New Strategies for Cysteine Bioconjugation and Protein Cross-Linking

Nafsika Forte

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Supervisor: Dr James R. Baker

June 2018
Declaration

I, Nafsika Forte, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Abstract

The connection of two or more proteins with different binding specificities or activities can lead to unique and powerful systems with significant potential in therapeutic applications, particularly oncology. While a plethora of active bifunctional proteins have been produced by fusion, this approach is not always viable, since modification of the N- or C- terminus can compromise the proteins’ bioactivities, with complications during folding and processing often being observed. Chemical cross-linking can offer a more versatile strategy towards the synthesis of protein-protein conjugates. However, current state-of-the-art approaches are either not selective, yielding heterogeneous mixtures or require mutations with natural or unnatural amino acids, which are often associated with low expression yields and limited applicability towards all protein systems.

Recently, a bis-dibromomaleimide cross-linking reagent was used to create a bispecific scFv-Fab antibody in a site-specific manner, by targeting and bridging disulfide bonds. In this work, the application of these cross-linking reagents was extended into the conjugation of proteins containing different reactive handles, as well as sterically hindered conjugation sites. Diiodomaleimides were shown to offer the ideal properties of rapid bioconjugation and importantly, reduced hydrolysis, which was identified as a significant factor affecting the conjugation efficiency. The optimized linkers were exploited to link human serum albumin to antibody fragments (scFv and Fab) as a prospective half-life extension platform, with retention of antigen binding and robust serum stability. The linker design was also extended to incorporate a third point of attachment for another biomolecule. An anti-CEA scFv homotrimer was formed, by disulfide bridging, which was shown to infer a combination of greater antigen avidity and increased in vivo half-life. Finally, with the view to developing a Bispecific T-cell Engager (BiTE) construct that could lead to more efficient cell-killing, an anti-CEA/anti-CD3 scFv heterotrimer of increased avidity towards CD3 antigen was developed, showing the potential of this strategy for the construction of a diverse range of bioconjugates.

The last part of this thesis describes investigations towards the development of a new methodology that selectively modifies specific lysine residues, via acyl transfer from the disulfide bond of Fab antibody fragment. This approach is expected to complement disulfide bridging as a strategy for selective modification, but also be advantageous for systems where this could result in detrimental changes to the antibody structure, offering a highly promising platform for bioconjugation.
Impact Statement

Cancer is becoming the leading cause of mortality in most developed countries, with an estimated number of new cancer cases being above 300,000 every year, in the UK. In the past decades, the potential of monoclonal antibodies in targeted therapy against cancer has been realized, with the market for therapeutic antibodies currently the fastest growing sector in the pharmaceutical industry. While monoclonal antibodies have demonstrated unprecedented activities and therapeutic windows in clinical and preclinical trials, their large size has been reported to limit tissue penetration, with more than 99% of the antibody being administered not reaching the tumour. In contrast, antibody fragments have the advantage of rapid tumour penetration, owing to their smaller size. However, this also leads to rapid clearance from circulation, which is unlikely to give sustained effects for therapy, due to high target retention required. The principal aim of the current project was to address this issue by creating antibody fragment conjugates of increased in vivo half-life and pharmacokinetic properties, which are expected to have a significant impact in the development of future therapeutics.

To achieve this, various cross-linking reagents based on the next-generation maleimide (NGM) scaffold were developed and evaluated in linking antibody fragments to Human Serum Albumin (HSA), as a prospective half-life extension platform. Through optimization, important factors affecting protein-protein conjugation were identified, which are anticipated to assist and direct future work in this area. Ultimately, this cross-linking strategy enabled access to the target conjugates in high yields, with retention of antigen binding and robust serum stability, overcoming significant limitations of previous literature approaches.

The versatility of this approach was also demonstrated in the creation of tri-protein conjugates, giving rise to a novel class of prospective therapeutics. An scFv homotrimer against carcinoembryonic antigen CEA, a common tumour marker, was initially obtained in an efficient one-pot procedure. Through a collaboration with the UCL Cancer Institute, the biodistribution of this construct was evaluated. A significant improvement in the in vivo half-life was observed, compared to the unmodified monomer, which in combination with the increase in valency could lead to the development of powerful constructs for future in vivo applications. Following this, a novel anti-CEA/anti-CD3 Bispecific T-cell Engager (BiTE) heterotrimer construct was developed to overcome the short half-life of equivalent heterodimer constructs that are currently in the clinic.
In addition, the increased avidity towards CD3 is expected to lead to more efficient cell-killing, which is currently under investigation in collaboration with the UCL Cancer Institute.

The last part of this project was focused on the development of a new bioconjugation methodology that selectively modifies specific lysine residues of an antibody fragment, in response to the heterogeneous product mixtures obtained by traditional lysine modification. It was envisaged that this could be achieved by acyl transfer from the cysteine residues of a disulfide bond to proximal lysine residues, demonstrating how standard native chemical ligation can be applied to a different context. The highly promising preliminary results obtained resulted in an EPSRC grant proposal subsequently being funded to further investigate the potential of this approach in therapeutic development.
For my mum
Acknowledgements

I would like to start by thanking my supervisor, Dr Jamie Baker for giving me the opportunity to work in this exciting project, along with the freedom to shape my PhD research. I am very grateful to Jamie for being an amazing supervisor, always helpful, supportive and very enthusiastic! I would like to extend my gratitude to Dr Vijay Chudasama for his advice and helpful discussions over the years.

I would like to give special thanks to Dr Kersti Karu and Dr Xiaoping Yang for their excellent mass spectrometry advice and help. In addition, thanks to Dr Abil Aliev for his invaluable help and guidance with NMR. I would also like to acknowledge our collaborator Professor Kerry Chester and members of the Chester group, particularly Dr Enrique Miranda and Maria Livanos for their help with ELISAs and for expressing proteins to use in this project.

I am very grateful to all past and present members of the Baker group; Dr Liz Hull, Dr Mauricio Morais, Dr Florian Kampmeier, Dr João Nunes, Dr Rosemary Huckvale, Dr Daniel Richards, Irene Benni, Archie Wall and Roshni Malde. A special thanks goes to Dr Liz Hull for training me in protein modification and to Dr Mauricio Morais for answering all of my questions. In addition, thanks to Archie Wall and Roshni Malde for proof-reading my thesis and being very fun desk neighbours. I would also like to thank the Howorka group in the chemical biology lab for their valuable help and fun discussions; Hafi Nguyen, Patrick Arnott, Jo Donnelly, Daniel Offenbartl and Conor Lanphere. Special thanks to Gen Pugh for her help during the endless hours of SEC optimization.

I would like to thank everyone in the KLB for the interesting discussions about science and life and for creating such a fun and “unconventional” work environment that no words can describe; Dr Antoine Maruani, André Shamsabadi, Dr Max Lee, Calise Bahou, Marcos Fernández, Faiza Javaid, Peter Szijj, Dr Valerie Karaluka, Dr Samantha Gibson, Charlotte Coomber, Dr Laure Benhamou, Dr Brian Sin, Dr Victor Laserna and Dr Marc Radigois. A very special thanks goes to Dr Marco Sabatini, Dr Raul Pereira and João Nogueira for being fantastic friends, always suggesting things that we “should” do.

Last, but not least I would like to thank my family. Constantine for being a great brother and friend, always up for adventures. I am deeply grateful to my dad for his support over the years and for
teaching us to always aim higher. Finally, I am eternally grateful to my mum, to whom this thesis is dedicated. Without her support, constant motivation and valuable life lessons, I would not be where I am today.

“Πάντα στὸ νοῦ σου νάχης τὴν Ἰθάκη. Τὸ φθάσιμον ἐκεῖ εἶν’ ὁ προορισμός σου. Αλλὰ μὴ βιάζῃς τὸ ταξείδι διόλου.”

Κωνσταντίνος Π. Καβάφης

“Have Ithaca always in your mind. Your arrival there is what you are destined for. But don’t in the least hurry the journey.”

Constantine P. Cavafy
Table of Contents

Declaration ..................................................................................................................... i
Abstract ......................................................................................................................... ii
Impact Statement ........................................................................................................ iii
Acknowledgements ....................................................................................................... vi
Abbreviations, formulae and symbols ........................................................................ xi
1. Background .................................................................................................................. 1
   1.1 Protein modification ............................................................................................... 1
       1.1.1 Modification of unnatural amino acids ......................................................... 3
       1.1.2 Bioorthogonal modification of inserted functionalities ................................. 4
       1.1.3 Modification of native amino acids .............................................................. 6
           1.1.3.1 Lysine modification ................................................................................... 7
           1.1.3.2 Cysteine modification ............................................................................... 9
           1.1.3.3 Disulfide bridging .................................................................................... 13
       1.1.4 Chemical cross-linking of biomolecules ......................................................... 19
       1.1.5 Native Chemical Ligation ................................................................................ 22
   1.2 Antibodies for targeted therapeutics .................................................................... 25
       1.2.1 Antibody fragments ....................................................................................... 27
       1.2.2 Multivalent antibody fragments ..................................................................... 28
       1.2.3 Antibody-Drug Conjugates ............................................................................ 30
       1.2.4 Site-selective conjugation of drugs to antibodies .......................................... 32
           1.2.4.1 Engineered antibody modification ............................................................ 32
           1.2.4.2 Native antibody modification ................................................................. 34
       1.2.5 Bispecific antibodies ...................................................................................... 37
   1.3 Human Serum Albumin ......................................................................................... 41
       1.3.1 Conjugation to Human Serum Albumin ....................................................... 42
   1.4 Aims and Objectives ............................................................................................ 45
2. Results and Discussion ............................................................................................. 48
   2.1 Albumin-antibody fragment conjugates .................................................................. 48
       2.1.1 Synthesis and application of bifunctional conjugation reagents ................... 49
           2.1.1.1 Bis-dibromomaleimide PEG linker .......................................................... 49
<table>
<thead>
<tr>
<th>Abbreviations, formulae and symbols</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Å</td>
<td>Angstrom</td>
</tr>
<tr>
<td>ADC</td>
<td>Antibody Drug Conjugate</td>
</tr>
<tr>
<td>Ac</td>
<td>Acetyl</td>
</tr>
<tr>
<td>Asn</td>
<td>Asparagine</td>
</tr>
<tr>
<td>aq.</td>
<td>Aqueous</td>
</tr>
<tr>
<td>AU</td>
<td>Arbitrary Units</td>
</tr>
<tr>
<td>BBS</td>
<td>Borate Buffered Saline</td>
</tr>
<tr>
<td>BiTE</td>
<td>Bispecific T-cell Engager</td>
</tr>
<tr>
<td>BME</td>
<td>β-mercaptoethanol</td>
</tr>
<tr>
<td>Boc</td>
<td>tert-butyloxyacarbonyl</td>
</tr>
<tr>
<td>Boc$_2$O</td>
<td>Di-tert-butyl dicarbonate</td>
</tr>
<tr>
<td>CD3</td>
<td>Cluster of Differentiation 3</td>
</tr>
<tr>
<td>CEA</td>
<td>Carcinoembryonic Antigen</td>
</tr>
<tr>
<td>CI</td>
<td>Chemical Ionization</td>
</tr>
<tr>
<td>CID</td>
<td>Collision Induced Dissociation</td>
</tr>
<tr>
<td>CLT</td>
<td>Cysteine-to-Lysine Transfer</td>
</tr>
<tr>
<td>CuAAC</td>
<td>Copper(I)-catalyzed Azide-Alkyne Cycloaddition</td>
</tr>
<tr>
<td>Cys</td>
<td>Cysteine</td>
</tr>
<tr>
<td>DAR</td>
<td>Drug-to-Antibody Ratio</td>
</tr>
<tr>
<td>DCM</td>
<td>Dichloromethane</td>
</tr>
<tr>
<td>DIMEA</td>
<td>$N,N$-Diisopropylethylamine</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-Dimethylaminopyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>ds-scFv</td>
<td>Disulfide-stabilised single-chain Fv</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDC</td>
<td>1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic Acid</td>
</tr>
<tr>
<td>EEDQ</td>
<td>2-Ethoxy-1-ethoxycarbonyl-1,2-dihydroquinoline</td>
</tr>
<tr>
<td>EI</td>
<td>Electron Ionization</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme Linked Immunosorbent Assay</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray Ionization</td>
</tr>
<tr>
<td>eq.</td>
<td>Equivalent</td>
</tr>
<tr>
<td>Fab</td>
<td>Fragment antigen-binding</td>
</tr>
<tr>
<td>FAR</td>
<td>Fluorophore-to-Antibody Ratio</td>
</tr>
<tr>
<td>Fc</td>
<td>Fragment crystallisable</td>
</tr>
<tr>
<td>FcRn</td>
<td>Neonatal Fc receptor</td>
</tr>
<tr>
<td>Fv</td>
<td>Fragment variable</td>
</tr>
<tr>
<td>g</td>
<td>Gram</td>
</tr>
<tr>
<td>h</td>
<td>Hour</td>
</tr>
<tr>
<td>HBTU</td>
<td>$O$-Benzotriazole-$N,N,N',N'$-tetramethyl-uronium hexafluorophosphate</td>
</tr>
<tr>
<td>HC</td>
<td>Heavy Chain</td>
</tr>
<tr>
<td>HER2</td>
<td>Human Epidermal growth factor Receptor 2</td>
</tr>
<tr>
<td>His</td>
<td>Histidine</td>
</tr>
<tr>
<td>HPLC</td>
<td>High Performance Liquid Chromatography</td>
</tr>
<tr>
<td>HRMS</td>
<td>High Resolution Mass Spectrometry</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish Peroxidase</td>
</tr>
<tr>
<td>HSA</td>
<td>Human Serum Albumin</td>
</tr>
<tr>
<td>ID/g</td>
<td>Injected Dose per gram</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilized Metal ion Affinity Chromatography</td>
</tr>
<tr>
<td>IR</td>
<td>Infrared</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>LC</td>
<td>Light Chain</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid Chromatography Mass Spectrometry</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>Liquid Chromatography Tandem Mass Spectrometry</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>LRMS</td>
<td>Low Resolution Mass Spectrometry</td>
</tr>
<tr>
<td>Lys</td>
<td>Lysine</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>Me</td>
<td>Methyl</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass per charge</td>
</tr>
<tr>
<td>mAb</td>
<td>Monoclonal Antibody</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>Matrix-Assisted Laser Desorption/ Ionization - Time-of-Flight</td>
</tr>
<tr>
<td>MESNa</td>
<td>Sodium-2-mercaptopethanesulfonate</td>
</tr>
<tr>
<td>MBq</td>
<td>Megabecquerel</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>min</td>
<td>Minute</td>
</tr>
<tr>
<td>mL</td>
<td>Millilitre</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>MMAE</td>
<td>Monomethyl Auristatin E</td>
</tr>
<tr>
<td>mmol</td>
<td>Millimole</td>
</tr>
<tr>
<td>mp</td>
<td>Melting point</td>
</tr>
<tr>
<td>MPAA</td>
<td>4-Mercaptophenylacetic Acid</td>
</tr>
<tr>
<td>MS</td>
<td>Mass Spectrometry</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular Weight</td>
</tr>
<tr>
<td>MWCO</td>
<td>Molecular Weight Cut-Off</td>
</tr>
<tr>
<td>NCL</td>
<td>Native Chemical Ligation</td>
</tr>
<tr>
<td>NGM</td>
<td>Next Generation Maleimide</td>
</tr>
<tr>
<td>NHS</td>
<td>N-hydroxysuccinimide</td>
</tr>
<tr>
<td>nM</td>
<td>Nanomolar</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometre</td>
</tr>
<tr>
<td>nmol</td>
<td>Nanomole</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Base</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene Glycol</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ppm</td>
<td>Parts per million</td>
</tr>
<tr>
<td>RT</td>
<td>Room Temperature</td>
</tr>
<tr>
<td>sat.</td>
<td>Saturated</td>
</tr>
<tr>
<td>scFv</td>
<td>Single chain fragment variable</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>SEC</td>
<td>Size Exclusion Chromatography</td>
</tr>
<tr>
<td>SPAAC</td>
<td>Strain Promoted Alkyne Azide Cycloaddition</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic Acid</td>
</tr>
<tr>
<td>THF</td>
<td>Tetrahydrofuran</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer Ribonucleic Acid</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin Layer Chromatography</td>
</tr>
<tr>
<td>TS</td>
<td>Transition State</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>UAA</td>
<td>Unnatural Amino Acid</td>
</tr>
<tr>
<td>UV/Vis</td>
<td>UV/Visible</td>
</tr>
<tr>
<td>μg</td>
<td>Microgram</td>
</tr>
<tr>
<td>μL</td>
<td>Microlitre</td>
</tr>
<tr>
<td>μM</td>
<td>Micromolar</td>
</tr>
<tr>
<td>μmol</td>
<td>Micromole</td>
</tr>
</tbody>
</table>
1. Background

This research project describes investigations towards new strategies for the chemical modification of proteins, including the development of cross-linking reagents to enable the efficient conjugation of proteins for potential therapeutic applications. This introduction will serve to contextualize the project within the field. The first section will give an overview on the current state-of-the-art in protein modification, with a particular focus on cysteine bioconjugation and the chemical cross-linking of proteins. The following sections will introduce the particular proteins which form the focus of this research project, i.e. antibodies and human serum albumin, along with existing strategies for their modification and challenges that need to be overcome.

1.1 Protein modification

The complete sequencing of the human genome, announced in 2003, ushered in a new era of understanding the chemistry of life.\(^1\) Whilst the number of protein coding genes was found to be surprisingly small (ca. 25,000), human complexity is much higher than the encoding genome would predict, as a result of the enormous variety of protein structures (i.e. >1 million molecular species of proteins).\(^2\) Nature imparts diversity in proteins initially at the transcriptional level, by mRNA splicing,\(^3\) and then upon translation through a range of chemical modifications to confer additional properties, which in turn mediate protein activity.\(^4\) Owing to such post-translational modifications (PTMs), proteins perform a magnificently diverse set of functions in living organisms, such as providing structure and stability, catalyzing reactions, controlling cellular function and signaling, redox regulation, as well as oxygen sensing.\(^5,6\) The five most common examples of post-translational modifications are shown in Figure 1 and include phosphorylation,\(^7\) acetylation,\(^8\) alkylation,\(^9\) glycosylation\(^10\) and oxidation,\(^11\) respectively.

![Figure 1 – Examples of post-translational modifications.](image-url)
Inspired by nature, there have been remarkable efforts to mimic such endogenous modifications, with the view to elucidating their role and providing insight into the function of proteins, which holds great promise for future pharmaceuticals. In addition, protein conjugation methodologies have been developed to impart exogenous modifications for a desired function. For instance, fluorescent tagging allows tracking of proteins both in vitro and in vivo, spin labels enable conformational studies, while affinity tagging allows otherwise difficult purifications. The success of synthetic methodologies has also driven the development of a variety of therapeutic protein conjugates to address unmet medical needs in different disease areas, such as cancer, malaria, HIV and pathogenic bacteria. Notably, 4 bioconjugate medicines (Prevnar®, Neulasta®, Victoza®, Levemir®) were ranked in the top 50 pharmaceutical products in 2016, based on annual global sales.

Over the years, the requirement of conjugation strategies for the design of tailored pharmaceuticals that ensure biocompatibility, safety and stability, as well as improved pharmacokinetic and pharmacodynamic profiles has prompted the development of numerous methodologies for protein modification. However, several challenges exist, which need to be overcome before fulfilling their promise. One of the key requirements for optimal efficacy is the preservation of the original protein structure and function, upon conjugation. As a result, conditions necessary to prevent protein denaturation must be employed, i.e. aqueous media, ambient temperature and near-physiological pH. The reactions should also tolerate salts and surfactants which are often required for protein stability. In addition, reactions should exhibit rapid reaction rates in order to achieve full conversion in reasonable timescales, since proteins are often available in low concentrations. A final challenge is that of site-selectivity, i.e. modifications should be introduced at a distinct site in the presence of hundreds of competing reactive groups, in order to generate a homogeneous and well-defined bioconjugate. Conceptually, this can be accessed by targeting the most differentiating natural amino acids or by pre-installing orthogonal functionalities by protein engineering.
1.1.1 Modification of unnatural amino acids

Recent advances in synthetic biology and protein engineering have made it possible to expand the genetic code and incorporate non-natural residues into proteins, possessing structures and functions not found in the canonical 20 amino acids.\textsuperscript{26} With this methodology, the possibilities for site-specific conjugation can be significantly enhanced, offering a solution to the issue of chemoselectivity. Several methods exist for the translational incorporation of unnatural amino acids (UAAs), which often rely on the reassignment of naturally occurring codons to a particular UAA.\textsuperscript{27} Methionine has frequently been employed for this technique, since it is one of the least common natural amino acids.\textsuperscript{28} Nevertheless, global replacement of methionine limits uses \textit{in vivo} and sometimes leads to cellular toxicity at the high concentrations required.\textsuperscript{13}

In 2001, Schultz \textit{et al.} pioneered a versatile strategy for the incorporation of UAAs into proteins.\textsuperscript{29} By introducing an orthogonal tRNA/aminocyl-tRNA synthetase pair from \textit{Methanocaldococcus jannaschii} into the \textit{E. coli} translational machinery, the synthetic amino acid \textit{O}-methyl-\textit{L}-tyrosine was incorporated into a protein, in response to suppression of the amber stop codon. In follow-up reports, this strategy has been extended to incorporate a diverse range of over 150 UUAs possessing varied structures and reactive handles.\textsuperscript{30,31} In particular, the introduction of amino acids containing azido- 1, alkynyl- 2 and keto- 3 side chains (Figure 2) has provided new avenues for the generation of conjugates through bioorthogonal chemical ligation strategies, such as 1,3-dipolar cycloadditions, Staudinger and oxime ligations (\textit{vide infra}).\textsuperscript{32,33}

![Figure 2](image)

\textbf{Figure 2} – Examples of unnatural amino acids; 1 azidohomoalanine, 2 homopropargylglycine, 3 \textit{p}-acetylphenylalanine.

Whilst this technology has enabled modification of biomolecules with an unprecedented degree of site-selectivity, it is not easily transferrable. Laborious screening procedures are required to identify suitable positions for modification and substitution for the UAAs can often be incomplete, leading to poor expression yields.\textsuperscript{34,35} In addition, more widespread access to the required plasmids
is needed, as well as global applicability towards all protein systems. In order to avoid such inherent limitations, an ideal site-directed conjugation technique would utilize residues natural to the protein that are revealed for modification only under defined conditions.

1.1.2 Bioorthogonal modification of inserted functionalities

Within the last two decades, bioorthogonal reactions have become essential tools for chemical biologists. A term originally coined by Bertozzi in 2003, bioorthogonal chemistry refers to reactions in which the coupling partners react selectively without being compromised by the heterogeneous milieu of the protein solution or living cell. Ideally, the reactions are fast and at the same time they do not perturb the biological setting in any way other than intended.

‘Click’ transformations (i.e. a set of near-perfect bond-forming reactions) reflect these underlying chemical principles and have gained momentum in the selective labeling of biomolecules. Historically, the first bioorthogonal ‘click’ reactions involved the condensation of ketones or aldehydes with α-effect amines, such as hydrazides or aminoxy compounds to afford hydrazone or oxime linkages (Scheme 1a).

Labeling of biomolecules in vivo has been somewhat limited, since the optimum pH for these transformations is 5-6, although notably Schultz et al. reported the intracellular labeling of a ketone-functionalized protein. The first truly bioorthogonal reaction that takes place at physiological conditions was reported in 2000 by Bertozzi et al. and is a modification of the classical Staudinger reaction of phosphines and azides. This transformation, termed the Staudinger ligation involves an intramolecular electrophile that traps the generated azalylide intermediate to afford an amide and the corresponding phosphine oxide (Scheme 1b).

Scheme 1 – Examples of bioorthogonal reactions; a. hydrazone and oxime ligation, b. Staudinger ligation.
A major breakthrough occurred in 2001, when the Sharpless\(^{44}\) and Meldal\(^{45}\) research groups reported a modification of a classical reaction of organic chemistry, the azide-alkyne Huisgen 1,3-dipolar cycloaddition. They discovered that the cycloaddition of azides with terminal alkynes to produce 1,4-disubstituted 1,2,3-triazoles can be dramatically accelerated by the use of Cu(I) catalyst, even at room temperature and aqueous conditions. Computation studies revealed that the copper-catalyzed azide-alkyne cycloaddition (CuAAC) is not a true cycloaddition, but proceeds \textit{via} a stepwise mechanism involving unprecedented metallacycle intermediates (Scheme 2).\(^{46}\) The first report of CuAAC as a bioconjugation strategy was demonstrated by Finn \textit{et al.} through the attachment of dyes to cowpea mosaic virus.\(^{47}\) Since then, the reaction has been rapidly embraced and found widespread use in labeling and imaging.

However, the usefulness of CuAAC in living systems has been hindered by the toxicity of copper.\(^{48}\) Oxidative stress and biological damage have been attributed to Cu(I)-promoted generation of reactive oxygen species, which can affect the structural and functional integrity of biomolecules.\(^{49}\) Research by Finn and coworkers has demonstrated that the use of ligands (such as THPTA) can intercept and quickly reduce reactive oxygen species without compromising the rate of CuAAC reaction very much.\(^{50}\)

\begin{center}
\textbf{Scheme 2} – Mechanistic overview of the copper-catalyzed azide-alkyne cycloaddition (CuAAC).
\end{center}
Building on work by Wittig and Krebs in the 1960s,\textsuperscript{51} who found that cyclooctyne reacts rapidly with phenylazide to afford a single triazole product, Bertozzi \textit{et al.} reported a strain-promoted azide-alkyne cycloaddition (SPAAC) in 2004 (Figure 3).\textsuperscript{52} This reaction requires no exogenous ligands or catalysts and is promoted by the bond angles of sp-hybridized carbons being 160°, i.e. less distortion energy is required to move towards the transition state.\textsuperscript{53} Still, the first generation of SPAAC was considerably slower than CuAAC. The reactivity of the cyclooctyne ring has since been modulated by e.g. appending electron withdrawing groups at the propargylic position (DIFO),\textsuperscript{54} or by augmentation of strain energy through aryl ring (DIBO)\textsuperscript{55} or cyclopropyl ring (BCN) fusion.\textsuperscript{56} The applicability of SPAAC has been realised in a variety of \textit{in vivo} context, including mammalian cells and animals.\textsuperscript{57} However, limitations remain, such as difficulty in synthesis and handling,\textsuperscript{13} while a degree of incompatibility towards cysteine residues has been reported.\textsuperscript{58}

\textbf{Figure 3} – Strain-promoted azide-alkyne cycloaddition (SPAAC); a. mechanism, b. examples of cyclooctynes.

\subsection*{1.1.3 Modification of native amino acids}

As applications of bioconjugates become more complex, the development of reliable chemoselective strategies for the introduction of functional moieties into proteins with high selectivity has become vital.\textsuperscript{59} Over the past decades, extensive research into the modification of native amino acids has been undertaken, since they offer immediate accessibility without the need for more specialist techniques.\textsuperscript{60} The side chains of proteinogenic amino acids present a wide range of functionalities with their reactivities varying considerably depending on their locations and the influence of nearby residues with which they interact. Most of these have been explored and strategies for the selective modification of cysteine,\textsuperscript{13} lysine,\textsuperscript{61} histidine,\textsuperscript{62} tyrosine,\textsuperscript{63} serine,\textsuperscript{64} arginine,\textsuperscript{65} tryptophan,\textsuperscript{66} as well as aspartic or glutamic acid\textsuperscript{67} and even methionine\textsuperscript{68} have been developed (Figure 4).
As a result of their strongly nucleophilic side chains, lysine and cysteine residues have been shown to be the most convenient targets for modification. To this end, research efforts have focused on developing reactions that exploit the reactivity of these two amino acids.

1.1.3.1 Lysine modification

Labeling via the primary amino groups on lysine residues has been the oldest and most straightforward method for protein modification.\(^6\)\(^1\) Whilst multiple accessible lysines are present on the surface of a protein resulting in efficient labeling, heterogeneous mixtures are inevitably obtained.\(^6\)\(^9\) The applicability of this method is therefore dependent upon the properties of the protein and the application required.

Over the years, a plethora of methods have been developed for lysine modification; a few examples are shown in Scheme 3. These include direct reaction with activated esters (e.g. \(N\)-hydroxysuccinimide NHS),\(^6\)\(^1\) iso(thio)cyanates,\(^7\)\(^0,7\)\(^1\) as well as reductive amination reactions with aldehydes (Scheme 3a-c).\(^7\)\(^2\) However, several issue of chemoselectivity have been identified with these methods. Reactions with alkylation/acylation agents are not always entirely selective for amines, i.e. cross-reactivity with surface accessible nucleophiles (e.g. Cys, Tyr, Ser, Thr side chains) has been observed.\(^6\)\(^0\) In addition, the use of cyanoborohydride in reductive amination has
the potential for reducing disulfide bonds within proteins. McFarland and Francis have reported an iridium-catalyzed alternative to this method, but with lower efficiency.\textsuperscript{73}

The recent needs of well-defined bioconjugates have prompted the development of next generation lysine modifications that avoid cross-reactivity with other residues. A highly selective protocol was reported by Tanaka \textit{et al.} in 2008,\textsuperscript{74} which is based on a rapid 6π-aza-electrocyclization (Scheme 3d). The reaction was shown to exhibit very rapid kinetics with selectivity towards solvent-accessible lysines (i.e. 10-30 min at 24 °C). A recent report by Carreira and coworkers described the use of diazonium terephthalates as an amine-selective method to afford hydrolytically stable triazin-4(3H)-ones (Scheme 3e).\textsuperscript{75} Gois \textit{et al.} have also developed a strategy for the modification of both the ε-amino group of lysines and the N-terminus, based on

\textbf{Scheme 3} – Examples of lysine modification methods; (a) amide formation, (b) urea/thiourea formation, (c) reductive amination, (d) 6π-aza-electrocyclization, (e) reaction with diazonium salts, (f) iminoboronate formation.
iminoboronate formation (Scheme 3f); this transformation is reversible upon the addition of fructose, dopamine or glutathione.\textsuperscript{76}

While recent advances can ensure chemoselectivity in lysine modification, site-selectivity remains a challenge that cannot be addressed as a result of the high natural abundance of lysine, although notably there are a few reported cases of singly modified lysines, but this approach is protein-specific and not general.\textsuperscript{77,78} This is particularly important when developing therapeutic protein conjugates; the wide distribution of loading levels afforded by lysine modification will result in heterogeneous mixtures that are associated with aggregation and inconsistent pharmacokinetic profiles and toxicity.\textsuperscript{79} In addition, lysine residues within or proximal to the active or binding site of a protein might be modified, which would result in loss of activity.\textsuperscript{80} In contrast, the low natural abundance of cysteine can offer a valuable handle for highly selective protein modification.

### 1.1.3.2 Cysteine modification

Free cysteine residues are extremely rare in proteins with an abundance of 0.2% relative to all naturally occurring amino acids.\textsuperscript{81} Their surface accessibility is limited and they are found mostly in hydrophobic environments.\textsuperscript{82} This is to ensure that the highly nucleophilic thiolate is preserved and hence able to fulfill its biological role. This involves being part of catalytic residues, e.g. in cysteine proteases and P-loop phosphatases that assist in the metabolism of various biochemicals.\textsuperscript{83,84} Cysteine residues can also be used in chelating metal ions, e.g. in metallo-proteins, iron-sulfur clusters are important in cellular events such as DNA synthesis or mediating one-electron redox processes.\textsuperscript{85}

In addition to cysteine’s low natural abundance, the ease of its introduction by site-directed mutagenesis allows access to protein assemblies with a single cysteine at a predetermined site.\textsuperscript{86} Even if multiple cysteines are present in a protein, this multiplicity is much smaller compared to lysines, making thiol-reactive labeling the preferred approach. It has been shown that the thiolate anion is 5x10\textsuperscript{10} times more reactive than the protonated thiol towards nucleophilic attack;\textsuperscript{87} therefore due to its relatively low pK\textsubscript{a} (i.e. 8.3, vs. 10.5 for lysine), at physiological pH cysteine is the most robustly nucleophilic of the 20 canonical amino acids.\textsuperscript{13} Cysteine-targeting chemistry is well established and widely employed to create bioconjugates; a few popular reactions are shown in Scheme 4. These involve alkylation with α-halocarboxyls,\textsuperscript{88} Michael acceptors (maleimides,
vinyl sulfones and related $\alpha,\beta$-unsaturated systems)\textsuperscript{11} and electron-deficient alkynes\textsuperscript{89} or disulfide formation with activated thiols.\textsuperscript{90} In addition, oxidative elimination of cysteine to dehydroalanine can provide a unique chemical handle for further conjugation with thiols or amines through thiol-or aza-Michael ligation.\textsuperscript{91,92}

![Scheme 4](image-url)

**Scheme 4** – Examples of cysteine modification methods; (a) alkylation with haloacetamides, (b) maleimides, (c) vinyl sulfones, (d) electron deficient alkynes, (e) disulfide formation, (f) oxidative elimination to dehydroalanine; MSH = $O$-mesitylenesulfonylhydroxylamine.

Maleimides are perhaps the most often used functional groups for bioconjugation. They were introduced by Friedmann in 1949, where it was shown that a cysteine-containing tripeptide (i.e. glutathione) was able to react quantitatively with maleimide in 2 minutes at room temperature.\textsuperscript{93} Ever since, maleimides have gained in popularity due to their exceptionally fast kinetics and high cysteine-selectivity; in comparison iodoacetamides are known to cross-react with methionine, histidine or tyrosine.\textsuperscript{94} To date, a plethora of maleimide-based modifying reagents containing various chemical probes (e.g. fluorophores, drugs, PEG chains, biotin) are commercially available with even more being synthesized in laboratories around the world.
Despite their widespread utilization, there are some limitations associated with the use of maleimides. It has recently come to light that the resultant thioether-succinimide adducts are not stable in plasma. They are shown to undergo a retro-Michael type reaction, releasing the payload which is then captured by blood thiols, particularly serum albumin. By hydrolyzing the thiosuccinimide ring of maleimide-based conjugates to a succinic acid improved stability and efficacy can be obtained (Scheme 5). However, the yield of the hydrolyzed thiol-stable conjugates might significantly decrease (often to 50-60%), due to a competing retro-Michael pathway during hydrolysis, which cannot be suppressed. In addition, it might be desirable for in vitro or in vivo applications to regenerate the unmodified protein by controlled disassembly of the conjugate. While maleimides are shown to undergo retro reactions, this is a very slow and inconvenient protocol for reversibility.

**Scheme 5** – Hydrolysis of the thiosuccinimide adduct to achieve thiol stability and competing retro-Michael reaction during hydrolysis.

In 2009-2011, Baker, Caddick and coworkers reported on the next generation maleimides (NGMs), which have greatly expanded the class of reagents for the rapid, selective and reversible modification of cysteine (Scheme 6). The incorporation of a leaving group (e.g. bromine) across the maleimide double bond allows an addition-elimination sequence to take place upon reaction with cysteine to afford a thiomaleimide moiety. Treatment of the L111C mutant of the Grb2 adaptor protein with monobromomaleimide at pH 8.0, 0 °C led to exclusive formation of species in 1 hour, even with 1 equivalent of reagent. The conjugation was reversible as the construct could be cleavable with excess thiol. In addition, the fact that the maleimide double bond is retained enables a second thiol addition to take place giving access to dithiosuccinimide.
successfully resolving another drawback of classical maleimide methodologies, i.e. the presence of only two points of attachment.

![Scheme](image)

**Scheme 6** – Protein modification using next generation maleimides (NGMs).

Importantly, it proved possible to convert a reversible thiomaleimide bioconjugate to an irreversible one, by quantitative hydrolysis to maleamic acid 9, which exhibits a dramatic decrease in its reactivity towards thiols. Hydrolysis is usually accomplished upon incubation of the construct at increased pH and temperature (e.g. pH 8, 37 °C).\(^\text{101}\)

Since their conception, next generation maleimides have been successfully employed in protein modification and their use has significantly expanded to a variety of applications, including reactive handles in polymers,\(^\text{102}\) fluorescence ‘turn-on’ reagents,\(^\text{103}\) protein PEGylation,\(^\text{104}\) nuclear imaging,\(^\text{105}\) photochemistry,\(^\text{106}\) photodynamic therapy,\(^\text{107}\) as well as the construction of antibody-drug conjugates and bispecific antibodies (*vide infra*).

Following on from the successful development of NGMs, Chudasama and Caddick *et al.* reported on the use of bromopyridazinediones (PDs) in 2011 for the preparation of cleavable bioconjugates (Scheme 7).\(^\text{108}\) These reagents demonstrate exceptional hydrolytic stability, which is crucial in preserving their thiol-cleavable property. In addition, they have the potential for four points of chemical attachment. Due to the decrease in the ring strain and the presence of two nitrogen atoms that donate more electron density into the ring, PDs are shown to be less reactive than NGMs. Even though initial protocols required 100 equivalents of reagent to achieve completion, recent advances have shown that due to their hydrolytic stability, less equivalents can be used for longer
reaction periods, with PDs finding important applications in the preparation of antibody-drug conjugates.\(^{109}\)

![Scheme 7 - Protein modification using pyridazinediones (PDs).](image_url)

### 1.1.3.3 Disulfide bridging

While the introduction of a single cysteine by site-directed mutagenesis can be straightforward, even in a full antibody, the presence of an unpaired cysteine can have a detrimental impact in the production yield and might result in disulfide scrambling, complicate re-folding or lead to aggregation.\(^{110,111}\) To avoid the need for genetic engineering, the modification of native disulfide bonds has attracted growing interest in recent years for site-selective modification. Disulfide bonds are present in most proteins to control their stability and biological activity with a low natural abundance of 1.5%; i.e. the total cysteine content is 1.7%, with 0.2% of free cysteine residues.\(^{112}\) Disulfide bridges contributing to the protein’s biological activity are usually buried within the protein’s hydrophobic regions to ensure that they are not altered by endogenous or exogenous modifications. In contrast, solvent accessible disulfides mainly impart stability to the protein and represent an opportunity for modification.\(^{113–115}\)

Disulfide bonds are considered to be weak links and can easily be reduced with small thiols (e.g. DTT) or phosphines (e.g. TCEP).\(^{116,117}\) Upon reduction of a disulfide bond, the two exposed cysteine residues could be targeted for modification as independent free thiols. However, depending on the protein, this might cause major disruption and destabilization of the native conformation.\(^{118}\) To this end, it is desired to covalently re-connect the free thiols with the view to retaining the 3D structure and function of the protein. The first successful attempt to modify a protein via disulfide bridging was reported by Broccolini and coworkers in 2006.\(^{119}\) Following disulfide bond reduction, the two thiols react in sequence with an enone-sulfonyl reagent via an
addition-elimination step, followed by a second addition reaction (Scheme 8). This eventually results in the two cysteines being re-connected via a three-carbon bridge.

![Scheme 8 – Mechanism of disulfide bridging with enone sulfanyl reagents.](image)

The mechanism has been further elucidated by NMR spectroscopy, where it was found that diastereomers are formed, due to the introduction of a new stereogenic centre. In addition, it has been shown that the rebridging is quite slow (up to 16 h) and does not always proceed with reagents bearing sterically hindered functionalities. For example, Weil and coworkers observed that reaction of reduced somatostatin with enone sulfanyl reagents containing either G2 PAMAM dendron or doxorubicin resulted in complex mixtures of products that still contained unconsumed starting materials. Nevertheless, it was shown that such bulky substituents can be introduced into the polypeptide by click chemistry after re-bridging or by direct attachment upon introducing a long PEG spacer between the reagent and the functionality.

Having successfully developed next-generation maleimides (NGMs) for single cysteine modification, Baker and Caddick et al. envisaged that the incorporation of two leaving groups on the maleimide double bond could enable two consecutive addition-elimination sequences to take place upon reaction with a reduced disulfide bond to reconnect it with a rigid two-carbon bridge (Scheme 9).

![Scheme 9 – Mechanism of disulfide bridging with next-generation maleimides (NGMs).](image)
Initial investigations were performed using the cyclic peptide hormone somatostatin that contains one single disulfide bond. Following reduction, bridging with dibromomaleimide was shown to proceed selectively and rapidly (i.e. within 2 minutes), without introducing a chiral centre and having the added advantage that the formed linker is potentially cleavable in a reversible manner.\textsuperscript{99,102} Through further reagent development it was observed that by substituting the leaving groups for thiophenol, the resultant dithiomaleimide exhibits no cross-reactivity with TCEP; unlike dibromomaleimide that slowly reacts with it.\textsuperscript{104} As a result, it can be used in tandem with the reducing agent, representing the first reagent for efficient \textit{in situ} bridging of disulfide bonds. This is particularly important for structurally sensitive proteins that might start to unfold in the time that has elapsed between disulfide cleavage and bridging. Disulfide bridging of proteins with NGMs has been shown to proceed with full retention of biological activity and a plethora of functionalities, such as PEG chains,\textsuperscript{104} fluorophores,\textsuperscript{122} drugs\textsuperscript{123} or even a spin label\textsuperscript{124} have successfully been incorporated into not only peptides, but also antibodies and antibody fragments.

Since its first inception, disulfide functionalization has gained increasing momentum for site-selective modification and various novel rebridging agents have been reported in recent years; an overview is shown in \textit{Scheme 10}. Dibromopyridazinediones, developed by Caddick and Chudasama \textit{et al}., are shown to efficiently bridge disulfide bonds with the advantage of two bioorthogonal handles for chemical functionalization (Scheme 10a).\textsuperscript{125} Wilson and Davis recently reported on trivalent arsenous acid (\textit{As}\textsubscript{III}) derivatives (Scheme 10b).\textsuperscript{126} Conjugation to the therapeutic peptide salmon calcitonin was shown to proceed rapidly (i.e. less than 2 min), with the products exhibiting low levels of toxicity compared to free arsenic acids. A potential drawback of \textit{As}\textsubscript{III} reagents might be their sensitivity to oxidation (i.e. to \textit{As}\textsubscript{V} species), especially during lengthy purifications. In addition, the authors suggest that these arsenous acid derivatives demonstrate enhanced specificity for disulfide bond bridging in the presence of free cysteine residues. However, this cannot be conclusive since the reaction was only performed with Bovine Serum Albumin (BSA), a protein that is known to contain a very hindered free cysteine residue.\textsuperscript{127} An interesting disulfide bridging strategy of using dichloro-s-tetrazine was developed by Smith and coworkers (Scheme 10c).\textsuperscript{128} Bridging of thioredoxin was shown to proceed in 1 minute and the reaction was reversible upon photochemical irradiation of the tetrazine-labeled construct to regenerate the disulfide bond. In addition, the incorporation of a tetrazine moiety as a
Scheme 10 – Examples of disulfide bridging reagents; (a) pyridazinediones, (b) organic arsenals, (c) dichloro-s-tetrazine, (d), dichloroacetone, (e) 3,3-bis(bromomethyl)oxetane, (f) alkynes (thiol-yne click). Bioorthogonal handle enabled later functionalisation via inverse electron demand Diels-Alder reaction. 1,3-dichloroacetone has also seen applicability in disulfide stapling to afford an acetone-linked bridge (Scheme 10d). Using this approach, Hu and Micoli et al. functionalized CRM197 protein carrier. This was further reacted via oxime ligation with an aminoxyl linker bearing an azide that enabled a subsequent SPAAC reaction with the Salmonella O-antigen with the view to investigating the impact of this protein-antigen coupling. In 2017, Bernardes and coworkers reported on the use of 3,3-bis(bromomethyl)oxetane as a disulfide bridging reagent (Scheme 10e). Even though further functionality was not imparted, oxetane grafts were shown to enhance stability and activity of the proteins and antibody fragments examined; it was suggested that this
was due to the ability of oxetanes to modulate parameters, such as solubility, basicity, lipophilicity and metabolic stability. However, a drawback of this approach is the long reaction times required for disulfide bridging to proceed (i.e. 24-72 h), which might increase the risk of structurally sensitive proteins starting to unfold. A photochemical disulfide stapling approach through a thiol-yne coupling has also been developed recently (Scheme 10f).\textsuperscript{131} Disulfide rebridging with terminal alkynes was achieved upon irradiation with UV light (365 nm) in the presence of a radical initiator. Despite the novelty of the approach and the commercial availability of various alkynes, conjugation yields were shown to be as low as 20%.

Notably, apart from imparting functionality in proteins, disulfide bridging has also been successfully employed in peptide “stapling”. This involves the cross-linking of cysteine residues, positioned in an $i, i+4/7$ fashion in the peptide backbone to form conformationally stable $\alpha$-helices, overcoming the intrinsically disordered nature of linear peptides.\textsuperscript{132} Using this macrocyclization strategy, a number of bioactive peptides derived from poor therapeutic candidates have been generated for various applications, including the targeting of “undraggable” intracellular protein-protein interactions.\textsuperscript{133,134}

A few examples of disulfide stapling reagents that have been used for the conformational stabilization of peptides are shown in Scheme 11. Interestingly, dibromomaleimide has been employed by Wilson and coworkers to constrain peptide sequences, with the resultant conjugates showing enhanced $\alpha$-helical conformation and resistance to proteolysis (Scheme 11a).\textsuperscript{135} Importantly, they exhibit retained or increased potency of inhibition of protein-protein interactions of the Bcl-2 family of regulatory proteins. Pentelute \textit{et al.} have reported a peptide stapling approach involving arylation with perfluoroaromatic reagents, such as a hexafluorobenzene species (Scheme 11b).\textsuperscript{136} The resultant staples were shown to increase the $\alpha$-helical content of the peptide, as well as the target binding and proteolytic stability. One limitation of this approach is the poor water solubility and low reactivity of perfluorobenzene reagents in aqueous environments, however, the presence of NMR-detectable $^{19}$F atoms can be exploited as an additional analytical tool in peptide structural investigations. Pentelute and Buchwald have also reported on a bis-palladium organometallic reagent for cross-linking two cysteine residues in a peptide to effect macrocyclization (Scheme 11c).\textsuperscript{137} Although further structural or biophysical characterization of the stapled peptide was not provided, it demonstrates the viable application of palladium-mediated stapling in biological environments. An interesting photochemical thiol-ene coupling approach
utilizing an $\alpha,\omega$-diene has been reported by Chou and coworkers (Scheme 11d).\textsuperscript{138} They synthesized stapled p53 tumour suppression protein mimetics, which were shown to inhibit p53-MDM2 interactions (which are responsible for degrading p53), leading to enhanced antitumour activity. However, peptide stapling was performed in N-methyl-2-pyrrolidine and hence the feasibility of this approach in aqueous media is yet to be determined.

**Scheme 11** – Examples of disulfide stapling reagents to form conformationally stable $\alpha$-helices to increase the bioactivity of peptides; (a) dibromomaleimides, (b) hexafluorobenzenes, (c) bis-palladium organometallic reagents, (d) $\alpha,\omega$-dienes.

Overall, disulfide bridging offers a robust and versatile strategy to impart functionality in proteins, as well as to staple peptides, with the view to creating bioactive analogues. The new generation of bridging reagents hold great promise for homogeneous and site-selective modifications, though challenges still need to be overcome.
1.1.4 Chemical cross-linking of biomolecules

The connection of two or more proteins with different binding specificities or activities can lead to unique and powerful systems, with significant potential in therapeutic applications. A plethora of active bifunctional proteins have been produced by genetic methods (e.g. fusion via their N- and C- termini) for various indications, including autoimmune diseases, inflammatory disorders, tissue repair and oncology. However, the modification of the N- or C- terminus can compromise the proteins’ bioactivities, whilst complications during folding and processing often result in low expression yields, limiting productivity and increasing costs. An alternative and more versatile approach is offered by chemical crosslinking. Homobifunctional and heterobifunctional reagents have been developed that target the most reactive lysine and cysteine residues.

To date, N-succinimidyl-4-(maleimidomethyl) cyclohexanecarboxylate (SMCC) has been one of the most frequently applied heterobifunctional reagents, used in crosslinking amine to thiol-containing proteins (Scheme 12). In a typical conjugation protocol, the most labile end of the linker (i.e. NHS ester) will firstly react with the amine groups of the first protein to afford a maleimide-activated intermediate. While maleimide groups hydrolyse in aqueous conditions, they display extended stability compared to NHS esters, therefore allowing purification of the activated maleimide intermediate before adding the sulphhydryl-containing protein into the conjugation.

![Scheme 12 – Cross-linking of an amine-containing protein to a thiol-containing protein using SMCC.](image-url)
Over the years a plethora of linkers have been developed with different reactive groups and spacers in response to the ever growing applications of protein-protein conjugation; a few examples are shown in Figure 5.

![Examples of bifunctional reagents.](image)

Figure 5 – Examples of bifunctional reagents.

Typically, heterobifunctional lysine-to-cysteine reagents consist of an amine-reactive activated ester, such as NHS, pentafluorophenol or p-nitrophenol. Efforts to increase the water solubility of activated esters have led to the development of sulfo-NHS or sulfo-tetrafluorophenol esters, used in biological media without the risk of precipitation. The other end of the linker usually contains a maleimide, an activated disulfide or a haloacetamide for thiol modification.

Interestingly, while disulfide exchange is not a stable modification, since it can be cleaved upon reaction with thiols, the incorporation of methyl groups (as in SMPT, Figure 5) make the resultant disulfide bond hindered and hence more stable in vivo. A very important component of all cross-linking reagents is the spacer between the two reactive functionalities. The length of this spacer can be crucial in determining the efficiency of the conjugation, especially if sterically hindered proteins are used, where usually a longer linker might be necessary. In addition, hydrophilic spacers (e.g. polyethylene glycol PEG or 1,3-dioxane) can confer water solubility not only to the reagent, but also to the resultant bioconjugate. A number of cross-linking reagents contain cleavable groups within their cross-bridge, such as a disulfide bond (as in DSP, Figure 5), lending greater flexibility to the experimental design. The constituents present in the spacer could also affect the reactivity of their functional groups. For instance, it is known that a maleimide group that has an aromatic ring immediately next to it (as in pPDM, Figure 5) is less stable to hydrolysis than a maleimide that has an aliphatic ring adjacent to it. In general, heterobifunctional reagents targeting two different reactive groups are known to provide better control over the conjugation,
compared to homobifunctional reagents (e.g. DSP, BM(PEG)₃, Figure 5). This is because the latter will target the same functionalities and hence upon reaction with the first protein, undesired dimerization or polymerization could be obtained.\textsuperscript{151} However, homobifunctional reagents have the advantage of a much easier chemical synthesis and potential dimerization can be avoided by adding an excess of the linker over the first protein.

As previously explained, lysine modification yields heterogeneous products of inconsistent pharmacokinetic profiles and toxicity. Consequently, thiol-to-thiol coupling appears to be advantageous for the preparation of structurally defined and homogeneous protein-protein conjugates.\textsuperscript{152} To date, most bifunctional reagents that enable thiol-to-thiol conjugation are homobifunctional, with bis-maleimide reagents being the most reactive. While cysteine-containing proteins are rare in nature, site-directed mutagenesis can give access to proteins with a single cysteine at a predetermined site, hence enabling site-specific protein-protein cross-linking to occur.\textsuperscript{153,154} Furthermore, selectivity could be achieved by targeting native disulfide bonds, without the need for prior engineering. However, apart from the structural integrity being compromised, it is not possible to obtain a homogeneous 1:1 protein-protein conjugate, using common cross-linking reagents, due to the presence of two cysteine residues per disulfide bond.\textsuperscript{155,156} If both cysteines are functionalized, a trimer could be obtained upon addition of the second protein. To this end, it is highly desirable to create protein-protein conjugates \textit{via} disulfide bridging.
1.1.5 Native Chemical Ligation

Further to conjugating biomolecules using cross-linking reagents, the development of native chemical ligation (NCL) by the Kent laboratory in the early 1990s was a breakthrough for protein synthesis, enabling the direct and chemoselective reaction of two unprotected peptide segments (Scheme 13).\(^\text{157}\) In the initial study, NCL was employed to synthesize human interleukin 8 from a C-terminal thioester peptide and another peptide containing an N-terminal cysteine residue. The mechanism of the reaction involves an initial reversible transthioesterification step, which is the rate determining step, followed by a spontaneous intramolecular S-to-N acyl transfer to form a native peptide bond at the ligation site, driven by the thermodynamic stability of an amide bond over a thioester.\(^\text{158}\) The key acyl transfer step is promoted by the favourable geometric orientation of the thioester and amine.\(^\text{159}\) Without this proximity effect other amines could react to afford a heterogeneous mixture.\(^\text{13}\)

\[\text{Scheme 13} \quad \text{Synthesis of human interleukin 8 (IL-8) by native chemical ligation.}\]

Since its introduction, NCL has been embraced by the scientific community and extensive optimization has been carried out by various research groups. It has been observed that the identity of the amino acid located at the C-terminus of the thioester can significantly affect the ligation rate.\(^\text{112}\) Dawson and coworkers reported that the fastest reaction rate was observed for glycine thioester, reacting quantitatively in less than 4 hours.\(^\text{160}\) In contrast, Val, Ile, Thr and Pro thioesters exhibit lower ligation rates (>24 h), which are shown to be accelerated by transforming the corresponding thioesters into selenoesters\(^\text{161}\) or by fine-tuning the reaction pH.\(^\text{162}\) In addition, the initial thiol-thioester exchange is observed to occur much more efficiently with aryl thioesters than alkyl thioesters. However, to facilitate synthesis and handling, peptide thioesters are usually synthesized as the less reactive alkyl derivative (e.g. MESNa thioester) and converted \textit{in situ} to the corresponding aryl thioester upon addition of aryl thiol.\(^\text{112}\) After examining fourteen different
thiol catalysts, Johnson et al. demonstrated that MPAA was the most efficient (Scheme 14a). In 2001, Yan and Dawson introduced a post-ligation desulfurization approach to afford alanine (Scheme 14b). This strategy not only gave access to cysteine-free peptides, but inspired the development of various ligation junctions, such as Phe, Gly, Val, Leu, Pro, Glu, and Lys, which after desulfurization can give access to a nascent residue of interest (e.g. for Phe, Scheme 14c).

![Scheme 14](image)

**Scheme 14** – a. Thiol-thioester exchange with aryl thiols; b. post-ligation desulfurization to afford alanine; c. post-ligation desulfurization to afford phenylalanine.

Apart from RANEY® nickel that was initially employed for desulfurization, various other conditions have been developed, such as nickel boride, Pd/ Al₂O₃, or even metal-free conditions. Prior to NCL, the only available method for peptide synthesis was solid phase peptide synthesis (SPPS), which was restricted to ca. 50 amino acids. To date, NCL, in combination with SPPS, has been successfully employed for the total synthesis of unmodified or post-translationally modified proteins, the selective introduction of modifications or labels, as well as conjugation of peptides with other macromolecules, such as DNA. Furthermore, Shultz et al. recently reported on the genetic incorporation of a thioester non-canonical amino acid into green fluorescent protein (GFP), significantly expanding the scope of NCL.

Interestingly, in 2009, Seitz and coworkers introduced the concept of internal cysteine ligation, involving acyl transfer from the cysteine to the N-terminal α-amino group via macrocyclic transition states (TS) (Scheme 15). While in NCL, the initial transthioesterification is the rate determining step (RDS), in ligations proceeding via larger ring sizes (>6 membered rings), the S-to-N acyl transfer becomes rate-determining. The ring size of the TS has been identified as a critical factor for the successful outcome of the process. In the original publication, Seitz et al. observed that 17- to 20- membered macrocyclic TS were optimal, while 8-, 11- or 14-membered
macrocycles were difficult to form, with slower amide bond formation resulting in increased competition from thioester hydrolysis.\textsuperscript{178} Overall their method required relatively long reaction times (48-72 h) and isolated yields of the ligation products were not reported. Since then, a series of mechanistic studies have determined that apart from the ring size of the TS, other factors are involved in the outcome of S-to-N acyl transfer. These include conformation, hydrogen-bonding, NH-π interactions, pH, solvent, reactivity/stability of the thioester and structure of the amino acids near the ligation site.\textsuperscript{159,180,181}

\begin{center}
\textbf{Scheme 15} – Peptide coupling via long range S-to-N acyl transfer.
\end{center}

Katritzky and coworkers investigated S-to-N acyl migrations proceeding \textit{via} 11- or 14-membered TS, under microwave irradiation in phosphate buffer (pH 7.3) at 50 °C.\textsuperscript{182} Under these conditions, shorter reaction times (1-3 h) were achieved, with 64% and 55% conversion to the target product being observed, respectively. Following HPLC purification, isolated yields of 26% and 23% were reported. Furthermore, S-to-N acyl transfer reactions have been observed to occur in nature over very large ring sizes, often with a high degree of site-specificity (e.g. in enzymatic ubiquitination, where the Cys-containing E2 enzyme forms a thioester bond with the C-terminal glycine residue of ubiquitin; the E3 ligase enzyme then catalyzes S-to-N acyl transfer to a lysine-containing substrate).\textsuperscript{183} In this case, acyl transfer is considered to be a proximity-induced process that occurs when the participating reactive groups are favourably positioned in space.
1.2 Antibodies for targeted therapeutics

The idea of a “magic bullet” that could be developed to selectively target disease was firstly hypothesized by Paul Ehrlich a century ago.184 This vision came closer to reality in the mid-1970s with the development of hybridoma technology by Köhler and Milstein that enabled the production of monoclonal antibodies (mAbs) with high specificity and affinity for their target antigens.185 In 1983, Muromonab-CD3 was the first monoclonal antibody to be approved by FDA for use in acute transplant rejection.186 However, the murine origins of this antibody resulted in complications not originally envisioned; high immunogenicity and poor ability to recruit immune effector functions, also known as ‘human anti-mouse antibody’ (HAMA) response.186,187 It was the advent of genetic engineering technology during the 1990s that helped to resolve many of these inherent limitations by enabling the antibody chimerization and even complete humanization.186,188 Since then antibodies have emerged as a powerful new class of therapeutics, with 21 antibodies approved for therapeutic use and more than 300 currently in development.189 Remarkably, adalimumab (Humira™), a monoclonal antibody indicated for the treatment of autoimmune diseases has been the world’s best-selling drug for years.190

Antibodies are macromolecular glycoproteins that are produced by B-cells in the blood upon activation of the adaptive immune system.191 They comprise a pair of identical heavy and light chains, linked by disulfide bonds. In mammals, antibodies are classified according to the heavy chain they contain into five main isotypes, i.e. IgA, IgD, IgE, IgG and IgM.189 These can differ in the sequence, number of constant domains, structure of the hinge region and valency. Antibody light chains also contain two isotypes, i.e. κ and λ, with κ being the more common. The IgG isotype is the most abundant in human serum (70-75% of the total immunoglobulin pool) and is further subdivided into four allotypes (IgG1, IgG2, IgG3 and IgG4).192 These are highly homologous, but differ in the architecture of the hinge region. Most therapeutic mAbs with human backbones are of the IgG1 type because it is arguably the most active in the immune system.193 IgG1 antibodies are monomers of ca. 150 kDa. Both heavy and light chains contain one variable domain (V_H, V_L), which is responsible for antigen binding. As such, they exhibit great variation in their amino acid composition. In addition, three constant domains (C_H1, C_H2, C_H3) are present in each heavy chain and one (C_L1) in each light chain. The four chain are held together by strong non-covalent interactions and four disulfide bonds (two in the hinge region and one in each switch region) giving rise to a Y-shaped molecule (Figure 6).187
The common hallmark of all antibodies is highly selective antigen binding, which is mediated via the complementarity-determining regions (CDRs), three of which are present in each of the V\textsubscript{H} and V\textsubscript{L} fragments.\textsuperscript{187,194} The action of antibodies though is twofold; apart from binding their target antigens on tumour cells, they should elicit cell death. Their mechanism of action can involve blocking an activation signal that is necessary for continued cell growth, signaling that leads to cell death or Fc mediated effector functions. These include complement-dependent cytotoxicity (CDC, upon binding to C1q receptor) or antibody dependent cell-mediated cytotoxicity (ADCC, upon binding to Fc\textgamma R receptor).\textsuperscript{187,195,196} In order for effector functions to occur, the presence of N-linked glycans (on Asn-297) is required, the structure of which can crucially influence binding.\textsuperscript{192} In addition, the Fc region confers a long \textit{in vivo} half-life to antibodies through binding to the neonatal Fc receptor (FcRn), providing a salvage mechanism for IgG.\textsuperscript{194,197} Even if this increased half-life leads to improved efficacy and localization to the target, it could also result in inappropriate activation of Fc receptor-expressing cells due to off-target binding, leading to massive cytokine release and associated toxicity.\textsuperscript{69} Moreover, the large size of IgG may significantly limit tissue penetration, with reports indicating that more than 99\% of the antibody being administered does not reach the tumour.\textsuperscript{198}
1.2.1 Antibody fragments

In contrast to full antibodies, antibody fragments (e.g. F(\(ab\))\(_2\), Fab, Fv, scFv, ds-scFv, Figure 7) have the advantage of improved tumour penetration, especially through solid tumours and homogeneous tissue distribution, owing to their smaller size.\(^{\text{187}}\) In addition, they can provide new binding specificities, particularly for targets that are inaccessible by full antibodies, such as cryptic antigens in enzyme active sites, G protein-coupled receptors or viral surface canyons.\(^{\text{199}}\) More importantly, due to the absence of the Fc region, fragments exhibit reduced toxicity and immunogenicity.\(^{\text{200}}\)

Originally, intact antibodies were dissected into constituent domains through proteolysis.\(^{\text{199}}\) Digestion of IgG with pepsin degrades the Fc region to small peptides to afford a F(\(ab\))\(_2\) fragment, which retains the bivalent binding affinity of IgG and confers high retention times on many cell-surface receptors and antigens. Alternatively, digestion with papain cleaves one or more peptide bonds in the hinge region to yield one Fc and two identical monovalent Fab fragments, each containing the \(V_H-C_H1\) and \(V_L-C_L1\) segments linked by a disulfide bond.

![Figure 7 – Structure of antibody fragments; green bars denote disulfide bonds.](image)

With the advent of recombinant techniques, it has become possible to produce a range of smaller fragments by direct expression from bacterial (\(E.\)\(coli\), \(Bacillus\) \(megaterium\))/ mammalian cells (CHO cells)\(^{\text{201}}\) or yeast (\(Pichia\) \(pastoris\)).\(^{\text{202}}\) Single-chain Fv (scFv) is the smallest fragment that retains full specificity and antigen binding activity.\(^{\text{200}}\) It is comprised of both \(V_H\) and \(V_L\) domains linked by a flexible polypeptide linker to prevent dissociation of the otherwise unstable Fv. The incorporation of a disulfide bond (ds-scFv – disulfide stabilized single-chain Fv) can further increase the stability of the construct.\(^{\text{203}}\)
While antibody fragments exhibit more favourable pharmacokinetics for tissue penetration, their molecular size below the renal threshold (ca. 50-60 kDa) leads to rapid clearance from circulation, limiting the percentage of injected dose being delivered at the tumour.\textsuperscript{204} This is beneficial for diagnostic applications, where high tumour to background ratios are required and three antibody fragment imaging agents are currently on the market (CEA-Scan, Verluma and Myoscint).\textsuperscript{205,206} However, for tumour therapy, such fragments are unlikely to give sustained effects due to high target retention required.\textsuperscript{204} In general, several strategies can be employed to tailor the half-life of such fragments, including PEGylation,\textsuperscript{207} affinity maturation,\textsuperscript{208} multimerization\textsuperscript{209} or fusion with albumin.\textsuperscript{210}

1.2.2 Multivalent antibody fragments

Systemic in vivo studies have confirmed that the pharmacokinetic (PK) properties and biodistribution of antibody molecules is greatly influenced by their molecular weight.\textsuperscript{211} It has been shown that the ideal tumour targeting reagent would exhibit rapid tumour penetration, high target retention and fast blood clearance to avoid off-target binding.\textsuperscript{204} Engineered scFv diabodies (52 kDa) and triabodies (78 kDa) have attracted considerable interest recently, since they seem to have the optimal size to afford a good balance between reduced clearance rates and increased tumour penetration compared to intact IgGs, as well as monomeric scFvs (Figure 8a).\textsuperscript{199,212} In addition, multivalent molecules show significant increase in functional binding affinity (avidity) to target antigens, which is primarily seen in reduced dissociation rates. This is a result of multiple binding to two or more antigens and the higher probability of rebinding as the molecule is retained near the surface when only one Fv dissociates.\textsuperscript{212} All of these reasons contribute to the increased tumour-to-blood ratios reported for these constructs, offering significant advantages for imaging, diagnosis and therapy (Figure 8b).\textsuperscript{213}
Multiple antigen binding has been shown to be dependent upon the correct alignment of Fv modules, as well as the spatial orientation of antigens on a surface.\textsuperscript{212,214} While in intact IgGs, the angle between the hinge and the Fab regions enables numerous cross-linking geometries, there are limits to this flexibility and IgGs are often unable to cross-link adjacent receptors.\textsuperscript{215,216} In contrast, engineered multivalent structures have been shown to be extremely flexible, with potential cross-linking angles from 60° to 180°.\textsuperscript{217}

The most successful strategy to genetically encode multimeric scFv constructs has been by shortening the peptide linker that connects the two variable domains (\textit{ca}. 12 residues in monomeric scFv).\textsuperscript{199} Molecular modelling and crystal structures have clearly shown that reducing the linker length to between 5 and 0 residues, the V\textsubscript{H} and V\textsubscript{L} domains come apart and re-assemble into non-covalently associated dimers or trimers.\textsuperscript{218,219} Alternative strategies to produce such multimers have included recombinant fusion,\textsuperscript{220} while the introduction of free cysteine residues by mutagenesis has enabled chemical cross-linking with bis-maleimide reagents. More specifically, scFv homodimers have been created in yields up to 44\% with bismaleimidohexane and CT52, a versatile reagent developed by Celltech Ltd. that contains two maleimide moieties and a DOTA chelator for radiolabeling (Figure 9).\textsuperscript{153,154}
In general, while the association of identical scFv molecules produces multimers of increased pharmacokinetic properties, the resultant conjugates are monospecific (i.e. bind only the same type of antigen). In recent years, extensive research has been conducted on the production of multispecific formats, capable of cross-linking different target antigens.\textsuperscript{140,212}

1.2.3 Antibody-Drug Conjugates

It has become apparent that even though clinical success has been witnessed with the use of unmodified mAbs, most antibodies do not have the cytotoxic potential to radically eliminate malignant cells.\textsuperscript{221} Antibodies are shown to be synergistically improved by arming them with cytotoxic drugs to create a new class of targeted therapy, i.e. antibody-drug conjugates (ADCs).\textsuperscript{18,222} This strategy has great potential to dramatically reduce the side-effects observed with standard chemotherapy, since the binding specificity of the antibody would facilitate selective drug accumulation at the site of disease, minimizing exposure to healthy tissues. In addition, this can enable the use of ultrapotent payloads, capable of killing cells at sub-nanomolar concentrations, which are too toxic to be used alone in conventional chemotherapy.\textsuperscript{198} These include tubulin inhibitors (e.g. auristatin and maytansine derivatives) or DNA intercalating agents, such as calicheamicins, duocarmycins and more recently the pyrrolobenzodiazepine dimer (PBD)
payload, which has exhibited sub-picomolar activity against a range of human tumour cell lines.\textsuperscript{222,223}

The third and important component of an ADC is the linker that connects the drug to the antibody, which must be stable in circulation to avoid drug release prior to antibody localization. At the same time, it needs to ensure active payload release following receptor-mediated endocytosis by the target cell.\textsuperscript{224} This can be achieved with both non-cleavable and cleavable linkers. An uncleavable linker could give active metabolites upon lysosomal antibody degradation.\textsuperscript{225} In contrast, cleavable linkers can enable payload release according to the unique microenvironment of various subcellular compartments. These include: 1) pH sensitive linkers,\textsuperscript{226} such as hydrazones that are cleavable in the low pH of endosomes (pH 5.0-6.5) or lysosomes (pH 4.5-5.0), 2) redox sensitive linkers,\textsuperscript{227} e.g. disulfide bonds that are reducible by the higher levels of glutathione in the cytosol compared to serum or 3) protease labile linkers,\textsuperscript{228} such as the widely-employed valine-citrulline linker, cleaved by cathepsin B or cathepsin S, as recently demonstrated.\textsuperscript{229} Interestingly, it has been shown that cathepsin B can also be secreted by tumour cells extracellularly to initiate proteolytic cascades and enable tumour proliferation.\textsuperscript{230} In addition, the progressive death of tumour cells releases increasing amounts of glutathione in the extracellular space.\textsuperscript{231} Neri and coworkers have demonstrated that non-internalizing ADCs, containing cleavable linkers, can release their payload in the extracellular space, which can then diffuse into surrounding tumour cells, as well as diseased neighbouring cells, through the so-called bystander effect.\textsuperscript{232}

In general, optimal ADC performance is shown to be a combination of factors, including drug potency, linker stability, drug release mechanism, conjugate composition, antigen-binding affinity, pharmacokinetics, internalization, as well as biodistribution.\textsuperscript{233} To date, out of more than 90 ADCs that have entered clinical trials, only 4 have been approved for clinical use (Figure 10); gemtuzumab ozogamicin (Mylotarg\textsuperscript{TM}), ado-trastuzumab emtansine (Kadcyla\textsuperscript{TM}), inotuzumab ozogamicin (Besponsa\textsuperscript{TM}) and brentuximab vedotin (Adcetris\textsuperscript{TM}).\textsuperscript{234} Generation of ADCs through lysine conjugation (as in Mylotarg, Kadcyla, Besponsa) has been heavily pursued due to the advantages of a biologically robust amide bond that attached the drug (i.e. NHS ester chemistry) and the limited antibody processing steps.\textsuperscript{224} However, this strategy is well documented to result in product heterogeneity. Wang \textit{et al.} have demonstrated that out of the 86 lysine residues present in an IgG1 antibody, drug molecules can distribute over 47\% of them, with $10^6$ distinct species being statistically possible when targeting drug to antibody ratios (DARs) of 2-4.\textsuperscript{235}
Figure 10 – ADCs currently in the market; a. Kadcyla, Mylotarg, Besponsa through lysine conjugation, b. Adcetris through cysteine conjugation; green bars denote disulfide bonds.

More importantly, a statistical mixture is obtained (0-8 drugs per antibody, i.e. following the Poisson distribution) resulting in a narrow therapeutic window. Unconjugated antibodies inhibit ADC activity, while higher-loaded species (i.e. >5) are prone to aggregation, fast clearance and are more likely to induce toxicity. The development of Adcetris, through cysteine conjugation was a significant improvement in terms of heterogeneity, with dozens of species being obtained when targeting a DAR of ca. 4. This was achieved by partial reduction (i.e. with 2 eq. TCEP) of the four interchain disulfide bonds and subsequent conjugation to afford products with DARs of 0, 2, 4, 6 and 8 as the major components. Overall, the development of next-generation ADCs has been focused on exploring technologies to producing homogeneous constructs, due to a more defined and improved pharmacokinetic profile.

1.2.4 Site-selective conjugation of drugs to antibodies

1.2.4.1 Engineered antibody modification

In order to reduce the heterogeneity associated with conventional ADC preparation techniques, novel conjugation methods have emerged. Most of these require recombinant antibody engineering in order to introduce bioorthogonal handles for site-specific conjugation (Scheme 16). In 2008, Junutula and coworkers reported on the generation of antibodies containing two engineered cysteine residues (referred to as THIOMABs). It was observed that the location of the
conjugation site can dramatically influence the ADC performance by modulating the stability of the antibody-linker interface.\textsuperscript{97} By using this strategy, ADCs were generated with an average DAR of 1.9 and more than 90\% homogeneity. Despite having half the average drug load, THIOMAB-drug conjugates were as efficacious \textit{in vivo}. In particular, they were shown to be ideal for conjugation with the highly potent PBD payload, which due to its poor water solubility results in aggregate formation at higher loadings.\textsuperscript{239} Remarkably, the THIOMAB platform has also been used to create a novel antibody-antibiotic conjugate (AAC) against intracellular \textit{Staphylococcus aureus}.\textsuperscript{240}

Unnatural amino acids (UAAs) have also been employed to provide site-specific conjugation sites and install functionality \textit{via} click chemistry. Schultz \textit{et al.} have recently reported the incorporation of $p$-acetylphenylalanine (pAcF) into an anti-CXCR4 antibody for the construction of ADCs with a DAR of \textit{ca.} 2, through oxime ligation.\textsuperscript{241} In a similar fashion, the incorporation of $p$-azidomethylphenylalanine (pAMF) into trastuzumab by Sato \textit{et al.} facilitated the formation of ADCs \textit{via} SPAAC.\textsuperscript{242}

\textbf{Scheme 16} – General scheme highlighting typical methods for engineered antibody modification; introduction of cysteine residues (THIOMABs) for alkylation; unnatural amino acid incorporation for click chemistry; formylglycine-generating enzyme to generate an aldehyde for Pictet-Spengler ligation; introduction of a glutamine tag for transglutaminase functionalization (adapted from Chudasama \textit{et al.}).\textsuperscript{243}
Alternative methods include the insertion of small peptide sequences into the antibody to form a “tag” that is selectively and irreversibly modified by enzymes. For example, Rabuka et al. have developed the SMARTTag™ technology. This involves a formylglycine-generating enzyme (FGE) that catalyzes the conversion of cysteine to aldehyde, when introduced in a CXPXR sequence. The resultant aldehyde is conjugated to a maytansine payload via hydrazino-iso-Pictet-Spengler (HIPS) ligation with 90% conjugation efficiency. In addition, a glutamine tag (LLQG) can serve as an acyl donor for enzymatic ligation to primary amines, catalyzed by a transglutaminase enzyme (TG). A transpeptidase enzyme (Sortase A) has also been used to catalyze transfer of polyglycine substrates to the C-terminus of the sequence motif LPETG, resulting in a stable amide linkage.

In general, although engineered ADCs have shown to outperform heterogeneous conjugates, they are not easily transferrable and might require unconventional expression systems, which are not yet fully clinically validated. Overall, they add an extra layer of complexity to the ADC production, which can increase the manufacturing cost.

1.2.4.2 Native antibody modification

The four conserved interchain disulfide bonds present in IgG1 represent an ideal target for site-specific modification. In addition, the fact that they are distal from the antigen binding or Fc regions does not compromise the biological activity of the antibody. Research at Seattle Genetics has led to the development of Adcetris, an anti-CD30 antibody conjugated to monomethyl auristatin E (MMAE) via a maleimide linker, indicated for hematological malignancies (vide supra, Figure 10, page 32). Initial studies had shown that even though reacting all 8 sulfhydryl groups can afford homogeneity (DAR 8), the resultant ADC was inferior in vivo, due to high plasma clearance rates and propensity to aggregate. The therapeutic index was shown to significantly increase by reducing the drug load from eight to four, which was achieved by partial reduction and subsequent conjugation of the four interchain disulfide bonds to afford an average DAR of 4.

In a follow up report, Lyon and coworkers suggested that the accelerated clearance of highly loaded ADCs is a result of increased hydrophobicity, which can be modulated through drug-linker design. By incorporating hydrophilic linkers and payloads, conjugates were generated with 8
drugs per antibody that showed decreased clearance and improved efficacy over conventional ADCs. However, this is not conclusive since all ADCs compared in that study contained different linkers and payloads, which could also have affected the ADC performance.

In general, apart from the structural integrity of the antibody being compromised, alkylation of cysteine residues with maleimides has been deemed undesired due to the resultant thiosuccinimide adducts undergoing retro-Michael reactions to release the payload. Notably, researchers at Seattle Genetics and Medimmune have designed maleimide reagents to contain electron withdrawing N-substituents that accelerate thiosuccinimide hydrolysis to under an hour, with the view to minimizing deconjugation. These include a basic amine on the N-alkyl substituent near the point of attachment to the thiosuccinimide or N-aryl substituents, respectively.

The functional re-bridging of disulfide bonds has the potential to yield homogeneous ADCs with controlled DARs, i.e. 4 drugs/ antibody, while maintaining the stabilizing effect of the disulfide bridge. Godwin and coworkers employed bis-cysteine selective sulfone reagents attached to MMAE, through a cleavable PEG spacer, to rebridge the disulfide bonds of trastuzumab. This strategy successfully afforded ADCs that were 78% DAR 4 and exhibited greater stability compared to conventional maleimide ADCs. In addition, superior potency compared to unconjugated trastuzumab was demonstrated in vitro. In vivo studies were also performed, where the disulfide-bridged ADC exhibited improved potency compared to Kadcyla. However the relative impact of this strategy on the overall therapeutic index could not be determined, since neither native antibody nor analogous heterogeneous controls were included in the study.

Next-generation maleimides have demonstrated exceptional efficiency in bridging disulfide bonds and have been used in a variety of protein conjugation applications. Their greatest potential has been realised in the formation of homogeneous ADCs. Using an in situ protocol, a dithiophenolmaleimide linked to the cytotoxic drug MMAE was employed for the immediate rebridging of disulfides in trastuzumab (Herceptin™) IgG1, upon TCEP reduction (Scheme 17).
Scheme 17 – Construction of an ADC using next generation maleimides via an in situ approach.

In order to appraise the blood serum stability of next generation maleimide ADCs, the construct was incubated in blood serum. Thiol exchange with albumin was shown to occur within four days, in a similar rate to that observed with classical maleimide conjugates, though via a different mechanism. However, by hydrolysing the maleimide moieties, robust serum stability was observed. Most importantly, hydrolysis to the maleamic acid was quantitative, without any loss of payload being observed. The resultant ADC had an average DAR of 3.89 and was shown to selectively target and kill HER2+ cells in vitro, in comparison to native Herceptin, where minimal reduction in cell survival was observed. Remarkably, a study was undertaken by Igenica Biotherapeutics, where ADCs prepared using the NGM platform were directly compared to analogous heterogeneous ADCs. All constructs contained identical linkers and payloads conjugated to identical sites, with the view to determining the relative contributions of homogeneity and disulfide bridging in ADC production. The results provided convincing evidence that homogeneous ADCs demonstrate superior efficacy and improved pharmacokinetics.

In general, while ADCs have shown considerable promise as cancer therapeutics, being the main interest of immunoconjugate research, the field would greatly benefit from a variety of advances in increasing drug load, improving conjugation methods, linker sensitivity and drug release. The conjugation of antibodies or antibody fragments with nanoparticles that encapsulate drugs might offer a variety of advantages to traditional ADCs. Of particular importance is the drug encapsulation stability in these carriers, as well as the opportunity to vastly increase the drug-antibody ratio. Recently, the conjugation of an anti-HER2 mAb to 30 nm-diameter iron oxide
magnetic nanoparticles was shown to successfully capture cancer cells with 1:10,000,000 enrichment of cancer cells over normal cells.\textsuperscript{255} However, the synthetic methods already reported are non-ideal, since they rely mainly on lysine modification with NHS ester derivatives or reduction of the interchain disulfide bonds and reaction with maleimide reagents.\textsuperscript{256,257} Chudasama and coworkers have recently reported on the site-specific functionalisation of nanoparticles through disulfide bridging, employing dibromopyridazinediones.\textsuperscript{258} Further to this, small molecule-drug conjugates (SMDCs) have been developed, representing an attractive alternative to ADC products. For instance, Neri and coworkers have developed high-affinity acetazolamide ligands, which bind to carbonic anhydrase IX (CAIX) on the surface of kidney cancer cells to selectively deliver payloads, sparing normal organs.\textsuperscript{259} Interestingly, on a direct comparative evaluation of an ADC against a SMDC, the latter showed a substantially higher tumour uptake (\textit{ca.} 40\% ID/ g) compared to the equivalent ADC (\textit{ca.} 5\% ID/ g).\textsuperscript{260}

\subsection*{1.2.5 Bispecific antibodies}

The restricted therapeutic ability of conventional monoclonal antibodies, as well as the heterogeneity of many ADCs have also led to the development of bispecific antibodies.\textsuperscript{142,190} More specifically, the association of two antibodies or antibody fragments into a single molecular entity can enable simultaneous binding to two different epitopes either on the same or different antigens. Over the past years, various bispecific formats have been developed for diagnosis, imaging and therapy. These include scFv-derived formats, such as diabodies,\textsuperscript{261} tandem diabodies,\textsuperscript{262} BiTEs (Bispecific T-cell Engagers)\textsuperscript{263} and DARTs (Dual Affinity Re-Targeting),\textsuperscript{264} as well as IgG-based formats, such as Triomab,\textsuperscript{265} DVD (Dual Variable Domain)\textsuperscript{266} and two-in-one antibodies.\textsuperscript{267} For most therapeutic applications, one antibody moiety serves as a targeting agent (i.e. to a tumour-associated antigen), while the second antibody is employed to recruit and activate a suitable leukocyte.\textsuperscript{268} The most sophisticated format is the Bispecific T-cell Engager (BiTE), which brings cancer cells in close proximity with T-cells, leading to a highly efficacious killing process. To date, only two BiTE constructs have been approved for clinical use; catumaxomab (Removab\textsuperscript{TM}) for the treatment of malignant ascites and blinatumomab (Blincyto\textsuperscript{TM}) for acute lymphoblastic leukemia.\textsuperscript{269,270} Catumaxomab is a hybrid between a mouse IgG2a and a rat IgG2a antibody formats that targets CD3 and the antigen epithelial cell adhesion molecule (EpCAM). It is
produced from a mouse/rat quadroma cell, resulting in 12.5% of the total products possessing the desired dual specificity.\textsuperscript{271} In contrast, blinatumomab is made by genetic fusion of two scFv fragments targeting CD3 and CD19, resulting in \textit{ca.} 55 kDa complex that actually suffers from short serum half-life.\textsuperscript{272} An alternative recombinant technique to produce bispecific antibodies involves introducing mutations into the CH3 domain of two different heavy chains against two respective antigens, which favour heterodimerization when co-expressed (“knob-into-hole” approach).\textsuperscript{268}

The construction of bispecifics through chemical modification techniques has been vastly explored through lysine/cysteine modification, employing bifunctional reagents (e.g. SMCC, bis-maleimides etc.).\textsuperscript{139,273} However, this results in sub-optimal cross-linking (\textit{vide supra}). Recently, site-specific conjugation approaches have started being explored in order to generate chemically defined homogeneous bispecific antibodies. This has been accomplished by either recombinant antibody engineering to introduce bioorthogonal handles or through native antibody modification, by disulfide bridging.

Schultz and coworkers achieved the formation of various anti-HER2/anti-CD3 BiTE formats (i.e. Tetra-IgG, Tri-IgG, TriFab, BiFab), through unnatural amino acid modification (Figure 11a).\textsuperscript{274} \textit{p}-acetylphenylalanine was introduced into each antibody, which was reacted \textit{via} oxime ligation with an aminooxy linker bearing an azide or a strained alkyne, respectively. Subsequent SPAAC afforded bispecific antibodies in 25-75\% yields and in timescales of 1-3 days. All bispecific antibodies demonstrated excellent \textit{in vivo} cytotoxicity. Interestingly, antigen-independent T-cell activation was observed with constructs containing an Fc domain (by cross-linking T-cells with FcRn-positive immune cells), which was reduced by mutating two residues in the Fc region to minimize FcRn-binding. Rabuka and Bertozzi \textit{et al.} have also employed the SMARTTag\textsuperscript{TM} technology to cross-link proteins (Figure 11b).\textsuperscript{275} Aldehydes were introduced to anti-HER2 IgG, which were functionalized with a strained alkyne through oxime ligation. The other partner, growth hormone (h-GH), was labeled with an azide, in a similar way. The target heterobifunctional protein was generated \textit{via} SPAAC, with 70\% overall conversion being observed, according to densitometry analysis.
The functional re-bridging of disulfide bonds has also been exploited for the construction of bispecifics, with two examples reported in the literature. Brochini and coworkers have synthesized an IgG mimic by employing a di(bis-sulfone) reagent. More specifically, in situ elimination of p-toluene sulfinic acid occurs to produce the reactive enone-sulfonyl moieties that react with two molecules of bevacizumab Fab, within 3 hours, to produce a homodimer in 12% yield (Scheme 18a). Baker and coworkers have shown that a bispecific construct can be generated by bridging native disulfide bonds with a bis-dibromomaleimide cross-linker (Scheme 18b). This approach enabled the fast production (i.e. within 1 h) of a homogeneous scFv-Fab conjugate, in 52% isolated yield.
Scheme 18 – Generation of bispecific antibodies by disulfide bridging; a. Bis-sulfones, b. Next-generation maleimides (NGMs).
1.3 Human Serum Albumin

Human serum albumin (HSA) is the most abundant protein in the vascular system with an average plasma concentration of 40 mg/mL.\textsuperscript{278} The long half-life of serum albumin ($t_{1/2} \sim 19$ days) has been a matter of debate for decades.\textsuperscript{270–281} Its large size (ca. 66.5 kDa) that is above the renal filtration threshold is definitely contributing to it. However, a molecular explanation for this extraordinarily long half-life came to light when Anderson and coworkers reported that albumin binds to the FcRn receptor, a feature shared with the IgG class of antibodies.\textsuperscript{282} Rescue occurs via a recycling pathway, where FcRn, predominantly located within acidified endosomes, binds its ligands taken up by pinocytosis. The binding is strictly pH dependent, due to the presence of conserved histidine residues that interact with acidic residues within the FcRn heavy chain. Ligand binding at acidic pH results in transport back to the cell surface, where exposure to the neutral pH of the blood triggers release of albumin back to the circulation.\textsuperscript{191,210,282}

The crystal structure of HSA, published in 1992 by Carter and coworkers revealed that the 585 amino acid molecule is heart-shaped, consisting of three repeating domains (Figure 12).\textsuperscript{283} Remarkably, 70-80\% of the circulating albumin possesses a free sulfhydryl group, i.e. Cys-34 that has a rather low pK\textsubscript{a} of 7.0, compared to 8.5 for cysteine.\textsuperscript{284,285} This is due to the thiolate ion being stabilised by a hydrogen bond formed with the neighbouring tyrosine-84.\textsuperscript{286} The conservation of the reduced form of Cys-34 that is attributed to its positioning within a 9.5 Å deep hydrophobic crevice on the surface of the protein raises the question as to the nature of its biological role.\textsuperscript{287} Cys-34 acts as the binding site for drugs (e.g. captopril, auranofin), metal ions (Cd\textsuperscript{2+}, Au\textsuperscript{2+}, Hg\textsuperscript{2+}, Ag\textsuperscript{2+}) and nitric oxide or reactive oxygen species. It can be considered, therefore, as a scavenger for thiols, heavy metals and oxidants.\textsuperscript{284,288}

![Figure 12](image_url) – Structure of human serum albumin highlighting key features (PDB file 1AO6).
1.3.1 Conjugation to Human Serum Albumin

In the middle of the 20th century the first reports appeared in the literature demonstrating that tumours are able to trap plasma proteins and use their degradation products for proliferation.\textsuperscript{289} This phenomenon has been termed “enhanced permeability and retention” in relation to passive tumour targeting – EPR effect, ideally propagating the therapeutic concept of drug targeting that was founded on Paul Ehrlich’s vision of the “magic bullet”.\textsuperscript{290,291} From the publications on albumin-based drug delivery systems that have appeared in the recent years, it is interesting to observe that various research groups have employed albumin as a drug carrier to achieve tumour accumulation. Examples of drugs include doxorubicin (DOXO-EMCH),\textsuperscript{284} methotrexate,\textsuperscript{292} MMAE\textsuperscript{293} and more. As expected, conjugation through lysine residues has been shown to yield heterogeneous mixtures, due to the multitude of lysine residues on the surface of HSA. Apart from the heterogeneity of the final conjugate, high loading has been reported to result in albumin being physiologically unstable and hence not attain half-lives similar to the native protein.\textsuperscript{294} Interestingly, Barbas and coworkers had reported the site-specific attachment of TAK-242 (a potent Toll-like receptor 4 inhibitor) to Lys-64 of HSA.\textsuperscript{295} In general, conjugation via Cys-34 has proved to be the most popular and selective strategy for attachment and importantly, it is located at a position remote from domain III (C-terminus), which is responsible for binding to the FcRn receptor, required for the albumin salvage pathway.\textsuperscript{281,296} The use of albumin as a drug carrier to enable passive tumor targeting has also been realised in the formation of albumin-drug nanoparticles, which can encapsulate and solubilize poorly water-soluble drugs.\textsuperscript{297} Abraxane\textsuperscript{®}, an albumin-paclitaxel nanoparticle was the first protein-nanoparticle to be approved by FDA for the treatment of metastatic breast cancer. This complex accumulates in the tumour, through the EPR effect.\textsuperscript{298} A further albumin transport pathway mediated by the 60 kDa gp60 glycoprotein, located on the endothelial cell surface seems to be responsible for tumour uptake and subsequent release of the drug.\textsuperscript{290} As a result of its exceptionally long circulatory half-life, HSA has also been used as a medium for the half-life extension of small therapeutic proteins/ peptides. In particular, insulin,\textsuperscript{299} ANP (Atrial Natriuretic Peptide),\textsuperscript{300} GLP-1 (Glucagon-Like Peptide 1),\textsuperscript{301} Exendin-4\textsuperscript{302} and C34 peptide (PC-1505)\textsuperscript{303} have successfully been conjugated to Cys-34 of HSA. Remarkably, the circulatory half-life of Exendin-4 was increased from 2-3 h in humans to over a week, representing almost a 200-fold improvement.\textsuperscript{279} This strategy has been used to extend the half-life of various other
proteinaceous drugs, including DARPin domain[^304], Kringle domain[^305], granulocyte colony stimulating factor (G-CSF)[^306], the opioid agonist dynorphin A (CJC-1008)[^307], and YY peptide[^308].

In recent years, there has been an increasing interest in the use of antibody fragments for targeted therapeutics. Even though such fragments exhibit quick biodistribution and good tumour penetration which is attractive for therapy, their short half-life results in rapid elimination from the body. PEGylation has been shown to increase their serum half-life, though there are a few reported cases where this led to reduced binding and activity, as well as immunogenicity with anti-PEG IgM antibodies being induced after the first dose.[^279][^309][^310] Genetic fusion to albumin is an appealing alternative that has been explored to successfully increase the serum half-life of scFv[^311], Fab[^312] fragments, as well as bispecific formats including BiTEs and diabodies.[^310][^313][^314] Preclinical evaluations of the pharmacokinetic properties of various fusions have been extensively performed and an improvement compared to non-fused counterparts has been shown. However, in a few examples, even if the half-life was increased, it was still far from that of the endogenous albumin of the animals used. A study undertaken by Novozymes Biopharma established that fusion to the N-terminal end of HSA can sometimes result in a minor reduction in binding to the FcRn receptor.[^210] In contrast, fusion to the C-terminus can significantly affect the binding affinity, with up to a 2-fold reduction. This might have a more pronounced effect in vivo, since fusions have to compete for receptor binding in the presence of large amounts of endogenous albumin.[^210] In addition, fusion to albumin for half-life extension might not be suitable for all proteins, since many require unmodified termini for optimal activity.[^142][^143]

Direct chemical conjugation to Cys-34 is proved to be effective for small proteins/ drugs. However, there is an obvious challenge in the conjugation of larger proteins/ antibody fragments to HSA. This could be attributed to Cys-34 being buried in a hydrophobic crevice that is partially protected from the solvent, thus posing steric constraints to the conjugation.[^127][^315] According to the literature, the best example of chemical conjugation of an antibody fragment (i.e. Fab) to Cys-34 of albumin was reported by Smith and coworkers with yields in the range 5-12%.[^316] This was accomplished by conjugating Cys-34 of albumin to Fab fragment containing a single cysteine mutant on the heavy chain, through a bis-maleimide cross-linking reagent with a hexane spacer.
In 2013, Weil and coworkers reported on the design of an orthogonal heterobifunctional cross-linking reagent that combines a bis-sulfone and a maleimide moiety to enable the step-wise conjugation of two biomolecules (Scheme 19).\textsuperscript{317} To test the virtue of this approach, they selected HSA and BSA (Bovine Serum Albumin) as model proteins that contain a single cysteine group. BSA was initially reacted with reagent 10, through the maleimide end at pH 6.0, to afford intermediate 11. Under these conditions, the bis-sulfone functionality remained intact. The pH was then raised to 8.0, where \textit{in situ} elimination of \textit{p}-toluene sulfinic acid occurred to generate enone-sulfone 12 that reacted with HSA. The resultant heterodimer 13 was isolated in 10% yield. Even though this approach is elegant and well-designed, it still suffers from low conjugation yield. It seems that the sterically hindered nature of Cys-34 of albumin may necessitate bifunctional reagents with both ends being of high reactivity (i.e. bissulfone is less reactive than maleimide) to enable efficient addition of the second biomolecule. Overall, methods for the high-yielding and site-specific conjugation of large proteins to HSA, \textit{via} Cys-34, are highly sought.

\begin{center}
\textbf{Scheme 19} – Stepwise orthogonal coupling of BSA and HSA, employing a bis-sulfone/ maleimide cross-conjugation reagent.
\end{center}
1.4 Aims and Objectives

Prior research has demonstrated that a bis-dibromomaleimide PEG cross-linking reagent can be used to create a bispecific scFv-Fab antibody in a site-specific and homogeneous manner. The principal aim of the current project is to extend this technology and determine the scope of these NGM cross-linkers in conjugating proteins containing different reactive handles (i.e. a disulfide to a single cysteine), as well as sterically hindered conjugation sites. In addition, to optimize the methodology for the generation of protein-protein conjugates which are robustly stable in serum and are thus suitable for in vivo applications.

To address these challenges, Human Serum Albumin (HSA) was selected as the ideal cysteine containing protein, in which the single free thiol (Cys-34) available for conjugation is positioned in a sterically hindered environment. In general, there has been considerable interest in the generation of albumin-antibody fragment constructs for therapeutic or imaging applications, with the products exhibiting significantly increased half-lives and tumour-to-blood ratios (vide supra). To this end, it was decided to synthesize NGM cross-linking reagents and evaluate their application in the synthesis of antibody fragment-albumin conjugates as a prospective half-life extension platform (Scheme 20).

**Scheme 20** – Proposed strategy for the construction of an antibody fragment – albumin protein conjugate through disulfide-to-cysteine modification.
Once the appropriate linker design has been developed to promote conjugation in a sterically hindered system, the linker technology will be extended to incorporate a third point of attachment for another biomolecule, giving rise to a novel class of triconjugates as potential therapeutics. Building on the HSA-scFv conjugation, the application of trifunctional NGM reagents in linking HSA to two scFv monomers will be initially examined with the view to obtaining a construct of increased half-life and functional binding affinity (Scheme 21). In addition, tri-scFv constructs have been identified as having an optimal size to afford a good balance between reduced clearance rates and increased tumour penetration. As a result, the application of the NGM platform towards the synthesis of tri-scFv conjugates will also be evaluated.

**Scheme 21** – Proposed strategy for the construction of a tri-protein conjugate.
Furthermore, a secondary aim of the present work is to develop a new bioconjugation methodology for the selective modification of lysine residues, in order to overcome the highly heterogeneous antibody conjugates obtained by existing techniques. This approach is expected to complement disulfide bridging as a strategy for selective modification and improved homogeneity, but also be advantageous in systems where disulfide bridging could result in detrimental changes of the antibody structure.

In general, site-selective lysine modification is a challenging task, due to the presence of multiple lysines on the surface of proteins. Inspired by native chemical ligation (NCL) occurring via macrocyclic transition states, it was envisaged that S-to-N acyl transfer from the disulfide bond of an antibody to proximal lysine residues could introduce site-selectivity. The proposed strategy, termed ‘cysteine-to-lysine transfer’ (CLT) is shown in Scheme 22 and involves an initial transthioesterification of the cysteine residues obtained upon reduction of the disulfide bond, followed by acyl transfer to proximal lysines. Crucially, this approach is expected to leave the disulfide bond unmodified, representing a traceless conjugation.

![Scheme 22 – Proposed strategy for cysteine-to-lysine transfer (CLT).](image-url)
2. Results and Discussion

2.1 Albumin-antibody fragment conjugates

The application of NGM bifunctional linkers towards the synthesis of albumin-antibody fragment conjugates was evaluated by employing two clinically relevant fragments that contain a single disulfide bond; trastuzumab (Herceptin) Fab, which targets the HER2/neu receptor overexpressed in 25-30% of breast cancers\textsuperscript{318} and anti-CEA scFv that targets the carcinoembryonic antigen (CEA), present in a wide range of cancers (particularly colorectal carcinoma).\textsuperscript{319} These two fragments were chosen because they contain disulfide bonds of distinct reactivity, which would enable to further test the scope of the methodology. More specifically, the disulfide bond in anti-HER2 Fab is solvent-accessible, while in anti-CEA scFv is more hindered. In order to identify the most robust cross-linking reagent, able to conjugate two sterically hindered proteins, the scFv fragment was used in initial optimization.

The choice of a homobifunctional reagent for the conjugation of different reactive centers (disulfide to cysteine) was made to preclude issues of regioselectivity. NGM handles were selected to be highly reactive in order to facilitate the addition of the second biomolecule in the sterically hindered system. This is due to literature precedent suggesting that a reagent designed to contain a functionality of lower reactivity for conjugation with the second protein, leads to lower efficiency.\textsuperscript{152,320}
2.1.1 Synthesis and application of bifunctional conjugation reagents

2.1.1.1 Bis-dibromomaleimide PEG linker

Initial investigations towards the scFv-HSA conjugation begun by employing the bis-dibromomaleimide PEG cross-linker DBM\textsubscript{CL}, since it was previously shown to be efficient in generating an scFv-Fab bispecific. The linker was synthesized in two steps, involving initial functionalization of the commercially available dibromomaleimide 14 with an N-methoxycarbonyl to give 15, which was then reacted with PEG(19) diamine (Scheme 23).

Scheme 23 - Synthesis of bis-dibromomaleimide PEG linker DBM\textsubscript{CL}; Reagents and conditions: (i) Me\textsubscript{3}Cl, N\textsubscript{3}Me morpholine, THF, RT, 20 min, (ii) PEG(19) diamine, CH\textsubscript{2}Cl\textsubscript{2}, silica, RT, 24 h.

Before testing the scFv-HSA conjugation, the generation of an scFv homodimer was trialed, using a previously developed protocol (Figure 13a).\textsuperscript{277} Addition of 0.42 equivalents of linker DBM\textsubscript{CL} to reduced scFv proceeded successfully to afford homodimer 16, observed as a distinct band at ca. 50 kDa by SDS-PAGE analysis (Figure 13b, lane 3).

Figure 13 - Generation of scFv homodimer using DBM\textsubscript{CL} a. Reagents and conditions: (i) DTT (20 eq.), pH 7.4, RT, 1 h, (ii) DBM\textsubscript{CL} (0.42 eq.), pH 7.4, RT, 1 h, b. SDS-PAGE analysis: M. molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Reduced scFv + linker conjugation mixture, after 1 h incubation.

An LC-MS spectrum of the commercial albumin (extracted from human serum, Sigma Aldrich) was obtained to reveal that it contains a mixture of mercaptalbumin (i.e. reduced thiol) and fractions consisting mainly of mixed disulfides between Cys-34 and low molecular weight thiols, e.g. cysteine. Reduction with 5 equivalents of DTT for 1 hour afforded native HSA that contained one free sulfhydryl group, as evidenced by LC-MS and Ellman’s assay (before reduction the
Sulfhydryl per Protein ratio, SPR was 0.18). While this albumin was used in the initial experiments with prior reduction being carried out, it was later replaced with recombinant HSA (expressed in *Saccharomyces cerevisiae*, Sigma Aldrich), which was shown to consist only of mercaptalbumin.

An optimization study was then carried out to determine the appropriate conditions for the reaction of HSA with DBM$_{\text{CL}}$ (Scheme 24). Completion was achieved within 15 minutes at RT, by using 5 equivalents of the linker, with no dimerization of HSA observed under these conditions. Interestingly, LC-MS analysis revealed that loss of a second bromine from the cross-linking reagent had occurred (Figure 14).

![Scheme 24 - Reaction of HSA with DBM$_{\text{CL}}$; Reagents and conditions: (i) DBM$_{\text{CL}}$ (5 eq.), pH 7.4, RT, 15 min; X refers to loss of the 2$^{\text{nd}}$ bromine atom, due to intramolecular cross-linkage.](image)

**Figure 14** – LC-MS spectrum of HSA-DBM$_{\text{CL}}$ conjugate 17: observed mass of 67,682 (calculated 68,677 for -2 Br).

In order to gain a better understanding of this unexpected debromination, HSA was reacted with $N$-methyl monobromomaleimide 18 and dibromomaleimide 19 (Figure15b). The LC-MS spectra of both adducts showed a nearly identical mass, suggesting that the loss of the second bromine was occurring from the same maleimide motif that Cys-34 of HSA had reacted with. The possibility of a reductive dehalogenation to a mono-thiomaleimide was initially considered, but it
was ruled out upon comparing the UV/Vis absorbance spectra of the 2 conjugates (Figure 15a). Conjugate 20 showed maximum absorbance at 357 nm, which is in accordance with that of a mono-thiomaleimide. In contrast, conjugate 21 exhibited a 19-nm bathochromic shift, absorbing at 376 nm, suggesting the presence of a different species.

**Figure 15** – Reaction of HSA with \(N\)-Me monobromo- and dibromomaleimide to afford constructs 20 and 21 respectively; a. UV/Vis spectra of resultant conjugates; b. Reaction scheme; Reagents and conditions: (i) \(N\)-Me monobromo- or dibromomaleimide (1.5 eq.), pH 7.4, RT, 20 min.

To further investigate this, the HSA conjugate 21 formed upon reaction of HSA with \(N\)-methyl dibromomaleimide was reacted with 1 equivalent \(\beta\)-mercaptoethanol (BME) and the absorbance was monitored (Figure 16). A new species 22 appeared with maximum absorbance at 392 nm, which is characteristic wavelength of a dithiomaleimide. When the construct was reacted with 100 equivalents of BME, quantitative reversion to native HSA was observed. These preliminary results suggest that displacement of the bromine by a nucleophilic residue is taking place, which is then acting as a good leaving group upon reaction with BME. It was anticipated that this debromination will not interfere with the desired conjugate formation and the point of cross-linkage was later discovered (see section 2.1.3).
Using the established protocol for the reduction of scFv and bridging with DBMCL\textsuperscript{,277} the scFv-HSA conjugation was then attempted (Scheme 25). The scFv was successfully bridged with linker DBM\textsubscript{CL} to afford intermediate 23, as shown by SDS-PAGE analysis (Figure 17). However, upon incubation of the intermediate with stoichiometric, as well as excess HSA (1-3 eq.), no band corresponding to the desired conjugate 24 was observed after 24 hours.

**Scheme 25** – Attempted conjugation of scFv to HSA using linker DBM\textsubscript{CL}; Reagents and conditions: (i) DTT (20 eq.), pH 7.4, RT, 1 h, (ii) Linker DBM\textsubscript{CL} (30 eq.), pH 7.4, RT, 15 min, (iii) HSA (1-5 eq.), pH 7.4, RT, 24 h.
**Figure 17** – SDS-PAGE analysis of attempted conjugation of scFv to HSA, using linker DBM<sub>CL</sub>: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Reduced scFv + linker, 4. Native HSA, 5. scFv-HSA conjugation mixture (1:1), 6. scFv-HSA conjugation mixture (1:2), 7. scFv-HSA conjugation mixture (1:3), after 24 h incubation.

The scFv-HSA conjugation was also repeated in the reverse order (Figure 18a). In particular, reduced scFv was added to the HSA-linker intermediate 17, where again no hint of product formation was witnessed after 24 hours, by SDS-PAGE analysis (Figure 18b).

**Figure 18** – Addition of scFv to HSA-linker intermediate 17; a. Reaction scheme; Reagents and conditions: (i) scFv (0.3-2.0 eq.), pH 7.4, RT, 24 h (prior scFv reduction with DTT (20 eq.), pH 7.4, RT, 1 h, b. SDS-PAGE analysis: M. Molecular marker, 1. Native HSA, 2. HSA + linker, 3. Native scFv, 4. Reduced scFv, 5. HSA-scFv conjugation mixture (3:1), 6. HSA-scFv conjugation mixture (2:1), 7. HSA-scFv conjugation mixture (1:1), 8. HSA-scFv conjugation mixture (1:2), after 24 h incubation.
To examine whether the unsuccessful conjugation was due to steric issues, the addition of a small disulfide-containing peptide (i.e. octreotide, ca. 1 kDa) to intermediate 17 was also attempted (Figure 19a). Adding 1 equivalent of octreotide afforded a mixture of unreacted intermediate 17 and native albumin, as observed by LC-MS (Figure 19b). This result was very interesting; the fact that native albumin is obtained suggests that octreotide has added on the maleimide that has reacted with the sterically hindered albumin. Therefore, the dibromomaleimide end of intermediate 17 is not available for reaction. It was hypothesized that hydrolysis of the second dibromomaleimide motif could be responsible for the unsuccessful conjugation.

Figure 19 - Addition of octreotide to HSA-linker intermediate 17; a. Reaction scheme; Reagents and conditions: (i) Octreotide (0.5-2.0 eq.), pH 7.4, RT, 1 h (prior reduction of octreotide with TCEP (1.5 eq.), 37 °C, 1 h), b. LC-MS spectrum after reaction of HSA-linker with 1 eq. octreotide: HSA observed mass of 66,462 (calculated 66,465), intermediate 17 observed mass of 67,686 (calculated 67,677).
2.1.1.2 NGM hydrolysis study

The susceptibility of maleimides towards hydrolysis has been well reported in the literature; this process is known to be temperature and pH dependent, yielding a maleamic acid that is unreactive towards nucleophiles.\textsuperscript{61} In addition, many studies demonstrate the effect that the $N$-substitution can have on the rate of hydrolysis, i.e. aliphatic substituents slow down hydrolysis, in contrast to electron-withdrawing-groups, which accelerate it.\textsuperscript{150,321}

In order to gain a better understanding of the NGM platform and hence develop robust cross-linking reagents, it was deemed necessary to study the effect that the maleimide leaving groups have on the rate of hydrolysis. The reaction of NGMs with proteins is very fast. However, when a second maleimide is present in the conjugation (i.e. as in the case of the cross-linking reagents used in the current project) it is important to ensure that it remains functional until conjugated to the second protein.

To investigate the rate of NGM hydrolysis, a small molecule UV study was pursued. The hypothesis was that the maleimide chromophore exhibits strong absorption in the UV/Vis region, which is lost upon conversion to the corresponding maleamic acid. The $N$-methyl derivatives of dichloro- 26, dibromo- 19, diiodo- 27, diphenoxy- 28, dithiophenol- 29 and unsubstituted maleimide 30 were chosen as model substrates. Their UV/Vis spectra were firstly obtained to determine their absorption maxima. The compounds were then intentionally hydrolysed (i.e. upon incubation at pH 8.6 for 2 days at 37 °C) to prove that the absorption at the corresponding maximum wavelengths is negligible. They were then incubated in phosphate buffer pH 7.4 (with 10% DMF) at RT and their absorbance over time was monitored (Figure 20).

The kinetics of hydrolysis were calculated using pseudo-first-order conditions. This was based on the assumption that the concentration of hydroxide anion effectively remains constant, since the reaction is performed under buffered conditions; as a result the reaction could be treated as unimolecular. Pseudo first-order rate constants ($k_{1,obs}$) were obtained from the slopes of curves generated from plotting $\ln[NGM]$ versus time and linear regression analysis (see experimental section 3.2). Hydrolysis half-lives ($t_{1/2}$) were then calculated following the equation:

$$t_{1/2} = \frac{\ln(2)}{k_{1,obs}}$$
Figure 20 – Hydrolysis of N-methyl maleimide derivatives over time: a. UV/Vis absorbance over time; data was not collected in triplicate, since a single measurement (n=1) was shown to afford a clear trend in hydrolysis half-life that could guide future reagent design. b. determination of pseudo-first-order rate constants ($k_{1,\text{obs}}$) and hydrolysis half-lives ($t_{1/2}$).

It was surprising to observe that all three halogen-substituted maleimides hydrolyse fast, with a direct correlation between increasing electronegativity and hydrolysis rate being observed. It seems that the electronegative nature of halogen atoms causes increased polarization of the imide carbonyl and hence increased susceptibility towards hydrolysis. Notably, dibromomaleimide exhibits a short half-life of 17.9 minutes, which likely explains the poor outcome in the case of the sterically encumbered HSA-antibody fragment conjugation.
Diphenoxymaleimide has previously been developed as an NGM reagent of attenuated reactivity. It had been hypothesized that the incorporation of this functionality in a cross-linker could enable the development of a heterobifunctional reagent, with the diphenoxymaleimide reacting with the second protein. However, the short hydrolysis half-life of this reagent (t_{1/2} = 23.1 min) does not make it appropriate for such application. Remarkably, N-methyl dithiophenolmaleimide was shown to exhibit the most robust hydrolytic stability, followed by N-methylmaleimide, with half-lives of ca. 9 and 8 hours respectively. It was envisaged that the development of a bis-dithiophenol cross-linking reagent could help overcome the short half-life of DBM\textsubscript{CL} that was identified as an important factor affecting the conjugation efficiency.

2.1.1.3 Bis-dithiophenolmaleimide PEG linker

Based on the outcome of the hydrolysis study and previous reports suggesting thiophenolmaleimides to be very efficient conjugation reagents\textsuperscript{104,123,124}, the thiophenol derivative of linker DBM\textsubscript{CL} was synthesized (Scheme 26). Bis-dithiophenolmaleimide PEG linker DTPM\textsubscript{CL} was obtained in excellent yield upon the addition of thiophenol to DBM\textsubscript{CL}.

![Scheme 26 – Synthesis of bis-dithiophenolmaleimide PEG linker DTPM\textsubscript{CL}; Reagents and conditions: (i) thiophenol, NaHCO\textsubscript{3}, MeOH/ CH\textsubscript{2}Cl\textsubscript{2}, RT, 2 h.](image)

With the new linker in hand, the scFv-HSA conjugation was attempted (Scheme 27). Bridging of reduced scFv with DTPM\textsubscript{CL} proceeded partially to afford intermediate 31, while re-oxidized scFv was also observed by SDS-PAGE analysis (Figure 21, lane 3). This partial bridging could be attributed to solubility issues of linker DTPM\textsubscript{CL}. Even though the 1 kDa PEG spacer was expected to confer water solubility to the cross-linking reagent, the presence of the 4 hydrophobic thiophenol rings resulted in DTMP\textsubscript{CL} precipitating out of the buffer solution, even with increasing amounts of DMF (up to 15%). Nevertheless, the small amount of intermediate 31 obtained did not react with HSA in the subsequent step (Figure 21, lanes 5-13).
Scheme 27 – Attempted conjugation of scFv to HSA using DTPM<sub>CL</sub>; Reagents and conditions: (i) DTT, pH 7.4, RT, 1 h, (ii) DTPM<sub>CL</sub> (30 eq.), 15 min, RT, (iii) HSA (0.1-10 eq.), pH 7.4, RT, 24 h.

Figure 21 – SDS-PAGE analysis of attempted conjugation of scFv to HSA using linker DTPM<sub>CL</sub>: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Reduced scFv + linker, 4. Native HSA, 5. scFv-HSA conjugation mixture (10:1), 6. scFv-HSA conjugation mixture (5:1), 7. scFv-HSA conjugation mixture (3:1), 8. scFv-HSA conjugation mixture (2:1), 9. scFv-HSA conjugation mixture (1:1), 10. scFv-HSA conjugation mixture (1:2), 11. scFv-HSA conjugation mixture (1:3), 12. scFv-HSA conjugation mixture (1:5), 13. scFv-HSA conjugation mixture (1:10), after 24 h incubation.
The scFv dimerization that successfully proceeded with the DBM<sub>CL</sub> was also attempted with DTPM<sub>CL</sub>. SDS-PAGE analysis suggested the presence of bridged intermediate 31 and oxidised scFv; no dimer formation was observed (Figure 22).

**Figure 22** – Attempted scFv dimerization using linker DTPM<sub>CL</sub>: a. Reaction scheme; Reagents and conditions: (i) DTT (20 eq.), pH 7.4, RT, 1 h, (ii) Linker DTPM<sub>CL</sub> (0.42 eq.), pH 7.4, RT, 1 h, b. SDS-PAGE analysis of attempted dimerization of scFv using linker DTPM<sub>CL</sub>: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Reduced scFv + linker conjugation mixture, after 24 h incubation.

It was envisaged that the lack of water solubility of the bis-dithiophenolmaleimide linker was contributing to the unsuccessful conjugation. However, the small amount of bridged scFv-linker intermediate obtained would still be expected to react with HSA (Scheme 27) or scFv (Figure 22). The fact that no conjugation is observed suggests that the other end of the linker is not available for reaction. It is known that fast hydrolysis is not an issue with this reagent; it is possible that due to its hydrophobic character, the 2nd dithiophenolmaleimide is not solvent accessible.

The reaction was also repeated in the reverse order, where HSA was firstly reacted with DTPM<sub>CL</sub>. The linker was again insoluble, however 10% DMF was shown to be enough to promote quantitative conversion to HSA-linker intermediate, as evidenced by LC-MS (Figure 23).

**Figure 23** – LC-MS spectrum after reaction of HSA with DTPM<sub>CL</sub>; observed mass 67,838 (calculated 67,844).
Interestingly, the reaction proceeded with displacement of only one thiophenol group, in contrast to DBM_{CL}, where cross-linkage with another amino acid was observed to occur, displacing a 2\textsuperscript{nd} bromine atom. Comparing the two substituents, it seems that the unexpected debromination is caused by the increased inductive withdrawal from the bromine substituent, which makes the maleimide ring more electron deficient.

Reduced scFv was then added to HSA-linker intermediate 32 in different ratios (Scheme 28). A faint band corresponding to the desired conjugate 33 could be seen by SDS-PAGE analysis at ca. 80 kDa after overnight incubation that did not become more intense after 48 hours (Figure 24). The addition of scFv was repeated under more harsh conditions, i.e. pH 8 or 37 °C. However, this did not improve the yield. When the reaction was performed at double concentration, the conjugate band was observed after 1 hour. Nevertheless, prolonged incubation did not exceed ca. 5% yield of conjugate.

**Scheme 28** – Attempted conjugation of HSA to scFv using linker DTPM_{CL}; Reagents and conditions: (i) DTPM_{CL} (5 eq.), pH 7.4, RT, 15 min, (ii) scFv (0.1-10 eq.), pH 7.4, RT, 48 h (prior scFv reduction with DTT (20 eq.), pH 7.4, RT, 1 h).

**Figure 24** – SDS-PAGE analysis of attempted HSA-scFv conjugation with reagent DTPM_{CL}: M.

As previously mentioned, Cys-34 of HSA lies within a 9.5 Å deep hydrophobic crevice. This appears to facilitate reaction with the hydrophobic dithiophenolmaleimide to obtain intermediate 32. Therefore, the decreased water solubility of the linker does not affect the outcome of this reaction. However, it is likely that the 2nd dithiophenolmaleimide gets stuck to the surface of HSA inside hydrophobic pockets and is unable to react in the subsequent conjugation. To circumvent this, the synthesis of water soluble equivalents of this linker was attempted. In addition, thiophenol linkers containing longer PEG spacers were synthesised to examine whether these could provide increased hydrophilicity.

2.1.1.4 Water soluble derivatives of the bis-dithiophenolmaleimide PEG linker

2.1.1.4.1 Bis-dithiosalicylic acid maleimide PEG linker

It was envisaged that, through the presence of carboxylic acids, a thiosalicylic acid maleimide analogue of the thiophenol linker would have increased solubility in the aqueous environment. Linker 34 was obtained in moderate yield upon the addition of thiosalicylic acid to DBMCl (Scheme 29).

![Scheme 29 – Synthesis of bis-dithiosalicylic acid maleimide linker 34; Reagents and conditions: thiosalicylic acid, KOAc, MeOH/ CH2Cl2, RT, 4 h.]

As expected, the thiosalicylic acid linker was completely water soluble. However, bridging of scFv with linker 34 (30 eq.) was found to be troublesome. The reaction was incomplete, with less than 50% functionalized scFv being obtained after 1 hour, as seen by SDS-PAGE analysis and LC-MS. When 0.42 equivalents of the linker were added to the antibody fragment to attempt dimerization, oxidized and reduced scFv were the main species observed in the gel, suggesting reduced reactivity.
of this reagent. Interestingly, HSA did not react at all with this linker, even with excess reagent after prolonged reaction period (20 eq., 5 hours).

According to the literature, the environment around Cys-34 is anionic; a glutamic acid residue is situated at the opening of the crevice, explaining the repulsion of the negatively charged linker 34. In a similar manner, it is likely that negatively charged amino acids are in proximity to the disulfide bond of scFv, causing repulsion of the linker and hence incomplete bridging. Crystal structures indicate that when five molecules of myristic acid are bound to HSA, a conformational change occurs to open up the Cys-34 crevice, exposing the thiol group. Following a protocol reported by Curry et al., HSA was complexed with myristic acid and then reacted with the thiosalicylic acid linker to examine whether this could promote the reaction. Nevertheless, no improvement was observed.

2.1.1.4.2 Bis-di-4-aminothiophenolmaleimide PEG linker

Having identified a problem arising from the use of a negatively charged linker, the attention was turned to the synthesis of a bis-di-4-aminothiophenolmaleimide PEG linker. Even though the amine groups of aminothiophenol will not be protonated at physiological pH (i.e. pKₐ ca. 5), previous results within the Chudasama group (UCL Chemistry) have shown that such leaving groups can confer increased water solubility to the bridging reagent, compared to thiophenol. This is believed to be due to hydrogen bonding between the amine groups and water molecules. Unfortunately, the synthesis of this linker proved problematic. Addition of 4-aminothiophenol to DBMₐ did not afford the expected product 35 (Scheme 30). A complex mixture of products was obtained, where isolation of compound 35 proved impossible.

Scheme 30 – Attempted synthesis of bis-di-4-aminothiophenolmaleimide PEG linker 35; Reagents and conditions: (i) 4-aminothiophenol, NaHCO₃, MeOH/CH₂Cl₂, RT, 2 h.
To this end, an alternative stepwise synthetic strategy was employed (Scheme 31). Dibromomaleimide hexanoic acid 36 was obtained in 91% yield upon the reaction of 6-aminohexanoic acid with dibromomaleic acid. Addition of 4-aminothiophenol to 36 resulted again in a complex mixture. Though, after lengthy purification, product 37 was isolated pure in 53% yield. EDC coupling of 37 with the PEG diamine followed to afford linker 38. The NMR of the crude product suggested that the reaction was successful. However, the linker could not be isolated pure by column chromatography. An alternative synthesis or purification was not sought because a more powerful bifunctional conjugation reagent was later developed.

![Chemical structure](image)

Scheme 31 – Attempted synthesis of bis-di-4-aminothiophenolmaleimide PEG linker 38; Reagents and conditions: (i) 6-aminohexanoic acid, AcOH, reflux, 14 h, (ii) 4-aminothiophenol, NaOAc, MeOH, RT, 1 h, (iii) PEG(19) diamine, EDC.HCl, DMAP, DMF, RT, 20 h.

### 2.1.1.5 Longer bis-dithiophenolmaleimide PEG linkers

Dithiophenolmaleimide was shown to exhibit robust hydrolytic stability, a feature that is highly desirable in a cross-linking reagent. However, the hydrophobic character of this reagent led to inefficient conjugation, probably due to the 2nd NGM not being solvent accessible. Literature precedent suggests that smaller PEG molecules (such as 1kDa) might be small enough to get inserted into hydrophobic parts of HSA.\(^{323}\) In contrast, due to their size, longer PEG molecules are not inserted into HSA, but excluded. In addition, long PEG linkers (i.e. up to 20 kDa) have previously been employed for the construction of protein conjugates.\(^{276}\) As a result, in tandem to the synthesis of water soluble analogues of the bis-dithiophenolmaleimide PEG linker, the effect of using a longer PEG spacer was also examined.
2.1.1.5.1 6 kDa PEG linker

A 6 kDa bis-dithiophenolmaleimide PEG linker was firstly synthesized. The synthetic strategy employed is outlined in Scheme 32. Addition of thiophenol to the commercially available dibromomaleimide 14 afforded dithiophenolmaleimide 39. This was functionalized with an \(N\)-(methoxycarbonyl) to produce 40, which was then reacted with PEG\(_{6000}\) diamine to yield the target linker 41. Purification by column chromatography and prep-TLC was attempted, but due to its highly polar nature, linker 41 was stuck to the silica stationary phase and could not be eluted even with pure methanol. A different purification approach was then trialed; the crude mixture was dissolved in acetone and centrifuged at 0 °C, where the product precipitated as a yellow solid and isolated in 76% yield.

```
\begin{align*}
14 & \rightarrow_{66\%} \text{Br} \quad \text{PhS} \quad \text{NH} \quad \text{O} \quad \text{Br} \quad \text{PhS} \\
39 & \rightarrow_{92\%} \text{PhS} \quad \text{NH} \quad \text{O} \quad \text{PhS} \quad \text{OMe} \\
40 & \rightarrow_{76\%} \text{PhS} \quad \text{NH} \quad \text{O} \quad \text{O} \quad \text{PhS} \\
41 & \quad \text{N} \quad \text{O} \quad \text{O} \quad \text{SPh} \\
\end{align*}
```

**Scheme 32** – Synthesis of 6 kDa bis-dithiophenolmaleimide PEG linker 41 \((n = \text{ca. } 135)\); Reagents and conditions: (i) thiophenol, \(\text{NaHCO}_3\), MeOH, RT, 40 min, (ii) Me-chloroformate, \(N\)-Me morpholine, EtOAc, RT, 1 h, (iii) PEG\(_{6000}\) diamine, \(p\)-TsOH cat., \(\text{CH}_2\text{Cl}_2\), RT, 16 h.

However, problems arising from the use of such a long linker were immediately encountered. The reaction of HSA with linker 41 was shown to be incomplete. Even with 10 equivalents of linker, two bands could be seen by SDS-PAGE, corresponding to HSA and HSA+linker (Figure 25). Exactly the same outcome was observed when HSA was reacted with the commercially available maleimide PEG\(_{5000}\) (i.e. up to 200 equivalents). In addition, significant undesired dimerization of HSA was observed with linker 41.
PEG molecules are known to contain hydrodynamic radii that are 5- to 10-fold greater than what would be predicted by their nominal molecular weight. In addition, the flexibility of the polyethylene glycol chain results in PEG adopting different conformations. Therefore, it is likely that one of these conformations cannot reach Cys-34 that lies inside a 9.5 Å deep hydrophobic crevice.

Partial bridging was also observed upon reaction of linker 41 with scFv, which is known to contain a hindered disulfide bond (Figure 26, lane 3). Furthermore, the excess linker used could not be efficiently removed via ultrafiltration (30 kDa MWCO). By staining the SDS-PAGE gel with barium iodide, a stain used to detect PEG, the presence of free linker could be observed, after purification of the scFv-linker intermediate by ultrafiltration (Figure 26, lanes 5 and 6). This remaining linker was shown to react with native HSA added in the conjugation mixture (Figure 26, lanes 8-11). Purification of the scFv-linker intermediate was also attempted by size-exclusion chromatography (SEC) with the view to isolating only the bridged scFv species (Figure 26, lane 12). Upon addition of HSA to bridged scFv, a stronger band corresponding to the target conjugate could be observed (Figure 26, lanes 13-14). However, a small amount of HSA-linker species, as well as HSA homodimer could be observed, suggesting that either purification by SEC did not fully remove the excess PEG reagent or that transfer of the PEG reagent between the proteins is occurring. Overall, this protocol was deemed inefficient and attention was turned to the use of a PEG linker smaller than 6 kDa.
**Figure 26** – SDS-PAGE analysis of scFv-HSA conjugation using linker 41: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. scFv + linker 41 (30 eq., RT, 20 min) before ultrafiltration, 4. scFv + linker 41 after ultrafiltration, 5. scFv-linker before ultrafiltration (BaI₂ stain), 6. scFv-linker after ultrafiltration (BaI₂ stain), 7. Native HSA, 8. scFv-linker + HSA (10:1), 9. scFv-linker + HSA (5:1), 10. scFv-linker + HSA (3:1), 11. scFv-linker + HSA (2:1), 12. scFv-linker purified by SEC, 13. scFv-linker (after SEC) + HSA (1:1), 14. scFv-linker (after SEC) + HSA (3:1), after 24 h incubation.

### 2.1.1.5.2 2.5 kDa PEG linker

The commercially available 6 kDa PEG diamine that was previously employed was a heterogeneous mixture of polymers, with a distribution of molecular weights (MW). In addition, its large MW resulted in troublesome reactions and purification of the intermediates. It was decided that a PEG linker shorter than 6 kDa and of defined mass needs to be employed in order to achieve homogeneous protein modification. A bis-dithiophenolmaleimide PEG linker containing a 2.5 kDa spacer was synthesized as shown in *Scheme 33*. Mono-Boc protection of the commercially available PEG(19) diamine 42 proceeded successfully to afford 43. EDC coupling of compound 43 with PEG(13) diacid followed to yield the di-addition product 44. The free diamine was then obtained *in situ*, upon TFA deprotection. This was reacted initially with N-(methoxycarbonyl)-3,4-dithiophenolmaleimide 40. However, inefficient ring closure of the maleimide motif was observed. To this end, reaction with the dithiophenolmaleimide NHS ester 45 was tried instead, to successfully afford linker 46 in 79% yield.
Scheme 33 – Synthesis of 2.5 kDa bis-dithiophenolmaleimide PEG linker 46; Reagents and conditions:
(i) Boc₂O (0.5 eq. over 4 h), CH₂Cl₂, RT, 16 h, (ii) PEG(13) diacid, EDC·HCl, DMAP, DMF, RT, 40 h,
(iii) TFA/CH₂Cl₂, RT, 16 h, (iv) NEt₃, DMF, RT, 30 min, (v) 45, DMF, RT, 22 h.

With linker 46 in hand, the HSA-scFv conjugation was attempted (Scheme 34). HSA was initially
reacted with the 2.5 kDa linker to give intermediate 47; 12 equivalents of linker 46 were shown to
be necessary for full conversion. Reduced scFv was then added to this intermediate to afford only
a faint band that corresponded to the desired conjugate 48, as seen by SDS-PAGE analysis after
overnight incubation; this did not become more intense after 48 hours or by performing the reaction
at 37 °C. The same outcome was also observed when the reaction was repeated in the reverse order,
i.e. bridging of scFv with linker 46, followed by addition of reduced HSA.

Scheme 34 – Attempted conjugation of HSA to scFv using linker 46; Reagents and conditions: (i) Linker
46 (12 eq.), pH 7.4, RT, 15 min, (ii) scFv (0.2-5 eq.), pH 7.4, RT, 48 h (prior reduction of scFv with DTT
(20 eq.), RT, 1 h).
At this stage it was realized that a longer PEG linker may not actually aid the conjugation. The 6 kDa linker 41 resulted in incomplete reaction with HSA or scFv, while the reactivity of the 2.5 kDa PEG linker 46 seems to resemble that of the 1 kDa bis-dithiophenolmaleimide PEG linker DTPM<sub>Cl</sub>. Even though the presence of the 2.5 kDa PEG spacer conferred increased water solubility to reagent 46, compared to DTPM<sub>Cl</sub>, the 2<sup>nd</sup> NGM functionality was still not available for conjugation. Overall, it was postulated that the hydrophobicity of the thiophenol rings was responsible for that, leading to the 2<sup>nd</sup> NGM moiety getting inserted into hydrophobic areas of the protein.

Literature suggests that sufficient linker length is necessary for conjugation to HSA. The Cys34-containing cleft is approximately 10 Å deep (PDB file 1AO6). A crystal structure of the disulfide-stabilised anti-CEA scFv does not exist to determine the exact location of the disulfide bond in the tertiary structure. It is known that this disulfide is partially protected from the solvent, but it is anticipated that it would be located in a crevice that is less than 10 Å deep. The 1 kDa PEG spacer employed is about 60 Å long, therefore by choosing the appropriate NGM substituents, this linker length should be sufficient to facilitate the sterically hindered conjugation.

### 2.1.1.6 Bis-diiodomaleimide PEG linkers

Having identified NGM hydrolysis and hydrophobicity as important factors determining the conjugation efficiency, attention was turned to the use of iodine substituents. Diiodomaleimide was shown to hydrolyse more slowly than dibromomaleimide by a factor of 2.7, as evidenced by the UV study (see section 2.1.1.2, Figure 20, page 56). In addition, previous stopped-flow analysis has shown that diiodomaleimide is a faster conjugation reagent than dibromomaleimide. This suggests that the decreased orbital overlap between carbon and iodine leads to reduced mesomorphic donation from the halogen into the maleimide ring, making it a better conjugate acceptor. As a result, with the view to developing robust NGM cross-linking reagents, two diiodomaleimide linkers were synthesized.

Bis-diiodomaleimide PEG linker DIM<sub>Cl</sub> was firstly synthesized in a three-step procedure, as outlined in Scheme 35. The synthesis involved an initial Finkelstein-type reaction to afford diiodomaleimide 49 from commercially available dibromomaleimide 14. This was then activated
with an N-methoxycarbonyl to give 50, and reacted with PEG(19) diamine to afford the target linker in 66% yield.

Scheme 35 – Synthesis of bis-diiodomaleimide PEG linker DIM$_{\text{CL}}$; Reagents and conditions: (i) NaI, AcOH, reflux, 2 h, (ii) Me-chloroformate, N-Me morpholine, THF, RT, 20 min, (iii) PEG(19) diamine, CH$_2$Cl$_2$, RT, 4 h, then silica, RT, 16 h.

While the identity of NGM substituents has been shown to be critical in determining the hydrolytic stability of the cross-linkers, N-substitution is also known to have a big effect. The presence of a hydrophilic PEG spacer, directly attached to the maleimide ring, can actually accelerate hydrolysis on its own. The incorporation of a hexane spacer between the maleimide and the PEG chain was also examined, in order to obtain a conjugation reagent of increased stability towards hydrolysis.

Both bis-diiodomaleimide and bis-dibromomaleimide hexane-PEG linkers were synthesized for comparison. The synthetic approach used is outlined in Scheme 36. EDC coupling of mono-Boc protected hexane diamine 51 with PEG(13) diacid afforded the diaddition product 52. Boc deprotection and reaction with the N-(methoxycarbonyl) dibromo- 15 or diiodomaleimide 50 afforded linkers DBM$_{\text{CL-C6}}$ and DIM$_{\text{CL-C6}}$ in 55% and 67% yield respectively.

Scheme 36 – Synthesis of DBM$_{\text{CL-C6}}$ and DIM$_{\text{CL-C6}}$ linkers; Reagents and conditions: (i) PEG(13)diacid, EDC.HCl, DMAP, HOBr, DMF, 0 °C to RT, 20 h, (ii) TFA/ CH$_2$Cl$_2$, RT, 12 h, (iii) DIPEA, CH$_2$Cl$_2$, RT, 15 min, (iv) N-(methoxycarbonyl)-3,4-dibromomaleimide 15 or N-(methoxycarbonyl)-3,4-diiodomaleimide 50, RT, 1 h, then p-TsOH, RT, 18 h.
It has to be noted that in previous reactions, a catalytic amount of silica was used to promote ring closure upon reaction of an amine with the maleimide carbamate. However, in the current synthesis, $p$-toluenesulfonic acid was shown to be more efficient.

Before testing the HSA-scFv conjugation with the new linkers, their hydrolytic stability was assessed by a UV study, similar to the one described previously for the $N$-methylmaleimide analogues (see Section 2.1.1.2, Figure 20, page 56). The bis-dibromomaleimide PEG linker DBM<sub>CL</sub> was also included in the study to allow for a direct comparison. All linkers were incubated in buffer pH 7.4 (10% DMF) and their absorbance over time was monitored. The result is shown in Figure 27.

**Figure 27** – UV/Vis absorbance of bifunctional dibromomaleimide and diiodomaleimide cross-linking reagents, incorporating PEG or hexane-PEG spacers, over time; data from a single measurement (n=1).
Following the trend previously observed, DBMCL exhibited the fastest hydrolysis, with a half-life ($t_{1/2}$) of 17 minutes. Remarkably, the half-life of this linker was extended by a factor of ca. 3 upon incorporation of a hexane spacer, with linker DBMCL-C6 having a half-life of 53 minutes. It was interesting to observe that DBMCL-C6 had a similar hydrolysis profile as the iodo PEG linker DIMCL. This suggests that the effect of incorporating either less electronegative leaving groups or aliphatic N-substituents can be comparable in slowing down hydrolysis. Gratifyingly, DIMCL-C6 was shown to be the most hydrolytically stable linker with a half-life of 119 minutes.

2.1.2 Optimization of the HSA-scFv conjugation

The HSA-scFv conjugation was attempted with all three new linkers in tandem, to allow for a side-by-side comparison, before further optimization was carried out. HSA was initially reacted with each linker; full conversion was achieved after 15 minutes with 6 equivalents of DIMCL and 3 equivalents of DIMCL-C6 or DBMCL-C6. Reduced scFv was then added in the conjugation mixture in different ratios. Pleasingly, SDS-PAGE analysis after 24 hours of incubation revealed the successful formation of the target conjugate (Figure 28). Mixing the proteins in a 1:1 ratio (Figure 28, lanes 10-12) gave a quite strong band corresponding to the desired conjugate, in each case. However, the effect of adding one protein in excess in order to drive the reaction to completion was examined. The best conversion was obtained by adding 3 equivalents of modified HSA to reduced scFv (Figure 28, lanes 4-6), judging by the disappearance of scFv limiting reagent. When the reaction was repeated in the reverse order, i.e. initial bridging of scFv, followed by addition of HSA, only a faint band corresponding to the target conjugate could be observed by SDS-PAGE. This result is interesting, because it can give a better understanding of the HSA reactivity; if hydrolysis is not such an important concern now (i.e. at the time-scale of reaction), it seems that due to steric or conformation issues, HSA does not add to the scFv-linker intermediate, which will start to hydrolyze over time.
Interestingly, while \( \text{DBM} \text{CL-C6} \) was found to have a similar hydrolysis rate as \( \text{DIM} \text{CL} \), conjugation using the latter was shown to be more efficient (Figure 28, lanes 4 and 6). This seems to be in accordance with diiodomaleimide being a faster conjugation reagent than dibromomaleimide.\(^{104}\)

The conjugation was then repeated with the 2 iodomaleimide linkers \( \text{DIM} \text{CL} \) and \( \text{DIM} \text{CL-C6} \) to determine the exact timescale of reaction and identify which one gives the best conversion. The results were analysed by SDS-PAGE and size-exclusion chromatography (SEC). In both cases, a faint band corresponding to the conjugate was visible after only 30 minutes of incubation at pH 7.4; this became most intense at 2 hours with \( \text{DIM} \text{CL} \) (Figure 29a, lane 6) and 6 hours with \( \text{DIM} \text{CL-C6} \) (Figure 29b, lane 10).
Figure 29 – HSA-scFv conjugation with iodomaleimide linkers $\text{DIM}_{\text{CL}}$ and $\text{DIM}_{\text{CL-C6}}$; a. SDS-PAGE analysis (with linker $\text{DIM}_{\text{CL}}$): M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Native HSA, 4-12. HSA-scFv (3:1), over time, b. SDS-PAGE analysis (with linker $\text{DIM}_{\text{CL-C6}}$): M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Native HSA, 4-12. HSA-scFv (3:1), over time.

The SDS-PAGE results suggested that a higher conversion was achieved with the more hydrolytically stable linker $\text{DIM}_{\text{CL-C6}}$, signifying that the sterically hindered conjugation requires more time to proceed. This was confirmed by comparing the SEC chromatograms of the two conjugations (Figure 30a and 30b). While coelution between the target conjugate and excess HSA was observed, the amount of scFv present in the conjugation was shown to be smaller with linker $\text{DIM}_{\text{CL-C6}}$, compared to linker $\text{DIM}_{\text{CL}}$. As a result, further optimization was carried out with reagent $\text{DIM}_{\text{CL-C6}}$.

Figure 30 – HSA-scFv conjugation with iodomaleimide linkers $\text{DIM}_{\text{CL}}$ and $\text{DIM}_{\text{CL-C6}}$; a. SEC chromatogram of conjugation (with linker $\text{DIM}_{\text{CL}}$), b. SEC chromatogram of conjugation (with linker $\text{DIM}_{\text{CL-C6}}$).
It was observed that the concentration of the reaction was very important, significantly affecting the outcome of the conjugation. More specifically, doubling the concentration of the final reaction mixture (i.e. to 74.4 μM) was shown to not only speed up the reaction (i.e. 4 hours), but also lead to higher conversion, as judged by SDS-PAGE analysis (Figure 31a) and the SEC chromatogram of the reaction mixture, which revealed that a smaller amount of scFv was present relative to the conjugate (Figure 31b).

**Figure 31** – HSA-scFv conjugation with linker DIM<sub>CL-C6</sub>, performed at 74.4 μM; a. SDS-PAGE analysis: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Native HSA, 4-10. HSA-scFv (3:1), over time, b. SEC chromatogram of conjugation.
2.1.2.1 Hydrolysis of the HSA-scFv conjugate

As mentioned earlier work with NGM ADCs has highlighted the importance of obtaining a conjugate of robust serum stability, which is achieved by hydrolyzing the maleimide bridge. To this end, a study was pursued to determine the conditions required for the hydrolysis of the HSA-scFv conjugate (Scheme 37). It was decided to hydrolyse the conjugate immediately after the reaction, so that upon subsequent purification, pure hydrolysed conjugate would be obtained.

![Scheme 37 – Hydrolysis of the maleimide bridges in the HSA-scFv conjugate to maleamic acids.](image)

The conditions required for the hydrolysis of the NGM ADCs require incubation of the construct at pH 8.4, 20 °C for 72 hours. However, it was envisaged that such pH might be too high for the HSA conjugate. More specifically, albumin is known to undergo a structural transition above pH 8 from the N (neutral) to B (basic) form of the protein, which involves a “loosening-up” of the entire protein with changes in the tertiary structure being observed. Such transition can be reversible for native albumin, by lowering the pH. However, the effect that this would have on albumin modified with scFv is not known. Therefore it was envisaged that in order to maintain the native protein structure, pH 8.0 should be the highest tested.

To this end, the effect of pH, temperature and time on the rate of hydrolysis was examined. The reaction mixture was incubated at pH 7.4, 7.8 and 8.0 at 25 °C and 37 °C. In all cases, the higher temperature seemed to significantly speed up the rate of hydrolysis. The results were analysed by SDS-PAGE under reducing conditions. More specifically, the samples were mixed with loading buffer containing BME, which would cleave the maleimide bridges resulting in fragmentation, while the hydrolysed maleamic acids would be stable. The best conditions were shown to be pH 8.0 at 37 °C. The hydrolysis was monitored over time by SDS-PAGE (Figure 32a). Significant hydrolysis was observed after overnight incubation, which was complete at 72 hours, i.e. no further change in the intensity of the conjugate band was observed (Figure 32a, lane 12). This was confirmed upon subsequent purification of the conjugate, which did not cleave upon incubation
with the reducing loading buffer (see section 2.1.2.2, Figure 33b, lane 6, page 77). It has to be noted that the hydrolysis conditions were optimized using the reaction mixture at 37.2 μM. These conditions were then applied to the reaction performed at 74.4 μM and the same outcome was witnessed, with full hydrolysis at 72 hours (Figure 32b, lane 5).

### Figure 32 – SDS-PAGE analysis of the HSA-scFv conjugate hydrolysis at pH 8.0, 37 °C:

a. Reaction performed at 37.2 μM; M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Native HSA, 4. Reaction mixture before hydrolysis, 5-12. Reaction mixture incubated at pH 8.0, 37 °C over time (reducing loading buffer); b. Reaction performed at 74.4 μM; M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Native HSA, 4. Reaction mixture before hydrolysis, 5. Reaction mixture incubated at pH 8.0, 37 °C for 72 h (reducing loading buffer).

### 2.1.2.2 Purification and characterization of the HSA-scFv conjugate

Having successfully identified the hydrolysis conditions, the reaction was repeated on a large scale and purification was attempted. It was previously observed that the conjugate and excess albumin could not be separated by size exclusion chromatography. To this end, it was decided to purify the mixture firstly by affinity chromatography, using a column that was precharged with Ni2+ ions (HisTrap™). The presence of a (His)6 tag at the C-terminus of scFv would cause only the scFv and conjugate to bind to the column, due to the high affinity of the hexahistidine tag to coordinate transition metals. In contrast, HSA would not bind and could be removed in subsequent washing steps. By increasing the imidazole concentration of the buffer (i.e. to 200 mM), the two (His)6-tagged proteins could be then eluted from the column, since they would no longer compete for binding sites on the nickel resin.

When the mixture was loaded on the nickel column, a small amount of unspecific binding of HSA was observed. This was minimized by adding 10 mM and 20 mM of imidazole in the binding and
washing buffers, respectively. Having successfully removed HSA, the reaction mixture was then purified by SEC to afford a pleasing 57% yield of pure conjugate, with respect to scFv limiting reagent (Figure 33). This yield was a significant improvement over previous reported yields on the direct conjugation of albumin to larger proteins.

**Figure 33** - Purification of HSA-scFv conjugate by size exclusion chromatography: a. SEC chromatogram post Ni-IMAC purification, b. SDS-PAGE analysis of HSA-scFv conjugation and purification: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Native HSA, 4. HSA-scFv conjugation mixture (3:1), 5. Purified HSA-scFv conjugate, 6. Purified HSA-scFv conjugate (reducing loading buffer).

The purity of the conjugate was confirmed by SDS-PAGE analysis (Figure 33b, lane 5), as well as SEC (Figure 34a). In addition, the conjugate was shown to be stable upon storage at -20 °C for 9 months, without any aggregation or cleavage being observed (Figure 34b).

**Figure 34** – SEC chromatograms of HSA-scFv conjugate; a. After purification, b. After storage at -20 °C for 9 months.
The purified conjugate was also analysed by LC-MS, where a peak corresponding to HSA-scFv conjugate was observed (Figure 35). Partial cleavage of the conjugate was shown to occur under the acidic conditions and increased temperature during LC-MS analysis. This is known for such maleamic acid linkers, with cleavage of the imide bond occurring to release the amine and form the corresponding anhydride, which can further hydrolyze to a maleic acid.\textsuperscript{329} MALDI-TOF analysis was also obtained to confirm that the conjugate was intact, and that the LC-MS conditions were responsible for the fragmentation (See Appendix 5.1, Figure 85, page 207).

LC-MS analysis revealed the loss of the second iodine from the maleimide motif that had reacted with HSA, in a similar manner to the debromination that had previously been discussed. Remarkably, this maleimide had not been hydrolysed; this explains why it did not cleave under the acidic LC-MS conditions. In order to gain a better insight into this unexpected reactivity, a tryptic digestion was performed.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure35}
\caption{LC-MS analysis of HSA-scFv conjugate; HSA-scFv observed mass of 94,237 (calculated 94,237 for hydrolysed NGM on scFv and cross-linked NGM on HSA), HSA-linker-\textsubscript{NH\textsubscript{3}}\textsuperscript{+} observed mass of 67,402 (calculated 67,405), scFv-maleic acid observed mass of 26,851 (calculated 26,851).}
\end{figure}

\subsection*{2.1.3 Tryptic digestion of HSA}

In general, trypsin protease is known to cleave the protein backbone on the C-terminal side of lysine and arginine residues.\textsuperscript{330} The smaller peptides produced can then facilitate protein identification by tandem mass spectrometry (MS/MS). More specifically, ions will be formed corresponding to the peptide mixture, which are separated by \textit{m/z} in the first stage of MS. The ions of interest can then be selected and fragmented by collision induced dissociation (CID), predominantly to \textit{b}- and \textit{y}- ions. These can then be detected in a second stage of MS. The end
result is a mass spectrum containing characteristic fragments of the selected peptide, from which information regarding the structure and amino acid sequence can be obtained. If a cross-linkage has occurred, the absence of the corresponding b- and y- ion fragments will be observed.

With the view to simplifying analysis, HSA was initially reacted with N-methyl diiodo-, dibromo-and dichloromaleimide. All dihalomaleimides were shown to react with HSA in identical manner, i.e. displacement of the second halogen atom was observed, suggesting that cross-linking with the same nucleophilic residue near Cys-34 is taking place. To identify this point of attachment, a tryptic digestion of HSA modified with commercially available N-methyl dibromomaleimide was performed. The analyzed conjugate was constructed under the same conditions as the cross-linked conjugates, including incubation at pH 8.0, 37 °C, 72 h to infer thiol stability, which was essential due to the presence of DTT in the digest protocol. LC-MS/MS analysis of the peptide mixture post digestion revealed the point of cross-linkage to be Lysine-41 (Figure 37). This is consistent with its position 9.5 Å from Cys-34 (PDB file 1AO6) (Figure 36). While trypsin cleaves peptide bonds C-terminal to lysine, it will not recognize a lysine residue that has undergone reaction. This is due to the absence of a positive charge that is necessary for interaction with the enzyme’s active site. As a result, a ‘missed-cleavage’ will be obtained that corresponds to a longer peptide, i.e. as observed in the current experiment.

**Figure 36** – PDB crystal structure of HSA (PDB file 1AO6), showing the distance (in Å) between Cys-34 and the proximal Lys-41
In addition, a second Cys-34 containing peptide was identified and characterized by LC-MS/MS (Figure 38). The mass of this species was shown to correspond to loss of the bromine atom from the maleimide motif and addition of hydroxide. However, complete coverage of the b- and y- ions was witnessed, suggesting that no cross-linkage was present. The structure of this species was postulated to be that of a hydroxy-thiomaleimide, as depicted in Figure 38. It was hypothesized to have formed via enamine hydrolysis of the cross-linked amino-thiomaleimide conjugate. The fact that the mass of this species was not present in the LC-MS of the HSA-scFv conjugate suggests that the enamine hydrolysis should have occurred during the digestion conditions. This is likely, since the protocol involved complete denaturation of HSA (i.e. with 6M Guanidine.HCl), which would result in the Cys-34 crevice opening up. As a result, it would be more susceptible to hydrolysis upon subsequent incubation at pH 8.0, 37 °C, which was required for disulfide reduction and alkylation, prior to digestion. The digestion was also performed with HSA reacted with dibromomaleimide containing a longer N-substituent (i.e. biotin) for a direct comparison. This molecule was synthesized as shown in Scheme 38. Initial HBTU-mediated coupling between mono-Boc protected hexane diamine and biotin afforded 53, which was then reacted with dibromomaleic acid to give dibromomaleimide-biotin 54 in 61% yield. Upon reaction of this molecule with HSA and subsequent trypsin digestion, identical digestion results to HSA reacted with N-methyl dibromomaleimide were obtained (see Appendix 5.2, Figures 87 and 88, page 208 and 209).

Scheme 38 – Synthesis of dibromomaleimide Biotin 54; Reagents and conditions: (i) D-Biotin, HBTU, DIPEA, DMF, RT, 24 h, (ii) TFA, RT, 18 h, (iii) dibromomaleic acid, AcOH, reflux, 3 h.
Figure 37 – LC-MS/MS spectrum of Cys-34/Lys-41 cross-linked peptide obtained by selecting m/z 919.16 [M+4H]^4+ as the precursor ion for CID; 3+ = trivalent ion (m/3), 2+ = bivalent ion (m/2), 0 = m-H₂O, * = m-NH₃.
Figure 38 – LC-MS/MS spectrum of Cys-34 peptide (N-Me maleimide derivative) obtained by selecting m/z 1280.52 [M+2H]^{2+} as the precursor ion for CID; 3+ = trivalent ion (m/3), 2+ = bivalent ion (m/2), 0 = m-H_{2}O, * = m-NH_{3}. Proposed structure of hydroxy-thiomaleimide – postulated to form via enamine hydrolysis of the cross-linked amino-thiomaleimide species.
By taking into account the digestion data and a thiol stability experiment performed before and after subjecting the conjugate to hydrolysis conditions (Scheme 39, see LC-MS spectra in Appendix 5.3, Figures 89-91, page 210), the following conclusion was reached. Upon initial reaction of HSA with N-Me dibromomaleimide (or any dihalomaleimide derivative), the 2\textsuperscript{nd} halogen atom is displaced. The product is thiol cleavable at this stage. It was postulated that cross-linking may initially occur with Tyrosine-84, which is only 2.7 Å away from Cysteine-34 (PDB file 1AO6). This would give rise to an aryloxy/thiomaleimide; these species are known to be cleavable with thiols.\textsuperscript{322} After incubation at pH 8.0, 37 °C for 72 h, Tyrosine-84 is displaced by Lysine-41, resulting in an amino/thiomaleimide species (detected by trypsin digestion), which is stable to thiols. This is expected, since lysine is known to mesomerically release electron density into the maleimide ring, deactivating it towards nucleophilic attack. In addition, further characterization of the intermediate species was not possible by digestion, due to the protocol requiring DTT for the cleavage of disulfide bonds; this would also result in cleavage of the maleimide motif. Overall, this unexpected cross-linking provided interesting results; it showed that thiol stability of NGMs cannot only be achieved by hydrolysis, but also upon reaction with a mesomerically donating functionality, which will deactivate the ring.

![Scheme 39](attachment:image.png)

**Scheme 39** – Modification of HSA with N-Me dibromomaleimide and treatment of the resultant conjugate with BME, before and after incubation at pH 8, 37 °C, 72 h; Reagents and conditions: (i) N-Me dibromomaleimide (1.5 eq.), pH 7.4, RT, 20 min, (ii) pH 8.0, 37 °C, 72 h, (iii) BME (100 eq.), pH 8.0, 37 °C, 2 h, (iv) BME (100 eq.), pH 8.0, 37 °C, 2 h.
2.1.4 HSA-Fab conjugation

Having developed a robust linker for the conjugation of HSA to scFv, it was decided to apply this technology to the conjugation of a different antibody fragment to HSA, with the view to further testing the scope of the methodology. The Fab fragment of the monoclonal IgG1 antibody trastuzumab (Herceptin™) was chosen, which was obtained upon a sequential enzymatic digest of the full antibody with pepsin and papain, in 78% overall yield.

The HSA-Fab conjugation was then attempted with all four iodo- and bromomaleimide cross-linkers, for a direct comparison (Scheme 40). In a similar manner to HSA-scFv conjugation, 3 equivalents of modified HSA over Fab were shown to be the best in promoting the reaction. By analyzing the SDS-PAGE results (Figure 39), it was observed that the conjugation was highly efficient with either iodomaleimide linker used, with maximum conversion being observed after only 1 hour (Figures 39a and 39b, lane 4).

**Scheme 40** – HSA-Fab conjugation (X = I or Br); Reagents and conditions: (i) Linker DIM\textsubscript{CL} (6 eq.) or Linker DIM\textsubscript{CL-C6} (3 eq.) or Linker DBM\textsubscript{CL} (6 eq.) or Linker DBM\textsubscript{CL-C6} (3 eq.), pH 7.4, RT, 15 min. (ii) Fab (0.33 eq.), pH 7.4, RT, 24 h (prior Fab reduction: TCEP (10 eq.), 37 °C, 1 h).

**Figure 39** – SDS-PAGE analysis of HSA-Fab conjugation; a. DIM\textsubscript{CL}, b. DIM\textsubscript{CL-C6}, c. DBM\textsubscript{CL}, d. DIM\textsubscript{CL-C6}; M. Molecular marker, 1. Native Fab, 2. Reduced Fab, 3. Native HSA, 4. HSA-Fab (3:1), 1 h, 5. HSA-Fab (3:1), 6 h; HC = heavy chain, LC = light chain.
As a result, the DIMCL linker was chosen to carry out further optimization. The more efficient conjugation observed was anticipated, due to the Fab disulfide bond being solvent accessible and hence exhibiting faster reactivity. A small amount of product was also formed with DBMCl and DBMCL-C6 (Figures 38c and 38d). However, the sterically hindered nature of HSA seems to predominate and since these two bromomaleimide linkers are of reduced hydrolytic stability or reactivity, further conversion was not witnessed.

Apart from the target conjugate band that appears at ca. 130 kDa, SDS-PAGE analysis also revealed the presence of a second species at ca. 70 kDa, which seems to correspond to the conjugate lacking a covalent linkage to the Fab light chain (FabLC). This was confirmed by obtaining an LC-MS spectrum of the conjugate (Figure 40). On closer inspection, it was observed that the light chain had undergone a complete decarboxylative cleavage of the C-terminal cysteine to afford an enamide. This light chain decarboxylation has previously been reported to occur on trastuzumab-thiomaleimide conjugates, when treated under photochemical conditions. More recently, it has also been shown to proceed even if light is excluded. In addition, the HSA-FabHC species was shown to consist of two signals; one at 91,711 which is in accordance with that of a thioenol species (X=SH), produced upon the decarboxylative cleavage of FabLC and one at 91,681 which corresponds to just loss of iodine leaving group. Based on previous data, it can be hypothesized that a nucleophilic residue of the heavy chain has reacted at this site. Since no free FabLC is observed, it is suggested that reaction should have occurred with the thioenol species,

**Figure 40** – LC-MS analysis of HSA-Fab conjugate to confirm loss of FabLC; FabLC observed mass of 23,362 (calculated 23,440 for native FabLC), HSA-FabHC observed mass of 91,681 (calculated 91,682 for loss of I leaving group) and 91,711 (calculated 91,714 for thioenol species (X=SH)), HSA-Fab observed mass of 115,136 (calculated 115,137 for hydrolysed maleimide on Fab), HSA-Linker-NH3+ observed mass of 67,403 (calculated 67,404), Fab-anhydride observed mass of 47,733 (calculated 47,733); the conjugate shown here was synthesized using DIMCL-C6; HC = heavy chain, LC = light chain.
which is very reactive. By performing the conjugation in buffer containing increased salt concentration, i.e. 500 mM NaCl, a significant improvement was observed (Figure 41, lane 5). The undesired decarboxylation was substantially minimized. This high salt concentration is expected to disrupt any non-covalent interactions. It is not obvious why this has led to such a substantial improvement, but one hypothesis is that the conformation of the carboxylate has been altered, reducing the efficiency of decarboxylation.

![Figure 41](image)

**Figure 41** – SDS-PAGE analysis of HSA-Fab conjugation (with DIMCl linker) at high salt concentration (500 mM NaCl): M. Molecular marker, 1. Native Fab, 2. Reduced Fab, 3. Native HSA, 4. HSA-Fab conjugation (20 mM NaCl), 5. HSA-Fab conjugation (500 mM NaCl).

The conjugation was then repeated on a larger scale, at 74.4 μM. After 1 hour, the reaction mixture was incubated at pH 8.0, 37 °C for 72 hours to effect hydrolysis. Subsequent purification by SEC afforded HSA-Fab in 49% yield, with respect to Fab limiting reagent (Figure 42a).

![Figure 42a](image)

**Figure 42a** – Purification of HSA-Fab conjugate by size exclusion chromatography: a. SEC chromatogram, b. SDS-PAGE analysis of HSA-Fab conjugation and purification: M. Molecular marker, 1. Native Fab, 2. Reduced Fab, 3. Native HSA, 4. HSA-Fab conjugation mixture (3:1), 5. Purified HSA-Fab conjugate, 6. Purified HSA-Fab conjugate (reducing loading buffer).
The purity of this conjugate was confirmed by SDS-PAGE (Figure 42b, lane 5), SEC (Figure 43), as well as LC-MS (Figure 44). In a similar manner to HSA-scFv, partial cleavage of the conjugate was shown to occur under the acidic conditions and increased temperature during LC-MS analysis. Gratifyingly, no signal corresponding to HSA-FabHC species was observed.

**Figure 43** – SEC chromatogram of HSA-Fab conjugate after purification.

**Figure 44** – LC-MS analysis of HSA-Fab conjugate; HSA-Fab observed mass 115,144 (calculated 115,147 for hydrolysed NGM on Fab and Lys-41 cross-linked NGM on HSA), HSA-linker-NH$_3^+$ observed mass of 67,412 (calculated 67,415), Fab-anhydride observed mass of 47,733 (calculated 47,734).
2.1.5 Analysis of the HSA-scFv and HSA-Fab conjugates by ELISA

In order to test the influence that the conjugation of albumin to scFv and Fab has on the immunoreactivity of the fragments, an Enzyme-Linked Immunosorbent Assay (ELISA) was carried out. A 96-well polystyrene plate was coated either with CEA or HER2 antigen. The HSA-scFv or HSA-Fab conjugates, as well as native scFv or Fab were applied to it. Following incubation with a detection and an HRP-linked antibody, o-phenylenediamine HCl was added. This was converted to 2,3-diaminophenazine, by HRP, that shows strong absorbance at 490 nm, allowing spectrophotometric quantification of the antigen-binding activity of the conjugate and native protein. As it can be observed in Figure 45, the two conjugates show comparable antigen binding activity to the native antibodies.

\[ \text{Figure 45 – ELISA analysis of conjugates; a. HSA-scFv (IC}_{50} = 3.82 \pm 1.12 \text{ nM}) \text{ and scFv (IC}_{50} = 3.68 \pm 1.13 \text{ nM}) (against CEA/ detection by anti-His antibody), b. HSA-Fab (IC}_{50} = 8.37 \pm 1.24 \text{ nM}) \text{ and Fab (IC}_{50} = 5.07 \pm 1.13 \text{ nM}) (against HER2/ detection by anti-HER2-Fab specific antibody).}\]

Following this, a sandwich ELISA was developed, in order to demonstrate the bivalent nature of the two conjugates. This involved binding to CEA or HER2, followed by detection with an anti-albumin antibody. A schematic illustration of the protocol used is shown in Figure 46.
As expected, only the HSA-Fab and HSA-scFv conjugates gave a signal, since they can bind simultaneously to the antigen and anti-albumin (Figure 47).

Figure 46 – Schematic illustration of the sandwich ELISA protocol used for analysis of the HSA-scFv conjugate; the same protocol was used for HSA-Fab, with HER2 instead of CEA antigen.

Figure 47 – Sandwich ELISA analysis of conjugates (against the corresponding antigen/ detection by anti-albumin antibody); a. HSA-scFv, native HSA and native scFv (against CEA), b. HSA-Fab, native HSA, native Fab (against HER2).
Having successfully developed the sandwich ELISA, it was then used to analyze the serum stability of the two conjugates (Figure 48). More specifically, the constructs were incubated in human serum at 37 °C for 7 days, at a final concentration of 40 nM. Aliquots were removed from the incubator each day and frozen at -80 °C. After the final time point, samples were analysed by sandwich ELISA, with absorbance as a function of time. The fact that the signal has been maintained throughout suggests the stability of HSA-scFv and HSA-Fab in vitro.

Figure 48 – Serum stability study, analysed by sandwich ELISA; a. HSA-scFv, b. HSA-Fab.
2.2 Trifunctional protein conjugates

Having successfully identified the appropriate NGM linker design to promote conjugation of two proteins in a sterically hindered system, this technology was further extended to allow for the creation of triprotein conjugates, which could give access to a novel class of therapeutics. Two trifunctional diiodomaleimide cross-linking reagents were synthesized, incorporating either PEG or hexane-PEG spacers for increased hydrolytic stability. The merit of these linkers in creating HSA-scFv-scFv, HSA-scFv-Fab and tri-scFv constructs was then evaluated.

2.2.1 Synthesis of trifunctional linkers

The two trifunctional reagents \( \text{tDIM}_{\text{CL}} \) and \( \text{tDIM}_{\text{CL-C6}} \) were synthesized, as outlined in Scheme 41. The synthesis commenced with an amide coupling between the commercially available tri-NHS ester 55 and PEG(7) mono-Boc protected diamine to give the trifunctional scaffold 56. After Boc deprotection, this was reacted either with \( N-\text{(methoxycarbonyl)}-3,4\text{-diiodomaleimide} \) 50 to afford tri-linker \( \text{tDIM}_{\text{CL}} \) or with Boc-6-aminohexanoic acid to yield the hexane-PEG trifunctional scaffold 57. Boc deprotection of 57 followed and subsequent reaction with \( N-\text{(methoxycarbonyl)}-3,4\text{-diiodomaleimide} \) 50 afforded reagent \( \text{tDIM}_{\text{CL-C6}} \). Interestingly, during this synthesis it was observed that addition of excess base (i.e. DIPEA) to the protonated triamines generated \textit{in situ} upon Boc deprotection of constructs 56 or 57 was efficient in enabling their subsequent reaction with the maleimide moiety. However, purification by column chromatography was not effective at separating the linkers from the DIPEA-TFA salt present in the reaction mixture. An alternative protocol was developed, in which the protonated triamine in each case was dissolved in water, the pH was raised to 11 and the linker extracted into DCM.
Scheme 41 – Synthesis of tris-diiodomaleimide PEG (tDIM_{CL}) and hexane-PEG (tDIM_{CL-C6}) cross-linking reagent; Reagents and conditions: (i) PEG(7) mono-Boc diamine, MeCN, RT, 20 h, (ii) TFA/CH₂Cl₂, RT, 18 h then H₂O, pH 11, extracted into CH₂Cl₂, (iii) N-(methoxycarbonyl)-3,4-diiodomaleimide 50, CH₂Cl₂, RT, 2 h, then silica, RT, 1 h, (iv) TFA/CH₂Cl₂, RT, 18 h then DIPEA, CH₂Cl₂, RT, 30 min, (v) Boc-6-aminohexanoic acid, EDC.HCl, DMAP, HOBr, DMF, 0°C to RT, 20 h, (vi) TFA/CH₂Cl₂, RT, 18 h, then H₂O, pH 11, extracted into CH₂Cl₂, (vii) N-(methoxycarbonyl)-3,4-diiodomaleimide 50, CH₂Cl₂, RT, 2 h, then silica, RT, 18 h.
2.2.2 Tri-protein conjugation with HSA

Building on prior work with the HSA-scFv conjugation, it was decided to evaluate the use of the trifunctional platform in linking HSA to two scFv monomers. Such a construct would increase the functional binding affinity of scFv as well as its half-life, something not achieved by simple scFv dimerization.

The trifunctional reagent with the hexane-PEG spacer, **tDIM**\(_{CL-C6}\) was employed to investigate the scope of this conjugation (Scheme 42). HSA was initially reacted with the linker; 15 equivalents of the reagent were shown to be necessary to minimize any unwanted HSA dimerization. After removal of excess reagent, reduced scFv was added into the conjugation in different equivalents (1-4 eq.) and the progress was monitored by SDS-PAGE (Figure 49). After 4 hours, a strong band corresponding to HSA-scFv species could be observed, while the desired tri-conjugate was only seen as a faint band at ca. 130 kDa. Prolonged incubation did not change this outcome.

**Scheme 42** – Conjugation of HSA to two scFv monomers; Reagents and conditions: (i) Linker **tDIM**\(_{CL-C6}\) (15 eq.), pH 7.4, RT, 15 min, (ii) scFv (1-4 eq.), pH 7.4, RT, 4 h or 24 h.
In previous experiments on HSA-scFv conjugation it was observed that higher conversion to the target product was obtained when a small excess of HSA-linker was added to reduced scFv. This is presumably due to partial hydrolysis of the maleimide motif occurring over the time that is required to purify the HSA-linker intermediate before reaction with scFv. It seems that the increased steric hindrance present in the current tri-protein system, in combination with maleimide hydrolysis prior to scFv addition are responsible for inhibiting the conjugation.

With the view to overcoming the presence of maleimide hydrolysis as much as possible, a stepwise conjugation was attempted, with the maleimide species being present in a slight excess in each step (Scheme 43). HSA was initially reacted with linker tDIM_{CL-C6}. Following purification of this intermediate, 0.5 equivalents of scFv were added in the conjugation to afford HSA-scFv (Figure 50, lane 6). After 1 h at RT, this mixture was purified by a short HisTrap™ column to remove excess HSA-linker species (Figure 50, lane 7). The pure HSA-scFv conjugate was then reacted with 0.5 equivalents of scFv. A band of higher intensity corresponding to the target trimer could be observed this time by SDS-PAGE analysis, however the conversion was still low (Figure 50, lane 8). Addition of Fab to the HSA-scFv species was also trialed, to examine whether a more accessible disulfide bond could facilitate the conjugation. Nevertheless, only a faint band corresponding to HSA-scFv-Fab could be witnessed after overnight incubation (Figure 50, lane 9).
Scheme 43 – Stepwise HSA-scFv-scFv and HSA-scFv-Fab conjugation, with purification of HSA-scFv intermediate by HisTrap™; Reagents and conditions: (i) Linker tDIM_{CL-C6} (15 eq.), pH 7.4, RT, 15 min, (ii) scFv (0.5 eq.), pH 7.4, RT, 1 h (prior scFv reduction with DTT (20 eq.), pH 7.4, RT, 1 h), purification of reaction mixture by HisTrap™, (iii) Reduced scFv (0.5 eq.), pH 7.4, RT, 24 h, (iv) Fab (0.5 eq.), pH 7.4, RT, 24 h (prior Fab reduction with TCEP (10 eq.), 37 °C, 1 h).
Figure 50 – SDS-PAGE analysis of stepwise HSA-scFv-scFv and HSA-scFv-Fab conjugation, with purification of HSA-scFv intermediate by HisTrap™: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Native Fab, 4. Reduced Fab, 5. HSA + tDIMCL-C6, 6. HSA-Linker + 0.5 eq. scFv (1 h), 7. HSA-scFv purified by HisTrap™, 8. HSA-scFv + 0.5 eq. scFv (24 h), 9. HSA-scFv + 0.5 eq. Fab (24 h).

In order to minimize the time of exposure of the linker in the aqueous solution and hence maleimide hydrolysis, the HSA-scFv-Fab conjugation was repeated using an in situ stepwise protocol, without HisTrap™ purification of the HSA-scFv intermediate. In addition the reactions were performed at a higher concentration of 100 μM to examine whether this could improve conversion. More specifically, following purification of the HSA-linker intermediate, 0.9 equivalents scFv were added in the conjugation mixture. After 1 hour, 1 equivalent Fab was added and the reaction was monitored by SDS-PAGE analysis. However, no significant improvement in the conjugation efficiency was witnessed. Overall, it seems that the increased steric hindrance of albumin predominates and does not enable addition of the third protein, with maleimide hydrolysis increasing over time, deactivating the reagent towards conjugate addition.
2.2.3 Tri-scFv conjugation

In tandem to the synthesis of albumin tri-protein conjugates, the application of the trifunctional reagents in linking three scFv fragments was also examined. In particular, tri-scFv constructs (ca. 80 kDa) were identified as ideal targets since they are considered within an optimal “tumour-target zone”, prospectively offering a good balance between reduced clearance rates (compared to scFvs) and increased tumour penetration (compared to intact immunoglobulins).\textsuperscript{199,212} In addition, the trivalency can confer a significant increase in avidity, which is beneficial for low affinity binders, since poor retention can result in diffusion out of the tumour tissue, requiring higher concentrations of the antibody in order to bind a significant fraction of the antigen. Efforts to reduce blood clearance have also been made by creating scFv dimers, which have exhibited increased tumour uptake compared to monomers.\textsuperscript{153,154} However, since their size (ca. 55 kDa) is just below the renal threshold, their circulating half-life has only been shown to increase by a factor of \textit{ca.} 2.5.

2.2.3.1 Generation of scFv homotrimer

The first target was to generate a homotrimer of the anti-CEA scFv, by disulfide bridging (Figure 51a). A one-pot procedure was developed, where linker t\textsc{dim}CL was shown to be efficient enough, without necessitating the use of the more hydrolytically stable linker. More specifically, the antibody fragment was initially reduced with DTT. Then, following removal of excess reducing agent, the scFv was incubated with 0.33, 0.30 and 0.28 equivalents of t\textsc{dim}CL to determine which stoichiometry gave more efficient conversion (Figure 51b). Gratifyingly, a strong band corresponding to the target trimer could be seen at \textit{ca.} 90 kDa after 1 hour, which did not become more intense after prolonged incubation. In addition, 0.33 equivalents of linker were shown to afford the best conversion, with gel densitometry analysis (using ImageJ) indicating that 75% of scFv monomer had been consumed. The presence of scFv dimer at 50 kDa was also observed, which was anticipated due to the increased steric hindrance of the system inhibiting addition of the 3\textsuperscript{rd} protein, in combination with maleimide hydrolysis over time. Performing the reaction at a higher concentration did not further improve the conversion and higher amounts of scFv aggregation started being observed.
Figure 51 – Tri-scFv conjugation; a. Reaction scheme; Reagents and conditions: (i) DTT (20 eq.), RT, 1 h, (ii) Linker tDIM<sub>CL</sub> (0.28-0.33 eq.), RT, 1-20 h. b. SDS-PAGE analysis of conjugation: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3-14. Conjugation mixture with different equivalents of linker at different timepoints (as marked on each lane on the figure).

The conjugation was then repeated on a bigger scale with 0.33 equivalents tDIM<sub>CL</sub>, at 37.2 μM. After 1 hour, the mixture was buffer exchanged to pH 8.0 and incubated at 37 °C for 72 hours to promote hydrolysis. Subsequent purification by SEC afforded tri-scFv in 30% yield (Figure 52).

Figure 52 – Purification of tri-scFv conjugate by size exclusion chromatography: a. SEC chromatogram, b. SDS-PAGE analysis of tri-scFv conjugation and purification: M. Molecular marker, 1. Native scFv, 2. Reduced scFv, 3. Tri-scFv conjugation mixture, 4. Purified tri-scFv conjugate, 5. Purified tri-scFv conjugate (reducing loading buffer).
The successful formation of the conjugate was confirmed by LC-MS (Figure 53) and MALDI-TOF analysis (see Appendix 5.1, Figure 86, page 207), with the latter showing only a signal for the target conjugate. This confirmed again that cleavage during LC-MS is due to the acidic conditions employed.

**Figure 53** – LC-MS analysis of tri-scFv conjugate; tri-scFv observed mass of 81,794 (calculated 81,800), di-scFv-linker-NH$_3^+$ observed mass of 54,958 (calculated 54,948), scFv-maleic acid observed mass of 26,851 (calculated 26,851).

The SEC chromatogram of the purified product showed only one peak corresponding to the trimer (Figure 54a), while ELISA analysis demonstrated that trimerization has successfully led to a decrease in the IC$_{50}$ (Figure 54b).

**Figure 54** – Characterization of tri-scFv conjugate; a. SEC chromatogram after purification, b. ELISA analysis of tri-scFv (IC$_{50}$ 0.65 ± 1.11 nM) and scFv IC$_{50}$ 3.50 ± 1.22 nM) against CEA.
2.2.3.2 Biodistribution of tri-scFv and scFv

Having successfully synthesized anti-CEA tri-scFv, its biodistribution in BALB\C mice was evaluated in comparison to the unmodified monomer, with the view to establishing whether the increased molecular weight of tri-scFv had conferred slower blood clearance. The biodistribution study was performed by collaborators at UCL Cancer Institute.

Tri-scFv and scFv were conjugated to commercially available N-hydroxysuccinimidyl DOTA for radiolabeling with $^{111}$In. ELISA analysis after DOTA conjugation demonstrated that immunoreactivity was identical to that before conjugation (see Appendix 5.4, Figure 92, page 211). In addition, size exclusion chromatograms of tri-scFv and scFv obtained before and after DOTA-labelling confirmed that the integrity of the conjugates had been maintained. Radiolabeling with $^{111}$In was achieved with 97% incorporation efficiency. The $^{111}$In-labeled conjugates were injected into BALB\C mice, by tail injection, in doses equivalent to 0.33 MBq of monomer and 0.44 MBq of trimer. Groups of three mice were then sacrificed at 3 h and 24 h; the tissues were analysed and the data for both conjugates plotted as % ID/g (Figure 55a-b). For clearance studies, blood was collected at 10, 20, 40, 60 and 180 minutes and plotted as % ID/g against time (Figure 55c).

The monomer demonstrated rapid uptake in the kidney, reaching 149.2 % ± 45.4 % ID/g in 3 h (Figure 55a – Note that values >100% ID/g are possible, when the tissue is smaller than 1 g$^{332}$). The blood clearance was shown to be extremely fast (Figure 55c), with the blood level of activity at 4.2 % ± 0.5 % ID/g in 3 h. In contrast, tri-scFv showed a much more even distribution (Figure 55b). Its blood clearance was markedly reduced (Figure 55c) compared to the monomer, with the blood level of activity exhibiting a 7-fold increase (29.0 % ± 4.8 % ID/g in 3 h). Kidney accumulation was also observed after 3 h, with the level of activity being at 44.7 % ± 4.6 % ID/g (ca. 3x lower than the monomer), however this seems to be consistent with previous results observed for fragments labelled with metallic isotopes, such as $^{111}$In.$^{154}$ Overall, the trimerization of scFv by disulfide bridging was shown to significantly increase the circulatory half-life of the construct, in comparison to the unmodified monomer. Although an scFv dimer was not included in this study, literature suggests that that such constructs can increase the half-life up to ca. 2.5x,$^{153,154}$ compared to the monomer. As a result, in combination with the further increase in valency, scFv trimerization represents a promising platform for antibody therapeutic development.
Figure 55 – Biodistribution of $^{111}$In-labeled scFv and tri-scFv in BALB/C mice; a. Biodistribution of monomer scFv at 3 h and 24 h, b. Biodistribution of trimer scFv at 3 h and 24 h, c. Blood clearance of scFv and tri-scFv. Each data point represents the mean of 3 mice ± SD.
2.2.3.3 Generation of bispecific scFv heterotrimer

Following the successful production of scFv homotrimer, it was decided to target the generation of a bispecific scFv heterotrimer, incorporating one anti-CEA scFv and two anti-CD3 scFv fragments, giving rise to a Bispecific T-cell Engager (BiTE) construct. As previously discussed, T-cell recruiting bispecific antibodies that simultaneously bind tumor-associated antigens and T-cell receptors (e.g. CD3) have shown excellent clinical efficacy in the treatment of hematological malignances and various solid tumors. A bispecific anti-CEA/anti-CD3 scFv heterodimer construct is currently in phase 1 clinical trials for the treatment of colorectal adenocarcinomas (Medi-565, MedImmune). It is produced by fusion of two scFv fragments and exhibits fast clearance and short half-life in vivo. It was anticipated that the production of a bispecific anti-CEA/anti-CD3 heterotrimer could represent a construct of increased half-life, due to its optimal size, as well as increased binding affinity to CD3 receptor that could lead to enhanced cell-killing ability.

The efficiency of the bispecific scFv heterotrimer conjugation was evaluated in a stepwise protocol, with anti-CEA scFv initially reacting with the linker, followed by addition of anti-CD3 in the conjugation mixture (Scheme 44). The more hydrolytically stable tDIMCL-C6 cross-linking reagent was employed for optimizing this conjugation. It was initially hypothesized that if bridging

![Scheme 44 – Synthesis of tri-scFv anti-CEA/anti-CD3 BiTE; Reagents and conditions: (i) DTT (20 eq.), pH 7.4, RT, 1 h, (ii) tDIMCL-C6 (30 eq.), pH 7.4, RT, 8 min, (iii) Reduced anti-CD3 scFv (2 eq.), pH 7.4, RT, 2 h (prior anti-CD3 scFv reduction with DTT (20 eq.), pH 7.4, RT, 1 h).](image)
of reduced anti-CEA scFv could be efficient with a very small excess of reagent (i.e. 1.05 eq.), then purification of this intermediate would not be required and anti-CD3 scFv could be added into the reaction mixture directly. This would minimize exposure of the reagent in the aqueous solution and hence hydrolysis. However, during optimization it was observed that this was not feasible, since reaction of scFv with 1.05 equivalents of tDIMCL-C6 was promoting scFv dimerization and trimerization to a great extent, even when the reaction was performed at 0 °C. This undesired scFv multimerization was shown to be minimized with 30 equivalents of reagent and only when scFv was added dropwise (i.e. within 2 min) to linker in solution to afford bridged anti-CEA scFv monomer (Figure 56b, lane 3). After 8 minutes at RT, the excess linker was removed by ultrafiltration and 2 equivalents of anti-CD3 scFv were added in the conjugation mixture. The progress of the reaction was monitored by SDS-PAGE analysis, where the highest possible conversion to the target trimer was observed after 2 hours (Figure 56b, lane 7). The mixture was then buffer exchanged to pH 8.0 and incubated at 37 °C for 72 hours to effect hydrolysis. Subsequent purification by column chromatography afforded the target trimer in 10% yield (Figure 56a). The purity of this construct was confirmed by SDS-PAGE analysis, where a single band at ca. 100 kDa was observed (Figure 56b, lane 8).

![Figure 56](image)

**Figure 56** – Tri-scFv anti-CEA/anti-CD3 BiTE conjugation; a. SEC purification, b. SDS-PAGE analysis conjugation and purification: M. Molecular marker, 1. Native anti-CEA scFv, 2. Reduced anti-CEA scFv, 3. Anti-CEA scFv + linker tDIMCL-C6, 4. Native anti-CD3 scFv, 5. Reduced anti-CD3 scFv, 6. Tri-scFv BiTE conjugation mixture (1 h), 7. Tri-scFv BiTE conjugation mixture (2 h), 8. Purified tri-scFv BiTE, 9. Purified tri-scFv BiTE (reducing loading buffer).
An LC-MS spectrum was also obtained (Figure 57). Due to low signal intensity, increased background noise was observed, however the peaks for the trimer and expected fragments could still be observed.

**Figure 57** – LC-MS analysis of tri-scFv conjugate; tri-scFv BiTE observed mass of 84,593 (calculated 84,318), Anti-CEA/anti-CD3 di-scFv-linker-NH$_3^+$ observed mass of 55,910 (calculated 56,366), Anti-CD3 scFv-maleic anhydride observed mass of 27,953 (calculated 27,953) – Note: Due to the heterogeneity of the two native scFv monomers, the expected molecular weights shown in this figure were calculated using the scFv monomer signal of highest intensity in each case, which results in the observed signals differing by ca. 300-500 Da from the expected values.

Overall, the isolated yield of the bispecific heterotrimer construct was rather low, suggesting that the limit of this NGM conjugation platform has been identified. It seems that the increased steric hindrance of the system predominates, with addition of the third protein occurring to a lower extent and also at a slower rate, at which point maleimide hydrolysis becomes significant, further inhibiting bioconjugation. Nevertheless, this approach enabled access to a novel anti-CEA/anti-CD3 scFv heterotrimer construct. A collaboration has been initiated with the UCL Cancer Institute and is currently underway to examine whether this construct of increased molecular weight and two arms binding CD3 could lead to more efficient T-cell killing of CEA-expressing cells, compared to an anti-CEA/anti-CD3 heterodimer.
2.3 Summary

In summary, a range of NGM cross-linking reagents were synthesized and their application in conjugating proteins containing different reactive handles (i.e. a single cysteine to a disulfide), sterically hindered conjugation sites, as well as linking together three proteins, was evaluated.

Human Serum Albumin (HSA) was selected as the ideal cysteine containing protein, since its single free thiol (Cys-34) available for conjugation is positioned within a 9.5 Å crevice, diminishing its solvent accessibility and hence reactivity. With the view to testing the scope of the methodology, the conjugation of HSA to scFv and Fab antibody fragments that contain disulfide bonds of distinct reactivity was explored. It was envisaged that this approach would enable access to a chemical platform for the half-life extension of antibody fragments, creating favourable constructs for therapeutic or imaging applications.

Overall, linker hydrolysis was identified as the most important factor affecting the conjugation efficiency. This seems to be due to the sterically hindered nature of this conjugation, requiring reagents which are stable for prolonged period, since premature hydrolysis would deactivate the second maleimide motif towards conjugate addition. Interestingly, the NGM substituents were shown to have a big effect on the rate of hydrolysis, with a correlation between increasing electronegativity of the leaving group and hydrolysis rate being observed. Notably, the bis-dibromomaleimide PEG cross-linking reagent that had previously been exploited in protein-protein conjugation was shown to have a short half-life of 17 minutes, resulting in poor yields of the target HSA-antibody fragment conjugates.

In contrast, dithiophenolmaleimide was identified as the most hydrolytically stable NGM, with a half-life of 9 hours. This led to the development of a bis-dithiophenolmaleimide cross-linking reagent DTPMcI to examine whether it can overcome the shortcomings of the dibromomaleimide analogue. The decreased water solubility of this reagent was immediately encountered as it was shown to precipitate out of the buffer solution and result in partial bridging of scFv. Reaction of the linker with HSA was observed to be efficient, presumably due to Cys-34 being located in a hydrophobic crevice. However, upon incubation of functionalized HSA with the antibody fragment, the desired conjugate formation never exceeded ca. 5% yield. Since hydrolysis was not a concern with this reagent, it was postulated that the hydrophobicity of the thiophenol rings might result in the 2nd NGM getting stuck to the surface of HSA inside hydrophobic pockets and hence
being unable to react in the subsequent conjugation. A water-soluble thiosalicylic acid analogue of this linker was then synthesized, but it did not react at all with HSA, potentially due to the anionic environment around Cys-34. Dithiophenolmaleimide cross-linking reagents incorporating longer PEG spacers were also synthesized to examine whether these can provide increased water solubility, but also due to literature precedent suggesting that longer PEG molecules are not inserted into HSA, but excluded. Nevertheless, these reagents did not offer any improvement in the conjugation yield. In addition, it was observed that the 6 kDa PEG linker was reacting only partially with HSA, potentially due its large hydrodynamic radius resulting in a conformation that cannot efficiently reach Cys-34 that lies inside a 9.5 Å deep crevice.

Dithiodiiodomaleimide cross-linking reagents were demonstrated to offer an optimized platform for the construction of albumin-antibody fragment conjugates. This was due to iodine substitution offering increased stability towards hydrolysis, compared to bromomaleimide reagents, but also decreased mesomeric donation into the maleimide ring, making it a better conjugate acceptor. Two bis-DIM PEG cross-linking reagents were synthesized; \textbf{DIM}_{\text{CL}} and \textbf{DIM}_{\text{CL-C6}} that incorporated a hexane spacer between maleimide and PEG to further increase the hydrolytic stability (Figure 58).

\textbf{Figure 58} – Bis-diiodomaleimide cross-linking reagents.

\textbf{DIM}_{\text{CL}} linker was shown to be highly efficient in conjugating the sterically hindered HSA to Fab antibody fragment that contains an accessible disulfide bond, with the target HSA-Fab conjugate being isolated in 49% yield, after purification (Scheme 45). In contrast, the scFv fragment that incorporates a more hindered disulfide bond necessitated the use of the more hydrolytically stable \textbf{DIM}_{\text{CL-C6}} linker in order to achieve high conversion to the target conjugate. Following purification, HSA-scFv was isolated in 57% yield (Scheme 45). Pleasingly, these yields represent
a substantial improvement over previous reported yields (i.e. up to 12%) on the direct conjugation of albumin to larger proteins.

**Scheme 45** – Generation of HSA-Fab and HSA-scFv conjugates, using the diiodomaleimide platform. Furthermore, both conjugates were hydrolyzed with the view to obtaining serum stable constructs. Their stability was confirmed upon incubation in human serum for 1 week, while their antigen-binding activity was maintained, as demonstrated from ELISA analysis. An interesting observation through these experiments was the loss of the second iodine of the maleimide motif that had reacted with HSA. LC-MS/MS analysis post tryptic digestion revealed the point of cross-linkage to be lysine-41, resulting in a serum stable amino/thiomaleimide species.

With the view to further extending the scope of NGM cross-linking reagents, two trifunctional diiodomaleimide linkers were developed, based on the previously optimized bifunctional reagent design; **tDIMCl** and **tDIMCl-c6** (Figure 59).
Using linker tDIM<sub>CL</sub>, the generation of an anti-CEA scFv homotrimer was achieved in 30% yield, in an efficient one-pot procedure (Scheme 47). ELISA analysis confirmed that the trimerization of scFv by disulfide bridging has successfully resulted in a decrease of the IC<sub>50</sub> value, consistent with the increase in avidity offered by the trivalency. In addition, a biodistribution study demonstrated that the circulatory half-life of the construct has significantly increased, compared to the unmodified monomer, suggesting that tri-scFv conjugates could lead to the development of powerful constructs for in vivo applications.

Following this, the generation of a bispecific anti-CEA/anti-CD3 scFv heterotrimer of increased avidity towards CD3 was targeted. This construct was isolated in 10% yield, through a stepwise conjugation approach using linker tDIM<sub>CL-C6</sub> (Scheme 46). In vivo experiments are currently underway to determine whether this novel BiTE conjugate containing two arms targeting the CD3 receptor, could lead to enhanced cell-killing ability, compared to a heterodimer.

Overall, iodomaleimide cross-linking reagents were shown to offer a robust and optimized platform for the construction of homogeneous protein-protein conjugates, as well as enabling access to novel trifunctional constructs of favourable in vivo characteristics. It seems that the limit of this platform was identified towards the stepwise creation of a trifunctional conjugate containing two different proteins, where the steric hindrance of the system appears to predominate, with maleimide hydrolysis over time becoming important. However, it is hoped that based on the current developments, future work in this area would facilitate the creation of such conjugates.
Scheme 46 – Generation of anti-CEA scFv homotrimer and anti-CEA/anti-CD3 scFv heterotrimer, using the diiodomaleimide platform.
2.4 Future work

Having demonstrated the facile production of albumin-antibody fragment conjugates using the optimized NGM platform, an exciting next step would be to alter the linker design to incorporate a clickable handle for the attachment of a functional moiety, such as a drug or a fluorophore (Figure 60).

**Figure 60** – Proposed bifunctional diiodomaleimide cross-linking reagent, incorporating a clickable handle for the synthesis of functionalized albumin-scFv conjugates; FM=functional moiety.

Further to this, the development of a bifunctional reagent containing two mutually orthogonal clickable handles could enable the dual functionalization of a protein-protein conjugate, giving access to powerful constructs (Figure 61). Possible applications of such a reagent could be in the creation of an scFv bispecific construct incorporating two drugs or a drug and a long PEG molecule for half-life extension.

**Figure 61** – Proposed bifunctional diiodomaleimide cross-linking reagent, incorporating two clickable handles for the synthesis of functionalized scFv heterodimer conjugates; FM=functional moiety.
While the conjugation of two proteins has been shown to work very efficiently, addition of a third protein to a tri-NGM scaffold was observed to occur to a lower extent, due to the increased steric hindrance of the system, with maleimide hydrolysis over time deactivating the reagent. To this end, the creation of a trifunctional conjugate could be attempted via click attachment of the third protein (Figure 62). This methodology can also enable selectivity between the second and third additions, allowing incorporation of a different third protein to create trispecific antibodies. Furthermore, the linker design could also be manipulated accordingly to include points of attachment for non-biological moieties.

![Figure 62](image)

**Figure 62** – Proposed bifunctional diiodomaleimide cross-linking reagent, incorporating a clickable handles for the synthesis of tri-protein conjugates; FM=functional moiety.

Finally, the unexpected Lys-41 cross-linkage that was observed to occur with Cys-34 of albumin and dihalomaleimides suggested that serum stability of NGMs can also be achieved upon reaction with a mesomerically donating functionality that will deactivate the ring. Based on this interesting result, reaction of a single cysteine-containing protein with a dibromomaleimide, followed by addition of a functionalized amine could afford a doubly labeled serum stable construct (Scheme 47). This approach might be particularly beneficial for maleimides containing N-substituents that are sensitive to the increased pH and temperature conditions, required for hydrolysis.

![Scheme 47](image)

**Scheme 47** – Proposed reaction of cysteine-containing protein with a dibromomaleimide, followed by addition of a functionalized amine to afford a doubly labeled serum stable construct; FM=functional moiety.
2.5 Cysteine-to-lysine transfer (CLT)

2.5.1 Background

A secondary aim of this work was to investigate a new bioconjugation methodology for the construction of homogeneous antibody conjugates, incorporating serum stable amides, in response to the heterogeneous product mixtures obtained by traditional NHS ester chemistry. It was envisaged that this could be achieved via acyl transfer from a disulfide bond to proximal lysine residues. The central hypothesis behind this idea was that acyl transfer would occur only to lysine residues that are proximally positioned within the protein structure.

It was decided to test the feasibility of this approach on the Fab fragment of trastuzumab, which contains a single disulfide bond. Interestingly, examination of the Fab crystal structure revealed that out of the 26 lysine residues present in total, only 2 are in proximity to the disulfide bond, within ca. 10 Å (Figure 63). Both Lys-228 and Lys-235 are located on the Fab heavy chain.

Since the two lysine residues are within an equal distance from the disulfide bond, it was deemed necessary to target both of them, with the view to achieving homogeneous modification. The proposed strategy is outlined in Scheme 48. After disulfide bond reduction, the two cysteine residues will be reacted with a small molecule thioester to afford the doubly labeled product, in an initial transthioesterification step. The acyl groups will then be transferred onto the two lysine residues of the Fab heavy chain, driven by the thermodynamic stability of the resultant amides over thioesters. It has to be noted that the Fab heavy and light chains are held together by strong
non-covalent interactions. Therefore, reduction of the disulfide bond is not expected to significantly alter the distance between the light chain cysteine and the lysine residues on the heavy chain. Finally, re-oxidation of the Fab disulfide bond will afford a traceless modification, with two lysine residues being selectively labeled.

**Scheme 48** – Proposed strategy for cysteine-to-lysine transfer (CLT).

In Native Chemical Ligation (NCL), peptide thioesters are usually synthesized as the less reactive alkyl derivatives and activated *in situ* to the more reactive aryl thioesters upon addition of a thiol catalyst, such as MPAA.\(^{112}\) As a result, an aryl thioester was the first target for achieving thiol-thioester exchange with the Fab disulfide bond. Preliminary work in this project was carried out by Irene Benni (Baker group, UCL), who synthesized aryl thioester 58 and evaluated conditions for transthioesterification (Scheme 49). The presence of an alkyne moiety would enable later functionalisation *via* click chemistry.

**Scheme 49** – Reaction of reduced Fab fragment with aryl thioester 58; Reagents and conditions: (i) TCEP (10 eq.), pH 7.4, 37 °C, 1 h, (ii) aryl thioester 58 (1-10 eq.), 37.2 μM, pH 6.0-8.0, RT, 15 min-24 h.
Optimization was carried out by differing the equivalents of aryl thioester added, reaction time, as well as pH. In general, the aryl thioester 58 was shown to be very reactive, promoting fast exchange with the cysteine residues. However, in all experiments it was observed that before transthioesterification was complete, unspecific lysine modification started appearing. An example is shown in Figure 64, for reaction with 10 equivalents of thioester, at pH 7.4 for 15 minutes. It seems that a reactive lysine residue is present on the light chain that gives rise to a second acyl addition at the timescale of reaction.

**Figure 64** – Reaction of reduced Fab with aryl thioester 58; Reagents and conditions: aryl thioester 58 (10 eq.), pH 7.4, RT, 15 min (prior disulfide bond reduction with TCEP (10 eq.), 37 ºC, 1 h); LC = light chain, HC = heavy chain, 0, +1 refer to the number of acyl groups added.

A control experiment was then performed, in which the Fab fragment was reacted with thioester 58 under identical conditions, but without prior disulfide bond reduction (Figure 65). Apart from native Fab, a second signal was observed that corresponds to addition of one acyl group, confirming that an intermolecular reaction with a lysine residue is taking place (Figure 65a). This sample was then treated with TCEP, in order to reduce the disulfide bond and identify in which chain the modification was present; it was confirmed to be on the light chain (Figure 65b).

**Figure 65** – Reaction of Fab with aryl thioester 58, without prior disulfide bond reduction; a. Reagents and conditions: aryl thioester 58 (10 eq.), pH 7.4, RT, 15 min. b. Reaction of Fab with aryl thioester 58 (10 eq.), pH 7.4, RT, 15 min, followed by treatment with TCEP (10 eq.), pH 7.4, 37 ºC, 1 h, LC = light chain, HC = heavy chain, 0, +1 refer to the number of acyl groups added.
In contrast, in a control reaction performed with an alkyl MESNa thioester, no unspecific lysine modification was witnessed with 10 equivalents of reagent after overnight incubation. Thus, it was a secondary aim of the current project to examine the scope of transthioesterification with an alkyl thioester and optimize conditions for transfer onto the lysine residues of the Fab fragment.

2.5.2 Optimization of the transthioesterification with MESNa thioester

In general, native chemical ligation with preformed thiophenyl thioesters has been shown to reach completion in 30 minutes, whereas the same ligation with MESNa alkyl thioesters might require more than 24 hours, which represents a ca. 50-fold reduction in the rate of transthioesterification.\(^{163}\) This is believed to be due to the increased electrophilicity of the thiophenyl thioester towards nucleophilic attack, as well as the better leaving group ability of thiophenol (i.e. thiophenol \(\text{pK}_a\) 6.6, MESNa \(\text{pK}_a\) 9.2), once the cysteine has reacted with the thioester.\(^{161}\) The decreased reactivity of MESNa thioester was shown to be beneficial in the current project, since it seems to avoid unspecific reaction of the thioester with lysine residues present on the Fab fragment. It was envisaged that through extensive optimization, the appropriate conditions for reaction with the Fab heavy and light chains would be discovered.

The MESNa thioester 59 was synthesized in 72% yield via an EEDQ coupling of the commercially available 4-pentynoic acid and MESNa (Scheme 50). Prior synthesis had involved the use of EDC coupling reagent. However, the target product was isolated in 26% yield, due to difficulty in separating the highly polar thioester from EDC, by column chromatography.

\[
\text{Scheme 50} \quad \text{Synthesis of alkyl MESNa thioester 59; Reagents and conditions: (i) EEDQ, MeCN/ DMF, 80 ^\circ \text{C}, 18 \text{ h}.}
\]
With reagent 59 in hand, conditions for thiol-thioester exchange with the two cysteine residues of the Fab fragment were evaluated (Scheme 51). More specifically, optimization was carried out by differing the equivalents of thioester added, concentration, temperature and reaction time. The pH was kept constant at 7.4; this was shown to provide a good balance between increased thiol reactivity and decreased lysine reactivity. Following disulfide bond reduction at pH 7.4, the thioester was added directly into the Fab solution, with other variables being altered.

Scheme 51 – Optimization of Fab transthioesterification with alkyl thioester 59: Reagents and conditions:
(i) TCEP (10 eq.), pH 7.4, 37 °C, 1 h, (ii) 59 (10-100 eq.), 37.2-200 μM, pH 7.4, RT-37 °C, 2 h-48 h.

Key examples are shown in Figure 66. The first experiment was performed with 10 equivalents of thioester, at 37.2 μM, RT for 24 hours, conditions identical to the control reaction that had previously been shown to afford no unspecific conjugation (Figure 66a). Under these conditions, partial transthioesterification with either chain was observed. In addition, the presence of a heavy chain species containing two thioester molecules was witnessed, but this could be attributed to acyl transfer onto one lysine of the FabHC. Increasing the temperature to 37 °C with the view to promoting faster exchange was attempted, but this was shown to afford higher loaded species (Figure 66b). The effect of performing the reaction either at a higher concentration or with a larger excess of thioester for a shorter reaction period was then evaluated (Figure 66c, 65d, respectively). In both cases, a significant improvement in thiol-thioester exchange was observed, with higher conversion being achieved when 100 equivalents of thioester were used. This is probably due to the excess reagent present pushing the equilibrium towards the product. In addition, in all cases it was observed that transthioesterification of the light chain cysteine was harder to achieve. This cysteine is C-terminal; as a result it is expected to have a higher pKₐ than the heavy chain cysteine and hence lower amount of thiolate present at physiological pH. In addition, it is possible that the carboxylate of the LC cysteine repels the negatively charged MESNa thioester.
Gratifyingly, when the reaction was performed with 100 equivalents of thioester at 150 μM, complete thioester exchange was witnessed after 4 hours at 22 °C (Figure 67a). The reaction temperature was shown to be very important, with slightly higher temperatures promoting the appearance of a heavy chain species containing two thioesters, potentially due to acyl transfer. A control reaction was also performed, in which the Fab fragment was incubated with the thioester under the same conditions, without disulfide bond reduction being carried out (Figure 67b). Under these conditions, no reaction with the thioester was witnessed.
2.5.3 Optimization of the acyl transfer

Having identified the conditions required for the initial transthioesterification step to occur, attention was turned to the optimization of the acyl transfer. Literature precedent suggests that in NCL the first transthioesterification step is the rate-determining step. In contrast, for ligations proceeding through macrocyclic transition states (>6 membered TS), the S-to-N acyl transfer becomes rate-determining. The ring size has been shown to become critical in determining the successful outcome of the process, with slower formation of the amide bond resulting in increased competition from thioester hydrolysis. However, these experiments were performed with small peptides and cannot be predictive for a protein system, since the tertiary structure can affect the positioning and proximity of the reactive groups.

It was anticipated that transfer would be faster at a higher pH, since an increased fraction of unprotonated lysine would be present, i.e. at physiological pH lysine is predominantly protonated (ca. 99.9%, pKₐ 10.5). However, due to the pH scale being logarithmic, a pH increase of 1 will correspond to ca. 1% unprotonated lysine, but at the same time the concentration of hydroxide anion will rise 10-fold, increasing the chance of competing hydrolysis of the thioester. Higher temperature can lead to faster acyl transfer, due to the increased thermal kinetic energy of reactants, but also promote hydrolysis. As a result, to examine the feasibility of acyl transfer, the Fab fragment after transthioesterification was incubated in buffers of different pH and temperature (Scheme 52).
Scheme 52 – Optimization of acyl transfer; Reagents and conditions: (i) TCEP (10 eq.,) pH 7.4, 37 °C, 1 h, (ii) thioester 59 (100 eq.), pH 7.4, 150 μM, 22 °C, 4 h, (iii) pH 7.4-9.5, 4 °C-37 °C, 24 h-72 h.

In previous reactions involving disulfide bond modification, EDTA was always added in the buffers to prevent metal-catalyzed disulfide re-oxidation. However, for the current experiments, buffers without EDTA were prepared, so that disulfide bond re-oxidation after acyl transfer is not intercepted. The acyl transfer was examined at pH 7.4, 7.7, 8.0, 8.4, 9.0 and 9.5 at 4, 12, 22 or 37 °C. The reactions were monitored up to 72 hours, by SDS-PAGE analysis (Figures 68-70). Gel samples were also run after reacting the mixtures with Ellman’s reagent, which has been shown to promote disulfide bond re-oxidation. In all cases, no difference was observed, suggesting that the heavy and light chain species shown in the gels are due to the presence of a thioester on either or both cysteine residues. As a result, complete transfer is expected to have been achieved when heavy and light chains are no longer visible.

**Figure 68** – SDS-PAGE analysis of acyl transfer reaction at pH 7.7 and 7.4; a. 24 h, b. 48 h, c. 72 h: M. Molecular marker, 1-3. pH 7.7, at 12 °C, 22 °C, 37 °C respectively, 4-6. pH 7.4, at 12 °C, 22 °C, 37 °C respectively.
Molecular marker, 1-3. pH 8.4, at 12 °C, 22 °C, 37 °C respectively, 4-6. pH 8.0, at 12 °C, 22 °C, 37 °C respectively.

Overall, the reaction was shown to be slow and thioester hydrolysis over time was unavoidable. Higher pH or temperature resulted in faster transfer, but also increased hydrolysis, as judged by LC-MS, where a significant amount of native Fab, as well as Fab modified with only one acyl group were witnessed. A few examples are shown in Figure 71a-c.
a. pH 8.0, 37 °C, 24 h

b. pH 7.7, 37 °C, 72 h

c. pH 9.0, 12 °C, 48 h

**Figure 71** – LC-MS of acyl transfer at: a. pH 8.0, 37 °C, 24 h, b. pH 7.7, 37 °C, 72 h, c. pH 9.0, 12 °C, 48 h; 0 = Native Fab, +1 acyl group observed mass 47,717-47,724 (calculated 47,718), +2 acyl groups observed mass 47,797-47,799 (calculated 47,798); * = These peaks do not represent additional acyl transfers; their existence seems to be due to the native Fab sample containing small extra species of higher molecular weight, which are now getting modified as well, either with one or two acyl groups (see Appendix 5.5, Figure 93, page 212 for details).
Pleasingly, one set of conditions was shown to afford a much better balance between increased acyl transfer and reduced thioester hydrolysis. Upon incubation at pH 8.4, 12 °C for 72 h, a loading of 1.4 acyl groups per Fab was calculated from the LC-MS spectrum (Figure 72).

![LC-MS of acyl transfer at pH 8.4, 12 °C, 72 h; 0 = Native Fab, +1 acyl group observed mass 47,718 (calculated 47,718), +2 acyl groups observed mass 47,798 (calculated 47,798).](image)

**Figure 72** – LC-MS of acyl transfer at pH 8.4, 12 °C, 72 h; 0 = Native Fab, +1 acyl group observed mass 47,718 (calculated 47,718), +2 acyl groups observed mass 47,798 (calculated 47,798).

### 2.5.4 Stability and activity of Fab after modification

Having successfully identified a pH and temperature combination that results in ca. 70% labeling of the two lysine residues, the stability of the resultant Fab construct towards thiols was then evaluated, in comparison to the Fab prior to acyl transfer (Scheme 53).

![Reaction of Fab fragment after transthioesterification and acyl transfer with BME; Reagents and conditions: (i) pH 8.4, 12 °C, 72 h, (ii) BME (100 eq.), pH 8.0, 37 °C, 2 h.](image)

**Scheme 53** – Reaction of Fab fragment after transthioesterification and acyl transfer with BME; Reagents and conditions: (i) pH 8.4, 12 °C, 72 h, (ii) BME (100 eq.), pH 8.0, 37 °C, 2 h.
Upon incubation with 100 equivalents of BME at pH 8.0, 37 °C for 2 hours, no change in the LC-MS spectrum of the Fab after acyl transfer was witnessed, suggesting its stability towards thiol attack (Figure 73).

**Figure 73** – LC-MS of Fab after acyl transfer, incubated with 100 eq. BME, pH 8.0, 37 °C, 2 h; 0 = Native Fab, +1 acyl group observed mass 47,719 (calculated 47,718), +2 acyl groups observed mass 47,799 (calculated 47,798).

In contrast, the LC-MS spectrum of the Fab fragment prior to acyl transfer, incubated with 100 equivalents of BME showed that the modification has been reversed, in accordance with the instability of thioesters towards thiols (Figure 74). On closer inspection, a small amount of acyl transfer seems to have occurred, during buffer exchange (to pH 8.0) and subsequent reaction at pH 8.0, 37 °C.

**Figure 74** – LC-MS of Fab before acyl transfer, incubated with 100 eq. BME, pH 8.0, 37 °C, 2 h; 0 = Native Fab, +1 acyl group observed mass 47,718 (calculated 47,718), +2 acyl groups observed mass 47,798 (calculated 47,798).
A size-exclusion chromatogram (SEC) of the native Fab and the Fab after acyl transfer was also obtained to determine whether incubation at pH 8.4 for 72 hours had resulted in any aggregation (Figure 75). The native Fab was shown to be of 97% purity; the other 3% seems to be due to residual F(ab’)2 species present after digestion with papain, during Fab preparation. A 94% purity was observed for the Fab fragment after acyl transfer, suggesting a maximum of 3% aggregation.

![SEC Chromatograms](image)

Figure 75 – SEC chromatograms; a. Native Fab, b. Modified Fab after acyl transfer at pH 8.4, 12 °C, 72 h.

The antigen-binding activity of Fab after acyl transfer was also evaluated, in comparison to the native antibody fragment. ELISA analysis showed that binding was comparable before and after modification (Figure 76).

![ELISA Analysis](image)

Figure 76 – ELISA analysis of native Fab and Fab conjugate after acyl transfer, against HER2.
2.5.5 Functionalisation of the Fab fragment after modification

Having demonstrated that stability and antigen-binding after acyl transfer has been maintained, the Fab fragment was reacted with Alexa Fluor 488 azide, under CuAAC conditions (Figure 77a). Visualization of the protein with a UV-transilluminator after CuAAC suggested the reaction of Alexa fluorophore with the Fab fragment (Figure 77b, lane 8).

**Figure 77** – CuAAC of Fab fragment with Alexa Fluor azide; a. Reaction scheme; Reagents and conditions: (i) Alexa Fluor 488 azide (20 eq.), CuSO₄ (200 μM), THPTA (1 mM), sodium ascorbate (10 mM), pH 8.0, 37 ºC, 4 h, b. SDS-PAGE analysis of transthioesterification, acyl transfer (pH 8.4, 12 ºC) and CuAAC: M. Molecular marker, 1. Native Fab, 2. Reduced Fab, 3. Fab after transthioesterification, 4. Acyl transfer (24 h), 5. Acyl transfer (48 h), 6. Acyl transfer (72 h), 7. Fab after acyl transfer + Alexa Fluor 488, 8. Fab after acyl transfer + Alexa Fluor 488 (picture taken on a UV-transilluminator).

The fluorophore-to-antibody ratio (FAR) was then determined photometrically by obtaining the UV/Vis absorption spectrum of the conjugate (Figure 78) and using the following formula, where \( C_f \) is the correction factor for the absorbance of Alexa Fluor at 280 nm:

\[
FAR = \frac{\frac{Abs_{495}}{\varepsilon_{495}}}{\frac{Abs_{280} - (C_f \cdot Abs_{495})}{\varepsilon_{280}}} = \frac{0.2688}{71,000} = 1.4
\]
Figure 78 – UV/Vis absorbance of Fab conjugate after CuAAC with Alexa Fluor 488.

An FAR value of 1.4 was obtained, matching the loading calculated from the LC-MS spectrum, prior to the click reaction. Overall, the data obtained though these experiments seem to be in accordance with the initial hypothesis of modifying two lysine residues via cysteine-to-lysine acyl transfer. More specifically, the alkyl MESNa thioester was shown to react selectively with the two cysteine residues of Fab fragment; in the absence of reducing agent, no reaction was observed. This product was cleavable upon thiol addition, consistent with the reversible nature of thioesters. Increasing the pH of the mixture was shown to result in transfer of the two acyl groups, followed by reformation of the disulfide bond. This suggests reaction of lysine residues, which at physiological pH are predominantly protonated. In addition, the resultant conjugate was robustly stable upon treatment with excess BME, which is indicative of amide bond formation. Furthermore, the Fab crystal structure (PDB file 1HZH) reveals the presence of only two lysine residues in proximity to the disulfide bond. Even if competing thioester hydrolysis is witnessed, two residues are being modified. If the acyl transfer was unselective, a statistical distribution of products would be observed. An enzymatic digestion will need to be carried out in the future to determine the exact points of attachment. If selectivity is confirmed, it is anticipated that the CLT strategy can give access to a novel bioconjugation approach of modifying lysine residues that are far from the antigen-binding site, in a site-specific manner.
2.6 Summary

Inspired by the concept of native chemical ligation (NCL), the idea of developing a new bioconjugation methodology that selectively modifies specific lysine residues, via acyl transfer from the disulfide bond of Fab antibody fragment, was examined (Scheme 54). This approach involved an initial transthioesterification reaction of the reduced disulfide bond with a small molecule thioester. While a highly reactive aryl thioester was shown to result in unspecific labeling, selectivity was achieved by employing a less reactive alkyl MESNa thioester. Thiol-thioester exchange was shown to successfully proceed to completion after 4 hours, with 100 equivalents of reagent at 22 °C and at 150 μM.

Scheme 54 – Cysteine-to-lysine transfer (CLT); Reagents and conditions: (i) TCEP (10 eq.), 37 °C, 1 h, (ii) thioester 59 (100 eq.), 150 μM, 22 °C, 4 h, (iii) pH 8.4, 12 °C, 72 h, (iv) Alexa Fluor 488 azide (20 eq.), CuSO4 (200 μM), THPTA (1 mM), sodium ascorbate (10 mM), pH 8.0, 37 °C, 4 h.

Following transthioesterification, the effect of pH, temperature and time on the rate of acyl transfer was studied. Overall, the reaction was shown to be slow. Increasing the buffer pH or temperature resulted in faster reaction, but thioester hydrolysis was identified as the main side-reaction, prohibiting quantitative transfer of both acyl groups. Pleasingly, a good balance between increased
transfer and reduced hydrolysis was obtained upon incubation at pH 8.4, 12 °C for 72 hours. Under these conditions, a loading of 1.4 was calculated from LC-MS that corresponds to ca. 70% labeling of the two lysine residues. Disulfide bond re-oxidation was observed to occur in situ after transfer, representing a traceless modification. The resultant conjugate was shown to remain intact after incubation with excess BME and antigen-binding activity was maintained. Finally, a CuAAC reaction with Alexa Fluor 488 azide proceeded successfully with a fluorophore-to-antibody ratio (FAR) of 1.4. Overall, it is anticipated that after gaining confirmation of the exact lysine residues that have reacted, this approach can become an exciting chemical tool for the production of next generation antibody conjugates, incorporating biologically robust amide bonds.

2.7 Future work

While these preliminary experiments have served as a proof-of-concept of the CLT approach, optimization of each step in the sequence should be carried out in the future with the view to developing a highly efficient strategy. More specifically, by modulating the reactivity of the thioester employed so that it has better balance between good leaving group ability and reduced nucleophilicity, less equivalents of reagent will be required to achieve transthioesterification. This could involve the use of alkyl thiols of slightly lower pKₐ than MESNa, such as methyl thioglycolate 60 (pKₐ 8.08) or develop a thioester that quenches itself upon release, such as 61 (Figure 79). In addition, work by Durek et al. has shown that in NCL at proline, a preformed selenoester afforded a rate enhancement of 350 times, compared to standard NCL, due to the superior leaving group ability of selenolates over structurally similar thiolates. As a result, a selenoester, such as 62 could be synthesized to examine the rate of thiol-selenoester exchange.

![Figure 79 – Examples of thio- or selenoesters that could be used to enhance the rate of initial transthioesterification; X = leaving group.](image-url)
Further to this, Dawson and coworkers have observed NCL at glycine to give the fastest ligation rate.\textsuperscript{160} Therefore, employing a more electrophilic thioester, such as the glycine-derived \textit{63} could increase the rate of acyl transfer (Figure 80). In addition, preliminary experiments have shown that addition of DTNB in the reaction mixture prior to transfer can reduce hydrolysis to a small extent. It is likely that after one acyl group has been transferred, the cysteine can react with DTNB to form a disulfide that contains a nitrobenzoic acid, which might repel hydroxide anions from attacking the other thioester. Along with this idea being further explored, the use of thioesters containing negatively charged groups, such as a carboxylate \textit{64} or sulfonate \textit{65} could be considered with the view to stabilizing the thioesters towards hydrolysis.

\textbf{Figure 80} – Examples of thioesters that could be used to enhance the rate of acyl transfer or stabilize the product towards hydrolysis.

A linear alkyne can be used for initial proof-of-concept experiments, but the incorporation of an azide or strained alkyne handle into the thioester motif would enable subsequent functionalization \textit{via} SPAAC. Furthermore, functional re-bridging of the disulfide bond after CLT could be attempted with the view to introducing a second modification in a site-specific manner.

Having successfully optimized CLT conditions on the Fab fragment that contains a single disulfide bond, modification of a full antibody should be attempted to determine the virtue of this approach in a more complex system. Finally, the creation of antibody-drug conjugates or radioimmunoconjugates can demonstrate the potential application of this platform in therapeutics and diagnostics.
3. Experimental section

3.1 Synthesis general remarks

All chemical reagents and solvents were purchased from Sigma, AlfaAesar, Santa Cruz Biotechnology or VWR and used as received, without any further purification. N-methyl diphenoxymaleimide 28 was kindly provided by Dr Christina Marculescu (Baker group, UCL). Dithiophenolmaleimide NHS ester 45 was kindly provided by Dr Elizabeth Hull (Baker group, UCL).

All petroleum ether used had a boiling point range of 40 to 60 °C. All reaction mixtures were stirred magnetically and reactions were carried out at atmospheric pressure, under argon. Room temperature is defined as between 15-25 °C. Reactions performed at 0 °C were cooled with an ice and water bath. The term in vacuo refers to solvent removal using Büchi rotary evaporation between 15-60 °C, at approximately 10 mm Hg.

Reactions were monitored by TLC, using TLC plates pre-coated with silica gel 60 F254 on aluminium (Merck KGaA). Detection was by UV (254 nm and 365 nm) or chemical stain (KMnO₄, ninhydrin, iodine). Column chromatography was carried out either manually using normal phase silica gel, 40-63 μm (BDH) and sand (VWR) or automatically using a Biotage Isolera with GraceResolv™ silica flash cartridges.

¹H and ¹³C NMR spectra were recorded at ambient temperature on a Bruker Advance AMX600 instrument, operating at 600 MHz for ¹H and at 150 MHz for ¹³C in the stated solvent, using CDCl₃ (δ = 7.26), CD₃OD (δ = 3.31) or DMSO (δ = 2.50) as the internal standard. Chemical shifts (δ) are reported in parts per million (ppm) and coupling constants (J) in Hertz (Hz). The multiplicity of each signal is indicated as s-singlet, d-doublet, t-triplet, q-quartet, quin-quintet, m-multiplet (i.e. complex peak obtained due to overlap), app-apparent or a combination of these. All assignments were made with the aid of DEPT, COSY, HSQC, HMBC or NOESY correlation experiments. Infra-red spectra were recorded on a Bruker ALPHA FT-IR spectrometer operating in ATR mode, with frequencies given in reciprocal centimeters (cm⁻¹). The absorptions are characterized as s (sharp), br (broad), m (medium), w (weak).

Melting points were taken on a Gallenkamp apparatus and are uncorrected.
High and low resolution mass spectra were recorded on a VG70 SE mass spectrometer, operating in modes ESI, EI or CI (+ or -) depending on the sample, at the Department of Chemistry, University College London or obtained by the EPSRC UK National Mass Spectrometry Facility (NMSF), Swansea.

Absorbance measurements were carried out on a Carry Bio 100 UV/Vis spectrophotometer (Varian) equipped with a temperature-controlled 12x sample holder in quartz cuvettes (Starna Scientific - path length 1 cm, volume 160 μL) at RT. Samples were baseline corrected. The UV/Vis data was analysed using Graphpad Prism 7.03 software.

3.1.1 Synthesis and characterization of compounds

**Methyl 3,4-dibromo-2,5-dioxo-2,5-dihydro-1H-pyrrole-1-carboxylate (15)**

![Structure](image)

To a solution of 3,4-dibromomaleimide (1.0 g, 3.9 mmol) and N-methyl morpholine (0.43 mL, 3.9 mmol) in THF (35 mL) was added methyl chloroformate (0.30 mL, 3.9 mmol). The resultant pink mixture was stirred at RT for 20 min. The solvent was then removed *in vacuo* and CH$_2$Cl$_2$ (50 mL) was added. The organic phase was washed with H$_2$O (2 × 40 mL), brine (30 mL), dried (MgSO$_4$) and concentrated *in vacuo* to afford the target compound as a purple solid (1.16 g, 3.72 mmol, 94%).

mp 115-117 °C (lit. mp 115-118 °C$^{335}$; $^1$H NMR (600 MHz, CDCl$_3$) δ$_H$ 4.00 (3H, s, OCH$_3$); $^{13}$C NMR (150 MHz, CDCl$_3$) δ$_C$ 159.4 (C(O)N), 147.1 (C(O)O), 131.6 (C), 55.0 (OCH$_3$); IR (solid) ν$_{max}$/ cm$^{-1}$ 2961 (w), 1807 (s), 1765 (s), 1724 (s), 1599 (s); LRMS (EI) m/z 315 ($^{81.81}$M$^+$, 49), 313 ($^{81.79}$M$^+$, 98), 311 ($^{79.79}$M$^+$, 49), 284 (56), 282 (57), 266 (100), 264 (99), 190 (62), 188 (59), 133 (47), 131 (51).
To a stirred solution of \(N\)-(methoxycarbonyl)-3,4-dibromomaleimide 15 (88 mg, 0.28 mmol) in \(\text{CH}_2\text{Cl}_2\) (4 mL) was added \(O,O'\)-Bis(2-aminoethyl)octadecaethylene glycol (100 mg, 0.111 mmol). After 4 h at RT, silica (ca. 100 mg) was added and the resultant mixture was stirred at RT for a further 20 h. The solvent was then removed \textit{in vacuo}. Purification by column chromatography (gradient elution from \(\text{CH}_2\text{Cl}_2\) to 7\% MeOH in \(\text{CH}_2\text{Cl}_2\) afforded the target compound as a pale yellow oil (89 mg, 0.065 mmol, 59\%).

\[^1\text{H}\] NMR (600 MHz, \(\text{CDCl}_3\)) \(\delta_H\) 3.77 (t, 4H, \(J = 5.7, 2 \times \text{NCH}_2\)), 3.63-3.55 (m, 76H, 38 \(\times \text{CH}_2\)); \[^{13}\text{C}\] NMR (150 MHz, \(\text{CDCl}_3\)) \(\delta_C\) 163.8 (C(O)N), 129.4 (C), 70.5 (CH\(_2\)), 70.0 (CH\(_2\)), 67.5 (CH\(_2\)), 38.9 (NCH\(_2\)); IR (oil) \(\nu_{\text{max}}\) cm\(^{-1}\) 2883 (s), 1784 (w), 1716 (s), 1099 (s); LRMS (ES+) \(m/z\) 1394 ([\(^{81,81,81}\text{M+NH}_4^+\], 20), 1392 ([\(^{81,81,79}\text{M+NH}_4^+\], 72), 1390 ([\(^{81,79,79}\text{M+NH}_4^+\], 100), 1388 ([\(^{79,79,79}\text{M+NH}_4^+\], 63), 1386 ([\(^{79,79,79}\text{M+NH}_4^+\], 15), 1346 (83), 1300 (30); HRMS (ES+) calcd for [\(\text{C}_{48}\text{H}_{80}\text{Br}_4\text{N}_2\text{O}_{22}\text{NH}_4^+\])\(^+\) [\(^{81,81,79}\text{M+NH}_4^+\)] 1390.2182, observed 1390.2175.
3-Bromo-1-methyl-1H-pyrrole-2,5-dione (18)

To N-methylmaleimide (976 mg, 8.78 mmol) in CHCl₃ (50 mL) was added Br₂ (0.990 mL, 19.3 mmol) and the resultant mixture was heated to reflux for 2 h. The solvent was then removed in vacuo. The residue was dissolved in EtOAc (40 mL) and washed with H₂O (20 mL), 15% aq. Na₂S₂O₃ (20 mL), dried (MgSO₄) and concentrated. The crude orange solid was dissolved in AcOH (50 mL). NaOAc (2.16 g, 26.4 mmol) was added and then resultant mixture was heated to reflux for 3 h. The solvent was then removed in vacuo, azeotroping with toluene (3 × 50 mL). EtOAc (50 mL) was added and the organic phase was washed with H₂O (30 mL), sat. aq. NaHCO₃ (40 mL), dried (MgSO₄) and concentrated. Purification by column chromatography (20% EtOAc in pet. ether) afforded the target product as a white crystalline solid (1.20 g, 6.30 mmol, 72%).

mp 83-85 °C (lit. mp 77-79 °C); ¹H NMR (600 MHz, CDCl₃) δH 6.89 (s, 1H, CH), 3.08 (s, 3H, NCH₃); ¹³C NMR (150 MHz, CDCl₃) δC 168.7 (C(O)N), 165.5 (C(O)N), 132.0 (C), 131.5 (C), 24.8 (NCH₃); IR (solid) νmax/ cm⁻¹ 3106, (w), 2850 (w), 1776 (w), 1698 (s), 1587 (s), 606 (m); LRMS (EI) m/z 191 ([⁺][M]+, 73), 189 ([⁺][M]+, 71), 134 (46), 132 (45), 106 (31), 104 (30), 82 (29), 66 (24), 53 (100); HRMS (EI) calcd for [C₅H₄O₂N⁺Br][M]+ 188.9420, observed 188.9419.

3,4-Dichloro-1-methyl-1H-pyrrole-2,5-dione (26)

To a solution of dichloromaleic anhydride (500 mg, 2.99 mmol) in AcOH (20 mL) was added methylamine.HCl (222 mg, 3.29 mmol). The resultant mixture was heated to reflux for 16 h. The solvent was then removed in vacuo, azeotroping with toluene (3 × 40 mL). Purification by column
chromatography (gradient elution from 10% EtOAc in pet. ether to 15% EtOAc in pet. ether) afforded the target compound as a white solid (429 mg, 2.38 mmol, 80%).

mp 80-82 (lit. mp 82-83)\(^{33}\); \(^1\)H NMR (600 MHz, CDCl\(_3\)) \(\delta\)H 3.12 (s, 3H, NCH\(_3\)) ; \(^{13}\)C NMR (150 MHz, CDCl\(_3\)) \(\delta\)C 163.2 (C(O)N), 133.5 (C), 25.5 (NCH\(_3\)); IR (solid) \(\nu_{\text{max}}/\text{cm}^{-1}\) 2948 (w), 1791 (w), 1708 (m), 1654 (m); LRMS (EI) \(m/z\) 183 ([\(^{37,37}\)M]+, 5), 181 ([\(^{37,35}\)M]+, 35), 179 ([\(^{35,35}\)M]+, 54), 87 (100); HRMS (EI) calcd for [C\(_5\)H\(_3\)I\(_2\)NO\(_2\)]\(^+\) \([^{35,35}\text{M}]^+\) 178.9535, observed 178.9534.

3,4-Diiodo-1-methyl-1\textit{H}-pyrrole-2,5-dione (27)

\[
\begin{array}{c}
\text{O} \\
\text{N} \\
\text{O} \\
\end{array}
\]

A mixture of 2,3-dibromo-\(N\)-methylmaleimide (150 mg, 0.558 mmol) and NaI (418 mg, 2.79 mmol) in AcOH (5 mL) was heated to reflux for 2 h. The reaction mixture was then allowed to cool to RT and H\(_2\)O (5 mL) was added. The yellow precipitate was filtered off, washed with H\(_2\)O (10 mL) and dried to afford the target compound as a yellow crystalline solid (142 mg, 0.391 mmol, 70%).

mp 140-142 °C; \(^1\)H NMR (600 MHz, CDCl\(_3\)) \(\delta\)H 3.16 (s, 3H, NCH\(_3\)); \(^{13}\)C NMR (150 MHz, CDCl\(_3\)) \(\delta\)C 166.6 (C(O)N), 117.3 (C), 26.2 (NCH\(_3\)); IR (solid) \(\nu_{\text{max}}/\text{cm}^{-1}\) 2936 (w), 1764 (m), 1696 (s), 1550 (m); LRMS (Cl+) \(m/z\) 381 ([M+NH\(_4\)]\(^+\), 100), 96 (22); HRMS (Cl+) calcd for [C\(_5\)H\(_3\)I\(_2\)NO\(_2\)NH\(_4\)]\(^+\) [M+NH\(_4\)]\(^+\) 380.85914, observed 380.85908.
1-Methyl-3,4-bis(phenylthio)-1\textit{H}-pyrrole-2,5-dione (29)\textsuperscript{337}

\[
\begin{tikzpicture}
  \pgfuseimage{image}
\end{tikzpicture}
\]

To a solution of 2,3-dibromo-N-methylmaleimide (230 mg, 0.866 mmol) and NaOAc (162 mg, 1.97 mmol) in MeOH (14 mL) was slowly added thiophenol (202 μL, 1.97 mmol) in MeOH (4 mL) resulting in an immediate orange colour change. After 30 min at RT, the solvent was removed \textit{in vacuo}. CH\textsubscript{2}Cl\textsubscript{2} (30 mL) was added and the organic phase washed with sat. aq. NaHCO\textsubscript{3} (40 mL), H\textsubscript{2}O (40 mL), dried (MgSO\textsubscript{4}) and concentrated. Purification by column chromatography (4\% EtOAc in pet. ether) afforded the target compound as a yellow solid (248 mg, 0.757 mmol, 89%).

mp 94-96 °C (lit. mp 102 °C)\textsuperscript{337}; \textsuperscript{1}H NMR (600 MHz, CDCl\textsubscript{3}) \( \delta \)\textsubscript{H} 7.29-7.28 (m, 2H, 2 \( \times \) Ar\textsubscript{C}H), 7.26-7.24 (m, 4H, 4 \( \times \) Ar\textsubscript{C}H), 7.22-7.20 (m, 4H, 4 \( \times \) Ar\textsubscript{C}H), 3.01 (s, 3H, N\textsubscript{C}H\textsubscript{3}); \textsuperscript{13}C NMR (150 MHz, CDCl\textsubscript{3}) \( \delta \)\textsubscript{C} 167.0 (C(O)N), 136.0 (C), 132.0 (ArCH), 129.1 (ArC), 129.1 (ArCH), 128.5 (ArCH), 24.8 (NCH\textsubscript{3}); IR (solid) \( \nu \)\textsubscript{max} \textsubscript{cm}^{-1} 3058 (w), 2925 (w), 1773 (m), 1698 (s), 1574 (m), 1502 (m); LRMS (EI) \( m/z \) 327 ([M]\textsuperscript{+}, 100), 218 (17), 161 (13); HRMS (EI) calcd for [C\(_{17}\)H\(_{13}\)NO\(_2\)S\(_2\)]\textsuperscript{+} [M]\textsuperscript{+} 327.0382, observed 327.0379.

1,1'-(3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57-Nonadecaioxanonapentacontane-1,59-diyl)bis(3,4-bis(phenylthio)-1\textit{H}-pyrrole-2,5-dione) (DTPM\textsubscript{CL})

\[
\begin{tikzpicture}
  \pgfuseimage{image}
\end{tikzpicture}
\]

To DBM\textsubscript{CL} (35 mg, 0.025 mmol) and NaHCO\textsubscript{3} (22 mg, 0.26 mmol) in CH\textsubscript{2}Cl\textsubscript{2} (1.1 mL) was slowly added thiophenol (13 μL, 0.12 mmol) in MeOH (1.3 mL). After 2 h at RT, the solvent was removed \textit{in vacuo}. CH\textsubscript{2}Cl\textsubscript{2} (20 mL) was added and the organic layer was washed with H\textsubscript{2}O (30
mL) and brine (30 mL). The combined organic extracts were dried (MgSO\(_4\)) and concentrated \textit{in vacuo} to afford the target compound as a bright yellow oil (38 mg, 0.025 mmol, 99%).

\(^1\)H NMR (600 MHz, CDCl\(_3\)) \(\delta\) 7.28-7.23 (m, 12H, 12 \(\times\) ArCH\(_2\)), 7.19 (d, 8H, \(J = 7.2\, \text{Hz}\), 8 \(\times\) ArCH\(_2\)), 3.70 (t, 4H, \(J = 5.7\, \text{Hz}\), 2 \(\times\) NCH\(_2\)), 3.64-3.58 (m, 76H, 38 \(\times\) CH\(_2\)); \(^{13}\)C NMR (150 MHz, CDCl\(_3\)) \(\delta\) C 166.9 (C(O)N), 135.7 (C), 132.0 (ArCH), 129.1 (ArCH), 129.1 (ArC), 128.5 (ArCH), 70.7 (CH\(_2\)), 69.9 (CH\(_2\)), 67.7 (CH\(_2\)), 38.0 (NCH\(_2\)); IR (oil) \(\nu_{\max}/\text{cm}^{-1}\) 2864 (s), 1710 (s), 1650 (w), 1580 (w), 1106 (s); LRMS (ES+) \(m/z\) 1507 ([M+NH\(_4\)]\(^+\), 100), 762 (62); HRMS (ES+) calcd for [C\(_{72}\)H\(_{100}\)N\(_{2}\)O\(_{23}\)S\(_4\)NH\(_4\)]\(^+\) [M+NH\(_4\)]\(^+\) 1506.5938, observed 1506.5925.

2,2',2'',2''''-(((3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57-Nonadecaoxanonapentacontane-1,59-diyl)bis(2,5-dioxo-2,5-dihydro-1H-pyrrole-1,3,4-triyl))tetrakis(sulfanediyl)) tetrabenzoic acid (34)

To a mixture of DBM\(_{\text{Cl}}\) (25 mg, 0.018 mmol) and KOAc (18 mg, 0.18 mmol) in CH\(_2\)Cl\(_2\) (2 mL) was slowly added thiosalicylic acid (12 mg, 0.080 mmol) in MeOH (1.5 mL), resulting in an immediate orange colour change. After 4 h at RT, the solvent was removed \textit{in vacuo}. EtOAc (20 mL) was added and the organic phase washed with 10% aq. citric acid (2 \(\times\) 10 mL). The aqueous layer was extracted with EtOAc (2 \(\times\) 30 mL) and the combined organic layers washed with H\(_2\)O (30 mL), brine (30 mL) and dried (MgSO\(_4\)). Purification by column chromatography (8% MeOH/1% AcOH in CH\(_2\)Cl\(_2\)) afforded the target compound as a bright yellow oil (15 mg, 0.0090 mmol, 50%).
\(^1\)H NMR (600 MHz, CDCl\(_3\)) \(\delta\) 7.92 (d, 4H, \(J = 7.2\), 4 × ArCH), 7.38 (t, 4H, \(J = 7.2\), 4 × ArCH), 7.30 (t, 4H, \(J = 7.5\), 4 × ArCH), 7.21 (d, 4H, \(J = 7.9\), 4 × ArCH), 3.73 (t, 4H, \(J = 6.0\), 2 × NC\(_2\)H), 3.67-3.61 (m, 76H, 38 × CH\(_2\)); \(^{13}\)C NMR (150 MHz, CDCl\(_3\)) \(\delta\) C 168.8 (C(O)OH), 166.8 (C(O)N), 166.7 (ArC), 137.1 (C), 132.2 (ArCH), 132.1 (ArCH), 131.8 (ArCH), 130.7 (ArC), 127.7 (ArCH), 70.5 (CH\(_2\)), 70.0 (CH\(_2\)), 67.6 (CH\(_2\)), 38.2 (NCH\(_2\)); IR (oil) \(\nu_{\text{max}}/ \text{cm}^{-1}\) 3471 (w, br), 2870 (s), 1708 (s), 1586 (w), 1464 (m), 1093 (s); LRMS (ES+) \(m/z\) 850 ([M+2(NH\(_4\))]\(^{2+}\), 100), 649 (12); HRMS (ES+) calcd for \([C_{76}H_{100}N_2O_5S_4(NH_4)_2]^{2+}\) 850.2935, observed 850.2938.

6-(3,4-Dibromo-2,5-dioxo-2,5-dihydro-1H-pyrro-1-yl)hexanoic acid (36)

![Image of the molecule](image)

A mixture of dibromomaleic acid (1.0 g, 3.7 mmol) and 6-aminohexanoic acid (430 mg, 3.28 mmol) in AcOH (40 mL) was heated to reflux for 14 h. The solvent was then removed \textit{in vacuo}, azeotroping with toluene (3 × 50 mL). Purification by column chromatography (50% EtOAc in pet. ether) afforded the target compound as a white crystalline solid (1.10 g, 2.98 mmol, 91%).

mp 126-127 °C (lit. mp 123-124 °C); \(^1\)H NMR (600 MHz, CDCl\(_3\)) \(\delta\) H 11.47 (br s, 1H, C(O)O), 3.62 (t, 2H, \(J = 7.2\), NCH\(_2\)), 2.36 (t, 2H, \(J = 7.4\), CH\(_2\)C(O)OH), 1.66 (overlapped quin, 4H, \(J = 7.4\), NCH\(_2\)CH\(_2\) and CH\(_2\)CH\(_2\)C(O)OH), 1.36 (quin, 2H, \(J = 7.6\), NCH\(_2\)CH\(_2\)CH\(_2\)); \(^{13}\)C NMR (150 MHz, CDCl\(_3\)) \(\delta\) C 179.6 (C(O)OH), 164.1 (C(O)N), 129.5 (C), 39.6 (NCH\(_2\)), 33.8 (CH\(_2\)C(O)OH), 28.2 (NCH\(_2\)CH\(_2\)), 26.1 (NCH\(_2\)CH\(_2\)CH\(_2\)), 24.1 (CH\(_2\)CH\(_2\)C(O)OH); IR (solid) \(\nu_{\text{max}}/ \text{cm}^{-1}\) 2937 (w), 2869 (w), 1720 (s), 1692 (s), 1587 (s); LRMS (ES-) \(m/z\) 369 ([\(^{81,81}\)M-2H]\(^+\), 38), 367 ([\(^{81,79}\)M-2H]\(^+\), 60), 365 ([\(^{79,79}\)M-2H]\(^+\), 39), 305 (45), 303 (45), 287 (98), 285 (100); HRMS (ES-) calcd for \([C_{10}H_9^{79,79}Br_2NO_4]^+\) [\(^{79,79}\)M-2H]\(^+\) 364.8898, observed 364.8899.
6-(3,4-Bis((4-aminophenyl)thio)-2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)hexanoic acid (37)

To a mixture of bis-dibromomaleimide hexanoic acid 36 (180 mg, 0.488 mmol) and NaOAc (240 mg, 2.93 mmol) in MeOH (8 mL) was added dropwise 4-aminothiophenol (147 mg, 1.17 mmol) in MeOH (2 mL). The resultant dark red mixture was stirred at RT for 1 h. The solvent was then removed in vacuo and purification by column chromatography (gradient elution from 2% MeOH/0.5% AcOH in CH₂Cl₂ to 4% MeOH/0.5% AcOH in CH₂Cl₂) afforded the target product as a red oil (118 mg, 0.258 mmol, 53%).

¹H NMR (600 MHz, CD₃OD) δH 7.07 (d, 4H, J = 8.8, 4 × ArCH), 6.61 (d, 4H, J = 8.8, 4 × ArCH), 3.40 (t, 2H, J = 7.1, NCH₂), 2.18 (t, 2H, J = 7.5, CH₂C(O)OH), 1.57 (app quin, 2H, J = 7.6, CH₂CH₂C(O)OH), 1.51 (quin, 2H, J = 7.5, NCH₂CH₂), 1.24 (quin, 2H, J = 7.9, NCH₂CH₂CH₂); ¹³C NMR (150 MHz, CD₃OD) δc 180.4 (C(O)OH), 168.6 (C(O)N), 150.4 (ArC), 137.4 (C), 135.8 (ArCH), 116.5 (ArCH), 116.4 (ArC), 39.3 (NCH₂), 37.0 (CH₂C(O)OH), 29.3 (NCH₂CH₂), 27.6 (NCH₂CH₂CH₂), 26.4 (CH₂CH₂C(O)OH); IR (oil) v_max/cm⁻¹ 3457 (w, br), 3226 (w, br), 3370 (w, br), 2934 (w), 2861 (w), 1695 (s), 1622 (s), 1595 (s), 1494 (s); LRMS (ES+) m/z 458 ([M+H]⁺, 100), 250 (8); HRMS (ES+) calcd for [C₂₂H₂₃N₃O₄S₂H]⁺ [M+H]⁺ 458.1208, observed 458.1227.

3,4-Bis(phenylthio)-1H-pyrrole-2,5-dione (39)

To 3,4-dibromomaleimide (302 mg, 1.18 mmol) and NaHCO₃ (604 mg, 7.19 mmol) in MeOH (20 mL) was slowly added thiophenol (304 μL, 2.96 mmol) in MeOH (2 mL). The resulting bright yellow solution was stirred at RT for 40 min. The solvent was then removed in vacuo. CH₂Cl₂ (30 mL) was added and the organic layer washed with H₂O (20 mL), brine (20 mL), dried (MgSO₄) and concentrated. The crude orange oil was purified by column chromatography (gradient elution
from 10% EtOAc in pet. ether to 30% EtOAc in pet. ether) to afford the target compound as a yellow solid (243 mg, 0.775 mmol, 66%).

mp 117-120 ℃ (lit. mp 102-104 ℃)\textsuperscript{335}; \textsuperscript{1}H NMR (600 MHz, CD\textsubscript{3}OD) \(\delta\) \textsubscript{H} 7.29-7.23 (m, 6H, 6 × ArCH), 7.15 (d, 4H, \(J = 7.5, 4 \times ArCH\)); \textsuperscript{13}C NMR (150 MHz, CD\textsubscript{3}OD) \(\delta\) \textsubscript{C} 169.3 (C(O)NH), 137.5 (C), 132.4 (ArCH), 130.7 (ArC), 130.1 (ArCH), 129.1 (ArCH); IR (solid) \(\nu_{\text{max}}/\text{cm}^{-1}\) 3228 (m, br), 1772 (m), 1688 (s), 1580 (m), 1519 (w); LRMS (ES\textsuperscript{+}) \(m/z\) 336 ([M+Na]\textsuperscript{+}, 38), 314 ([M+H]\textsuperscript{+}, 100), 179 (10); HRMS (ES\textsuperscript{+}) calcd for \([\text{C}_{16}\text{H}_{11}\text{NO}_2\text{S}_2\text{H}]^+\) [M+H]\textsuperscript{+} 314.0309, observed 314.0300.

**Methyl 2,5-dioxo-3,4-bis(phenylthio)-2,5-dihydro-1H-pyrrole-1-carboxylate (40)**\textsuperscript{335}

![](image)

To a solution of dithiophenolmaleimide 39 (100 mg, 0.319 mmol) and N-methyl morpholine (71 \(\mu\text{L}, 0.65 \text{ mmol}) in EtOAc (4 mL) was added methyl chloroformate (62 \(\mu\text{L}, 0.80 \text{ mmol). The resultant yellow mixture was stirred at RT for 1 h. The reaction mixture was then washed with H}_2\text{O} (4 × 20 mL), dried (MgSO\textsubscript{4}) and concentrated \textit{in vacuo} to afford the target compound as an orange crystalline solid (110 mg, 0.296 mmol, 92%).

mp 110-113 ℃ (lit. mp 110-112)\textsuperscript{335}; \textsuperscript{1}H NMR (600 MHz, CDCl\textsubscript{3}) \(\delta\) \textsubscript{H} 7.34-7.26 (m, 10H, 10 × ArCH), 3.92 (s, 3H, OCH\textsubscript{3}); \textsuperscript{13}C NMR (150 MHz, CDCl\textsubscript{3}) \(\delta\) \textsubscript{C} 161.9 (C(O)N), 147.9 (C(O)O), 137.2 (C), 132.7 (ArCH), 129.3 (ArCH), 129.1 (ArC), 128.2 (ArCH), 54.5 (OCH\textsubscript{3}); IR (solid) \(\nu_{\text{max}}/\text{cm}^{-1}\) 2961 (w), 1802 (s), 1760 (s), 1713 (s), 1576 (m); LRMS (ES\textsuperscript{+}) \(m/z\) 394 ([M+Na]\textsuperscript{+}, 30), 372 ([M+H]\textsuperscript{+}, 100), 301 (28); HRMS (ES\textsuperscript{+}) calcd for \([\text{C}_{18}\text{H}_{13}\text{NO}_4\text{S}_2\text{H}]^+\) [M+H]\textsuperscript{+} 372.0370, observed 372.0364.
Bis-dithiophenolmaleimide PEG\textsubscript{6000} (41)

![Chemical structure](image.png)

To a stirred solution of 40 (8 mg, 0.02 mmol) in CH\textsubscript{2}Cl\textsubscript{2} (2 mL) was added \textit{O,O′}-Bis(2-aminoethyl)octadecaethylene glycol (50 mg, 0.0083 mmol) and \textit{p}-TsOH (\textit{ca.} 2 mg). The resultant mixture was stirred at RT for 16 h. CH\textsubscript{2}Cl\textsubscript{2} (20 mL) was then added and the organic phase washed with H\textsubscript{2}O (30 mL), sat. aq. NaHCO\textsubscript{3} (30 mL) and 10\% aq. citric acid (50 mL). The combined organic extracts were dried (MgSO\textsubscript{4}) and concentrated \textit{in vacuo}. Acetone (2.5 mL) was added and the resultant solution was placed in an ice bath for 20 min, during which time a precipitate began to form. The suspension was centrifuged at 4000 \times g for 40 min at 4°C. The supernatant was removed, acetone (2 mL) was added and the mixture centrifuged for further 40 min at 4°C. The supernatant was removed and the precipitate was dried to afford the target product as a yellow solid (42 mg, 0.00035 mmol, 76\%).

mp 49-52 °C; \textsuperscript{1}H NMR (600 MHz, CDCl\textsubscript{3}) \textit{δ}\textsubscript{H} 7.26 (d, xH, \textit{J} = 7.2, ArCH), 7.23 (t, xH, \textit{J} = 7.5, ArCH), 7.18 (d, xH, \textit{J} = 7.2, ArCH), 3.75 (t, xH, \textit{J} = 4.5, CH\textsubscript{2}), 3.69 (t, xH, \textit{J} = 5.7, CH\textsubscript{2}), 3.64-3.63 (m, xH, CH\textsubscript{2}), 3.52 (t, xH, \textit{J} = 4.9, CH\textsubscript{2}); \textsuperscript{13}C NMR (150 MHz, CDCl\textsubscript{3}) \textit{δ}\textsubscript{C} 166.9 (C(O)N), 135.7 (C), 131.9 (ArCH), 129.1 (ArC), 129.0 (ArCH), 128.5 (ArCH), 70.7 (CH\textsubscript{2}), 69.9 (CH\textsubscript{2}), 67.7 (CH\textsubscript{2}), 38.0 (CH\textsubscript{2}), 30.0 (CH\textsubscript{2}); IR (solid) \textit{ν}_{\text{max}} \text{ cm}^{-1} 2880 (s), 1710 (w), 1465 (s), 1099 (s); LRMS (MALDI-TOF ultrafleXtreme) \textit{M}_{w} = 6883, \textit{M}_{n} = 6624, PDI = 1.04; see Figure 81.

\* An exact integration could not be obtained, because the starting material PEG\textsubscript{6000} was a mixture of polymers with an average molecular weight of 6000.
Figure 81 – MALDI-TOF ultrafleXtreme spectrum of bis-dithiophenolmaleimide PEG$_{6000}$ 41, showing the polymer distribution.

Tert-butyl(59-amino-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57-nonadecaoxanona pentacontyl)carbamate (43)

![Structure](structure.png)

To a solution of $O,O′$-Bis(2-aminoethyl)octadecylene glycol (262 mg, 0.292 mmol) in CH$_2$Cl$_2$ (2 mL) was added dropwise a solution of Boc$_2$O (32 mg, 0.15 mmol) in CH$_2$Cl$_2$ (4 mL) over 4 h. The resulting mixture was stirred at RT for 16 h. The solvent was then removed in vacuo and the residue purified by column chromatography (6% MeOH/ 1% NEt$_3$ in CH$_2$Cl$_2$) to yield the target compound as a colourless oil (109 mg, 0.109 mmol, 75%).

$^1$H NMR (600 MHz, CDCl$_3$) $^\delta$H 5.37 (br s, 2H, NH$_2$), 5.06 (br s, 1H, NH), 3.73 (t, 2H, J = 4.9, NH$_2$CH$_2$CH$_2$), 3.64-3.53 (m, 72H, 36 × CH$_2$), 3.46 (t, 2H, J = 5.3, CH$_2$CH$_2$NHC(O)), 3.23 (q, 2H, J = 4.9, CH$_2$NHC(O)), 3.03 (t, 2H, J = 4.9, NH$_2$CH$_2$), 1.37 (s, 9H, C(CH$_3$)$_3$); $^{13}$C NMR (150 MHz, CDCl$_3$) $^\delta$C 156.0 (OC(O)NH), 79.0 (C(CH$_3$)$_3$), 70.5 (CH$_2$), 70.4 (CH$_2$), 70.3 (CH$_2$), 70.2 (CH$_2$), 70.0 (CH$_2$), 69.9 (CH$_2$), 40.6 (CH$_2$), 40.3 (CH$_2$), 28.4 (C(CH$_3$)$_3$); IR (oil) $\nu_{\text{max}}$ / cm$^{-1}$ 3367 (w, br), 2866 (s), 1707 (m), 1640 (m), 1521 (m), 1100 (s); LRMS (ES+) $m/z$ 998 ([M+H]$^+$, 100); HRMS (ES+) calcd for [C$_{45}$H$_{92}$N$_2$O$_{21}$H]$^+$ [M+H]$^+$ 997.6271, observed 997.6276.
Di-tert-butyl(61,103-dioxo-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57,64,67,70,73, 76,79,82,85,88,91,94,97,100,107,110,113,116,119,122,125,128,131,134,137,140,143,146,149, 152,15,158,161-henpentacontaoxa-60,104-diazatrihexacontahectane-1,163-diyl)dicarbamate (44)

To a mixture of \(O,O'\)-Bis(2-carboxyethyl)dodecaethylene glycol (18 mg, 0.026 mmol), EDC.HCl (11 mg, 0.057 mmol) and DMAP (0.65 mg, 0.0053 mmol) in DMF (2 mL) was added mono-Boc diamine 43 (55 mg, 0.055 mmol). The resultant mixture was stirred at RT for 40 h. DMF was then removed \textit{in vacuo}, azeotroping with toluene (4 \times 20 mL). The residue was dissolved in CH\(_2\)Cl\(_2\) (40 mL) and the organic phase washed with H\(_2\)O (20 mL), dried (MgSO\(_4\)) and concentrated \textit{in vacuo}. Purification by column chromatography (gradient elution from CH\(_2\)Cl\(_2\) to 10\% MeOH in CH\(_2\)Cl\(_2\)) afforded the target compound as a pale yellow oil (44 mg, 0.017 mmol, 64\%).

\(^1\)H NMR (600 MHz, CDCl\(_3\)) \(\delta\)H 3.75-3.73 (m, 4H, 2 \times CH\(_2\)), 3.76-3.58 (m, 196H, 98 \times CH\(_2\)), 3.55-3.52 (m, 8H, 4 \times CH\(_2\)), 3.50-3.43 (m, 4H, 2 \times CH\(_2\)), 3.33-3.29 (m, 4H, 2 \times CH\(_2\)), 1.44 (s, 18H, 2 \times C(CH\(_3\)_3)); \(^{13}\)C NMR (150 MHz, CDCl\(_3\)) \(\delta\)C 161.7 (C(O)NH), 156.1 (OC(O)NH), 79.3 (C(CH\(_3\)_3)), 72.7 (CH\(_2\)), 70.7 (CH\(_2\)), 69.8 (CH\(_2\)), 67.4 (CH\(_2\)), 61.8 (CH\(_2\)), 40.5 (CH\(_2\)), 39.3 (CH\(_2\)), 37.9 (CH\(_2\)), 29.8 (CH\(_2\)), 28.6 (C(CH\(_3\)_3)); IR (oil) \(v\text{max}/\text{cm}^{-1}\) 2866 (s), 1708 (s), 1524 (w), 1094 (s); LRMS (ES+) \(m/z\) 906 ([M+3Na]\(^{3+}\), 7), 510 (93), 449 (100); HRMS (ES+) calcd for [C\(_{120}\)H\(_{238}\)N\(_4\)O\(_{57}\)Na\(_3\)]\(^{3+}\) [M+3Na]\(^{3+}\) 905.8519, observed 905.8518.
N1,N43-bis(64-(2,5-dioxo-3,4-bis(phenylthio)-2,5-dihydro-1H-pyrrol-1-yl)-61-oxo-3,6,9,12,15,18, 21,24,27,30,33,36,39,42,45,48,51,54,57-nonadecaosa-60-azatetrahexacontyl)-4,7,10,13,16,19,22, 25,28,31,34,37,40-tridecaoxatritetracontanediamide (46)

To a solution of compound 44 (16 mg, 0.0060 mmol) in CH$_2$Cl$_2$ (1.3 mL) was added TFA (0.8 mL) and the reaction mixture was stirred at RT for 16 h. The solvent was then removed in vacuo, azeotroping with toluene (3 × 30 mL). The crude residue was dissolved in DMF (1.2 mL) and NEt$_3$ (4 μL, 0.03 mmol) was added. After 30 min at RT, dithiophenolmaleimide NHS ester 45 (7 mg, 0.01 mmol) in DMF (0.2 mL) was added. The resultant yellow solution was stirred at RT for 22 h. The solvent was then removed in vacuo. CH$_2$Cl$_2$ (40 mL) was added and the organic phase washed with H$_2$O (20 mL), sat. aq. NaHCO$_3$ (40 mL), 10% aq. citric acid (40 mL), dried (MgSO$_4$) and concentrated in vacuo to afford the target product as a yellow oil (15 mg, 0.0047 mmol, 79%).

$^1$H NMR (600 MHz, CDCl$_3$) $\delta$H 7.28 (m, 4H, 4 × ArCH), 7.25 (t, 8H, J = 7.9, 8 × ArCH), 7.19 (d, 8H, J = 7.2, 8 × ArCH), 3.78-3.72 (m, 6H, 3 × CH$_2$), 3.67-3.59 (m, 192H, 96 × CH$_2$), 3.56 (overlapped t, 4H, J = 6.4, 2 × NCH$_2$), 2.17 (t, 4H, J = 7.3, 2 × NCH$_2$CH$_2$CH$_2$), 1.92 (quin, 4H, J = 6.8, 2 × NCH$_2$CH$_2$); $^{13}$C NMR (150 MHz, CDCl$_3$) $\delta$C 172.0 (C(O)NH), 167.9 (C(O)N), 167.2 (NHC(O), 135.6 (C), 131.9 (ArCH), 129.1 (ArC), 129.0 (ArCH), 128.5 (ArC), 70.7 (CH$_2$), 70.3 (CH$_2$), 70.0 (CH$_2$), 39.4 (NCH$_2$), 38.4 (CH$_2$), 33.7 (NCH$_2$CH$_2$CH$_2$), 24.8 (NCH$_2$CH$_2$); IR (oil) $\nu_{max}$/ cm$^{-1}$ 3330 (w, br), 2868 (s), 1706 (s), 1659 (m), 1536 (m), 1096 (s); LRMS (ES+) $m/z$ 1329 (26), 830 (24), 804 ([M+4H]$^{4+}$, 7), 654 (29), 317 (100); HRMS (ES+) calcd for [C$_{150}$H$_{252}$N$_6$O$_{59}$S$_4$H$_3$]$^{3+}$ [M+3H]$^{3+}$ 1071.5352, observed 1071.5363.
3,4-Diiodo-1H-pyrrole-2,5-dione (49)\textsuperscript{338}

\[\text{\begin{tikzpicture}
  \draw [thick] (0,0) -- (0,1) -- (1,1) -- (1,0) -- (0,0);
  \draw [thick] (0.5,0) -- (0.5,1);
  \draw [thick] (0,0.5) -- (1,0.5);
  \draw [thick] (0,0.5) -- (0,0.5) -- (1,0.5);
  \draw [thick] (0,0) -- (0,0.5);
  \draw [thick] (1,0) -- (1,0.5);
  \draw [thick] (0,1) -- (0,0.5);
  \draw [thick] (1,1) -- (1,0.5);
  \draw [thick] (0,0) -- (0,0.5);\end{tikzpicture}}\]

A mixture of 3,4-dibromomaleimide (220 mg, 0.863 mmol) and NaI (388 mg, 2.59 mmol) in AcOH (5 mL) was heated to reflux for 2 h. The reaction mixture was then allowed to cool to RT and H\textsubscript{2}O (5 mL) was added. The yellow precipitate was filtered off, washed with water (10 mL) and dried to afford the target compound as a yellow crystalline solid (230 mg, 0.659 mmol, 76%). mp 242-245 °C (lit. mp 254-255 °C)\textsuperscript{338}; \textsuperscript{1}H NMR (600 MHz, CDCl\textsubscript{3}) \(\delta_H\) no signals; \textsuperscript{13}C NMR (150 MHz, CDCl\textsubscript{3}) \(\delta_C\) 165.7 (C(O)N), 118.4 (C); IR (solid) \(\nu_{\text{max}}/\text{cm}^{-1}\) 3235 (m, br), 1755 (m), 1706 (s), 1561 (m), 1543 (s); LRMS (EI) \(m/z\) 349 ([M]\textsuperscript{+}, 100), 179 (92); HRMS (EI) calcd for [C\textsubscript{4}HI\textsubscript{2}NO\textsubscript{2}]\textsuperscript{+} [M]\textsuperscript{+} 348.8091, observed 348.8090.

Methyl 3,4-diiodo-2,5-dioxo-2,5-dihydro-1H-pyrrole-1-carboxylate (50)

\[\text{\begin{tikzpicture}
  \draw [thick] (0,0) -- (0,1) -- (1,1) -- (1,0) -- (0,0);
  \draw [thick] (0.5,0) -- (0.5,1);
  \draw [thick] (0,0.5) -- (1,0.5);
  \draw [thick] (0,0.5) -- (0,0.5) -- (1,0.5);
  \draw [thick] (0,0) -- (0,0.5);
  \draw [thick] (1,0) -- (1,0.5);
  \draw [thick] (0,1) -- (0,0.5);
  \draw [thick] (1,1) -- (1,0.5);
  \draw [thick] (0,0) -- (0,0.5);\end{tikzpicture}}\]

To a solution of 3,4-diiodomaleimide 49 (200 mg, 0.573 mmol) and N-methyl morpholine (64.0 \(\mu\)L, 0.582 mmol) in THF (10 mL) was added methyl chloroformate (45.0 \(\mu\)L, 0.582 mmol). The resultant mixture was stirred at RT for 20 min. The solvent was then removed \textit{in vacuo} and CH\textsubscript{2}Cl\textsubscript{2} (30 mL) was added. The organic phase was washed with H\textsubscript{2}O (30 mL), sat. aq. NaHCO\textsubscript{3} (20 mL), dried (MgSO\textsubscript{4}) and concentrated \textit{in vacuo} to afford the title compound as a pale yellow solid (222 mg, 0.546 mmol, 95%). mp 170-172 °C; \textsuperscript{1}H NMR (600 MHz, CDCl\textsubscript{3}) \(\delta_H\) 4.00 (s, 3H, OCH\textsubscript{3}); \textsuperscript{13}C NMR (150 MHz, CDCl\textsubscript{3}) \(\delta_C\) 161.9 (C(O)N), 147.1 (C(O)O), 119.9 (C), 54.9 (OCH\textsubscript{3}); IR (solid) \(\nu_{\text{max}}/\text{cm}^{-1}\) 2953 (w), 1794 (s), 1757 (s), 1712 (s), 1574 (m), 1556 (s); LRMS (EI) \(m/z\) 407 ([M]\textsuperscript{+}, 96), 179 (100); HRMS (EI) calcd for [C\textsubscript{6}H\textsubscript{3}NOI\textsubscript{2}]\textsuperscript{+} [M]\textsuperscript{+} 406.8146, observed 406.8145.
To a stirred solution of \( \text{N-}(\text{methoxycarbonyl})-3,4\text{-diiodomaleimide} \) \( 50 \) (59 mg, 0.14 mmol) in \( \text{CH}_2\text{Cl}_2 \) (4 mL) was added \( \text{O,O'\text{-Bis}(2\text{-aminoethyl})octadecaethylene glycol} \) (50 mg, 0.056 mmol). After 4 h at RT, silica (ca. 200 mg) was added and the resultant mixture was stirred for further 16 h at RT. The solvent was then removed \textit{in vacuo}. Purification by column chromatography (gradient elution from \( \text{CH}_2\text{Cl}_2 \) to 5% MeOH in \( \text{CH}_2\text{Cl}_2 \)) afforded the target compound as a yellow oil (57 mg, 0.037 mmol, 66%).

\[ ^1\text{H NMR (600 MHz, CDCl}_3) \delta \text{H} 3.82 (t, 4H, J = 5.6, 2 \times \text{NCH}_2), 3.65-3.59 (m, 76H, 38 \times \text{CCH}_2); ^{13}\text{C NMR (150 MHz, CDCl}_3) \delta \text{C} 166.5 (\text{C(O)N}), 117.6 (\text{C}), 70.6 (\text{CH}_2), 70.5 (\text{CH}_2), 70.1 (\text{CH}_2), 67.8 (\text{CH}_2), 39.4 (\text{NCH}_2); \text{IR (oil)} \nu_{\text{max}} \text{ cm}^{-1} 2869 (s), 1773 (w), 1712 (s), 1091 (s); \text{LRMS (ES+)} m/z 1578 ([M+NH}_4]^+, 49), 798 (100); \text{HRMS (ES+)} \text{calcd for [C}_{48}\text{H}_{80}\text{I}_4\text{N}_2\text{O}_{23}\text{NH}_4]^+ [\text{M+NH}_4]^+ 1578.1669, \text{observed 1578.1654.} \]

**Di-tert-butyl(8,50-dioxo-11,14,17,20,23,26,29,32,35,38,41,44,47-tridecaoxa-7,51-diazahepta pentacontane-1,57-diyl)dicarbamate (52)**

To a stirred mixture of \( \text{O,O'\text{-Bis}(2\text{-carboxyethyl})dodecaethylene glycol} \) (445 mg, 0.644 mmol), EDC.HCl (272 mg, 1.42 mmol), HOBt hydrate (217 mg, 1.42 mmol) and DMAP (23.0 mg, 0.188 mmol) in DMF (10 mL) was added \( \text{N-Boc-1,6-diaminohexane} \) (347 \( \mu \text{L}, 1.55 \text{mmol}) at 0 °C. The resultant mixture was brought to RT and stirred for 20 h. The solvent was then removed \textit{in vacuo}, azeotroping with toluene (3 \( \times \) 50 mL). The residue was dissolved in \( \text{CH}_2\text{Cl}_2 \) (60 mL) and washed with \( \text{H}_2\text{O} \) (100 mL), sat. aq. \( \text{NaHCO}_3 \) (100 mL) and 10% aq. citric acid (100 mL). The combined
organic extracts were dried (MgSO₄) and concentrated. Purification by column chromatography (gradient elution from 2% MeOH in CH₂Cl₂ to 10% MeOH in CH₂Cl₂) afforded the target compound as a colourless oil (550 mg, 0.506 mmol, 79%).

¹H NMR (600 MHz, CDCl₃) δH 6.57 (br s, 2H, 2 × C(O)N), 4.64 (br s, 2H, 2 × NHC(O)O), 3.70 (t, 4H, J = 5.8, 2 × CH₂CH₂NHC(O)), 3.63-3.61 (m, 48H, 24 × CH₂), 3.20 (q, 4H, J = 6.9, 2 × C(O)NHC₃H₃), 3.08 (q, 4H, J = 6.4, 2 × CH₂NHC(O)), 2.44 (t, 4H, J = 5.8, 2 × CH₂C(O)NH), 1.48-1.43 (m, 8H, 2 × CH₂CH₂NHC(O) and 2 × CH₂CH₂NHC(O)O), 1.41 (s, 18H, 2 × C(C₃H₃)), 1.30 (quin, 8H, J = 7.2, 2 × CH₂NHCH₂CH₂ and 2 × CH₂CH₂CH₂NHC(O)O); ¹³C NMR (150 MHz, CDCl₃) δC 171.7 (NH(C(O))), 156.1 (NH(C(O)O), 79.1 (C(CH₃)₃), 70.6 (CH₂), 70.4 (CH₂), 70.3 (CH₂), 67.5 (CH₂), 40.4 (C(O)NHCH₂), 39.3 (CH₂NHC(O)), 37.1 (CH₂C(O)NH), 30.1 (CH₂NHC₂CH₂), 29.6 (CH₂NHC(O)), 28.6 (C(CH₃)₃), 26.5 (C(O)NHCH₂CH₂CH₂), 26.4 (CH₂CH₂NHC(O)); IR (oil) νmax/cm⁻¹ 3339 (w, br), 2925 (m), 2866 (m), 1690 (m), 1094 (s); LRMS (ES+) m/z 1105 ([M+NH₄]⁺, 49), 553 (100), 444 (49); HRMS (ES+) calcd for [C₅₂H₁₀₂N₄O₁₉NH₄]⁺ [M+NH₄]⁺ 1104.7477, observed 1104.7459.

N1,N43-Bis(6-(3,4-dibromo-2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)hexyl)- 4,7,10,13,16,19,22, 25,28,31,34,37,40-tridecaoxatritetracontanediamide (DBMCl-C₆)

To a solution of compound 52 (100mg, 0.0920 mmol) in CH₂Cl₂ (5 mL) was added TFA (5 mL) and the reaction mixture was stirred at RT for 12 h. The solvent was then removed in vacuo, azeotroping with toluene (3 × 50 mL). The crude residue was dissolved in CH₂Cl₂ (5 mL) and DIPEA (48 μL, 0.276 mmol) was added. After 15 min at RT, N-(methoxycarbonyl)-3,4-dibromomaleimide 15 (77 mg, 0.25 mmol) was added and the mixture stirred for 1 h. p-TsOH (ca. 50 mg) was then added and the mixture stirred for a further 18 h at RT. The solvent was removed in vacuo and purification by column chromatography (gradient elution from 2% MeOH in CH₂Cl₂ to 5% MeOH in CH₂Cl₂) afforded the target compound as a pale yellow oil (69 mg, 0.051 mmol, 55%).

146
$^1$H NMR (600 MHz, CDCl$_3$) $\delta$H 6.53 (br s, 2H, 2 × C(O)NH), 3.72 (t, 4H, $J = 5.7$, 2 × CH$_2$CH$_2$C(O)NH), 3.67-3.63 (m, 48H, 24 × CH$_2$), 3.60 (t, 4H, $J = 7.2$, 2 × NCH$_2$), 3.22 (q, 4H, $J = 7.0$, 2 × C(O)NHCH$_2$), 2.46 (t, 4H, $J = 5.7$, 2 × CH$_2$C(O)NH), 1.60 (quin, 4H, $J = 7.3$, 2 × CH$_2$CH$_2$N), 1.48 (quin, 4H, $J = 7.5$, 2 × C(O)NHCH$_2$CH$_2$), 1.32 (m, 8H, 2 × C(O)NHCH$_2$CH$_2$ and CH$_2$CH$_2$CH$_2$N); $^{13}$C NMR (150 MHz, CDCl$_3$) $\delta$c 171.7 (NHC(O)), 164.1 (C(O)N), 129.5 (C), 70.6 (CH$_2$), 70.4 (CH$_2$), 70.3 (CH$_2$), 67.5 (CH$_2$CH$_2$C(O)NH), 39.7 (CH$_2$N), 39.3 (C(O)NHCH$_2$), 37.1 (CH$_2$C(O)NH), 29.5 (C(O)NHCH$_2$CH$_2$), 28.5 (CH$_2$CH$_2$N), 26.4 (C(O)NHCH$_2$CH$_2$), 26.4 (CH$_2$CH$_2$CH$_2$N); IR (oil) $\nu_{max}$ cm$^{-1}$ 3301 (w), 2864 (m), 1784 (w), 1721 (s), 1685 (s), 1094 (s); LRMS (ES+) m/z 1384 ([$^{81,81,81,81}$M+NH$_4]^+$, 3), 1382 ([$^{81,81,81,79}$M+NH$_4]^+$, 5), 1380 ([$^{81,81,79,79}$M+NH$_4]^+$, 8), 1378 ([$^{81,79,79,79}$M+NH$_4]^+$, 6), 1376 ([$^{79,79,79,79}$M+NH$_4]^+$, 2), 658 (90), 633 (100), 505 (38), 334 (88); HRMS (ES+) calcd for [C$_{50}$H$_{82}$]$^{81,81,79,79}$Br$_4$N$_4$O$_{19}$H$_2$]$^{2+}$ [$^{81,81,79,79}$M+2H]$^{2+}$ 682.1209, observed 682.1196.

**N1,N43-Bis(6-(3,4-diiodo-2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)hexyl)-4,7,10,13,16,19,22,25,8,31,34,37,40-tridecaoxatritetracontanediamide (DIM$_{CL-C6}$)**

![Diagram of DIM$_{CL-C6}$]

To a solution of compound 52 (142 mg, 0.131 mmol) in CH$_2$Cl$_2$ (5 mL) was added TFA (5 mL) and the reaction mixture was stirred at RT for 12 h. The solvent was then removed in vacuo, azeotroping with toluene (3 × 50 mL). The crude residue was dissolved in CH$_2$Cl$_2$ (7 mL) and DIPEA (68 μL, 0.39 mmol) was added. After 15 min at RT, N-(methoxycarbonyl)-3,4-diiodomaleimide 50 (127 mg, 0.312 mmol) was added and the mixture stirred for 1 h at RT. $p$-TsOH (ca. 50 mg) was then added and the mixture stirred for a further 18 h at RT. The solvent was removed in vacuo and purification by column chromatography (gradient elution from 2% MeOH in CH$_2$Cl$_2$ to 6% MeOH in CH$_2$Cl$_2$) afforded the target compound as a pale yellow oil (136 mg, 0.0877 mmol, 67%).

$^1$H NMR (600 MHz, CDCl$_3$) $\delta$H 6.52 (br s, 2H, 2 × C(O)NH), 3.72 (t, 4H, $J = 5.8$, 2 × CH$_2$CH$_2$C(O)NH), 3.66-3.63 (m, 48H, 24 × CH$_2$), 3.60 (t, 4H, $J = 7.2$, 2 × NCH$_2$), 3.21 (q, 4H, $J$
= 6.8, 2 × C(O)NHCH₂, 2.47 (t, 4H, J = 5.7, 2 × CH₂C(O)NH), 1.59 (quin, 4H, J = 7.3, 2 × CH₂CH₂N), 1.48 (quin, 4H, J = 7.3, C(O)NHCH₂CH₂CH₂), 1.32 (m, 8H, 2 × C(O)NHCH₂CH₂ and 2 × CH₂CH₂CH₂N); ¹³C NMR (150 MHz, CDCl₃) δC 171.7 (NH C(O)), 166.6 (C(O)N), 117.4 (C), 70.6 (CH₂), 70.4 (CH₂), 70.3 (CH₂), 67.5 (CH₂CH₂C(O)NH), 40.2 (CH₂N), 39.3 (C(O)NHCH₂), 37.0 (CH₂CH₂C(O)NH), 29.5 (C(O)NHCH₂CH₂CH₂), 28.5 (CH₂CH₂N), 26.4 (C(O)NHCH₂CH₂) 26.4 (CH₂CH₂CH₂N); IR (oil) νmax/ cm⁻¹: 3314 (w), 2930 (m), 2861 (m), 1772 (w), 1710 (s), 1645 (s), 1096 (s); LRMS (ESI) m/z 1568 ([M+NH₄]⁺, 74), 793 (100); HRMS (ESI) calcd for [C₅₀H₈₂I₄N₄O₁₉NH₄]⁺ [M+NH₄]⁺ 1568.2090, observed 1568.2068.

*Tert*-butyl(6-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl) pentanamido) hexyl)carbamate (53)

D-Biotin (221 mg, 0.905 mmol), HBTU (360 mg, 0.949 mmol) and DIPEA (236 μL, 1.36 mmol) were dissolved in DMF (5 mL) and stirred for 20 min at RT. N-Boc-1,6-hexanediamine (205 mg, 1.16 mmol) in DMF (2 mL) was then added and the resultant mixture was stirred for 24 h at RT. After this period, the solvent was removed *in vacuo*. Purification by column chromatography (gradient elution from CH₂Cl₂ to 10% MeOH in CH₂Cl₂) afforded the target compound as a white crystalline solid (396 mg, 0.895 mmol, 99%).

mp 184-186 °C; ¹H NMR (600 MHz, CD₃OD) δH 4.49 (dd, 1H, J = 7.8, 4.6, SCH₂CH), 4.30 (dd, 1H, J = 7.9, 4.5, SCHCH), 3.22-3.18 (m, 1H, SCH), 3.16 (td, 2H, J = 6.9, 2.8, CH₂NH(O)CH₂), 3.02 (t, 2H, J = 7.0, OC(O)NHCH₂), 2.92 (dd, 1H, J = 12.7, 5.0, SCHH), 2.70 (d, 1H, J = 12.7, SCHH), 2.19 (t, 2H, J = 7.2, NHC(O)CH₂), 1.76-1.56 (m, 4H, SCHCH₂ and SCHCH₂CH₂CH₂), 1.52-1.44 (m, 6H, OC(O)NHCH₂CH₂, CH₂C(O)NHCH₂CH₂ and SCHCH₂CH₂), 1.43 (s, 9H, (C(CH₃)₃), 1.35-1.32 (m, 4H, OC(O)NHCH₂CH₂CH₂ and CH₂C(O)NHCH₂CH₂CH₂); ¹³C NMR (150 MHz, CD₃OD) δC 176.0 (CH₂C(O)), 166.1 (NHC(O)NH), 158.6 (OC(O)NH), 79.8 (C(CH₃)₃), 63.4 (SCHCH), 61.6 (SCH₂CH), 57.0 (SCH), 41.2 (OC(O)NHCH₂), 41.0 (SCH₂), 40.2
(CH₂NHC(O)CH₂, 36.8 (NHC(O)CH₂), 30.9 (SCHCH₂), 30.4 (OC(O)NHCH₂CH₂), 29.8 (CH₂C(O)NHCH₂CH₂), 29.5 (SCH₂CH₂), 28.8 (C(CH₃)₃), 27.7 (CH₂C(O)NHCH₂CH₂CH₂), 27.5 (OC(O)NHCH₂CH₂CH₂), 27.0 (SCH₂CH₂CH₂); IR (solid) ν_max/cm⁻¹ 3386 (w), 3252 (w, br), 2932 (w), 1687 (s), 1621 (m), 1516 (s); LRMS (ES⁺) m/z 465 ([M+Na]+, 17), 443 ([M+H]+, 100), 343.2 (45); HRMS (ES⁺) calcd for [C₂₁H₃₈N₄O₅S]⁺ [M+H]+ 443.2687, observed 443.2675.

N-(6-(3,4-Dibromo-2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)hexyl)-5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (54)

Compound 53 (180 mg, 0.407 mmol) was dissolved in TFA (8 mL). After 18 h at RT, the solvent was removed in vacuo. The crude residue was dissolved in AcOH (18 mL) and dibromomaleic acid (112 mg, 0.409 mmol) was added. The reaction was heated to reflux for 3 h. All volatile material was then removed in vacuo. Purification by column chromatography (gradient elution from CH₂Cl₂ to 12% MeOH in CH₂Cl₂) afforded the target product as a pale yellow solid (143 mg, 0.247 mmol, 61%).

mp 158-160 °C; ¹H NMR (600 MHz, CD₃OD) δH 5.86 (br s, 2H, NHC(O)NH), 5.05 (br s, 1H, NHC(O)), 4.53 (dd, 1H, J = 7.5, 5.1, SCH₂CH), 4.33 (dd, 1H, J = 7.3, 4.9, SCHCH), 3.61 (t, 2H, J = 7.2, NCH₂), 3.22 (m, 2H, NHCH₂), 3.16 (m, 1H, SCH), 2.93 (dd, 1H, J = 12.9, 5.0, SCHH), 2.74 (d, 1H, J = 12.8, SCHH), 2.21 (m, 2H, NHC(O)CH₂), 1.78-1.66 (m, 4H, SCHCH₂ and C(O)CH₂CH₂), 1.64-1.58 (m, 2H, NCH₂CH₂), 1.52-1.42 (m, 4H, NHCH₂CH₂ and SCHCH₂CH₂), 1.37-1.28 (m, 4H, NCH₂CH₂CH₂ and NHCH₂CH₂CH₂); ¹³C NMR (150 MHz, CDCl₃) δc 173.1 (C(O)CH₂), 164.2 (NHC(O)NH), 163.5 (C(O)N), 129.5 (CBr), 61.9 (SCHCH), 60.2 (SCH₂CH), 55.5 (SCH), 40.7 (SCH₂), 39.7 (NHCH₂), 39.4 (NCH₂), 36.1 (C(O)CH₂), 29.5 (NHCH₂CH₂), 28.4 (NCH₂CH₂), 28.2 (SCHCH₂), 28.2 (SCH₂CH₂), 26.3 (C(O)CH₂CH₂), 26.3 (NHCH₂CH₂CH₂), 25.7 (NCH₂CH₂CH₂); IR (solid) ν_max/cm⁻¹ 3289 (w, br), 2931 (w), 2859 (w), 1715 (s), 1693 (s), 149
1641 (m), 1544 (m); LRMS (ES+) m/z 605 ([\(^{81,81}\text{M+Na}\])\(^{+}\), 8), 603 ([\(^{81,79}\text{M+Na}\])\(^{+}\), 16), 601 ([\(^{79,79}\text{M+Na}\])\(^{+}\), 8), 583 ([\(^{81,81}\text{M+H}\])\(^{+}\), 54), 581 ([\(^{81,79}\text{M+H}\])\(^{+}\), 100), 579 ([\(^{79,79}\text{M+H}\])\(^{+}\), 49); HRMS (ES+) calcd for [C\(_{20}\)H\(_{28}\)\(^{81,79}\text{Br}_{2}\)N\(_{4}\)O\(_{4}\)S\(_{4}\)]\(^{+}\) [M+H]\(^{+}\) 581.0251, observed 581.0242.

**Di-tert-butyl (27-(2,2-dimethyl-4,30-dioxo-3,8,11,14,17,20,23,26-octaoxa-5,29-diazahentriacontan-31-yl)-25,29-dioxo-3,6,9,12,15,18,21,34,37,40,43,46,49,52-tetradecaoxa-24,27,30-triazatetrapentacontane-1,54-diyl)dicarbamate (56)**

\[
\begin{align*}
\text{HN} & \quad \text{O} \\
\text{O} & \quad \text{O} \\
\text{HN} & \quad \text{O} \\
\text{O} & \quad \text{O} \\
\end{align*}
\]

To *O-* (2-Aminoethyl)-*O*'-[2-(Boc-amino)ethyl]hexaethylene glycol (142 mg, 0.303 mmol) in MeCN (3 mL) was added nitrilotriacetic acid tri(N-succinimidyl) ester (44.0 mg, 0.0912 mmol). After 20 h at RT, the solvent was removed *in vacuo*. CH\(_2\)Cl\(_2\) (50 mL) was added and the organic phase was washed with sat. aq. NaHCO\(_3\) (30 mL), 10% aq. citric acid (30 mL), dried (MgSO\(_4\)) and concentrated. Purification by column chromatography (gradient elution from CH\(_2\)Cl\(_2\) to 10% MeOH in CH\(_2\)Cl\(_2\)) afforded the target compound as a colourless oil (114 mg, 0.0739, 81%).

\(^1\text{H NMR (600 MHz, CDCl}_3\)) \(\delta\) 7.67 (br s, 3H, 3 × NHC(O)), 5.09 (br s, 3H, 3 × NHC(O)O), 3.62-3.60 (m, 54H, 27 × CH\(_2\)), 3.59-3.57 (m, 18H, 9 × CH\(_2\)), 3.53 (t, 6H, \(J = 5.5\), 3 × C(O)NHCH\(_2\)CH\(_2\)), 3.51 (t, 6H, \(J = 5.1\), 3 × OC(O)NHCH\(_2\)CH\(_2\)), 3.43 (q, 6H, \(J = 5.5\), 3 × C(O)NHCH\(_2\)CH\(_2\)), 3.28 (t, 6H,
$J = 5.3, 3 \times \text{OC(O)NHCH}_2$, 3.27 (s, 6H, 3 $\times \text{NCH}_2$), 1.41 (s, 27H, 3 $\times \text{C(CH}_3)_3$); $^{13}$C NMR (150 MHz, CDCl$_3$) $\delta$ 170.1 (C(O)NH), 156.1 (OC(O)NH), 79.2 (C(CH$_3$)$_3$), 70.6 (CH$_2$), 70.3 (CH$_2$), 70.2 (CH$_2$), 69.7 (CH$_2$), 59.4 (NCH$_2$), 40.4 (OC(O)NHCH$_2$), 39.1 (C(O)NHCH$_2$), 28.6 (C(CH$_3$)$_3$); IR (oil) $\nu_{\text{max}}$/ cm$^{-1}$ 3324 (w, br), 2868 (m), 1705 (m), 1659 (m), 1089 (s); LRMS (ES+) $m/z$ 1565 ([M+Na]$^+$, 21), 1543 ([M+H]$^+$, 27), 780 (76), 622 (14), 521 (26), 415 (100); HRMS (ES+) calcd for [C$_{69}$H$_{135}$N$_7$O$_{30}$H]$^+$ [M+H]$^+$1542.9326, observed 1542.9313.

2,2',2''-Nitrilotris(N-(23-(3,4-diiodo-2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)-3,6,9,12,15,18,21-heptaaxatricosyl)acetamide) (tDIMCl)

To compound 56 (67.0 mg, 0.0435 mmol) in CH$_2$Cl$_2$ (2 mL) was added TFA (2 mL). The resultant mixture was stirred for 18 h at RT. The solvent was then removed in vacuo, azeotroping with toluene (3 $\times$ 30 mL). The residue was dissolved in H$_2$O (4 mL) and the pH was adjusted to 11 with aq. NaOH (0.05 M). The aqueous phase was extracted into CH$_2$Cl$_2$ (3 $\times$ 50 mL), dried (MgSO$_4$) and concentrated. The residue was then dissolved in CH$_2$Cl$_2$ (6 mL) and N-(methoxycarbonyl)-3,4-diiodomaleimide 50 (67.0 mg, 0.165 mmol) in CH$_2$Cl$_2$ (4 mL) was added to give a yellow solution that was stirred for 2 h at RT. Silica (ca. 400 mg) was then added and after 1h at RT, the solvent was removed in vacuo. Purification by column chromatography (gradient elution from
CH₂Cl₂ to 10% MeOH in CH₂Cl₂) afforded the target compound as a yellow oil (68.0 mg, 0.0296 mmol, 68%).

¹H NMR (600 MHz, DMSO) δ_H 8.16 (t, 3H, J = 5.7, 3 × NH), 3.63 (t, 6H, J = 5.7, 3 × CH₂), 3.51-3.48 (m, 66H, 33 × CH₂), 3.47-3.44 (m, 12H, 6 × CH₂), 3.41 (t, 6H, J = 5.8, 3 × CH₂), 3.24 (q, 6H, J = 5.7, 3 × NHCH₂), 3.18 (s, 6H, 3 × NCH₂C(O)); ¹³C NMR (150 MHz, DMSO) δ_C 170.1 (C(O)NH), 167.3 (C(O)N), 119.2 (C), 69.8 (CH₂), 69.6 (CH₂), 69.5 (CH₂), 57.8 (NCH₂C(O)), 38.3 (CH₂); IR (oil) v_max/cm⁻¹ 3302 (w, br), 2868 (w), 1710 (s), 1657 (m), 1096 (s); LRMS (MALDI-TOF ultrafleXtreme) m/z 2260 ([M+Na]^+, 100), 2134 (42), 2008 (10), 1535 (11); HRMS (MALDI-TOF ultrafleXtreme) calcd for [C₆₆H₁₀₅I₆N₇O₃₀Na]⁺ [M+Na]^+ 2260.1066, observed 2260.1021.

To compound 56 (106 mg, 0.0687 mmol) in CH$_2$Cl$_2$ (4 mL) was added TFA (4 mL). The resultant mixture was stirred for 18 h at RT. The solvent was then removed in vacuo, azeotroping with toluene (3 × 50 mL). The residue was dissolved in DMF (3 mL). DIPEA (64.0 μL, 0.368 mmol) was added and the mixture was stirred for 30 min at RT. This was then added to a stirred solution of Boc-6-aminohexanoic acid (79.0 mg, 0.342 mmol), EDC.HCl (65.0 mg, 0.339 mmol), HOBt hydrate (52.0 mg, 0.340 mmol) and DMAP (5.00 mg, 0.0409 mmol) in DMF (10 mL) at 0 °C. The resultant mixture was brought to RT and stirred for 20 h. The solvent was then evaporated in vacuo. The crude residue was dissolved in CHCl$_3$ (50 mL) and washed with H$_2$O (50 mL), sat. aq.
NaHCO₃ (50 mL), 10% aq. citric acid (50 mL), dried (MgSO₄) and concentrated. Purification by column chromatography (gradient elution from CH₂Cl₂ to 10% MeOH in CH₂Cl₂) afforded the target compound as a colourless oil (109 mg, 0.0579 mmol, 85%).

¹H NMR (600 MHz, CDCl₃) δH 7.65 (br s, 3H, 3 × NCH₂C(O)NH), 6.27 (br s, 3H, 3 × NHC(O)), 4.68 (br s, 3H, 3 × NHC(O)O), 3.64-3.61 (m, 60H, 30 × CH₂), 3.60-3.58 (m, 12H, 6 × CH₂), 3.55-3.53 (m, 12H, 6 × CH₂), 3.43 (m, 12H, 3 × NCH₂C(O)NHCH₂ and 3 × CH₂NHC(O)), 3.30 (s, 6H, 3 × NCH₂), 3.09 (q, 6H, J = 6.5, 3 × CH₂NHC(O)O), 2.17 (t, 6H, J = 7.5, 3 × NHC(O)CH₂), 1.63 (quin, 6H, J = 7.7, 3 × NHC(O)CH₂CH₂), 1.47 (quin, 6H, J = 7.3, 3 × CH₂CH₂NHC(O)O), 1.42 (s, 27H, 3 × C(CH₃)₃), 1.32 (quin, 6H, J = 7.3, 3 × NHC(O)CH₂CH₂CH₂CH₂); ¹³C NMR (150 MHz, CDCl₃) δc 173.1 (NH(C(O)), 170.8 (NCH₂C(O)), 156.1 (NHC(O)O), 79.1 (C(CH₃)₃), 70.6 (CH₂), 70.4 (CH₂), 70.3 (CH₂), 70.0 (CH₂), 69.6 (CH₂), 68.8 (CH₂), 59.4 (NCH₂), 40.5 CH₂NHC(O)O, 39.9 (CH₂), 39.2 (CH₂), 39.1 (CH₂), 36.5 (NHC(O)CH₂), 29.9 (CH₂CH₂NHC(O)O), 28.6 (C(CH₃)₃), 26.6 (NHC(O)CH₂CH₂CH₂), 25.4 (NHC(O)CH₂CH₂); IR (oil) νmax/ cm⁻¹ 3309 (br, w), 2866 (m), 1650 (s), 1538 (s); LRMS (ES+) m/z 1904 ([M+Na]⁺, 6), 951 (56), 892 (52), 628 (88), 528 (100); HRMS (ES+) calcd for [C₈₇H₁₆₈N₁₀O₃₃Na₂][M+2Na]²⁺ 963.5780, observed 963.5764.
To compound 57 (55 mg, 0.029 mmol) in CH$_2$Cl$_2$ (2 mL) was added TFA (2 mL). The resultant mixture was stirred for 18 h at RT. The solvent was then removed in vacuo, azeotroping with toluene (3 × 30 mL). The residue was dissolved in H$_2$O (4 mL) and the pH was adjusted to 11 with aq. NaOH (0.05 M). The aqueous phase was extracted into CH$_2$Cl$_2$ (3 × 50 mL), dried (MgSO$_4$) and concentrated. This was then dissolved in CH$_2$Cl$_2$ (6 mL) and N-(methoxycarbonyl)-3,4-diiodomaleimide 50 (42.0 mg, 0.104 mmol) in CH$_2$Cl$_2$ (4 mL) was added to give a yellow solution that was stirred for 2 h at RT. Silica (ca. 400 mg) was then added. After 18h at RT, the solvent was removed in vacuo. Purification by column chromatography (gradient elution from CH$_2$Cl$_2$ to 15% MeOH in CH$_2$Cl$_2$) afforded the target compound as a yellow oil (17 mg, 0.0066 mmol, 22%).
$^1$H NMR (600 MHz, CDCl$_3$) $\delta$ 7.70 (br s, 3H, 3 × NCH$_2$(O)NH), 6.34 (br s, 3H, 3 × NHC(O)), 3.66-3.63 (m, 6H, 3 × CH$_2$CH$_2$N and 30 × CH$_2$), 3.62-3.60 (m, 12H, 6 × CH$_2$), 3.57-3.54 (m, 12H, 6 × CH$_2$), 3.44 (2 overlapped q, 12H, $J = 5.4, 5.1, 3 ×$ NCH$_2$(O)NHCH$_2$ and 3 × CH$_2$NHC(O)), 3.30 (s, 6H, 3 × NCH$_2$(O)C), 2.17 (t, 6H, $J = 7.5, CH_2$NHC(O)CH$_2$), 1.65 (quin, 6H, $J = 7.4, 3 ×$ NHC(O)CH$_2$CH$_2$), 1.61 (quin, 6H, $J = 7.9, 3 ×$ CH$_2$CH$_2$NC(O)), 1.32 (quin, 6H, $J = 6.5, 3 ×$ NHC(O)CH$_2$CH$_2$CH$_2$); $^{13}$C NMR (150 MHz, CDCl$_3$) $\delta$C 172.9 (NHCO), 170.8 (NCH$_2$(O)), 166.5 (C(O)N), 117.5 (C), 70.7 (CH$_2$), 70.6 (CH$_2$), 70.3 (CH$_2$), 70.1 (CH$_2$), 69.7 (CH$_2$), 59.4 (NCH$_2$(O)), 40.2 (CH$_2$NC(O)), 39.3 (CH$_2$), 39.2 (CH$_2$), 36.3 (NHC(O)CH$_2$), 28.4 (CH$_2$CH$_2$NC(O)), 26.4 (NHC(O)CH$_2$CH$_2$CH$_2$), 25.2 (NHC(O)CH$_2$CH$_2$); IR (oil) $\nu_{\text{max}}$/ cm$^{-1}$ 3277 (br w), 2938 (m), 2870 (m), 1720 (s), 1653 (s); LRMS (ES$^+$) m/z 1311 ([M+2Na]$^{2+}$, 66), 1281 (76), 1255 (80), 882 (100), 844 (96), 765 (28); HRMS (ES$^+$) calcd for [C$_{84}$H$_{138}$I$_6$N$_{10}$O$_{33}$Na$_2$]$^{2+}$ [M+2Na]$^{2+}$ 1311.1740, observed 1311.1716.

**Sodium 2-(pent-4-ynoylthio)ethane-1-sulfonate (59)**

![Structure](image)

A solution of 4-pentynoic acid (200 mg, 2.04 mmol) and EEDQ (604 mg, 2.45 mmol) in MeCN (15 mL) and DMF (2 mL) was stirred at RT for 30 min. Then, sodium 2-mercaptopoethanesulfonate (335 mg, 2.04 mmol) was added and the resultant mixture was stirred at 80 °C for 18 h. The solvent was then removed in vacuo and purification by column chromatography (gradient elution from CH$_2$Cl$_2$ to 20% MeOH in CH$_2$Cl$_2$) afforded the target compound as a white solid (357 mg, 1.46 mmol, 72%).

mp 240 °C (decomposition); $^1$H NMR (600 MHz, CD$_3$OD) $\delta$H 3.27-3.24 (m, 2H, C(O)SCH$_2$), 3.01-2.98 (m, 2H, CH$_2$SO$_3$Na), 2.78 (t, 2H, $J = 7.2, C$(O)CH$_2$), 2.50 (td, 2H, $J = 7.2, 2.5, C$(O)CH$_2$CH$_2$), 2.27 (t, 1H, $J = 2.3, CH$); $^{13}$C NMR (150 MHz, CD$_3$OD) $\delta$C 198.4 (C(O)), 82.8 (C), 70.5 (CH), 52.2 (CH$_2$SO$_3$Na), 43.2 (C(O)CH$_2$), 25.0 (C(O)SCH$_2$), 15.2 (C(O)CH$_2$CH$_2$); IR (solid) $\nu_{\text{max}}$/ cm$^{-1}$ 3291 (m), 3237 (m), 2993 (w), 2938 (w), 1679 (s); LRMS (ES$-$) m/z 221 ([M-H]$^-$, 100); HRMS (ES$-$) calcd for [C$_7$H$_8$O$_4$S$_2$]$^-$ [M-H]$^-$ 220.9948, observed 220.9945.
3.2 UV/Vis hydrolysis study

N-Me maleimide derivatives (20 mM solutions in DMF) were incubated in conjugation buffer (40 mM phosphate, 10 mM NaCl, 6 mM EDTA, pH 7.4) at the following final concentrations: N-Me dichloromaleimide (3.41 mM), N-Me dibromomaleimide (1.70 mM), N-Me diiodomaleimide (0.60 mM), N-Me diphenoxymaleimide (0.90 mM), N-Me dithiophenolmaleimide (0.175 mM) and N-Me maleimide (1.80 mM). The concentration of DMF was adjusted to 10% and the absorbance of each sample was monitored over time.

Kinetic data was further analysed to obtain the pseudo first order rate constants and half-lives. The extinction coefficients of N-Me maleimide derivatives were determined experimentally from the slopes of best fit lines generated by plotting absorbance versus concentration (Figure 82 and 83).

Pseudo first-order rate constants were then obtained from the slopes of best fit lines generated from plotting ln [NGM] versus time and linear regression analysis (Figure 84). Hydrolysis reaction half-lives ($t_{1/2}$) were calculated from the pseudo first order rate constants using the following equation: 

$$t_{1/2} = \frac{\ln(2)}{k_{1,obs}}$$
Figure 82: Absorption spectra of $N$-Me maleimide derivatives at different concentrations; a. $N$-Me dichloromaleimide ($\lambda_{\text{max}} = 312$ nm), b. $N$-Me dibromomaleimide ($\lambda_{\text{max}} = 326$ nm), c. $N$-Me diiodomaleimide ($\lambda_{\text{max}} = 360$ nm), d. $N$-Me diphenoxymaleimide ($\lambda_{\text{max}} = 346$ nm), e. $N$-Me dithiophenolmaleimide ($\lambda_{\text{max}} = 430$ nm), f. $N$-Me maleimide ($\lambda_{\text{max}} = 300$ nm).
Figure 83: Determination of extinction coefficients of \(N\)-Me maleimide derivatives; a. \(N\)-Me-dichloromaleimide (\(\varepsilon = 290.0 \text{ M}^{-1} \text{ cm}^{-1}\)), b. \(N\)-Me dibromomaleimide (\(\varepsilon = 530.9 \text{ M}^{-1} \text{ cm}^{-1}\)), c. \(N\)-Me diiodomaleimide (\(\varepsilon = 1789 \text{ M}^{-1} \text{ cm}^{-1}\)), d. \(N\)-Me diphenoxymaleimide (\(\varepsilon = 1067 \text{ M}^{-1} \text{ cm}^{-1}\)), e. \(N\)-Me dithiophenol maleimide (\(\varepsilon = 6392 \text{ M}^{-1} \text{ cm}^{-1}\)), \(N\)-Me maleimide (\(\varepsilon = 638.9 \text{ M}^{-1} \text{ cm}^{-1}\)).
Figure 84: Determination of pseudo first order rate constants of $N$-Me maleimide derivatives; a. $N$-Me dichloromaleimide ($k_{1,obs} = 0.07659 \text{ min}^{-1}$), b. $N$-Me dibromomaleimide ($k_{1,obs} = 0.03871 \text{ min}^{-1}$), c. $N$-Me diiodomaleimide ($k_{1,obs} = 0.01444 \text{ min}^{-1}$), d. $N$-Me diphenoxymaleimide ($k_{1,obs} = 0.02999 \text{ min}^{-1}$), e. $N$-Me dithiophenolmaleimide ($k_{1,obs} = 0.00125 \text{ min}^{-1}$), f. $N$-Me maleimide ($k_{1,obs} = 0.001425 \text{ min}^{-1}$).
3.3 Bioconjugation general remarks

Albumin, extracted from human serum was purchased from Sigma-Aldrich. Recombinant albumin, expressed in *Saccharomyces cerevisiae* was purchased from Sigma-Aldrich. Anti-CEA shmFe ds-scFv was kindly provided by the Chester group (UCL Cancer Institute) and expressed following a literature procedure. Anti-CD3 scFv was kindly provided by the Chester group (UCL Cancer institute); procedure not published. Trastuzumab (Herceptin™) was purchased from UCLH in its clinical formulation (Roche, lyophilized). Octreotide was purchased from Generon in its clinical formulation (lyophilized).

Reagents and solvents were purchased from commercial sources and used as supplied. All buffer solutions were prepared with double-deionized water and filter-sterilized. All buffer solutions were degassed prior to use, where the term ‘degassed’ refers to the process of removing O₂ from a solution by bubbling argon through it. Conjugation buffer was 40 mM phosphates, 20 mM NaCl, 6 mM EDTA at pH 7.4. Phosphate-buffered saline (PBS) was 12 mM phosphates, 140 mM NaCl at pH 7.4. BBS buffer was 80 mM boric acid, 20 mM NaCl at pH 8.0-9.5 (as specified under each experiment). Ultrapure DMF was purchased from Sigma-Aldrich and stored under dry conditions. DTT and TCEP.HCl were purchased from Sigma-Aldrich and stored at 4 °C.

Conjugation experiments were carried out in standard polypropylene Eppendorf® safe-lock tubes (1.5 or 2.0 mL) at atmospheric pressure with mixing at 20 °C, unless otherwise stated. Ultrafiltration was carried out in Amicon® Ultra-4 Centrifugal Filter Units with a molecular weight cut-off (MWCO) of 30 or 10 kDa or in Vivaspin® 500 centrifugal concentrators (5 or 10 kDa MWCO). Centrifugation was carried out on an Eppendorf 5415R fixed angle rotor centrifuge operating at 14000 rcf at 20 °C or in an Eppendorf® 5810 swing-bucket rotor centrifuge operating at 3220 rcf at 20 °C. Temperature controlled experiments were carried out on an Eppendorf® thermomixer. Concentration of peptide samples following trypsin digestion was carried out using an Eppendorf® centrifugal vacuum concentrator 5301 operating at 240 rcf at 45 °C. Ni-IMAC purification was carried out with a HisTrap™ HP column, operated with a peristaltic pump. Purification by size exclusion chromatography (SEC) was carried out on an ÄKTA FPLC system (GE Healthcare), equilibrated in PBS, unless otherwise stated. Detection was by absorption at 280 nm.

Protein concentrations were determined by UV/Vis absorbance using a ThermoScientific NanoDrop 2000C running in A280 mode, using molecular extinction coefficients of: \( \varepsilon_{\text{Fab}} = 68,590 \)
M⁻¹ cm⁻¹, \( \varepsilon_{\text{(anti-CEA scFv)}} = 48,735 \) M⁻¹ cm⁻¹, \( \varepsilon_{\text{(anti-CD3 scFv)}} = 57,215 \) M⁻¹ cm⁻¹ (calcd from sequence), \( \varepsilon_{\text{HSA}} = 36,500 \) M⁻¹ cm⁻¹, \( \varepsilon_{\text{HSA-Fab}} = 105,090 \) M⁻¹ cm⁻¹ (calcd from monomers), \( \varepsilon_{\text{(anti-CEA trimer)}} = 146,205 \) M⁻¹ cm⁻¹ (calcd from monomer), \( \varepsilon_{\text{(anti-CEA/anti-CD3 trimer)}} = 163,165 \) M⁻¹ cm⁻¹ (calcd from monomers).

Protein conjugation reactions were monitored by 12% glycine-SDS-PAGE with a 6% stacking gel under non-reducing conditions, unless otherwise stated. Samples were mixed 1:1 with SDS non-reducing loading buffer (composition for 6 x SDS: 1 g SDS, 3 mL glycerol, 6 mL 0.5 M Tris buffer pH 6.8, 2 mg bromophenol blue) or reducing loading buffer (composition for 4 x SDS: 0.8 mL β-mercaptoethanol (BME), 0.8 g SDS, 4 mL glycerol, 2.5 mL 0.5 M Tris buffer pH 6.8, 2.5 mL H₂O, 1 mg bromophenol blue) and heated at 75 °C for 5 min before applied to the gel. Samples were run at constant current (30 mA) for 40 minutes in 1 x SDS running buffer. Gels were stained with Coomassie G-250 (0.05% w/v) in 49.95% H₂O, 40% MeOH, 10% AcOH and de-stained with 10% MeOH, 10% AcOH, 80% H₂O solution.

MALDI-TOF analysis was performed on a MALDI micro MX – TOF Mass Spectrometer (Waters Corporation) in positive linear mode using a 337 nm nitrogen laser (laser power set at 180, pulse at 2000, detector at 2500). Protein samples (in H₂O) were mixed 1:1 with matrix (sinapinic acid, prepared at a concentration of 10 mg/mL in 70:30 v/v MeCN:H₂O) at a final concentration of ca. 0.3-0.4 mg/mL. 3 μL of each sample were deposited onto a stainless steel MALDI-TOF plate (Waters Corporation) and air-dried before inserting the plate into the instrument. Spectral acquisition was performed with MassLynx software.

The LC-MS spectra shown in section 2.1.1 were obtained using a Waters Acquity uPLC system connected to Waters Acquity Single Quad Detector (SQD) with the following parameters. Hypersil Gold C4, 1.9 μm, 2.1 mm × 50 mm column was used, which was maintained at 50 °C. 10 μL of a protein sample (at ca. 2-4 μM) was separated on the column using an eluting gradient of 5% MeCN/ 0.1% formic acid in H₂O/ 0.1% formic acid to 95% MeCN/ 0.1% formic acid in H₂O/ 0.1% formic acid over 5 min at a flowrate of 600 μL/min. MS Mode: ES+. Scan Range: \( m/z = 250 – 2000 \). Scan time: 0.25 s. Data obtained in continuum mode. The electrospray source of the MS was operated with a capillary voltage of 3.5 kV and a cone voltage of 50V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 L/h. Ion series were generated by integration of the total ion chromatogram (TIC) over the appropriate range. The raw data was then
analyzed by deconvoluting a spectrum to a zero charge mass spectrum using the MaxEnt1 deconvolution algorithm within the MassLynx software.

All other LC-MS spectra were obtained using an Agilent 6510 QTOF LC-MS system (Agilent, UK). Agilent 1200 HPLC system was equipped with an Agilent PLRP-S, 1000Å, 8 µm, 150 mm × 2.1 mm column. 10 µL of a protein sample (at ca. 2-4 µM) was separated on the column using an eluting gradient of 15% MeCN/ 0.1% formic acid in H₂O/ 0.1% formic acid to 95% MeCN/ 0.1% formic acid in H₂O/ 0.1% formic acid over 20 min, at a flow rate of 300 µL/min. The oven temperature was maintained at 60 °C. Agilent 6510 QTOF mass spectrometer was operated in a positive polarity mode, coupled with an ESI ion source. The ion source parameters were set up with a VCap of 3500 V, a gas temperature at 350 °C, a dry gas flow rate at 10 L/min and a nebulizer of 30 psig. MS Tof was acquired under conditions of a fragmentor at 350 V, a skimmer at 65 V and an acquisition rate at 0.5 spectra/s in a profile mode, within a scan range between 700 and 5000 m/z. The raw data was then analysed by deconvoluting a spectrum to a zero charge mass spectrum using a maximum entropy deconvolution algorithm within the MassHunter software version B.07.00.

Albumin peptide sequence identification post tryptic digestion was carried out by LC-MS/MS, using a Finnigan LTQ mass spectrometer coupled with an Accela 600 HPLC system (Thermo Scientific, UK). The Accela 600 HPLC system was fitted with a Hypersil Gold C₁₈ column, 1.9 µm, 150 mm × 2.1 mm (Thermo Scientific, UK). Peptides were separated using an eluting gradient of 5% MeCN/ 0.1% formic acid in H₂O/ 0.1% formic acid to 95% MeCN/ 0.1% formic acid in H₂O/ 0.1% formic acid over 40 min. The flow rate was at 200 µL/min and the column temperature at RT. In a general, 10 µl of a sample (equivalent to ca. 1 µg of peptides) was loaded onto the column for each run. The Finnigan LTQ mass spectrometer was operated in a positive polarity mode within an ESI source. The ion source parameters were set up with a spray voltage of 4.5 V, a capillary voltage of 49 V, an ion source temperature at 280 °C and sheath/auxiliary gas flow rates at 40/10 arb. The MS full scan was acquired at a normal scan rate within a mass range between 400 and 2000 m/z, and centroid data were collected. The peptides of interest were fragmented simultaneously at a CID-type normalized collision energy of 70 with a max mass range. The iso-width was set up as 2 m/z and Act Q as 0.25. The raw data was analysed using the Thermo Xcalibur software and the peptides of interest were assigned manually according to the fragmented spectra.
3.3.1 Bioconjugation experiments

3.3.1.1 Albumin-antibody fragment conjugation

**General method for the preparation of anti-CEA scFv homodimer**

To anti-CEA scFv (300 μL, 0.011 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was added DTT (6.0 μL, 0.22 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MiniTrap G-25, GE Healthcare) and the reduced scFv was buffer exchanged into conjugation buffer, via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (220 μL, 1.0 mg/mL). DBMCl or DTPMCl (9.2 μL, 0.0034 μmol, 372 μM solution in DMF, 0.42 eq.) was then added to reduced scFv (220 μL, 0.0082 μmol, 37.2 μM, 1.0 mg/mL). After 1 h or 24 h at RT, the mixture was analyzed by SDS-PAGE.

**General procedure for the reduction of commercial HSA, extracted from human serum**

To HSA (400 μL, 0.015 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added DTT (2.0 μL, 0.074 μmol, 37.2 mM solution in conjugation buffer, 5 eq.) and the resultant mixture was incubated for 1 h at RT. The excess DTT was then removed using a desalting column (PD MiniTrap G-25, GE Healthcare) and the reduced HSA was buffer exchanged into conjugation buffer, via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (280 μL, 0.010 μmol, 2.47 mg/mL). The reduction was confirmed by LC-MS.

**Reaction of HSA with N-Me dibromomaleimide, followed by addition of BME (1.0 eq.)**

To HSA (1000 μL, 0.037 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added N-Me dibromomaleimide (1.5 μL, 0.056 μmol, 37.2 mM solution in DMF, 1.5 eq.) and the reaction was incubated for 20 min at RT. The excess reagent was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (940 μL, 0.035 μmol, 2.47 mg/mL). To this solution was added BME (0.94 μL, 0.035 μmol, 37.2 mM solution in conjugation buffer, 1 eq.) and the resultant mixture was incubated for 2 h at RT. The UV/Vis absorbance spectrum of the sample before and after BME
addition was recorded. (The same protocol was used for reaction of HSA with $N$-Me monobromomaleimide).

**General procedure for the modification of anti-CEA scFv with DBM$_{CL}$ (or DTPM$_{CL}$ or bis-dithiophenolmaleimide 2.5 kDa PEG linker 46), followed by conjugation to HSA**

Anti-CEA scFv (320 μL, 0.012 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (6.4 μL, 0.24 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was removed using a desalting column (PD Mini Trap G-25, GE Healthcare) and the reduced scFv was buffer exchanged into conjugation buffer, via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (240 μL, 1.0 mg/mL). Linker DBM$_{CL}$ or DTPM$_{CL}$ or 46 (7.2 μL, 0.27 μmol, 37.2 mM solution in DMF, 30 eq.) was then added to reduced scFv (240 μL, 0.0089 μmol, 37.2 μM) and the resultant mixture was incubated at RT for 15 min (For linker DTPM$_{CL}$ or 46, the concentration of DMF was adjusted to 15%). The excess linker was then removed via ultrafiltration (10 kDa MWCO for DBM$_{CL}$ and DTPM$_{CL}$ or 30 kDa MWCO for 46). The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (150 μL, 1.0 mg/mL). To modified scFv (37.2 μM, 1.0 mg/mL) was then added HSA (37.2 μM, 2.47 mg/mL) in conjugation buffer in the following ratios, by volume: (HSA: scFv) 10:1 (18.2 μL: 1.8 μL), 5:1 (16.7 μL: 3.3 μL), 3:1 (15.0 μL: 5.0 μL), 2:1 (13.4 μL: 6.6 μL), 1:1 (10.0 μL: 10.0 μL), 1:2 (6.6 μL: 13.4 μL), 1:3 (5.0 μL: 15.0 μL), 1:5 (3.3 μL: 16.7 μL), 1:10 (1.8 μL: 18.2 μL). The mixtures were incubated at RT and analyzed at different timepoints by SDS-PAGE.

**General procedure for the modification of HSA with DBM$_{CL}$ (or DTPM$_{CL}$ or bis-dithiophenol maleimide 2.5 kDa PEG linker 46), followed by conjugation to scFv**

To HSA (600 μL, 0.022 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added linker DBM$_{CL}$ or DTPM$_{CL}$ (3.0 μL, 0.11 μmol, 37.2 mM solution in DMF, 5 eq.) or 46 (7.2 μL, 0.27 μmol, 37.2 mM solution in DMF, 12 eq) and the reaction was incubated for 15 min at RT (For DTPM$_{CL}$ or 46 the concentration of DMF was adjusted to 10%). The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was
determined by UV/Vis absorbance and adjusted to 37.2 μM (560 μL, 2.47 mg/mL). Meanwhile, anti-CEA scFv (400 μL, 0.015 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (8.0 μL, 0.30 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MiniTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (340 μL, 1.0 mg/mL). To modified HSA (37.2 μM, 2.47 mg/mL) was then added reduced scFv (37.2 μM, 1.0 mg/mL) in the following ratios, by volume: (HSA:scFv) 10:1 (18.2 μL: 1.8 μL), 5:1 (16.7 μL: 3.3 μL), 3:1 (15.0 μL: 5.0 μL), 2:1 (13.4 μL: 6.6 μL), 1:1 (10.0 μL: 10.0 μL), 1:2 (6.6 μL: 13.4 μL), 1:3 (5.0 μL: 15.0 μL), 1:5 (3.3 μL: 16.7 μL), 1:10 (1.8 μL: 18.2 μL). The mixtures were incubated at RT and analyzed at different timepoints by SDS-PAGE.

The scFv addition was also performed at 37 °C with linker DTPMCL or 46, or at pH 8.0 (BBS buffer), where both HSA-linker intermediate and reduced scFv were buffered exchanged to BBS buffer (pH 8.0) via ultrafiltration (10 kDa MWCO) and then mixed at RT in the ratios previously mentioned.

**Modification of HSA with DBMCL, followed by conjugation to octreotide**

To HSA (350 μL, 0.013 μmol, 2.47 mg/mL, 37.2 μM) in conjugation buffer was added linker DBMCL (1.8 μL, 0.067 μmol, 37.2 mM solution in DMF, 5 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 15.0 μM (678 μL, 1.0 mg/mL). Meanwhile, octreotide (50 μL, 0.075 μmol, 1.53 μg/mL, 1.5 mM) in conjugation buffer was reduced with TCEP (3.75 μL, 0.11 μmol, 30.0 mM solution in H2O, 1.5 eq.) for 1 h at 37 °C. Modified HSA was then split into 3 aliquots (150 μL each, 15.0 μM, 1 eq.) and different equivalents of reduced octreotide (used without further purification) were added into each one, i.e. (0.75 μL, 1.5 mM, 0.5 eq.), (1.5 μL, 1.5 mM, 1 eq.), (3.0 μL, 1.5 mM, 2 eq.). After 1 h at RT, the reaction mixtures were analyzed by LC-MS.
Modification of anti-CEA scFv with bis-dithiophenolmaleimide 6 kDa PEG linker 41, followed by conjugation with HSA

Anti-CEA scFv (400 μL, 0.015 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (8.0 μL, 0.30 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was removed using a desalting column (PD MiniTrap G-25, GE Healthcare) and the reduced scFv was buffer exchanged into conjugation buffer, via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (310 μL, 1.0 mg/mL). Linker 41 (18 μL, 0.33 μmol, 18.6 mM solution in DMF, 30 eq.) was then added to reduced scFv (300 μL, 0.011 μmol, 37.2 μM). After 20 min at RT, the scFv-linker solution was purified by size exclusion chromatography into conjugation buffer (24 mL Superdex™ 200 10/300 GL, GE Healthcare, sample volume 300 μL, loop volume 1 mL, flowrate 0.5 mL/min) to separate only the bridged scFv species (110 μL, 37.2 μM, 1.0 mg/mL). The modified scFv was then split into 2 aliquots (50 μL each, 37.2 μM, 1 eq.) and different equivalents of HSA were added into each one, i.e. (50 μL, 37.2 μM, 1 eq.), (16.6 μL, 37.2 μM, 0.33 eq.). The mixture was incubated at RT and analyzed at different timepoints by SDS-PAGE.

Modification of anti-CEA scFv with DIMCL, followed by conjugation with HSA

Anti-CEA scFv (400 μL, 0.015 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (8.0 μL, 0.30 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was removed using a desalting column (PD MiniTrap G-25, GE Healthcare) and the reduced scFv was buffer exchanged into conjugation buffer, via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (280 μL, 1.0 mg/mL). To reduced scFv (200 μL, 0.0074 μmol, 37.2 μM) was then added linker DIMCl (6.0 μL, 0.22 μmol, 37.2 mM solution in DMF, 30 eq.). After 10 min at RT, the excess linker was removed via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (140 μL, 1.0 mg/mL). Modified scFv (37.2 μM, 1.0 mg/mL) was then mixed with HSA (37.2 μM, 2.47 mg/mL) in conjugation buffer in the following ratios: (HSA:scFv) 5:1 (16.6 μL: 3.3 μL), 3:1 (15.0 μL: 5.0 μL), 2:1 (13.4 μL: 6.6 μL), 1:1 (10.0 μL: 10.0 μL), 1:2 (6.6 μL: 13.4 μL), 1:3 (5.0 μL: 15.0 μL), 1:5 (3.3 μL: 16.6 μL). The mixture was incubated at RT and analyzed at different timepoints by SDS-PAGE.
Modification of HSA with DIM_{CL}, followed by conjugation with scFv

To HSA (380 μL, 0.014 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added linker DIM_{CL} (2.3 μL, 0.086 μmol, 37.2 mM solution in DMF, 6 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (340 μL, 2.47 mg/mL). Meanwhile, anti-CEA scFv (200 μL, 0.0074 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (4.0 μL, 0.15 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MiniTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (140 μL, 1.0 mg/mL). Modified HSA (37.2 μM) was then mixed with reduced scFv (37.2 μM) in the following ratios: (HSA:scFv) 5:1 (16.6 μL: 3.3 μL), 3:1 (15.0 μL: 5.0 μL), 2:1 (13.4 μL: 6.6 μL), 1:1 (10.0 μL: 10.0 μL), 1:2 (6.6 μL: 13.4 μL), 1:3 (5.0 μL: 15.0 μL), 1:5 (3.3 μL: 16.6 μL). The mixture was incubated at RT and analyzed at different timepoints by SDS-PAGE.

General procedure for the modification of HSA with DIM_{CL-C6} (or DBM_{CL-C6}) followed by conjugation with scFv (conjugation at 37.2 μM)

To HSA (300 μL, 0.011 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added linker DIM_{CL-C6} or DBM_{CL-C6} (0.9 μL, 0.033 μmol, 37.2 mM solution in DMF, 3 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (280 μL, 2.47 mg/mL). Meanwhile, anti-CEA scFv (400 μL, 0.015 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (8.0 μL, 0.30 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MiniTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (220 μL, 1.0 mg/mL). To modified HSA (37.2 μM, 2.47 mg/mL) was then added reduced scFv (37.2 μM, 1.0 mg/mL) in the following ratios, by
volume: (HSA:scFv) 3:1 (15.0 μL: 5.0 μL), 2:1 (13.4 μL: 6.6 μL), 1:1 (10.0 μL: 10.0 μL). The mixtures were incubated at RT and analysed at different timepoints by SDS-PAGE.

**Optimized synthesis of HSA-scFv conjugate**

To HSA (3000 μL, 0.11 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added linker **DIMCL-C6** (9.0 μL, 0.33 μmol, 37.2 mM solution in DMF, 3 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 111.6 μM (973 μL, 7.55 mg/mL). Meanwhile, anti-CEA scFv (1000 μL, 0.037 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (20.0 μL, 0.74 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MidiTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (905 μL, 1.0 mg/mL). To modified HSA (900 μL, 0.10 μmol, 111.6 μM, 3 eq.) was added reduced scFv (900 μL, 0.033 μmol, 37.2 μM, 1 eq.) and the reaction was incubated for 4 h at RT. After this period, the mixture was buffer exchanged to BBS buffer (pH 8.0) via ultrafiltration (10 kDa MWCO) and incubated for 72 h at 37 °C.

The reaction mixture was then subjected to Ni-IMAC purification (HisTrap™ HP 1 mL column (GE Healthcare), binding buffer (20 mM phosphates, 500 mM NaCl, 10 mM imidazole, pH 7.5), washing buffer (20 mM phosphates, 500 mM NaCl, 20 mM imidazole, pH 7.5), elution buffer (20 mM phosphates, 500 mM NaCl, 200 mM imidazole, pH 7.5). The collected fractions were then purified by size exclusion chromatography (125 mL Superdex™ 75 prep grade XK 26/100, GE Healthcare, sample volume 800 μL, loop volume 2 mL, flowrate 1 mL/min) to afford HSA-scFv, the amount of which was determined by UV/Vis absorbance (1.8 mg, 0.019 μmol, 57%). The resultant conjugate was characterized by SDS-PAGE, SEC, LC-MS and ELISA.
**Proteomics of HSA-dibromomaleimide (N-Me and N-Biotin) conjugates**

**Preparation of HSA conjugates for tryptic digestion**

To HSA (2000 μL, 0.074 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added dibromomaleimide (N-Me or N-Biotin 54; 3.0 μL, 0.11 μmol, 37.2 mM solution in DMF, 1.5 eq.) and the reaction was incubated for 20 min at RT. The excess reagent was then removed *via* ultrafiltration (30 kDa MWCO) into conjugation buffer and the protein was buffer exchanged into BBS buffer (pH 8.0). The reaction mixture was incubated for 72 h at 37 °C.

**Tryptic digestion – general protocol**

Native HSA or HSA modified with N-Me or N-Biotin dibromomaleimide 54 were buffered exchanged into H$_2$O (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 752 μM (90 μL, 50 mg/mL). 20 μL (0.015 μmol) of this solution were diluted to 100 μL with Tris buffer (50 mM Tris, 6M Guanidine.HCl, 2 mM EDTA, pH 8.0). DTT (5 μL, 1.0 μmol, 200 mM solution in Tris buffer (100 mM Tris, pH 8.0)) was then added. After 75 min at 37 °C, iodoacetamide (10 μL, 2.0 μmol, 200 mM solution in Tris buffer (100 mM Tris, pH 8.0)) was added and the mixture was incubated for further 75 min at 37 °C, in the dark. The reaction was diluted with 100 μL H$_2$O and 400 μL Tris buffer (50 mM Tris, pH 8.0), before trypsin (Pierce™ Trypsin Protease, MS Grade, 10 μL, 0.00043 μmol, 42.9 μM, 1.0 mg/mL solution in 10 mM HCl) was added. The resultant mixture was incubated for 16 h at 37 °C with shaking (300 rpm). After this period, the reaction was centrifuged and stopped by the addition of TFA (1% in H$_2$O, 625 μL). The peptide mixture was then purified using a C18 cartridge (Sep-Pak C18, 360 mg sorbent); the sample was firstly washed with 10 mL of 98% H$_2$O, 2% MeCN, 0.1% formic acid solution and eluted into 3 mL of 30% H$_2$O, 70% MeCN, 0.1% formic acid solution. It was then concentrated using a Speedvac concentrator and analysed by LC-MS and LC-MS/MS.
BME treatment of HSA modified with N-Me dibromomaleimide, before and after incubation at pH 8, 37 °C, 72 h

HSA modified with N-Me dibromomaleimide before and after incubation at pH 8, 37 °C, 72 h (120 μL, 0.0046 μmol, 37.2 μM, 2.47 mg/mL) in BBS buffer (pH 8.0) was treated with BME (12.0 μL, 0.446 μmol, 37.2 mM solution in BBS buffer (pH 8.0), 100 eq.) for 2 h at 37 °C. After this period, the samples were buffer exchanged to H2O (10 kDa MWCO) and submitted for LC-MS analysis.

General procedure for the preparation of Trastuzumab (Herceptin™) Fab fragment

Trastuzumab (0.50 mL, 44.0 μM, 6.41 mg/mL) was buffer exchanged into NaOAc buffer (20 mM NaOAc, pH 3.1) via ultrafiltration (10 kDa MWCO). Immobilized pepsin (0.15 mL) was also washed with NaOAc buffer (20 mM NaOAc, pH 3.1), using a Pierce™ centrifuge column. Trastuzumab (0.50 mL) was then added to the pepsin and the mixture was incubated at 37 °C for 5 h, under constant agitation (1,100 rpm). The resin was then removed from the digest using a Pierce™ centrifuge column and washed with digest buffer (3 × 0.4 mL, 50 mM phosphate, 150 mM NaCl, 1 mM EDTA, pH 6.8). The digest was combined with the washes and the volume was adjusted to 0.5 mL. Next, immobilized papain (0.65 mL) was activated with 10 mM DTT in digest buffer (50 mM phosphate, 150 mM NaCl, 1 mM EDTA, pH 6.8). This was incubated at 37 °C for 90 min under constant agitation (1,100 rpm). The resin was washed with digest buffer (4 × 0.4 mL, 50 mM phosphate, 150 mM NaCl, 1 mM EDTA, pH 6.8) without DTT, using a Pierce™ centrifuge column. F(ab′)2 (0.5 mL) was added to the washed papain and the mixture incubated at 37 °C for 20 h under constant agitation (1,100 rpm). The resin was separated and washed with conjugation buffer (4 × 0.4 mL, 40 mM phosphate, 20 mM NaCl, 6 mM EDTA, pH 7.4). The digest was combined with the washes and the buffer was exchanged completely for conjugation buffer via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 150 μM. Aliquots of Fab were stored at -20 °C for up to 6 months.
General procedure for the modification of HSA with DIM<sub>CL</sub> (or DBM<sub>CL</sub> or DBM<sub>CL-C6</sub> or DIM<sub>CL-C6</sub>) followed by conjugation to Fab (under low salt conditions (20 mM NaCl) and 37.2 μM)

To HSA (300 μL, 0.011 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added linker DIM<sub>CL</sub> or DBM<sub>CL</sub> (1.8 μL, 0.066 μmol, 37.2 mM solution in DMF, 6 eq.) or DIM<sub>CL-C6</sub> or DBM<sub>CL-C6</sub>, (0.9 μL, 0.033 μmol, 37.2 mM solution in DMF, 3 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (240 μL, 2.47 mg/mL). Meanwhile, anti-HER2 Fab (200 μL, 0.0074 μmol, 37.2 μM, 1.76 mg/mL) in conjugation buffer was reduced with TCEP (2.0 μL, 0.074 μmol, 37.2 mM solution in H<sub>2</sub>O, 10 eq.) for 1 h at 37 °C. The excess TCEP was then removed using a desalting column (PD MiniTrap™ G-25, GE Healthcare) and the reduced Fab was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (150 μL, 1.76 mg/mL). Meanwhile, anti-HER2 Fab (500 μL, 0.019 μmol, 37.2 μM, 1.76 mg/mL) in conjugation buffer was reduced with TCEP (5.0 μL, 0.19 μmol, 37.2 mM solution in H<sub>2</sub>O, 10 eq.) for 1 h at 37 °C. The excess TCEP was then removed using a desalting column (PD MiniTrap™ G-25, GE Healthcare) and the reduced Fab was concentrated via ultrafiltration (10 kDa MWCO) and buffer exchanged into conjugation buffer containing 500 mM NaCl. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (479 μL, 1.76 mg/mL). To modified HSA (450 μL, 0.050 μmol, 111.6 μM, 3 eq.) was

Optimized synthesis of HSA-Fab conjugate

To HSA (1500 μL, 0.056 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added linker DIM<sub>CL</sub> (9.0 μL, 0.33 μmol, 37.2 mM solution in DMF, 6 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer containing 500 mM NaCl. The concentration was determined by UV/Vis absorbance and adjusted to 111.6 μM (467 μL, 7.55 mg/mL). Meanwhile, anti-HER2 Fab (500 μL, 0.019 μmol, 37.2 μM, 1.76 mg/mL) in conjugation buffer was reduced with TCEP (5.0 μL, 0.19 μmol, 37.2 mM solution in H<sub>2</sub>O, 10 eq.) for 1 h at 37 °C. The excess TCEP was then removed using a desalting column (PD MiniTrap™ G-25, GE Healthcare) and the reduced Fab was concentrated via ultrafiltration (10 kDa MWCO) and buffer exchanged into conjugation buffer containing 500 mM NaCl. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (479 μL, 1.76 mg/mL). To modified HSA (450 μL, 0.050 μmol, 111.6 μM, 3 eq.) was
added reduced Fab (450 μL, 0.017 μmol, 37.2 μM, 1 eq.) and the reaction was incubated for 1 h at RT. After this period, the mixture was buffer exchanged to BBS buffer (pH 8.0) via ultrafiltration (10 kDa MWCO) and incubated for 72 h at 37 °C. The reaction mixture was then purified by size exclusion chromatography (24 mL Superdex™ 200 10/300 GL, GE Healthcare, sample volume 200 μL, loop volume 500 μL, flowrate 0.2 mL/min) to afford HSA-Fab, the amount of which was determined by UV/Vis absorbance (0.93 mg, 0.0081 μmol, 49%). The resultant conjugate was characterized by SDS-PAGE, SEC, LC-MS and ELISA.

Serum stability study

HSA-scFv and HSA-Fab conjugates (80 nM in conjugation buffer) were diluted 1:1 with human serum (male AB plasma, Sigma Aldrich) at a final concentration of 40 nM in sterile tubes. Each conjugate mixture was split into 8 equal aliquots (500 μL each). One aliquot was immediately frozen at -80 °C. The remaining samples were incubated at 37 °C. Aliquots were removed from the incubator after 1, 2, 3, 4, 5, 6 and 7 days and transferred to a -80 °C freezer. After the final time point, samples were analysed by sandwich ELISA, using the protocols described in section 3.3.1.4; the only difference was that the samples were placed directly on the plate, without any dilutions being carried out.

3.3.1.2 Trifunctional conjugation

Attempted conjugation of HSA to two scFv monomers

To HSA (200 μL, 0.0074 μmol, 2.47 mg/mL, 37.2 μM) in conjugation buffer was added tDIMCL-c6 (3.0 μL, 0.11 μmol, 37.2 mM solution in DMF, 15 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (150 μL, 2.47 mg/mL). Meanwhile, anti-CEA scFv (280 μL, 0.010 μmol, 1.0 mg/mL, 37.2 μM) in conjugation buffer was reduced with DTT (5.6 μL, 0.21 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MiniTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by
UV/Vis absorbance and adjusted to 37.2 μM (200 μL, 1.0 mg/mL). Modified HSA was then split into 4 aliquots (20 μL each, 37.2 μM, 1 eq.) and different equivalents of reduced scFv were added into each one, i.e. (20 μL, 37.2 μM, 1 eq.), (40 μL, 37.2 μM, 2 eq), (60 μL, 37.2 μM, 3 eq.), (80 μL, 37.2 μM, 4 eq.). The reaction mixtures were analyzed at different timepoints by SDS-PAGE.

**Attempted stepwise conjugation of HSA to two scFv monomers (or scFv-Fab), with purification of HSA-scFv intermediate by HisTrap™**

To HSA (1500 μL, 0.056 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added tDIMCl-c6 (22.5 μL, 0.84 μmol, 37.2 mM solution in DMF, 15 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 74.4 μM (700 μL, 4.94 mg/mL). Meanwhile, anti-CEA scFv (850 μL, 0.032 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (17.0 μL, 0.63 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MidiTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (800 μL, 1.0 mg/mL). To modified HSA (700 μL, 0.052 μmol, 74.4 μM, 4.94 mg/mL, 2 eq.) was then added reduced scFv (700 μL, 0.026 μmol, 37.2 μM, 1.0 mg/mL, 1 eq.) and the reaction was incubated for 1 h at RT. The reaction mixture was then subjected to Ni-IMAC purification (HisTrap HP 1 mL column (GE Healthcare), binding buffer (20 mM phosphates, 500 mM NaCl, 10 mM imidazole, pH 7.5), washing buffer (20 mM phosphates, 500 mM NaCl, 20 mM imidazole, pH 7.5), elution buffer (20 mM phosphates, 500 mM NaCl, 200 mM imidazole, pH 7.5). The collected fractions were then buffer exchanged to conjugation buffer via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 20.6 μM (600 μL, 1.9 mg/mL). Meanwhile, anti-HER2 Fab (150 μL, 0.0056 μmol, 37.2 μM, 1.76 mg/mL) in conjugation buffer was reduced with TCEP (1.5 μL, 0.056 μmol, 37.2 mM solution in H2O, 10 eq.) for 1 h at 37 °C. The excess TCEP was then removed using a desalting column (PD MiniTrap™ G-25, GE Healthcare) and the reduced Fab was concentrated via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 20.6 μM (150 μL, 0.98 mg/mL). To HSA-scFv after HisTrap purification (300 μL,
0.0062 μmol, 20.6 μM, 1.9 mg/mL, 2 eq.) was added either freshly reduced scFv (150 μL, 0.0031 μmol, 20.6 μM, 0.55 mg/mL, 1 eq.) or reduced Fab (150 μL, 0.0031 μmol, 20.6 μM, 0.98 mg/mL, 1 eq.) and the resultant mixtures were incubated at RT. The reaction mixtures were analyzed at different timepoints by SDS-PAGE analysis.

**Attempted stepwise conjugation of HSA to scFv and Fab with the 2nd and 3rd additions at 100 μM**

To HSA (800 μL, 0.030 μmol, 37.2 μM, 2.47 mg/mL) in conjugation buffer was added tDIMcL-c6 (12 μL, 0.45 μmol, 37.2 mM solution in DMF, 15 eq.) and the reaction was incubated for 15 min at RT. The excess linker was then removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 100 μM (298 μL, 6.65 mg/mL). Meanwhile, anti-CEA scFv (800 μL, 0.030 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (16 μL, 0.60 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MidiTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 100 μM (240 μL, 2.67 mg/mL). To modified HSA (245 μL, 0.025 μmol, 100 μM, 6.65 mg/mL, 1 eq.) was then added reduced scFv (220 μL, 0.022 μmol, 100 μM, 2.67 mg/mL, 0.9 eq.) and the reaction was incubated for 1 h at RT. Meanwhile, anti-HER2 Fab (800 μL, 0.030 μmol, 37.2 μM, 1.76 mg/mL) in conjugation buffer was reduced with TCEP (8.0 μL, 0.30 μmol, 37.2 mM solution in H2O, 10 eq.) for 1 h at 37 °C. The excess TCEP was then removed using a desalting column (PD MidiTrap™ G-25, GE Healthcare) and the reduced Fab was concentrated via ultrafiltration (10 kDa MWCO). The concentration was determined by UV/Vis absorbance and adjusted to 100 μM (250 μL, 4.76 mg/mL). To the HSA-scFv mixture was then added reduced Fab (245 μL, 0.025 μmol, 100 μM, 4.76 mg/mL, 1 eq.) and the resultant solution was incubated at RT. The reaction mixture was analyzed at different timepoints by SDS-PAGE.
Optimized synthesis of anti-CEA tri-scFv conjugate

Anti-CEA scFv (2800 μL, 0.10 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (56.0 μL, 2.08 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MidiTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 37.2 μM (2550 μL, 0.095 μmol, 1.0 mg/mL). To this was added linker tDIMcl (8.42 μL, 0.031 μmol, 3.72 mM solution in DMF, 0.33 eq.) and the reaction was incubated for 1 h at RT. After this period, the mixture was buffer exchanged into BBS buffer (pH 8.0) via ultrafiltration (10 kDa MWCO) and incubated for 72 h at 37 °C. The reaction mixture was then purified by size exclusion chromatography (24 mL Superdex™ 75 10/300 GL and 24 mL Superdex™ 200 10/300 GL in series, GE Healthcare, sample volume 200 μL, loop volume 500 μL, flowrate 0.05 mL/min) to afford tri-scFv, the amount of which was determined by UV/Vis absorbance (0.77 mg, 0.0094 μmol, 30%). The resultant conjugate was characterized by SDS-PAGE, SEC, LC-MS and ELISA.

Tri-scFv and scFv biodistribution protocol

Performed by Dr Vineeth Rajkumar at UCL Cancer Institute

Tri-scFv and scFv samples were incubated with Chelex100 (BioRad) for 16 h at 4 °C. The samples were then filtered (Corning™ Costar™ Spin-X™ Centrifuge Tube Filters) and buffered exchanged (10 kDa MWCO) to phosphate buffer (0.1 M Na₂HPO₄, pH 7.5, pre-incubated with Chelex100). To tri-scFv (215 μL, 0.0061 μmol, 28.5 μM, 2.33 mg/mL) or scFv (658 μL, 0.019 μmol, 28.5 μM, 0.76 mg/mL) was added DOTA-NHS ester (Macrocyclics, 4.73 μL (0.13 μmol) or 14.5 μL (0.41 μmol) respectively, 28.5 mM solution in DMF, 22 eq.). The reactions were incubated at 4 °C for 24 h with continuous mixing (300 rpm). After this period, the conjugates were buffer exchanged (10 kDa MWCO) to ammonium acetate buffer (0.2 M NH₄OAc, pH 6.0) and diluted to 2.0 mg/mL. The radiolabeling procedure was a modification of the protocol described by Cooper et al. Conjugates were incubated with an excess of ¹¹¹InCl₃ (Mallinckrodt) for 1 h at 37 °C. The reaction was stopped by addition of 0.1 M EDTA and free ¹¹¹InCl₃ was removed using a PD-10 desalting column (GE Healthcare). Radiochemical purity was analysed by instant thin layer chromatography.
(iTLC) using iTLC-silica gel (SG) strips (Varian), with a resulting labelling efficiency of above 95% for both compounds. A single dose of 330 picomoles of $^{111}$In-labelled conjugates was injected into BALB/c mice (Charles River) by tail injection (3 mice per group), equivalent to 0.33 MBq of monomer and 0.44 MBq of trimer. For clearance studies blood was collected at 10, 20, 40, 60 and 180 min; organs were collected 3 h and 24 h after injection to assess radiation uptake. All experiments were conducted with approval of the UK Home Office under PPL 70/6677.

**Optimized synthesis of anti-CEA/ anti-CD3/ anti-CD3 scFv heterotrimer**

Anti-CEA scFv (1400 μL, 0.052 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (28.0 μL, 1.04 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MidiTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 62.0 μM (650 μL, 1.7 mg/mL). To conjugation buffer (650 μL) was added linker tDIMCt-C6 (32.5 μL, 1.21 μmol, 37.2 mM solution in DMF, 30 eq.), followed by DMF (40 μL) to make up a final DMF concentration of 10%. Anti-CEA scFv (650 μL, 0.040 μmol, 62.0 μM, 1.7 mg/mL) was then mixed with DMF (70 μL) to make up a final DMF concentration of 10%. This solution was added dropwise to tDIMCt-C6 in the buffer/ DMF solution, over 2 min at RT. After further 6 min at RT, the excess linker was removed via ultrafiltration (30 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 50.0 μM (750 μL, 1.4 mg/mL). Meanwhile, anti-CD3 scFv (2000 μL, 0.074 μmol, 37.2 μM, 1.0 mg/mL) in conjugation buffer was reduced with DTT (40 μL, 1.49 μmol, 37.2 mM solution in conjugation buffer, 20 eq.) for 1 h at RT. The excess DTT was then removed using a desalting column (PD MidiTrap™ G-25, GE Healthcare) and the reduced scFv was concentrated via ultrafiltration (10 kDa MWCO) into conjugation buffer. The concentration was determined by UV/Vis absorbance and adjusted to 50.0 μM (1300 μL, 1.4 mg/mL). To bridged anti-CEA scFv (650 μL, 0.033 μmol, 50 μM, 1 eq.) was added reduced anti-CD3 scFv (1300 μL, 0.065 μmol, 50 μM, 2 eq.) and the reaction was incubated for 2 h at RT. After this period, the mixture was buffer exchanged to BBS buffer (pH 8.0) via ultrafiltration (10 kDa MWCO) and incubated for 72 h at 37 °C. The reaction mixture was then purified by size exclusion chromatography (125 mL Superdex™ 75 prep grade XK 26/100, GE Healthcare, sample volume
900 μL, loop volume 2 mL, flowrate 0.2 mL/min) to afford anti-CEA/anti-CD3/anti-CD3 scFv heterotrimer, the amount of which was determined by UV/Vis absorbance (0.28 mg, 0.0033 μmol, 10%). The resultant conjugate was characterized by SDS-PAGE and LC-MS.

### 3.3.1.3 Cysteine-to-lysine transfer (CLT)

#### Optimized conditions for transthioesterification of anti-HER2 Fab with thioester 59

To anti-HER2 Fab (40 μL, 0.0060 μmol, 150 μM, 7.15 mg/mL) in conjugation buffer was added TCEP (0.4 μL, 0.060 μmol, 150 mM solution in H₂O, 10 eq.). After 1 h at 37 °C, thioester 59 (2.0 μL, 0.60 μmol, 300 mM solution in DMF, 100 eq.) was added to the conjugation and the resultant mixture was incubated at 22 °C for 4 h. The excess thioester was then removed via ultrafiltration (10 kDa MWCO) into H₂O and the sample submitted for LC-MS analysis, where it was kept at 4 °C until analyzed.

#### Control reaction of anti-HER2 Fab with thioester 59

To anti-HER2 Fab (40 μL, 0.0060 μmol, 150 μM, 7.15 mg/mL) in conjugation buffer was added thioester 59 (2.0 μL, 0.60 μmol, 300 mM solution in DMF, 100 eq.) and the resultant mixture was incubated at 22 °C for 4 h. The excess thioester was then removed via ultrafiltration (10 kDa MWCO) into H₂O and the sample submitted for LC-MS analysis.

#### BME treatment of anti-HER2 Fab after transthioesterification with thioester 59

To anti-HER2 Fab (52.0 μL, 0.0078 μmol, 150 μM, 7.15 mg/mL) in conjugation buffer was added TCEP (0.52 μL, 0.078 μmol, 150 mM solution in H₂O, 10 eq.). After 1 h at 37 °C, thioester 59 (2.6 μL, 0.78 μmol, 300 mM solution in DMF, 100 eq.) was added to the conjugation and the resultant mixture was incubated at 22 °C for 4 h. The excess thioester was then removed via ultrafiltration (10 kDa MWCO) into conjugation buffer and then buffer exchanged into BBS buffer (pH 8.0). The concentration was determined by UV/Vis absorbance and adjusted to 20 μM (335 μL, 0.0067 μmol, 0.95 mg/mL). To this solution of Fab was added BME (33.5 μL, 0.67 μmol, 20 mM solution
in BBS buffer (pH 8.0), 100 eq.). After 2 h at 37 °C, the excess BME was removed via ultrafiltration into H₂O (10 kDa MWCO) and the sample submitted for LC-MS analysis.

**Optimized conditions for acyl transfer**

To anti-HER2 Fab (330 μL, 0.0495 μmol, 150 μM, 7.15 mg/mL) in conjugation buffer was added TCEP (3.30 μL, 0.495 μmol, 150 mM solution in H₂O, 10 eq.). After 1 h at 37 °C, thioester 59 (16.5 μL, 4.95 μmol, 300 mM solution in DMF, 100 eq.) was added in the conjugation and the resultant mixture was incubated at 22 °C for 4 h. The excess thioester was then removed via ultrafiltration (10 kDa MWCO) into conjugation buffer and buffer exchanged into BBS buffer (pH 8.4). The concentration was determined by UV/Vis absorbance and adjusted to 20 μM (2068 μL, 0.95 mg/mL). The resultant solution was incubated for 72 h at 12 °C. After this period, the mixture was buffer exchanged into BBS (pH 8.0), via ultrafiltration (10 kDa MWCO). A small amount of sample was buffer exchanged into H₂O for LC-MS analysis and size exclusion chromatography (24 mL Superdex™ 200 10/300 GL, GE Healthcare, sample volume 200 μL, loop volume 1 mL, flowrate 0.25 mL/min).

**BME treatment of anti-HER2 Fab after acyl transfer**

To anti-HER2 Fab after acyl transfer (50.0 μL, 0.001 μmol, 0.95 mg/mL, 20 μM) in BBS buffer (pH 8.0) was added BME (5.0 μL, 0.10 μmol, 20 mM solution in BBS buffer (pH 8.0), 100 eq.). After 2 h at 37 °C, the excess BME was removed via ultrafiltration (10 kDa MWCO) into H₂O and the sample submitted for LC-MS analysis.

**CuAAC of anti-HER2 Fab after acyl transfer with Alexa Fluor 488**

To a solution of anti-HER2 Fab after acyl transfer (150 μL, 0.0030 μmol, 20 μM, 0.95 mg/mL) in BBS buffer (pH 8.0) was added THPTA ligand (1.8 μL, 0.18 μmol, 100 mM solution in H₂O), followed by CuSO₄ (1.8 μL, 0.036 μmol, 20 mM solution in H₂O). To this mixture was added Alexa Fluor 488 azide (6.0 μL, 0.060 μmol, 10 mM solution in DMF, 20 eq.), followed by sodium ascorbate (18.2 μL, 1.82 μmol, 100 mM solution in H₂O). The resultant mixture was incubated for 4 h at 37 °C. Following this period, excess reagents were removed using a desalting column (PD
MiniTrap™ G-25, GE Healthcare) and repeated ultrafiltration (10 kDa MWCO) into conjugation buffer. The sample was then analyzed by UV/Vis spectroscopy and SDS-PAGE.

### 3.3.1.4 Enzyme Linked Immunosorbent Assay (ELISA)

#### Protocol for CEA ELISA

A 96-well plate was coated for 1 h at RT with full length human CEA (Calbiochem, 100 μL/well, 1 μg/mL solution in PBS). After washing (3 × 0.1% Tween® 20 in PBS, followed by 3 × PBS), the wells were blocked for 1 h at RT with 5% Marvel milk powder (Premier foods) in PBS (200 μL/well). The wells were then washed and the following dilutions of HSA-scFv and scFv or tri-scFv and scFv: 270 nM, 90 nM, 30 nM, 10 nM, 3.33 nM, 1.11 nM, 0.37 nM, 0.123 nM, 0.0412 nM, 0.0137 nM, prepared in 1% Marvel solution in 0.1% Tween® 20 in PBS (100 μL/well). The assay was then incubated at RT for 1 h, washed and a primary antibody (anti-tetra-His mouse IgG1, Quiagen, 1:1,000 in 1% Marvel solution in 0.1% Tween® 20 in PBS) was added (100 μL/well). After 1 h at RT, the plate was washed and the detection antibody (ECL anti-mouse sheep IgG1 HRP linked, GE Healthcare, 1:1,000 in 1% Marvel solution in 0.1% Tween® 20 in PBS) was added (100 μL/well) and incubated at RT for 1 h. The plate was washed again and o-phenylenediamine hydrochloride (Sigma-Aldrich, 100 μL/well, 0.5 mg/mL in a phosphate-citrate buffer with sodium perborate) was added. Once a yellow-orange colour was observed, the reaction was stopped by addition of HCl (4M, 50 μL/well). Absorbance was immediately measured at 450 nm and was corrected by subtracting the average of negative controls (i.e. PBS had been added to some of the wells instead of CEA or instead of the samples). Each sample was tested in triplicate and errors are shown as the standard deviation of the average. ELISA data was analysed with Graphpad Prism 7.03 (using equation Sigmoidal, 4PL, X is log(concentration)) and the values have been normalized. For calculation of IC₅₀, concentration values (in nM) were transformed into log(concentration).

#### Protocol for Sandwich CEA/ anti-albumin ELISA

Two 96-well plates were coated for 1 h at RT with full length human CEA (Calbiochem, 100 μL/well, 1 μg/mL solution in PBS). After washing (3 × 0.1% Tween® 20 in PBS, followed by 3
× PBS), the wells were blocked for 1 h at RT with 5% Marvel milk powder (Premier foods) in PBS (200 μL/well). The wells were then washed and the following dilutions of HSA-scFv, HSA and scFv were applied: 270 nM, 90 nM, 30 nM, 10 nM, 3.33 nM, 1.11 nM, 0.37 nM, 0.123 nM, 0.0412 nM, 0.0137 nM, prepared in 1% Marvel solution in 0.1% Tween® 20 in PBS (100 μL/well). The assay was then incubated at RT for 1 h, washed and a primary antibody (anti-albumin mouse IgG1, ThermoFisher Scientific, 1:1000 in 1% Marvel solution in 0.1% Tween® 20 in PBS) was added (100 μL/well). After 1 h at RT, the plates were washed and the detection antibody (ECL anti-mouse sheep IgG1 HRP linked, GE Healthcare, 1:1,000 in 1% Marvel solution in 0.1% Tween® 20 in PBS) was added (100 μL/well) and incubated at RT for 1 h. The plate was washed again and o-phenylenediamine hydrochloride (Sigma-Aldrich, 100 μL/well, 0.5 mg/mL in a phosphate-citrate buffer with sodium perborate) was added. Once a yellow-orange colour was observed, the reaction was stopped by addition of HCl (4 M, 50 μL/well). Absorbance was immediately measured at 450 nm and was corrected by subtracting the average of negative controls (i.e. PBS had been added to some of the wells instead of CEA or instead of the samples). Each sample was tested in triplicate and errors are shown as the standard deviation of the average. ELISA data was analysed with Graphpad Prism 7.03 (using equation Sigmoidal, 4PL, X is log(concentration)) and the values have been normalized. For calculation of IC₅₀, concentration values (in nM) were transformed into log(concentration).

**Protocol for HER2 ELISA**

A 96-well plate was coated for 1 h at RT with HER2 (Sino Biological, 100 μL/well, 0.25 μg/mL solution in PBS). After washing (3 × 0.1% Tween® 20 in PBS, followed by 3 × PBS), the wells were blocked for 1 h at RT with 5% Marvel milk powder (Premier foods) in PBS (200 μL/well). The wells were then washed and the following dilutions of HSA-Fab and Fab or Fab after acyl transfer and Fab were applied: 270 nM, 90 nM, 30 nM, 10 nM, 3.33 nM, 1.11 nM, 0.37 nM, 0.123 nM, 0.0412 nM, 0.0137 nM, prepared in 1% Marvel solution in 0.1% Tween® 20 in PBS (100 μL/well). The assay was then incubated at RT for 1 h, washed and the detection antibody (Anti-Human IgG, Fab specific-HRP antibody, Sigma Aldrich, 1:5000 in 1% Marvel solution in 0.1% Tween® 20 in PBS) was added (100 μL/well). After 1 h at RT, the plates were washed and o-phenylenediamine hydrochloride (Sigma-Aldrich, 100 μL/well, 0.5 mg/mL in a phosphate-citrate buffer with sodium perborate) was added. Once a yellow-orange colour was observed, the reaction
was stopped by addition of HCl (4M, 50 μL/well). Absorbance was immediately measured at 450 nm and was corrected by subtracting the average of negative controls (i.e. PBS had been added to some of the wells instead of HER2 or instead of the samples). Each sample was tested in triplicate and errors are shown as the standard deviation of the average. ELISA data was analysed with Graphpad Prism 7.03 (using equation Sigmoidal, 4PL, X is log(concentration)) and the values have been normalized. For calculation of IC$_{50}$, concentration values (in nM) were transformed into log(concentration).

**Protocol for Sandwich HER2/ anti-albumin ELISA**

Two 96-well plates were coated for 1 h at RT with HER2 (Sino Biological, 100 μL/well, 0.25 μg/mL solution in PBS). After washing (3 × 0.1% Tween® 20 in PBS, followed by 3 × PBS), the wells were blocked for 1 h at RT with 5% Marvel milk powder (Premier foods) in PBS (200 μL/well). The wells were then washed and the following dilutions of HSA-Fab, HSA and Fab were applied: 270 nM, 90 nM, 30 nM, 10 nM, 3.33 nM, 1.11 nM, 0.37 nM, 0.123 nM, 0.0412 nM, 0.0137 nM, prepared in 1% Marvel solution in 0.1% Tween® 20 in PBS (100 μL/well). The assay was then incubated at RT for 1 h, washed and a primary antibody (anti-albumin mouse IgG1, ThermoFisher Scientific, 1:1000 in 1% Marvel solution in 0.1% Tween® 20 in PBS) was added (100 μL/well). After 1 h at RT, the plates were washed and the detection antibody (ECL anti-mouse sheep IgG1 HRP linked, GE Healthcare, 1:1,000 in 1% Marvel solution in 0.1% Tween® 20 in PBS) was added (100 μL/well) and incubated at RT for 1 h. The plates were washed again and o-phenylenediamine hydrochloride (Sigma-Aldrich, 100 μL/well, 0.5 mg/mL in a phosphate-citrate buffer with sodium perborate) was added. Once a yellow-orange colour was observed, the reaction was stopped by addition of HCl (4M, 50 μL/well). Absorbance was immediately measured at 450 nm and was corrected by subtracting the average of negative controls (i.e. PBS had been added to some of the wells instead of HER2 or instead of the samples). Each sample was tested in triplicate and errors are shown as the standard deviation of the average. ELISA data was analysed with Graphpad Prism 7.03 (using equation Sigmoidal, 4PL, X is log(concentration)) and the values have been normalized. For calculation of IC$_{50}$, concentration values (in nM) were transformed into log(concentration).
4. References


189


132 L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner,


D. Luo, N. Man, M. Krantz, K. Wilde, D. Wishart, Y. Zhang, F. Jacobs and L. Martin, *J.
Biochem., 1995, **831**, 825–831.


N. G. Caculitan, J. dela Cruz Chuh, Y. Ma, D. Zhang, K. R. Kozak, Y. Liu, T. H. Pillow,


317  T. Wang, A. Pfisterer, S. L. Kuan, Y. Wu, O. Dumele, M. Lamla, K. Müllen and T. Weil,


5. Appendix

5.1 MALDI spectra

![Figure 85 - MALDI-TOF analysis of HSA-scFv conjugate.](image)

![Figure 86 – MALDI-TOF analysis of tri-scFv conjugate.](image)
5.2 LC-MS/MS spectra of tryptic digestion of HSA modified with N-Biotin dibromomaleimide 54

Figure 87 – LC-MS/MS spectrum of Cys-34/Lys-41 cross-linked peptide (N-Biotin maleimide derivative) obtained by selecting m/z 996.83 [M+4H]^{4+} as the precursor ion for CID; 3+ = trivalent ion (m/3), 2+ = bivalent ion (m/2), 0 = m-H_2O, * = m-NH_3.
Figure 88 – LC-MS/MS spectrum of Cys-34 peptide (N-Biotin maleimide derivative) obtained by selecting m/z 957.70 [M+3H]^3+ as the precursor ion for CID; 3+ = trivalent ion (m/3), 2+ = bivalent ion (m/2), 0 = m-H_2O, * = m-NH_3. Proposed structure of hydroxy/thiomaleimide – postulated to form via enamine hydrolysis of the cross-linked amino-thiomaleimide species.
5.3 LC-MS spectra of HSA modified with N-Me dibromomaleimide, treated with BME before and after incubation at pH 8.0, 37 °C, 72 h

Figure 89 – LC-MS analysis of HSA modified with N-Me dibromomaleimide; observed mass of 66,549 (calculated 66,549 for HSA modified with N-Me dibromomaleimide missing both Br atoms).

Figure 90 – LC-MS analysis of HSA modified with N-Me dibromomaleimide treated with 100 eq. BME (before incubation at pH 8, 37 °C, 72 h); observed mass of 66,443 (calculated 66,440 for native HSA).

Figure 91 – LC-MS analysis of HSA modified with N-Me dibromomaleimide treated with 100 eq. BME (after incubation at pH 8, 37 °C, 72 h); a; observed mass of 66,550 (calculated 66,549 for HSA modified with N-Me dibromomaleimide missing both Br atoms).
5.4 ELISA analysis of tri-scFv and scFv before and after DOTA conjugation

Figure 92 – ELISA analysis of tri-scFv and scFv (against CEA) before and after reaction with DOTA-NHS.
5.5 LC-MS spectrum of native Fab highlighting additional species present, in comparison to Fab after acyl transfer

Figure 93 – LC-MS spectra of native Fab and Fab after acyl transfer, highlighting modification of additional species present in the native Fab sample; a. Native Fab sample, b. Fab sample after acyl transfer (conditions: pH 8.4, 12 °C, 72 h); +1, +2 refers to the number of acyl groups added.
5.6 Publications

The following articles have been published as a result of the work reported in this thesis:
