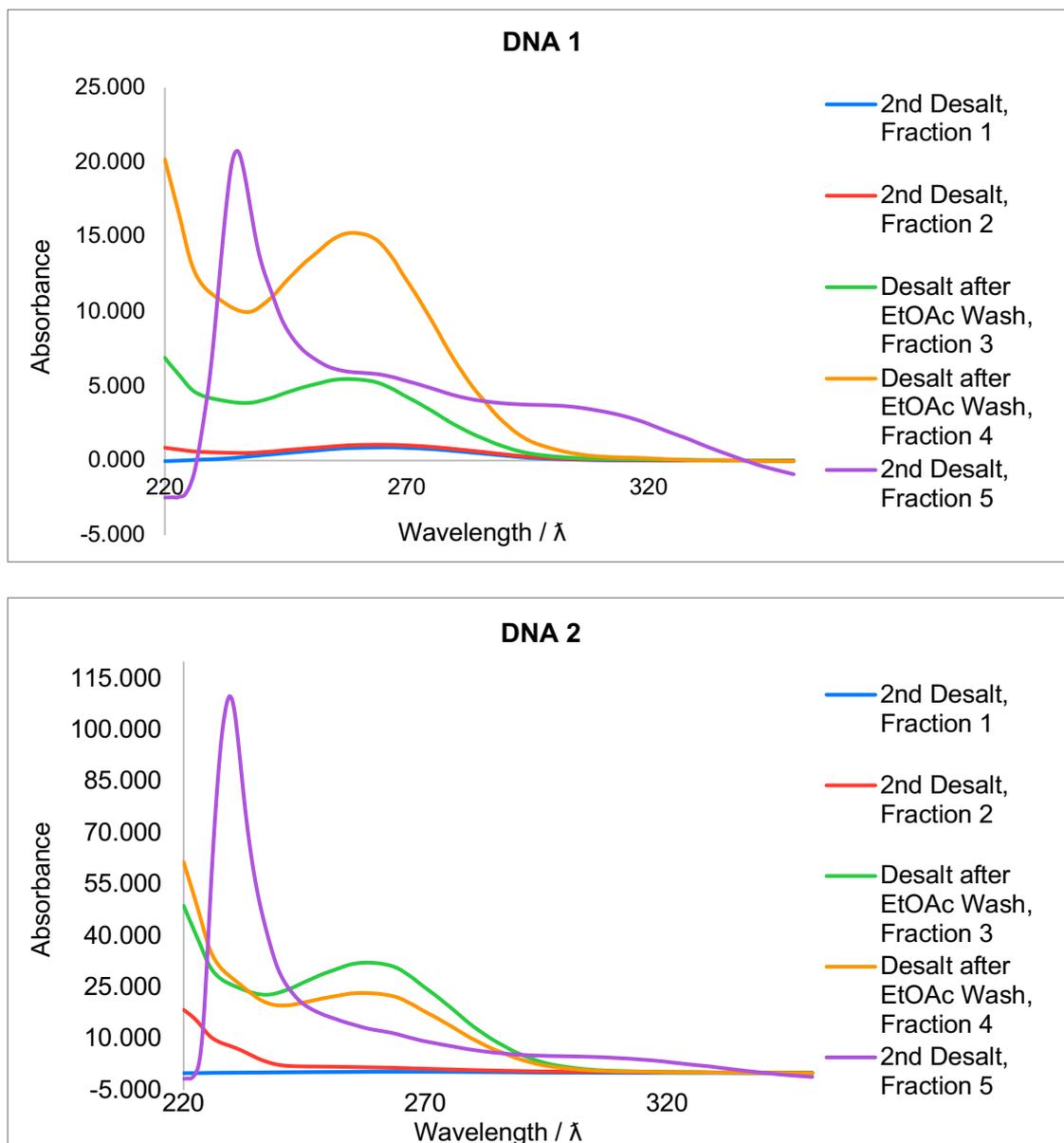


Figure 46: UV-Vis spectra of the linker-DNA conjugates from the first conjugation attempt (compound **11**)
 Above: DNA1-Linker
 Below: DNA2-Linker

It was later realised that the mass of linker used in the conjugation attempts had been far too high, with a 1000 times excess being used. The mass of linker was reduced to 0.6 mg, which produced cleaner UV-Vis data and concentrations that were in the expected range for the latter conjugations. From the UV-Vis spectra, fraction 4 contained the highest concentration of DNA for DNA1, whilst for DNA2 fraction 3 contained the highest concentration. All of the fractions for both DNA1 and DNA2, exhibited moderate purity as there were no other, unexpected peaks visible in the spectrum, displayed in figure 47.

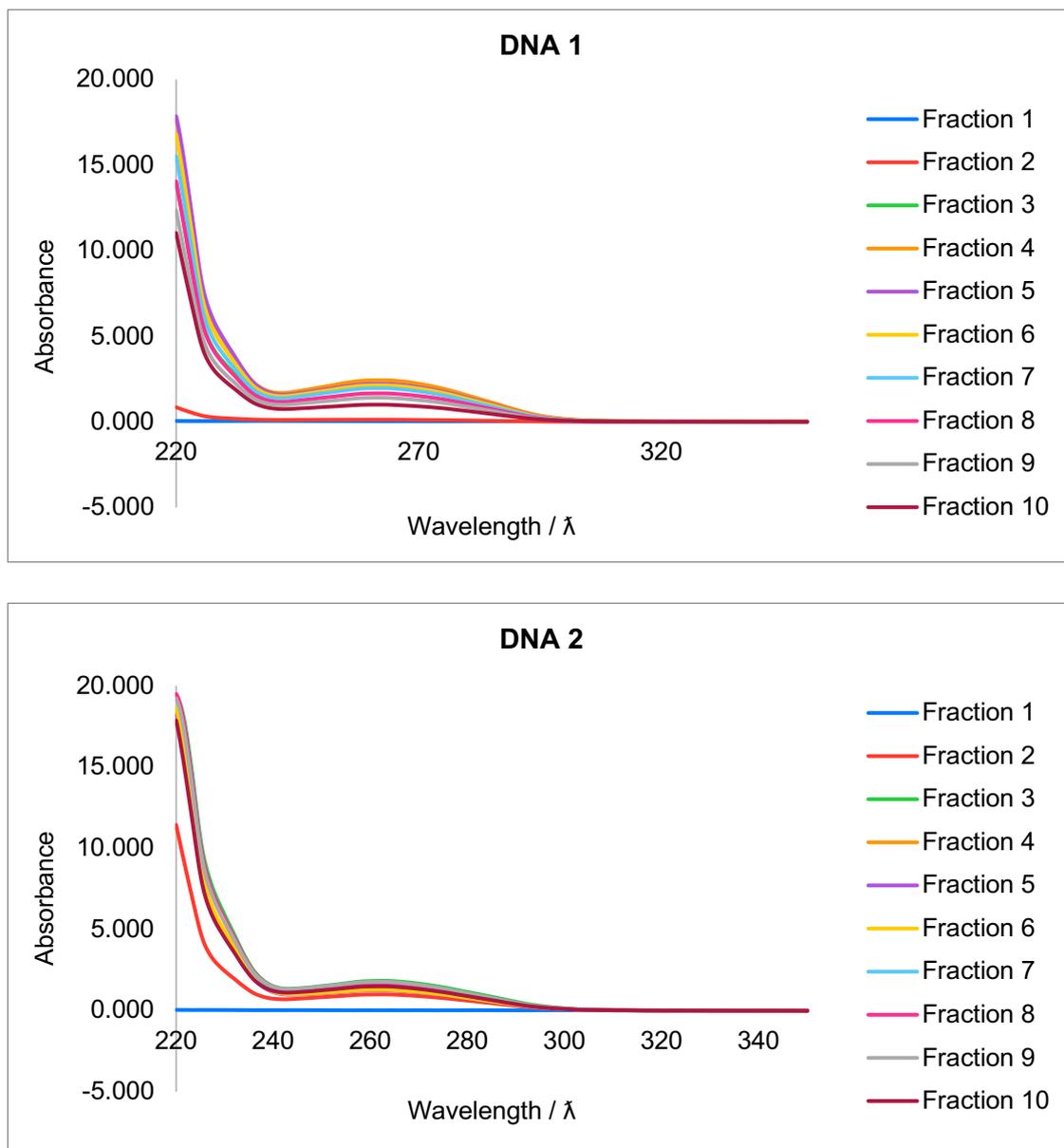


Figure 47: UV-Vis spectra of the linker-DNA conjugates from the latest conjugation attempt (compound **11**)
 Above: DNA1-Linker
 Below: DNA2-Linker

As there was a delay in obtaining some MS data confirming whether the linker had conjugated with the DNA, the Kaiser test was also carried out on the latter reaction attempts. A positive result, in which the solution turns blue, would have indicated a free amine was present but when carried out, the solution remained colourless. This indicated that the amine had been functionalised with, hopefully, the linker.

Phosphoramidite Maleimide Alcohol Conjugation with DNA

The conjugation of the phosphoramidite maleimide alcohol with DNA has been attempted twice, with both attempts being unsuccessful. The UV-Vis data collected indicated fractions 3 and 4 contained high concentrations of DNA for DNA1. Although the spectrum implied these fractions had reasonable purity, fraction 5 could have been contaminated as a change in the expected spectral shape indicates contaminants. For DNA2, fractions 3, 4 and 5 all displayed a high, pure concentration of DNA, as specified in figure 48.

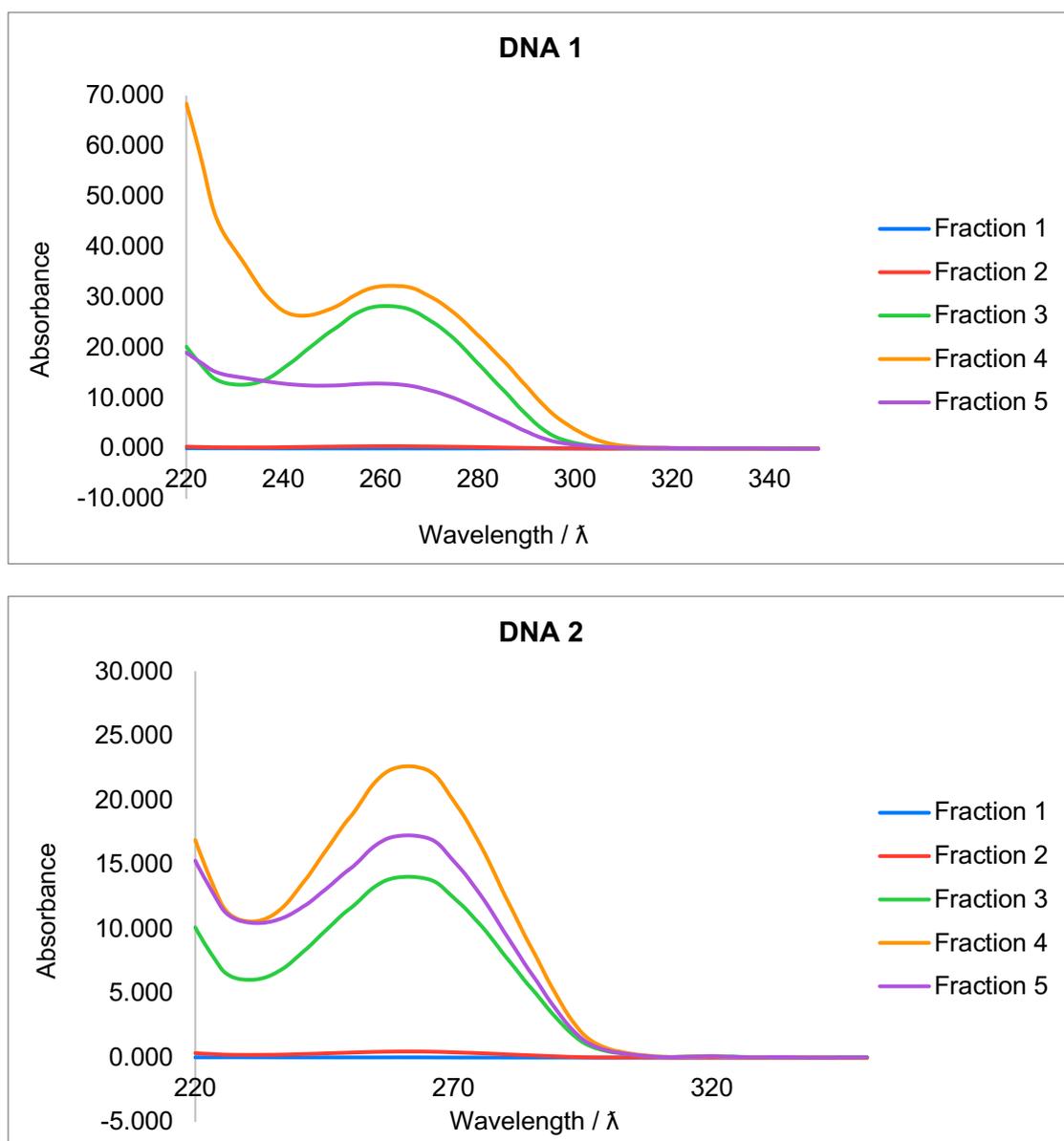


Figure 48: UV-Vis spectra of the linker-DNA conjugates (compound 12)

Above: DNA1-Linker

Below: DNA2-Linker

The DNA 1 strand has a MW of 6101.03 Da and DNA 2 has a MW of 6079.05 Da. The linker would have a MW of 315.15 Da, resulting in a DNA1-linker conjugate with a MW of 6414.08 Da and a DNA2-linker conjugate with a MW of 6394.10 Da. However, the MS data showed that the linker had not attached to the DNA and only the DNA remained in the solution. This could be due to a number of differing reasons. Firstly, the phosphorous would have continued to oxidise whilst the ^{31}P NMR spectrum was obtained and during the time when the DNA strands were washed with acetonitrile and deblock; these washes were carried out in order to remove the 4,4'-dimethoxytrityl protecting group on the 5'-hydroxyl of the DNA strands. There may also have been issues with the columns containing the DNA strands as they had been stored in the freezer.

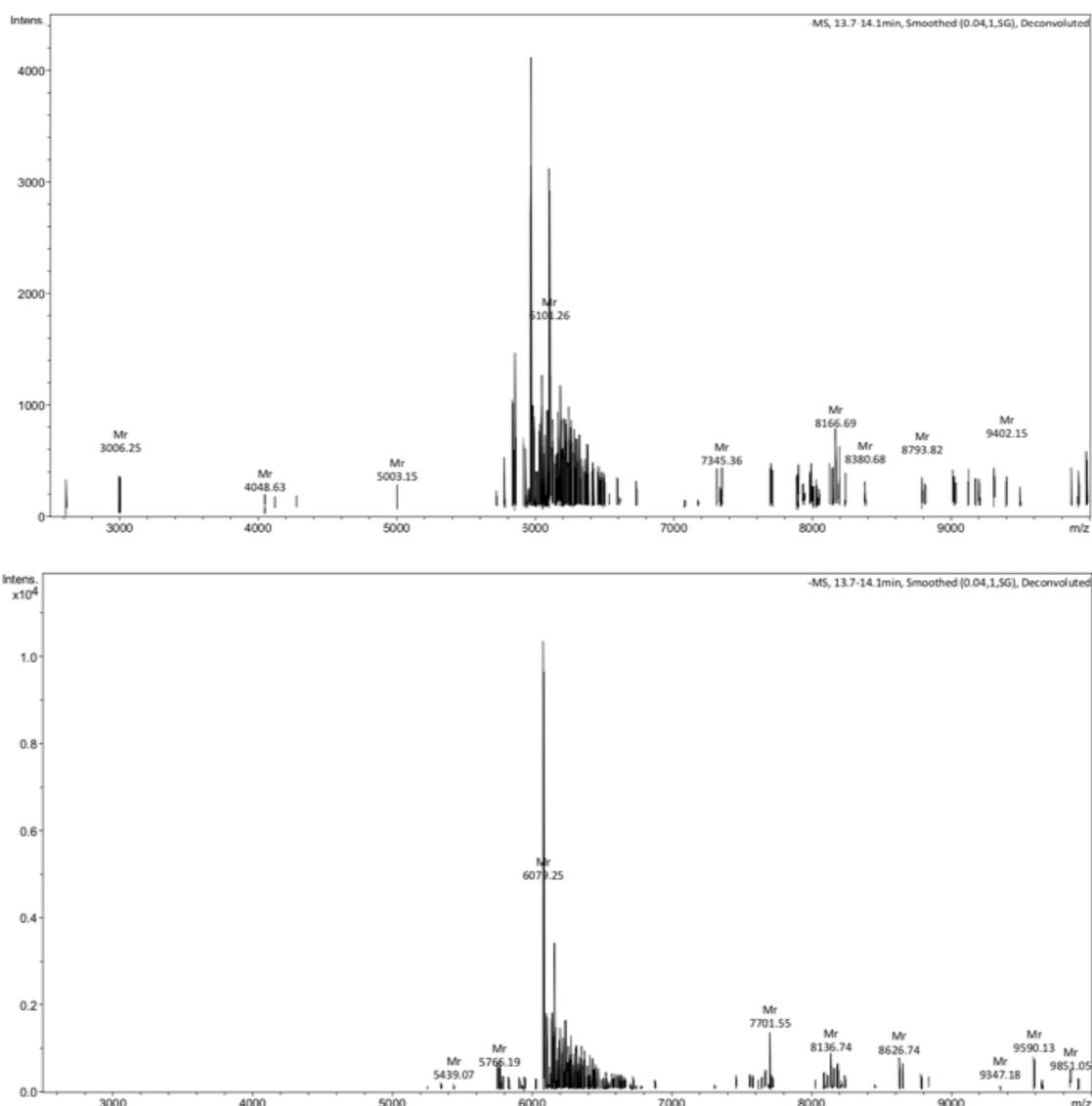


Figure 49: MALDI-TOF MS data of the linker-DNA conjugates (compound **12**)
Above: DNA1-Linker
Below: DNA2-Linker

4.3.4 Conjugation of DNA-Linker and Protein

6-Maleimidohexanoic Acid N-Hydroxysuccinimide Ester-DNA Conjugation with Protein

It should be noted that this experiment was run before MS data was obtained for the conjugation of the linker (**linker 2A**) and DNA. MS data has since shown that the conjugation of the linker and DNA was unsuccessful and consequently this work would not have attained any beneficial results.

The gel produced showed that the conjugation of DNA-linker with Sup35NM LCys was unsuccessful. Conjugation of the protein to the fluorescent chloroacetamide was successful, as the gel showed three bands; these three bands represented conjugation to the protein monomer, protein dimer and also no conjugation. However, after staining with SYBR safe, the only bands visible for DNA were right at the bottom of the gel, corresponding to just the DNA-linker. It would be expected that there would also be a conjugation band higher up on the gel, which would represent DNA-LCys. Likewise, once the gel was stained with Instant Blue, the only visible bands corresponded to the protein monomer and there were no bands visible for the conjugated DNA-LCys. The gel also did not show LCys-PEG-Maleimide.

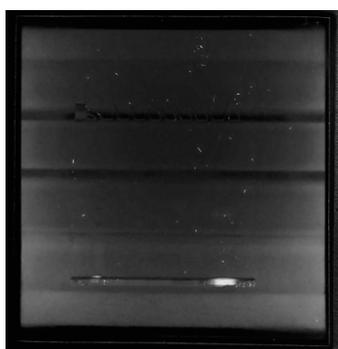


Figure 50: The gel after DNA staining

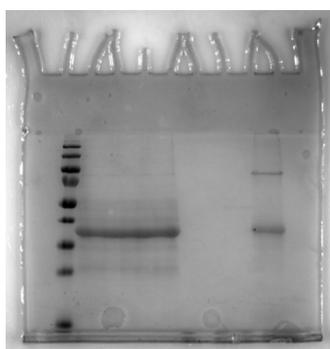


Figure 51: The gel after protein staining

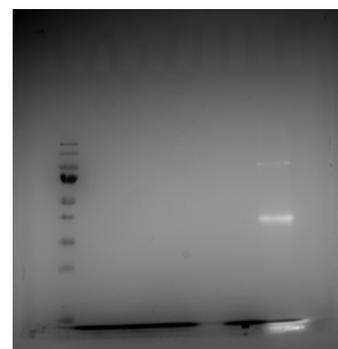


Figure 52: The gel viewed on the transilluminator

The above protocol was determined to be the most suitable produced so far. A previous attempt of the conjugation followed a slightly different method, as described below; these

conjugations have been carried out in order to find the most suitable conditions and method for future conjugations.

A PD10 column was run, using an Elution G buffer; in order to exchange the buffer that the protein was in, and also to remove the DTT used to reduce the cysteine on the protein. The protein concentration was determined and thus samples of LCys-DNA1 and LCys-DNA2 were prepared, comprising of protein, DNA, DTT and Elution G buffer. These two samples were combined and then ran through PD10 columns and UV-Vis absorbance collected. However, during these steps it was assumed the reduction of the thiol-group on the cysteine residue by DTT and the conjugation of DNA-linker to the protein were both immediate; future conjugations will allow time for these reactions to occur before the subsequent steps are carried out. The absorbance was also too high to be calculated and thus the final conjugated protein concentration was not estimated; this would mean the molar equivalents of DNA and linker would not be accurate and hence the reaction would not be repeatable.

Samples were also prepared to be run on an SDS-PAGE gel (12 % polyacrylamide) with the following quantities,

LCys-DNA1: Protein (50 μ L), DNA1 (19.16 μ L) ,DTT (10 μ L), Elution G buffer (20.84 μ L)

LCys-DNA2: Protein (50 μ L), DNA2 (10.24 μ L), DTT (10 μ L), Elution G buffer (29.76 μ L)

LCys-DNA1-DNA2: Protein (50 μ L), DNA1 (9.58 μ L), DNA2 (5.12 μ L), DTT (10 μ L),
Elution G buffer (25.30 μ L)

The samples were precipitated by the ethanol precipitation protocol before being loaded onto the gel. The gel was run for 2 hours and then stained with SYBR safe and Instant Blue. It was evident from the gel that the conjugation did not work. From the SYBR safe stain, there was a thin band across all three samples lane where the DNA was; this corresponded to the DNA itself which, as it is so small, ran to the end of the gel. There were no visible bands representative of the protein-DNA conjugation. Similarly, the stain with Instant Blue

showed the protein as both a monomer and a dimer, but there was no conjugation band present.

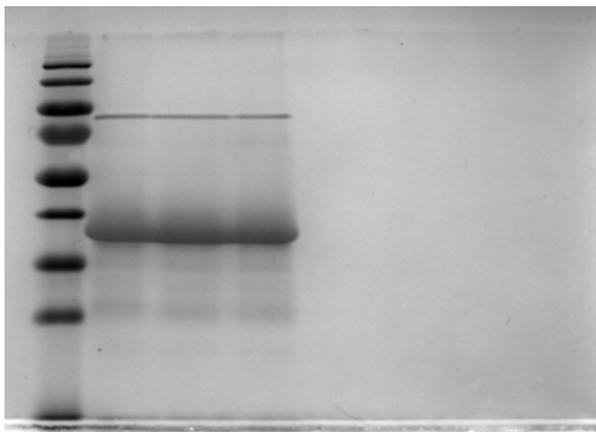


Figure 53: The gel after protein staining

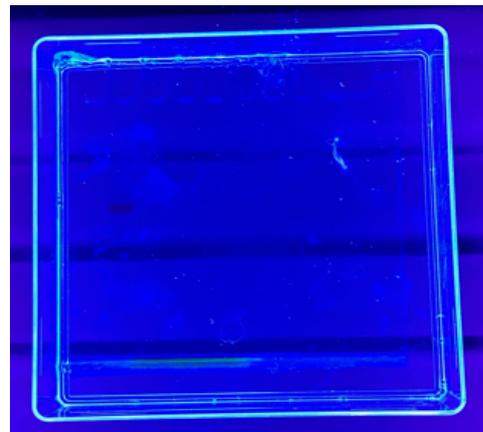


Figure 54: The gel after DNA staining

An amyloid formation assay was also run over ~72 hours, using a 96 well microplate. The well plate contained varying proportions of Sup35NM LCys and wild type (WT) protein, as well as some wells containing ThT dye, as shown below in figure 55; ThT was added in order to study the kinetics of fibril formation, whilst the wells that contained only the eluted solution would enable AFM imaging of the fibrils. However, upon discovery that the conjugation was unsuccessful from the SDS-PAGE gel, downstream experiments (ThT assay, AFM imaging) were aborted, but their initial impression was to confirm there was no change in formation compared to that of the wild type protein.

During the buffer exchange of Sup35NM LCys into the fibril formation buffer, 2 peaks were observed when the A280 nm value was measured. As a result, both protein eluents were used. However, as the conjugation did not work, the first peak shown in the A280 nm measurement would have been representative of the protein, whilst the second peak would have corresponded to just the linker-DNA eluting separately. Consequently, the fibrils that could be seen from the amyloid formation assay were fibrils of the wild type protein aggregating on its own. This, coupled with the SDS-PAGE gel, indicated that the conjugates had not been produced and consequently, the ThT assay was stopped and AFM imaging of the fibrils was not carried out.

		Eluent from Peak 1			Eluent from Peak 2								
		W	W	W	W	W	W	W	W	W	W	W	W
With ThT	W	100% LCys	50% LCys 50% WT	10% LCys 90% WT	100% LCys	50% LCys 50% WT	10% LCys 90% WT	W					W
	W	100% LCys	50% LCys 50% WT	10% LCys 90% WT	100% LCys	50% LCys 50% WT	10% LCys 90% WT	W					W
	W	100% LCys	50% LCys 50% WT	10% LCys 90% WT	100% LCys	50% LCys 50% WT	10% LCys 90% WT	W					W
Without ThT	W	100% LCys	50% LCys 50% WT	10% LCys 90% WT	100% LCys	50% LCys 50% WT	10% LCys 90% WT	W					W
	W	100% LCys	50% LCys 50% WT	10% LCys 90% WT	100% LCys	50% LCys 50% WT	10% LCys 90% WT	W					W
	W	100% LCys	50% LCys 50% WT	10% LCys 90% WT	100% LCys	50% LCys 50% WT	10% LCys 90% WT	W					W
		W	W	W	W	W	W	W	W	W	W	W	W

Figure 55: The 96 well microplate

4.4 Conclusion

In general, the synthesis of both the chemical linkers has been achieved, although with varying degrees of success. The synthesis of **linker 2A**, shown below in figure 56, 6-maleimidohexanoic acid *N*-hydroxysuccinimide ester, was completed with the structure of the linker confirmed by both ^1H and ^{13}C NMR spectra and MS data. Subsequently, the linker was conjugated to DNA strands, that had been purchased from IDT technologies. Despite the UV-Vis absorbance graphs and the Kaiser test indicating that the conjugation had occurred, MS data repeatedly showed the conjugation had been unsuccessful. The conjugation of the **linker 2A-DNA** with the protein had been carried out before the MS data had been obtained, which consequently provided futile results.

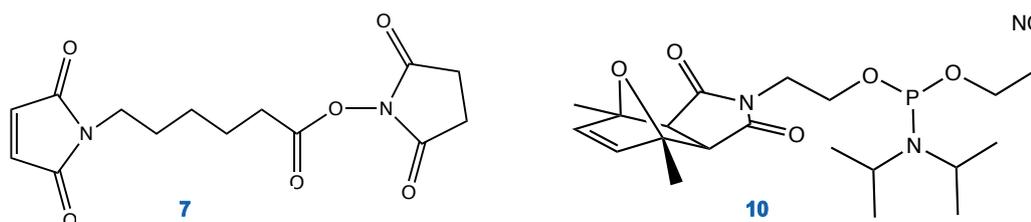


Figure 56: The two chemical linkers
 Left: Compound **7**, which is **linker 2A**
 Right: Compound **10**, which is **linker 2B**

Linker 2B, shown above in figure 56, was also partially successful. The three-step synthesis was undertaken and NMR spectra showed that the desired molecule had been produced at each step; MS data also confirmed the presence of the target molecules for the first two reactions. Nevertheless, when the linker was hand-coupled to the synthesised DNA strands, the coupling was unsuccessful, as confirmed by MS data. Consequently, the conjugation of **linker 2B-DNA** with protein has not been attempted.

5.1 Conclusion

In summary, this research set out to synthesise two different chemical linkers, which would be utilised to conjugate a known amyloid-forming protein and a strand of DNA. Chapter 2 focused on the synthesis and characterisation of two maleimide linkers, whose structure only differed in the length of the carbonyl chain between the nitrogen-containing aromatic group and the maleimide ring. Originally the project set out to synthesise only **linker 1A**. However, due to multiple unsuccessful attempts a new method was devised, producing the bigger **linker 1B**; the structures of these are shown below in figure 57 for reference. Once NMR spectra and MS data showed whether the desired product had been produced, conjugations with the protein were carried out; MS data showed protein-linker conjugates were obtained.

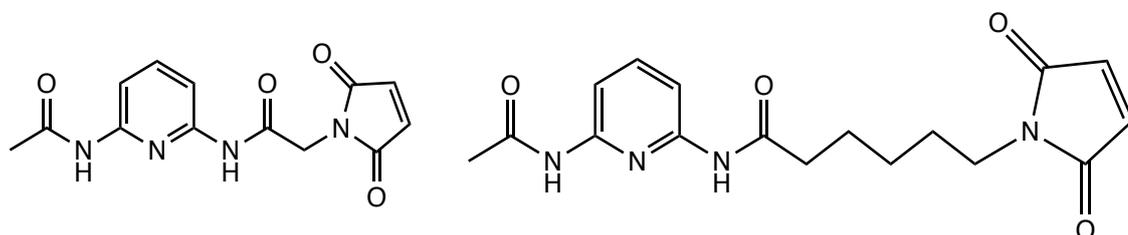


Figure 57: The two maleimide chemical linkers
Left: **Linker 1A** (compound 4)
Right: **Linker 1B** (compound 5)

Chapter 3 reviews the methods used to conjugate **linker 1B** synthesised in chapter 2 with a Waltz peptide, Ac-CGSGHYFNIF-NH₂. Many attempts at the conjugation were carried out before the method was refined and the presence of conjugates were seen by TLC analysis and later by MS data. Once achieved, the reaction was attempted again on a larger scale in the hope to isolate the conjugate in order for purification and more analysis to be carried out but without any positive results.

Chapter 4 then discussed the techniques to produce and characterise two phosphoramidite maleimide linkers, whose structures were slightly different. The first method led to the successful synthesis of **linker 2A** (compound **7**), shown below in figure 58, which was thus bonded to a strand of DNA by a solution-based method. NMR spectra was used to check the structure of the linker, before UV-Vis measurements confirmed the presence of the DNA and were used to calculate its concentration. Once completed, the linker-DNA strands were conjugated to a protein, Sup35NM LCys, with polyacrylamide gels, MS data and analysis of 96 well microplates used to confirm the presence of the conjugates and to monitor the growth of amyloid fibrils. The first attempts at the conjugation was unsuccessful, which led to method development to find the most suitable conditions and method for the succeeding conjugations. **Linker 2B** (compound **10**), shown below in figure 58, was successfully synthesised, but unfortunately the conjugation with the DNA was unsuccessful. NMR spectra confirmed that the linker had been produced, although the phosphorous atom was present in both its oxidised and unoxidised state, with more oxidised phosphorous present. As a result, when coupled to the DNA, using solid beads containing the DNA, the reaction was unsuccessful. UV-Vis measurements confirmed the presence of the DNA, but MS data showed that the DNA was not bound to the linker.

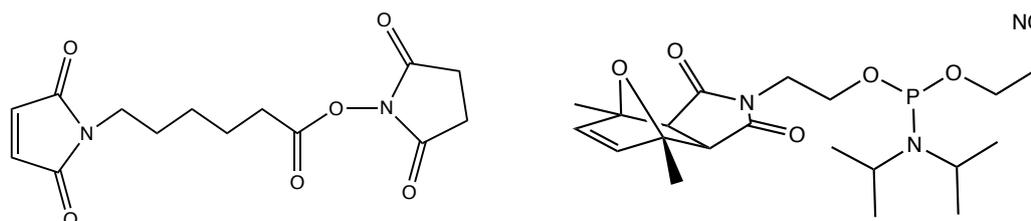


Figure 58: The two chemical linkers
Left: **Linker 2A** (compound **7**)
Right: **Linker 2B** (compound **10**)

5.2 Future Work

If this research was continued, the attempts at synthesising **linker 1A** would have been pursued. More conjugations of **linker 1B** would have also been carried out, with analysis of amyloid formation assays incorporating the use of ThT dye to monitor the kinetics and image any fibrils produced. If achieved, variation of the DNA strands would have also been investigated; differing the length and sequence of DNA strands could alter the assembly of Sup35NM LCys into fibrils, as well as the ratio of protein to linker and DNA strand.

Future work of this project would continue the analysis of the phosphoramidite maleimide **linker 2A** with DNA strands and protein. The conjugation of DNA-linker with protein would be continued, with the level of crosslinking between the complementary DNA strands controlled by altering the amount of fibrils in the amyloid that are covalently bonded to the linker and thus to a DNA strand, once of the conditions of the conjugation had been optimised.

Furthermore, the conjugation of **linker 1B** with the Waltz peptide would be continued. The conjugates of the linker and peptide were seen in MS data and future work would thus repeat this conjugation on a larger scale for more analysis to be continued. If successful, a linker-peptide conjugate would be isolated, purified and, once its presence confirmed by MS, AFM imaging would be carried out to study the formation and assembly of the linker-peptide fibrils.

5.3 References

- 1 L. Lutter, C. J. Serpell, M. F. Tuite and W.-F. Xue, *Biochim. Biophys. Acta - Proteins Proteomics*, 2019, **1867**, 140257.
- 2 J. D. Sipe and A. S. Cohen, *J. Struct. Biol.*, 2000, **130**, 88–98.
- 3 T. P. J. Knowles, M. Vendruscolo and C. M. Dobson, *Nat. Rev. Mol. Cell Biol.*, 2014, **15**, 384–396.
- 4 K. E. Marshall, R. Marchante, W. F. Xue and L. C. Serpell, *Prion*, 2014, **8**, 192–196.
- 5 M. Verma, A. Vats and V. Taneja, *Ann. Indian Acad. Neurol.*, 2015, **18**, 138–145.
- 6 K. K. Frederick, G. T. Debelouchina, C. Kayatekin, T. Dorminy, A. C. Jacavone, R. G. Griffin and S. Lindquist, *Chem. Biol.*, 2014, **21**, 295–305.
- 7 E. Evangelisti, R. Cascella, M. Becatti, G. Marrazza, C. M. Dobson, F. Chiti, M. Stefani and C. Cecchi, *Sci. Rep.*, 2016, **6**, 1–14.
- 8 K. L. Stewart and S. E. Radford, *Biophys. Rev.*, 2017, **9**, 405–419.
- 9 J. Cummings, *Clin. Transl. Sci.*, 2018, **11**, 147–152.
- 10 J. R. Harrison and M. J. Owen, *Br. J. Psychiatry*, 2016, **208**, 1–3.
- 11 S. Chakradhar, *Nat. Med.*, 2018, **24**, 1785–1787.
- 12 S. Makin, *Nature*, **559**, 54–57.
- 13 J. N. Onuchic, N. D. Socci, L.-S. Zaida and P. G. Wolynes, *Fold. Des.*, 1996, **1**, 441–450.
- 14 S.-Q. Liu, X.-L. Ji, Y. Tao, D.-Y. Tan, K.-Q. Zhang and Y.-X. Fu, in *Protein Engineering*, ed. P. Kaumaya, InTech, London, 2012, ch. 10, pp. 207–252.
- 15 J. N. Onuchic and P. G. Wolynes, *Curr. Opin. Struct. Biol.*, 2004, **14**, 70–75.
- 16 S. Y. Ow and D. E. Dunstan, *Protein Sci.*, 2014, **23**, 1315–1331.
- 17 M. G. Iadanza, M. P. Jackson, E. W. Hewitt, N. A. Ranson and S. E. Radford, *Nat. Rev. Mol. Cell Biol.*, 2018, **19**, 755–773.
- 18 M. P. Jackson and E. W. Hewitt, *Biomolecules*, 2017, **7**, 1–13.
- 19 E. Chatani and N. Yamamoto, *Biophys. Rev.*, 2018, **10**, 527–534.
- 20 T. Härd, *J. Phys. Chem. Lett.*, 2014, **5**, 607–614.

- 21 R. Nelson, M. R. Sawaya, M. Balbirnie, A. Ø. Madsen, R. Grothe and D. Eisenberg, *Nih*, 2006, **435**, 773–778.
- 22 S. Linse, *Biophys. Rev.*, 2017, **9**, 329–338.
- 23 G. Meisl, J. B. Kirkegaard, P. Arosio, T. C. T. Michaels, M. Vendruscolo, C. M. Dobson, S. Linse and T. P. J. Knowles, *Nat. Protoc.*, 2016, **11**, 252–272.
- 24 B. H. Meier, R. Riek and A. Böckmann, *Trends Biochem. Sci.*, 2017, **42**, 777–787.
- 25 L. C. Serpell, *Biochim. Biophys. Acta - Mol. Basis Dis.*, 2000, **1502**, 16–30.
- 26 D. S. Eisenberg and M. R. Sawaya, *Annu. Rev. Biochem.*, 2017, **86**, 69–95.
- 27 J. D. Sipe, M. D. Benson, J. N. Buxbaum, S.-I. Ikeda, G. Merlini, M. J. M. Saraiva and P. Westermark, *Amyloid*, 2014, **21**, 221–224.
- 28 C. Xue, T. Y. Lin, D. Chang and Z. Guo, *R. Soc. Open Sci.*, 2017, **4**, 1–12.
- 29 A. Schmidt, M. Schmidt, K. Annamalai, N. Grigorieff and M. Fändrich, *Proc. Natl. Acad. Sci.*, 2016, **113**, 6200–6205.
- 30 C. Liu, X. Gui, X. Zhang, Z. Luo, Z. Liu, X. Li, F. Luo, C. Zhao, Y. Li and D. Li, *Cell Res.*, 2018, **28**, 897–903.
- 31 F. S. Ruggeri, J. Habchi, A. Cerreta and G. Dietler, *Curr. Pharm. Des.*, 2016, **22**, 3950–3970.
- 32 S. Kim, J. H. Kim, J. S. Lee and C. B. Park, *Small*, 2015, **11**, 3623–3640.
- 33 H. H. Susapto, O. U. Kudu, R. Garifullin, E. Yilmaz and M. O. Guler, *ACS Appl. Mater. Interfaces*, 2016, **8**, 17421–17427.
- 34 E. R. Draper, M. Wallace, D. Honecker and D. J. Adams, *Chem. Commun.*, 2018, **54**, 10977–10980.
- 35 L. Adler-Abramovich and E. Gazit, *Chem. Soc. Rev.*, 2014, **43**, 6873–7238.
- 36 A. Onoda, H. Harada, T. Uematsu, S. Kuwabata, R. Yamanaka, S. Sakurai and T. Hayashi, *RSC Adv.*, 2017, **7**, 1089–1092.
- 37 X. Q. Dou and C. L. Feng, *Adv. Mater.*, 2017, **29**, 1604062–1604082.
- 38 J. J. Panda and V. S. Chauhan, *Polym. Chem.*, 2014, **5**, 4418–4436.
- 39 M. Sleutel, I. Van Den Broeck, N. Van Gerven, C. Feuillie, W. Jonckheere, C. Valotteau, Y. F. Dufrêne and H. Remaut, *Nat. Chem. Biol.*, 2017, **13**, 902–908.
- 40 T. P. J. Knowles and R. Mezzenga, *Adv. Mater.*, 2016, **28**, 6546–6561.

- 41 S. G. Biesecker, L. K. Nicastro, R. Paul Wilson and Ç. Tükel, *Biomolecules*, 2018, **8**, 1–15.
- 42 J. Y.-J. Shyy, S. Chien, Z. Chen, S. Subramaniam and S. Li, *Annu. Rev. Biomed. Eng.*, 2017, **19**, 195–219.
- 43 O. N. Tikhodeyev, *Biol. Rev.*, 2018, **93**, 1987–2005.
- 44 G. Lee, W. Lee, H. Lee, C. Y. Lee, K. Eom and T. Kwon, *Sci. Rep.*, 2015, **5**, 1–8.
- 45 C. Ha and C. B. Park, *Biotechnol. Bioeng.*, 2005, **90**, 848–855.
- 46 R. S. Jacob, S. Das, S. Ghosh, A. Anoop, N. N. Jha, T. Khan, P. Singru, A. Kumar and S. K. Maji, *Sci. Rep.*, 2016, **6**, 1–18.
- 47 B. Xu, J. Shi, X. Du, H. Wang, Z. Feng and J. Li, *Chem. Commun.*, 2016, **52**, 6332–6335.
- 48 Z. G. Wang and B. Ding, *Acc. Chem. Res.*, 2014, **47**, 1654–1662.
- 49 Y. Tian, T. Wang, W. Liu, H. L. Xin, H. Li, Y. Ke, W. M. Shih and O. Gang, *Nat. Nanotechnol.*, 2015, **10**, 637–644.
- 50 P. W. K. Rothmund, *Nature*, 2006, **440**, 297–302.
- 51 C. Hou, S. Guan, R. Wang, W. Zhang, F. Meng, L. Zhao, J. Xu and J. Liu, *J. Phys. Chem. Lett.*, 2017, **8**, 3970–3979.
- 52 W. Wang, Y. Han, Y. Fan and Y. Wang, *Langmuir*, 2019, **35**, 2334–2342.
- 53 W. Zhou, Q. Y. Lin, J. A. Mason, V. P. Dravid and C. A. Mirkin, *Small*, 2018, **14**, 1–7.
- 54 C. J. Serpell, M. Barlóg, K. Basu, J. F. Fakhoury, H. S. Bazzi and H. F. Sleiman, *Mater. Horizons*, 2014, **1**, 348–354.
- 55 Q. Y. Lin, J. A. Mason, Z. Li, W. Zhou, M. N. O'Brien, K. A. Brown, M. R. Jones, S. Butun, B. Lee, V. P. Dravid, K. Aydin and C. A. Mirkin, *Science (80-.)*, 2018, **359**, 669–672.
- 56 P. S. Arora, W. B. Sherman, S. L. Gras, N. C. Seeman, M. N. Bongiovanni, J. W. Canary, A. Udomprasert, R. Sha and T. Wang, *Nat. Nanotechnol.*, 2014, **9**, 537–541.
- 57 M. Surin, *Polym. Chem.*, 2016, **7**, 4137–4150.
- 58 K. Schwartz, M. Ganesan, D. E. Payne, M. J. Solomon and B. R. Boles, *Mol.*

- Microbiol.*, 2016, **99**, 123–134.
- 59 J. N. Abraham, N. Gour, S. Bolisetty, R. Mezzenga and C. Nardin, *Eur. Polym. J.*, 2015, **65**, 268–275.
- 60 M. Balbirnie, R. Grothe and D. S. Eisenberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 2375–2380.
- 61 S. A. Bondarev, V. V. Shchepachev, A. V. Kajava and G. A. Zhouravleva, *J. Biol. Chem.*, 2013, **288**, 28503–28513.
- 62 S. Lindquist, J. R. Glover, J.-J. Liu, A. S. Kowal, E. C. Schirmer and M. M. Patino, *Cell*, 2004, **89**, 811–819.
- 63 C. Bailly and M. J. Waring, *Nucleic Acids Res.*, 1998, **26**, 4309–4314.
- 64 IDT Oligoanalyser Tool, <https://www.idtdna.com/pages/tools/oligoanalyzer>, (accessed 10 September 2019).
- 65 D. Wan, K. Satoh and M. Kamigaito, *Macromolecules*, 2006, **39**, 6882–6886.
- 66 J. R. Dunetz, J. Magano and G. A. Weisenburger, *Org. Process Res. Dev.*, 2016, **20**, 140–177.
- 67 K. J. Jensen, P. T. Shelton and S. L. Pedersen, *Peptide Synthesis and Applications*, Humana Press, 2nd edn., 2013, vol. 1047.
- 68 G. T. Hermanson, *Bioconjugate Techniques*, Academic Press, Third., 2013.
- 69 A. J. Sinclair, V. Del Amo and D. Philp, *Org. Biomol. Chem.*, 2009, **7**, 3308–3318.
- 70 Z. S. Al-Garawi, K. L. Morris, K. E. Marshall, J. Eichler and L. C. Serpell, *Interface Focus*, , DOI:10.1098/rsfs.2017.0027.
- 71 K. L. Morris, A. Rodger, M. R. Hicks, M. Debulpaep, J. Schymkowitz, F. Rousseau and L. C. Serpell, *Biochem. J.*, 2013, **450**, 275–283.
- 72 R. Freeman, M. Han, Z. Álvarez, J. A. Lewis, J. R. Wester, N. Stephanopoulos, M. T. McClendon, C. Lynsky, J. M. Godbe, H. Sangji, E. Luijten and S. L. Stupp, *Science (80-.)*, 2018, **362**, 808–813.
- 73 A. Sánchez, E. Pedroso and A. Grandas, *Org. Lett.*, 2011, **13**, 4364–4367.
- 74 F. A. Aldaye and H. F. Sleiman, *Angew. Chemie - Int. Ed.*, 2006, **45**, 2204–2209.
- 75 C. A. G. N. Montalbetti and V. Falque, *Tetrahedron*, 2005, **61**, 10827–10852.