

Mass Spectrometric Methods for the Analysis of Chemically Modified Proteins

Katherine Grace Stevens
B Med Sci, B Biomed Res (Hons.)

A thesis submitted in fulfilment of the requirements for the
degree of Doctor of Philosophy

Department of Chemistry
School of Physical Sciences
University of Adelaide, Australia
November 2020



Contents

Abstract	4
Declaration	6
Acknowledgements	7
CHAPTER ONE	8
Conjugating immunoassays to mass spectrometry: Solutions to contemporary challenges in clinical diagnostics	8
Foreword	8
Introduction	10
Combined immunochemical and mass spectrometric approaches to biomolecule analysis.....	12
Summary	17
CHAPTER TWO	20
Synthesis of modifiable photocleavable MALDI mass tags using copper click chemistry	20
Foreword	20
Abstract	24
Introduction	25
Results and Discussion	26
Experimental Procedures	32
References.....	36
Supporting Information.....	39
CHAPTER THREE.....	1
Retro Diels-Alder fragmentation of fulvene-maleimide bioconjugates for mass spectrometric detection of biomolecules	1
Foreword	1
Abstract	5
Introduction	6
Experimental Methods.....	8
Results and Discussion	12
Conclusions	19
References.....	20
Supporting Information.....	23
CHAPTER FOUR	41
Application of bottom-up proteomics for identifying sites of oxidative modification and chemical labelling in equine heart myoglobin	41
4.1 Introduction	41
4.2 Aims.....	43
4.3 Results and Discussion	44
4.4 Conclusions	51

4.6 Experimental	53
4.7 References.....	55
CHAPTER FIVE	58
Development of a modular synthetic route for protein chemical cross-linking reagents.....	58
Foreword.....	58
Abstract.....	62
Introduction	63
Results and Discussion	65
Conclusion	71
References.....	76
Supplementary Information.....	79
CHAPTER SIX	85
Conclusions and future directions.....	85
References.....	87
Appendix	89
Table A1. Carbonylated peptide spectrum matches identified from Sequest HT targeted search of unmodified myoglobin.	89
Table A2. Carbonylated peptide spectrum matches identified from Sequest HT targeted search of oxidised myoglobin.....	90
Table B1. DPANzine-labelled peptide spectrum matches identified from Sequest HT targeted search of oxidised myoglobin incubated with fluorophore.	91
Table B2. Peptide spectrum matches identified from Sequest HT targeted search of oxidised myoglobin incubated with fluorophore.	91
Table C1. Mass shifts identified from Sequest HT error-tolerant search of unmodified myoglobin.	92
Table C2. Mass shifts identified from Sequest HT error-tolerant search of DPANzine-labelled unoxidised myoglobin.....	93
Table C3. Mass shifts identified from Sequest HT error-tolerant search of DPANzine-labelled unoxidised myoglobin after treatment with sodium cyanoborohydride.	94
Table C4. Mass shifts identified from Sequest HT error-tolerant search of oxidised myoglobin detected using DeltaMass.....	95
Table C5. Mass shifts identified from Sequest HT error-tolerant search of DPANzine-labelled oxidised myoglobin detected using DeltaMass.	96

Abstract

Mass spectrometry has critical roles in analytical chemistry, biomedical research and the diagnosis and treatment of human disease. Its ability to unambiguously identify and quantify a diverse range of biomolecules, from small molecule metabolites to large intact protein complexes in complex biological matrices, has solidified this technique's place in the clinical chemists' laboratory. However, mass spectrometry's wider adoption by biochemists and clinicians is hindered by the inherent volume and complexity of the data it can generate from a biological context, and the subsequent level of specialist equipment and knowledge required to interpret results in a reliable and meaningful way. It is therefore apparent that a potential solution to this issue is to develop new methodologies, which adapt simple biochemical techniques into mass spectrometry sample preparation workflows that can be performed by biochemists using standard laboratory equipment.

This thesis begins by reviewing some recent advances in clinical mass spectrometry and the current limitations of this technology. In doing so, we establish the importance of hyphenated techniques, which merge the concept of immunoassays with mass spectrometric detection to enable its adoption in high-throughput clinical and biomedical research applications. Many of these new approaches are only possible because of the development of efficient biocompatible chemical reactions, which facilitate the covalent modification of antibodies to improve their analysis via mass spectrometry.

The concept of chemical labelling with *mass tags*, or synthetic linkers that fragment during ionisation and/or other stages of gas-phase analysis, has been widely explored as a strategy for detecting large heterogeneous biomolecules, such as antibodies. Chapter 2 explores the design, synthesis and application of an ultra-violet-cleavable linker for labelling intact antibodies to simplify their detection via matrix-assisted laser desorption/ionisation mass spectrometry. The novel aspect of this project is the utilisation of copper-catalysed azide-alkyne cycloaddition, or *copper click chemistry*, to enable straightforward modification of the detected photodissociation product, thereby offering the potential for multiplexed analysis of biomolecules.

Despite the established utility of the copper click reaction, there are several limitations for its application in biological contexts. We therefore sought to identify alternative routes for modifying mass tags. Chapter 3 explores the utilisation of Diels-Alder cycloaddition chemistry for synthesising mass-tagged biomolecules, which conveniently facilitate efficient gas-phase fragmentation, even with the absence of a photocleavable moiety.

The ability to easily detect proteins within biological specimens is incredibly valuable for diagnostics; however, the aetiology of many diseases involves modifications to proteins that occur after biosynthesis, or *post-translational modifications*. Chapter 4 explores the application of bottom-up proteomics methods to identify amino acid modifications to equine heart myoglobin in response to a model for oxidative stress, and characterise the products of its labelling with a novel fluorophore.

Whereas the ability to analyse primary protein sequences and post-translational modifications is undoubtedly valuable, the function of many proteins is closely linked to their three-dimensional higher order structures, including transient interactions with other biomolecules. Reagents known as *chemical crosslinkers*, can be utilised to stabilise interactions between nearby amino acids prior to bottom-up proteomic analysis, thereby retaining tertiary and quaternary protein structural information. Chapter 5 describes some of the recent efforts from our research group to develop efficient strategies for synthesising and applying chemical cross-linkers with improved features for the purification and identification of cross-linked proteins.

This thesis therefore explores various potential implementations of mass spectrometry for investigating the dynamic and complex roles of proteins in the context of human disease, from altered expression levels to post-translational modifications and protein-protein interactions.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Katherine G Stevens

1/12/20

Acknowledgements

Thank you to my supervisor Tara Pukala, for providing me with the guidance, support and friendship I needed to get this far. To my past and present colleagues and friends in the Pukala research group, I am grateful to all of you for accepting me as part of the group, your friendship and for all the fun times we shared together.

To my co-supervisor Paul Trim, thank you for teaching me more about MALDI in a few months than I managed to learn in 3 years by myself. I wish I had more opportunity to work with you and your great team at SAHMRI. Thank you also to my co-supervisor Andrew Abell and his lab group, who made me feel welcome and gave me so much great advice, which helped me to achieve as much as I did in my time at Adelaide.

Thank you to my mentor, Mathew Palmer, for continuing to be a voice of reason and giving me a sense of direction at times when I started to get off track.

To my collaborators, Aimee Horsfall from the Abell Lab, Kym McNicholas and Robert Rush from Biosensis, Anton Blencowe and Kirsten Platts from the University of South Australia, Yuning Hong and Siyang Ding from La Trobe University and everyone else involved, thank you for making this work possible. Thank you also to Mark Condina, Parul Mittal and Clifford Young from UniSA for generously sharing some of their proteomics wisdom with me.

Thank you to everyone in the Adelaide Chemistry department who accepted me as one of their own. I have made some great friends over the years and learnt so much from so many of you. Likewise, to all the wonderful friends I have made in Molecular Life Sciences, for being such a welcoming, fun and generous group of people. A special thanks to the amazing technical and support staff in the Faculty of Sciences who have been indispensable throughout my PhD. I would especially like to thank Matthew Bull, for spending so much of his own time repairing the instruments I have needed for my research.

Lastly, to my parents, Fiona and George, thank you for teaching me the value of hard work and knowledge. Inheriting your love of learning has made me who I am today. And to my friends Stefania, Saurabh, Pdraig, BJ, Robyn, Sherry, Naomi, Parul and Tara, thank you for being there when I needed someone to listen and supporting me when times were tough.

Conjugating immunoassays to mass spectrometry: Solutions to contemporary challenges in clinical diagnostics

Foreword

The focus of this thesis is on the development of mass spectrometric (MS) methods for analysing proteins in ways that produce meaningful data in relation to human health, with an emphasis on their application in the context of biomarker detection. While the overarching aim was to establish novel methodologies for achieving this, which often make use of state-of-the-art technologies, it is important to first discuss the important developments that preceded them. After summarising some of the basic concepts of clinical diagnostic tests and MS methodology, the following review emphasises some of the important advances in MS technology that have enabled the exploration of hyphenated immunoassay-MS techniques. Although still somewhat in their infancy, these hyphenated approaches have found value and recognition in applications for which the limitations of conventional MS and immunoassay methods preclude the effective diagnosis and treatment of various significant human diseases. In doing so, we also introduce the concept of 'bioconjugation', that is, the covalent modification of biomolecules, and its potential for aiding the detection of molecules with properties that are otherwise less than ideal for MS analysis. Importantly, this review also provides an in-depth examination of the limitations and challenges faced in previous attempts at developing clinically feasible methods for detecting antibodies, and therefore informs the design of novel reagents that could bring this technology closer to eventual adoption in clinical setting. Although many are a long way from approval for use in clinical settings, these examples set a firm precedent for the continued exploration of such approaches, which becomes the foundation for the work we present in the following chapters.

Statement of Authorship

Title of Paper	Conjugating immunoassays to mass spectrometry: Solutions to contemporary challenges in clinical diagnostics
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Stevens, K. G.; Pukala, T. L., Conjugating immunoassays to mass spectrometry: solutions to contemporary challenges in clinical diagnostics. <i>TrAC Trends in Analytical Chemistry</i> 2020 , 116064.

Principal Author

Name of Principal Author (Candidate)	Katherine G Stevens		
Contribution to the Paper	Author of main body text, decision on article content		
Overall percentage (%)	85		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	1/12/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Tara L Pukala		
Contribution to the Paper	Contributions to article structure and content		
Signature		Date	4/12/20



Conjugating immunoassays to mass spectrometry: Solutions to contemporary challenges in clinical diagnostics

Katherine G. Stevens, Tara L. Pukala*

The University of Adelaide, Adelaide, Australia



ARTICLE INFO

Article history:
Available online 7 October 2020

Keywords:
Antibodies
Bioconjugation
Biomarkers
Diagnostic tests
Electrospray ionisation-mass spectrometry (ESI-MS)
Immunoassay
Mass cytometry
Mass spectrometry imaging
Matrix-assisted laser desorption/ionisation (MALDI)
Proteomics

ABSTRACT

Developments in immunoassays and mass spectrometry have independently influenced diagnostic technology. However, both techniques possess unique strengths and limitations, which define their ability to meet evolving requirements for faster, more affordable and more accurate clinical tests. In response, hybrid techniques, which combine the accessibility and ease-of-use of immunoassays with the sensitivity, high throughput and multiplexing capabilities of mass spectrometry are continually being explored. Developments in antibody conjugation methodology have expanded the role of these biomolecules to applications outside of conventional colorimetric assays and histology. Furthermore, the range of different mass spectrometry ionisation and analysis technologies has enabled its successful adaptation as a detection method for numerous clinically relevant immunological assays. Several recent examples of combined mass spectrometry-immunoassay techniques demonstrate the potential of these methods as improved diagnostic tests for several important human diseases. The present challenges are to continue technological advancements in mass spectrometry instrumentation and develop improved bioconjugation methods, which can overcome their existing limitations and demonstrate the clinical significance of these hybrid approaches.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Biomarkers and clinical analytical techniques

The term *biomarker* can denote any biological molecule or combination of factors that indicate a particular biological state, which are often used to differentiate normal or abnormal processes or conditions [1]. Whereas new biomarkers are identified using untargeted, semi-quantitative (comparative) analytical approaches, to develop a viable clinical test there needs to be a reproducible method for their absolute quantitation. For implementation in clinical laboratories, analytical tests designed for diagnostic applications need to meet performance requirements with respect to their accuracy and predictive capabilities [2]. Diagnostic accuracy is determined based on a test's abilities to positively identify individuals who have a condition and eliminate those who do not, or its *sensitivity* and *specificity*, respectively. A test's predictive value is evaluated by calculating the proportions of correct diagnoses out of

the total positive and total negative test results, or *positive* and *negative predictive values*, respectively [2]. In the context of biomarkers, these factors are dependent on the dynamic range, accuracy, and reproducibility of whichever analytical method is used to detect changes in their abundance.

Most common clinical laboratory tests make use of spectrophotometric and/or immunologic detection methods [3]. Of these, *immunoassays*, which take advantage of the highly selective interactions between specific immunoglobulins and their target antigens, are some of the most clinically relevant techniques [4,5]. Furthermore, the chemical composition of these large proteins enables a variety of strategies for modifying their structure with limited perturbation of their antigen-binding activity (discussed in detail in Section 2.2). This feature allows for immunoassays to be coupled to a range of different detection methods, including radiometric, fluorescent, colorimetric, chemiluminescent, non-labelled (light scattering) and electrochemical detection [4,6]. Detection can be further enhanced using enzymatic, polymerase chain reaction (PCR), liposome and nanomaterial-based signal amplification strategies [7–9].

In many instances, absolute levels of biomarkers in biological fluids or tissue biopsies are insufficient for determining a reliable

* Corresponding author. School of Physical Sciences, The University of Adelaide, 5005, Australia.

E-mail address: tara.pukala@adelaide.edu.au (T.L. Pukala).

Abbreviations		m/z	mass-to-charge ratio (m/z)
CPD	carboxypeptidase D	MERS	Middle East respiratory syndrome
DESI	desorption electrospray ionisation	MS	mass spectrometry
EDC	<i>N</i> -ethyl- <i>N'</i> -(3-dimethylaminopropyl)carbodiimide	MSI	mass spectrometry imaging
ELISA	enzyme-linked immunosorbent assay	MSIA	mass spectrometry-immunoassay
ESI	electrospray ionisation	MS/MS	tandem mass spectrometry
FDA	US Food and Drug Administration	NHS	<i>N</i> -hydroxysuccinimide
FITC	fluorescein isothiocyanate	PCR	polymerase chain reaction
GC-MS	gas chromatography-mass spectrometry	P-PC	photocleavage product
ICAT	isotope-coded affinity tag	PSMS	paper spray mass spectrometry
ICPL	isotope-coded protein labelling	PTM	post-translational modification
ICP-MS	inductively coupled plasma-mass spectrometry	SARS	severe acute respiratory syndrome
IHC	immunohistochemistry	SIMS	secondary-ion mass spectrometry
iMALDI	immunoMALDI	SISCAPA	stable isotope standards and capture by anti-peptide antibodies
IMC	imaging mass cytometry	SPR	surface plasmon resonance
iTRAQ	isobaric tag for relative and absolute quantitation	TFP	tetrafluorophenyl
LC-MS	liquid chromatography-mass spectrometry	TMT	tandem mass tag
LDI	matrix-free laser desorption/ionisation	UV	ultraviolet
MALDI	matrix-assisted laser desorption/ionisation		

diagnosis, particularly for diseases characterised by complex changes in tissue morphology and localised changes in protein expression [10]. More relevant information can therefore be obtained by comparing the spatial distribution of biomarkers in normal and diseased specimens. Immunohistochemistry (IHC) involves the labelling of specific antigens in tissue sections with antibodies, which are then visualised using some combination of staining and imaging techniques(s). Owing to its simplicity, affordability and versatility, this technique is commonly employed in diagnostic pathology [10,11].

Developments in immunoassay and IHC technologies, such as automated enzyme-linked immunosorbent assays (ELISAs), microfluidics, lab-on-a-chip technologies, and computer-assisted image analysis, have resulted in significant reductions in analysis time and complexity, sample volumes and specialised equipment or expertise required [5,10]. However, many of these methods still utilise some form of spectrophotometric detection and are therefore limited in the number of analytes that can be detected in a single experiment due to overlaps in the emission ranges of different fluorophores and narrow dynamic range [12]. The development of immunoassays that utilise detection methods not constrained by the inherent limitations of spectrophotometric measurements has therefore become an important goal for modern diagnostic medicine.

1.2. Evolution of biomolecular and clinical mass spectrometry

Since its inception in the early 20th century, mass spectrometry (MS) has developed into an important tool for biomedical researchers and clinicians [13–15]. *Soft ionisation* methods, such as electrospray ionisation (ESI), matrix-assisted laser desorption/ionisation (MALDI) and chemical ionisation, enable ionisation of molecules with minimal fragmentation. These methods are therefore very useful for the MS analysis of intact biomolecules, with MALDI and ESI commonly utilised in both research and clinical settings [3,13,14,16]. Compared to spectrophotometric methods for detecting biomolecules, MS differentiates analytes based on the mass-to-charge ratio (m/z) of intact molecules and/or the characteristic products of their gas-phase fragmentation, and therefore provides high specificity and sensitivity and enables the detection of different isoforms [13,17]. The ability of MS to detect many different analytes simultaneously, or *multiplex*, is useful for analysing complex

biological mixtures as entire proteomes, lipidomes, or metabolomes can be investigated for a single sample [13,18]. This technology has obvious applications in diagnostic medicine; hence, renewed enthusiasm for advancement in MS methodology is now aimed at developing clinically viable platforms [17,19].

Widespread adoption of MS in clinical laboratories did not occur until the 1980's, after the limitations of immunoassays for illicit drug screening and detection of steroids became apparent. Subsequent acceptance of gas chromatography-MS (GC-MS) for immunoassay validation in clinics lead to mainstream use of the more versatile liquid chromatography-MS (LC-MS) [17,18,20]. Despite its waning popularity due to the requirement for extensive sample preparation and often chemical derivatisation, GC-MS still has an important place in the clinical analysis of selected compounds such as excreted steroid metabolites [21,22].

The development of *tandem* MS (MS/MS), which enables the unambiguous assignment of biomolecules based on their unique fragmentation patterns, widened the boundaries of MS within the clinical laboratory to include protein and peptide biomarker detection, multi-analyte therapeutic drug monitoring, drug abuse screening, toxin analysis, endocrinology and screening for metabolic diseases [15,23]. The capabilities of this technology are exemplified by the now wide-spread adoption of MS/MS-based blood-spot screening for congenital metabolic diseases in newborns, which was previously limited to individual metabolites but can now detect in excess of 40 analytes simultaneously [24,25].

Another revolutionary development in clinical MS was the discovery of a method for identifying bacterial molecular fingerprints using MALDI, hence providing a faster and easier alternative to other time-consuming laboratory tests for identifying pathogenic microorganisms [17,26]. This demonstration of MALDI-MS as a clinically viable technique, which could be performed with limited sample preparation, was followed by US Food and Drug Administration (FDA)-approval of two MALDI systems for identifying gram-negative bacteria [3,17,27]. The multiplexing capabilities of MALDI have since been exploited to develop rapid PCR-based MS assays for screening of human coronaviruses, including severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS) viruses. In this study, the authors concluded that multiplexed analysis reduced false negative results and has the potential to detect novel viruses [28].

Until 1997, immunochemical detection still had a distinct advantage over MS: the capacity to determine the spatial distribution of proteins in tissues. The invention of an innovative new way to acquire MALDI spectra led to the establishment of a new research field: MS imaging (MSI). During MSI experiments, individual mass spectra are obtained sequentially across the surface of a biological specimen and converted into an intensity map showing the localisation of ions with specific m/z [29]. With the development of additional ionisation techniques, such as desorption electrospray ionisation (DESI) and secondary-ion mass spectrometry (SIMS), MSI can now be used for the analysis of proteins and protein complexes, small molecules, lipids, metabolites, oligonucleotides and sugars, many of which cannot be detected using immunochemical methods [30–33]. A form of laser ablation MS, rapid evaporative ionisation mass spectrometry (iKnife; Waters Corporation, Milford, US), has even been developed to assist surgeons during tumour removal by providing real-time analysis of patient tissue components to identify cancerous tissue margins [30].

1.3. Current challenges for the development of clinically viable MS methods

For analysis of biological samples, MS workflows generally include a sequence of sample preparation, separation and MS analysis procedures and their clinical viability is dependent on the complexity, cost and duration of each step [17,34]. Some inherent limitations to common MS technologies also have a significant influence on their effectiveness in clinical diagnostic applications.

1.3.1. Matrix effects and ion suppression

Another significant issue for MS analysis of complex biological samples is *ion suppression*, which refers to the reduction in ionisation of target analytes because of interference from other components within the biological matrix, such as salts, detergents or other non-volatile compounds [34,35]. For ESI-based analyses, this problem is usually addressed during sample preparation and separation steps or additional enrichment steps, for example by using LC to separate non-volatile components [34,36].

Comparatively, sample preparation for MALDI-MS is a lot simpler as this ionisation method is more tolerant of biological sample components, such as buffers [3]. However, the absence of pre-analysis enrichment steps also makes it more difficult to detect low-abundance ions, as MALDI spectra are often dominated by signals from more concentrated, albeit less clinically significant, biomolecules, a phenomenon sometimes described using the analogy of a *needle in a haystack* [31].

For imaging experiments, MALDI also requires additional sample preparation steps to that of immunohistochemistry, such as enzymatic digestion or chemical release of proteins and glycans [13]. Some common tissue conservation techniques, such as para-formaldehyde fixation followed by long-term storage, can result in incompatibility with MSI analysis [33]. Furthermore, in most instances these experiments are limited to qualitative assessments of analyte spatial distribution due to the influence of tissue-specific ion suppression unless isotopically labelled internal standards or specialised MALDI matrices are used [37]. For these reasons, MALDI-MSI is far from replacing immunohistochemistry as a primary diagnostic technique; however, it has established an important place within clinical research as a tool for novel biomarker discovery, tumour classification and staging, and treatment monitoring [38].

1.3.2. Relative and absolute quantification of biomarkers using MS

For clinical applications, a major caveat of soft ionisation techniques, including ESI and MALDI, is that some analytes will ionise

more efficiently than others, meaning that absolute concentrations cannot be determined based on ion counts alone. Absolute quantification therefore requires the use of structurally similar or isotope-labelled internal standards, which may not be commercially available or economically viable for large scale clinical testing [16,34]. For example, the analysis of complex protein digests would require labelled peptides for each target, making multiplexed experiments expensive if commercially synthesised standards are used. Proteomics experiments can also introduce additional variability in preanalytical steps, such as enzymatic digestion, so ideally make use of labelled intact proteins and therefore require the availability of recombinant protein expression systems [39–41].

Alternatively, various approaches to generating labelled standards using simple chemical labelling reagents are available and enable relative quantification between samples using MS. These methods utilise common amino acid modification strategies (Table 1) to label either proteins or peptides with stable isotopes, hetero-elements (e.g. halogens), MS-cleavable tags and/or affinity handles [42]. Significant examples include the biotinylated isotope-coded affinity tag (ICAT) reagents for labelling peptides, amine-directed isotope-coded protein labelling (ICPL), and the isobaric tandem mass tag (TMT) and isobaric tag for relative and absolute quantitation (iTRAQ) systems [42–44]. Despite their obvious clinical utility, approval of these workflow for routine testing is impeded by their inherent variability and the absence of effective bioinformatics and data analysis platforms for interpreting the complex data they generate [42].

Absolute quantification can be achieved in a more straightforward manner using inductively coupled plasma (ICP)-MS, which involves atomisation of molecules using extremely high temperatures (7000–10,000 K) to detect hetero-elements (any element other than C, N, O and H) [12,45,46]. This ionisation technique is highly sensitive, has a wide dynamic range, and produces signals that are directly proportional to the sample concentration of a given element, irrespective of the solvent, analyte and sample matrix [42]. Hence, the technique has been useful for quantitative analysis of biomolecules with naturally occurring trace for phosphoproteins, metalloproteins, and selenoproteins [12,30,45]. However, quantification of intact complex biomolecules, such as peptides and antibodies, is also possible through chemical or metabolic labelling with hetero-element-containing reagents [42,45]. This technology has important implications for the implementation of ICP-MS and other MS platforms in clinical setting and is discussed in more detail in Sections 2.3.2 and 2.4.1.

1.3.3. Accessibility of MS in the clinical laboratory

Most MS experiments require specialist knowledge of correct sample preparation methods and instrumentation, in addition to dedicated software programs to acquire, analyse and interpret data [47]. Hence, despite their significant advantages, establishing robust clinical MS methods is often more complicated, expensive and/or time-consuming compared to conventional approaches, such as immunohistochemistry and ELISAs [47]. Due to these limitations, MS technologies are still underutilised in clinical settings in favour of immunochemical methods [19].

2. Combined immunochemical and mass spectrometric approaches to biomolecule analysis

Approaches combining immunoassay methodology with MS detection have been developed in efforts to overcome some of the individual limitations of these techniques. However, owing to their relatively high molecular weight and substantial heterogeneity, direct MS analysis of antibody-antigen complexes poses additional challenges [48,49]. ESI of intact complexes requires samples to be

Table 1
Examples of bioconjugation chemistries used for modifying proteins.

Modification/reagent	Target residue(s)	Reference(s)
Non-specific conjugation		
Amide coupling using coupling reagent(s), such as EDC	Carboxylic acids (C termini, glutamine and asparagine residues) and free amines (lysine residues and N termini)	[66]
Reactive esters (NHS, TFP, etc.), aldehydes and isothiocyanates	Free amines, including N termini and lysine residues	[33,43,66–68]
Maleimide and haloacetamide reagents, and aryl palladium complexes	Reduced thiols (cysteine residues) or thiolated amines	[12,42,44,66,67,69]
Selenocysteine conjugation (maleimide and iodoacetamide reagents)	Selenocysteine	[67]
Site-directed conjugation		
Hydrazides and alkoxyamines	Aldehydes or ketones on oxidised glycans or non-natural amino acids, respectively	[66,67]
Enzymatic ligation	Various genetically encoded recognition peptide sequences	[67,70]
Copper click chemistry	Non-natural alkyne or azide-containing amino acids	[67]
Copper-free click chemistry	Various non-natural amino acids	[67]
Indole-3-butyric acid photoactivated ligation	Endogenous nucleotide binding sites	[67,71]
Non-covalent conjugation		
Biotin-avidin/streptavidin	Biotinylated residues targeted by avidin/streptavidin or vice versa	[12,44,63,70,72–74]
Protein A affinity capture	Immunoglobulin F _c region	[54]
Nickel-chelate affinity	Metal coordination sites on IgG class antibodies	[66]

Abbreviations: EDC, *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide; NHS, *N*-hydroxysuccinimide; TFP, tetrafluorophenyl.

prepared in volatile aqueous buffers to preserve non-covalent interactions; these conditions favour formation of high *m/z* ions and salt adducts, which reduces ion transmission and mass accuracy, and therefore sensitivity and selectivity [49]. Similarly, MALDI analysis requires specialised matrices or chemical cross-linking to analyse intact complexes and generates singly charged, high *m/z* ions with poor mass resolution [48,49]. To address these issues, several novel methods for detecting biomarkers have been proposed that involve chemically modifying antibodies to facilitate their analysis using standard MS instrumentation.

2.1. Immunoaffinity-based solutions in quantitative proteomics

Given the limitations of MS for direct analysis of complex mixtures, the features of immunoassays that enable isolation of analytes from biological samples offer opportunities for adapting MS detection to clinical sample analysis by improving selectivity and sensitivity. Various conjugation methods are available for immobilising antibodies on different solid supports (covered in detail in Section 2.2). These combined immunoaffinity capture-MS methods, or MS-immunoassays (MSIAs) utilise immobilised antibodies to enrich target proteins from complex biological samples and thus improve detection limits for low-abundance analytes. Similar technology has also been adapted to develop functional assays of biomarkers and therefore providing important insight into pathological biomolecule interactions and treatment monitoring [50,51].

2.1.1. MS immunoassays

The high specificity of MS makes it possible to evaluate different protein isoforms and PTMs, which typically cannot be differentiated by antibodies alone [30,52,53]. This concept was first demonstrated with MALDI, for the detection of snake venom proteins in human whole blood after enrichment using antibodies immobilised on agarose beads [54]. The MSIA has since been developed into a pipette tip-based commercial platform (MSIA D.A.R.T.SM; Thermo Fisher Scientific, Waltham, US) compatible with ESI systems, and various iterations have now been used to quantify clinically relevant protein isoforms involved in insulin resistance, Alzheimer's disease, renal function, endocrine function, lung cancer, and cardiovascular disease, among others [52,53,55–58]. This combination of immunoaffinity capture and MS detection can provide substantial improvements in sensitivity

and specificity by exploiting the enrichment capabilities of antibodies, while eliminating issues related to non-specific binding by using direct measurement of unique antigen *m/z*'s [59]. MSIA workflows are also applicable to both bottom-up and top-down proteomics analyses, the latter of which can identify novel proteoforms without prior knowledge of the complete amino acid sequence and PTMs, therefore providing more complete sequence coverage [60–62].

Alternatively, the stable isotope standards and capture by anti-peptide antibodies (SISCAPA; SISCAPA Assay Technologies, Washington DC, USA) platform enables multiplexed analysis of pre-digested protein samples using peptide-reactive antibodies immobilised onto capillary columns, which are incorporated into an autosampler LC system. Samples are spiked with isotopically labelled peptide standards before on-line washing and enrichment, hence retaining the high-throughput and quantitative capabilities of LCMS proteomics workflows [63,64]. In analogous immuno-MALDI (iMALDI) workflows, peptides are eluted from the capture antibody using an acidic MALDI matrix solution directly onto the MALDI target plate [61,65]. For these peptide-centric assays, the process for generating biomarker-reactive antibodies and isotope-labelled standards is relatively simple, compared to protein-centric MSIAs [61].

2.1.2. Surface plasmon resonance MS

Surface plasmon resonance (SPR) is another technique that has evolved from a method for studying antibody-antigen interactions into a clinically relevant quantitative tool for detecting biomarkers [49,51]. SPR measurements are often a product of a continuous flow of analyte across immobilised antibodies on a metal surface, which induces a change in the surface's refractive index for each instance of analyte binding. This technique has the advantages of being able to determine absolute analyte concentrations and binding kinetics from very small volumes at low sample concentrations. However, the mode of detection prevents delineation of antibody-antigen complex stoichiometry, protein variants, complexes, and non-specific binding [49,50]. Coupling this technology with subsequent MS analysis has since overcome these limitations. Moreover, the development of SPR imaging now enables direct multiplexed MALDI analysis of SPR chip arrays, without the need for recovery of analytes from the SPR surface. This approach not only improves the dynamic range of MS analysis, but also significantly reduces sample

preparation time, making it useful for detecting a range of human biomarkers, in addition to providing insight into their higher-order structure and function [50].

2.2. Bioconjugation strategies and challenges

The broad range of reactive chemical groups on antibodies makes them amenable to various conjugation methods (Fig. 1 and Table 1) [66]. In addition to conventional methods for modifying amino acids via free amines and carboxyl groups, modifications at selectively reduced cysteine residues and glycosylation sites are commonly used for functionalising antibodies. Unlike modifications to lysine and acidic amino residues, these site-directed modifications are less likely to lead to loss of activity from changes to the antigen-binding (F_{ab}) region [66,67]. Alternatively, antibodies can be engineered to incorporate non-natural amino acids to allow site-specific conjugations [67]. However, this approach requires additional resources and expertise, whereas several amine and carboxylate targeting reagents are commercially available and can be used to modify validated antibodies. The ongoing development of new bioconjugation chemistries will likely lead to more simple and effective conjugation strategies for modifying antibodies and thus make them more appropriate for routine use in clinical settings.

2.3. MS detection of labelled antibodies for immunoassays

Given the limitations of common MS instrumentation for analysing intact antibodies, there is a growing interest in developing conjugation reagents designed to improve their ionisation and detection. Many of these approaches utilise labelling reagents that are cleaved at some point during MS analysis to release a fragment of known m/z , which is detected as a proxy for the intact antibody-biomarker complex [59,75]. These fragments, or *mass tags*, can be designed to fall within optimum m/z ranges, facilitate ionisation and undergo signal amplification to overcome limitations in dynamic range, poor ionisation, stability and sensitivity. Altering the chemical structure of the mass tag can also be used to generate different m/z values and therefore enable multiplexed analysis using different antibodies in the same sample [76].

2.3.1. MALDI mass tags

MALDI instruments have been popular choices for demonstrating mass tag concepts, because the most common ionisation

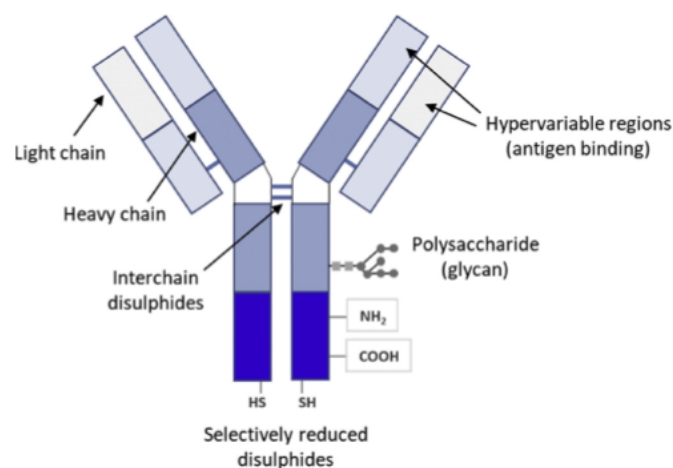


Fig. 1. Immunoglobulins possess multiple structural features that make them amenable to a variety of conjugation methods. Adapted from Ref. [66].

methods involve excitation of samples using an ultraviolet (UV) laser; mass tag release can therefore be triggered by UV-induced cleavage of photodegradable linkers covalently linked to antibodies (Fig. 2) [73,76]. An added benefit of this approach is that fixed charges can be incorporated into mass tags (or formed from their photodegradation products), enabling matrix-free laser desorption/ionisation (LDI) and thus eliminating matrix background signals [76]. Antibody mass tags have been effectively used for the development of MALDI/LDI-MS ELISA and microarray platforms. These early examples demonstrate the two most prominent mass tag design strategies: *ortho*-nitrobenzyl derivatives that can be incorporated into solid-phase peptide syntheses; and triphenylmethyl (trityl) tags, which fragment into cationic reporter ions, which can be detected in the absence of a MALDI matrix [73,77,78]. Comparatively more work has focussed on the application of these concepts for specific MSI, which is discussed in detail in Section 2.4.2.

2.3.2. ICP-MS and mass cytometry

Another type of MS instrumentation that has gained popularity in clinical settings is ICP-MS; the extremely high resolution, sensitivity and dynamic range of these instruments are desirable features when designing clinical assays. The development of new antibody conjugation reagents provided a way to apply ICP-MS for detecting fragile antibody molecules by labelling them with metal ions that provided unique isotopic signals in mass spectra, called *mass cytometry* (Fig. 3) [12,79]. Mass cytometry involves the detection of heteroatom labelled antibodies attached to specific antigens on or within individual cells in a suspension (a detailed explanation and protocol for the generation of heavy-metal-labelled antibodies for mass cytometry has been provided by Nolan and co-workers [69]). The detection of multiple markers enables separation of cells in complex biological mixtures such as blood, providing quantitative measurements of diagnostic cell populations (Fig. 4) [46].

Even the earliest example of an ICP-MS-based immunoassay for thyroid stimulating hormone showed improved sensitivity, compared to radioimmunoassay results for the same clinical samples [72]. More recent examples have explored the use of different labelling methods and signal amplification strategies, such as conjugated nanoparticles or metal-chelating polymers, to develop sensitive multi-parameter sandwich ELISAs and microarrays [46,80]. This multiplexing ability and compatible antibody labelling strategies of ICP-MS ELISAs also make this detection platform ideally suited for improving the diagnostic capabilities of existing flow cytometry technology [12,46]. The most recent commercially available mass cytometers (CyTOF3; Fluidigm Corporation, South San Francisco, US) can detect up to 40 markers per cell, compared to typically less than 20 for fluorescence-based flow cytometry [30,46,81].

2.4. Targeted MSI and IHC-MSI

The continued clinical utilisation of IHC and other imaging techniques emphasises the importance of spatial information in modern diagnostic pathology [10]. Although this technique has evolved to complement a wide range of imaging platforms, the exploration of IHC as an adjunct to MSI, or antibody-based *targeted MSI*, only began recently [33,79]. Development of ICP-MS and MALDI-MS into imaging platforms has led to the exploration of several targeted MSI methods, which are analogous to mass cytometry and MALDI/LDI mass tag strategies, respectively. However, these methods provide additional spatial information, which would prove indispensable if MS detection were to replace traditional antibody detection methods in clinical settings.

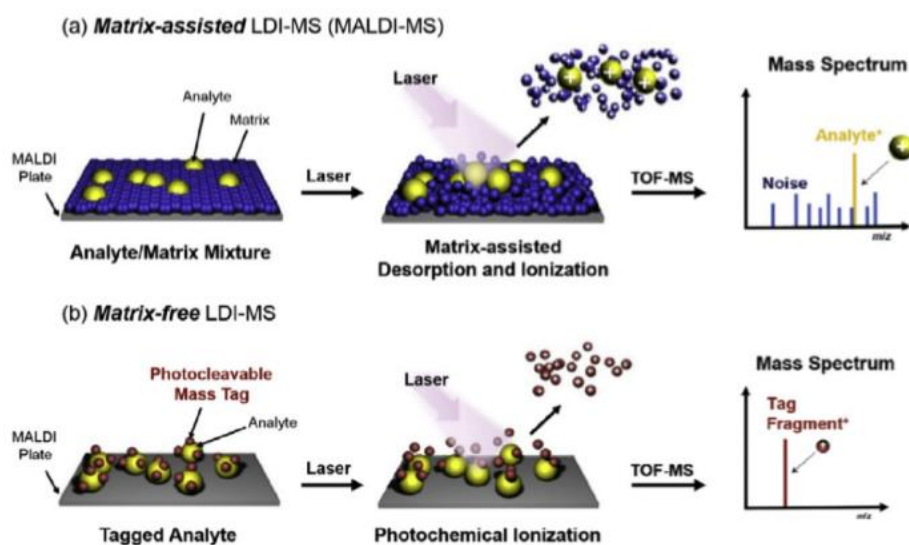


Fig. 2. Matrix-assisted laser desorption/ionisation mass spectrometry (MS) results in significant background signals from matrix adducts (A). Mass tags that form positively charged ions following photocleavage by an ultraviolet laser can be used for matrix-free laser desorption/ionisation (LDI)-MS (B). Reprinted with permission from Ref. [76].

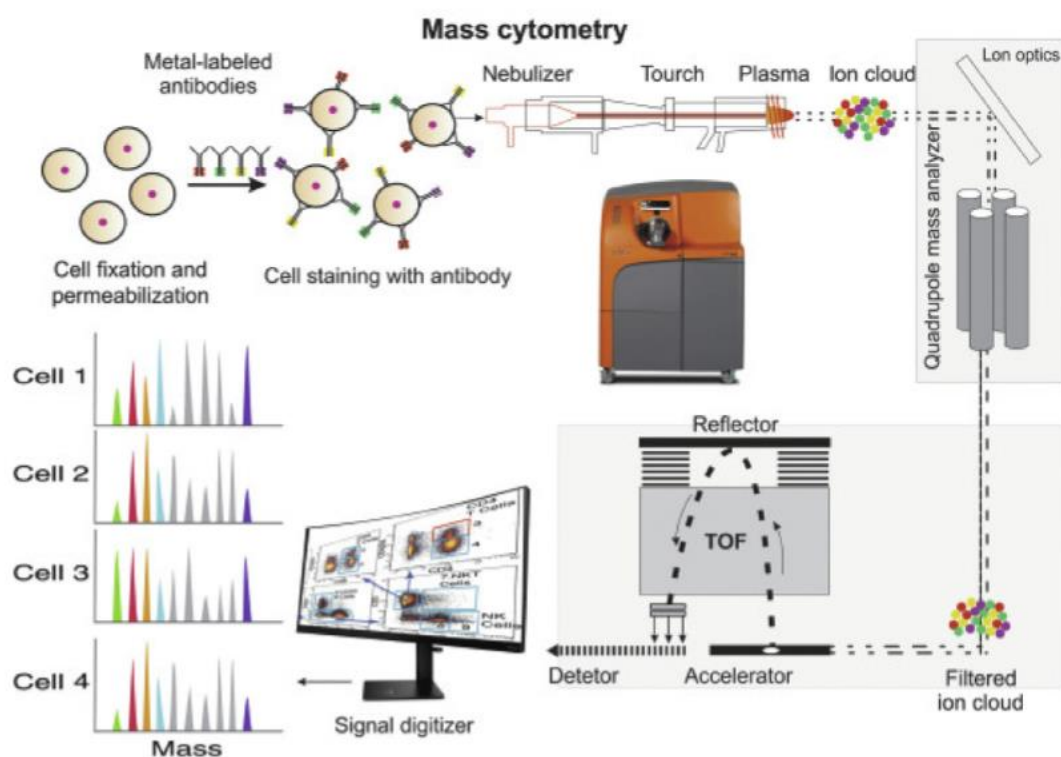


Fig. 3. General mass cytometry workflow for identifying cell populations using metal-conjugated antibodies and inductively coupled plasma-mass spectrometry. Reprinted with permission from Ref. [101].

2.4.1. Imaging mass cytometry

Imaging mass cytometry (IMC) applies the principles of mass cytometry in combination with a form of laser ablation ionisation to create raster images of metal ions from antibody conjugates (Fig. 5) [79]. The technology was originally used to demonstrate the potential of multiplexed IMC experiments as a revolutionary quantitative biomarker detection tool for breast cancer with subcellular spatial resolution [82]. This study was followed by an exploration of patient responses to trastuzumab treatment for breast cancer using IMC, which proves that this technique can also serve as an

important prognostic tool, which could be used to inform treatment regimens [83]. Since, IMC has been applied to biomarker detection for an extensive range of human diseases, with a focus on applications involving immune cells, including the study tumour-immune cell interactions, autoimmune diseases and immunophenotyping [69].

2.4.2. Targeted MALDI/LDI imaging

Targeted MALDI/LDI-MSI employs similar principles to those of mass tag-based immunoassays, whereby low *m/z* fragments are

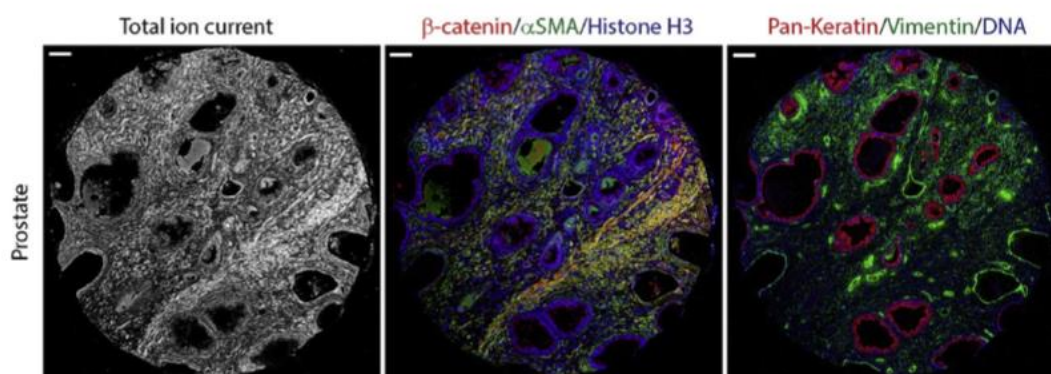


Fig. 4. Mass cytometry imaging of various antigens in normal human prostate tissue. Reprinted with permission from Ref. [79].

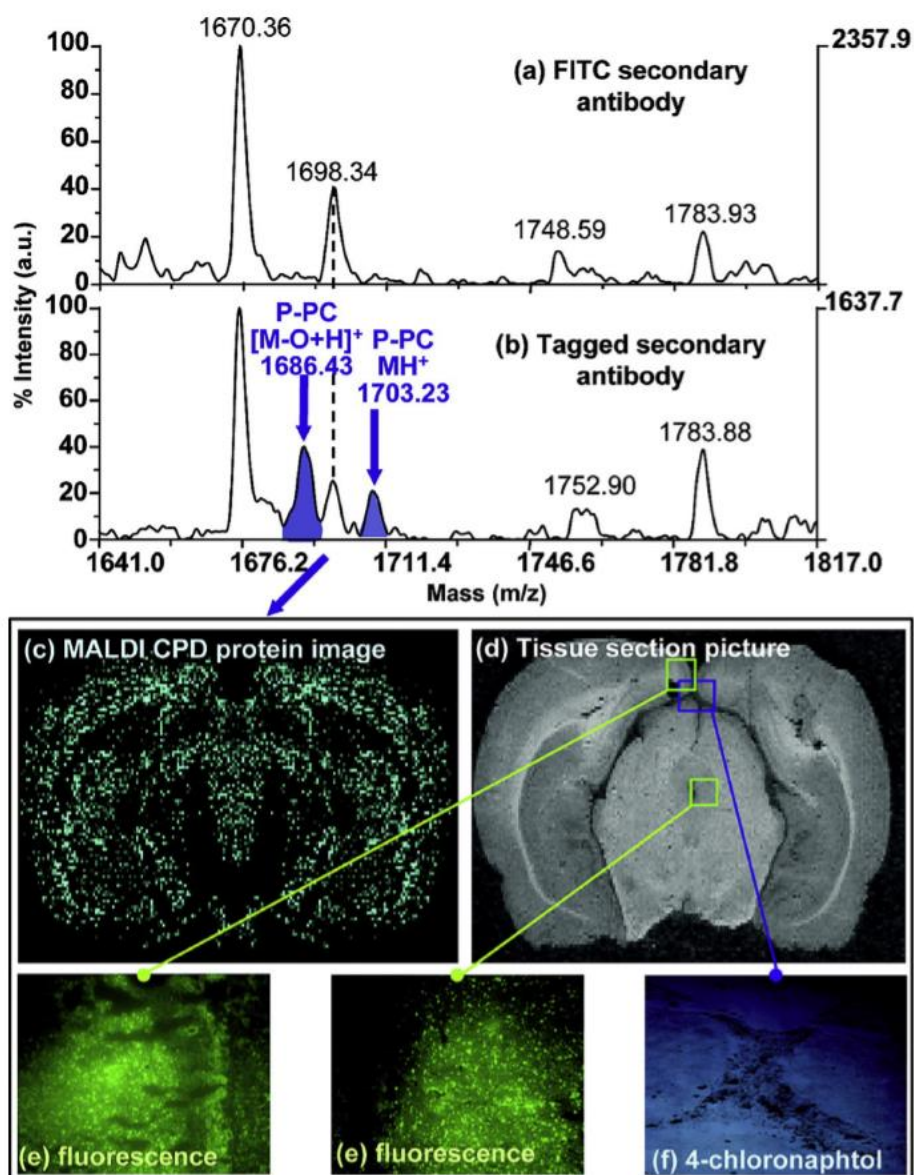


Fig. 5. Targeted matrix-assisted laser desorption/ionisation mass spectrometry (MALDI-MS) has been used to localise carboxypeptidase D protein in rat brain tissue sections. MALDI spectra of fluorescein isothiocyanate (FITC)-labelled (a) and mass-tagged (b) secondary antibodies corresponding to MALDI-MS (c), photographic (d) and fluorescence (e,f) images showing colocalisation of the mass tag photocleavage product (P-PC) m/z with anti-carboxypeptidase D (CPD) antibody binding. Reprinted with permission from Ref. [84].

cleaved from labelled antibodies by a scanning MALDI laser, thus enabling localisation of antibody binding sites on tissue sections (Fig. 5) [33,70,84,85]. Because MALDI/LDI methods can detect non-metal mass tags, the predicted costs of generating a wide range of antibody conjugates could be significantly lower compared to IMC. However, a present limitation to their wider application is the requirement for individual chemical syntheses for each mass tag of different m/z [33,68]. An effective solution is to incorporate photocleavable groups into peptide mass tags, which can then be easily modified by changing the amino acid or nucleotide sequence of the cleaved fragment. This tactic has been demonstrated using both solid-phase and automated peptide synthesis for duplexed imaging of ovarian cancer biomarkers [78,84].

Alternatively, notable examples of targeted LDI-MSI have focused on signal-enhancement strategies, such as the utilisation of multiple mass tags conjugated to avidin. These conjugates enabled labelling of biotinylated antibodies to breast cancer antigens with a higher number of tags without disrupting antigen-binding activity [70]. In this system, horseradish peroxidase and alkaline phosphatase-conjugated streptavidin/avidin can also be used for dual light microscopy and MS imaging because these enzymes catalyse formation of both coloured precipitates and amplified m/z signals from conventional IHC substrates [68,70]. Dendrimer-based mass tags for amplifying LDI signals from an activity-based probe, which binds to a specific enzyme receptor, have also been reported. Although this amplification strategy has yet to be demonstrated for antibodies; however, the authors pre-empted the imminent investigation of click chemistry as a tool for mass tag conjugation and amplification [33,86].

2.5. Challenges and outlook for labelled antibodies in clinical MS

More recent examples of immunoassays with MS detection have aimed to develop low-cost, transportable platforms that can be used for point-of-care diagnostic tests. For example, fluorescent mass-tagged aptamers and conductive chips have been utilised for multiplexed sensitive detection of cancer antigens using ESI-MS [87]. Other examples make use of simple mass tags that hydrolyse under mild alkaline conditions have been used to detect malaria and cancer biomarkers via a degradable paper-based device and nano-ESI-MS [88]. Paper spray MS (PSMS) methods such as this have notable clinical utility as they can be used to analyse small volumes, such as pin-prick blood samples, using handheld mass spectrometers [89].

Meanwhile, significant advancements in MALDI-MSI instrumentation, such as improved spatial resolution for cellular and sub-cellular localisation, will hopefully lead to the eventual matching of MSI capabilities with that of fluorescence microscopy [90]. The coupling of MALDI sources to high resolution mass analysers and additional separation techniques, such as ion mobility, has enabled localisation of proteins with improved selectivity, which will likely be a useful feature for targeted multiplexed imaging applications [91,92]. A range of additional ambient ionisation techniques, which require minimal sample preparation and have already been utilised for cell and plasma-based immunoassays, allude to new possibilities for MSI as a means of rapidly detecting biomarkers in clinical samples [87,92,93].

While these new developments offer promising solutions as detection methods for multiplexed immunoassays and IHC, such platforms are still limited by the cost and availability of commercially available labelling reagents, which often necessitates individual custom syntheses [33,87,94]. And like conventional immunoassays, each antibody must be carefully validated and to ensure specificity for target antigens and appropriate sensitivity for the given detection method [95]. Traditional detection methods

will therefore continue have an important role in the validation of potential antibody panels for MS biomarker detection. Importantly, much of the existing literature is focussed on method development and the demonstration of novel techniques, with fewer examples applying them to broader clinical applications or aiming to establish standardised workflows [33,96,97]. As such, there is a need for larger translational and preclinical studies to establish practical and reproducible ways of utilizing these approaches as viable diagnostic tools.

3. Summary

Experienced pathologists can easily diagnose many important human diseases using basic histological techniques [31]. Hence, the development of MS methodology for clinical applications should be focused on conditions for which accurate clinical assessment requires greater specificity and/or multiplexed analysis. The potential for MS technology to detect biomolecules that are not amenable to immunochemical detection and analyse multiple biomarkers simultaneously makes it an ideal platform for diagnosing multifactorial diseases that have proven challenging using conventional approaches. However, the inherent limitations of common MS detection methods for complex biological samples preclude their wider clinical adoption. Hyphenated immunochemical-MS approaches have offered promising solutions to some of these limitations, such as increased sensitivity, specificity, dynamic range, in addition to facilitating affordable and quantitative multiplexed analyses.

There is a growing emphasis on delivering *personalised medicine*, which acknowledges the multitude of confounding factors—genetic, environmental or otherwise—that can influence patients' individual risks of developing a disease and responses to treatment [98]. Diagnostic tests that integrate different technologies and evaluate multiple biomarkers will be required to facilitate improved clinical outcomes for a range of human diseases [99]. Of these, multiplexed MSIA technologies offer the potential for improved diagnostic performance, as such tests will not be reliant on a single marker or proteoform and therefore more effective for assessing diseases with high patient-to-patient variability. Likewise, the simultaneous quantification and localisation of low-abundance biomarkers, which is difficult using IHC methods based on staining intensities or cell counts, is practically feasible using MSI and is increasingly recognised as an important tool in clinical research [30,100].

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. KGS acknowledges financial support from a University of Adelaide Faculty of Sciences Divisional Scholarship.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] N.L. Henry, D.F. Hayes, Cancer biomarkers, *Mol. Oncol.* 6 (2) (2012) 140–146.
- [2] N. Rifai, M.A. Gillette, S.A. Carr, Protein biomarker discovery and validation: the long and uncertain path to clinical utility, *Nat. Biotechnol.* 24 (8) (2006) 971–983.
- [3] C.A. Crutchfield, et al., Advances in mass spectrometry-based clinical biomarker discovery, *Clin. Proteomics* 13 (1) (2016) 1.

- [4] D.S. Hage, Immunoassays, *Anal. Chem.* 71 (12) (1999) 294–304.
- [5] S.K. Vashist, J.H. Luong, *Handbook of Immunoassay Technologies: Approaches, Performances, and Applications*, Academic Press, 2018.
- [6] J. Smolsky, et al., Surface-enhanced Raman scattering-based immunoassay technologies for detection of disease biomarkers, *Biosensors* 7 (1) (2017) 7.
- [7] H.A.H. Rongen, A. Bult, W.P. van Bennekom, Liposomes and immunoassays, *J. Immunol. Methods* 204 (2) (1997) 105–133.
- [8] L. Chang, J. Li, L. Wang, Immuno-PCR: an ultrasensitive immunoassay for biomolecular detection, *Anal. Chim. Acta* 910 (2016) 12–24.
- [9] Z. Tang, Z. Ma, Multiple functional strategies for amplifying sensitivity of amperometric immunoassay for tumor markers: a review, *Biosens. Bioelectron.* 98 (2017) 100–112.
- [10] L.L. De Matos, et al., Immunohistochemistry as an important tool in biomarkers detection and clinical practice, *Biomark. Insights* 5 (2010) S2185. *BMI*.
- [11] J. Duraiyan, et al., Applications of immunohistochemistry, *J. Pharm. BioAllied Sci.* 4 (Suppl 2) (2012) S307–S309.
- [12] R. Liu, et al., Inductively coupled plasma mass spectrometry-based immunoassay: a review, *Mass Spectrom. Rev.* 33 (5) (2014) 373–393.
- [13] A.W. Fung, et al., Emerging role of clinical mass spectrometry in pathology, *J. Clin. Pathol.* 73 (2) (2020) 61–69.
- [14] J. Griffiths, A brief history of mass spectrometry, *Anal. Chem.* 80 (15) (2008) 5678–5683.
- [15] S.K. Grebe, R.J. Singh, LC-MS/MS in the clinical laboratory—where to here? *Clin. Biochem. Rev.* 32 (1) (2011) 5.
- [16] Y.-K. Chong, et al., Clinical mass spectrometry in the bioinformatics era: a hitchhiker's guide, *Comput. Struct. Biotechnol. J.* 16 (2018) 316–334.
- [17] P.J. Jannetto, R.L. Fitzgerald, Effective use of mass spectrometry in the clinical laboratory, *Clin. Chem.* 62 (1) (2016) 92–98.
- [18] A.H.B. Wu, D. French, Implementation of liquid chromatography/mass spectrometry into the clinical laboratory, *Clin. Chim. Acta* 420 (2013) 4–10.
- [19] U.A. Kiernan, et al., Quantitative mass spectrometry evaluation of human retinol binding protein 4 and related variants, *PLoS One* 6 (3) (2011), e17282.
- [20] R.L. Fitzgerald, D. Herold, Serum total testosterone: immunoassay compared with negative chemical ionization gas chromatography-mass spectrometry, *Clin. Chem.* 42 (5) (1996) 749–755.
- [21] C. Shackleton, O.J. Pozo, J. Marcos, GC/MS in recent years has defined the normal and clinically disordered steroidome: will it soon be surpassed by LC/tandem MS in this role? *JES* 2 (8) (2018) 974–996.
- [22] N. Krone, et al., Gas chromatography/mass spectrometry (GC/MS) remains a pre-eminent discovery tool in clinical steroid investigations even in the era of fast liquid chromatography tandem mass spectrometry (LC/MS/MS), *J. Steroid Biochem.* 121 (3) (2010) 496–504.
- [23] B. Shushan, A review of clinical diagnostic applications of liquid chromatography–tandem mass spectrometry, *Mass Spectrom. Rev.* 29 (6) (2010) 930–944.
- [24] J.G. Villoria, et al., Neonatal screening for inherited metabolic diseases in 2016, *Semin. Pediatr. Neurol.* 23 (4) (2016) 257–272.
- [25] U. Garg, M. Dasouki, Expanded newborn screening of inherited metabolic disorders by tandem mass spectrometry: clinical and laboratory aspects, *Clin. Biochem.* 39 (4) (2006) 315–332.
- [26] R. Holland, et al., Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.* 10 (10) (1996) 1227–1232.
- [27] R. Patel, MALDI-TOF MS for the diagnosis of infectious diseases, *Clin. Chem.* 61 (1) (2015) 100–111.
- [28] L. Xiu, et al., Establishment and application of a universal coronavirus screening method using MALDI-TOF mass spectrometry, *Front. Microbiol.* 8 (2017) 2017.
- [29] R.M. Caprioli, T.B. Farmer, J. Gile, Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS, *Anal. Chem.* 69 (23) (1997) 4751–4760.
- [30] J.Y. Yang, D.A. Herold, Chapter 13 - evolving platforms for clinical mass spectrometry, in: H. Nair, W. Clarke (Editors), *Mass Spectrometry for the Clinical Laboratory*, Academic Press, San Diego, 2017, pp. 261–276.
- [31] F. Leung, et al., Mass spectrometry-based tissue imaging: the next frontier in clinical diagnostics? *Clin. Chem.* 65 (4) (2020) 510–513.
- [32] A.R. Buchberger, et al., Mass spectrometry imaging: a review of emerging advancements and future insights, *Anal. Chem.* 90 (1) (2018) 240–265.
- [33] H. Gagnon, et al., Targeted mass spectrometry imaging: specific targeting mass spectrometry imaging technologies from history to perspective, *Prog. Histochem. Cytochem.* 47 (3) (2012) 133–174.
- [34] F.G. Strathmann, A.N. Hoofnagle, Current and future applications of mass spectrometry to the clinical laboratory, *Am. J. Clin. Pathol.* 136 (4) (2011) 609–616.
- [35] T.M. Annesley, Ion suppression in mass spectrometry, *Clin. Chem.* 49 (7) (2003) 1041–1044.
- [36] A.N. Hoofnagle, M.H. Wener, The fundamental flaws of immunoassays and potential solutions using tandem mass spectrometry, *J. Immunol. Methods* 347 (1) (2009) 3–11.
- [37] I. Rzagalinski, et al., Toward higher sensitivity in quantitative MALDI imaging mass spectrometry of CNS drugs using a nonpolar matrix, *Anal. Chem.* 90 (21) (2018) 12592–12600.
- [38] P.-M. Vaysse, et al., Mass spectrometry imaging for clinical research—latest developments, applications, and current limitations, *Analyst* 142 (15) (2017) 2690–2712.
- [39] S. Lehmann, et al., Clinical mass spectrometry proteomics (cMSP) for medical laboratory: what does the future hold? *Clin. Chim. Acta* 467 (2017) 51–58.
- [40] L. Sylvain, et al., Quantitative clinical chemistry proteomics (qCCP) using mass spectrometry: general characteristics and application, *CCLM* 51 (5) (2013) 919–935.
- [41] M. Mann, Functional and quantitative proteomics using SILAC, *Nat. Rev. Mol. Cell Biol.* 7 (12) (2006) 952–958.
- [42] O. Chahrour, D. Cobice, J. Malone, Stable isotope labelling methods in mass spectrometry-based quantitative proteomics, *J. Pharmaceut. Biomed.* 113 (2015) 2–20.
- [43] S. Wiese, et al., Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research, *Proteomics* 7 (3) (2007) 340–350.
- [44] S.P. Gygi, et al., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, *Nat. Biotechnol.* 17 (10) (1999) 994–999.
- [45] A. Sanz-Medel, et al., ICP-MS for absolute quantification of proteins for heteroatom-tagged, targeted proteomics, *Trac. Trends Anal. Chem.* 40 (2012) 52–63.
- [46] O. Ornatsky, et al., Highly multiparametric analysis by mass cytometry, *J. Immunol. Methods* 361 (1–2) (2010) 1–20.
- [47] D. Nedelkov, Human proteoforms as new targets for clinical mass spectrometry protein tests, *Expert Rev. Proteomics* 14 (8) (2017) 691–699.
- [48] Z. Zhang, H. Pan, X. Chen, Mass spectrometry for structural characterization of therapeutic antibodies, *Mass Spectrom. Rev.* 28 (1) (2009) 147–176.
- [49] C. Bich, et al., Characterization of antibody–antigen interactions: comparison between surface plasmon resonance measurements and high-mass matrix-assisted laser desorption/ionization mass spectrometry, *Anal. Biochem.* 375 (1) (2008) 35–45.
- [50] D. Nedelkov, R.W. Nelson, Surface plasmon resonance mass spectrometry: recent progress and outlooks, *Trends Biotechnol.* 21 (7) (2003) 301–305.
- [51] J.-F. Masson, Surface plasmon resonance clinical biosensors for medical diagnostics, *ACS Sens.* 2 (1) (2017) 16–30.
- [52] M.-S. Gauthier, et al., A semi-automated mass spectrometric immunoassay coupled to selected reaction monitoring (MSIA–SRM) reveals novel relationships between circulating PCSK9 and metabolic phenotypes in patient cohorts, *Methods* 81 (2015) 66–73.
- [53] U.A. Kiernan, et al., Comparative phenotypic analyses of human plasma and urinary retinol binding protein using mass spectrometric immunoassay, *Biochem. Biophys. Res. Commun.* 297 (2) (2002) 401–405.
- [54] R.W. Nelson, et al., Mass spectrometric immunoassay, *Anal. Chem.* 67 (7) (1995) 1153–1158.
- [55] K. Ueda, et al., Antibody-coupled monolithic silica microtips for high-throughput molecular profiling of circulating exosomes, *Sci. Rep.* 4 (2014) 6232.
- [56] H. Yassine, et al., Mass spectrometric immunoassay and MRM as targeted MS-based quantitative approaches in biomarker development: potential applications to cardiovascular disease and diabetes, *Proteomics Clin. Appl.* 7 (7–8) (2013) 528–540.
- [57] B. Krastins, et al., Rapid development of sensitive, high-throughput, quantitative and highly selective mass spectrometric targeted immunoassays for clinically important proteins in human plasma and serum, *Clin. Biochem.* 46 (6) (2013) 399–410.
- [58] F. Klont, et al., Assuring consistent performance of an insulin-like growth factor 1 MALDI immunoassay by monitoring measurement quality indicators, *Anal. Chem.* 89 (11) (2017) 6188–6195.
- [59] D. Nedelkov, Mass spectrometry-based immunoassays for the next phase of clinical applications, *Expert Rev. Proteomics* 3 (6) (2006) 631–640.
- [60] O. Trenchevska, R.W. Nelson, D. Nedelkov, Mass spectrometric immunoassays in characterization of clinically significant proteoforms, *Proteomes* 4 (1) (2016) 13.
- [61] F. Weiß, et al., Catch and measure—mass spectrometry-based immunoassays in biomarker research, *BBA-Protein Proteomics* 1844 (5) (2014) 927–932.
- [62] O. Trenchevska, R.W. Nelson, D. Nedelkov, Mass spectrometric immunoassays for discovery, screening and quantification of clinically relevant proteoforms, *Bioanalysis* 8 (15) (2016) 1623–1633.
- [63] N.L. Anderson, et al., Mass spectrometric quantitation of peptides and proteins using stable isotope standards and capture by anti-peptide antibodies (SISCAPA), *J. Proteome Res.* 3 (2) (2004) 235–244.
- [64] E. Kuhn, et al., Developing multiplexed assays for troponin I and interleukin-33 in plasma by peptide immunoaffinity enrichment and targeted mass spectrometry, *Clin. Chem.* 55 (6) (2009) 1108–1117.
- [65] E.N. Warren, et al., Absolute quantitation of cancer-related proteins using an MS-based peptide chip, *BioTechniques* 38 (6S) (2005) S7–S11.
- [66] G.T. Hermanson, in: *Bioconjugate Techniques*, second ed., Academic Press, New York, 2008.
- [67] K. Tsuchikama, Z. An, Antibody–drug conjugates: recent advances in conjugation and linker chemistries, *Protein Cell* 9 (1) (2018) 33–46.
- [68] G. Thiery, et al., Improvements of TArgeted multiplex mass spectrometry IMaging, *Proteomics* 8 (18) (2008) 3725–3734.
- [69] G. Han, et al., Metal-isotope-tagged monoclonal antibodies for high-dimensional mass cytometry, *Nat. Protoc.* 13 (10) (2018) 2121–2148.

- [70] G. Thiery, et al., Targeted multiplex imaging mass spectrometry with single chain fragment variable (SCFV) recombinant antibodies, *J. Am. Soc. Mass Spectrom.* 23 (10) (2012) 1689–1696.
- [71] J.R. McCombs, S.C. Owen, Antibody drug conjugates: design and selection of linker, payload and conjugation chemistry, *AAPS J.* 17 (2) (2015) 339–351.
- [72] C. Zhang, et al., A novel combination of immunoreaction and ICP-MS as a hyphenated technique for the determination of thyroid-stimulating hormone (TSH) in human serum, *J. Anal. Atom. Spectrom.* 16 (12) (2001) 1393–1396.
- [73] M. Lorey, et al., Mass-tag enhanced immuno-laser desorption/ionization mass spectrometry for sensitive detection of intact protein antigens, *Anal. Chem.* 87 (10) (2015) 5255–5262.
- [74] G. Thiery-Lavenant, A.I. Zavalin, R.M. Caprioli, Targeted multiplex imaging mass spectrometry in transmission geometry for subcellular spatial resolution, *J. Am. Soc. Mass Spectrom.* 24 (4) (2013) 609–614.
- [75] D.R. Bandura, et al., Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry, *Anal. Chem.* 81 (16) (2009) 6813–6822.
- [76] N. Kang, et al., Design and synthesis of new mass tags for matrix-free laser desorption ionization mass spectrometry (LDI-MS) based on 6,11-dihydrothiochromeno[4,3-b]indole, *Tetrahedron* 72 (36) (2016) 5612–5619.
- [77] J. Stauber, et al., Polymerase chain reaction and immunoassay–matrix assisted laser desorption mass spectrometry using Tag-Mass technology: new tools to break down quantification limits and multiplexes, *Anal. Chem.* 81 (22) (2009) 9512–9521.
- [78] J. Stauber, et al., Specific MALDI-MSI: tag-Mass, in: *Mass Spectrometry Imaging*, Springer, 2010, pp. 339–361.
- [79] Q. Chang, et al., Imaging mass cytometry, *Cytometry Part A* 91 (2) (2017) 160–169.
- [80] C. Giesen, et al., History of inductively coupled plasma mass spectrometry-based immunoassays, *Spectrochim. Acta B* 76 (2012) 27–39.
- [81] R. Gadalla, et al., Validation of CyTOF against flow cytometry for immunological studies and monitoring of human cancer clinical trials, *Front. Oncol.* 9 (415) (2019).
- [82] C. Giesen, et al., Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry, *Nat. Methods* 11 (4) (2014) 417–422.
- [83] D.E. Carvajal-Hausdorf, et al., Multiplexed (18-plex) measurement of signaling targets and cytotoxic T cells in trastuzumab-treated patients using imaging mass cytometry, *Clin. Cancer Res.* 25 (10) (2019) 3054–3062.
- [84] R. Lemaire, et al., Tag-Mass: specific molecular imaging of transcriptome and proteome by mass spectrometry based on photocleavable tag, *J. Proteome Res.* 6 (6) (2007) 2057–2067.
- [85] G. Thiery, et al., Multiplex target protein imaging in tissue sections by mass spectrometry – TAMSIM, *Rapid Commun. Mass Spectrom.* 21 (6) (2007) 823–829.
- [86] J. Yang, et al., Activity-based probes linked with laser-cleavable mass tags for signal amplification in imaging mass spectrometry: analysis of serine hydrolase enzymes in mammalian tissue, *Anal. Chem.* 84 (8) (2012) 3689–3695.
- [87] S. Xu, et al., Ultrasensitive ambient mass spectrometry immunoassays: multiplexed detection of proteins in serum and on cell surfaces, *JACS* 141 (1) (2019) 72–75.
- [88] S. Chen, Q. Wan, A.K. Badu-Tawiah, Mass spectrometry for paper-based immunoassays: toward on-demand diagnosis, *J. Am. Chem. Soc.* 138 (20) (2016) 6356–6359.
- [89] H. Wang, et al., Paper spray for direct analysis of complex mixtures using mass spectrometry, *Angew. Chem.* 49 (5) (2010) 877–880.
- [90] K. Ščupáková, et al., Cellular resolution in clinical MALDI mass spectrometry imaging: the latest advancements and current challenges, *CCLM* 58 (6) (2020) 914.
- [91] I. Piga, et al., Ultra-high resolution MALDI-FTICR-MSI analysis of intact proteins in mouse and human pancreas tissue, *Int. J. Mass Spectrom.* 437 (2019) 10–16.
- [92] M. Sans, C.L. Feider, L.S. Eberlin, Advances in mass spectrometry imaging coupled to ion mobility spectrometry for enhanced imaging of biological tissues, *Curr. Opin. Chem. Biol.* 42 (2018) 138–146.
- [93] N. Li, et al., Recent advances of ambient ionization mass spectrometry imaging in clinical research, *J. Separ. Sci.* 43 (15) (2020) 3146–3163.
- [94] R. Longuespée, et al., Spectroimmunohistochemistry: a novel form of MALDI mass spectrometry imaging coupled to immunohistochemistry for tracking antibodies, *OMICS* 18 (2) (2014) 132–141.
- [95] P. Brodin, The biology of the cell—insights from mass cytometry, *FEBS J.* 286 (8) (2019) 1514–1522.
- [96] Q. Baca, et al., The road ahead: implementing mass cytometry in clinical studies, one cell at a time, *Cytometry Part B Clin. Cytometry* 92 (1) (2017) 10–11.
- [97] M. El Ayed, et al., MALDI imaging mass spectrometry in ovarian cancer for tracking, identifying, and validating biomarkers, *Med. Sci. Mon.* 16 (8) (2010) BR233–BR245.
- [98] T.T. Duarte, C.T. Spencer, Personalized proteomics: the future of precision medicine, *Proteomes* 4 (4) (2016) 29.
- [99] B. Zhang, B. Kuster, Proteomics is not an island: multi-omics integration is the key to understanding biological systems, *Mol. Cell. Proteomics* 18 (8 suppl 1) (2019) S1.
- [100] S.A. Bogen, A root cause analysis into the high error rate in clinical immunohistochemistry, *Appl. Immunohistochem. Mol. Morphol.* 27 (5) (2019) 329–338.
- [101] P. Minakshi, et al., Chapter 14 - single-cell proteomics: technology and applications, in: D. Barh, V. Azevedo (Editors), *Single-Cell Omics*, Academic Press, 2019, pp. 283–318.

Synthesis of modifiable photocleavable MALDI mass tags using copper click chemistry

Foreword

In Chapter 1, we established that a significant limitation of existing mass spectrometry (MS)-based immunoassay platforms is the amount of time and effort required to modify the structure of mass tags, which is necessary for simultaneous analysis of multiple biomarkers. We hypothesised that a simple way to overcome this issue is to design mass tags that can be modified by conjugation to peptides via simple bio-orthogonal chemistries. Our prototype design for a MS-cleavable linker therefore included of an alkyne handle for copper-catalysed azide-alkyne cycloaddition ('copper click') to an azide-containing peptide. The linker also featured an *ortho*-nitrobenzyl ether moiety, to enable cleavage of the conjugated peptide by the 355 nm laser of a matrix-assisted laser desorption/ionisation (MALDI)-MS. Attachment of these mass tags to proteins was facilitated by an *N*-hydroxysuccinimide (NHS) ester; however, we also demonstrated the capacity to easily modify the linker structure to include various other reactive groups, which could be used to incorporate other biomolecules by exploiting their various reactivities. The NHS linker was then used to label an antibody reactive to a biomarker associated with several human neurodegenerative and psychiatric diseases, brain-derived neurotrophic factor, and we were able to detect the cleaved mass tag using MALDI-MS. The success of these preliminary experiments prompted our continued investigations of different conjugation chemistries and MS-cleavable groups, which could improve the utility of this platform.

Statement of Authorship

Title of Paper	Synthesis of modifiable photocleavable MALDI mass tags using copper click chemistry
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input checked="" type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Article formatted for intended publication in <i>Bioconjugate Chemistry</i> peer-reviewed journal.

Principal Author

Name of Principal Author (Candidate)	Katherine G Stevens		
Contribution to the Paper	Collected majority of experimental data, author of main body text		
Overall percentage (%)	75		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	1/12/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Aimee J Horsfall		
Contribution to the Paper	Solid-phase synthesis of peptide, collection of peptide synthetic characterization data, authorship of relevant experimental section		
Signature		Date	04/12/20

Name of Co-Author	River J Pachulicz		
Contribution to the Paper	Conducted parts of antibody conjugation, copper click and MALDI experiments under supervision of KGS		
Signature		Date	04/12/20

Please cut and paste additional co-author panels here as required.

Statement of Authorship

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Principle supervisor to AJH, co-supervisor to KGS, provided advice and financial support for small-molecule and solid-phase-peptide syntheses		
Signature		Date	3/12/2020

Name of Co-Author	Tara L Pukala		
Contribution to the Paper	Provided advice related to experimental design and project aims, editing of manuscript		
Signature		Date	3/12/2020

Please cut and paste additional co-author panels ~~here~~ as required.

Synthesis of modifiable photocleavable MALDI mass tags using copper click chemistry

Katherine G. Stevens¹, Aimee J. Horsfall¹, River J. Pachulicz¹, Andrew A. Abell¹ and Tara L. Pukala^{1*}

¹The University of Adelaide, Adelaide, Australia

*Corresponding author:

Tara Pukala

School of Physical Sciences

The University of Adelaide, Australia 5005

Ph: +61 8 8313 5497

E-mail: tara.pukala@adelaide.edu.au

Word count: 4090

Figures: 5

Abstract

Growing recognition of the independent limitations of both conventional immunoassays and mass spectrometry for biomarker detection and quantification has led to the development of hyphenated techniques, which combine antibody-based biomarker targeting with mass spectrometric detection. One branch of this broad area of research is focused on the development of chemical labels, or 'mass tags,' which improve the detection of large biomolecules, such as antibodies, to better enable their ionization and quantification using standard mass spectrometry instrumentation. A key consideration for this strategy is the ease with which these mass tags can be modified to enable variation in their structure and subsequent mass-to-charge ratio during analysis. The ability to easily synthesize a wide array of different tags would enable the simultaneous detection of multiple biomarkers. In the present study, we employ copper-catalyzed azide-alkyne cycloaddition chemistry to conjugate an azide-labelled peptide or fluorophore mass tag to proteins via a photocleavable linker, which releases the tag during matrix-assisted laser desorption/ionization (MALDI) mass spectrometric analysis. This linker is successfully used for the fluorescence and MALDI detection of an antibody, demonstrating its potential utility for identifying large intact biomolecules in mass spectrometry-based immunoassays.

Keywords: Mass spectrometry, Matrix-assisted laser desorption/ionization (MALDI), Copper click, Mass tag, Peptides, Antibodies, Immunoassay

Introduction

Mass spectrometry (MS), in its various forms, has become an important tool in modern clinical diagnostics.¹ Recently, chemical labelling with cleavable reporter ions, or *mass tags*, has been utilized to expand the scope of biomolecules that can be analyzed and quantified using widely available MS instruments.²⁻⁷ For liquid-chromatography (LC)-MS applications, mass tag reagents often include a collision-induced dissociation (CID)-labile linker, which generates a reporter ion of a given m/z that can be easily identified and tailored to fit within the dynamic range of most instruments.³ This concept has been extended to inductively coupled plasma (ICP), and matrix-assisted and matrix-free laser desorption/ionization (MALDI and LDI, respectively)-MS, which can retain the localization of analytes on a tissue section or array.⁸⁻¹⁰

Antibodies are useful detection agents for a wide range of significant biomarkers, which can be difficult to detect directly via MS owing to poor signal/noise ratios in biological samples.¹¹ However, heterogeneous proteins pose additional challenges for conventional MALDI instruments, which have limited mass resolution for high m/z ions.¹²⁻¹³ Mass tags are therefore useful for differentiating between multiple antibodies in a single experiment and enabling multiplexed biomarker detection. The most prevalent example of this strategy is the use of heavy metal-labelled antibodies for detection of biomarkers using ICP-MS.¹⁴ Nevertheless, this approach requires expensive reagents and the harsh ionization conditions retain limited information about intact biomolecules prior to atomization.¹⁵ An alternative strategy is to instead utilize organic photocleavable linkers, which produce characteristic fragments when exposed to the ultra-violet (UV) light of a MALDI laser.⁵⁻⁶

One major factor limiting the wider uptake of organic mass tags is the time and effort required to individually synthesize tags of different m/z , which precludes their application in high-throughput multiplexed experiments. Some examples make use of photocleavable linkers that can be incorporated into solid-phase and automated peptide synthesis, in which peptide-based mass tag m/z ratios can be modified by altering the amino acid sequence. However, these approaches still require additional time-consuming synthesis and purification steps. We proposed a solution to this issue involving the covalent attachment of peptide mass tags using the bio-orthogonal copper-catalyzed azide-alkyne cycloaddition (CuAAC) reaction. This approach enables the modification of a single alkyne-containing linker using azide-containing mass tags, which can be requested as part of commercial custom peptides, as well as found in many existing reagents. Furthermore, we demonstrate a workflow for antibody conjugation and purification that can be performed using common high-throughput protein purification techniques. This simplified mass tag analysis procedure thereby demonstrates the potential for targeted MALDI-MS as a clinical biomarker detection method.

Results and Discussion

The *ortho*-nitrobenzyl group was an obvious starting point for our photocleavable linker, given this group demonstrates efficient cleavage at wavelengths close to that of the high-frequency neodymium-doped yttrium aluminum garnet (Nd:YAG) lasers of most newer MALDI imaging systems (355 nm; **Figure 1**).¹⁶

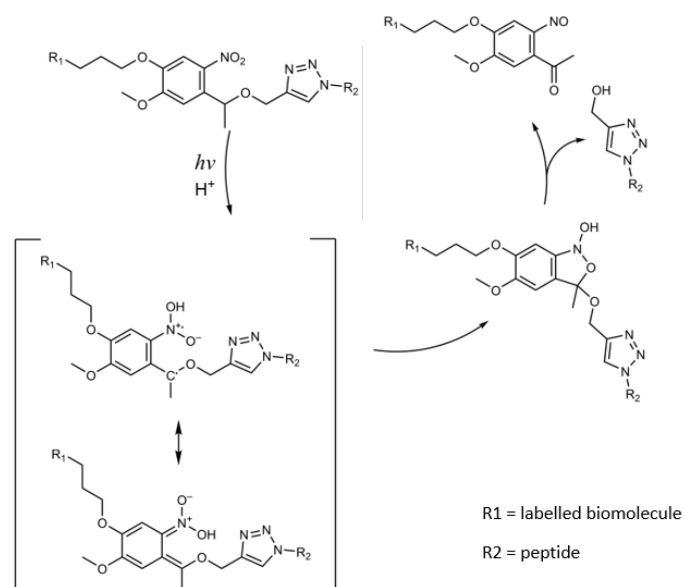


Figure 1. Photodissociation mechanism for an *o*-nitrobenzyl ether derivative showing release of 1,4-triazole moiety formed during copper-catalyzed azide-alkyne cycloaddition.¹⁷

O-nitrobenzyl derivatives also have highly tunable absorption maxima, which are predominantly influenced by altering substituents on the aromatic ring.¹⁸ The photoreactivity of these linkers is further enhanced by the attachment of mass tags via ether linkages, which are also stable in buffered aqueous solutions and therefore suitable for applications using native proteins and sensitive biomolecules.¹⁹ For these reasons, *o*-nitrobenzyl ethers have already been used as MALDI probes for targeted mass spectrometry imaging (MSI) of proteins.²⁰ For our novel linker design, we modified this structure to include a terminal alkyne, which enables conjugation to azide-containing peptides via copper-catalyzed azide-alkyne cycloaddition (CuAAC or *copper click*; **Figure 2**).^{21,22}

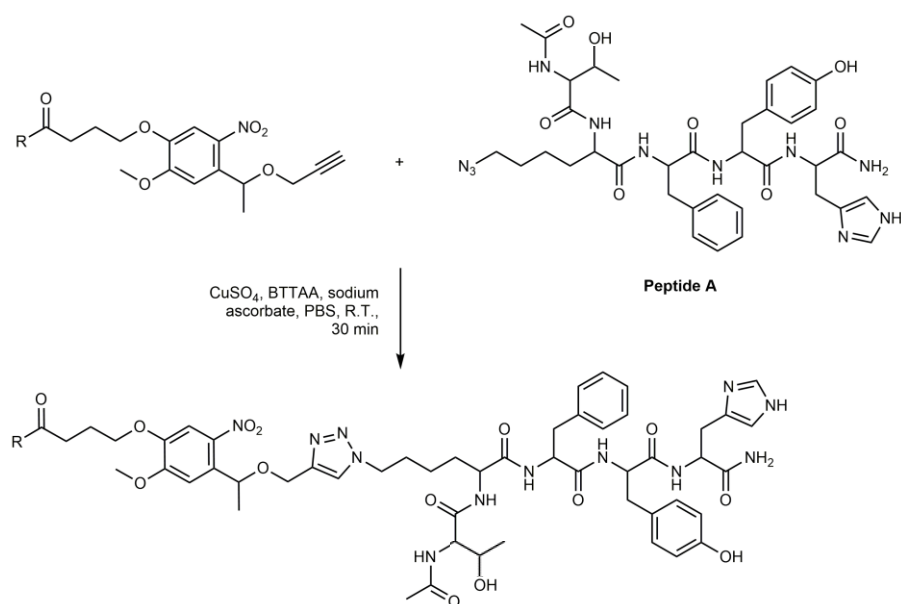


Figure 2. Conjugation of alkyne *o*-nitrobenzyl linker to an azide-labelled peptide via copper-catalyzed azide-alkyne cycloaddition.

Synthesis of these CuAAC linkers can easily be achieved by starting from the commercially available hydroxyethyl photolinker **2** (Merck, Darmstadt, Germany). However, we synthesized a large quantity of **2** using affordable and readily available reagents, by adapting a previously described method (see Supporting information).²³ Functionalization with a terminal alkyne was accomplished through subsequent alkylation with propargyl bromide. Amine-reactive *N*-hydroxysuccinimide (NHS) **4** and pentafluorophenyl (PFP) ester **6** linkers were synthesized via esterification of acid **3** with *N,N'*-disuccinimidyl carbonate (DSC) or *bis*(pentafluorophenyl) carbonate (PFP₂O), respectively. Alternatively, aldehyde-reactive hydrazide **5** was synthesized by methyl-esterification and subsequent hydrazinolysis (**Figure 3**).

Lysozyme was used as a model protein for optimization of bioconjugation experiments, owing to its affordability and 14.3 kDa molecular weight, which is conveniently within a suitable range for obtaining MALDI spectra of the intact protein with sufficient resolution to identify modifications, and also above the molecular weight cut-offs (MWCO) of common commercial protein purification devices. The protein was labelled with NHS linker **4**, which was then conjugated to an azide-labelled peptide, peptide **A** (ACTK[N₃]FYH-NH₂) via CuAAC under non-denaturing conditions. MALDI-MS spectra of the purified conjugate shows modifications to the intact protein, evident as additional peaks at intervals approximately equal to the 282 Da remaining fragment of the photocleaved linker after loss of the peptide (**Figure 4A**). The expected modified peptide photocleavage product was also observable at 818.5469, 840.5319 and 856.5111 *m/z* as its respective protonated molecular ion, sodium adduct and potassium adduct, indicating that both NHS coupling and CuAAC steps were successful (**Figure 4B**). The additional absence of peaks corresponding to unmodified peptide **A** suggests the purification procedures for removing unreacted reagents were also effective.

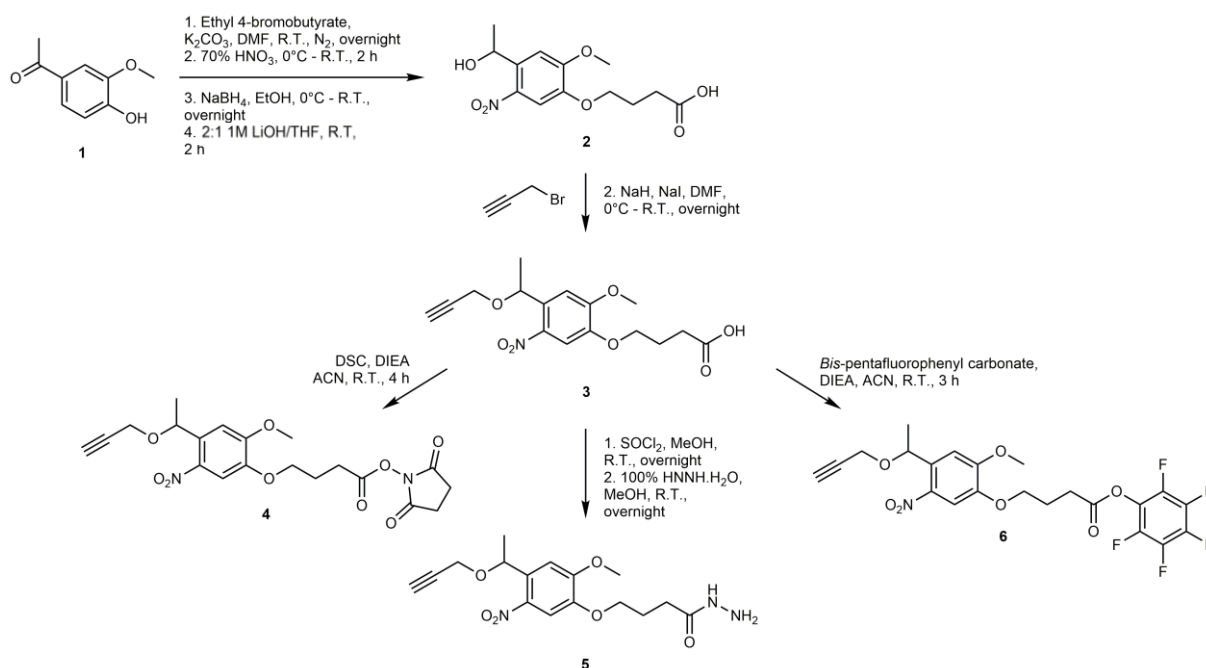


Figure 3. Synthesis of *o*-nitrobenzyl derivatives with amine (**4** and **6**) and oxidized glycan (**5**)-reactive functional groups.

These conjugation methods were then adapted for the generation of a mass-tagged biomarker-specific antibody for the detection of brain-derived neurotrophic factor (BDNF). BDNF was an attractive candidate for our immunoassay method development, being considered a circulating marker of various neurodegenerative and psychiatric disorders, yet lacking a validated clinical assay.²⁴

Successful conjugation of linker **4** to an anti-BDNF antibody was visualized using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) following CuAAC to the azide-labelled fluorophore cyanine-3 (Cy3 azide; **Figure 5A**). The Cy3-labelled band appears fainter when visualized using Coomassie Brilliant Blue staining, which could result from reduced interactions between the anionic Coomassie dye and lysine residues after their modification with the negatively charged Cy3 label.²⁵ However, the presence of Cy3-labelled bands at approximately 25, 50 and 150 kDa, which correspond to the light and heavy chain fragments and residual non-reduced immunoglobulin, respectively, indicates successful conjugation of the fluorescent label to the antibody. We were unable to identify the intact molecular ion for the cleaved fluorophore using MALDI-MS, possibly owing to decomposition of the cyanine fluorophore resulting from the acidic matrix conditions and laser desorption.²⁶ However, these results also suggest that this linker could be used to create multimodal probes if fluorophores with suitable properties for MS detection are used, or following further optimization of sample preparation and analysis conditions.

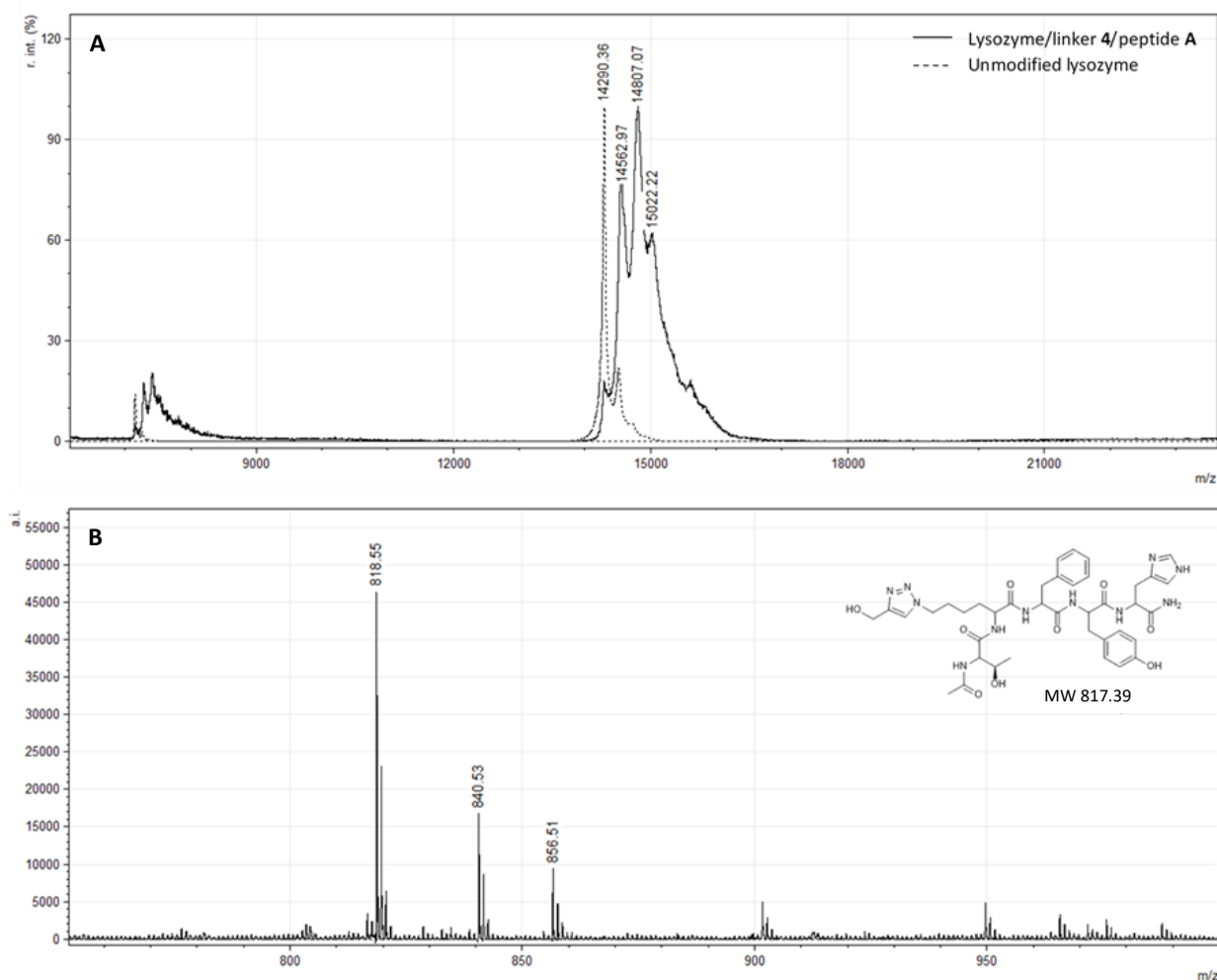


Figure 4. Matrix-assisted laser desorption/ionization mass spectra of lysozyme labelled with photocleavable linker **4** and conjugated to peptide **A**, showing unmodified and modified intact protein (A) and photocleaved peptide (B).

Next, the anti-BDNF antibody was modified with NHS linker **4** and peptide **A**, to facilitate its detection using MALDI-MS, resulting in a spectrum that, above m/z 700 (the lower threshold for most CHCA matrix background ions), was dominated by ions corresponding to the photocleaved peptide (**Figure 5B**).

These preliminary results demonstrate that NHS-based linker **4** can be applied to the MALDI-MS detection of antibodies. However, further investigation is required to determine the effects that this labelling strategy will have on the antigen-binding activity of antibodies, with modification of lysine residues in antigen-binding regions potentially leading to loss of activity.²⁷ As alternatives, linkers **5** and **6** employ alternative respective labelling chemistries: (1) hydrazides, which are often used for labelling antibodies owing to their specificity for oxidized glycan residues and are therefore less likely to alter antigen binding regions, and (2) PFP esters, which are reportedly less prone to aqueous hydrolysis compared to NHS esters.²⁸⁻²⁹ The conjugation of mass tags to antibodies using these other conjugation chemistries has yet to be optimized; however, we consider the application of our NHS linker to be a suitable preliminary study, owing to the common use of reactive esters in the existing literature.²⁷

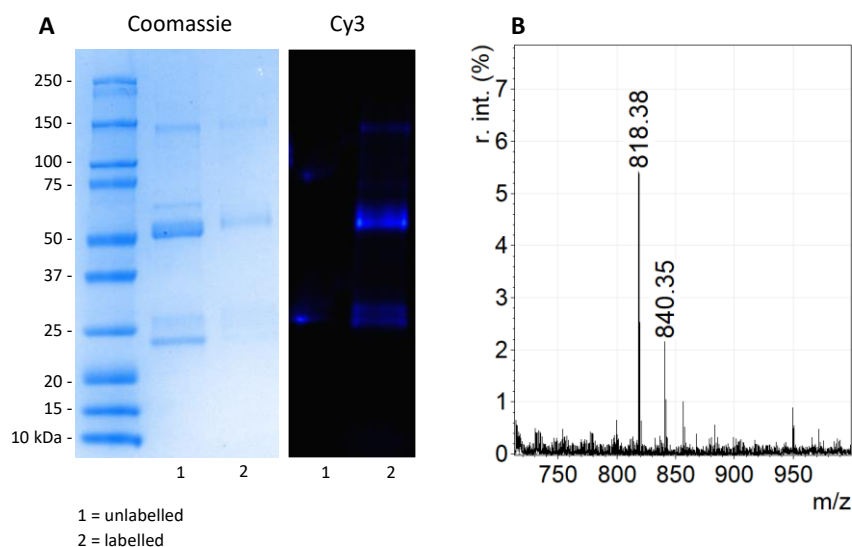


Figure 5. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (A) and matrix-assisted laser desorption/ionization spectra (B) of anti-brain-derived neurotrophic factor antibody labelled with linker 4 and conjugated to cyanine 3 (Cy3) dye and peptide A, respectively.

A major challenge for conventional immunoassay development is overcoming the costs, time and sample quantities required for individual biomarker analysis.³⁰ Coupled with the fact that many multifactorial diseases, such as cancers, require assessment of multiple markers for effective treatment monitoring, there is unsurprisingly a growing focus on developing and validating multiplexed assays.³¹ This interest in analyzing multiple biomarkers simultaneously has contributed to the increasing uptake of LC-MS in clinical settings.^{30, 32} However, several requirements of LC-MS limit its suitability for detecting protein biomarkers, including extensive sample preparation and chromatography steps.³³⁻³⁴ In terms of sample throughput, MALDI-MS overcomes many of the shortcomings of ESI-MS thanks to its tolerance for sample contaminants such as salts and detergents.³⁵ However, the tendency of MALDI to produce predominantly singly charged ions, combined with limited resolving power of the most common (linear TOF) mass analyzers at high m/z , prohibits its wider utility in the analysis of complex samples without additional offline fractionation or enrichment steps.³⁶⁻³⁸ The commercial pipette-tip-based mass spectrometric immunoassay (MSIATM) developed by Nelson and coworkers selectively enriches biomarkers using immobilized antibodies and can be used for detection of intact proteins eluted directly into MALDI matrix.^{33, 39} This approach thus eliminates the need for prior enzymatic digestion or time-consuming chromatography steps, but is limited to proteins that can be effectively ionized and have sufficient m/z resolution to differentiate between them using MALDI-MS. Alternatively, mass-tagged antibodies could be used to detect these immobilized biomarkers, followed by MALDI-MS analysis of the matrix eluent in the m/z range relevant to the photocleavage product. The number of biomarkers which could practically be measured simultaneously would therefore be limited by the number of different mass tags available, rather than the limited ability of MALDI to ionize and resolve intact high molecular-weight biomolecules.⁴⁰

Our CuAAC approach could also enable absolute quantitation of biomarkers by simplifying the generation of stable isotope-labelled ('heavy') analytical standards into the photocleavage product by incorporating isotopes into the linker structure, and therefore making individual syntheses of heavy peptides unnecessary. Furthermore, the versatility of the CuAAC reaction means that our strategy is not limited to the use of peptide mass tags, with the potential for modification of the photolinker structure using any other azide-containing molecule. The methods demonstrated here therefore provide valuable precedent for the application of click chemistry in antibody labelling and its potential to generate multimodal mass tags with tailored m/z and ionization properties.

Experimental Procedures

Synthesis of MALDI-cleavable linkers

General. Light-sensitive reactions and intermediates (compounds **2-6**) were protected from long-term exposure to light. All reagents were purchased from Sigma-Aldrich (Castle Hill, Australia) or Chem-Supply (Gillman, Australia) unless stated otherwise. Column chromatography was performed on 0.2-0.5 mm (Scharlau Science Group, Barcelona, ES) or 0.04-0.063 mm (Carl Roth GmbH & Co. KG, Karlsruhe, GE) silica gel 60 deactivated by washing with dilute triethylamine for acid-sensitive compounds **4** and **6**. 500 MHz ¹H and ¹³C nuclear magnetic resonance (NMR) spectra were obtained using an Agilent 500/54 Premium Shielded NMR spectrometer (Agilent Technologies, Santa Clara, USA). 600 MHz H and ¹³C NMR spectra were acquired using an Oxford 600 NMR spectrometer (Oxford Instruments, Abingdon, UK). ¹⁹F NMR spectra were obtained using a Bruker Avance III 500 MHz NMR spectrometer (Bruker Daltonics, Billerica, USA). High-resolution mass spectrometry (HRMS) data were obtained using an Agilent 6230 TOF LC/MS equipped with an Infinity 1260 LC system (Agilent Technologies).

4-(2-Methoxy-5-nitro-4-{1-[(prop-2-yn-1-yl)oxy]ethyl}phenoxy)butanoic acid (**3**). The synthesis of compound **2** was adapted from previously published methods (see Supporting Information).²³ 4-[4-(1-Hydroxyethyl)-2-methoxy-5-nitrophenoxy]butanoic acid (**2**; 6.59 g, 22 mmol) was dissolved in dry dimethylformamide (100 mL) and cooled to 0 °C. Sodium hydride (3.37 g, 140 mmol) was added to the stirred solution, followed by distilled propargyl bromide (4.71 g, 40 mmol) and sodium iodide (3.32 g, 22 mmol). The reaction was stirred overnight at room temperature, before quenching with water (100 mL). The suspension was then adjusted to pH 3 via addition of 1 M hydrochloric acid and extracted with 3:1 chloroform/isopropanol (3 x 50 mL). The combined organic extracts were washed with 10 % aqueous sodium thiosulfate (50 mL), then brine (50 mL) and dried over magnesium sulfate. The solvent was removed under reduced pressure and the crude residue purified via silica flash chromatography (20-50 % ethyl acetate in hexanes with 0.1-1 % formic acid) and recrystallized from hot ethanol to obtain the product as a yellow crystalline solid (3.55 g, 48 %). ¹H NMR (600 MHz, CDCl₃): δ 10.7 (br s, COOH), 7.59 (s, 1H, aromatic), 7.20 (s, 1H, aromatic), 5.38 (q, *J* = 6.2 Hz, 1H, CH(OCH₂CCH)CH₃), 4.14 (td, *J* = 6.2, 1.5 Hz, 2H, OCH₂CH₂CH₂COOH), 4.05 (dd, *J* = 15.5, 2.5 Hz, 1H, CH₂CCH), 3.96 (dd, *J* = 15.6, 2.4 Hz, 1H, CH₂CCH), 3.96 (s, 3H, OCH₃), 2.63 (t, *J* = 7.2 Hz, 2H, CH₂COOH), 2.40 (t, *J* = 2.4 Hz, 1H, CCH), 2.25 – 2.16 (m, 2H, CH₂CH₂COOH), 1.54 (d, *J* = 6.2 Hz, 3H, CH(OCH₂CCH)CH₃); ¹³C NMR (126 MHz, CDCl₃): δ 180.5 (COOH), 157.0 (aromatic), 149.7 (aromatic), 143.1 (aromatic), 137.3 (aromatic), 111.8 (aromatic), 111.3 (aromatic), 81.9 (CH₂CCH), 77.3 (CH₂CCH), 75.5 (CH(OCH₂CCH)CH₃), 70.7 (OCH₂CH₂CH₂COOH), 59.1 (OCH₃ and CH₂CCH), 32.8 (OCH₃ and CH₂CCH), 26.7 (OCH₂CH₂CH₂COOH), 26.1 (CH(OCH₂CCH)CH₃); HRMS: calcd. for (C₁₆H₁₉NO₇)Na⁺: 360.1054, found 360.1055.

N-Hydroxysuccinimidyl 4-(2-methoxy-5-nitro-4-{1-[(prop-2-yn-1-yl)oxy]ethyl}phenoxy)butanoate (**4**). 4-(2-Methoxy-5-nitro-4-{1-[(prop-2-yn-1-yl)oxy]ethyl}phenoxy)butanoic acid (**3**; 33 mg, 0.10 mmol), *N,N'*-diisopropylethylamine (DIEA; 68 μL, 0.39 mmol) and DSC (28 mg, 0.11 mmol) were dissolved in dry acetonitrile

(5 mL) and the reaction was stirred for 4 hours at room temperature. The reaction was cooled to -20 °C, then syringe-filtered and the solvent removed under a gentle stream of nitrogen gas. The crude residue was purified via silica flash chromatography (50-60 % ethyl acetate in petroleum benzene) to afford the product as a viscous yellow oil (21 mg, 48 %). ¹H NMR (500 MHz, CDCl₃): δ 7.61 (s, 1H, aromatic), 7.22 (s, 1H, aromatic), 5.38 (q, *J* = 6.2 Hz, 1H, CH(OH)CH₃), 4.19 (t, *J* = 6.0 Hz, 2H, OCH₂CH₂CH₂COOH), 4.06 (dd, *J* = 15.5, 2.4 Hz, 1H, CH₂CCH), 3.98 (s, 3H, OCH₃), 3.96 (dd, *J* = 15.5, 2.4 Hz, 1H, CH₂CCH), 2.95-2.83 (m, 6H, CH₂COO and NHS), 2.41 (t, *J* = 2.3 Hz, 1H, CCH), 2.31 (m, *J* = 6.7 Hz, 2H, CH₂CH₂COO), 1.55 (d, *J* = 6.2 Hz, 3H, CH(OH)CH₃); ¹³C NMR (126 MHz, CDCl₃): δ 171.7 (NHS), 170.8 (COO), 157.1 (aromatic), 149.6 (aromatic), 143.1 (aromatic), 137.5 (aromatic), 112.1 (aromatic), 111.3 (aromatic), 81.9 (CH₂CCH), 77.4 (CH₂CCH), 75.5 (CH(OCH₂CCH)CH₃), 70.1 (OCH₂CH₂CH₂COO), 59.1 (CH₂CCH), 59.0 (OCH₃), 30.2 (OCH₂CH₂CH₂COO), 28.3 (NHS CH₂), 26.8 (OCH₂CH₂CH₂COO), 26.1 (CH(OH)CH₃); HRMS: calcd. for (C₂₀H₂₂N₂O₉)Na⁺: 457.1218, found 457.1215.

4-(2-Methoxy-5-nitro-4-{1-[(prop-2-yn-1-yl)oxy]ethyl}phenoxy)butanehydrazide (5). 4-(2-Methoxy-5-nitro-4-{1-[(prop-2-yn-1-yl)oxy]ethyl}phenoxy)butanoic acid (**3**; 277 mg, 0.82 mmol) was dissolved in dry methanol (10 mL). Thionyl chloride (200 μL, 2.7 mmol) was added dropwise while stirring. The reaction was stirred overnight at room temperature, then the solvent removed *in vacuo* and the residue co-evaporated with methanol (3 x 10 mL) to obtain a waxy yellow solid. The crude product was redissolved in dry methanol (10 mL) before addition of 100 % hydrazine hydrate (100 μL, 0.82 mmol) and the reaction stirred at room temperature overnight. The solvent was removed *in vacuo* to obtain a viscous yellow oil (267 mg, 93 %). ¹H NMR (500 MHz, CDCl₃): δ 7.57 (s, 1H, aromatic), 7.21 (s, 1H, aromatic), 7.08 (br s, 1H, NH), 5.37 (q, *J* = 6.3 Hz, 1H, CH(OH)CH₃), 4.11 (td, *J* = 6.1, 2.0 Hz, 2H, OCH₂CH₂CH₂COO), 4.04 (dd, *J* = 15.5, 2.4 Hz, 1H, CH₂CCH), 3.98-3.94 (m, *J* = 2.4 Hz, 4H, OCH₃ and CH₂CCH), 2.44-2.38 (m, 3H, CCH and CH₂COO), 2.25 – 2.15 (m, 2H, CH₂CH₂COOH), 1.53 (d, *J* = 6.2 Hz, 3H, CH(OH)CH₃); ¹³C NMR (126 MHz, CDCl₃): δ 173.4 (COO), 154.6 (aromatic), 147.4 (aromatic), 140.9 (aromatic), 135.1 (aromatic), 109.5 (aromatic), 109.0 (aromatic), 79.6 (CH₂CCH), 75.1 (CH₂CCH), 73.2 (CH(OCH₂CCH)CH₃), 68.8 (OCH₂CH₂CH₂COO), 56.9-56.8 (OCH₃ and CH₂CCH), 31.3 (OCH₂CH₂CH₂COO), 25.0 (OCH₂CH₂CH₂COO), 23.8 (CH(OCH₂CCH)CH₃); HRMS: calcd. for (C₁₆H₁₉NO₇)Na⁺: 360.1054, found 360.1055; HRMS calcd. for (C₁₆H₂₁N₃O₆)H⁺: 352.1509, found 352.1511.

2,3,4,5,6-pentafluorophenyl 4-{2-methoxy-5-nitro-4-[1-(prop-2-yn-1-yloxy)ethyl]phenoxy}butanoate (6). 4-(2-Methoxy-5-nitro-4-{1-[(prop-2-yn-1-yl)oxy]ethyl}phenoxy)butanoic acid (**3**; 48 mg, 0.14 mmol), DIEA (104 μL, 0.60 mmol) and *bis*-pentafluorophenyl carbonate (64 mg, 0.16 mmol) were dissolved in dry acetonitrile (10 mL) and the reaction was stirred for 3 hours at room temperature. The solvent was removed *in vacuo* and the crude residue purified via silica flash chromatography (25-60 % ethyl acetate in petroleum benzene) to obtain a viscous red/brown oil (11.2 mg, 27%). ¹H NMR (599 MHz, CDCl₃): δ 7.60 (s, 1H, aromatic), 7.22 (s, 1H, aromatic), 5.38 (q, *J* = 6.2 Hz, 1H, CH(OH)CH₃), 4.18 (td, *J* = 6.1, 2.2 Hz, 2H, OCH₂CH₂CH₂COO), 4.05 (dd, *J* = 15.6, 2.4 Hz, 1H, CH₂CCH), 3.98-3.95 (m, 4H, OCH₃ and CH₂CCH), 2.95 (t, *J* = 7.3 Hz, 2H, CH₂COOH), 2.40 (t, *J* = 2.4 Hz, 1H, CCH), 2.35-2.30 (m, 2H, CH₂CH₂COO), 1.54 (d, *J* = 6.3 Hz, 3H, CH(OH)CH₃); ¹³C NMR (151 MHz, CDCl₃): δ 169.0 (COO) 154.57 (aromatic), 147.08 (aromatic), 142.5-141.7 (m, PFP), 140.6-140.3 (m, CNO₂ and PFP), 138.9-138.6

(m, PFP), 137.3-137.0 (m, PFP), 135.1 (aromatic), 109.5 (aromatic), 108.8 (aromatic), 79.3 (CH₂CCH), 74.8 (CH₂CCH), 72.9 (CH(OCH₂CCH)CH₃), 67.9 (OCH₂CH₂CH₂COO), 56.6 (OCH₃ and CH₂CCH), 30.0 (OCH₂CH₂CH₂COO), 24.3 (OCH₂CH₂CH₂COO), 23.6 (CH(OCH₂CCH)CH₃); HRMS calcd. for (C₂₂H₁₈F₅NO₇)H⁺: 504.1076, found 504.1026.

Synthesis and characterization of azide-labelled peptide

Unless otherwise indicated, all starting materials were purchased from commercial sources and used without further purification. All L-amino acids purchased from Chem Impex International. RP-HPLC solvents were buffer A: H₂O with 0.1 % TFA, and buffer B: ACN with 0.1 % TFA; 0.2 μm filtered. Purity of all compounds was determined by analytical RP-HPLC on an Agilent 1260 Infinity HPLC equipped with a Phenomenex Luna C18(2) column (250 x 4.6 mm) over a linear gradient of 0-100 % B (15 min). Mass spectra for characterisation were collected using a Bruker HCT Ultra (Bruker Daltonics) via direct injection, and ACN with 0.1 % formic acid as the running buffer.

Solid-phase synthesis of peptide Ac-TK(N₃)FYH-NH₂ (peptide A). The following Fmoc-protected amino-acids were used for all peptides unless otherwise specified. All amino-acids were purchased from Chem-Impex Int'l.: Fmoc-L-Thr(tBu)-OH Fmoc-L-Lys(ε-N₃)-OH, Fmoc-Phe-OH, Fmoc-Tyr(tBu)-OH, Fmoc-His(Trt)-OH. Rink Amide PL resin (0.1 mmol, 322 mg, 0.31 mmol/g, Agilent) was swollen in 1:1 *N,N*-dimethylformamide (DMF)/dichloromethane (10 mL) for 15 min. The mixture was transferred to a fritted syringe, the solution drained, and the resin washed with DMF (3 x 5 mL). The Fmoc-protecting group was removed by treatment of the resin with a solution of 20 % piperidine in DMF (5 mL) for 15 min. The solution was drained and the resin washed with DMF (3 x 5 mL). Amino-acid couplings were achieved by addition of a solution of Fmoc-protected amino-acid (5 equivalents), 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU; 5 equivalents) and DIEA (10 equivalents, 174 μL) in DMF (5 mL), to the resin and stirred intermittently for 1 h. The solution was drained and the resin washed with DMF (5 x 5 mL). The *N*-terminal Fmoc-protecting group was removed by treatment of the resin with a solution of 20 % piperidine and in DMF (5 mL) for 10 min, the solution was drained and the resin washed with DMF (5 x 5 mL). A TNBS test was used to verify each coupling (negative/colourless) and deprotection (positive/red) step, with steps repeated as necessary. Successive couplings and Fmoc-deprotections were repeated to achieve the desired sequence. After the final Fmoc-deprotection, the *N*-terminal amine was protected with an acetyl functionality by reaction with acetic anhydride (470 μL) and DIEA (870 μL) in DMF (10 mL) for 15 min. The resin was washed with DMF (3 x 5 mL), and dichloromethane (5 x 5 mL) then dried with Et₂O (3 x 5 mL). The peptide was cleaved from the resin by addition of 95:2.5:2.5 TFA/TIPS/H₂O (6 mL) to the resin and rocked for 1 h. The TFA solution was pipetted from the resin and concentrated to 0.5-1 mL under a nitrogen stream, then the peptide was precipitated with diethyl ether (10 mL) and the mixture cooled to -20°C. The precipitate was pelleted by centrifugation (7600 rpm, 10 min) and the supernatant decanted. The pellet was dried under a nitrogen stream, and then dissolved in 1:1 ACN/H₂O, before syringe filtering (0.2 μm) and lyophilised to give the crude peptide as a white fluffy powder (~40 % crude purity). ESI-MS calcd. for (C₃₆H₄₇N₁₁O₈)H⁺: 762.8, found: 762.6; R_t (C₁₈, 0-100 %, 20 min) 12.0 min.

TNBS Test. A small spatula of swollen resin was taken out and 1 drop each of TNBS (100 μ L 5 % w/v picrylsulfonic/trinitrobenzenesulfonic acid in H₂O added to 900 μ L of DMF) and DIEA solutions (100 μ L in 900 μ L of DMF) added and allowed to develop for 1 min. Clear/yellow beads indicated no free amine (negative), while red/orange beads showed free amine was present (positive).

Conjugation of photocleavable linkers to biomolecules

N-hydroxysuccinimide ester coupling to proteins. Twenty molar equivalents of NHS photolinker **4** in DMSO (5.26 μ L, \leq 5 % final concentration) was added to 1 mg/mL protein in 1x phosphate-buffered saline (PBS; 100 μ L) and the mixture incubated at room temperature in the dark for 30 min. The labelled protein was purified using Pierce™ Dye Removal Column resin (Thermo Fisher Scientific, Waltham, USA) and stored at -20°C until further use. NHS ester coupling to primary amines on proteins was first optimized using 1 mg/mL lysozyme (from chicken egg white; Sigma Aldrich, St Louis, USA). Conjugation to antibodies was then demonstrated using 1 mg/mL anti-BDNF antibody (Biosensis, Thebarton, Australia).

General procedure for CuAAC reactions. The CuAAC procedure was based on previously published methods.⁴¹⁻⁴³ Briefly, purified and undiluted protein conjugates labelled with NHS photolinker **4** (20 μ L) were diluted with PBS (90 μ L), followed by addition of aqueous solutions of 2.5 mM azide-containing crude peptide **A** (ACTK[N₃]FYH-NH₂; 20 μ L) or Cy3-azide (Sigma Aldrich), 100 mM tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (BTAA; 10 μ L; Click Chemistry Tools, Scottsdale, USA), 20 mM copper sulfate (10 μ L) and 300 mM sodium ascorbate (10 μ L). The reaction was incubated at room temperature for 30 min, then purified and concentrated into PBS (100 μ L) using a 10 kDa MWCO Amicon® Ultra-0.5 Centrifugal Filter Device (Merck Millipore, Burlington, USA).

SDS-PAGE analysis of Cy3-labelled anti-BDNF antibody. Protein SDS-PAGE was performed using a Mini-PROTEAN® TGX™ 10-well precast gel and Mini-PROTEAN® Tetra vertical electrophoresis cell (both Bio-Rad, Hercules, US). Bands were visualized using a ChemiDoc XRS+ imaging system (Bio-Rad) before and after staining with Coomassie Brilliant Blue stain (Thermo Fisher Scientific) with the Cy3 and Stain Free Gel application settings, respectively.

Matrix-assisted laser desorption ionization profiling of protein conjugates. Matrix-assisted laser desorption/ionization (MALDI) spectra were acquired using an ultraFlextreme MALDI-TOF/TOF mass spectrometer in reflector positive mode, controlled by FlexControl software v3.4 (Bruker Daltonics). Protein samples were prepared for MALDI analysis by mixing with an equal volume of either 10 mg/mL α -cyano-4-hydroxycinnamic acid (HCCA) or sinapinic acid (SA) matrix in 70 % acetonitrile plus 1 % trifluoroacetic acid (TFA). This solution was deposited on a Bruker MTP AnchorChip™ or ground steel 384 target plate for analysis of intact proteins or photocleavage products, respectively. Spectra were converted to mzXML file format using FlexAnalysis v3.4 (Bruker Daltonics), then processed, including smoothing, cropping and baseline subtraction, using mMass v5.5.0 (<http://www.mmass.org/>).⁴⁴

Acknowledgement

The authors would like to thank Robert Rush and Kym McNicholas at Biosensis Pty Ltd., Adelaide for their expert advice and generous donation of the anti-BDNF antibody. KGS received financial support from a University of Adelaide Faculty of Sciences Divisional Scholarship.

References

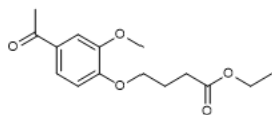
1. Crutchfield, C. A.; Thomas, S. N.; Sokoll, L. J.; Chan, D. W., Advances in mass spectrometry-based clinical biomarker discovery. *Clin. Proteom.* 2016, *13* (1), 1.
2. Thompson, A.; Schäfer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C., Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* 2003, *75* (8), 1895-1904.
3. Xu, S.; Ma, W.; Bai, Y.; Liu, H., Ultrasensitive ambient mass spectrometry immunoassays: multiplexed detection of proteins in serum and on cell surfaces. *JACS* 2019, *141* (1), 72-75.
4. Isola, N. R.; Allman, S. L.; Golovlev, V. V.; Chen, C. H., MALDI-TOF Mass spectrometric method for detection of hybridized DNA oligomers. *Anal. Chem.* 2001, *73* (9), 2126-2131.
5. Lorey, M.; Adler, B.; Yan, H.; Soliymani, R.; Ekström, S.; Yli-Kauhaluoma, J.; Laurell, T.; Baumann, M., Mass-tag enhanced immuno-laser desorption/ionization mass spectrometry for sensitive detection of intact protein antigens. *Anal. Chem.* 2015, *87* (10), 5255-5262.
6. Lemaire, R.; Stauber, J.; Wisztorski, M.; Van Camp, C.; Desmons, A.; Deschamps, M.; Proess, G.; Rudlof, I.; Woods, A. S.; Day, R.; Salzet, M.; Fournier, I., Tag-Mass: specific molecular imaging of transcriptome and proteome by mass spectrometry based on photocleavable tag. *J. Prot. Res.* 2007, *6* (6), 2057-2067.
7. Sauer, S., Typing of single nucleotide polymorphisms by MALDI mass spectrometry: Principles and diagnostic applications. *Clin. Chim. Acta* 2006, *363* (1), 95-105.
8. Caprioli, R. M.; Farmer, T. B.; Gile, J., Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* 1997, *69* (23), 4751-4760.
9. Chang, Q.; Ornatsky, O. I.; Siddiqui, I.; Loboda, A.; Baranov, V. I.; Hedley, D. W., Imaging mass cytometry. *Cytom. Part A* 2017, *91* (2), 160-169.
10. Gagnon, H.; Franck, J.; Wisztorski, M.; Day, R.; Fournier, I.; Salzet, M., TARGETED MASS spectrometry Imaging: Specific Targeting Mass Spectrometry imaging technologies from history to perspective. *Prog. Histochem. Cytochem.* 2012, *47* (3), 133-174.
11. Luque-Garcia, J. L.; Neubert, T. A., Sample preparation for serum/plasma profiling and biomarker identification by mass spectrometry. *J. Chromatogr. A* 2007, *1153* (1), 259-276.
12. Zhang, Z.; Pan, H.; Chen, X., Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom. Rev.* 2009, *28* (1), 147-176.
13. Sanz-Medel, A.; Montes-Bayón, M.; Luisa Fernández Sánchez, M., Trace element speciation by ICP-MS in large biomolecules and its potential for proteomics. *Anal. Bioanal. Chem.* 2003, *377* (2), 236-247.
14. Giesen, C.; Waentig, L.; Panne, U.; Jakubowski, N., History of inductively coupled plasma mass spectrometry-based immunoassays. *Spectrochim. Acta B* 2012, *76*, 27-39.
15. Zhang, T.; Warden, A. R.; Li, Y.; Ding, X., Progress and applications of mass cytometry in sketching immune landscapes. *Clin. Transl. Med.* 2020, *10* (6), e206.

16. Holle, A.; Haase, A.; Kayser, M.; Höhndorf, J., Optimizing UV laser focus profiles for improved MALDI performance. *J. Mass Spectrom.* 2006, *41* (6), 705-716.
17. Edler, M.; Mayrbrugger, S.; Fian, A.; Trimmel, G.; Radl, S.; Kern, W.; Griesser, T., Wavelength selective refractive index modulation in a ROMP derived polymer bearing phenyl- and ortho-nitrobenzyl ester groups. *J. Mater. Chem.* 2013, *1* (25), 3931-3938.
18. Hansen, M. J.; Velema, W. A.; Lerch, M. M.; Szymanski, W.; Feringa, B. L., Wavelength-selective cleavage of photoprotecting groups: strategies and applications in dynamic systems. *Chem. Soc. Rev.* 2015, *44* (11), 3358-3377.
19. Šolomek, T.; Mercier, S.; Bally, T.; Bochet, C. G., Photolysis of ortho-nitrobenzylic derivatives: the importance of the leaving group. *Photochem. Photobiol. Sci.* 2012, *11* (3), 548-555.
20. Stauber, J.; El Ayed, M.; Wisztorski, M.; Salzet, M.; Fournier, I., Specific MALDI-MSI: tag-Mass. In *Mass Spectrometry Imaging*, Springer: 2010; pp 339-361.
21. Rostovtsev, V. V., Green, L. G., Fokin, V. V., & Sharpless, K.B., A stepwise Huisgen cycloaddition process: copper (I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem.* 2002 *114* (14), 2708-2711.
22. Presolski, S. I.; Hong, V. P.; Finn, M. G., Copper-catalyzed azide-alkyne click chemistry for bioconjugation. *Curr. Prot. Chem. Biol.* 2011, *3* (4), 153-162.
23. Kloxin, A. M.; Tibbitt, M. W.; Anseth, K. S., Synthesis of photodegradable hydrogels as dynamically tunable cell culture platforms. *Nat. Protoc.* 2010, *5* (12), 1867-1887.
24. Polacchini, A.; Metelli, G.; Francavilla, R.; Baj, G.; Florean, M.; Mascaretti, L. G.; Tongiorgi, E., A method for reproducible measurements of serum BDNF: comparison of the performance of six commercial assays. *Sci. Rep.* 2015, *5* (1), 17989.
25. Cao, S.-L.; Hao, J.-H.; Ou-Yang, W.-D.; Chen, Z.-S.; Lv, Y.-S.; Zeng, L.-W., Introductory Chapter: Color Detection. In *Color Detection*, IntechOpen: 2020.
26. Kameyama, A.; Kaneda, Y.; Yamanaka, H.; Yoshimine, H.; Narimatsu, H.; Shinohara, Y., Detection of oligosaccharides labeled with cyanine dyes using matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 2004, *76* (15), 4537-4542.
27. Gautier, V.; Boumeester, A. J.; Lössl, P.; Heck, A. J., Lysine conjugation properties in human IgGs studied by integrating high-resolution native mass spectrometry and bottom-up proteomics. *Proteomics* 2015, *15* (16), 2756-2765.
28. Berg, E. A.; Fishman, J. B., Labeling antibodies. *Cold Spring Harb. Protoc.* 2020, *2020* (7), pdb.top099242.
29. Das, A.; Theato, P., Activated Ester Containing Polymers: Opportunities and Challenges for the Design of Functional Macromolecules. *Chem. Rev.* 2016, *116* (3), 1434-1495.
30. Fu, Q.; Schoenhoff, F. S.; Savage, W. J.; Zhang, P.; Van Eyk, J. E., Multiplex assays for biomarker research and clinical application: translational science coming of age. *Proteomics Clin. Appl.* 2010, *4* (3), 271-284.
31. Tighe, P. J.; Ryder, R. R.; Todd, I.; Fairclough, L. C., ELISA in the multiplex era: potentials and pitfalls. *Proteomics Clin. Appl.* 2015, *9* (3-4), 406-422.
32. Rauh, M., LC-MS/MS for protein and peptide quantification in clinical chemistry. *J. Chromatogr. B* 2012, *883-884*, 59-67.
33. Parker, C. E.; Pearson, T. W.; Anderson, N. L.; Borchers, C. H., Mass-spectrometry-based clinical proteomics—a review and prospective. *Analyst* 2010, *135* (8), 1830-1838.

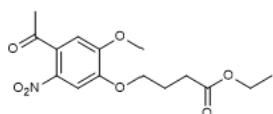
34. Kellie, J. F.; Tran, J. C.; Lee, J. E.; Ahlf, D. R.; Thomas, H. M.; Ntai, I.; Catherman, A. D.; Durbin, K. R.; Zamdborg, L.; Vellaichamy, A., The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput. *Mol. BioSyst.* 2010, 6 (9), 1532-1539.
35. El-Aneed, A.; Cohen, A.; Banoub, J., Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Appl. Spectrosc. Rev.* 2009, 44 (3), 210-230.
36. Gevaert, K.; Vandekerckhove, J., Protein identification methods in proteomics. *Electrophoresis* 2000, 21 (6), 1145-1154.
37. Duncan, M. W.; Nedelkov, D.; Walsh, R.; Hattan, S. J., Applications of MALDI Mass Spectrometry in Clinical Chemistry. *Clin. Chem.* 2016, 62 (1), 134-143.
38. Piga, I.; Heijs, B.; Nicolardi, S.; Giusti, L.; Marselli, L.; Marchetti, P.; Mazzoni, M. R.; Lucacchini, A.; McDonnell, L. A., Ultra-high resolution MALDI-FTICR-MSI analysis of intact proteins in mouse and human pancreas tissue. *Int. J. Mass Spectrom.* 2019, 437, 10-16.
39. Nedelkov, D.; Niederkofler, E. E.; Oran, P. E.; Peterman, S.; Nelson, R. W., Top-down mass spectrometric immunoassay for human insulin and its therapeutic analogs. *J. Proteom.* 2018, 175, 27-33.
40. Trenchevska, O.; Nelson, R. W.; Nedelkov, D., Mass spectrometric immunoassays in characterization of clinically significant proteoforms. *Proteomes* 2016, 4 (1), 13.
41. Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V., Polytriazoles as copper (I)-stabilizing ligands in catalysis. *Org. Lett.* 2004, 6 (17), 2853-2855.
42. Presolski, S. I.; Hong, V. P.; Finn, M., Copper-catalyzed azide-alkyne click chemistry for bioconjugation. *Curr. Protoc. Chem. Biol.* 2011, 3 (4), 153-162.
43. Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M., Labeling live cells by copper-catalyzed alkyne-azide click chemistry. *Bioconjug. Chem.* 2010, 21 (10), 1912-1916.
44. Strohm, M.; Hassman, M.; Kořata, B.; Kodíček, M., mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun. Mass Spectrom.* 2008, 22 (6), 905-908.

Supporting Information

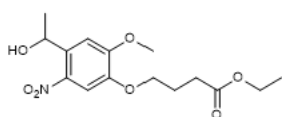
Synthesis of hydroxyethyl photolinker



Ethyl 4-(4-acetyl-2-methoxyphenoxy)butanoate. Acetovanillone (**1**; 9.97 g, 60 mmol) and ethyl-4-bromobutyrate (10.8 mL, 75 mmol) were dissolved in dry dimethylformamide (100 mL) while stirring and the solution was purged with N₂ for 10 min. Potassium carbonate (12.5 g, 90 mmol) was added, resulting in a light yellow suspension, which was stirred overnight at room temperature. The reaction mixture was added slowly to water (800 mL) and the precipitate collected via vacuum filtration. The precipitate was washed with water (3 x 20 mL) and dried under high vacuum to obtain a white powder (15.6 g, 93 %). ¹H NMR (500 MHz, CDCl₃): δ_H 7.56 (dd, *J* = 8.3, 2.1 Hz, 1H, aromatic), 7.54 (d, *J* = 2.0 Hz, 1H, aromatic), 6.91 (d, *J* = 8.3 Hz, 1H, aromatic), 4.18-4.13 (m, 4H, OCH₂CH₂ and COOCH₂CH₃), 3.92 (s, 3H, OCH₃), 2.57 (s, 3H, COCH₃), 2.55 (t, *J* = 7.2 Hz, 2H, CH₂COOCH₂), 2.23-2.17 (m, 2H, OCH₂CH₂CH₂), 1.27 (t, *J* = 7.1 Hz, 2H, COOCH₂CH₃).

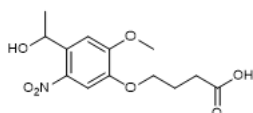


Ethyl 4-(4-acetyl-2-methoxy-5-nitrophenoxy)butanoate. Ethyl 4-(4-acetyl-2-methoxyphenoxy)butanoate (14.8 g, 53 mmol) was added in small portions to a stirred solution of 70 % nitric acid (150 mL) at 0 °C over 30 min. The reaction was stirred for 2 h at 0 °C, then quenched by gradually diluting the reaction mixture with water (800 mL) while stirring. The resulting yellow precipitate was collected via vacuum filtration and washed with water (800 mL). The crude product was dried under high vacuum, then recrystallised from hot ethanol to obtain a yellow crystalline solid (9.97 g, 56 %). ¹H NMR (500 MHz, CDCl₃): δ_H 7.62 (s, 1H, aromatic), 6.76 (s, 1H, aromatic), 4.21-4.14 (m, 4H, OCH₂CH₂ and CH₂COOCH₂CH₃), 3.97 (s, 3H, OCH₃), 2.56 (t, *J* = 7.2 Hz, 2H, CH₂COOCH₂), 2.51 (s, 3H, COCH₃), 2.21 (m, *J* = 6.8 Hz, 2H, OCH₂CH₂CH₂), 1.28 (t, *J* = 7.1 Hz, 3H, COOCH₂CH₃).



Ethyl 4-[4-(1-hydroxyethyl)-2-methoxy-5-nitrophenoxy]butanoate. Ethyl 4-(4-acetyl-2-methoxy-5-nitrophenoxy)butanoate (9.95 g, 31 mmol) was suspended in 100 % ethanol (150 mL) and cooled to 0 °C. Sodium borohydride (591 mg, 15.6 mmol) was added in small portions over 10 min while stirring. The reaction was

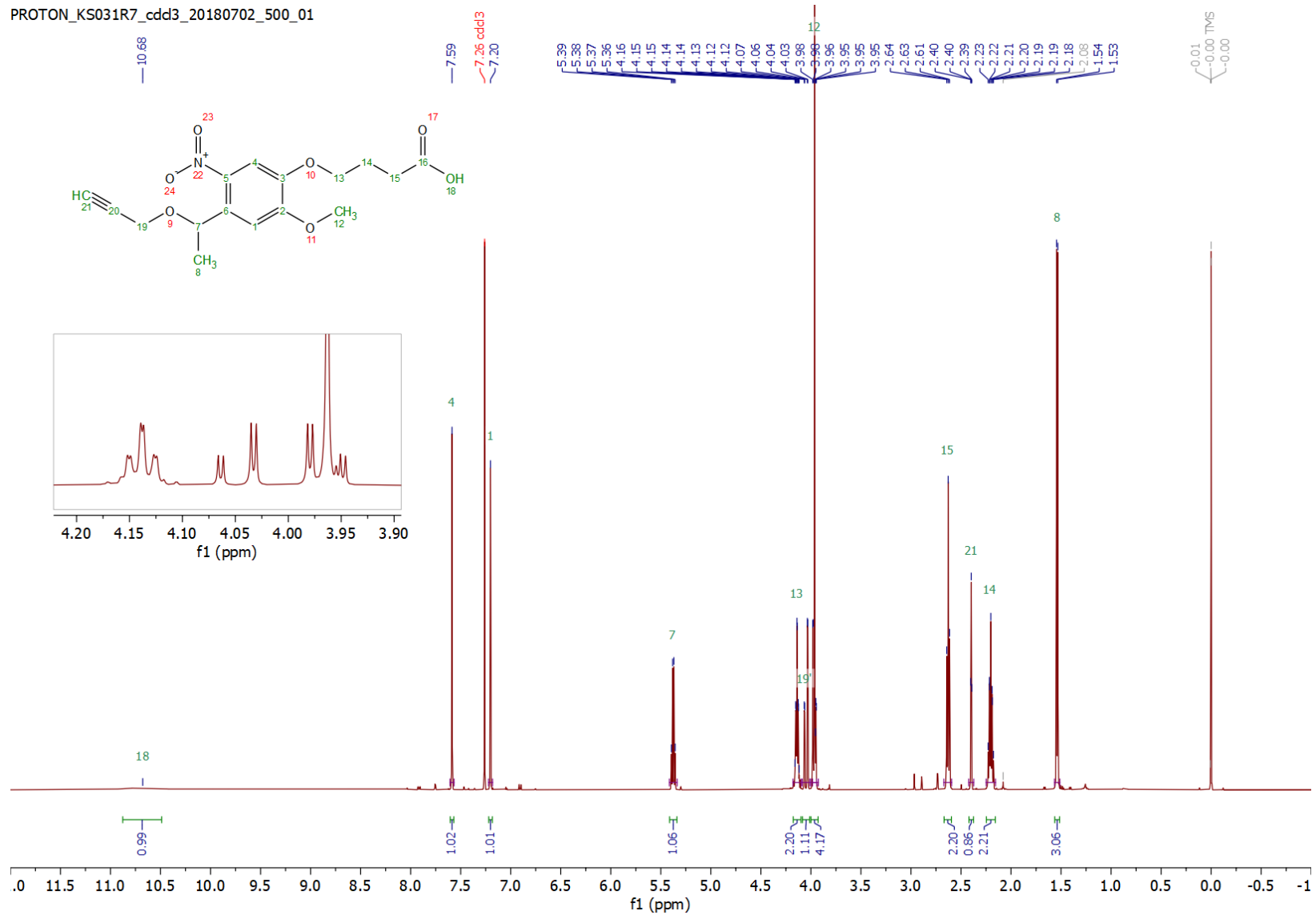
allowed to warm to room temperature while stirring overnight, then quenched with water (100 mL) and the ethanol evaporated under reduced pressure. The resulting precipitate was collected via vacuum filtration, washed with water (150 mL), and dried under high vacuum to afford an off-white crystalline solid (8.87 g, 87 %). ^1H NMR (500 MHz, CD_3OD): δ_{H} 7.57 (s, 1H, aromatic), 7.29 (s, 1H, aromatic), 5.60-5.53 (m, 1H, $\text{CH}(\text{OH})\text{CH}_3$), 4.20-4.08 (m, 4H, OCH_2CH_2 and $\text{CH}_2\text{COOCH}_2\text{CH}_3$), 3.98 (s, 3H, OCH_3), 2.54 (t, $J = 7.2$ Hz, 2H, $\text{CH}_2\text{COOCH}_2$), 2.26 (d, $J = 3.7$ Hz, 1H, OH), 2.18 (m, $J = 6.8$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{COOCH}_2$), 1.56 (s, 3H, $\text{CH}(\text{OH})\text{CH}_3$), 1.27 (t, $J = 7.1$ Hz, 3H, $\text{COOCH}_2\text{CH}_3$).



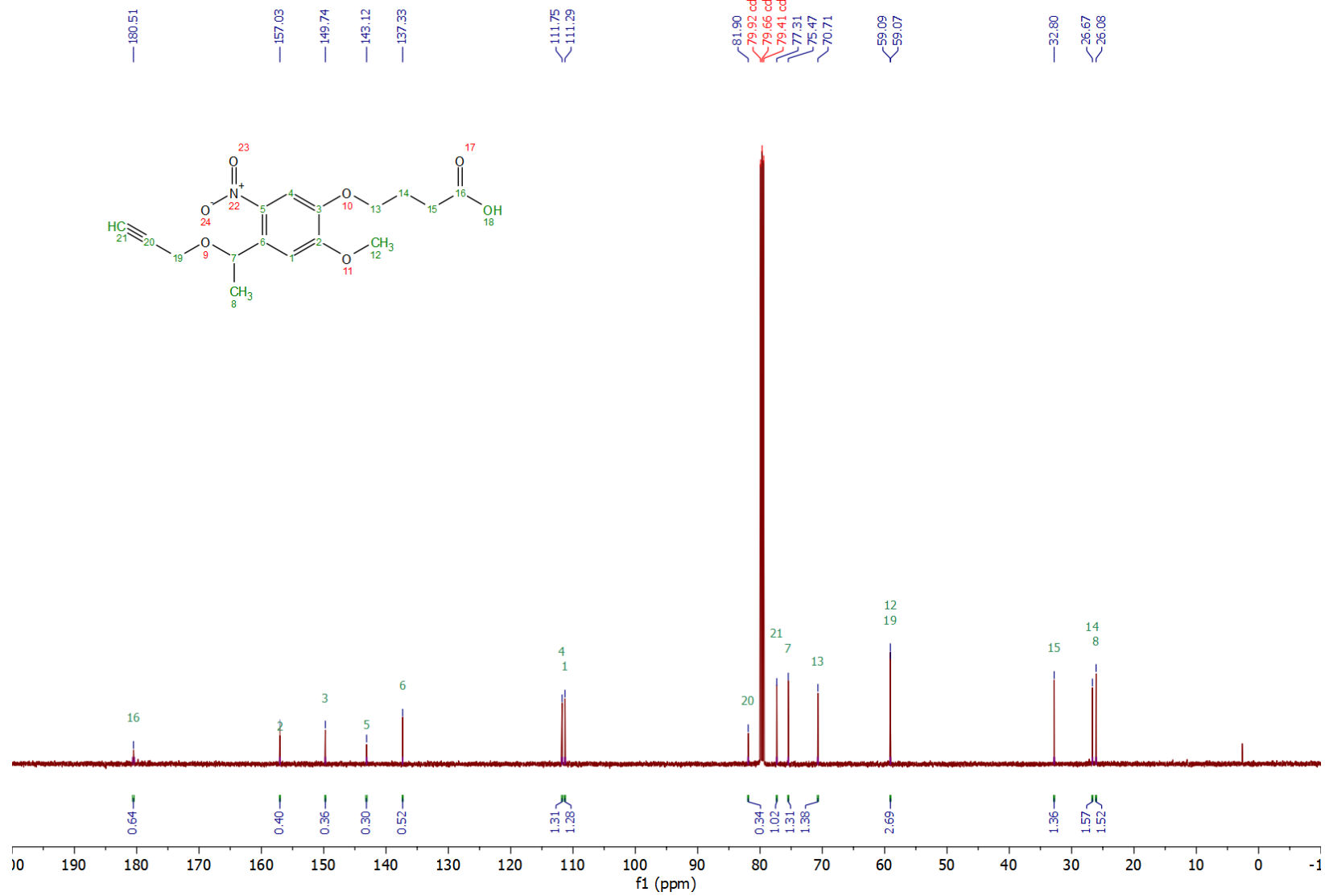
4-[4-(1-Hydroxyethyl)-2-methoxy-5-nitrophenoxy]butanoic acid (2). Ethyl 4-[4-(1-hydroxyethyl)-2-methoxy-5-nitrophenoxy]butanoate (8.87 g, 27 mmol) was dissolved in tetrahydrofuran (50 mL). 1 M lithium hydroxide (100 mL) was added and the reaction stirred at room temperature for 2 h. The tetrahydrofuran was removed under reduced pressure and the remaining aqueous solution adjusted to pH 3 via addition of 1 M hydrochloric acid. The resulting precipitate was collected via vacuum filtration and recrystallised from hot ethanol to obtain the product as a yellow powder (6.59 g, 82 %). ^1H NMR (500 MHz, CD_3OD): δ_{H} 7.59 (s, 1H, aromatic), 7.40 (s, 1H, aromatic), 5.46 (q, $J = 6.3$ Hz, 1H, $\text{CH}(\text{OH})\text{CH}_3$), 4.10 (t, $J = 6.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{COOH}$), 3.97 (s, 3H, OCH_3), 2.52 (t, $J = 7.3$ Hz, 2H, CH_2COOH), 2.10 (m, 2H, $\text{CH}_2\text{CH}_2\text{COOH}$), 1.48 (d, $J = 6.3$ Hz, 3H, $\text{CH}(\text{OH})\text{CH}_3$). ^1H NMR spectra was consistent with previous reports.¹

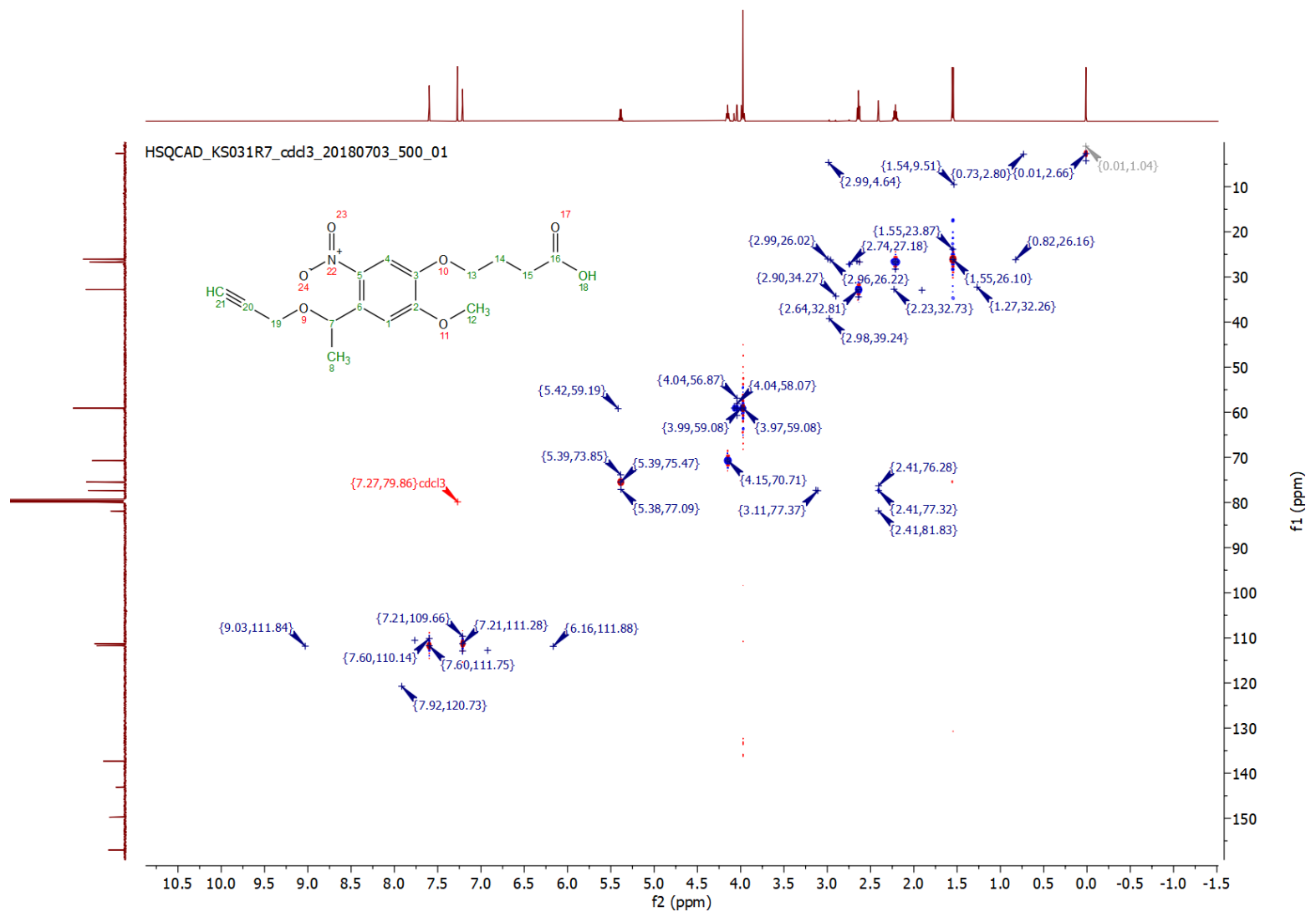
NMR spectra

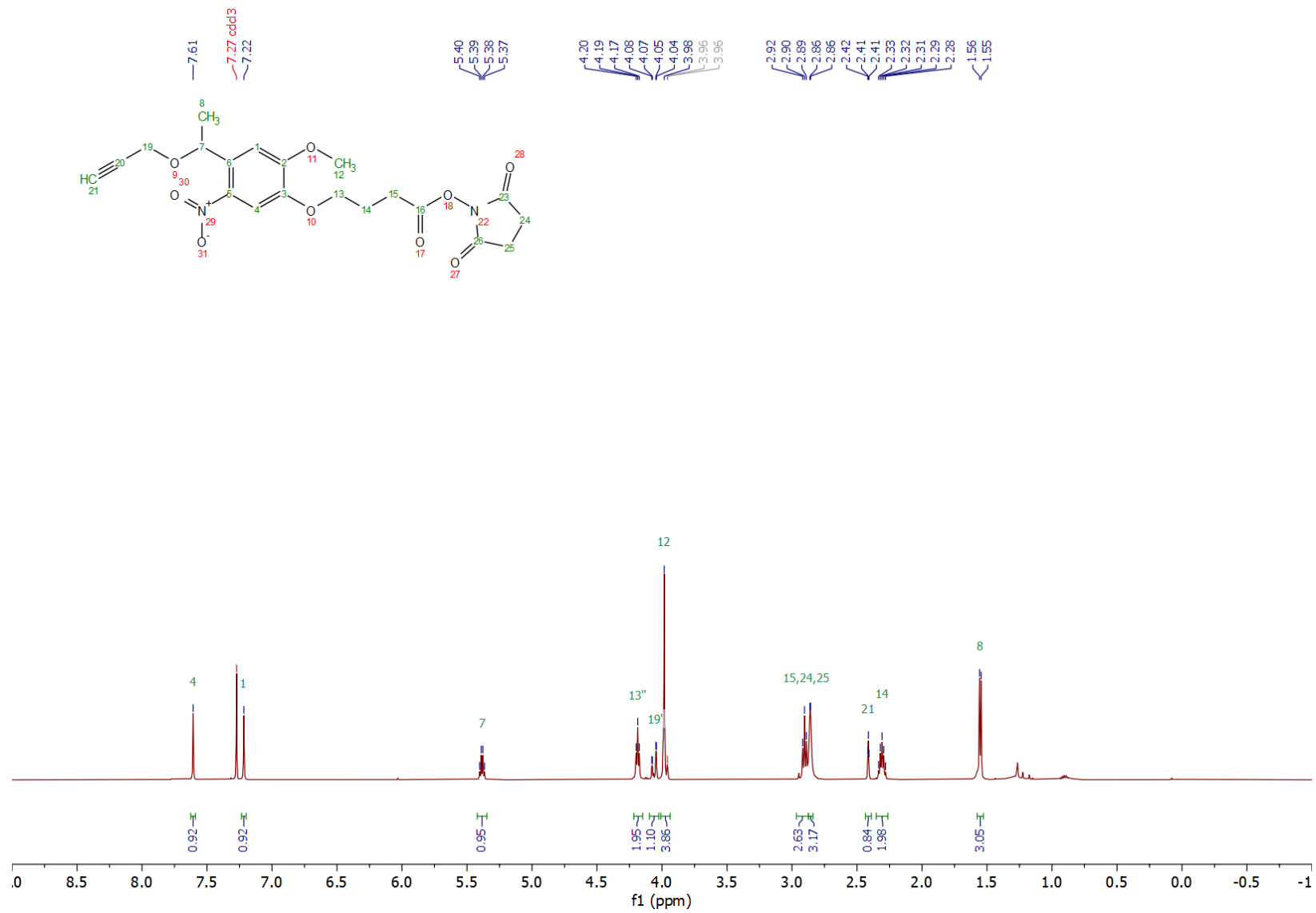
PROTON_KS031R7_cdd3_20180702_500_01

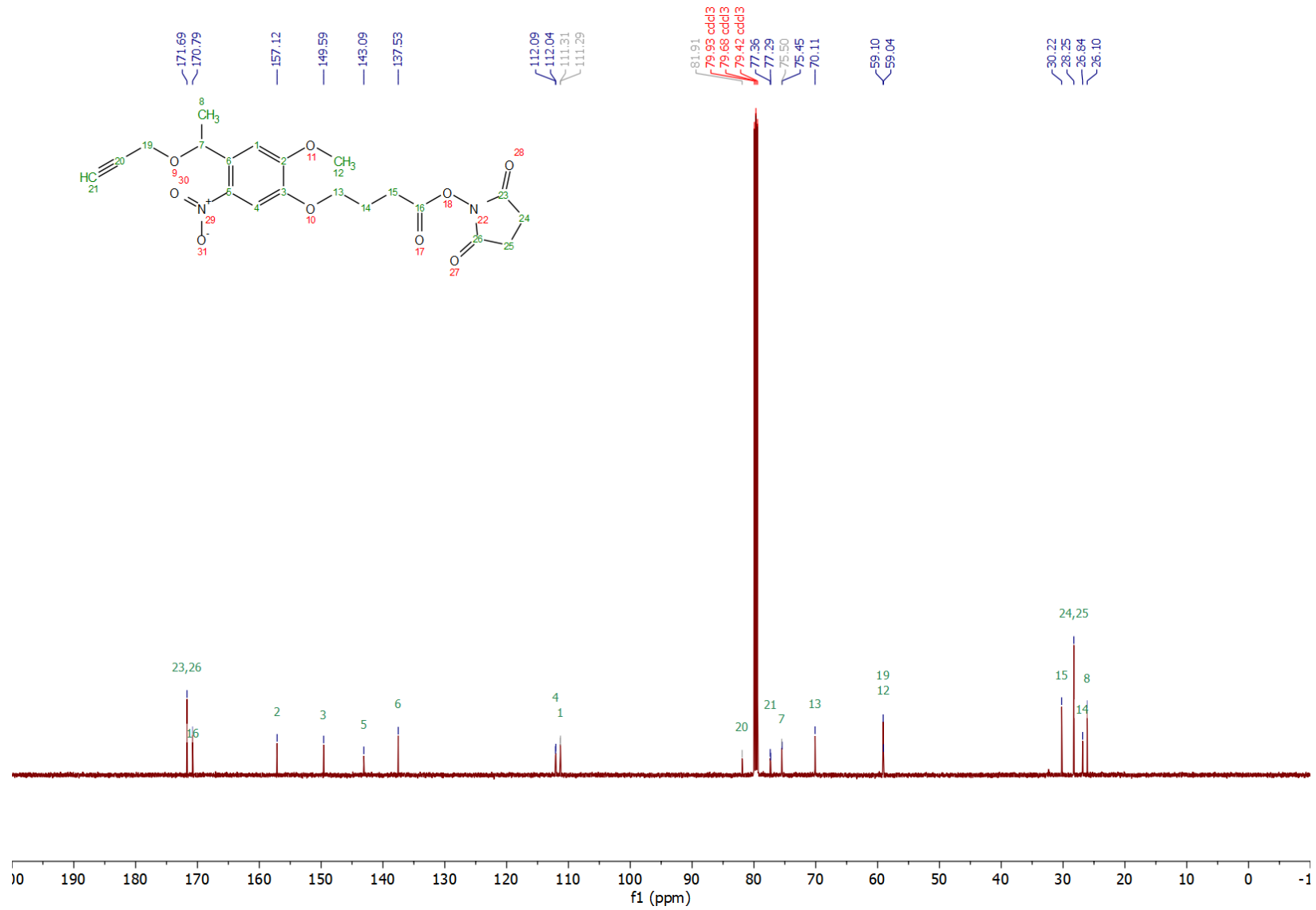


CARBON_KS031R7_cdd3_20180703_500_01

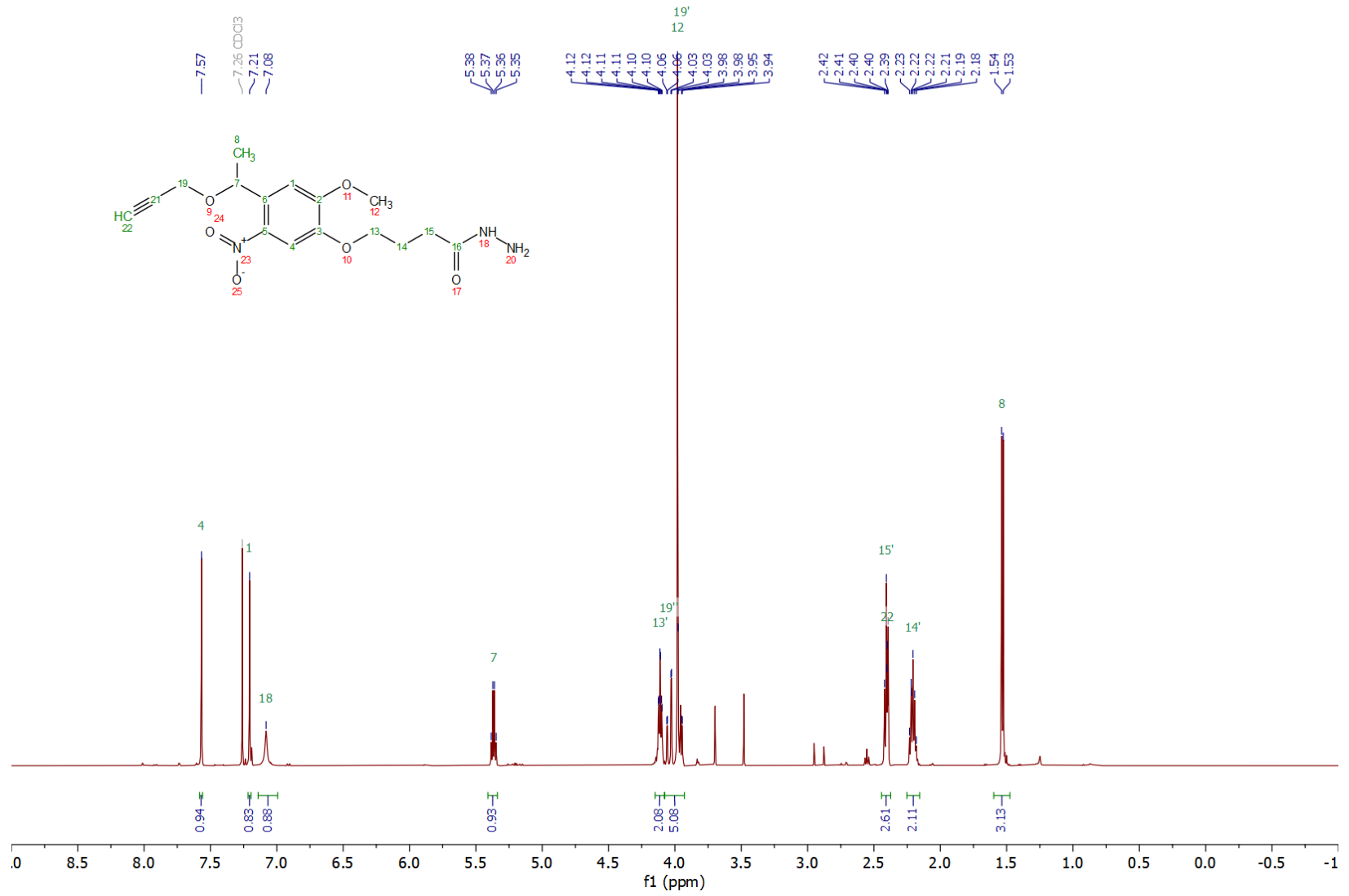




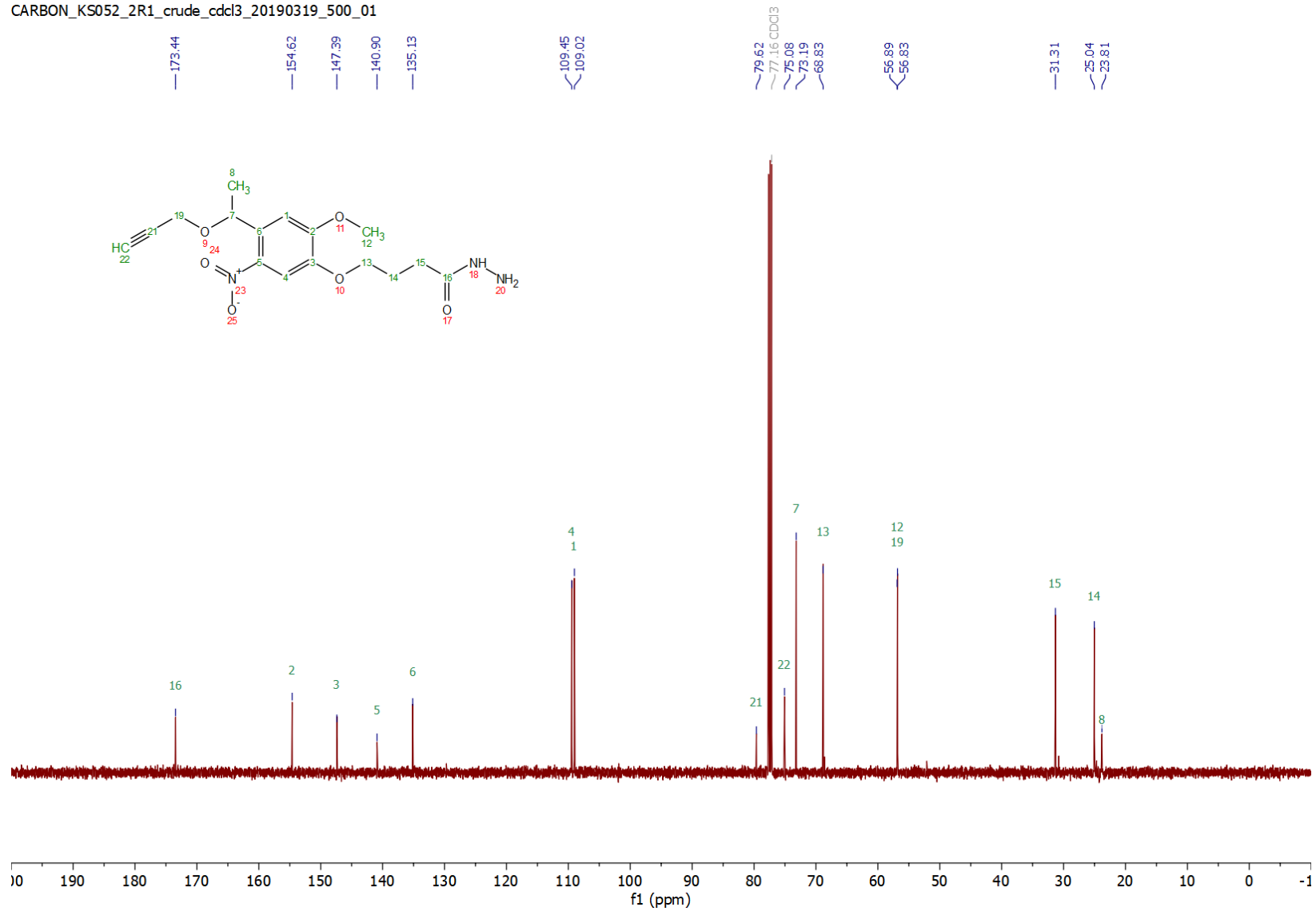




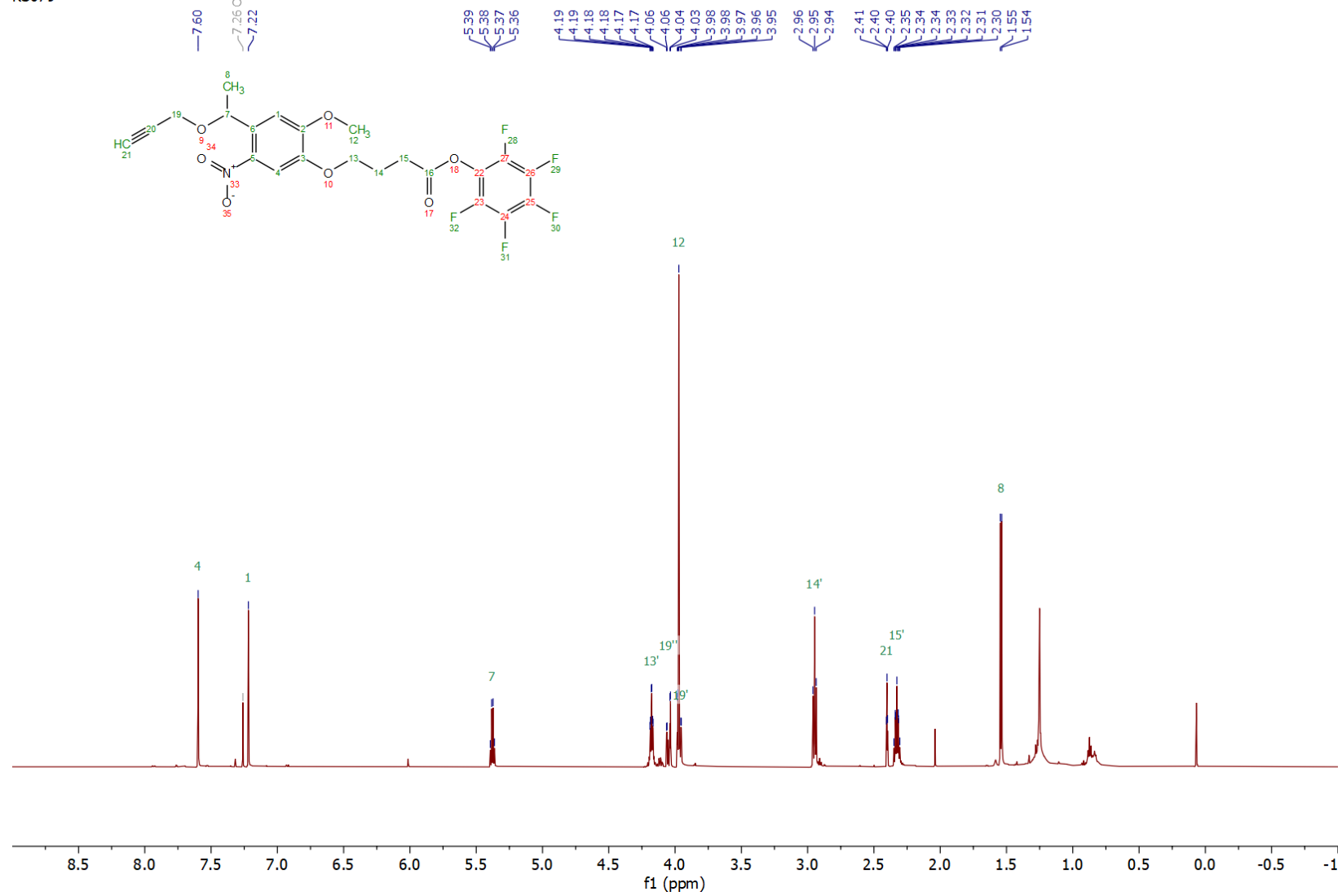
PROTON_KS052_2R1_crude_add3_20190319_500_01



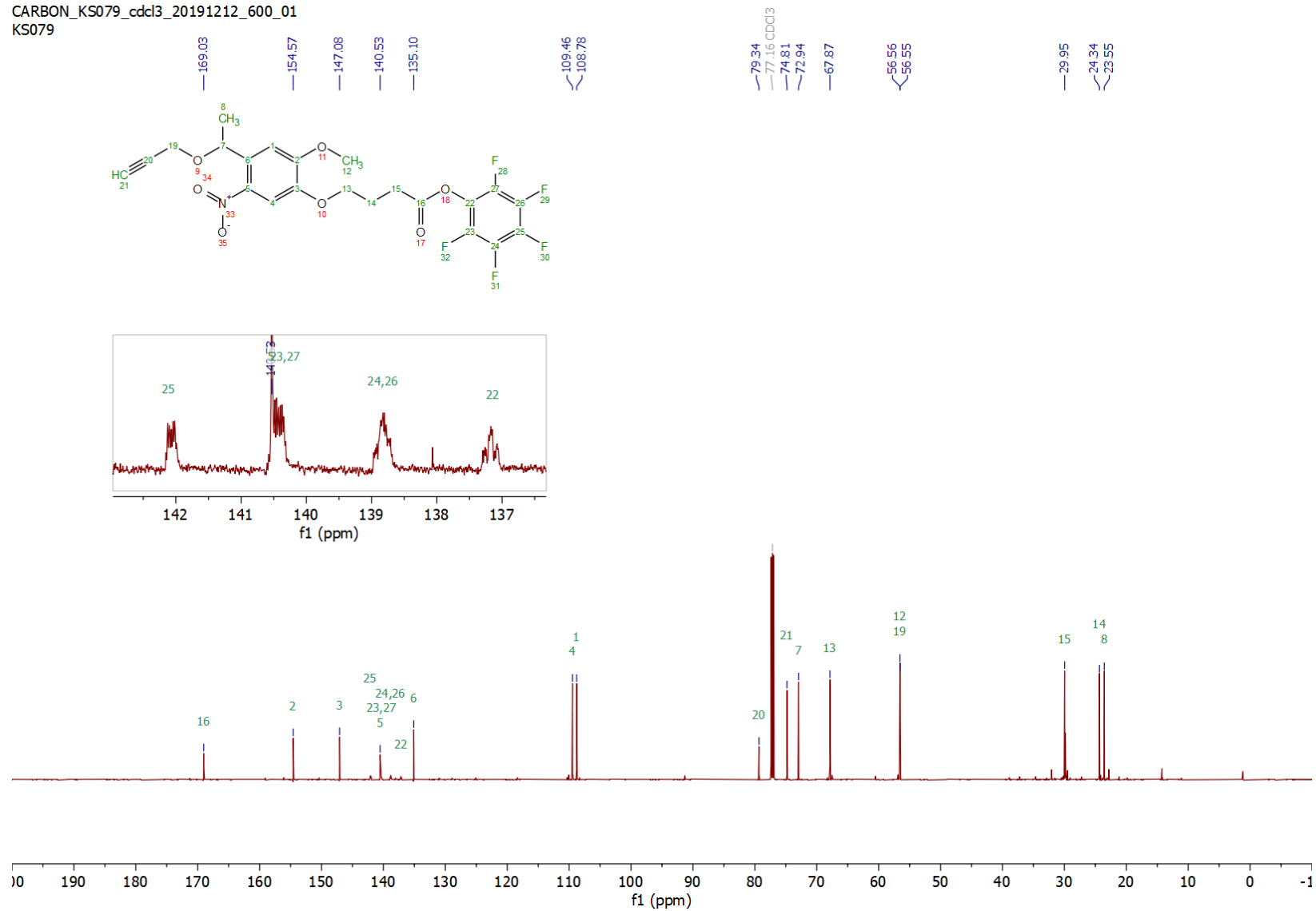
CARBON_KS052_2R1_crude_cdcl3_20190319_500_01



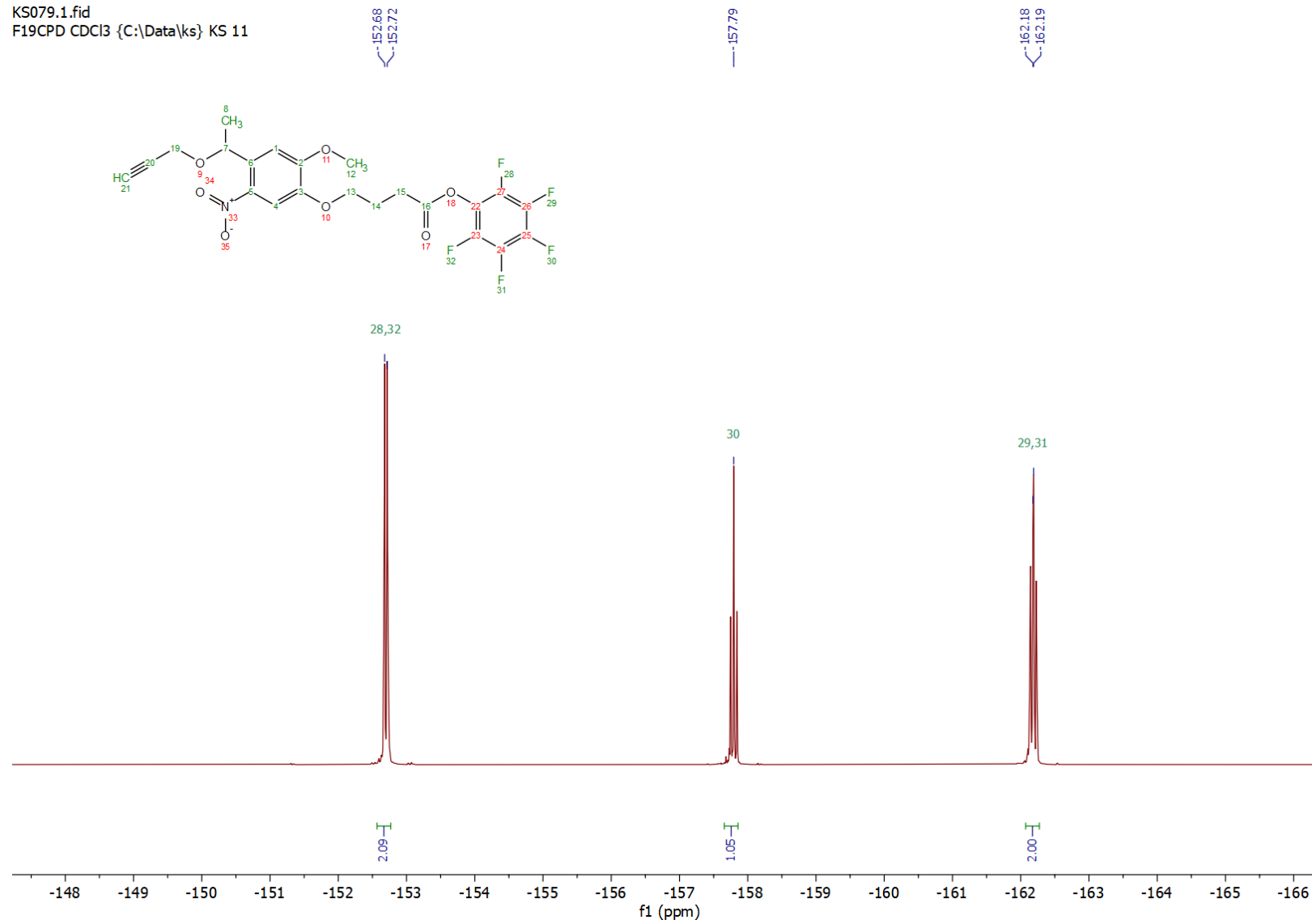
PROTON_KS079_cdcl3_20191212_600_01
KS079



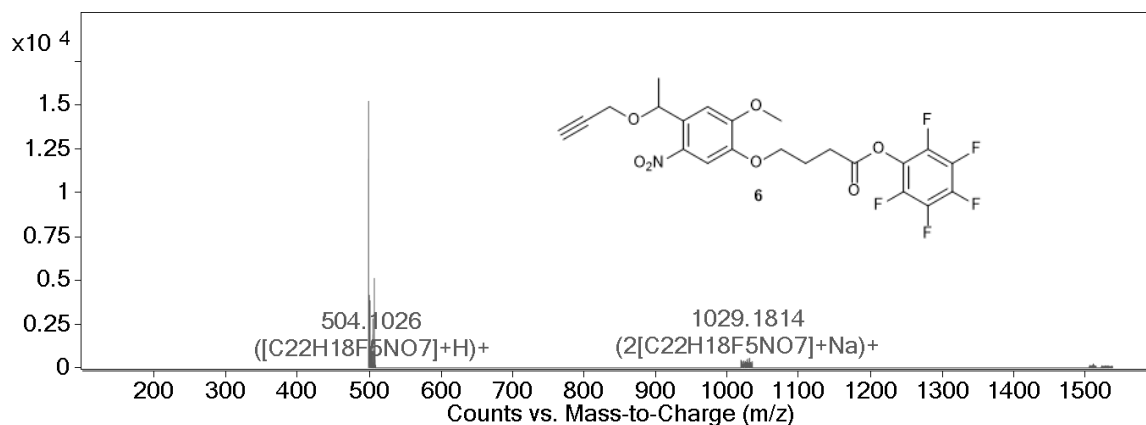
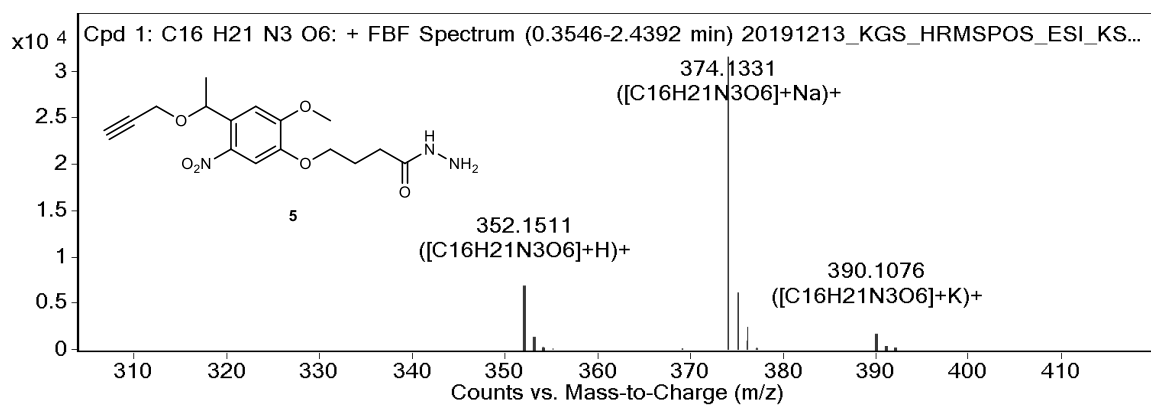
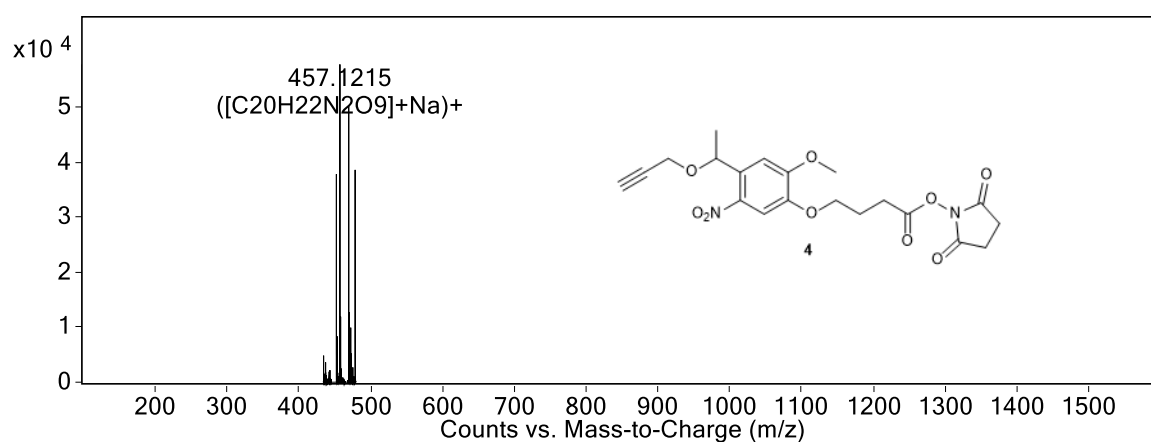
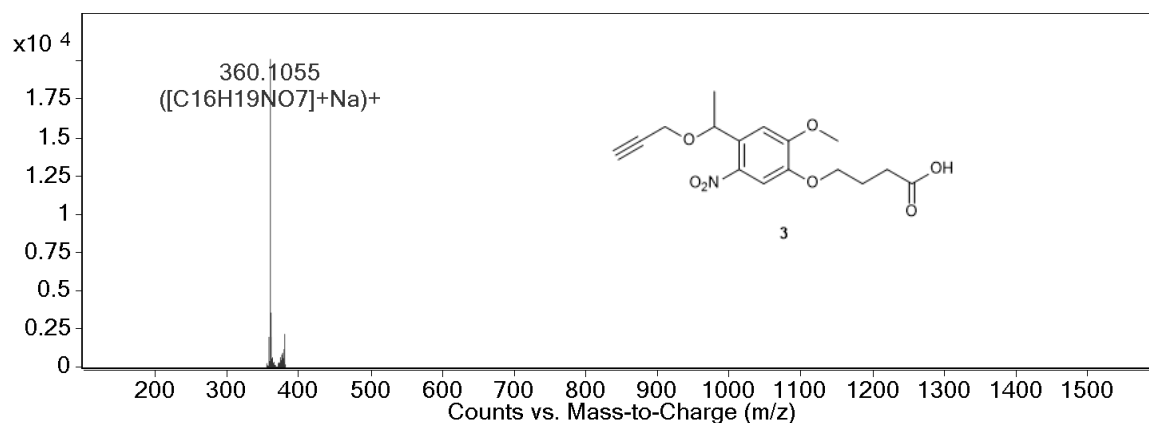
CARBON_KS079_cdc13_20191212_600_01
KS079



KS079.1.fid
F19CPD CDCl3 {C:\Data\ks} KS 11



HRMS spectra



References

1. Klinger, D. and Landfester K., Enzymatic- and light-degradable hybrid nanogels: Crosslinking of polyacrylamide with acrylate-functionalized dextrans containing photocleavable linkers. *J. Polym. Sci. A Polym. Chem.* **2012**, 50 (6), 1062-1075.

Retro Diels-Alder fragmentation of fulvene-maleimide bioconjugates for mass spectrometric detection of biomolecules

Foreword

Following successful demonstration of the application of copper-catalysed azide-alkyne cycloaddition (CuAAC) for synthesising mass tags, we decided to explore additional bio-orthogonal reactions, which could enable conjugation to peptides without the need for a copper catalyst or additional reagents, such as ligands and reducing agents. One possibility, the Diels-Alder (DA) cycloaddition reaction between maleimides and dienes, such as pentafulvene, has been widely reported in the context of materials synthesis, but has limited examples of its use for generating mass tag-labelled biomolecules. One explanation for the underutilisation of these reactions could be the labile nature of these DA adducts at high temperatures or in the gas phase; however, we exploit this property for the synthesis bioconjugates that display efficient fragmentation following both matrix-assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI) mass spectrometry (MS). This new mass tag design therefore expands on our previous mass tag designs by being applicable to both ESI analysis of liquid samples and MALDI analysis of immobilised analytes or tissue imaging without the inclusion of a bulky photocleavable moiety.

Statement of Authorship

Title of Paper	Retro Diels-Alder fragmentation of fulvene-maleimide bioconjugates for mass spectrometry detection of biomolecules
Publication Status	<input type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input checked="" type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Article formatted for intended publication in <i>Analytical Chemistry</i> peer-reviewed journal.

Principal Author

Name of Principal Author (Candidate)	Katherine G Stevens		
Contribution to the Paper	Collected majority of experimental data, author of main body text		
Overall percentage (%)	70		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	1/12/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Kirsten J Platts*		
Contribution to the Paper	Synthesis of compound 4, optimisation of amide coupling conditions.		
Signature		Date	4/12/20

Name of Co-Author	NEIL M O BRIEN-SIMPSON		
Contribution to the Paper	Peptide design, synthesis and characterisation		
Signature		Date	1/12/20

Please cut and paste additional co-author panels here as required.

*Unavailable for signed statement so TLP signed in lieu of co-author.

Statement of Authorship

Name of Co-Author	Wenyi Li		
Contribution to the Paper	Peptide design, synthesis and characterisation		
Signature		Date	1/12/20

Name of Co-Author	Paul J Trim		
Contribution to the Paper	Acquisition of timsTOF fleX data, editing of manuscript		
Signature		Date	1/12/20

Name of Co-Author	Anton Blencowe		
Contribution to the Paper	Principle supervisor to KJP, reporter peptide design, editing of manuscript		
Signature		Date	1/12/20

Name of Co-Author	Tara L Pukala		
Contribution to the Paper	Provided advice related to experimental design and project aims, editing of manuscript		
Signature		Date	4/12/20

Please cut and paste additional co-author pan

Retro Diels-Alder fragmentation of fulvene-maleimide bioconjugates for mass spectrometric detection of biomolecules

Running title: *Retro Diels-Alder of fulvene-maleimide adducts*

Katherine Stevens¹, Kirsten Platts², Neil O'Brien-Simpson³, Wenyi Li³, Anton Blencowe², Paul J. Trim⁴ and Tara Pukala^{1*}

¹ Department of Chemistry, Faculty of Sciences, The University of Adelaide, Adelaide, South Australia 5000, Australia

² Applied Chemistry and Translational Biomaterials Group, Clinical and Health Sciences, The University of South Australia, Adelaide, South Australia 5000, Australia

³ Centre for Oral Health Research, The Melbourne Dental School and the Bio21 Institute, The University of Melbourne, 720 Swanston Street, Carlton, Melbourne, Victoria, Australia. 3010.

⁴ Proteomics, Metabolomics and MS Imaging, South Australian Health and Medical Research Institute, Adelaide, South Australia 5000, Australia.

Address reprint requests to:

Tara Pukala

School of Physical Sciences

The University of Adelaide, Australia 5005

Ph: +61 8 8313 5497

E-mail: tara.pukala@adelaide.edu.au

Word count:

Figures: 5

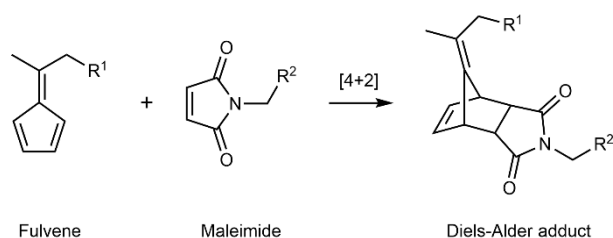
Schemes: 4

Abstract

Diels-Alder chemistry is a well-explored avenue for the synthesis of bioactive materials; however, its potential applications have recently expanded following the development of reactions that can be performed in buffered aqueous environments at low temperatures, including fulvene-maleimide [4+2] cycloadditions. In this study, we synthesized two novel amine-reactive fulvene linkers to demonstrate the application of this chemistry for generating mass spectrometry-cleavable labels ('mass tags'), which can be used for the labelling and detection of proteins. Successful conjugation of these linkers to maleimide-labelled peptides was observed at low temperatures in phosphate-buffered saline, allowing the non-destructive modification of proteins. The labile nature of fulvene-maleimide adducts in the gas phase also makes them suitable for both matrix-assisted laser desorption/ionization and electrospray ionization mass spectrometric analysis. Unlike previous examples of matrix-assisted laser desorption ionization (MALDI) mass tags, fulvene-maleimide cycloaddition adducts fragment predictably upon gas-phase activation without the need for bulky photocleavable groups. Further exploration of this chemistry could lead to new approaches for mass spectrometry-based bioassays.

Introduction

The non-destructive chemical modification of biomolecules, broadly labelled as ‘bioconjugation’ chemistry, has underpinned many important developments in biomedical research, including new diagnostic methods, imaging reagents, industrial catalysis and drug design.¹ The copper-catalyzed azide-alkyne cycloaddition reaction is a well-established example of such chemistry; however, continued exploration of catalyst-free alternatives to the ‘copper click’ reaction is a common motivation for present-day organic chemists.² Popular examples include the strain-promoted azide-alkyne cycloaddition (SPAAC) and inverse-electron-demand Diels–Alder (IEDDA), which both have appropriate kinetics at ambient temperatures and in aqueous conditions to make them applicable for the non-denaturing bioconjugation of proteins.³⁻⁴ The Diels–Alder (DA) cycloaddition reactions of maleimides and furans are another attractive subclass of biorthogonal chemistry, owing to their high yields under mild aqueous conditions, but are yet to be widely explored for applications outside materials synthesis.⁵⁻⁶ This reaction is considered reversible, as the resulting adduct exists in a dynamic equilibrium with the diene and dienophile reactants, the position of which is dependent on temperature, solvent, and concentration.⁷ While this is a useful property for many applications, e.g. the manufacture of self-healing polymers, it is undesirable in scenarios where long-term stability and fast reaction kinetics are important.⁷⁻⁹ In contrast, the DA adducts of fulvene and maleimide cycloadditions (**Scheme 1**) exhibit significantly enhanced kinetics and stability under physiological conditions.⁸ Recently, this reactivity has been exploited for the synthesis of antibacterial polymers¹⁰ and cell-encapsulating hydrogels⁸; however, to the best of our knowledge it has yet to be utilized as a strategy for the direct covalent functionalization of biomolecules.



Scheme 1.

The present study sought to demonstrate application of the maleimide-fulvene DA reaction to generate stable protein conjugates using a pentafulvene-based chemical crosslinker with an amine-reactive pentafluorophenyl (PFP) ester and maleimide-labelled synthetic peptides.¹¹ A photocleavable *ortho*-nitrobenzyl linker was included in the design, which was intended to result in release of the modified peptide on excitation by a 355 nm laser, such as that of modern matrix-assisted laser desorption/ionization mass spectrometer (MALDI-MS) sources.¹² This conjugate could therefore act as a reporter ion, or ‘mass tag,’ used for the MS detection of biomolecules that are otherwise difficult to accurately identify using MALDI, as previously demonstrated for antibodies, lectins and oligonucleotides.¹³⁻¹⁵

This approach has a wide range of biomedical applications, including MALDI tissue imaging, genotyping of single nucleotide polymorphisms, protein kinase assays, immunoassays and protein interaction reporters.¹⁶⁻²¹ For example, the direct analysis of tryptic products of unlabeled primary antibodies has been used to detect single antigens within tissues and co-localize them with endogenous proteins.²² However, the detection of multiple antigens using this approach has not been demonstrated, presumably because it relies on identification of antibody-specific peptides that may be shared between primary antibodies of different specificities. Instead, various strategies using photocleavable mass tags have been used to label biomolecules so that cleaved fragments generated *in situ* during the MALDI ionization process can be used as a proxy for their targets, which otherwise make poor MS analytes. Previous examples of MALDI mass tags have utilized varying tactics for modifying the mass-to-charge ratios (m/z) of photocleaved fragments so that multiple labelled biomolecules can be detected in the same experiment.^{14, 18, 21} Here we propose the maleimide-fulvene DA reaction as a simple method for modifying the structure of mass tags, provide an assessment of the reactivity of this linker and its gas-phase fragmentation, and identify a novel strategy for synthesizing MS-cleavable conjugates.

Experimental Methods

Synthesis of fulvene linkers

Materials. All reagents were purchased from Sigma-Aldrich (Castle Hill, Australia) or Chem-Supply (Gillman, Australia) unless stated otherwise. Light-sensitive reactions and intermediates (compounds **2-9**) were protected from prolonged exposure to light by covering in aluminum foil. Column chromatography was performed on 0.04-0.063 mm (Carl Roth GmbH & Co. KG, Karlsruhe, Germany) silica gel 60 deactivated by washing with dilute triethylamine or on 0.2-0.5 mm silica gel (Scharlau Science Group, Barcelona, Spain), for acid-sensitive compounds **5** and **6**, and compound **7**, respectively.

General methods. Nuclear magnetic resonance (NMR) spectra were obtained using an Agilent 500/54 Premium Shielded NMR spectrometer (Agilent Technologies, Santa Clara, US) operating at 500, 125 and 470 MHz for ^1H , ^{13}C and ^{19}F , respectively. High-resolution mass spectrometry (HRMS) data were obtained using an Agilent 6230 TOF spectrometer equipped with an Infinity 1260 LC system (Agilent Technologies). The syntheses of compounds **2** and **3** were adapted from methods described previously.²³⁻²⁴ Synthesis of 4-(cyclopenta-2,4-dien-1-ylidene)pentanoic acid (**4**) was adapted from a previously reported method (see Supplemental Information [SI] for full details).²⁵

*4-(4-{1-[4-(Cyclopenta-2,4-dien-1-ylidene)pentanamido]ethyl}-2-methoxy-5-methylphenoxy)butanoic acid (**5**).* 4-(Cyclopenta-2,4-dien-1-ylidene)pentanoic acid (**5**; 167 mg, 1.0 mmol), NHS (114 mg, 0.99 mmol) and *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (188 mg, 0.98 mmol) were added to dichloromethane (5 mL) and stirred at room temperature for 30 min. 4-[4-(1-Aminomethyl)-2-methoxy-5-nitrophenoxy]butanoic acid hydrochloride (**3**; 165 mg, 0.49 mmol) was added and the reaction stirred under a N_2 atmosphere for 44 h at ambient temperature. The reaction was quenched with 1 M acetic acid (5 mL) and the aqueous layer re-extracted with dichloromethane (2 x 5 mL). The combined organic extracts were washed with saturated sodium chloride solution (5 mL), dried over anhydrous MgSO_4 and the solvent was removed under reduced pressure. The crude residue was purified via silica flash chromatography (0-5 % methanol in dichloromethane) to obtain a yellow powder (63 mg, 29 %). ^1H NMR (599 MHz, CD_3OD): δ_{H} 7.54 (s, 1H), 7.00 (s, 1H), 6.44 (dt, $J = 5.3, 1.9$ Hz, 1H), 6.41 (dt, $J = 5.4, 1.8$ Hz, 1H), 6.37 (dt, $J = 5.4, 1.8$ Hz, 1H), 6.32 (dt, $J = 5.4, 1.8$ Hz, 1H), 5.52 (q, $J = 6.9$ Hz, 1H), 4.08 (t, $J = 6.2$ Hz, 2H), 3.90 (s, 3H), 2.77 (q, $J = 7.2$ Hz, 2H), 2.49-2.37 (m, 4H), 2.17 (s, 3H), 2.14 – 2.07 (m, 1H), 1.46 (d, $J = 6.9$ Hz, 3H); ^{13}C NMR (151 MHz, CD_3OD): δ_{C} 173.9, 155.3, 151.1, 148.3, 144.9, 142.0, 136.1, 132.1, 131.8, 121.6, 121.4, 110.3, 110.2, 69.8, 56.8, 46.4, 36.5, 33.5, 32.8, 26.29, 21.9, 20.7; HRMS calcd. for $(\text{C}_{23}\text{H}_{29}\text{N}_2\text{O}_7+\text{H})^+$: 446.2045, found 446.2036 (error = 7.92 ppm).

*2,3,4,5,6-Pentafluorophenyl 4-(4-{1-[4-(cyclopenta-2,4-dien-1-ylidene)pentanamido]ethyl}-2-methoxy-5-nitrophenoxy)butanoate (**6**).* 4-(4-{1-[4-(Cyclopenta-2,4-dien-1-ylidene)pentanamido]ethyl}-2-methoxy-5-methylphenoxy)butanoic acid (**5**; 37 mg, 0.083 mmol) was dissolved in anhydrous *N,N*-dimethylformamide (5 mL). Bis(pentafluorophenyl) carbonate (65 mg, 0.16 mmol) was added, followed by triethylamine (35 μL , 0.25 mmol) and the reaction was stirred at room temperature for 4 h. The mixture was then diluted with

dichloromethane (20 mL) and washed with saturated sodium chloride solution (3 x 5 mL), concentrated under reduced pressure and the remaining solvent removed under a gentle stream of N₂ gas. The crude residue was purified via silica flash chromatography (40:60 ethyl acetate/hexanes). The product was redissolved in dichloromethane (1 mL) and washed with saturated sodium chloride solution (3 x 1 mL), dried over anhydrous MgSO₄ and the solvent removed under reduced pressure to obtain a viscous yellow oil (11 mg, 22 %). ¹H NMR (500 MHz, CDCl₃): δ_H 7.55 (s, 1H), 6.86 (s, 1H), 6.47-6.42 (m, 4H), 6.36 (d, *J* = 7.7 Hz, 1H), 5.45 (t, *J* = 7.3 Hz, 1H), 4.16 (t, *J* = 6.0 Hz, 2H), 3.93 (s, 3H), 2.93 (t, *J* = 7.3 Hz, 2H), 2.86-2.79 (m, 2H), 2.49-2.37 (m, 2H), 2.36-2.28 (m, 2H), 2.17 (s, 3H), 1.50 (d, *J* = 7.1 Hz, 3H) ppm; ¹³C NMR (126 MHz, CDCl₃): δ_C 171.08, 169.03, 154.02, 150.69, 147.03, 143.58, 140.8, 142.3-137.0 (m), 133.7, 131.6, 120.8, 120.2, 111.7, 110.7, 67.9, 56.5, 48.2, 36.0, 32.4, 30.0, 29.9, 24.4, 20.8, 20.7; ¹⁹F NMR (470 MHz, CDCl₃): δ_F -162.3 (m, 2F), -157.9 (m, 1F), -152.8 (m, 2F); HRMS calcd. for (C₂₉H₂₇N₂O₇F₅+H)⁺: 611.1808, found 611.1815 (error = 0.64 ppm).

2,5-Dioxopyrrolidin-1-yl 4-(cyclopenta-2,4-dien-1-ylidene)pentanoate (7). 4-(Cyclopenta-2,4-dien-1-ylidene)pentanoic acid (**4**; 200 mg, 1.2 mmol), *N,N'*-disuccinimidyl carbonate (624 mg, 2.5 mmol) and *N,N*-diisopropylethylamine (DIEA; 844 μL, 4.9 mmol) were dissolved in dichloromethane (10 mL) and the reaction was stirred at room temperature for 4 h. The solvent was removed under reduced pressure and the crude mixture was purified via silica flash chromatography (100 % ethyl acetate) to obtain a waxy yellow solid (306 mg, 98 %). ¹H NMR (500 MHz, CDCl₃): δ_H 6.49 (s, 4H), 2.99-2.92 (m, 2H), 2.88-2.81 (m, 6H), 2.22 (s, 3H); ¹³C NMR (126 MHz, CDCl₃) δ 169.4, 168.1, 148.4, 144.3, 132.3, 132.2, 31.5, 30.7, 26.0, 20.9; HRMS calcd. for (C₁₄H₁₅NO₄+H)⁺: 262.1071, found 262.1075 (error = 0.57 ppm).

Synthesis of Maleimide-beta-Ala-Peptides (Mal-peptides)

Materials. Ethyl cyanohydroxyiminoacetate (oxyma), O-Benzotriazole-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU), 1-hydroxybenzotriazole (HOBt), DIEA, *N,N*-dimethylformamide (DMF), piperidine, trifluoroacetic acid (TFA) and 9-fluorenylmethoxycarbonyl (Fmoc) amino acids were obtained from Auspep Pty Ltd (Australia). Triisopropylsilane (TIPS) and *N,N'*-diisopropylcarbodiimide (DIC) were obtained from Sigma-Aldrich (Australia). Diethyl ether and dichloromethane were obtained from BDH (UK). 2,4,6-Trinitrobenzene sulfonic acid (TNBSA) was kindly provided in powder form by Dr Denis Scanlon, University of Adelaide. Unless otherwise stated chemicals were of peptide synthesis grade or its equivalent.

Procedures. The peptide, GSGSGRGS GSG, was chemically synthesized on a CEM Liberty-Blue microwave-assisted synthesizer (Ai Scientific, Australia). The peptide was synthesized as a C-terminal carboxamide assembled from Rink amide-resin (Merck Millipore, Australia), in the Fmoc/tBu mode of synthesis. For a 0.1 mmol reaction scale, Fmoc-deprotection was performed in two stages by initial treatment with 20 % piperidine/DMF (v/v) under microwave radiation for 30 s (40 W, 40 °C), followed by filtration and a second addition of the above solution (45 W, 90 °C; 3 min). The peptide-resins were then rinsed with dimethylformamide (4 x 7 mL). Acylation was achieved by the addition of a solution containing Fmoc-amino acid (4 eq, relative to reaction scale), oxyma (4 eq) and DIC (8 eq) in DMF (v/v; 4 mL) to the N α -deprotected peptide-resin and the mixture agitated under

microwave radiation for 10 min (30 W, 90 °C). Dichloromethane (5 × 2 min) was used to rinse the peptide-resins as a final washing step.

After the removal of the final Fmoc group beta-alanine-maleimide was coupled to the N α -terminal amino group of each peptide by the manual addition of a solution containing beta-alanine-maleimide (2 eq, relative to reaction scale), HBTU (2 eq) and DIEA (4 eq) in DMF (2 mL, pre-activated 5 mins) to the N α -deprotected peptide-resin and the mixture gently agitated for 1 h. Coupling was monitored by the TNBSA test and following a negative TNBSA test,²⁶ the resin was washed with DMF (1 x 10 mL) and dichloromethane (5 x 10 mL).

The Mal-peptides were cleaved from the resin support by the addition of TFA/TIPS/water (95:3:2, % v/v/v/v; 5 mL) for 2.5 h, after which the cleavage filtrates were evaporated under nitrogen flow and the crude product was isolated by precipitation in cold ether and washed in cold ether (4,500 g, 4 × 30 mL). The dried crude Mal-peptides were then dissolved in 0.1 % (v/v) TFA in milliQ water and purified using a semi-preparative ZORBAX 300 SB-C18 column (9.4 mm x 25 cm) installed in a High Performance Liquid Chromatography (HPLC) 1200 system (Agilent Technologies, Australia) under a flow rate of 2 mL/min using buffer A (0.1 % v/v TFA in milliQ water) and buffer B (0.1 % TFA in 90 % acetonitrile, 10 % milliQ water, v/v) as the limiting solvent. Peptide detection was performed by absorbance at 214 nm. The desired fraction and analysis of the purified peptide and Mal-peptides were identified by Thermo Orbitrap Exactive liquid chromatography–mass spectrometry (LC-MS; ThermoFisher Scientific, Scoresby, Australia). ESI-MS calcd. for (M+H)⁺: 1015.4 Da, found 1015.3 Da.

Conjugation of photocleavable linker to peptide

A 2.5 mM stock solution of maleimide-labelled peptide in DMSO (20 μ L) was diluted in phosphate-buffered saline (PBS; 169.8 μ L), followed by the addition of a 182 mM stock solution of photocleavable linker **6** in dry dimethyl sulfoxide (DMSO; 0.2 μ L) and the mixture was incubated at 37 °C for 2 h in a Thermomixer comfort shaking incubator (Eppendorf, Germany) set to 300 rpm. 2 μ L aliquots were taken at 30, 60, and 120 min and 24 h timepoints to monitor the reaction progress and immediately used for MALDI-MS analysis.

Labelling of lysozyme with peptide mass tag via non-photocleavable linker

A 27 mM stock solution of non-photocleavable linker **7** in dry DMSO (5.26 μ L) was added to a 1 mg/mL solution of lysozyme in PBS (100 μ L), followed immediately by the 2.5 mM maleimide-peptide stock solution in water (20 μ L), and the mixture was incubated at 37 °C for 2 h in a Thermomixer comfort shaking incubator set to 300 rpm. For MALDI-MS, the modified protein (**7**-labelled lysozyme-peptide conjugate) was separated from unreacted/hydrolyzed linker using a 10 kDa MWCO Amicon[®] Ultra-0.5 Centrifugal Filter Device (Merck Millipore, USA) to a final volume of 100 μ L in PBS and diluted 10-fold in PBS prior to MS analysis. Alternatively, the amide coupling reaction was quenched with 200 mM ammonium acetate (375 μ L) and stored at 4 °C until concentration using a 10 kDa MWCO Amicon[®] Ultra-0.5 Centrifugal Filter Device into 200 mM ammonium acetate (100 μ L) and analysis via nanoelectrospray ionization (nanoESI)-MS analysis.

MS analysis of linker-peptide conjugates and labelled proteins

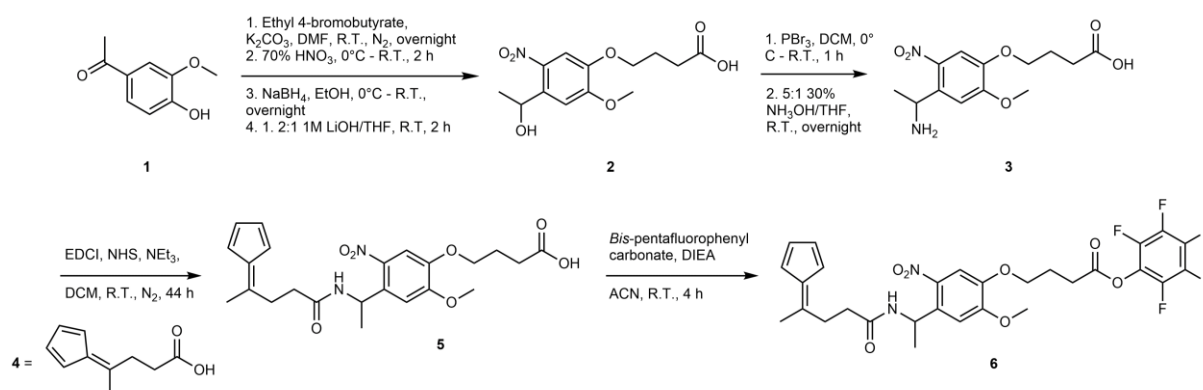
MALDI-MS. Linker **6**-peptide conjugates and **7**-labelled lysozyme-peptide conjugate were prepared for analysis via MALDI-MS by mixing with equal volumes of 10 mg/mL α -cyano-4-hydroxycinnamic acid (CHCA; Bruker Daltonics, USA) matrix in 70 % (v/v) acetonitrile with 1 % (v/v) trifluoroacetic acid. The pre-mixed solutions (0.5-1 μ L) were spotted onto an MTP AnchorChip™ 384 MALDI target (Bruker Daltonics). Profiling and LIFT spectra were acquired using an UltraFlex™ MALDI-TOF/TOF mass spectrometer (Bruker) in positive reflectron mode. Collision-induced dissociation (CID) spectra and trapped ion mobility spectra (TIMS) were acquired on a timsTOF flex MALDI-quadrupole (Q)-TOF mass spectrometer (Bruker). MALDI spectra were converted to either .mzXML using FlexAnalysis v3.4 (Bruker Daltonics), or .xy file format using DataAnalysis v5.3 (Bruker Daltonics), respectively, then processed (including smoothing, cropping and baseline subtraction) using mMass v5.5.0 (<http://www.mmass.org/>).²⁷

NanoESI-MS. Spectra were acquired using a 6560 Ion Mobility LC/Q-TOF mass spectrometer fitted with a nanospray source, calibrated using the Agilent LC/MS tuning mix (both Agilent, Santa Clara, USA). The purified solution of **7**-labelled lysozyme-peptide conjugate in 200 mM ammonium acetate was diluted 1:1 with acetonitrile to denature the protein and nanosprayed using platinum-coated borosilicate glass capillaries prepared in-house using a P-97 Flaming/Brown micropipette puller (Sutter Instrument Company, Novato, US). A low-pressure N₂ gas supply was fitted to the capillary housing to maintain sample flow through the nanospray source. Spectra were acquired, analyzed and deconvoluted using the MassHunter Workstation software vB.07.00 with the BioConfirm Software add-on (Agilent).

Results and Discussion

Design and synthesis of photocleavable linker

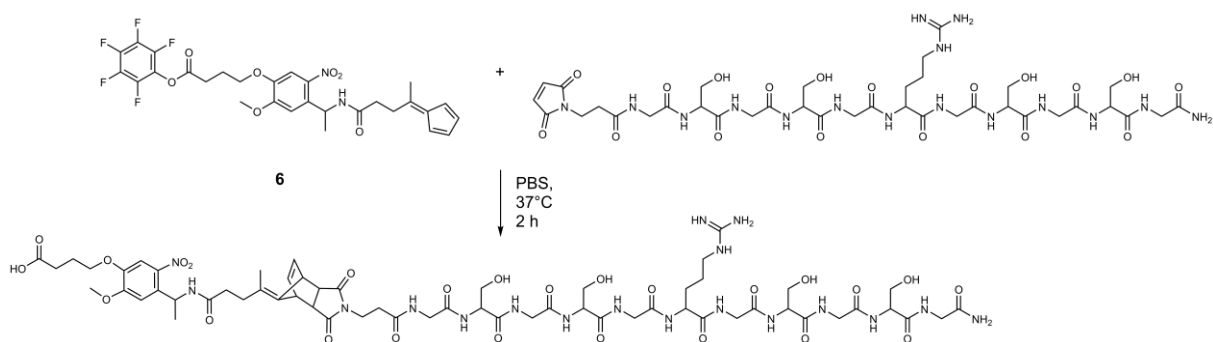
A number of different photocleavable groups have been used as the basis for designing photocleavable mass tags, including 6,11-dihydrothiochromeno[4,3-b]indole, *o*-nitrobenzyl, triphenylmethyl (trityl), Ru(II) polypyridine complex, phenacyl and coumarin-based linkers.²⁸⁻³² Our linkers were based on the *o*-nitrobenzyl group, owing to their straightforward synthesis, appropriate absorption maxima (≈ 350 nm) and easily modified structure (for general photocleavage mechanism see SI, **Scheme S1**).³³ Starting from 4'-hydroxy-3'-methoxyacetophenone **1**, nitrobenzyl linker **6** was synthesized in seven steps (**Scheme 2**). 4-(2,4-Cyclopentadien-1-ylidene)pentanoic acid **4** was synthesized as described previously.²⁵ The final linker **6** contains a pentafluorophenyl ester for amide coupling to primary amines, such as lysine residue side chains and protein N-termini, a photocleavable nitrobenzyl amide, and a pentafulvene for [4+2] cycloaddition with maleimide derivatives.^{8, 11}



Scheme 2.

MALDI-MS analysis of photocleavable linker-peptide conjugates

Synthetic peptides make effective MALDI mass tags, owing to their modular and simple synthesis, and tunable and precise molecular weight, charge and solubility.^{13, 30, 34} We demonstrated the successful conjugation of pentafulvene linker **6** to a maleimide-labeled peptide (Maleimide-beta-Ala-GSGSGRGS-GSG-NH₂) under mild aqueous conditions (**Scheme 3**; $37^\circ C$ in PBS), with the molecular ion and neutral loss ($M-H_2O+H$)⁺ for the expected product detected using MALDI-MS, depending on the instrument used for analysis (**Figure 1**). However, even after overnight incubation, the predominant signal ($m/z = 1015.4$) corresponded to the peptide without the additional mass of the linker.



Scheme 3.

Initial analysis of the reaction mixture using a Bruker UltrafleXtreme MALDI-time of flight (TOF) instrument, which features a Smartbeam-II™ 1kHz 355 nm Nd:YAG laser, did not produce detectable amounts of the expected photodissociation product; instead, the neutral loss fragment ion for the intact cycloaddition adduct with hydrolyzed PFP ester (m/z 1441.6) was observed (**Figure 1A**). Conversely, analysis of the same samples using the timsTOF fleX MALDI-Q-TOF produced the desired photodissociation product (m/z 1178.5) even at lower laser powers (**Figure 1B**). We rationalized these observations as a result of differences in the design and performance of the MALDI sources for each instrument. The timsTOF fleX incorporates a 355 nm Smartbeam 3-D™ laser, which has a higher repetition rate (≥ 10 kHz versus 1 kHz for the Smartbeam-II™), a factor which has been shown to significantly influence the optimum laser energy per pulse and therefore UV-induced fragmentation. Furthermore, this instrument utilizes an approach called ‘beam steering,’ which essentially focuses the laser beam to a narrower rectangular acquisition area, thus improving sensitivity.³⁵⁻³⁹ The timsTOF fleX also operates under intermediate pressure, as opposed to the high vacuum of the UltrafleXtreme, which therefore results in collisional cooling and a potential reduction in in-source fragmentation.⁴⁰ These differences in source configuration could explain why the same photodissociation pattern was not observed for spectra obtained using the UltrafleXtreme, in addition to possible deterioration in the performance and sensitivity of this older instrument over time.

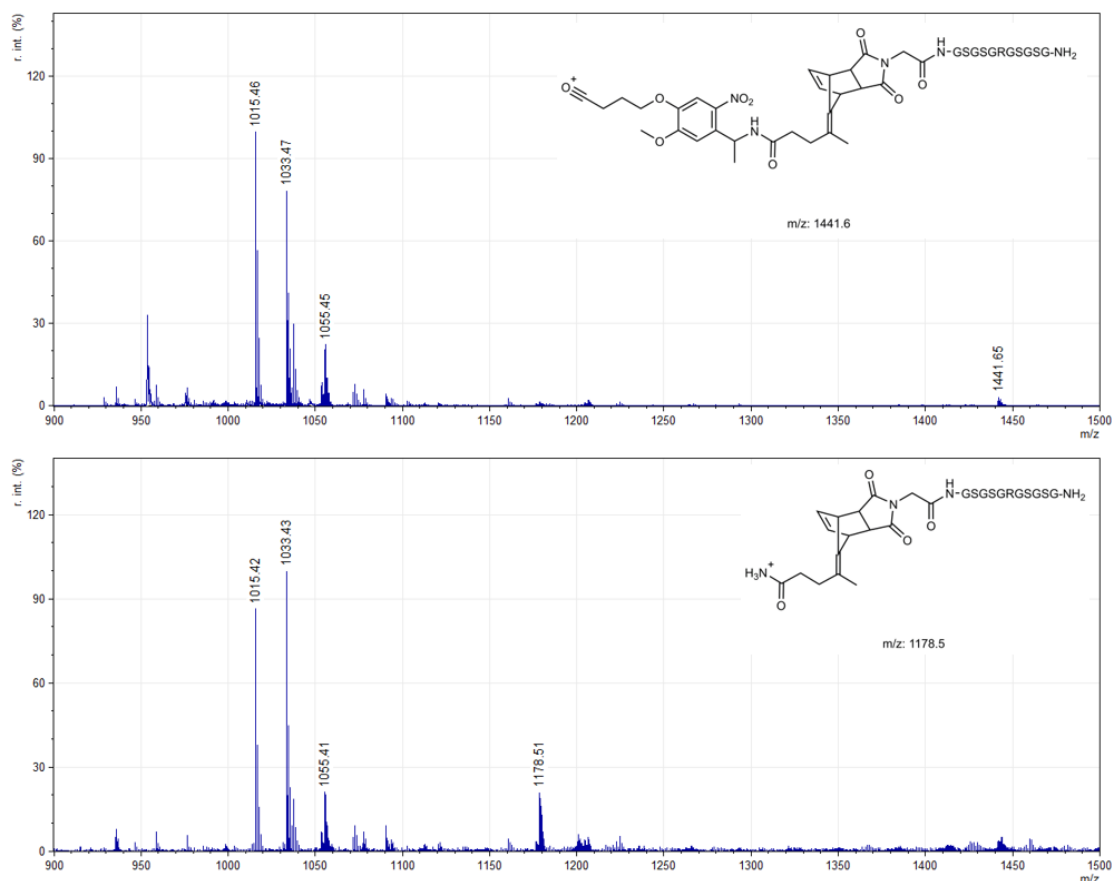


Figure 1. Matrix-assisted laser desorption/ionization mass spectrometric analysis of fulvene linker and maleimide-labeled peptide cycloaddition product using UltrafleXtreme (A) and timsTOF flex (B) mass spectrometers.

Both spectra were dominated by apparently unreacted peptide (with and without hydrolysis of the maleimide to the open maleamic acid form: $m/z \approx 1033$ and 1015 , respectively), for which we hypothesized two potential reasons: (1) some precipitation of the linker was observed on addition to the PBS reaction buffer, suggesting that the majority may have remained unreacted, and/or (2) that the cycloaddition product was fragmenting back into the maleimide precursor. Although addressing poor aqueous solubility required re-evaluation of the linker design, the possibility that an alternative fragmentation mechanism was occurring required further investigation.

MALDI-MS analysis of photocleavable linker-peptide conjugate gas-phase fragmentation

To further investigate the fragmentation pathways of these conjugates, we identified the products of CID for the respective ions identified using the UltrafleXtreme, operated in 'LIFT' mode, and timsTOF flex operated in MS/MS mode (Figures 2A and B, respectively). For both ions (m/z 1441.6 and 1178.5) the major fragments corresponded to the expected m/z of the protonated unmodified maleimide-labelled peptide. The higher mass resolution and narrower precursor ion selection of the timsTOF flex allowed less ambiguous assignment of the peptide, with only 1.2 ppm deviation between the observed m/z 1015.42 ion and the predicted peptide m/z (1015.4188).

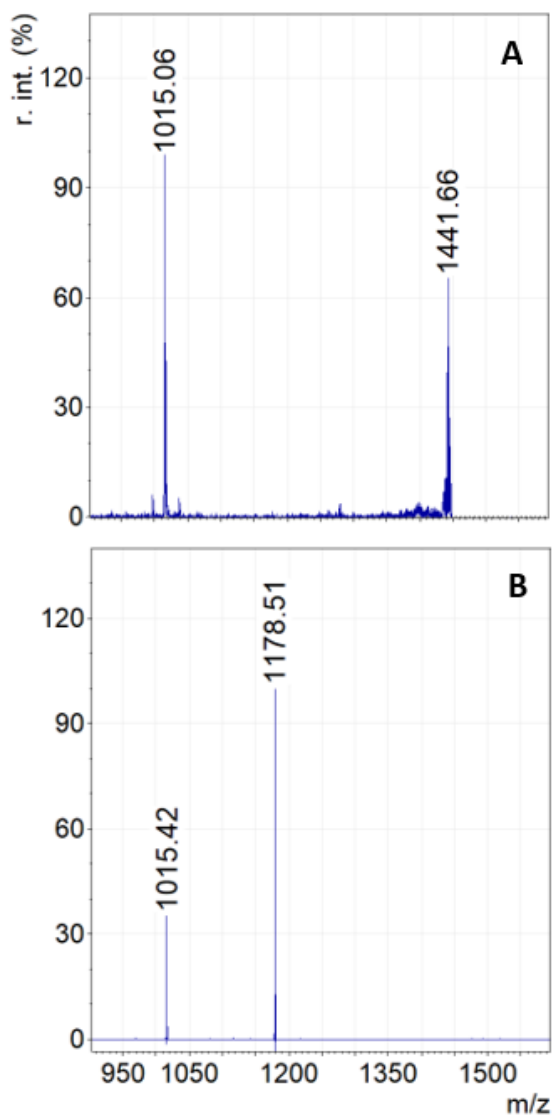


Figure 2. Matrix-assisted laser desorption/ionization mass spectrometric analysis of cycloaddition product gas-phase fragmentation using UltrafleXtreme (A) and timsTOF flex (B) mass spectrometers operated in LIFT (Lift 1 = 18.98 kV, Lift 2 = 3.70 kV) and collision-induced dissociation (CID, collision voltage = 35eV) modes, respectively.

Additional confirmation of this hypothesis was obtained by using TIMS to demonstrate that both precursor (m/z 1178.5164) and CID product ions (m/z 1015.4165) had identical inverse reduced mobility ($1/K_0$) values (**Figure 3**). The configuration of the timsTOF flex, with ions directed through the TIMS tunnel prior to the quadrupole and collision cell, means that TIMS separation occurs prior to CID; therefore, ions originating from different precursors are expected to have different $1/K_0$ values. The value of $1/K_0$ for this precursor ion also differed from that of the unmodified peptide in precursor ion spectra (**Figure S1**).

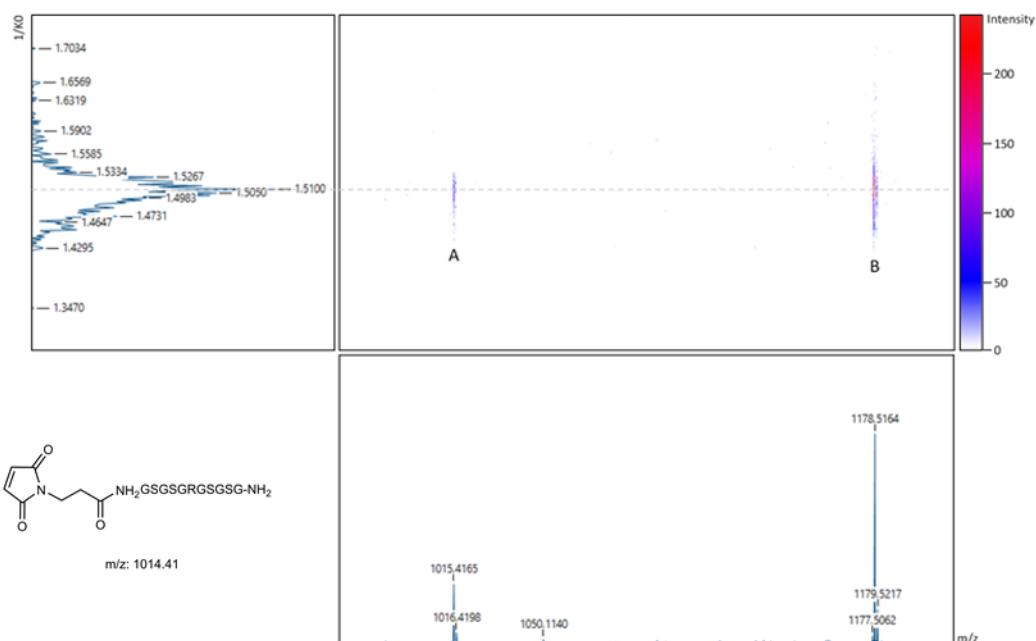
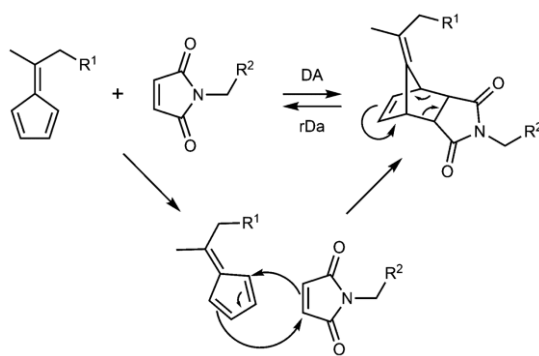


Figure 3. Trapped ion mobility-matrix-assisted laser desorption/ionization mass spectra and mobiligram of the photocleaved cycloaddition product showing identical $1/K_0$ distributions for selected precursor ion (A) and resulting CID fragment ion (B), which corresponds to the unmodified maleimide-labelled peptide.

Together, these results strongly support our proposal that the covalent bonds formed by the cycloaddition reaction between maleimide and fulvene derivatives are somewhat labile in the gas phase, despite its reported stability in aqueous solutions.⁷⁻⁹ Based on similar fragmentation behavior observed for furan-maleimide adducts, this transition likely occurs via a retro-DA (rDA) mechanism (**Scheme 4**).⁴¹

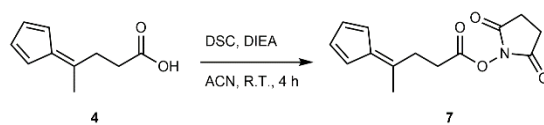


Scheme 4.

Given this apparent labile nature of fulvene-maleimide adducts in the gas phase, we sought to determine if a similar fragmentation pattern would occur for conjugates without the presence of a photolabile group that might enable development of a new class of mass tag reagents for biomolecule detection.

MALDI-MS analysis of protein-peptide conjugates

The removal of the photocleavable *o*-nitrobenzyl group enabled the one-step synthesis of fulvene linker **7**, which featured an *N*-hydroxysuccinimide (NHS) ester for amide coupling to lysine residues and protein N termini (**Scheme 5**).⁴²



Scheme 5.

The application of this linker for labelling large biomolecules was demonstrated using the 14.3 kDa protein lysozyme. This model protein, which contains six lysine residues in the mature protein, was incubated with a mixture of **7** and maleimide-labelled peptide, resulting in simultaneous reaction of the NHS ester with the protein, and cycloaddition of the fulvene group to the peptide.⁴³ The unreacted (hydrolyzed) linker and peptide could be subsequently removed by centrifugal diafiltration; hence, detection of the unmodified peptide in MALDI spectra of the purified conjugate (**Figure 4**) likely resulted from fragmentation of the desired cycloaddition product.

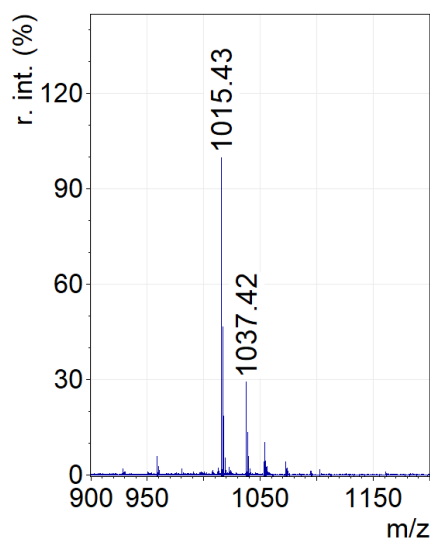


Figure 4. Matrix-assisted laser desorption/ionization mass spectrometric analysis of lysozyme labelled with non-photocleavable fulvene linker and maleimide-labeled peptide cycloaddition product.

ESI-MS analysis of protein-peptide conjugates

Given the promising results of the MALDI analysis of mass tag-labelled lysozyme, and the enhanced rDA fragmentation of fulvene **6**-maleimide adducts under mild CID conditions (**Figure 2B**), it seemed reasonable to assume that **7** could also be utilized for enhanced detection of proteins via ESI-MS. Analysis of the intact lysozyme conjugate using nanoESI-MS and subsequent spectral deconvolution to identify different species amongst

overlapping charge states revealed modifications consistent with the addition of ≤ 3 linker-peptide adducts to the protein (predicted mass shift per peptide label m/z 1161.49, deconvoluted $\Delta m/z$ 1161.12 \pm 0.27; **Figure 5A**). The base peak at m/z 1934.1629, which corresponded to a single mass tag addition to the 8+ charge state, was then subjected to CID at 30 V, producing fragment ions consistent with the rDA peptide product, and the 7+ charge state for lysozyme with the residual fragment of the linker (m/z 2065.2054, or 1.4456 kDa; **Figure 6B**). These results suggest a broader utility for this linker as a dual MALDI/ESI mass tag reagent.

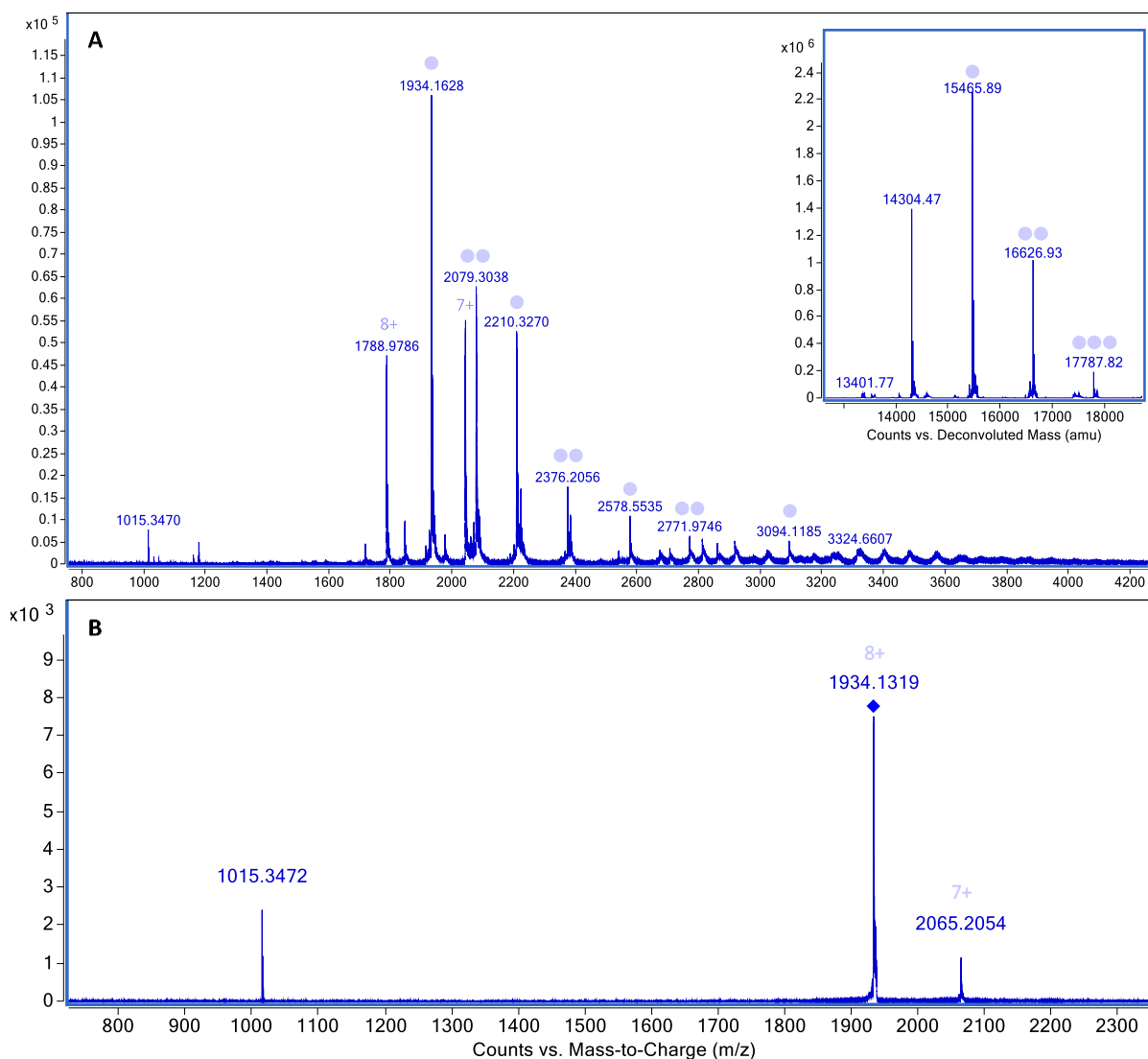


Figure 5. Nano-electrospray ionization-mass spectra of intact lysozyme labelled with non-photocleavable fulvene linker **7**. Intact precursor spectra (**A**) and subsequent deconvolution (inset) shows modification to the protein with up to 3 peptide mass tags. Collision-induced dissociation of the 8+ protein charge state with a single mass tag (**B**) produced fragments assigned to the maleimide-labelled peptide, and lysozyme with the residual linker fragment after stripping of the singly charged peptide. Blue circles indicate number of peptide mass tags determined from mass shifts in deconvoluted spectra. Blue diamond indicates precursor selected for CID fragmentation.

At present, we could only find one other example where MS-induced rDA fragmentation has been exploited for the generation of mass tags in which furan-maleimide adducts were utilized for monitoring enzyme-

catalyzed reactions of immobilized substrates on silicon chips using a MALDI-TOF instrument.⁴¹ Despite their efficient fragmentation, the subdued reaction kinetics of furan-maleimide cycloadditions and reduced stability under physiological conditions limit their utility outside of organic synthesis. In contrast, we have shown that the fulvene-maleimide DA reaction can be used to rapidly generate protein bioconjugates in aqueous buffer at low temperatures and has facile fragmentation properties that enable mass tag detection in both MALDI and ESI experiments.^{8,41}

Conclusions

This study demonstrated the application of two novel MS-cleavable fulvene-based linkers for labelling proteins with maleimide-containing mass tags. The initial design was based on previous examples of photocleavable linkers intended for the detection of peptide mass tags using MALDI-MS; however, the addition of a fulvene group provided a convenient handle for the attachment of maleimide-containing groups and thereby modifying the mass of the detected photocleavage product.¹³ Furthermore, MS analysis of photocleavable linker-peptide adducts identified efficient rDA fragmentation, thus the bulky *o*-nitrobenzyl moiety was deemed redundant for this application. The photocleavable linker could instead be employed as a bifunctional reagent in other applications for which combined photocleavable and bio-orthogonal reactivities are desirable, such as the synthesis of photoreactive caged protein conjugates.⁴⁴⁻⁴⁵

These versatile reagents could be potentially useful for derivatizing large intact proteins containing reactive amines, such as antibodies, and thereby enabling their differentiation via MS, despite overlapping charge states or poor mass resolution.⁴⁶ The simple synthetic route for this linker could also make this strategy applicable to a broader variety of biomolecules or synthetic targets, as the NHS ester could easily be substituted for other reactive groups. Continued optimization and hence better control over conjugation efficiency would enable straightforward quantitation of biomolecules using isotope-labelled peptides, thereby negating the need for expression or synthesis of isotope-labeled standards. Further investigation of bio-orthogonal DA chemistries could reveal promising new avenues for developing multiplexed MS bioassays, in applications such as multiplexed MALDI/ESI-MS immunoassays, MS imaging, mutation screening, enzyme activity assays and identifying protein interactions.^{16-21, 29, 47}

Acknowledgements

KGS acknowledges financial support from a University of Adelaide Faculty of Sciences Divisional Scholarship. KP acknowledges scholarship support from the Cooperative Research Centre for Cell Therapies Manufacturing.

References

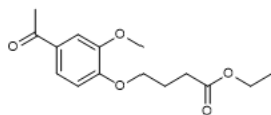
1. Kalia, J.; Raines, R. T., Advances in Bioconjugation. *Curr. Org. Chem.* 2010, *14* (2), 138-147.
2. Sletten, E. M.; Bertozzi, C. R., Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* 2009, *48* (38), 6974-6998.
3. Jewett, J. C.; Bertozzi, C. R., Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev.* 2010, *39* (4), 1272-1279.
4. Becer, C. R.; Hoogenboom, R.; Schubert, U. S., Click chemistry beyond metal-catalyzed cycloaddition. *Angew. Chem. Int. Ed.* 2009, *48* (27), 4900-4908.
5. Sanyal, A., Diels–Alder cycloaddition-cycloreversion: A powerful combo in materials design. *Macromol. Chem. Phys.* 2010, *211* (13), 1417-1425.
6. Gabano, E.; Perin, E.; Bonzani, D.; Ravera, M., Conjugation between maleimide-containing Pt(IV) prodrugs and furan or furan-containing drug delivery vectors via Diels–Alder cycloaddition. *Inorganica Chim. Acta* 2019, *488*, 195-200.
7. Boutelle, R. C.; Northrop, B. H., Substituent effects on the reversibility of furan–maleimide cycloadditions. *J. Org. Chem.* 2011, *76* (19), 7994-8002.
8. Madl, C. M.; Heilshorn, S. C., Rapid Diels–Alder cross-linking of cell encapsulating hydrogels. *Chem. Mater.* 2019, *31* (19), 8035-8043.
9. Gandini, A., The furan/maleimide Diels–Alder reaction: A versatile click–unclick tool in macromolecular synthesis. *Prog. Polym. Sci.* 2013, *38* (1), 1-29.
10. Ilker, M. F.; Nüsslein, K.; Tew, G. N.; Coughlin, E. B., Tuning the hemolytic and antibacterial activities of amphiphilic polynorbornene derivatives. *JACS* 2004, *126* (48), 15870-15875.
11. So, W. H.; Zhang, Y.; Kang, W.; Wong, C. T. T.; Sun, H.; Xia, J., Site-selective covalent reactions on proteinogenic amino acids. *Curr. Opin. Biotechnol.* 2017, *48*, 220-227.
12. Hansen, M. J.; Velema, W. A.; Lerch, M. M.; Szymanski, W.; Feringa, B. L., Wavelength-selective cleavage of photoprotecting groups: strategies and applications in dynamic systems. *Chem. Soc. Rev.* 2015, *44* (11), 3358-3377.
13. Lemaire, R.; Stauber, J.; Wisztorski, M.; Van Camp, C.; Desmons, A.; Deschamps, M.; Proess, G.; Rudlof, I.; Woods, A. S.; Day, R.; Salzet, M.; Fournier, I., Tag-Mass: specific molecular imaging of transcriptome and proteome by mass spectrometry based on photocleavable tag. *J. Prot. Res.* 2007, *6* (6), 2057-2067.
14. Thiery, G.; Shchepinov, M. S.; Southern, E. M.; Audebourg, A.; Audard, V.; Terris, B.; Gut, I. G., Multiplex target protein imaging in tissue sections by mass spectrometry – TAMSIM. *Rapid Commun. Mass Spectrom.* 2007, *21* (6), 823-829.
15. Dai, C.; Cazares, L. H.; Wang, L.; Chu, Y.; Wang, S. L.; Troyer, D. A.; Semmes, O. J.; Drake, R. R.; Wang, B., Using boronolactin in MALDI-MS imaging for the histological analysis of cancer tissue expressing the sialyl Lewis X antigen. *Chem. Comm.* 2011, *47* (37), 10338-10340.
16. Thiery, G.; Anselmi, E.; Audebourg, A.; Darii, E.; Abarbri, M.; Terris, B.; Tabet, J. C.; Gut, I. G., Improvements of TArgeted multiplex mass spectrometry IMaging. *Proteomics* 2008, *8* (18), 3725-3734.
17. Hammond, N.; Koumi, P.; Langley, G. J.; Lowe, A.; Brown, T., Rapid mass spectrometric identification of human genomic polymorphisms using multiplexed photocleavable mass-tagged probes and solid phase capture. *Org. Biomol. Chem.* 2007, *5* (12), 1878-1885.

18. Zhou, G.; Khan, F.; Dai, Q.; Sylvester, J. E.; Kron, S. J., Photocleavable peptide–oligonucleotide conjugates for protein kinase assays by MALDI-TOF MS. *Mol. Biosyst.* 2012, 8 (9), 2395-2404.
19. Yang, L.; Tang, X.; Weisbrod, C. R.; Munske, G. R.; Eng, J. K.; von Haller, P. D.; Kaiser, N. K.; Bruce, J. E., A photocleavable and mass spectrometry identifiable cross-linker for protein interaction studies. *Anal. Chem.* 2010, 82 (9), 3556-3566.
20. Lorey, M.; Adler, B.; Yan, H.; Soliymani, R.; Ekström, S.; Yli-Kauhaluoma, J.; Laurell, T.; Baumann, M., Mass-tag enhanced immuno-laser desorption/ionization mass spectrometry for sensitive detection of intact protein antigens. *Anal. Chem.* 2015, 87 (10), 5255-5262.
21. Stauber, J.; El Ayed, M.; Wisztorski, M.; Day, R.; Fournier, I.; Salzet, M., Polymerase chain reaction and immunoassay–matrix assisted laser desorption mass spectrometry using Tag-Mass technology: new tools to break down quantification limits and multiplexes. *Anal. Chem.* 2009, 81 (22), 9512-9521.
22. Longuespée, R.; Boyon, C.; Desmons, A.; Kerdraon, O.; Leblanc, E.; Farré, I.; Vinatier, D.; Day, R.; Fournier, I.; Salzet, M., Spectroimmunohistochemistry: a novel form of MALDI mass spectrometry imaging coupled to immunohistochemistry for tracking antibodies. *OMICS* 2014, 18 (2), 132-141.
23. Kloxin, A. M.; Tibbitt, M. W.; Anseth, K. S., Synthesis of photodegradable hydrogels as dynamically tunable cell culture platforms. *Nat. Protoc.* 2010, 5 (12), 1867-1887.
24. Griffin, D. R.; Schlosser, J. L.; Lam, S. F.; Nguyen, T. H.; Maynard, H. D.; Kasko, A. M., Synthesis of photodegradable macromers for conjugation and release of bioactive molecules. *Biomacromolecules* 2013, 14 (4), 1199-1207.
25. Coşkun, N.; Erden, I., An efficient catalytic method for fulvene synthesis. *Tetrahedron* 2011, 67 (45), 8607-8614.
26. Hancock, W.; Battersby, J., A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2, 4, 6-trinitrobenzene-sulphonic acid. *Anal. Biochem.* 1976, 71 (1), 260-264.
27. Strohalm, M.; Hassman, M.; Kořata, B.; Kодиček, M., mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun. Mass Spectrom.* 2008, 22 (6), 905-908.
28. Kang, N.; Lee, J.-M.; Jeon, A.; Oh, H. B.; Moon, B., Design and synthesis of new mass tags for matrix-free laser desorption ionization mass spectrometry (LDI-MS) based on 6,11-dihydrothiochromeno[4,3-b]indole. *Tetrahedron* 2016, 72 (36), 5612-5619.
29. Gagnon, H.; Franck, J.; Wisztorski, M.; Day, R.; Fournier, I.; Salzet, M., Targeted mass spectrometry imaging: specific targeting mass spectrometry imaging technologies from history to perspective. *Prog. Histochem. Cyto.* 2012, 47 (3), 133-174.
30. Olejnik, J.; Ludemann, H.-C.; Olejnik, E. K.; Berkenkamp, S.; Hillenkamp, F.; Rothschild, K. J., Photocleavable peptide-DNA conjugates: synthesis and applications to DNA analysis using MALDI-MS. *Nucleic Acids Res.* 1999, 27 (23), 4626-4631.
31. Erkan, H.; Telci, D.; Dilek, O., Design of Fluorescent Probes for Bioorthogonal Labeling of Carbonylation in Live Cells. *Scientific Reports* 2020, 10 (1), 7668.
32. Maki, T.; Ishida, K., Photocleavable molecule for laser desorption ionization mass spectrometry. *J. Org. Chem.* 2007, 72 (17), 6427-6433.

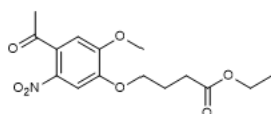
33. Gaplovsky, M.; Il'ichev, Y. V.; Kamdzhilov, Y.; Kombarova, S. V.; Mac, M.; Schwörer, M. A.; Wirz, J., Photochemical reaction mechanisms of 2-nitrobenzyl compounds: 2-Nitrobenzyl alcohols form 2-nitroso hydrates by dual proton transfer. *Photochem. Photobiol. Sci.* 2005, 4 (1), 33-42.
34. Napoli, A.; Athanassopoulos, C. M.; Moschidis, P.; Aiello, D.; Di Donna, L.; Mazzotti, F.; Sindona, G., Solid phase isobaric mass tag reagent for simultaneous protein identification and assay. *Anal. Chem.* 2010, 82 (13), 5552-5560.
35. Trim, P. J.; Djidja, M.-C.; Atkinson, S. J.; Oakes, K.; Cole, L. M.; Anderson, D. M.; Hart, P. J.; Francese, S.; Clench, M. R., Introduction of a 20 kHz Nd: YVO4 laser into a hybrid quadrupole time-of-flight mass spectrometer for MALDI-MS imaging. *Anal. Bioanal. Chem.* 2010, 397 (8), 3409-3419.
36. Spraggins, J. M.; Djambazova, K. V.; Rivera, E. S.; Migas, L. G.; Neumann, E. K.; Fuetterer, A.; Suetering, J.; Goedecke, N.; Ly, A.; Van de Plas, R., High-performance molecular imaging with MALDI trapped ion-mobility time-of-flight (timsTOF) mass spectrometry. *Anal. Chem.* 2019, 91 (22), 14552-14560.
37. Soltwisch, J.; Heijs, B.; Koch, A.; Vens-Cappell, S.; Höhndorf, J.; Dreisewerd, K., MALDI-2 on a trapped ion mobility quadrupole time-of-flight instrument for rapid mass spectrometry imaging and ion mobility separation of complex lipid profiles. *Anal. Chem.* 2020, 92 (13), 8697-8703.
38. Schafer, R., Ultraflexextreme: redefining MALDI mass spectrometry performance. *LCGC North Am.* 2009, 14-15.
39. Steven, R. T.; Dexter, A.; Bunch, J., Investigating MALDI MSI parameters (Part 1)—a systematic survey of the effects of repetition rates up to 20kHz in continuous raster mode. *Methods* 2016, 104, 101-110.
40. Garrett, T. J.; Yost, R. A., Analysis of Intact Tissue by Intermediate-Pressure MALDI on a Linear Ion Trap Mass Spectrometer. *Anal. Chem.* 2006, 78 (7), 2465-2469.
41. Meng, J. C.; Averbuj, C.; Lewis, W. G.; Siuzdak, G.; Finn, M., Cleavable linkers for porous silicon-based mass spectrometry. *Angew. Chem.* 2004, 116 (10), 1275-1280.
42. Mädler, S.; Bich, C.; Touboul, D.; Zenobi, R., Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J. Mass Spectrom.* 2009, 44 (5), 694-706.
43. Canfield, R. E., The amino acid sequence of egg white lysozyme. *J. Biol. Chem.* 1963, 238, 2698-707.
44. Dormán, G.; Prestwich, G. D., Using photolabile ligands in drug discovery and development. *Trends Biotechnol.* 2000, 18 (2), 64-77.
45. Wouters, S. F. A.; Wijker, E.; Merckx, M., Optical control of antibody activity by using photocleavable bivalent peptide-DNA locks. *ChemBioChem.* 2019, 20 (19), 2463-2466.
46. Ezan, E.; Dubois, M.; Becher, F., Bioanalysis of recombinant proteins and antibodies by mass spectrometry. *Analyst* 2009, 134 (5), 825-834.
47. Chen, S.; Wan, Q.; Badu-Tawiah, A. K., Mass spectrometry for paper-based immunoassays: toward on-demand diagnosis. *J. Am. Chem. Soc.* 2016, 138 (20), 6356-6359.

Supporting Information

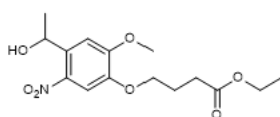
Synthesis of fulvene linkers



Ethyl 4-(4-acetyl-2-methoxyphenoxy)butanoate. Acetovanillone (9.97 g, 60 mmol) and ethyl-4-bromobutyrate (10.8 mL, 75 mmol) were dissolved in dry dimethylformamide (100 mL) while stirring and the solution was purged with N₂ for 10 min. Potassium carbonate (12.5 g, 90 mmol) was added, resulting in a light yellow suspension, which was stirred overnight at room temperature. The reaction mixture was added slowly to water (800 mL) and the precipitate collected via vacuum filtration. The precipitate was washed with water (3 x 20 mL) and dried under high vacuum to obtain a white powder (15.6 g, 93 %). ¹H NMR (500 MHz, CDCl₃): δ_H 7.56 (dd, *J* = 8.3, 2.1 Hz, 1H, aromatic), 7.54 (d, *J* = 2.0 Hz, 1H, aromatic), 6.91 (d, *J* = 8.3 Hz, 1H, aromatic), 4.18-4.13 (m, 4H, OCH₂CH₂ and COOCH₂CH₃), 3.92 (s, 3H, OCH₃), 2.57 (s, 3H, COCH₃), 2.55 (t, *J* = 7.2 Hz, 2H, CH₂COOCH₂), 2.23-2.17 (m, 2H, OCH₂CH₂CH₂), 1.27 (t, *J* = 7.1 Hz, 2H, COOCH₂CH₃).

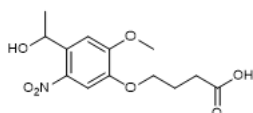


Ethyl 4-(4-acetyl-2-methoxy-5-nitrophenoxy)butanoate. Ethyl 4-(4-acetyl-2-methoxyphenoxy)butanoate (14.8 g, 53 mmol) was added in small portions to a stirred solution of 70 % nitric acid (150 mL) at 0 °C over 30 min. The reaction was stirred for 2 h at 0 °C, then quenched by gradually diluting the reaction mixture with water (800 mL) while stirring. The resulting yellow precipitate was collected via vacuum filtration and washed with water (800 mL). The crude product was dried under high vacuum, then recrystallised from hot ethanol to obtain a yellow crystalline solid (9.97 g, 56 %). ¹H NMR (500 MHz, CDCl₃): δ_H 7.62 (s, 1H, aromatic), 6.76 (s, 1H, aromatic), 4.21-4.14 (m, 4H, OCH₂CH₂ and CH₂COOCH₂CH₃), 3.97 (s, 3H, OCH₃), 2.56 (t, *J* = 7.2 Hz, 2H, CH₂COOCH₂), 2.51 (s, 3H, COCH₃), 2.21 (m, *J* = 6.8 Hz, 2H, OCH₂CH₂CH₂), 1.28 (t, *J* = 7.1 Hz, 3H, COOCH₂CH₃).

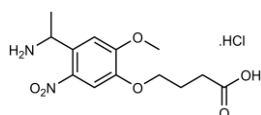


Ethyl 4-[4-(1-hydroxyethyl)-2-methoxy-5-nitrophenoxy]butanoate. Ethyl 4-(4-acetyl-2-methoxy-5-nitrophenoxy)butanoate (9.95 g, 31 mmol) was suspended in 100 % ethanol (150 mL) and cooled to 0 °C. Sodium borohydride (591 mg, 15.6 mmol) was added in small portions over 10 min while stirring. The reaction was

allowed to warm to room temperature while stirring overnight, then quenched with water (100 mL) and the ethanol evaporated under reduced pressure. The resulting precipitate was collected via vacuum filtration, washed with water (150 mL), and dried under high vacuum to afford an off-white crystalline solid (8.87 g, 87 %). ^1H NMR (500 MHz, CD_3OD): δ_{H} 7.57 (s, 1H, aromatic), 7.29 (s, 1H, aromatic), 5.60-5.53 (m, 1H, $\text{CH}(\text{OH})\text{CH}_3$), 4.20-4.08 (m, 4H, OCH_2CH_2 and $\text{CH}_2\text{COOCH}_2\text{CH}_3$), 3.98 (s, 3H, OCH_3), 2.54 (t, $J = 7.2$ Hz, 2H, $\text{CH}_2\text{COOCH}_2$), 2.26 (d, $J = 3.7$ Hz, 1H, OH), 2.18 (m, $J = 6.8$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{COOCH}_2$), 1.56 (s, 3H, $\text{CH}(\text{OH})\text{CH}_3$), 1.27 (t, $J = 7.1$ Hz, 3H, $\text{COOCH}_2\text{CH}_3$).

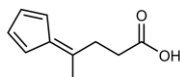


4-[4-(1-Hydroxyethyl)-2-methoxy-5-nitrophenoxy]butanoic acid. Ethyl 4-[4-(1-hydroxyethyl)-2-methoxy-5-nitrophenoxy]butanoate (8.87 g, 27 mmol) was dissolved in tetrahydrofuran (50 mL). 1 M lithium hydroxide (100 mL) was added and the reaction stirred at room temperature for 2 h. The tetrahydrofuran was removed under reduced pressure and the remaining aqueous solution adjusted to pH 3 via addition of 1 M hydrochloric acid. The resulting precipitate was collected via vacuum filtration and recrystallised from hot ethanol to obtain the product as a yellow powder (6.59 g, 82 %). ^1H NMR (500 MHz, CD_3OD): δ_{H} 7.59 (s, 1H, aromatic), 7.40 (s, 1H, aromatic), 5.46 (q, $J = 6.3$ Hz, 1H, $\text{CH}(\text{OH})\text{CH}_3$), 4.10 (t, $J = 6.2$ Hz, 2H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{COOH}$), 3.97 (s, 3H, OCH_3), 2.52 (t, $J = 7.3$ Hz, 2H, CH_2COOH), 2.10 (m, 2H, $\text{CH}_2\text{CH}_2\text{COOH}$), 1.48 (d, $J = 6.3$ Hz, 3H, $\text{CH}(\text{OH})\text{CH}_3$). ^1H NMR spectra was consistent with previous reports.¹



4-[4-(1-aminoethyl)-2-methoxy-5-nitrophenoxy]butanoic acid hydrochloride (3). 4-[4-(1-Hydroxyethyl)-2-methoxy-5-nitrophenoxy]butanoic acid (1.71 g, 5.7 mmol) was suspended in dry dichloromethane (30 mL) and cooled to 0 °C. Phosphorous tribromide (1.6 mL, 17 mmol) was added and the reaction allowed to warm to ambient temperature while stirring for 1 h. The reaction was quenched with water (30 mL) and the organic layer washed with water (2 x 30 mL) and dried over anhydrous MgSO_4 . The solvent was removed under reduced pressure and the residue redissolved in tetrahydrofuran (10 mL) and 30 % aqueous ammonia solution (50 mL) added. The solution was stirred overnight at ambient temperature, then the ammonia removed (pH 7) by bubbling N_2 gas through the solution. The remaining solution was acidified (pH 1) with 1 M hydrochloric acid and washed with ethyl acetate (3 x 20 mL). The aqueous extract was dried under a gentle stream of N_2 gas, then the crude residue resuspended in ethyl acetate (30 mL) and stirred vigorously for 10 min. The precipitate was collected via vacuum filtration and washed with ethyl acetate (30 mL) and dried under reduced pressure to obtain an off-white powder (1.03 g, 54 %). ^1H NMR (500 MHz, $(\text{CD}_3)_2\text{SO}$): δ_{H} 7.70 (s, 1H, aromatic), 7.59 (s, 1H,

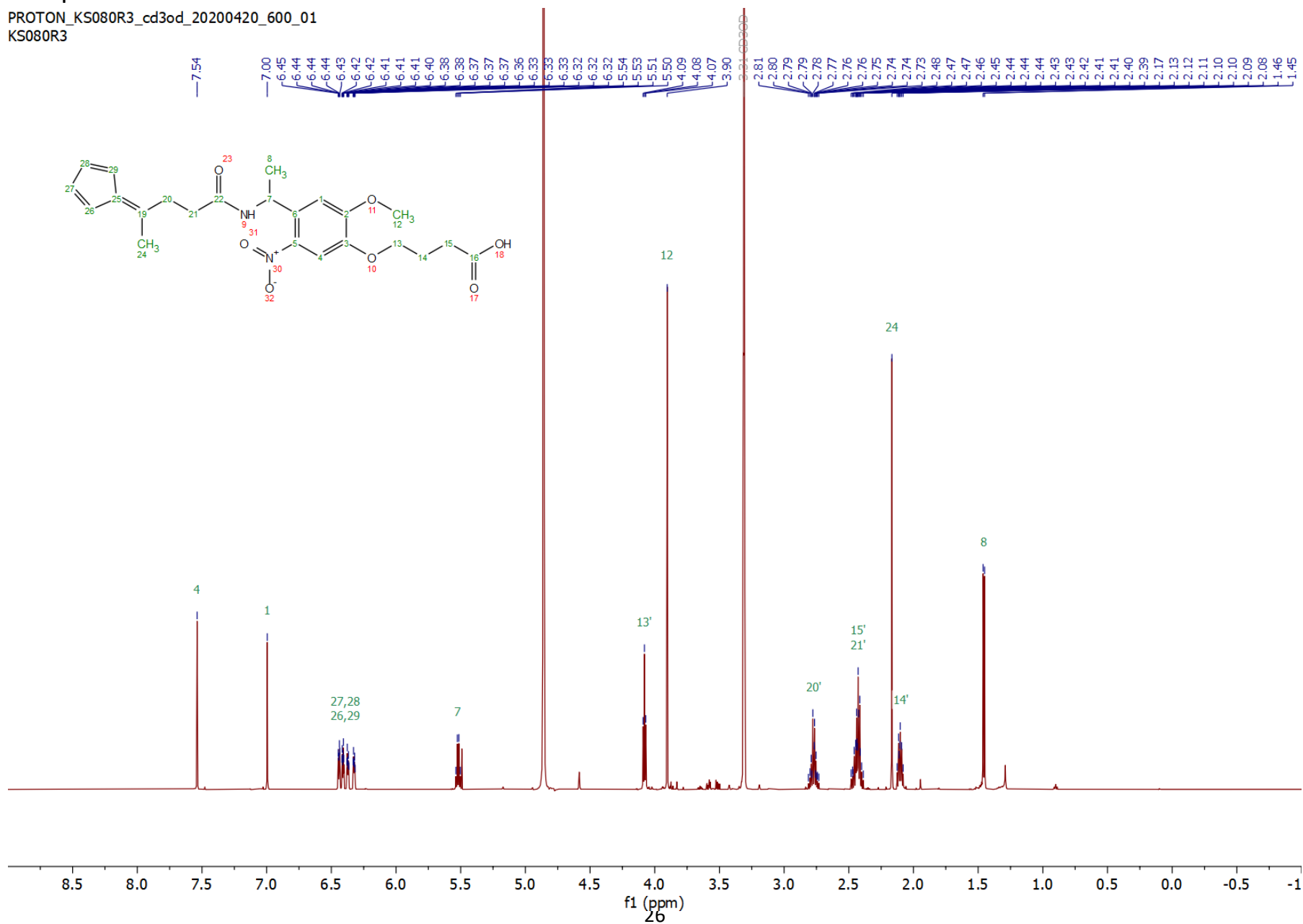
aromatic), 7.47 (s, 4H), 5.75 (s, 3H, NH₃), 4.87 (q, *J* = 6.7 Hz, 1H, CH(NH₂)CH₃), 4.09 (t, *J* = 6.5 Hz, 2H, OCH₂CH₂CH₂COOH), 3.96 (s, 3H, OCH₃), 2.38 (t, *J* = 7.3 Hz, 2H, CH₂COOH), 1.95 (m, *J* = 6.9 Hz, 2H, CH₂CH₂COOH), 1.60 (d, *J* = 6.7 Hz, 3H, CH(NH₂)CH₃). Spectra were consistent with previously reported structure.²

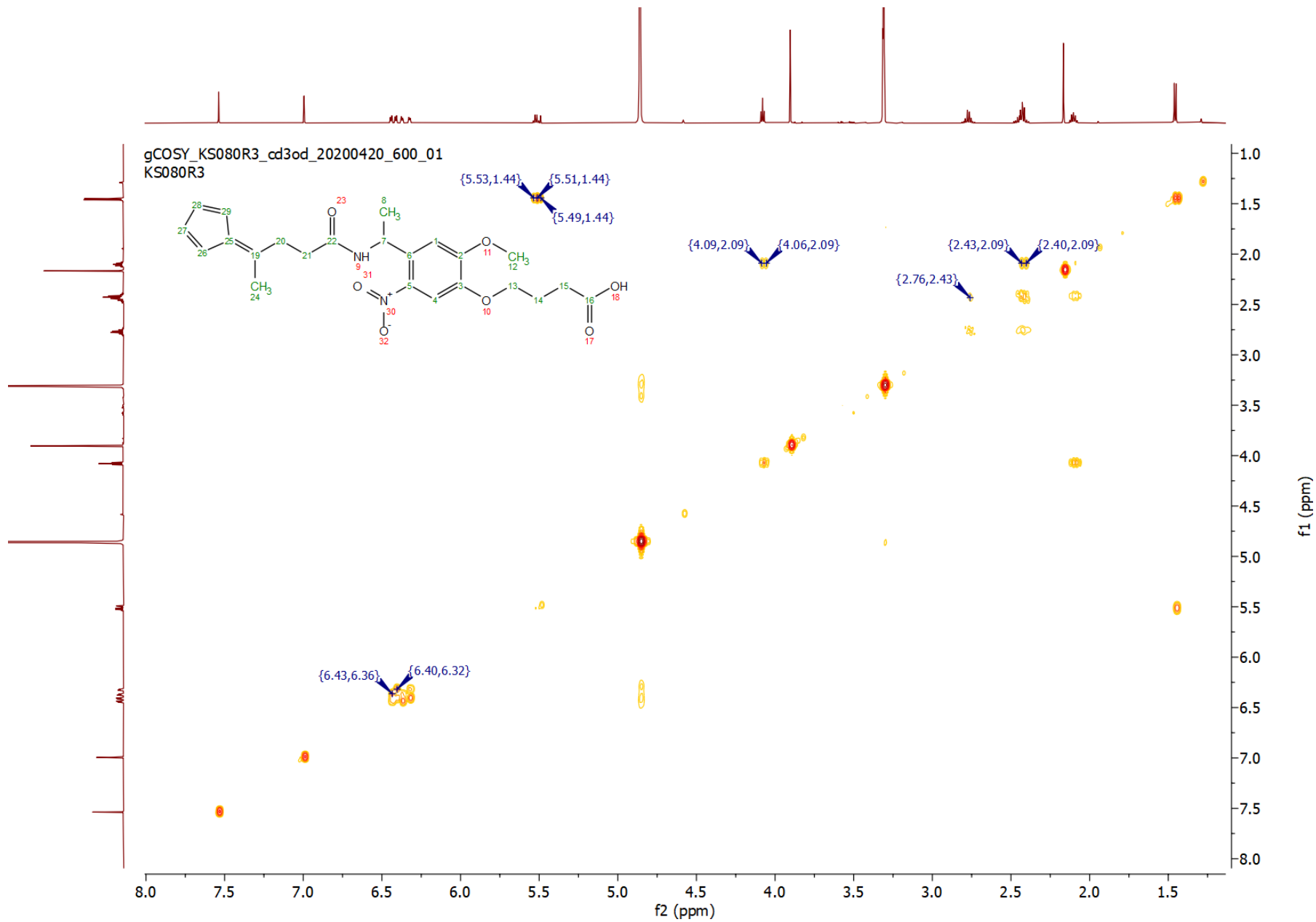


4-(2,4-Cyclopentadien-1-ylidene)pentanoic acid. To a solution of levulinic acid (4.0 g, 34 mmol) and freshly distilled cyclopentadiene (7.1 mL, 92 mmol) in MeOH (35 mL) under argon, pyrrolidine (0.60 mL, 6.9 mmol) in triethylamine (7.0 mL, 50 mmol) was added dropwise over 10 min with stirring. The solution was stirred under argon at ambient temperature in the dark for 24 h. The mixture was acidified using acetic acid (1.5 mL) and the methanol was removed in vacuo. Diethyl ether (100 mL) and water (100 mL) were added to the residue and the organic layer was removed and washed with water (5 x 50 mL) and saline (3 x 50 mL), dried over anhydrous MgSO₄ and filtered. The filtrate was concentrated in vacuo and the resulting residue was recrystallised from hexane to afford yellow crystals, 5.35 g (94.5 %). ¹H NMR (500 MHz, (CD₃)₂SO): δ_H 1.91 (s, 3H, CH₃), 2.44 (t, *J* = 2.4 Hz, 2H, CH₂CO), 2.75 (t, *J* = 2.4 Hz, 2H, CH₂CC), 6.43-6.39 (m, 2H, 2CH), 6.53-6.08 (m, 2H, 2CH). ¹³C NMR (125 MHz, (CD₃)₂SO): δ_C 173.1 (1C, COOH), 152.0 (1C, CH(CC)CH), 142.5 (1C, CH₃(CC)CH₂), 130.4 (2C, CH(CC)CH), 120.6 (2C, CHCHCHCH), 33.1 (1C, CH₂COOH), 31.3 (1C, CH₂CC), 20.2 (1C, CH₃).

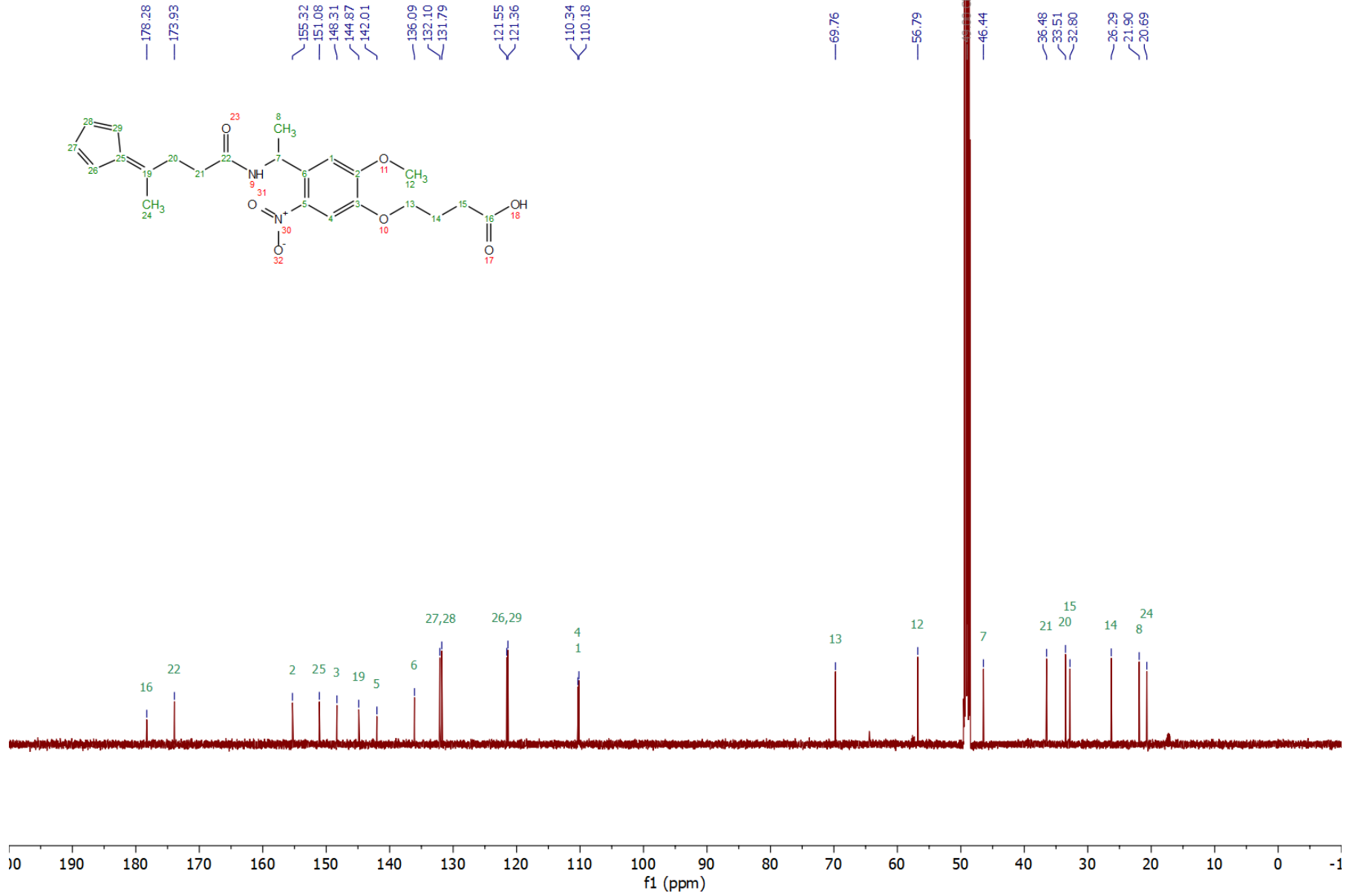
NMR spectra

PROTON_KS080R3_cd3od_20200420_600_01
KS080R3

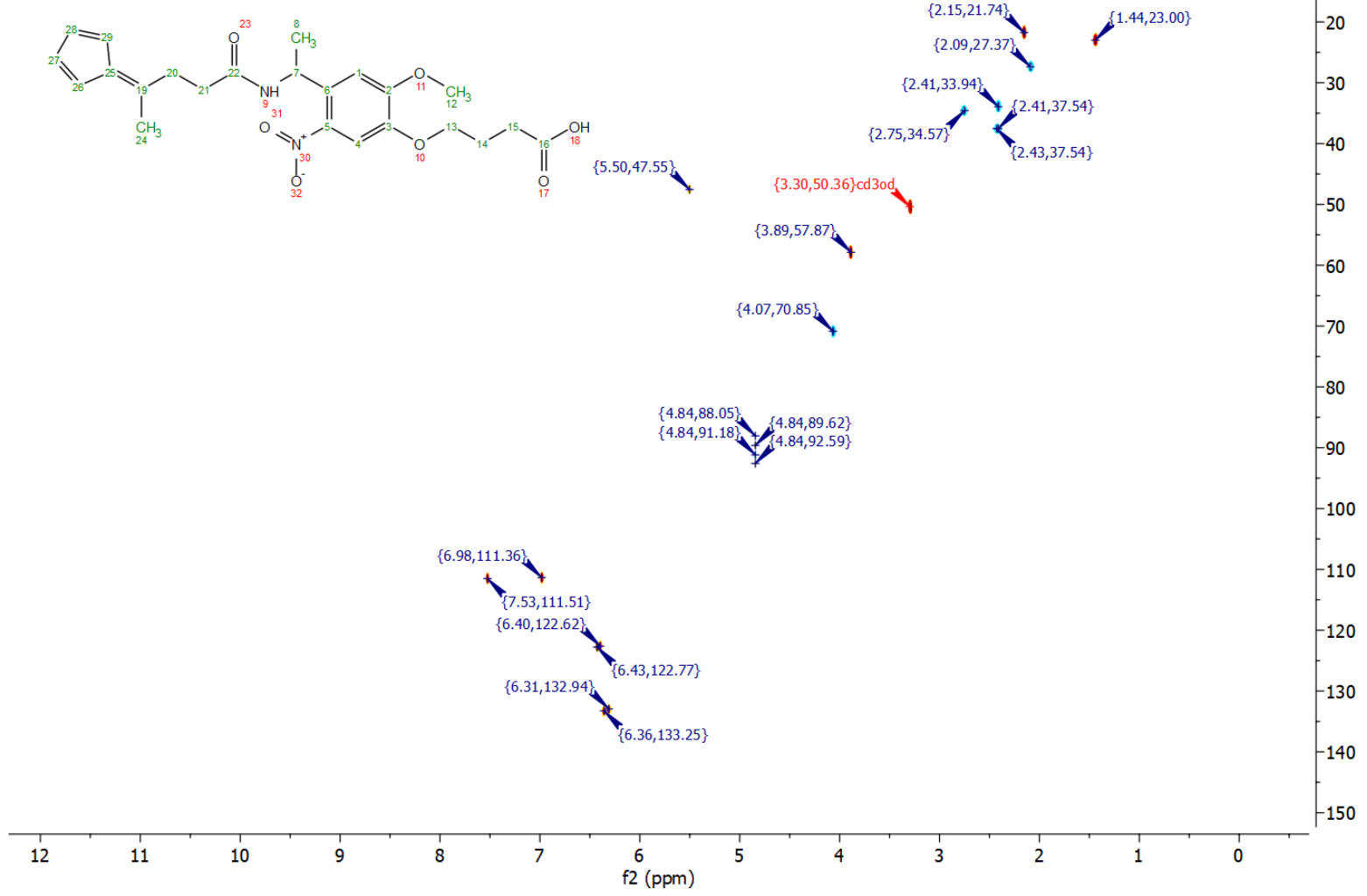




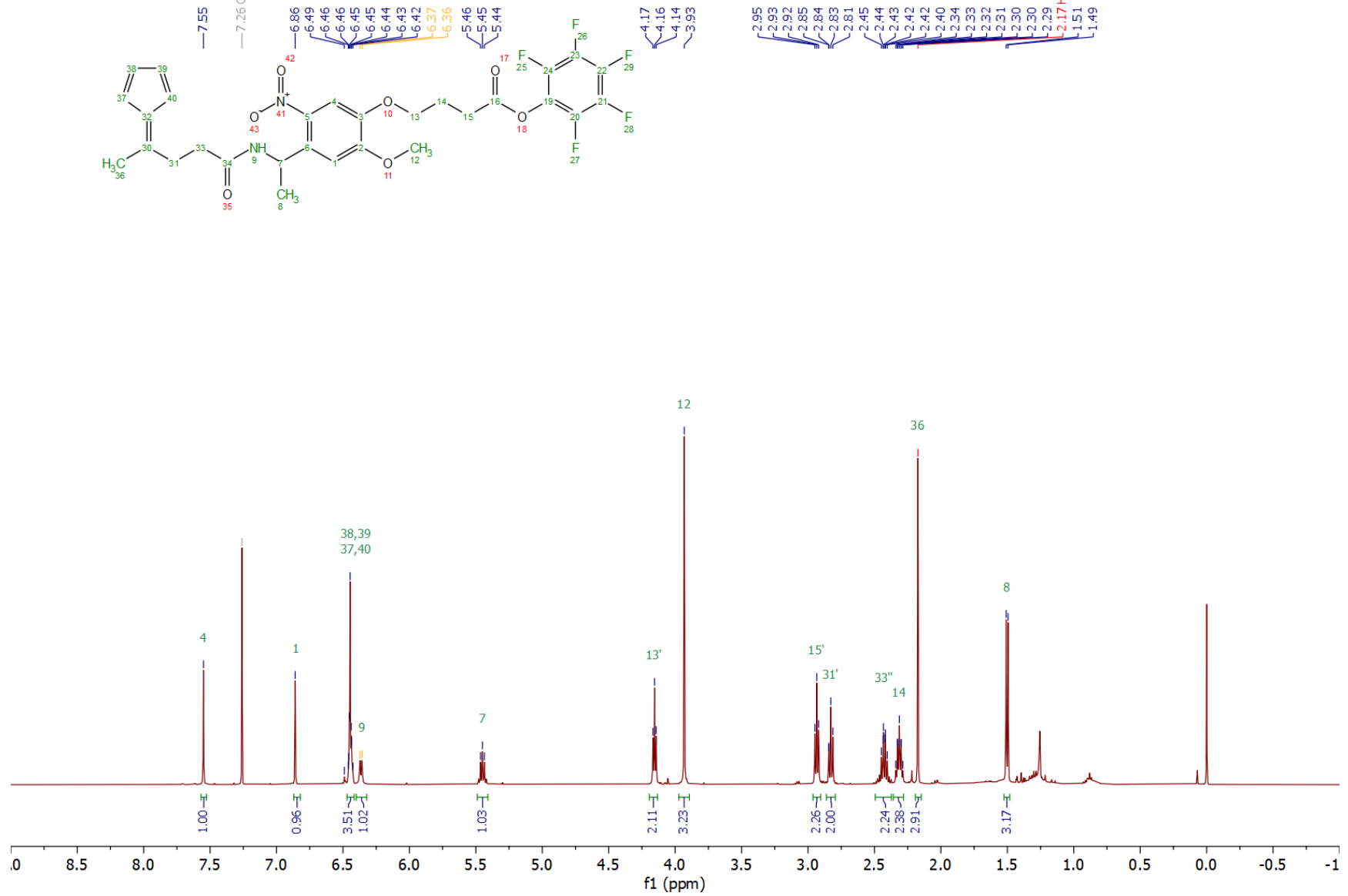
CARBON_KS080R3_cd3od_20200420_600_01
KS080R3

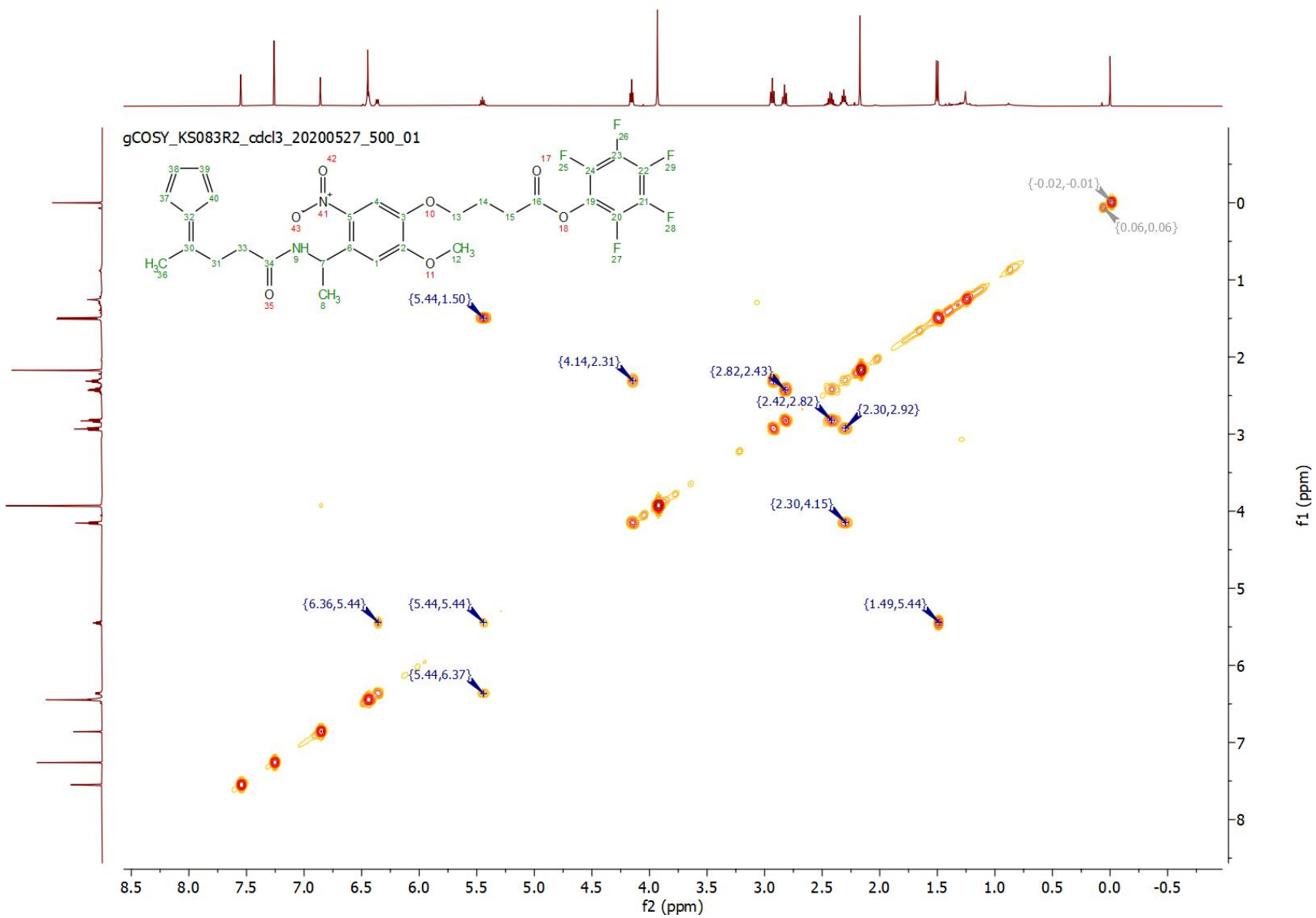


HSQCAD_KS080R3_cd3od_20200420_600_01
KS080R3

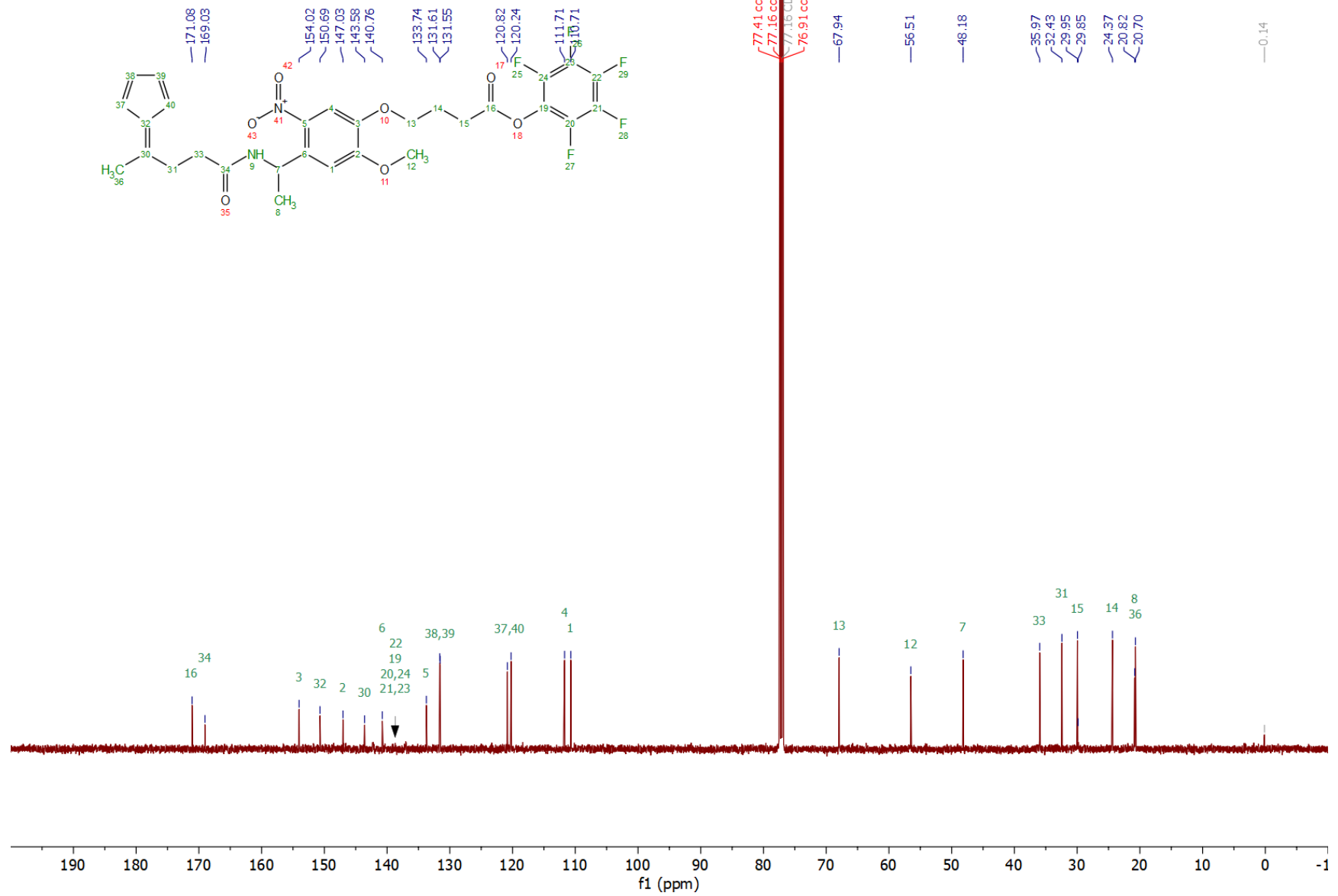


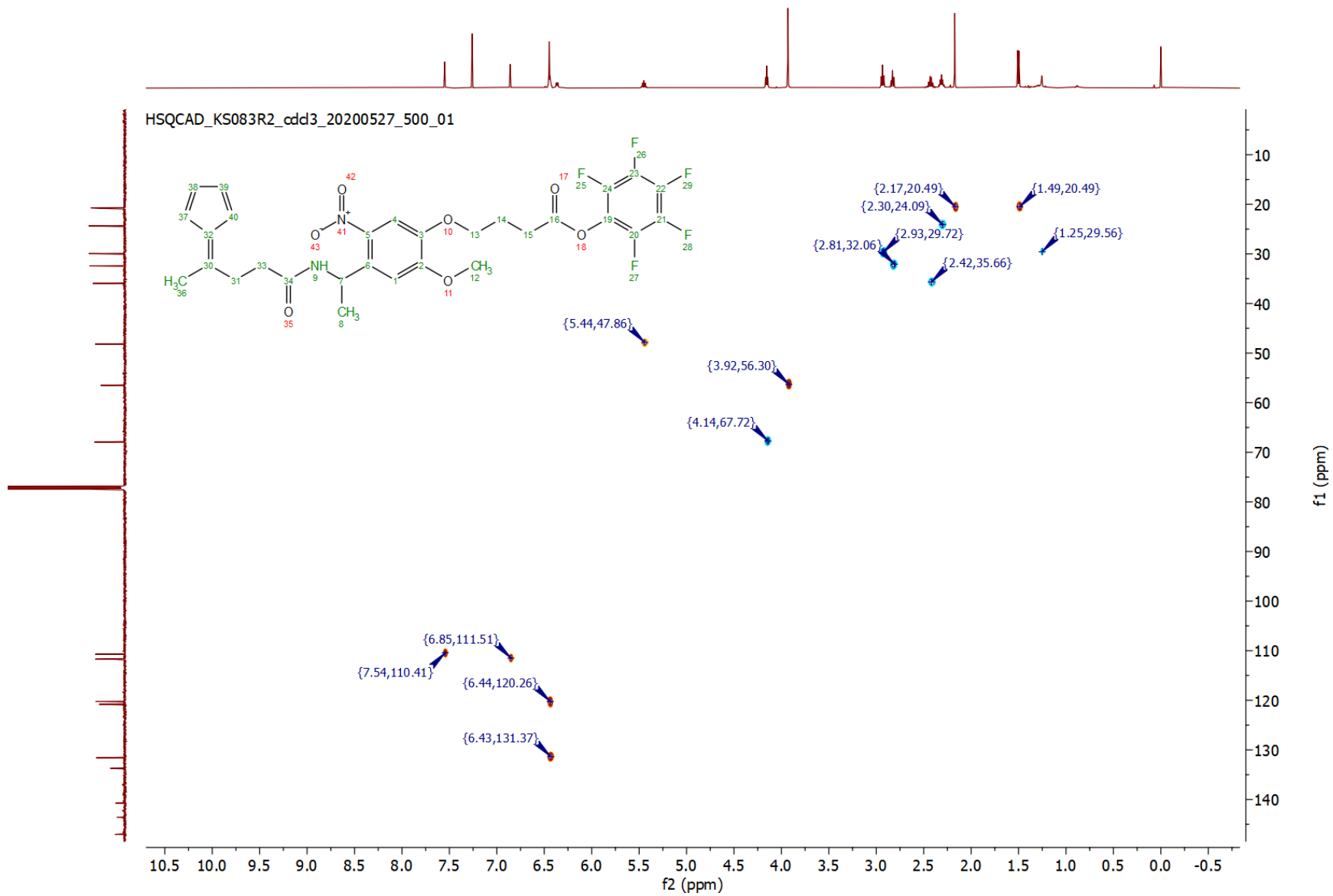
PROTON_KS083R2_cdd3_20200527_500_02



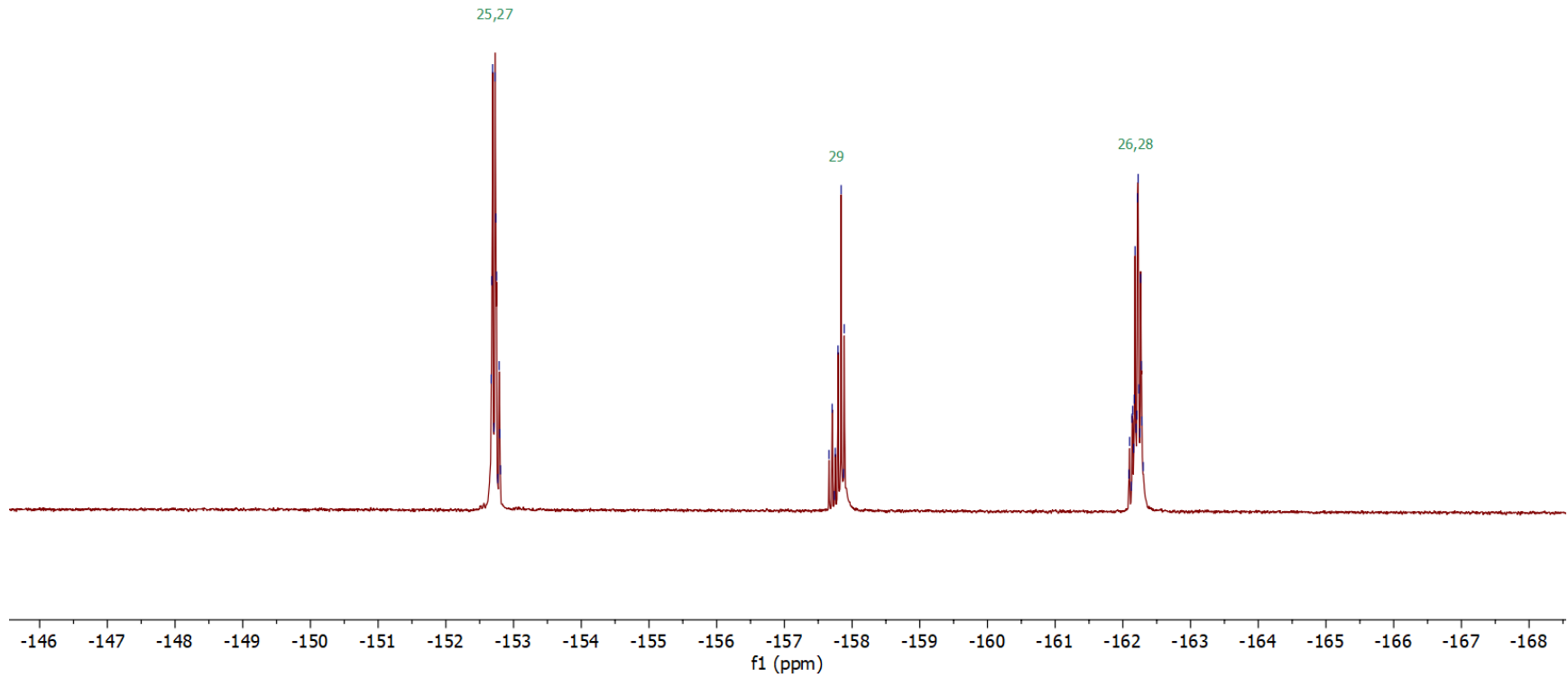
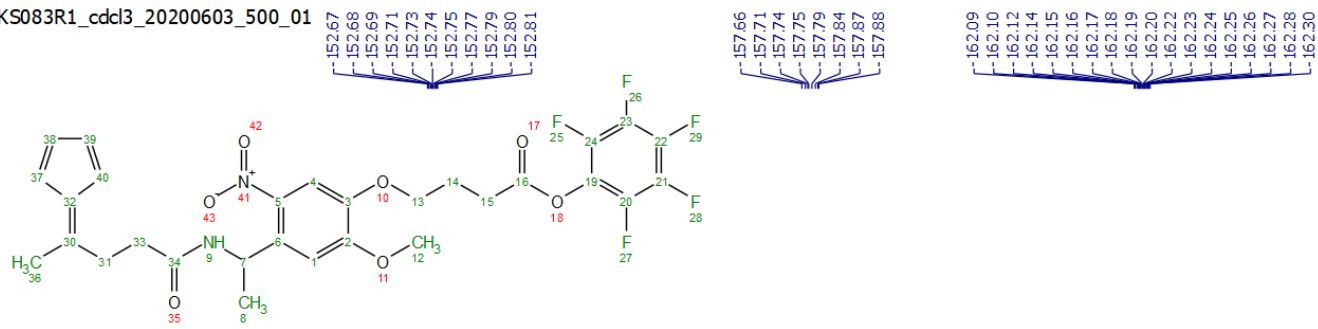


CARBON_KS083R2_cdcl3_20200527_500_01

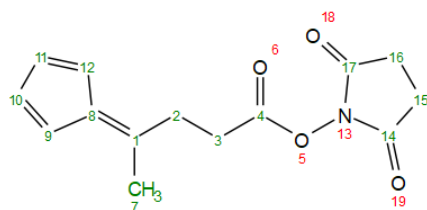




FLUORINE_KS083R1_cdc13_20200603_500_01



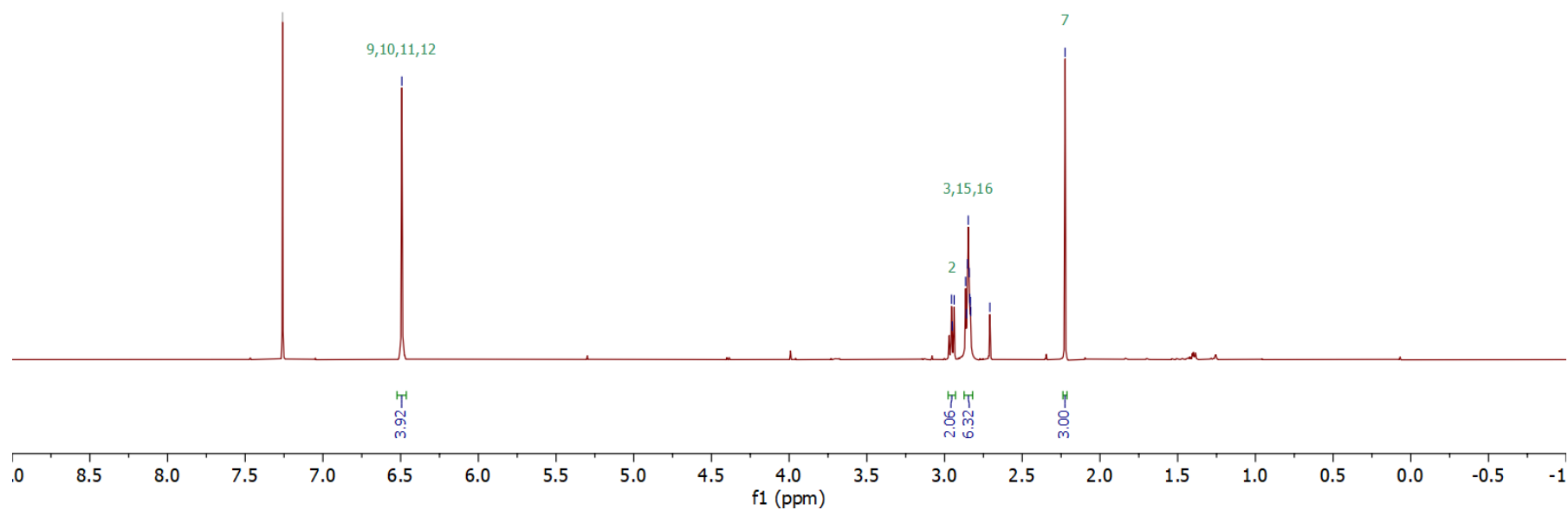
PROTON_KS078_cdcl3_20191210_500_01



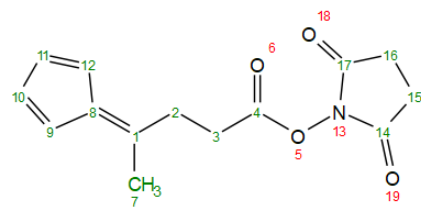
7.26 CDCl₃

6.49

2.96
2.95
2.94
2.86
2.86
2.85
2.85
2.84
2.84
2.83
2.83
2.71
2.22



CARBON_KS078_cdcl3_20200915_500_01



169.44
168.10

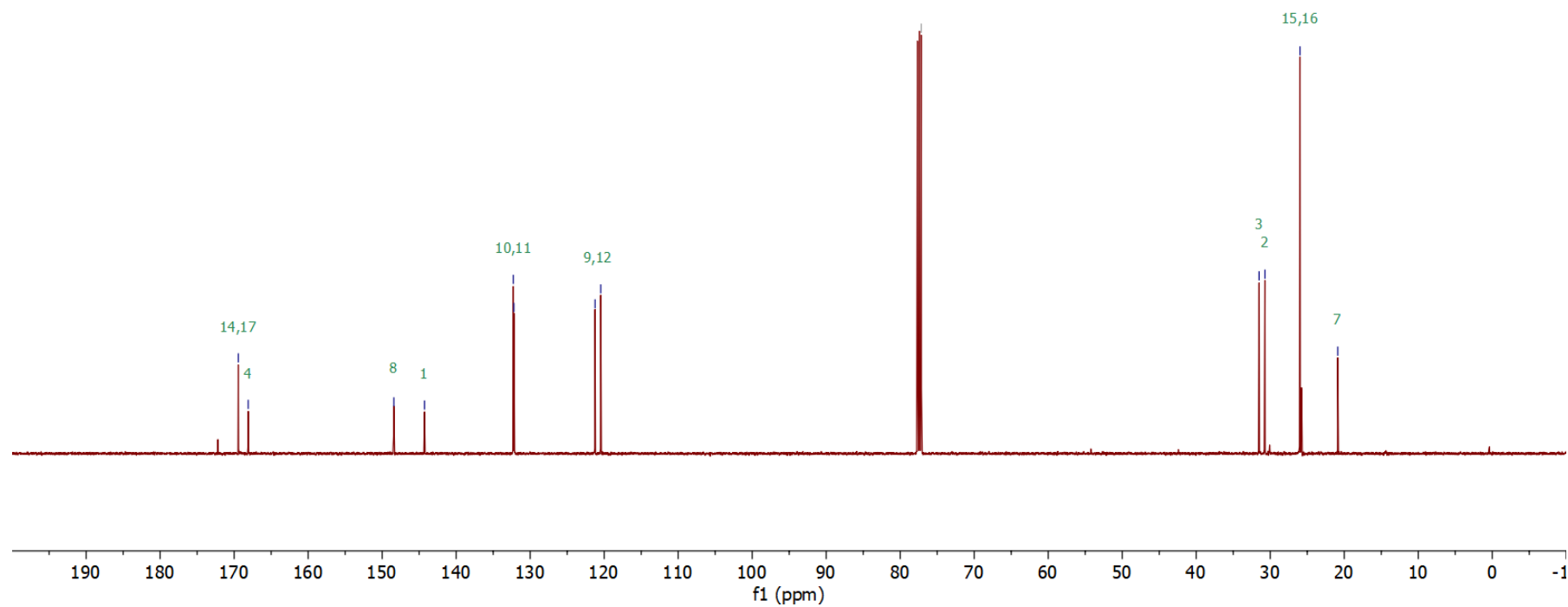
148.42
144.28

132.28
132.21

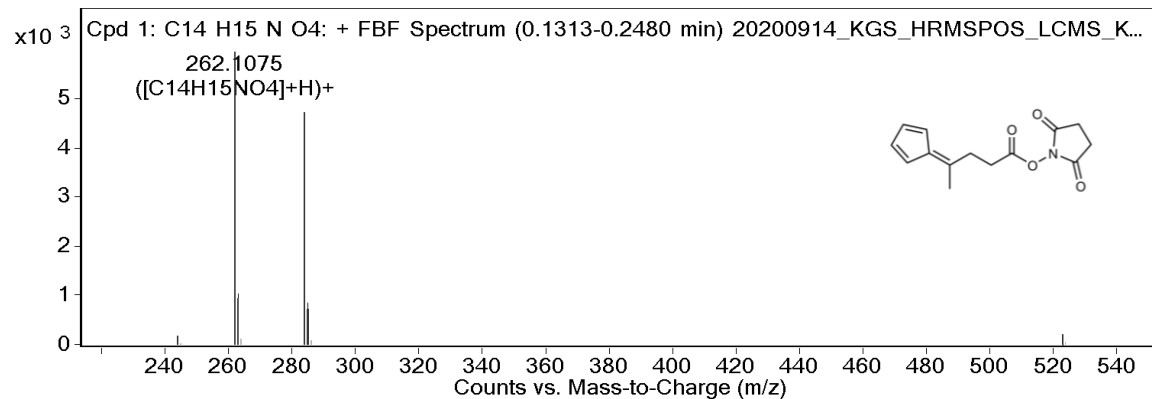
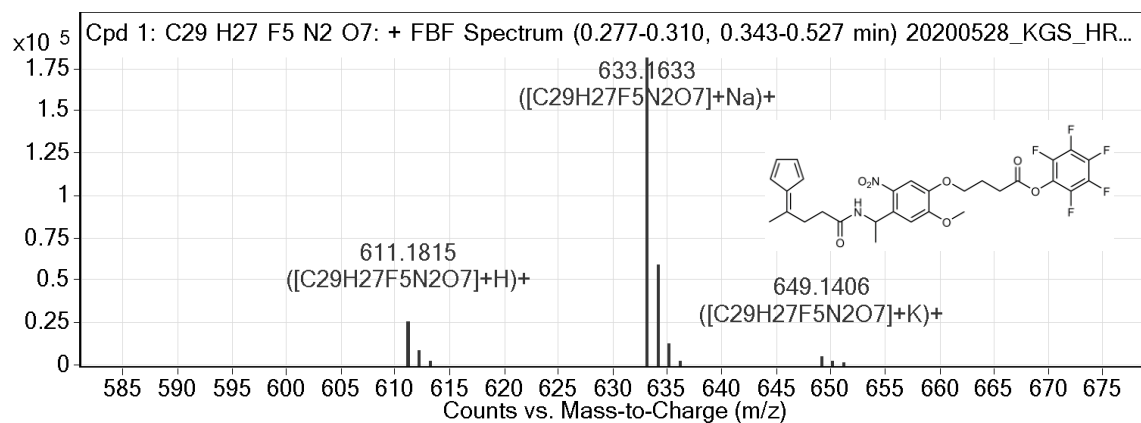
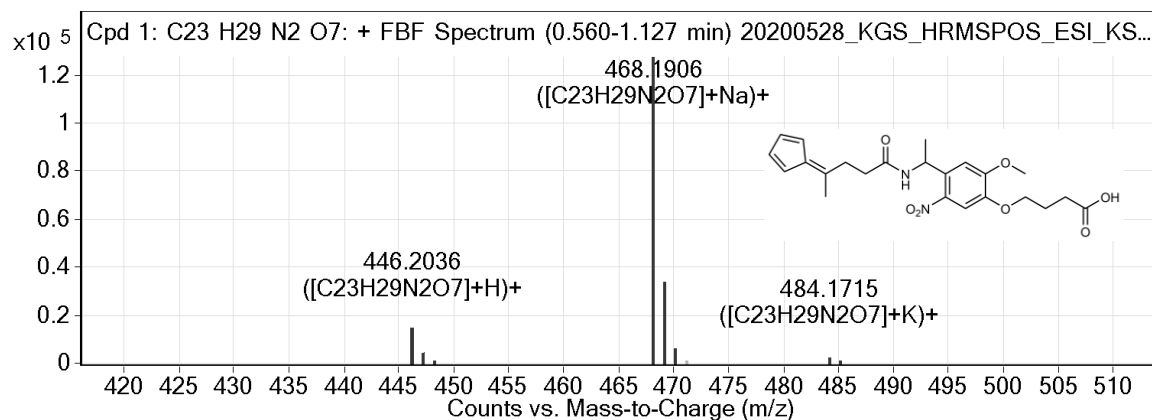
121.23
120.46

77.16 CDCl₃

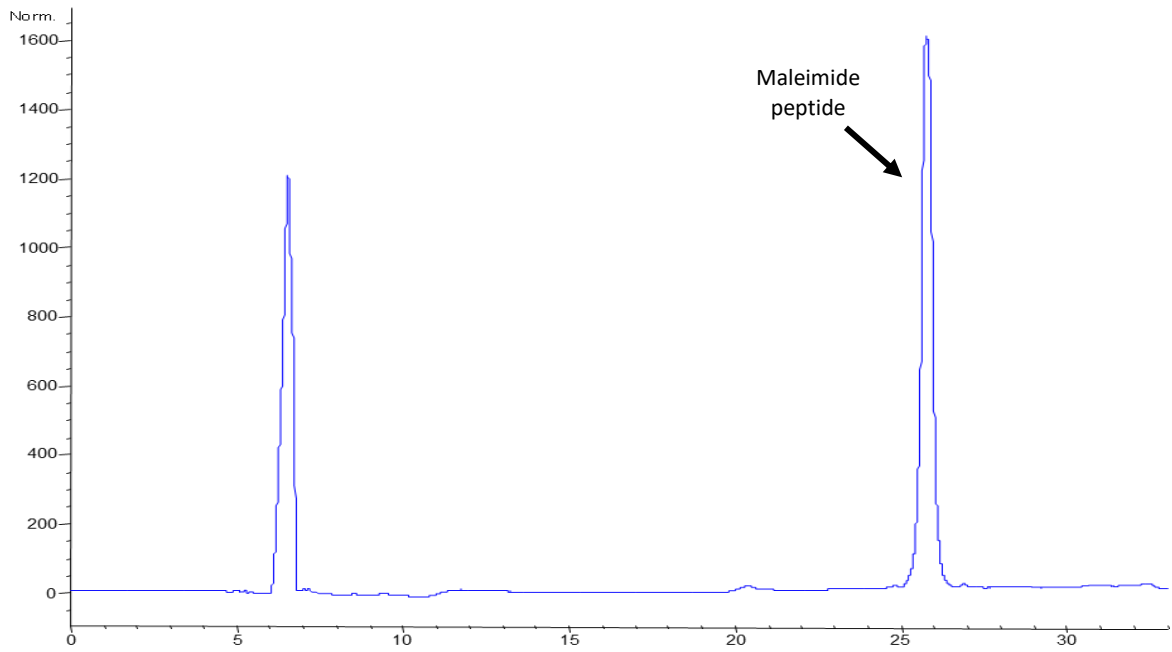
31.51
30.72
25.97
20.88



HRMS



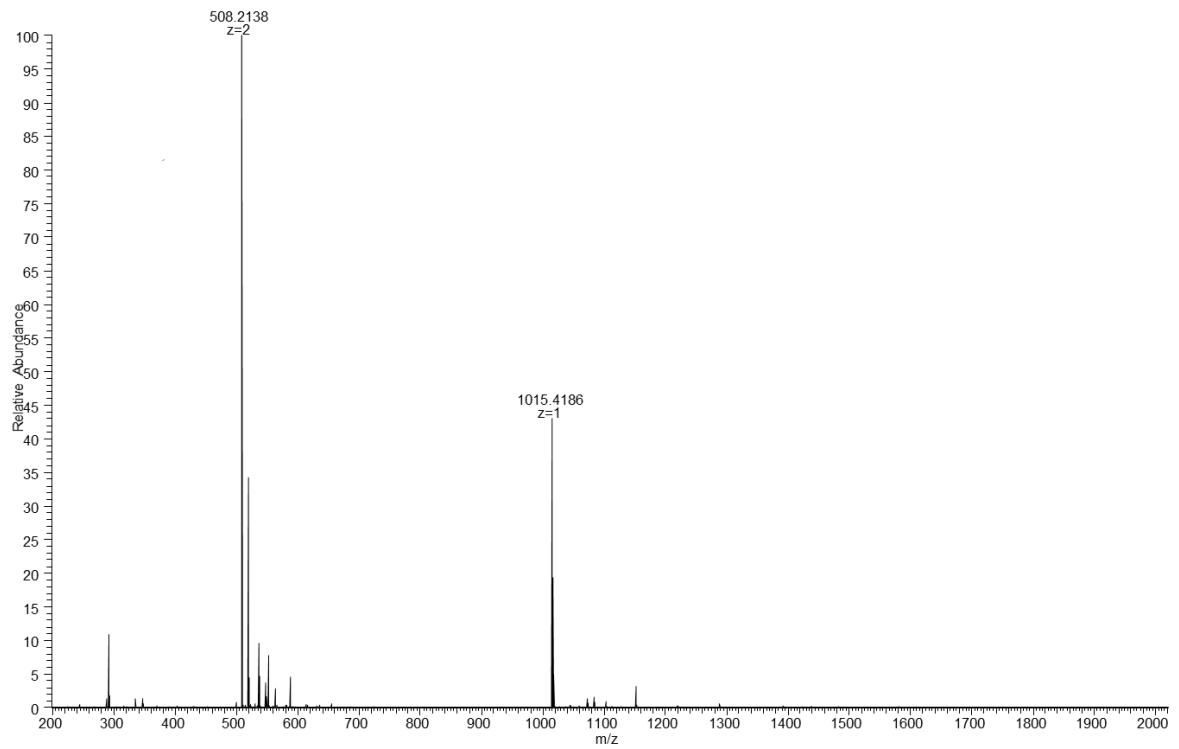
Peptide HPLC chromatogram

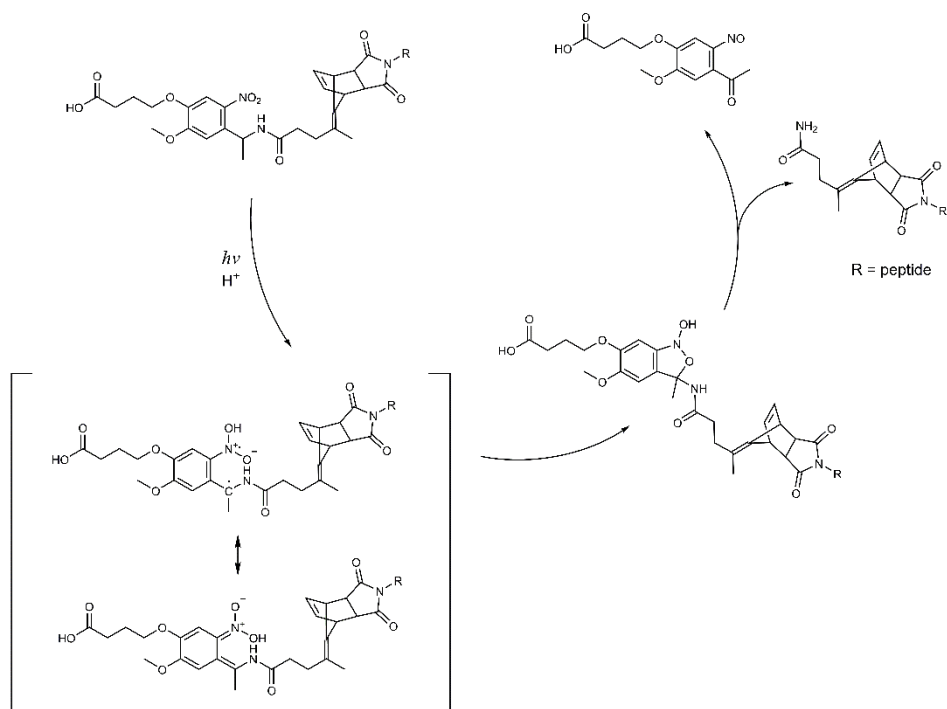


Peptide ESI-MS

Maleimide-beta-Ala-GSGSGRGS-G-NH₂

Calculated M+H: 1015.3, Found M+H: 1015.4





Scheme S1.3

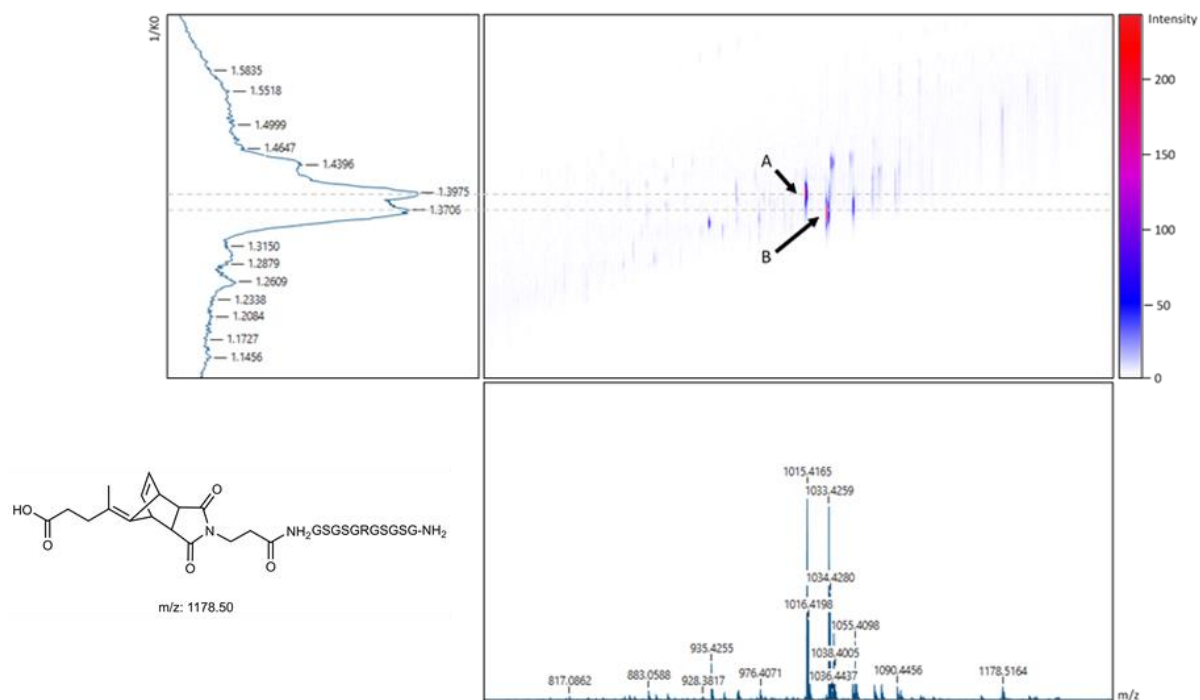


Figure S1. Trapped ion mobility-matrix-assisted laser desorption/ionization mass spectra of photocleaved cycloaddition product (structure inset, m/z 1178.50) showing prominent peaks for unmodified (A) and hydrolyzed (B) maleimide-labelled peptide.

References

1. Klinger, D.; Landfester, K., Enzymatic- and light-degradable hybrid nanogels: Crosslinking of polyacrylamide with acrylate-functionalized Dextrans containing photocleavable linkers. *Journal of Polymer Science Part A: Polymer Chemistry* **2012**, *50* (6), 1062-1075.
2. Griffin, D. R.; Schlosser, J. L.; Lam, S. F.; Nguyen, T. H.; Maynard, H. D.; Kasko, A. M., Synthesis of photodegradable macromers for conjugation and release of bioactive molecules. *Biomacromolecules* **2013**, *14* (4), 1199-1207.
3. Gaplovsky, M.; Il'ichev, Y. V.; Kamdzhilov, Y.; Kombarova, S. V.; Mac, M.; Schwörer, M. A.; Wirz, J., Photochemical reaction mechanisms of 2-nitrobenzyl compounds: 2-Nitrobenzyl alcohols form 2-nitroso hydrates by dual proton transfer. *Photochem. Photobiol. Sci.* **2005**, *4* (1), 33-42.

Application of bottom-up proteomics for identifying sites of oxidative modification and chemical labelling in equine heart myoglobin

4.1 Introduction

4.1.1 *Proteoforms: missing links between the genome and the proteome*

Mass spectrometry (MS)'s progress towards soft ionisation methods and their integration with separation techniques have been driving factors in its coup against conventional techniques for biomolecule analysis.¹ One of the most powerful applications of this technology has been the development of workflows for identifying, sequencing and quantifying proteins, collectively constituting the broad field of modern *proteomics*. Close inspection of the human genome implies the existence of approximately 20 000 human proteins based on direct interpretation of amino acid sequences. However, the actual estimated numbers of proteins that make up the human proteome vary from 10 000 to 1 billion.²⁻³ The large discrepancy between these numbers stems from the vast number of different modifications a protein can undergo between its initial biosynthesis and eventual analysis.² Importantly, many of these alternative species, or *proteoforms*, are thought to have important roles in both the regulation of biological function and pathogenesis of a significant number of diseases.³⁻⁴

4.1.2 *Oxidative modifications to proteins as disease markers*

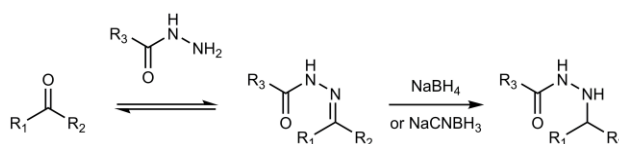
While sample preparation, data acquisition and data analysis technology has conceived sophisticated MS workflows that address the current focus on *in vivo* post-translational modifications (PTMs), there is a growing interest in applying and refining this methodology for the analysis of non-enzymatic chemical modifications to proteins.⁴ Chemical modifications to amino acids can occur *in vivo* and reflect changes in cellular environments, or during sample preparation, either intentionally by specific reagents or unintentionally, as 'artifacts'.⁵ Oxidative modifications are an important class of chemical modifications, which result from the (often non-reversible) reactions of reactive oxygen species (ROS), including superoxide anion radicals, hydrogen peroxide and hydroxyl radicals, with proteins, lipids, sugars and nucleic acids. Chronic high levels of oxidative stress are involved in many diseases, including diabetes mellitus, neurodegenerative diseases, inflammatory diseases, atherosclerosis and cancer.⁶⁻⁸

Of the many ROS products, carbonylated proteins represent one of the most diverse categories and are commonly associated with metal-catalysed oxidation (MCO), a process in which reduced metal ions, such as Fe²⁺, react with hydrogen peroxide (H₂O₂) to form hydroxyl radicals. The addition of carbonyl groups to amino acids can then occur via various mechanisms, including direct oxidation of amino acid side chains or reaction between amino acids with the carbonylated oxidation products of lipids, amino acids, reducing sugars and other

metabolites.^{6,9-11} These carbonylated residues can further react with other amino acids to form protein-protein crosslinks, for example in the pathologically significant generation of advanced glycation end-products (AGEs).⁹ With over 35 possible ways in which proteins can be oxidised, compounded by the number of modifiable residues in a given protein, the complexity of these species creates a significant analytical challenge.^{6, 12}

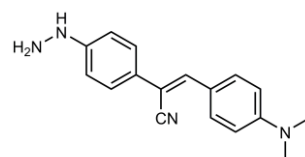
4.1.3 Bioconjugation methods for targeting carbonylated proteins

Analytical strategies for detecting and quantifying carbonylated proteins frequently involve their derivatisation using hydrazine-like reagents, which react with aldehydes and ketones to form hydrazones, a form of stable Schiff base (**Scheme 1**). This reaction is reversible, but the products can be stabilised by reduction with borohydride reagents, such as sodium borohydride (NaBH₄) or sodium cyanoborohydride (NaCNBH₃).¹³⁻¹⁵



Scheme 4.1.1. Conjugation of hydrazide reagents to carbonyls, followed by reduction using sodium borohydride or sodium cyanoborohydride.¹⁴⁻¹⁵

One of the oldest derivatising agents is dinitrophenylhydrazine (DNPH), which is commonly used to visualise carbonylated proteins following their separation using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).^{6, 16} Carbonylated proteins can be visualised and the degree of labelling and therefore carbonylation can be determined spectrophotometrically by measuring absorption at 360 nm. DNPH-specific antibodies can also be used for immunoblot detection, immunoprecipitation, immunofluorescent imaging and enzyme-linked immunosorbent assays of derivatised proteins.¹⁶⁻¹⁷ Despite its widespread use, DNPH can react non-specifically with sulfenic acids and is unreliable for measuring heme-containing proteins, which are known to absorb at similar wavelengths.^{12-13, 16, 18} As such, there has been substantial investment in developing alternative detection reagents, usually in the form of biotin or fluorescent hydrazides.^{16, 19} Conjugation to biotin enables selective enrichment of carbonylated proteins and peptides, and is therefore extremely useful for both gel electrophoretic and MS analyses, whereas fluorescent labels are valuable for applications including live cell imaging and viability assays.^{6, 16, 19} We have recently investigated the conjugation of a novel hydrazine-containing α -cyanostilbene fluorophore, DPANzine (**Figure 4.1.1**). Previous unpublished work led by collaborators at La Trobe University (Yuning Hong, Siyang Ding) has confirmed that this fluorophore effectively labels MCO-treated proteins, including bovine serum albumin, cytochrome c and myoglobin, although the site-specificity and labelling efficiency is yet to be determined.



DPANzine

Chemical Formula: C₁₇H₁₈N₄
Molecular Weight: 278.36

Figure 4.1.1. The novel hydrazine-containing α -cyanostilbene fluorophore, DPANzine, is designed to react with carbonylated biomolecules.

4.1.4 Proteomics methods for analysing carbonylated proteins

Obtaining more specific information regarding the structure and location of modified residues requires high-resolution MS analysis, most commonly in the form of *bottom-up* proteomics, a term describing analytical workflows in which proteins are chemically or enzymatically processed into shorter peptides prior to analysis (Figure 1).^{6, 12, 20} MS analysis of these peptides can then be performed using various combinations of ionisation methods and mass analysers.¹² There are comparably fewer examples describing *top-down* proteomics of carbonylated proteins, which instead involves MS analysis of intact proteins followed by gas-phase fragmentation of peptide bonds, possibly as a result of the difficulty in obtaining sufficient ionisation, fragmentation and sensitivity from complex and heterogeneous protein mixtures.^{4, 12, 21-22}

Bottom-up proteomics of carbonylated proteins are often preceded by a purification or enrichment step, for example by streptavidin affinity chromatography after conjugation to biotin hydrazide or two-dimensional gel electrophoresis.²³ Derivatised or unmodified peptides can then be analysed via tandem MS (MS/MS) to identify which amino acid residues are modified based on the differences between observed and theoretical m/z values (*mass shifts*).^{12, 23} Chemical labelling with carbonyl-reactive fluorophores can also be assessed using this bottom-up proteomics, which could therefore be a useful means of studying the interactions between novel reagents, such as DPANzine, and oxidised proteins to better understand the results of cell-based assays and imaging experiments.^{12, 24}

4.2 Aims

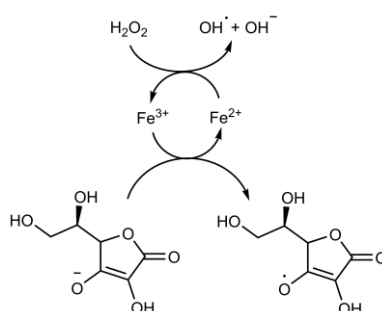
4.2.1 Establish a bottom-up proteomics approach for assessing amino acid carbonylation and chemical labelling in an in vitro model for oxidative stress

4.2.2 Determine the amino acid specificity of labelling by the novel hydrazine-containing fluorophore, DPANzine, for equine heart myoglobin

4.3 Results and Discussion

4.3.1 In vitro model for protein carbonylation

To investigate the site specificity of carbonylation and DPANzine labelling on proteins, we chose a model system using FeCl_3 /ascorbic acid and the model heme protein, myoglobin. This system is designed to mimic the *in vivo* reaction that occurs under physiological conditions between endogenous iron and ascorbate monoanion, in which Fe^{3+} is reduced to Fe^{2+} , which catalyses the decomposition of H_2O_2 to highly reactive hydroxide radicals (the 'Fenton reaction', **Scheme 4.3.1**).²⁵



Scheme 4.3.1. Metal-catalysed oxidation under physiological conditions occurs through reduction of Fe^{3+} to Fe^{2+} by ascorbic acid. Fe^{2+} catalyses the decomposition of hydrogen peroxide to hydroxyl radicals.²⁵

Myoglobin is a popular choice for studying protein oxidation and carbonylation, owing to its well-characterised structure and important roles in oxygen transport and nitric oxide scavenging in skeletal and heart muscle tissue. These functions are enabled by an iron-bound heme prosthetic group, which can also catalyse the Fenton reaction and lead to the generation of intracellular ROS.²⁶⁻²⁸ Interaction of these ROS with other biomolecules, such as lipids, can lead to aberrant signalling pathways and subsequently various complications, including vasoconstriction, ischemia and acidosis.²⁷ These processes can also lead to oxidative modifications to myoglobin itself, although a large portion of the existing literature related to human pathology is focused on the functional and structural effects related to glycation by sugars, rather than amino acid carbonylation.²⁸⁻³⁰ Furthermore, increased levels of circulating and excreted myoglobin are used as clinical biomarkers for early acute myocardial infarction or acute renal failure, as these changes can be indicative of injury to muscle tissue, in addition to contributing to the pathological progression of these diseases.³¹⁻³³ Investigating the direct amino acid carbonylation products of oxidised myoglobin and establishing effective labelling strategies will potentially provide valuable insights into the role of oxidative stress in cardiac and skeletal myopathies and inform the design of novel detection strategies.

4.3.2 Proteomic assessment of MCO-induced myoglobin carbonylation

To confirm the efficacy of our MCO model, it was necessary to first evaluate the extent of carbonylation of myoglobin that had been exposed to the sodium ascorbate/ FeCl_3 oxidation conditions. Efficient protease

activity is particularly important for heavily modified proteins, for which modified residues at the cleavage site (e.g. lysine and arginine residues for trypsin) can inhibit protease activity, thereby resulting in increased missed cleavages and reduced sequence coverage.³⁴ Alternative proteases, such as chymotrypsin, Asp-N and Glu-C, are suitable alternatives to trypsin; however, we opted for a well-established commercial trypsin/Lys-C protocol, which cleaves peptide bonds at the same sites as trypsin alone, albeit with fewer missed cleavages due to the higher tolerance of Lys-C to strong denaturing conditions.³⁵⁻³⁷ Retaining C-terminal cleavage at basic arginine and lysine residues is also advantageous in that it generally yields more efficient peptide ionisation and predictable fragmentation.³⁸

Independent proteomic analyses of the oxidised myoglobin digest confirmed that this protease combination was effective and provided 79 % coverage of equine heart myoglobin and identified 17 unique peptides, for which 12 sites were identified as potential carbonylated amino acids (**Figure 4.3.1**).

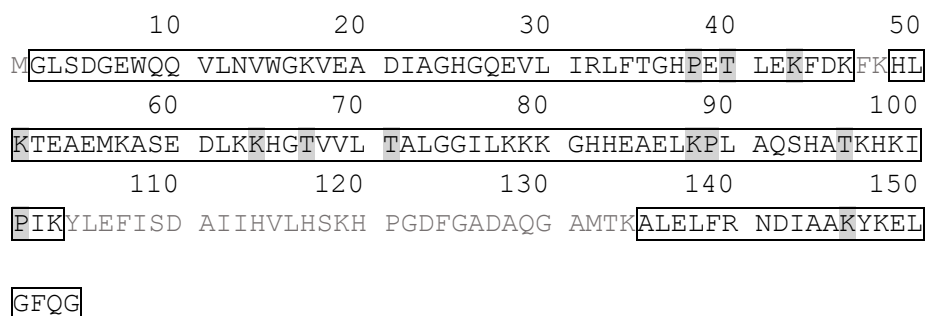


Figure 4.3.1. Sequence coverage (framed portions) and carbonylation sites (shaded) for oxidised equine myoglobin (<https://www.uniprot.org/uniprot/P68082>).

Of the many amino acids that are sensitive to MCO, only lysine, arginine, proline and threonine are known to produce reactive carbonyls.¹¹ These modifications were the basis used for a Sequest HT search of the oxidised myoglobin MS/MS data for carbonylated amino acids, which could potentially be modified by DPANzine (Table 4.3.1). The present analysis of oxidised myoglobin found potential carbonylation products to lysine, proline and threonine residues (**Table 4.3.2**), although it is important to note that differentiation between proline oxidation to the hydrazine-reactive carbonyl glutamic semialdehyde and unreactive hydroxyproline is not possible based on mass shifts alone (both +15.995 Da).¹² Interestingly, carbonylation of these residues was also observed for the untreated control protein, suggesting that similar modifications also occur *in vivo*, during isolation and/or purification or as artifacts of the digestion and analysis workflows.

Mapping of carbonylation sites only observed in the oxidised samples to the X-ray crystallographic structure for myoglobin shows these modifications clustered around the heme prosthetic group, which suggests that it may be involved in the generation of ROS required for the oxidation of these residues (**Figure 4.3.2**). Similar concentration of carbonylated lysine residues around heme groups has previously been observed during top-down proteomic analysis of the electron transport protein, cytochrome *c* after treatment with the mild

oxidising agent, chloramine-T.^{22, 39} In contrast, carbonylation products formed in the absence of ascorbic acid and FeCl₃ did not appear to be localised to these domains, suggesting a different mechanism may be responsible for these modifications.

Table 4.3.1. Metal-catalysed oxidative modifications resulting in amino acid carbonylation used for Sequest HT database search of digested myoglobin.¹¹

Amino acid residue	Modification	Exact Δ mass (Da)	Average Δ mass (Da)
M	Oxidation	+15.9949	+15.9994
N,Q	Deamidated	+0.9840	+0.9848
P	Glutamic semialdehyde	+15.9949	+15.9994
R	Glutamic semialdehyde	-43.0534	-43.0711
K	Amino adipic semialdehyde (allysine)	-1.0316	-1.0311
T	2-amino-3-ketobutyric acid	-2.0156	-2.0157

Table 4.3.2. Amino acid carbonylation products identified for myoglobin.

Treatment	Modification Name	Amino Acid	Position in Peptide	Sequence	Modified PSMs	Unmodified PSMs	Position
No oxidation	Thr->AminoKetobutyricAcid	T	14	HPGDFGADAQGAMTK	14	53	133
	Pro->GlutamicSemialdehyde*	P	2	HPGDFGADAQGAMTK	1	66	121
	Lys->Allysine	K	15	HPGDFGADAQGAMTK	2	65	134
MCO	Thr->AminoKetobutyricAcid	T	7	HGTVVLTALGGILK	5	99	71
			3	HGTVVLTALGGILKK	4	43	67
			17	KGHHEAELKPLAQSHATK	1	66	96
			8	LFTGHPETLEK	1	13	40
	Pro->GlutamicSemialdehyde*	P	4	HKIPK	1	9	101
			10	KGHHEAELKPLAQSHATK	2	65	89
			6	LFTGHPETLEK	7	7	38
	Lys->Allysine	K	12	ALELFRNDIAAKYK	1	3	146
			3	HLKTEAEMK	2	3	51
			9	KGHHEAELKPLAQSHATK	1	66	88
			1	KHGTVVLTALGGILKK	1	6	64
11			LFTGHPETLEKFDK	1	1	43	

*Pro->GlutamicSemialdehyde modifications may also include Pro->Hydroxyproline modifications (both $\Delta m/z$ +15.995).¹² MCO, metal-catalysed oxidation.

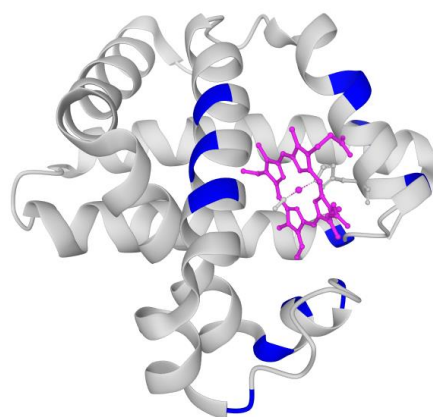

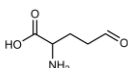
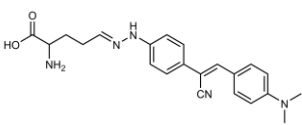
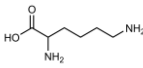
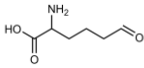
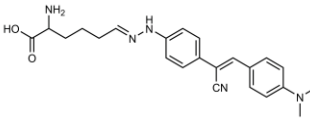
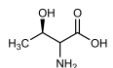

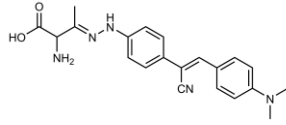


Figure 4.3.2. Crystal structure for equine myoglobin (<https://www.rcsb.org/structure/1dwr>) showing carbonylated residues (blue) found via bottom-up proteomic analysis of the oxidised protein that are not present in unmodified samples and their proximity to the heme prosthetic group (pink).⁴⁰⁻⁴¹

4.3.3 Identification of oxidised peptide labelling with DPANzine fluorophores

Based on the findings from the analysis of oxidised myoglobin, modifications to lysine, threonine and proline residues resulting from DPANzine conjugation give rise to mass shifts of +276.1375, +259.1110 and +258.1269 Da, respectively (**Table 4.3.3**). Given the absence of arginine modifications for the unlabelled oxidised myoglobin and the limited number of dynamic modifications allowed by Sequest, we decided to exclude DPANzine-arginine modifications from the analysis.⁴² Despite the high number of carbonylated residues found prior to labelling, a Sequest HT search of oxidised myoglobin incubated with DPANzine identified only one labelled residue, assigned to either Lysine 14 or 15 of peptide HGTVVLTALGGILKK (amino acid 79 or 80 with respect to the protein sequence), although exact assignment was not possible due to the absence of $y_1/b_{15/14}$ fragment ions (**Figure 4.3.3**). This modification was also unique to DPANzine-labelled oxidised myoglobin and not found in unmodified (with or without incubation with DPANzine) or unlabelled samples.

Table 4.3.3. Modified amino acid residues used for Sequest HT database search of digested myoglobin labelled with DPANzine.

Amino acid residue	Carbonylation product	Modified residue structure	Δ mass (Da)	Δ mass (Da), NaCNBH ₃
 Proline Exact mass: 115.0633 Average mass: 115.1320	 glutamic semialdehyde		Exact: 276.1375 Average: 276.3430	Exact: 278.1531 Average: 278.3586
 Lysine Exact mass: 146.1055 Average mass: 146.1900	 amino adipic semialdehyde (allysine)		Exact: 259.1110 Average: 259.3120	Exact: 261.1266 Average: 261.1320
 Threonine Exact mass: 119.0582 Average mass: 119.1200	 2-amino-3-ketobutyric acid		Exact: 258.1269 Average: 258.3279	Exact: 260.1426 Average: 260.3436

An additional modification to the N-terminal lysine of this same peptide (with an additional N-terminal missed cleavage) was identified with a mass shift consistent with that of DPANzine conjugated to amino adipic semialdehyde (+259.1110 Da); however, nearly identical peptide-spectrum matches (PSMs), varying only in the number of missed cleavages, were found in oxidised unlabelled, labelled, and oxidised labelled myoglobin

digests (Figure 4.3.4). Interestingly, this modification, which was located at the N-terminus of the peptide, appeared to suppress fragmentation, as no unambiguous b ions were observed in the fragment spectra. We therefore hypothesised that this mass shift instead resulted from another modification of similar mass, such as a previously unreported PTM, non-enzymatic glycation product, or sample preparation artifact, but were unable to find relevant literature examples to confidently assign it. Modification of myoglobin by reducing sugars and reactive aldehydes, such as glucose, fructose, ribose, 3-deoxyglucosone and methylglyoxal, have been reported previously.^{29-30, 43-47} These modifications are a result from a three-step process known as the Maillard reaction: (1) reaction between the carbonyl groups of these molecules with the amines of amino acid residues to form Schiff bases, (2) rearrangement of aldose sugars to form ketoamine (Amadori) adducts and (3) degradation to more reactive carbonyls, which can form crosslinks to other amino acid residues.⁴⁸ Because of the wide range of potential Maillard reaction products, we hypothesised that this modification could result from an unexpected fragmentation, degradation or reaction product of this pathway.

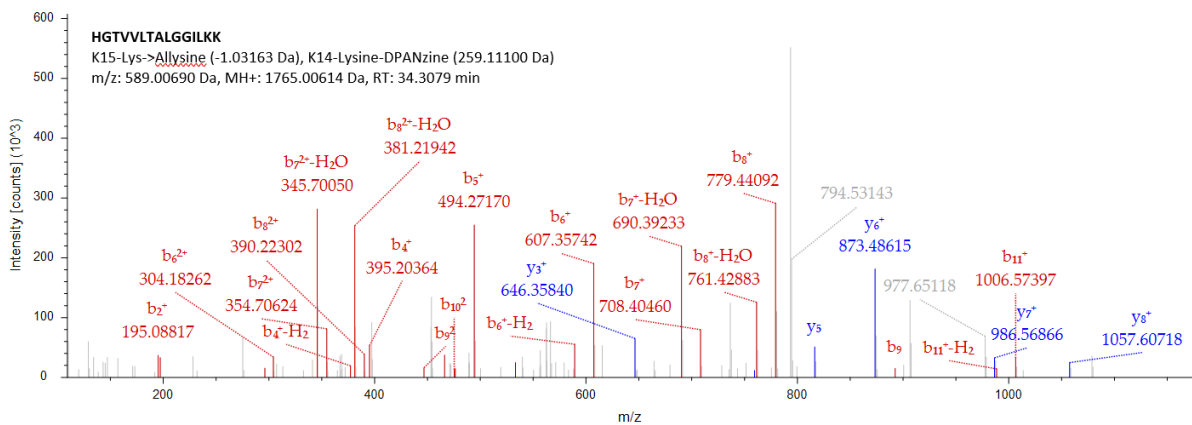


Figure 4.3.3. Fragment spectrum for m/z 589.0069 Da identified as peptide HGTVVLTALGGILKK with modifications corresponding to amino adipic semialdehyde (allysine) and DPANzine labelling at lysines 15 and 14, respectively.

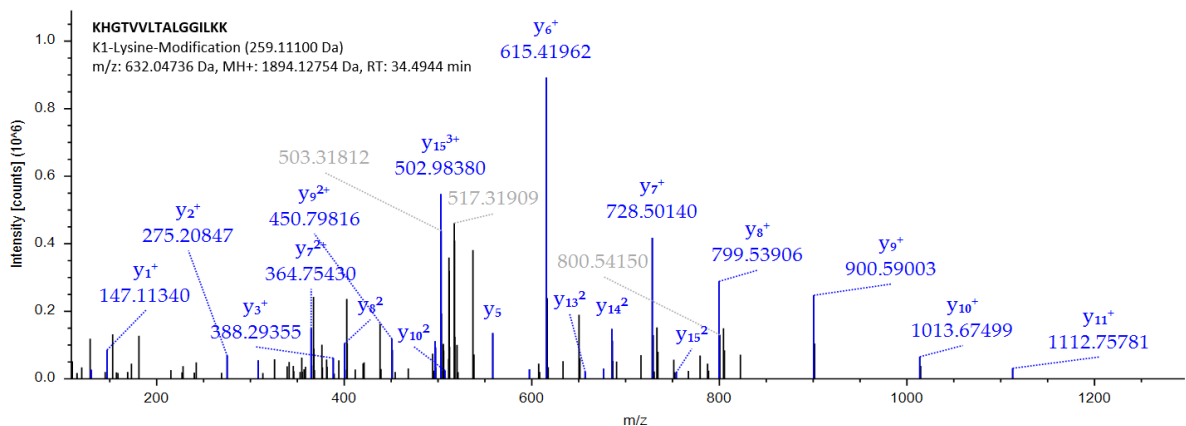


Figure 4.3.4. Fragment spectrum for m/z 532.04736 Da identified as peptide KHGTVVLTALGGILKK with an unknown modification detected for lysine 1.

The reaction between hydrazones and carbonyls to form Schiff bases is reversible; therefore, labelling with hydrazone derivatives is often followed by a reduction step, using either NaBH₄ or NaCNBH₃.¹⁵ Because the limited labelling of oxidised myoglobin by DPANzine could have resulted from decomposition of conjugates during sample preparation, duplicate samples of DPANzine-labelled oxidised and unmodified myoglobin were treated by an additional reduction step using NaCNBH₃. Against expectations, NaCNBH₃ reduction resulted in no detectable DPANzine labelling. It is possible that this reduction step may have somehow modified other unlabelled residues, leading to mass shifts that were not included as potential modifications in the Sequest search, although the reduction of aldehydes and ketones by NaCNBH₃ at neutral pH is expected to be negligible.⁴⁹ Unexpected combinations of modifications would lead to mass shifts that are outside the narrow mass error range (and therefore these peptides would be missed during targeted searches).

4.3.4 Error-tolerant search for oxidative and chemical modifications

Initial proteomic analysis of these samples established that the complexity of MS data derived from tryptic digests of MCO-treated and chemically labelled proteins may prohibit the effective identification of PSMs resulting from unexpected combinations of modifications. An alternative strategy, known as an 'error-tolerant' or 'open' search, can be used to instead identify peptides with large deviations from their expected m/z , by setting the search parameters to include wide precursor mass errors. Previously unmatched MS/MS spectra can therefore be assigned to a theoretical peptide sequence to determine mass shifts corresponding to PTMs or other deviations from the database amino acid sequence.⁵⁰ An error-tolerant search of the data for DPANzine-labelled oxidised myoglobin for mass shifts of +200-1000 Da found several modifications that were not present in the control and oxidised unlabelled samples. Of these, mass shifts of +241.987 - 242.053 Da were of interest because they possibly resulted from modification of a carbonylated lysine residue by DPANzine with simultaneous loss of ammonia (-17.0265 Da), which was not included as a dynamic modification in the original search (**Figure 4.3.5**). On inspection of the relevant fragment spectra (**Figure 4.3.6**), several of these peptides also had b and y ion fragments consistent with addition of an \approx 260 Da modification, from which the labelled amino acids could also be predicted. Specifically, a +260.1184 Da mass shift for the b₉⁺ ion of peptide NDIAAKYKELGFQG (**Figure 4.3.6B**) would likely result from modification of a lysine residue, whereas a mass shift of +258.1927 Da for the b₁₂²⁺ ion of peptide HGTVVLTALGGILK suggests modification of a threonine residue (**Figure 4.3.6C**).

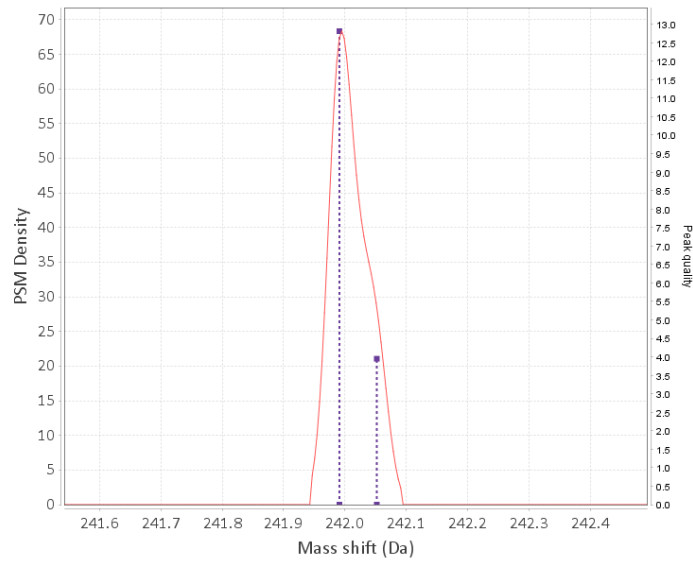
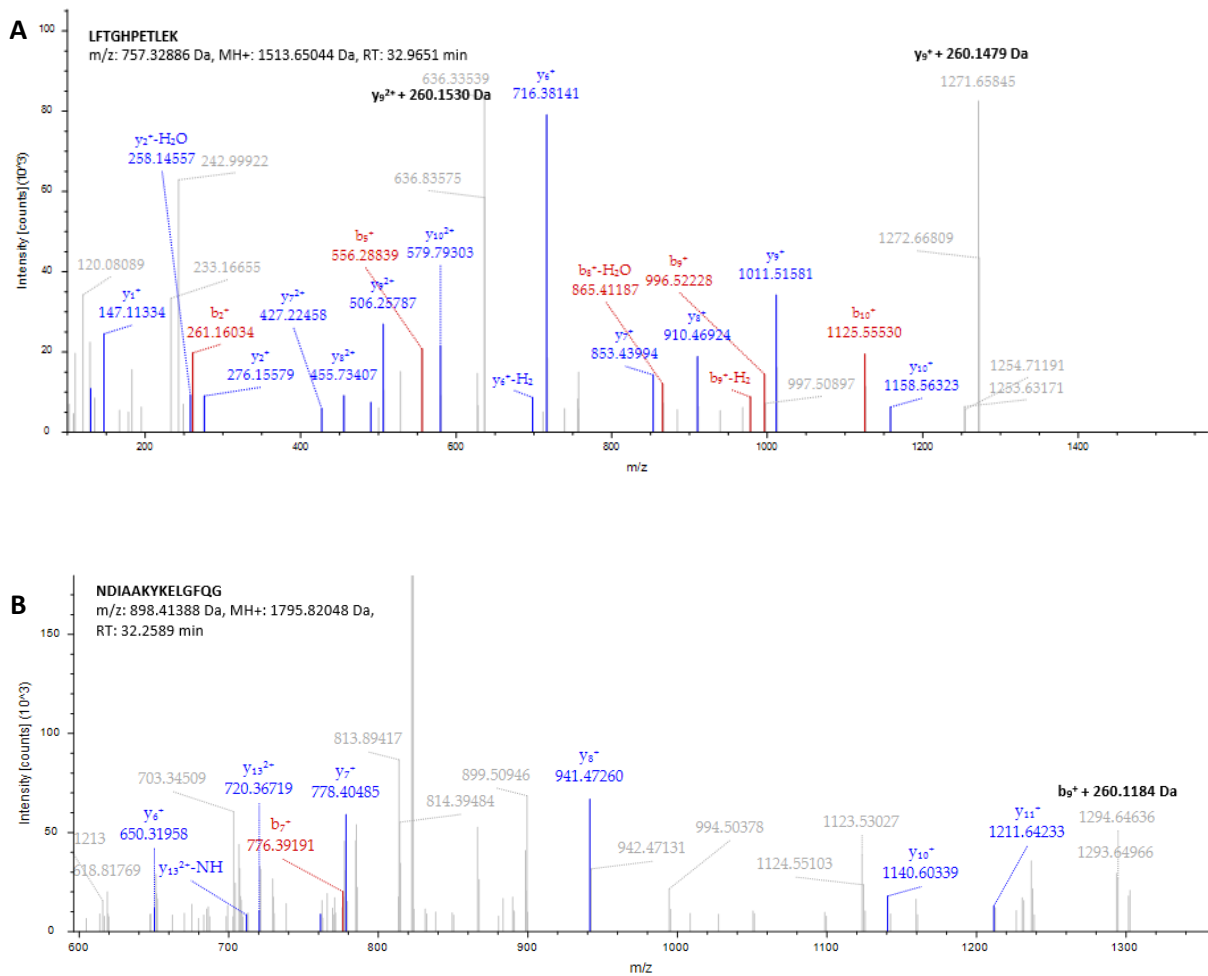


Figure 4.3.5. Mass shift distribution of a +241.987-242.053 Da modification identified for DPANzine-labelled oxidised myoglobin. The calculated kernel density estimate (KDE) is shown as a red line. Peak qualities (second derivative of KDE) are indicated by purple dotted lines.⁵¹



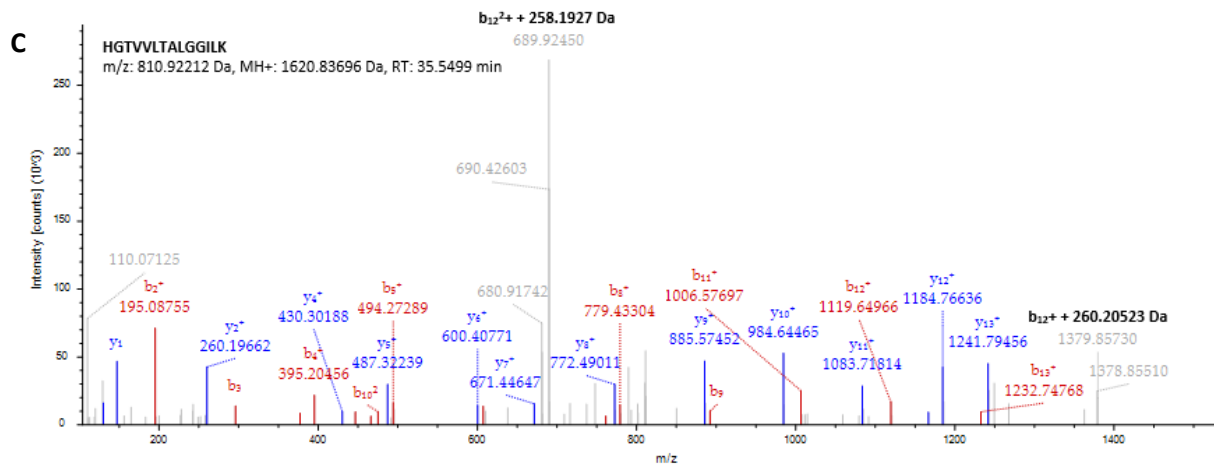


Figure 4.3.6. Fragment spectrum for peptides with mass shifts of +242 Da detected using a DeltaMass error-tolerant search. (A) M/z 757.32886 Da identified as peptide LFTGHPETLEK with a modification of +242.025 Da. (B) Fragment spectrum for m/z 898.41388 Da identified as peptide NDIAAKYKELGFQG with a modification of +241.987 Da. (C) Fragment spectrum for m/z 810.92212 Da identified as peptide HGTVVLTALGGILK with a modification of +241.987.

Based on the results from this search, it appears that several DPANzine modifications may have been missed during the targeted search. Many additional mass shifts specific to the DPANzine-labelled oxidised myoglobin sample were not assigned to known modifications, but possibly arose from fragmentation or degradation of this label, or combinations of modifications, which are difficult to interpret without individual manual analysis of each fragment spectrum.²⁴

4.4 Conclusions

Our preliminary assessment of myoglobin carbonylation suggests that the mechanisms behind *in vitro* carbonylation, in this case induced by exposure to a combination of ascorbic acid and FeCl₃, may differ from those involved in carbonylation resulting from the lower levels of oxidative stress encountered under physiological conditions or during sample preparation.²² These differences could be significant in the pathogenesis of a range of human diseases.²⁸⁻³³ However, broader investigations into the effects of other ROS model systems and comparison of different heme and nonheme proteins is required to better explain these findings.²²

A disadvantage of Sequest searches is the limited number of dynamic modifications that can be defined, which can potentially lead to misassigned peptides or undetected modifications.^{35, 42} This limitation is clearly evident from the present study, in which the complexity and heterogeneity of potential *in vivo*, *in vitro* and chemical modifications strongly influenced the optimum data analysis strategy for identifying individual modifications. An initial targeted Sequest search returned few viable examples of DPANzine labelling. In contrast, the utilisation of an error-tolerant search provided a convenient means of comparing mass shift profiles between experimental and control groups. The identification of unique mass shifts for DPANzine-labelled

oxidised myoglobin confirmed that this approach is an easy way to screen for unpredicted combinations of modifications, which result in discrepancies between the precursor and fragment ion spectra. From these data, we were able to deduce that DPANzine primarily modifies carbonylated lysine and threonine residues, which is consistent with the fact that MCO-induced carbonylation also occurred more frequently at these residues.

Undoubtedly, the complexity of modified peptides that result from MCO is a significant challenge for MS proteomics. Even for a small protein such as myoglobin, manual interpretation of fragment spectra to correlate error-tolerant search results with mass shifts for individual fragment ions is a time-consuming process. The establishment of an ideal approach to analysing chemical labelling of carbonylated residues will require further investigation into alternative models of protein oxidation, sample preparation and analysis workflows, and data analysis methods. For example, sequence libraries generated from additional optimised data-dependent acquisitions could be used in combination with data-independent analysis strategies to obtain a more comprehensive analysis of these modifications and even be used for analysing complex labelled samples, such as cell lysates.^{52, 53} However, this study provided useful insights into potential strategies for overcoming the significant challenges of analysing heavily modified proteins and presents a simple data processing strategy for utilising pre-existing data acquired using a common shotgun proteomics approach.

4.5 Acknowledgements

I would like to acknowledge the contributions of Siyang Ding and Yuning Hong from Latrobe University, Melbourne for kindly providing the modified myoglobin samples that were used for my bottom-up proteomics analysis. Orbitrap MS spectra was acquired by Clifford Young of the Future Industries Institute, the University of South Australia, Adelaide.

4.6 Experimental

4.6.1 MCO protocol. Equine heart myoglobin (10 mg/mL) was solubilised in oxidising buffer (25 mM 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid [HEPES], 25 mM ascorbic acid, 100 μ M FeCl₃, pH 7.2) and incubated for at 37 °C for 1 h. The protein solution was then applied to a gel filtration chromatography column in HE buffer (50 mM HEPES, 1 mM EDTA, pH 7.2). The resulting oxidised protein was quantified using a Bradford protein assay and aliquoted to store at -80 °C until further use.

4.6.2 Labelling with DPANzine. Samples of oxidised and unmodified myoglobin (10 μ g per sample, $\approx 1.5 \times 10^{-4}$ μ mol) were allowed to react with DPANzine (10 equivalents, 1.5×10^{-3} μ mol) in HE buffer containing <5 % DMSO (v/v) for 20 min at room temperature in the dark. Additional labelled samples (oxidised and unmodified) were reduced by addition of NaCNBH₃ (60 mM, diluted from a 600 mM stock) and incubation at room temperature for 30 min. Lyophilised aliquots were stored at -20 °C until further use.

4.6.3 Trypsin/Lys-C digestion. Following oxidation, labelling and/or reduction, myoglobin samples (10 μ g) were reconstituted in 7 M urea in 50 mM Tris with 5 mM dithiothreitol (100 μ L) and incubated at 37 °C for 30 min. Iodoacetamide was added to a final concentration of 15 mM and the solution incubated in the dark at room temperature for 30 min. Following reduction and alkylation, myoglobin samples were digested using Mass Spectrometry Grade Trypsin/Lys-C (Promega, Madison, USA) according to the manufacturer's two-step in-solution digestion protocol. Briefly, 1/25 equivalents protease:protein (w/w; 10 μ L of a 40 ng/ μ L stock) were added to each sample and the reaction incubated at 37 °C for 3 h 45 min while shaking. Solutions were then diluted with 50 mM Tris (pH 8, 700 μ L) and incubated overnight at 37 °C while shaking. Digestions were quenched with trifluoroacetic acid (8 μ L) and samples lyophilised before reconstitution in MilliQ water (90 μ L) and purification using Pierce™ C18 Spin Columns (Thermo Fisher Scientific, Scoresby, Australia).

4.6.4 LC-MS/MS. LC-MS analysis of myoglobin samples was performed using an Ultimate 3000 RSLCnano HPLC connected to an Orbitrap Exploris 480 mass spectrometer (Thermo Scientific, Bremen, Germany). Peptides (250 ng) in 7 % acetonitrile were loaded onto a 25 cm analytical column (75 μ m internal diameter packed with 1.9 μ m C18 particles) heated to 50 °C. A 10-minute linear gradient (3 to 28 % acetonitrile in 0.1 % formic acid) was run at 300 nL/min to facilitate peptide separation and eluted over a total run time of 45 minutes. Compensation voltages of -50 and -70 V were generated from a FAIMS Pro interface (Thermo Scientific) to regulate the entry of eluted peptides into the mass spectrometer. MS scans (m/z 300 to 1500) were acquired at resolution 60 000 (m/z 200) in positive ion mode. MS/MS scans (15 000 resolution) of multiply charged precursors were obtained in data-dependent mode, with fragmentation (27.5 % high-energy collisional dissociation [HCD] energy) performed if the precursors surpassed a 1×10^6 intensity threshold. The dynamic exclusion period was set to 10 seconds.

4.6.5 Data Analysis. Proteome Discoverer v2.5 (Thermo Scientific) was used with processing workflows based on the basic Sequest analysis template ('PWF_OT_Basic_SequestHT') included in the software package, which included the following nodes: (1) 'Spectrum Files'; (2) 'Spectrum Selector' to extract MS2 scans, together

with a precise precursor m/z and charge; (3) 'Sequest HT', which generates virtual spectra from a database (in this case the FASTA file for *Equus caballus* (horse) myoglobin (<https://www.uniprot.org/uniprot/P68082.fasta>) and matches them to experimental spectra; and (4) 'Fixed Value PSM Validator' to eliminate peptides with Delta Correlation (Delta Cn) >0.05 .⁵⁴ All node parameters were kept as their default values, with the exception of chemical modifications, which included dynamic N-terminal modifications acetylation (+42.001 Da), methionine loss (-131.040 Da) and methionine loss with acetylation (-89.030 Da); static cysteine carbamidomethylation (+57.021 Da); and modifications outlined in **Tables 4.3.1** and **4.3.3** for assessing carbonylation, and labelling with DPANzine with or without subsequent NaCNBH₃ reduction, respectively.

An error-tolerant search was also performed by setting the Sequest HT precursor mass tolerance to ± 1000 Da and included the same chemical modifications, modules and parameters described above. The analysis results for each file were exported separately as .pep.xml files for analysis using DeltaMass v1.2.2 (<https://github.com/chhh/deltamass>).⁵⁴ The Sequest HT output files were then searched against forward sequences generated from the horse myoglobin FASTA database with zero-peak correction, for mass shifts of 200-1000 Da. The minimum data ('Min Data') and minimum peptide spectrum matches ('Min PSMs') were set to 5 and 1, respectively. All other peak detection settings were left as their default values.

4.7 References

1. El-Aneed, A.; Cohen, A.; Banoub, J., Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Appl. Spectrosc. Rev.* 2009, *44* (3), 210-230.
2. Ponomarenko, E. A.; Poverennaya, E. V.; Ilgisonis, E. V.; Pyatnitskiy, M. A.; Kopylov, A. T.; Zgoda, V. G.; Lisitsa, A. V.; Archakov, A. I., The Size of the Human Proteome: The Width and Depth. *Int. J. Anal. Chem.* 2016, *2016*, 7436849.
3. Smith, L. M.; Kelleher, N. L.; Linial, M.; Goodlett, D.; Langridge-Smith, P.; Goo, Y. A.; Safford, G.; Bonilla, L.; Kruppa, G.; Zubarev, R., Proteoform: a single term describing protein complexity. *Nat. Methods* 2013, *10* (3), 186-187.
4. Lisitsa, A.; Moshkovskii, S.; Chernobrovkin, A.; Ponomarenko, E.; Archakov, A., Profiling proteoforms: Promising follow-up of proteomics for biomarker discovery. *Expert Rev. Proteomic.* 2014, *11* (1), 121-129.
5. Karty, J. A.; Ireland, M. M. E.; Brun, Y. V.; Reilly, J. P., Artifacts and unassigned masses encountered in peptide mass mapping. *J. Chromatogr. B* 2002, *782* (1), 363-383.
6. Madian, A.; Regnier, F., Proteomic identification of carbonylated proteins and their oxidation sites. *J. Proteome Res.* 2010, *9* (8), 3766-3780.
7. Griffiths, H. R., Chemical modifications of biomolecules by oxidants. In *React., Process.*, Springer: 2005; pp 33-62.
8. Welch, K. D.; Davis, T. Z.; Van Eden, M. E.; Aust, S. D., Deleterious iron-mediated oxidation of biomolecules¹ This article is part of a series of reviews on "Iron and Cellular Redox Status." The full list of papers may be found on the homepage of the journal. ⁶ Guest Editor: Mario Comporti. *Free Radic. Biol. Med.* 2002, *32* (7), 577-583.
9. Singh, R.; Barden, A.; Mori, T.; Beilin, L., Advanced glycation end-products: a review. *Diabetologia* 2001, *44* (2), 129-146.
10. Henning, C.; Glomb, M. A., Pathways of the Maillard reaction under physiological conditions. *Glycoconj. J.* 2016, *33* (4), 499-512.
11. Møller, I. M.; Rogowska-Wrzesinska, A.; Rao, R. S. P., Protein carbonylation and metal-catalyzed protein oxidation in a cellular perspective. *J. Proteom.* 2011, *74* (11), 2228-2242.
12. Rogowska-Wrzesinska, A.; Wojdyla, K.; Nedić, O.; Baron, C. P.; Griffiths, H., Analysis of protein carbonylation-pitfalls and promise in commonly used methods. *Free Radic. Res.* 2014, *48* (10), 1145-1162.
13. Baraibar, M. A.; Ladouce, R.; Friguet, B., Proteomic quantification and identification of carbonylated proteins upon oxidative stress and during cellular aging. *J. Proteom.* 2013, *92*, 63-70.
14. Carrico, I. S., Chemoselective modification of proteins: Hitting the target. *Chem. Soc. Rev.* 2008, *37* (7), 1423-1431.
15. Zatsepin, T. S.; Stetsenko, D. A.; Gait, M. J.; Oretskaya, T. S., Use of carbonyl group addition–Elimination reactions for synthesis of nucleic acid conjugates. *Bioconjug. Chem.* 2005, *16* (3), 471-489.
16. Yan, L.-J.; Forster, M. J., Chemical probes for analysis of carbonylated proteins: A review. *J. Chromatogr. B* 2011, *879* (17), 1308-1315.
17. Bollineni, R. C.; Hoffmann, R.; Fedorova, M., Identification of protein carbonylation sites by two-dimensional liquid chromatography in combination with MALDI- and ESI-MS. *J. Proteom.* 2011, *74* (11), 2338-2350.

18. Adams, S.; Green, P.; Claxton, R.; Simcox, S.; Williams, M. V.; Walsh, K.; Leeuwenburgh, C., Reactive carbonyl formation by oxidative and non-oxidative pathways. *Front. Biosci.* 2001, *6*, A17-A24.
19. Erkan, H.; Telci, D.; Dilek, O., Design of fluorescent probes for bioorthogonal labeling of carbonylation in live cells. *Sci. Rep.* 2020, *10* (1), 1-11.
20. Zhang, Q.; Ames, J. M.; Smith, R. D.; Baynes, J. W.; Metz, T. O., A perspective on the Maillard reaction and the analysis of protein glycation by mass spectrometry: probing the pathogenesis of chronic disease. *J. Proteome Res.* 2009, *8* (2), 754-769.
21. Chait, B. T., Mass spectrometry: Bottom-up or top-down? *Science* 2006, *314* (5796), 65-66.
22. Yin, V.; Holzschlager, D.; Konermann, L., Delineating heme-mediated versus direct protein oxidation in peroxidase-activated cytochrome c by top-down mass spectrometry. *Biochemistry* 2020, *59* (42), 4108-4117.
23. Madian, A. G.; Regnier, F. E., Proteomic identification of carbonylated proteins and their oxidation sites. *J. Proteom. Rev.* 2010, *9* (8), 3766-3780.
24. Fedorova, M.; Bollineni, R. C.; Hoffmann, R., Protein carbonylation as a major hallmark of oxidative damage: Update of analytical strategies. *Mass Spectrom. Rev.* 2014, *33* (2), 79-97.
25. Jomova, K.; Baros, S.; Valko, M., Redox active metal-induced oxidative stress in biological systems. *Transit. Met. Chem.* 2012, *37* (2), 127-134.
26. Kamga, C.; Krishnamurthy, S.; Shiva, S., Myoglobin and mitochondria: A relationship bound by oxygen and nitric oxide. *Nitric Oxide* 2012, *26* (4), 251-258.
27. Reeder, B. J.; Wilson, M. T., Hemoglobin and myoglobin associated oxidative stress: from molecular mechanisms to disease states. *Curr. Med. Chem.* 2005, *12* (23), 2741-2751.
28. Boadi, W. Y.; Johnson, D., Effects of low doses of quercetin and genistein on oxidation and carbonylation in hemoglobin and myoglobin. *J. Diet. Suppl.* 2014, *11* (3), 272-287.
29. Roy, A.; Sil, R.; Chakraborti, A. S., Non-enzymatic glycation induces structural modifications of myoglobin. *Mol. Cell. Biochem.* 2010, *338* (1), 105-114.
30. Bokiej, M.; Livermore, A. T.; Harris, A. W.; Onishi, A. C.; Sandwick, R. K., Ribose sugars generate internal glycation cross-links in horse heart myoglobin. *Biochem. Biophys. Res. Commun.* 2011, *407* (1), 191-196.
31. Qureshi, A.; Gurbuz, Y.; Niazi, J. H., Biosensors for cardiac biomarkers detection: A review. *Sens. Actuators, A* 2012, *171-172*, 62-76.
32. Lippi, G.; Plebani, M., Serum myoglobin immunoassays: Obsolete or still clinically useful? *Clin. Chem. Lab. Med.* 2016, *54* (10), 1541-1543.
33. Pur, M. R. K.; Hosseini, M.; Faridbod, F.; Ganjali, M. R., Highly sensitive label-free electrochemiluminescence aptasensor for early detection of myoglobin, a biomarker for myocardial infarction. *Microchimica Acta* 2017, *184* (9), 3529-3537.
34. Tran, D. T., Engineering proteases for mass spectrometry-based post translational modification analyses. *Proteomics* 2019, *19* (10), 1700471.
35. Verrastro, I.; Pasha, S.; Jensen, K. T.; Pitt, A. R.; Spickett, C. M., Mass spectrometry-based methods for identifying oxidized proteins in disease: advances and challenges. *Biomolecules* 2015, *5* (2), 378-411.
36. Saveliev, S.; Bratz, M.; Zubarev, R.; Szapacs, M.; Budamgunta, H.; Urh, M., Trypsin/Lys-C protease mix for enhanced protein mass spectrometry analysis. *Nat. Methods* 2013, *10* (11), 1134.
37. Giansanti, P.; Tsiatsiani, L.; Low, T. Y.; Heck, A. J., Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat. Protoc.* 2016, *11* (5), 993-1006.

38. Trevisiol, S.; Ayoub, D.; Lesur, A.; Ancheva, L.; Gallien, S.; Domon, B., The use of proteases complementary to trypsin to probe isoforms and modifications. *Proteomics* 2016, *16* (5), 715-728.
39. Yin, V.; Mian, S. H.; Konermann, L., Lysine carbonylation is a previously unrecognized contributor to peroxidase activation of cytochrome c by chloramine-T. *Chem. Sci.* 2019, *10* (8), 2349-2359.
40. Chu, K.; Vojtchovský, J.; McMahon, B. H.; Sweet, R. M.; Berendzen, J.; Schlichting, I., Structure of a ligand-binding intermediate in wild-type carbonmonoxy myoglobin. *Nature* 2000, *403* (6772), 921-923.
41. Sehnal, D.; Rose, A. S.; Koča, J.; Burley, S. K.; Velankar, S., Mol*: towards a common library and tools for web molecular graphics. In *Proceedings of the Workshop on Molecular Graphics and Visual Analysis of Molecular Data*, Eurographics Association: Brno, Czech Republic, 2018; pp 29–33.
42. Rykær, M.; Svensson, B.; Davies, M. J.; Hagglund, P., Unrestricted mass spectrometric data analysis for identification, localization, and quantification of oxidative protein modifications. *J. Proteome Res.* 2017, *16* (11), 3978-3988.
43. Banerjee, S.; Maity, S.; Chakraborti, A. S., Methylglyoxal-induced modification causes aggregation of myoglobin. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 2016, *155*, 1-10.
44. Roy, A.; Sen, S.; Chakraborti, A. S., In vitro nonenzymatic glycation enhances the role of myoglobin as a source of oxidative stress. *Free Radic. Res.* 2004, *38* (2), 139-146.
45. Sen, S.; Bose, T.; Roy, A.; Chakraborti, A. S., Effect of non-enzymatic glycation on esterase activities of hemoglobin and myoglobin. *Mol. Cell. Biochem.* 2007, *301* (1), 251-257.
46. Ghelani, H.; Razmovski-Naumovski, V.; Pragada, R. R.; Nammi, S., (R)- α -Lipoic acid inhibits fructose-induced myoglobin fructation and the formation of advanced glycation end products (AGEs) in vitro. *BMC Complement. Altern. Med.* 2018, *18* (1), 13.
47. Tang, R.; Faisal, M.; Alatar, A. A.; Alsaleh, A. N.; Saeed, M.; Ahmad, S., Glycation of heme-protein, "myoglobin" by 3-deoxyglucosone: Implications in immunogenicity. *J. King Saud Univ. Sci.* 2020, *32* (5), 2598-2602.
48. Horvat, Š.; Jakas, A., Peptide and amino acid glycation: new insights into the Maillard reaction. *J. Pept. Sci.* 2004, *10* (3), 119-137.
49. Clinton, F., Sodium cyanoborohydride-A highly selective reducing agent for organic functional groups. *Synthesis* 3.
50. Chick, J. M.; Kolippakkam, D.; Nusinow, D. P.; Zhai, B.; Rad, R.; Huttlin, E. L.; Gygi, S. P., A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. *Nat. Biotechnol.* 2015, *33* (7), 743-749.
51. Avtonomov, D. M.; Kong, A.; Nesvizhskii, A. I., DeltaMass: Automated detection and visualization of mass shifts in proteomic open-search results. *J. Proteome Res.* 2018, *18* (2), 715-720.
52. Sajic, T.; Liu, Y.; Aebersold, R., Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. *Proteomics Clin. Appl.* 2015, *9* (3-4), 307-321.
53. Hansen, F. M.; Tanzer, M. C.; Brüning, F.; Bludau, I.; Stafford, C.; Schulman, B. A.; Robles, M. S.; Karayel, O.; Mann, M., Data-independent acquisition method for ubiquitinome analysis reveals regulation of circadian biology. *Nat. Comm.* 2021, *12* (1), 254.
54. Tabb, D. L.; Eng, J. K.; Yates, J. R., Protein identification by SEQUEST. In *Proteome Res.: Mass Spectrom.*, Springer: 2001; pp 125-142.

Development of a modular synthetic route for protein chemical cross-linking reagents

Foreword

Whereas post-translational modifications are widely acknowledged as important factors in the pathogenesis of many human diseases, there is an increasing interest in the study of protein quaternary structure and protein-protein interactions and how they influence protein function. In Chapter 4, we demonstrated that bottom-up proteomics can be an effective method for studying the interactions between novel chemical reagents and individual amino acid residues in proteins. This strategy can also be applied for the analysis of proteins modified by *chemical crosslinking*, a term used to denote the covalent modification of proteins by reagents ('crosslinkers') with two reactive groups that are able to form a link between nearby amino acid residues. Crosslinked peptides can provide information related to the position of amino acids within proteins, and therefore how these residues were positioned in the intact protein or complex. Recent developments in crosslinkers design have gained a lot of attention from mass spectrometrists, who are interested in incorporating additional functionalities, such as affinity tags and cleavable groups, to facilitate the purification and identification of crosslinked peptides. The following study describes work by our research group to develop modular synthetic routes towards novel crosslinkers, which include various combinations of these features.

Statement of Authorship

Title of Paper	Development of a modular synthetic route for protein chemical cross-linking reagents		
Publication Status	<input type="checkbox"/> Published	<input type="checkbox"/> Accepted for Publication	<input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
	<input checked="" type="checkbox"/> Submitted for Publication		
Publication Details			

Principal Author

Name of Principal Author (Candidate)	Henry Sanders		
Contribution to the Paper	Majority of synthesis and experimental work, interpreted data and wrote manuscript.		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	23-Nov-20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Kayla Downey		
Contribution to the Paper	Baseline synthesis, data interpretation and experimental methods used from published thesis.		
Signature		Date	27-11-2020

Name of Co-Author	Emily R Bubner		
Contribution to the Paper	Assisted in synthesis, analysis and experimentation.		
Signature		Date	

Please cut and paste additional co-author panels here as required.

Statement of Authorship

Name of Co-Author	Katherine G Stevens		
Contribution to the Paper	Contributed to design, synthesis and characterisation of photoreactive crosslinkers, click chemistry experiment design		
Signature		Date	6/12/20

Name of Co-Author	Andrew D Abell		
Contribution to the Paper	Experimental design, editing of manuscript		
Signature		Date	6/12/20

Name of Co-Author	Tara L Pukala		
Contribution to the Paper	Provided advice related to experimental design and project aims, editing of manuscript		
Signature		Date	4/12/20

Please cut and paste additional co-author panels ~~here~~ as required.

Development of a modular synthetic route for protein chemical cross-linking reagents

Henry M Sanders¹, Emily R Bubner¹, Katherine G Stevens¹, Kayla M Downey¹, Andrew D Abell¹ and Tara L Pukala^{1*}

¹School of Physical Sciences, The University of Adelaide, Adelaide, SA 5005, Australia

*Correspondence: Tara L Pukala; School of Physical Sciences, The University of Adelaide, Adelaide, SA 5005, Australia; tara.pukala@adelaide.edu.au; Tel. +61 8 8313 5497

Abstract

Cross-linking mass spectrometry is a rapidly emerging technique that gives information on protein structure and subunit architecture. Typically in such an experiment, chemical reagents covalently link protein sites together, which, when followed by proteolytic digestion and detection of linkage sites by mass spectrometry analysis, gives information on inter- and intra-protein contacts. While many cross-linking reagents are commercially available, their specific functionalities may not always satisfy the diverse chemical and analytical requirements best suited for a particular system or experiment of interest. Here we describe a modular synthetic protocol allowing customisation of reactive groups and spacer arms, giving flexibility in linker design to enable the vast variety of potential protein cross-linking experiments. Implementing the general synthetic method allowed for the production of 8 unique homo- and heterobifunctional cross-linkers, including 4 different reactive groups, positive and negative mode mass spectrometry-cleavable spacer arms and sites for post-linkage derivatisation, for example with enrichment and identification tags. Optimisation and application of these reagents to model systems exemplifies the potential for this modular protocol to address some of the analytical challenges associated with protein structure determination by cross-linking mass spectrometry.

Introduction

The function of proteins is inherently tied to their three-dimensional structure and how they interact with other molecules. Consequently, the ongoing development of analytical methods to probe protein organisation is critical for advancements in structural biology. Many protein systems are not amenable to structure determination by traditional high-resolution techniques such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, due to inherent limitations of the methods. As a result, alternative low-resolution approaches such as chemical cross-linking combined with mass spectrometry (XL-MS) have emerged as powerful alternatives. XL-MS is a highly versatile tool to interrogate protein structure and map protein-protein binding interfaces.¹ Unlike these traditional methods, XL-MS does not require a homogenous analyte structure and can give information on highly dynamic proteins. Only micrograms of sample is required and it can be performed both *in vitro* and *in vivo*,^{2,3} with either targeted or proteome-wide applications.⁴⁻⁶

In a standard XL-MS workflow, proteins are chemically ligated in their native state using small molecule cross-linkers and then enzymatically digested. The resultant peptide mixture is sequenced by liquid chromatography tandem MS (LC-MS/MS) to locate linkage sites. The simplest cross-linker design includes two reactive units, which allow the compound to covalently link amino acids, commonly at basic or sulfhydryl residues, separated by a spacer arm (**Fig. 1**). For example, *N*-hydroxysuccinimide (NHS), the most common reactive group, targets lysine or N-terminal residues, while the maleimide reactive group targets cysteines.^{3,7} In the case of certain proteins, amino acids of the same type are scarce, leading to the development of heterobifunctional linkers such as *N*-(α -maleimidoacetoxy) succinimide ester (AMAS), a linker that combines NHS and maleimide reactive groups. Building upon the concept of heterobifunctional reagents, non-specific linkers combine a 'standard reactive group' with a photoreactive group, typically aryl azides or diazirines, capable of indiscriminately reacting with heteroatom bonds to greatly increase potential linkage opportunities.

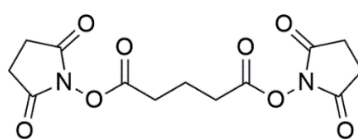


Figure 1. An example of a simple, commercially available cross-linker, disuccinimidyl glutarate (DSG), comprised of a non-cleavable 5 carbon chain between two lysine-reactive NHS residues.

Following covalent cross-linking, as the length of the spacer arm is known, inter- and intra-protein upper distance constraints can be defined through site-specific localisation of the cross-links.³ These distance constraints allow for the interrogation of protein-binding interfaces, intramolecular contacts and interaction stoichiometries.⁸ Furthermore, while techniques like crystallography require rigid structures for accurate models to be constructed, XL-MS can also capture regions that exhibit structural differences between conformational

states, and can give information on the rearrangements and transient interactions of proteins, which are often critical in controlling cellular processes.^{2,3}

Despite great advances in the field of XL-MS, analytical limitations remain. For example, due to low reaction efficiencies, unlinked peptides and 'dead-end' modifications, which have linked on one end and are hydrolysed on the other, significantly outnumber useful intra- and inter-protein cross-links. This has led to an increased interest in the introduction of enrichment tags capable of extracting modified peptides through a variety of common methods such as avidin- or immobilised-metal-affinity chromatography.⁹⁻¹² In addition, peptides joined with non-cleavable linkers must be analysed as covalent pairs, giving rise to more complex MS/MS data as fragmentation occurs simultaneously on both chains. This has spurred the development of cleavable spacer arms capable of fragmenting under low-energy gas-phase activation regimes (such as collision induced dissociation, CID) during MS.^{13,14} Here, linked peptides initially fragment at the labile spacer arm yielding characteristic fragmentation patterns, distinguishing them from unlinked peptides, and enabling peptide pairs to be selected independently by MS³ analysis to more easily identify the residue-specific site of modification.¹⁵ Finally, alternative strategies such as the use of isotopically labelled linkers or linkers with reporter tag motifs, which show characteristic mass differences during MS analysis, can improve linked peptide identification.¹⁶

The structural diversity of proteins also creates unique challenges for each individual protein system. For example, having access to only a few reactive amino acids can conceal important structural information. Differences in the amounts and types of available amino acids necessitates different linker reactivity, while larger distances between reactive groups may require different lengths for spacer arms. This has driven the development of new reactive groups like aromatic glyoxal cross-linkers (ArGOs), dihydrazide sulfoxide and 1,1'-carbonyldiimidazole, capable of reacting with arginine, acidic amino acids and hydroxyl groups, respectively.¹⁷⁻¹⁹ While the wealth of new cross-linker options is beneficial to the field, in most applications, these are typically confined to commercially available compounds that lack flexibility in terms of spacer arm length and the analytical features described above.

The ability to easily access a diverse range of cross-linking reagents with different reactivities, structures and properties would greatly expand the capabilities of XL-MS. This study describes the development of a modular synthetic strategy, which allows for the straight-forward production of cross-linkers that can incorporate various functionalities, including enrichment and fluorescent tags, varied-length or CID-cleavable spacer arms and hetero- or homobifunctional reactive groups, to enable researchers to readily adapt linker selection to the specific needs of the experiment or system of interest. We exemplify the synthetic utility of this strategy through construction of a small linker library and demonstrate application of the linkers to example peptide and protein systems.

Results and Discussion

Development of a modular synthetic protocol

To address the lack of diversity in commercially available cross-linker options, a modular synthetic protocol was developed based around amino-acid building blocks to ensure the components were readily available, low cost and easily adaptable for functional group tuneability (**Fig. 2**). The synthesis was designed to allow for reactive groups and spacer arms that can be exchanged depending on the demands of the experiment, while also including a motif to which enrichment or luminescent tags (or other functionalities) can be attached after cross-linking.

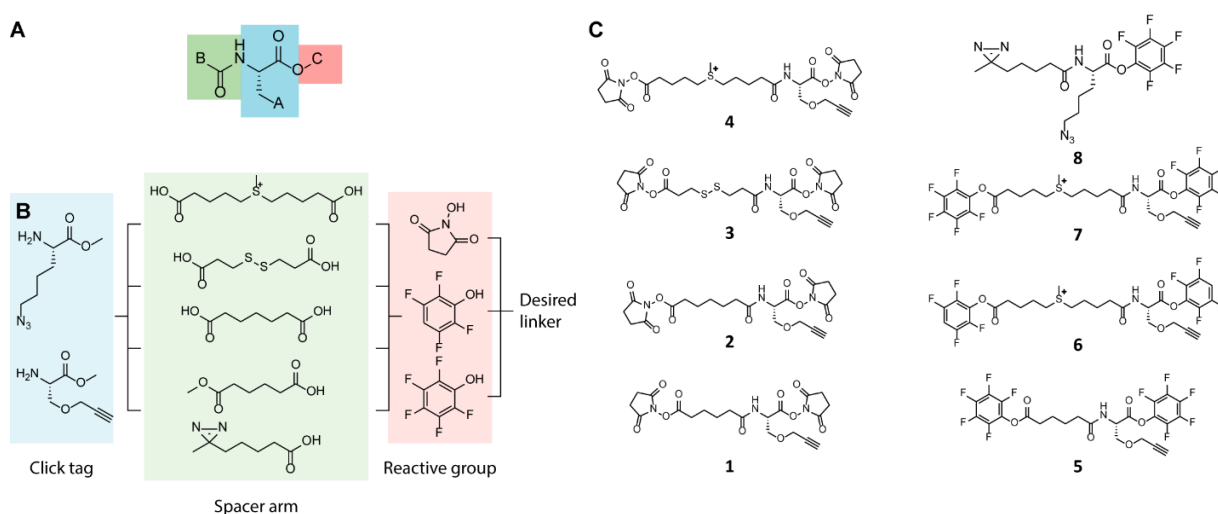
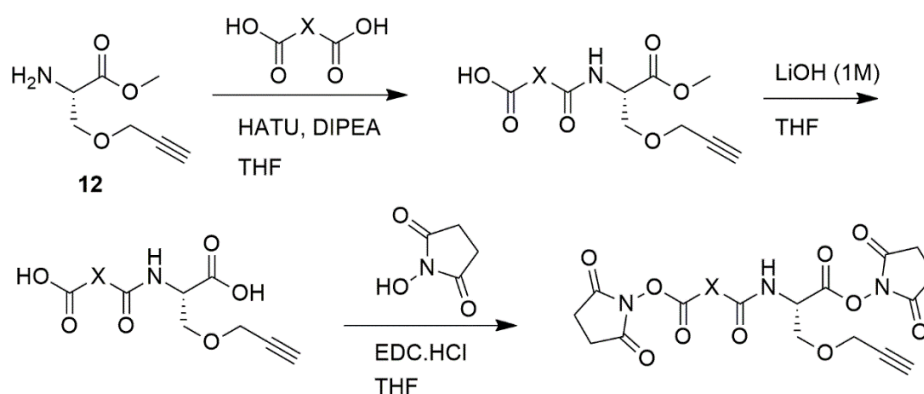


Figure 2. (A) The general structure of the modular linker includes 3 components: the 'click'-modified amino acid core (blue), the spacer arm (green) and the reactive group (red), which can be synthesised in a combinatorial fashion (B). (C) Structures contained in the cross-linker library synthesised using the modular protocol.

At the core of the synthetic strategy is a base unit that includes a 'click' tag capable of undergoing derivatisation under biocompatible conditions. This click chemistry is mediated by an alkyne or azide tag which allows for the inclusion of post-linkage modifications via a copper-catalysed Huisgen cycloaddition. Currently, there are numerous examples of alkyne/azide reagents that are commercially available, including modifications such as dyes, FLAG tags and agarose or magnetic beads. Two base units were proposed with an alkyne and an azide click tag (**Fig. 2B**), which allow for the same sample to be aliquoted and derivatised with different modifications, which would be impossible if modifications were introduced to the linker prior to cross-linking. The alkyne base unit (**12**) was synthesised by *O*-alkylation of commercially available *tert*-butyloxycarbonyl (*N*-Boc)-protected serine with propargyl bromide, followed by a one pot *N*-Boc cleavage and methyl esterification (**Supp. Scheme 1**). The azide base unit (**10**) was synthesized by a diazo transfer reaction with an *N*-Boc protected lysine, followed by the same one pot *N*-Boc cleavage and methyl esterification reaction (**Supp. Scheme 1**).

Following synthesis of the two base units, all subsequent reactions followed the same general reaction scheme wherein in the spacer arm is coupled to the amino acid base unit by a peptide coupling, followed by a methyl ester hydrolysis and finally coupling of the reactive groups to the exposed acids (**Scheme 1**). For the easy incorporation of different spacer arms, the synthesis was designed around a hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU)-mediated peptide coupling so that any spacer arm with a carboxylic acid could be incorporated into the final compound (**Fig. 2B**). This allows different spacer arm lengths to be chosen by selecting the appropriate length diacid for coupling. In this study two examples were chosen, monomethyl adipate and heptanedioic acid, to showcase the simplicity by which different carbon-chain lengths can be included.



Scheme 1. General synthetic protocol highlighting the modularity of the approach. Introducing a diacid and reactive group of choice allows for synthesis of a custom linker.

To demonstrate the incorporation of gas-phase labile bonds into the cross-linker spacer arm, two additional diacids were chosen for coupling: 5-methyl 5,5'-thiodipentanoic acid, capable of fragmentation under low-energy CID conditions in positive-ion mode,²⁰ and 3,3'-dithiopropionic acid, capable of facile fragmentation under negative-ion CID conditions.²¹ It is possible to perform XL-MS analysis under both positive and negative ion MS conditions, and therefore it is advantageous to incorporate moieties that undergo selective fragmentation in both of these modes. For linkers **4**, **6** and **7** an additional step was necessary wherein the penultimate NHS-coupled compound undergoes methylation with methyl iodide to form the fixed-charge sulfonium ion required for positive-ion mode fragmentation.

Following peptide coupling of the spacer arm, alkaline hydrolysis of the methyl ester yielded two free carboxylic acids with which a variety of reactive groups are available for *N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC-HCl) coupling to generate the final reactive linker. Further customisation is possible to introduce UV-reactive heterobifunctional cross-linking through the coupling of 3-methyl-diazirine-3-propanoic acid, instead of a diacid, to yield linker **8**. Subsequent steps follow the same protocol, with the exception of reducing the equivalents of NHS coupling reagents to reflect that only one NHS is coupled to the final linker.

Based on this general synthetic protocol, a small library of 8 linkers was synthesised with differing spacer-arm lengths, fragmentable moieties and reactive groups, all of which included an alkyne or azide tag available for further modification post-linkage (**Fig. 2C**).

Cross-linking and fragmentation characterisation

To demonstrate the reactivity of the synthesised library, a selection of amine-reactive linkers was applied to a test peptide, acetylated AAKA (AcAAKA), which is a simple model system with a single primary amine available for cross-linking. Incubation of cross-linkers with AcAAKA at a high concentration allows for the simulation of the peptide linkages that could be expected in a proteolytic digest. Initially, compound **1** was used to form a cross-link between AcAAKA peptides and MS/MS analysis was performed on the cross-linked peptide, the structure of which was confirmed by the expected fragmentation products (**Supp. Fig. 1**). Following this, the facile fragmentation of CID-cleavable linkers was validated using a negative- and positive-ion-mode cleavable linker (compounds **3** and **4** respectively). Both compounds were reacted with AcAAKA and MS/MS analysis revealed spectra consistent with the expected dissociation patterns (**Fig. 3**). Compound **4** produced 2 major fragmentation products separated by 125 *m/z*, suggesting formation of a 6-membered oxazoline ring under low-energy CID, preferring the non-serine side (**Fig. 3A**).²⁰ Capable instruments may then scan for this characteristic 125 *m/z* doublet during MS/MS under lower fragmentation energies and isolate each ion for higher-energy MS sequencing experiments, allowing each linked peptide to be more easily identified and then sequenced, greatly improving site-specific localisation of the modification.²² Similarly, the negative-ion-mode linker cleaved as expected, forming a characteristic quartet fragmentation pattern, also suited to subsequent MS³ sequencing experiments.²³ The facile fragmentation about the disulfide bond in the linker chain is effected by either an enolate anion or by an anion situated directly adjacent to the disulfide.²⁴ There are therefore four potential possible products as fragmentation can take place on the serine or non-serine side of the disulfide bond, although the relative abundance of these signals in the MS/MS spectra of the linked AcAAKA peptides did not reveal significant preference for either side in this example (**Fig. 3B**).

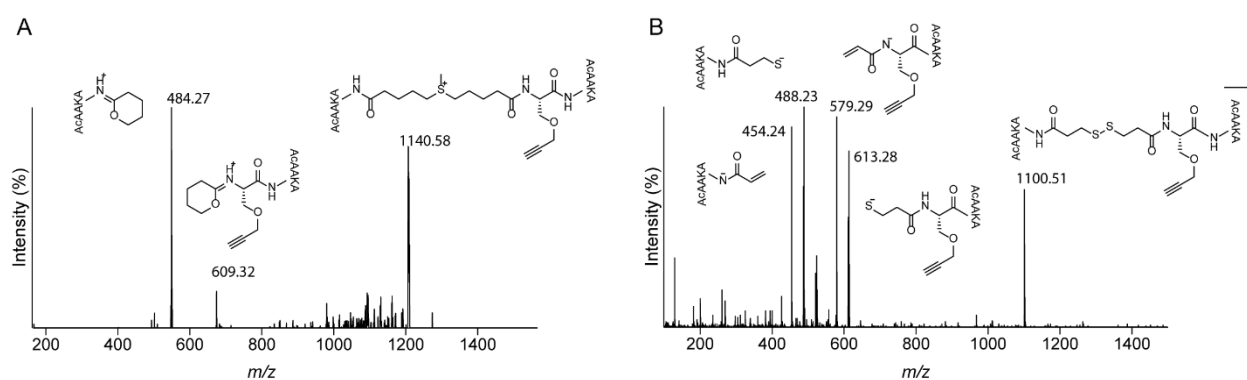


Figure 3. (A) Low-energy CID (20–40 V) of linker **3** cross-linked to AcAAKA yield two major fragment ions corresponding to 6-member oxazoline rings on the non-serine (484.27 *m/z*) or serine side (609.32 *m/z*) from the precursor ion (1140.58 *m/z*). (B) The linker **4** precursor

ion (1100.51 m/z) yielded 4 fragment ions upon CID at a low energy: an enone product on the non-serine (454.24 m/z) and serine (579.29 m/z) side and a thiol product on the non-serine (488.23 m/z) and serine (613.28 m/z).

Overall, these results confirm that the alkyne-modified serine is a suitable base unit to combine with diacid spacer arms in order to successfully form covalent links between lysine residues, and the resultant linkers are able to effectively fragment under CID conditions.

Post cross-link derivatisation of the modular linker

A feature of the modular linker design is the included alkyne/azide tag, which can be derivatised subsequent to the reaction with proteins. This is advantageous since it avoids the inclusion of bulky groups such as biotin in the reactive linker, which can potentially restrict access to reactive sites on the protein during the cross-linking reaction. It also allows for flexibility in the choice of 'clickable' motif to include a range of functionalities. Alkyne and azide groups on biotin, fluorescent tags, agarose UV-cleavable beads, peptides and oligos, amongst many other examples, are becoming increasingly available commercially, and the flexibility to apply several tags to a single cross-linked sample is attractive. After validation of covalent linkage, the ability for post-cross-linking derivatisation of the linkers through a copper-catalysed Huisgen cycloaddition reaction was demonstrated with both biotin and cyanine 3 (Cy3) for enrichment and fluorescent visualisation, respectively.

Following covalent linkage between compound **1** and AcAAKA, a simple copper catalysed click reaction facilitated the addition of biotin azide. Evidence for the successful derivatisation was provided by MS, with MS/MS analysis of the expected $(M+H)^+$ ion of the product at 1294 m/z giving rise to the expected fragment ions (**Fig. 4A**). Biotin affinity purification following this reaction allows for enrichment in a manner which is familiar and available to many biochemistry laboratories.

The flexibility offered by post-cross-linking derivatisation was further showcased using two different linkers, compound **5** and **6**. Both compounds and a commercially available control, disuccinimidyl sulfoxide (DSSO), were used to cross-link hen egg lysozyme and the resulting products were separated by SDS-PAGE and visualised using Coomassie Brilliant Blue stain (**Fig 4B**). Modification of the lysozyme monomer by cross-linkers was confirmed by MS (**Supp. Fig. 2**). Despite differences in linker length that may affect lysine accessibilities, both custom linkers showed comparable cross-linking efficiency with the DSSO control, and dimer band densities were reflective of the monomer to dimer ratio in buffered solution.²⁵ Before SDS-PAGE, a sample of protein linked with compounds **5** and **6** was further covalently modified with Cy3-azide by copper-catalysed cycloaddition, and modifications were evidenced by an upward band shift in the gel due to the added mass of the cy3 tag. The gels were then visualised by Cy3 fluorescence emission at 556 nm, following excitation at 553 nm. To emphasize how fluorescence tags could be used to increase visibility of hidden covalently modified protein complexes, smaller amounts of Cy3-modified proteins were loaded onto the gel to reduce the dimer intensity below limits of detection of the Coomassie stain (**Fig 4B**). Notably, the band width of lysozyme-Cy3-linker complexes increased, giving an indication of heterogeneous mass changes due to attached cy3 tags. Upon

irradiation and fluorescence detection, the dimeric species that were undetectable using Coomassie staining alone were revealed by the cy3 dye.

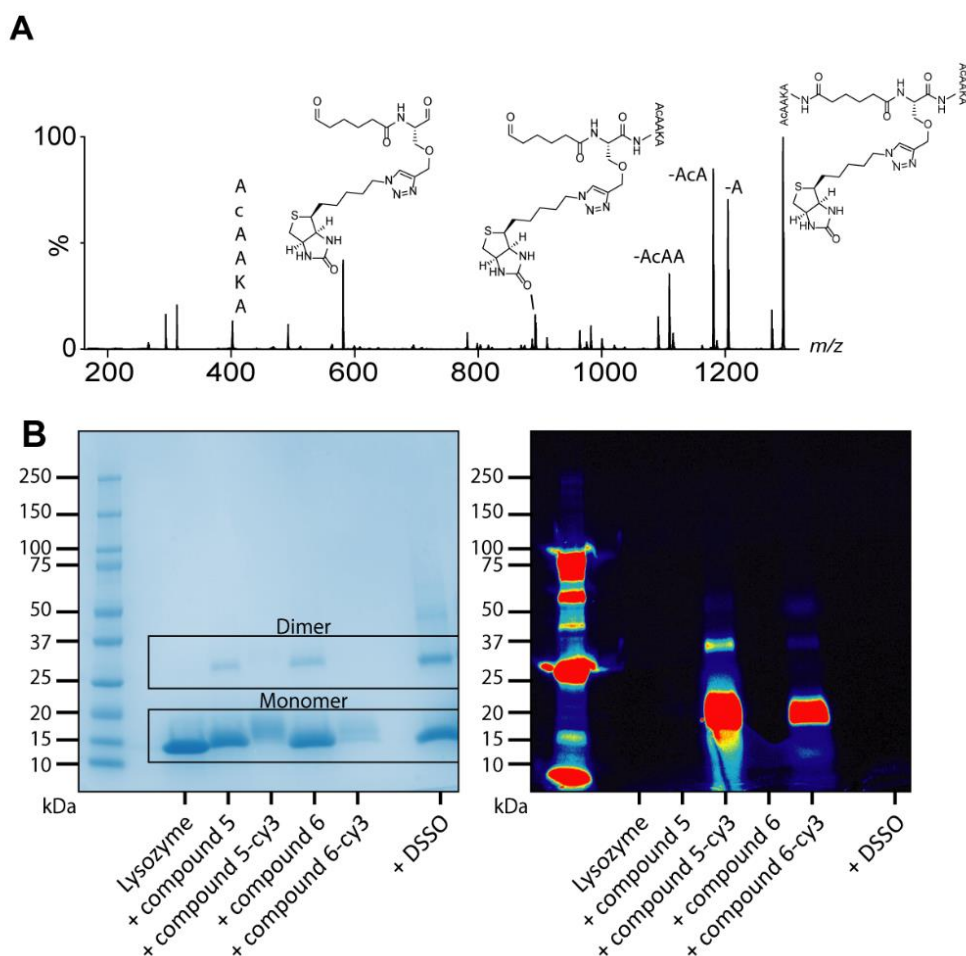


Figure 4. (A) CID was used to characterise cross-linked AcaAKA using linker **1** followed by post-linkage modification with biotin-azide using a Cu-catalysed click cycloaddition reaction. The precursor ion was identified (1293 m/z) and MS/MS analysis revealed characteristic fragment ions at 892 and 581 m/z that represent the precursor ion minus one and two AcaAKA peptides, respectively. (B, left) SDS-PAGE revealed successful covalent linkage of lysozyme dimers using compounds **5** and **6** by comparison to a commercially available linker (DSSO), as well as post-linkage modification with Cy3 (B, right). Fluorescence imaging highlights the ability to uncover hidden oligomers through dye derivatisation of linked samples.

The application of XL-MS has in recent years gone beyond *in vitro* samples and small numbers of interacting proteins, with cellular and tissue samples now subject to XL-MS for proteome-wide interactomic analysis. The avenue of *in vivo* cross-linking and the increased complexity of these samples has led to a greater dependence on enrichment tags and cleavable spacer arms. For example, cross-linking of mammalian cells with azide-A-DSBSO, a cross-linker capable of undergoing biotin click chemistry, was able to enrich and sequence 136 intrasubunit and 104 intersubunit protein cross-links.²⁶ Following this, new enrichment techniques allow for covalent linkage directly to alkyne or azide-containing beads, skipping biotin-streptavidin enrichment techniques and improving enrichment efficiency by up to 5 times.²⁷ While impressive, the structural information obtained could be expanded upon further by introducing different spacer arm lengths and reactive motifs using the

modular approach presented. While this study explores 4 potential reactive groups, cysteine-reactive bismaleimide and acid-reactive dihydrazide groups offer alternative linking chemistry that can be easily derived in one step from NHS-containing linkers as shown by Gutierrez, *et al.*^{7,18} Furthermore, the wide availability of isotopically labelled amino acids would allow for the simple inclusion of a heavy serine/lysine base unit for synthesis, potentially yielding greater cross-link identification and even relative quantitation via the detection of characteristic isotope patterns during MS.¹⁶

Conclusion

This study presents a modular synthetic protocol for the synthesis of protein cross-linking reagents capable of tackling the individual challenges associated with protein structure determination by XL-MS. The ability to exchange reactive groups and spacer arms as needed was showcased through the synthesis of a library containing 8 unique cross-linkers. The reagents analysed here displayed covalent linkages with model peptides and lysozyme with a comparable efficiency to that of commercially available cross-linkers. The introduction of cleavable spacer arms allows for improved identification strategies through diagnostic fragmentation peaks. In addition, centring the synthesis around an alkyne-/azide-modified amino acid allows for introduction of additional functionality through copper click chemistry, as we demonstrated through post-linkage modification of biotin for avidin-affinity purification or Cy3 for improved visualisation through fluorescence. Overall, this synthetic protocol aims to reduce the complexity of XL-MS by offering a scaffold for greatly improving the diversity of potential cross-linking reagents, intentionally designed for the highly variable requirements of the field.

Materials and Methods

Materials and reagents

Chemicals were purchased from Merck (Kenilworth, NJ, USA) or AK Scientific (Union City, CA, USA). The peptide AcAAKA was synthesised in-house using standard Fmoc solid-phase methods on 2-chlorotrityl chloride resin (GL Biochem, Shanghai, China) and purified by high-performance liquid chromatography (HPLC) to greater than 95% purity as described previously.²⁴

Synthesis of cross-linkers

All linkers were synthesised from a general synthetic procedure beginning with amide coupling between the desired diacid spacer arm (this includes coupling with 3-methyl-diazirine-3-propanoic acid) and either the alkyne base unit (**12**) or the azide base unit (**10**), followed by deprotection by methyl ester hydrolysis. Finally, esterification with N-hydroxysuccinimide (NHS), tetrafluorophenol (TFP) or pentafluorophenol (PFP) and EDC-HCl yielded the final linkers (**Fig. 2C**). 500 MHz ¹H and ¹³C nuclear magnetic resonance (NMR) spectra were obtained using an Agilent 500/54 Premium Shielded NMR spectrometer (Agilent Technologies, Santa Clara, CA, USA). High-resolution mass spectrometry (HRMS) data were obtained using an Agilent 6230 TOF LC/MS equipped with an Infinity 1260 LC system (Agilent Technologies). MS/MS data were obtained using a Micromass QTOF2 mass spectrometer (Milford, MA, USA).

Synthesis of compound 10. This procedure was based on a previously published synthesis.²⁹ Triflic anhydride (25 mL, 146 mmol) was added to a solution of sodium azide (26.6 g, 410 mmol) in water (100 mL) while stirring at 0 °C. The reaction was allowed to warm to room temperature while stirring for 2 h, then extracted with dichloromethane (2 x 50 mL) and the combined organic extracts washed with saturated Na₂CO₃ (100 mL). The organic extract was added dropwise to a stirred solution of *N*- α -Boc-L-lysine (10.1 g, 41.1 mmol), CuSO₄·5H₂O (1.11 g, 4.45 mmol) and K₂CO₃ (8.68 g, 62.8 mmol) in 2:1 (v/v) methanol/water (450 mL) at 0 °C. The solvent was removed under reduced pressure and the resulting residue diluted with water (75 mL) and acidified by dropwise addition of 6 N HCl until the formation of a precipitate. The suspension was then diluted with potassium phosphate buffer (pH 6.2, 150 mL of 1:1 (v/v) 0.25 M K₂HPO₄/ 0.25 M KH₂PO₄) before further acidifying to pH 3 with 6 N HCl. The mixture was then extracted with ethyl acetate (2 x 150 mL). The combined organic extract was washed with water (100 mL) and brine (100 mL), then dried (NaSO₄) and the solvent removed under reduced pressure. The residue was dissolved chloroform (100 mL), filtered by suction filtration and the solvent removed under reduced pressure to obtain a yellow oil (10.4 g, 93%). ¹H NMR (500 MHz, CD₃OD, δ): 4.11-4.08 (dd, *J* = 5.0 Hz, 1H), 3.33-3.30 (m, 2H), 1.88-1.81 (m, 1H), 1.72-1.64 (m, 1H), 1.65-1.60 (m, 2H), 1.54-1.47 (m, 2H), 1.46 (s, 9H). ¹³C NMR (125 MHz, CD₃OD, δ): 176.2, 158.1, 80.5, 79.4, 54.7, 52.2, 32.4, 29.4, 28.7, 24.1. Compound **9** (2.04 g, 7.5 mmol) was dissolved in dry methanol (75 mL) at 0°C while stirring under N₂ atmosphere. Thionyl chloride (1 mL, 15.0 mmol) was added dropwise to the stirred solution and the reaction allowed to warm to room temperature for 24 h. The mixture was concentrated *in vacuo* and the residue co-evaporated with methanol (3 x 15 mL). Lastly, the residue was redissolved in chloroform, filtered and the solvent

removed under reduced pressure to obtain a pale-yellow oil (1.6 g, 115 % with residual solvent, used without further purification). ¹H NMR (500 MHz, CD₃OD, δ): 4.07 (t, *J* = 6.4 Hz, 1H), 3.86 (s, 3H), 3.36 (t, *J* = 6.6 Hz, 2H), 3.32 (p, *J* = 1.7 Hz, 1H), 1.94 (m, *J* = 19.7, 16.5, 14.2, 8.9, 5.6 Hz, 2H), 1.65 (p, *J* = 7.0 Hz, 2H), 1.61-1.53 (m, 1H), 1.53-1.42 (m, 1H). HRMS (ESI) *m/z*: M⁺ calculated for C₇H₁₄N₄O₂, 186.1117, found [M+H]⁺: 187.1193.

Synthesis of compound 12. Compound **12** was synthesised by dissolving Boc-L-Serine (2.01 g, 9.7 mmol) in DMF (100 mL) and stirring on ice for 15 min. A 60 % w/w dispersion of sodium hydride (995 mg, 24.8 mmol) was added and the solution was stirred on ice for 45 min. An 80 % solution of propargyl bromide (1.2 mL, 10.7 mmol) was added dropwise to the solution and the reaction stirred at 0°C for 15 min. The reaction mixture was allowed to warm to room temperature and then stir overnight under a nitrogen atmosphere. Water (15 mL) was added and stirred for 5 minutes. The solution was washed with diethyl ether (20 mL), acidified (pH <3) with 1 M KHSO₄ (95 mL) and then extracted with ethyl acetate (3 x 30 mL). The combined extracts were washed with water (3 x 100 mL) and dried over Na₂SO₄. The solvent was removed *in vacuo* to yield **11** as a yellow oil which was used without further purification (2.16 g, 91 %). ¹H NMR (500 MHz, DMSO-*d*₆, δ): 5.38 (d, *J* = 8.2 Hz, 1H), 4.49 (d, *J* = 8.1 Hz, 1H), 4.23-4.14 (m, 2H), 4.00 (dd, *J* = 8.8, 2.2 Hz, 1H), 3.81 (dd, *J* = 9.4, 3.5 Hz, 1H), 2.47 (t, *J* = 2.3 Hz, 1H), 1.46 (s, 9H). ¹³C NMR (125 MHz, CDCl₃, δ): 175.1, 155.9, 80.6, 78.8, 75.4, 69.6, 58.8, 53.8, 28.41. All NMR data is consistent with literature.²⁸ The acid (**11**) (2.00 g, 8.22 mmol) was dissolved in methanol (100 mL) and stirred on ice for 15 min. Thionyl chloride was then added dropwise (1.20 mL, 16.4 mmol) under a nitrogen atmosphere. The solvent was removed *in vacuo* and the residue was co-evaporated with methanol three times to yield compound **12** (1.64 g, 97 %) as an orange oil. ¹H NMR (500 MHz, DMSO-*d*₆, δ): 8.64 (br s, 3H), 4.37 (t, *J* = 3.3 Hz, 1H), 4.24 (dd, *J* = 16.1, 2.2 Hz, 1H), 4.20 (dd, *J* = 16.1, 2.2 Hz, 1H), 3.91 (dd, *J* = 10.5, 4.1 Hz, 1H), 3.85 (dd, *J* = 10.6, 3.0 Hz, 1H), 3.76 (s, 3H) 3.56 (t, *J* = 2.1 Hz, 1H). ¹³C NMR (125 MHz, CD₃OD, δ): 168.7, 79.4, 77.2, 67.6, 59.4, 54.3, 53.9. All NMR data is consistent with literature.³⁰ HRMS (ESI⁺) calculated for C₇H₁₁NO₃ M⁺ = 157.0739, found [M+H]⁺: 158.0359.

General modular synthetic protocol. **10** or **12** was dissolved in dry tetrahydrofuran (5 mL per 100 mg) with *N,N*-diisopropylethylamine (DIEA, 4.0 mol eq) while stirring. The desired spacer arm diacid (1.0 mol eq) was added to the solution followed by HATU (1.2 mol eq). The mixture was stirred overnight at room temperature under a nitrogen atmosphere. 1 M HCl (5 mL per 100 mg) was added and the solution was extracted with ethyl acetate (1.5 mL per 100 mg, 3x). The organic phase was washed with water (3 x) and brine then dried over Na₂SO₄ and the solvent was removed *in vacuo*. The product was then purified by flash chromatography on normal phase silica to yield the desired amide.

To hydrolyse the methyl ester, the amide was dissolved in THF (1 mL per 100 mg) and stirred on ice for 15 minutes. 1 M LiOH (4 mol eq.) was added and the reaction was stirred until completion as determined by thin layer chromatography. The solvent was removed *in vacuo* and the residue was partitioned between dichloromethane and water. The organic layer was isolated and the aqueous layer was extracted with dichloromethane (2x). The combined organic extracts were washed with brine and dried over Na₂SO₄. The solvent was removed *in vacuo* to yield the desired diacid as a yellow oil.

The diacid was dissolved in anhydrous THF (10 mL per 100 mg) with EDC-HCl (3.4 mol eq) and stirred at room temperature for 10 minutes. NHS, TFP or PFP (6 eq.) were added and the reaction was stirred at room temperature under a nitrogen atmosphere overnight. The solvent was removed *in vacuo* and the resulting residue was partitioned between dichloromethane and water. The aqueous layer was extracted with dichloromethane (3x) and the combined organic extracts were dried over MgSO₄ and filtered. The solvent was removed from the filtrate *in vacuo* to afford the ester, which was used for cross-linking reactions without further purification.

Synthesis of compounds 4, 6 and 7. Synthesis of positive mode cleavable linkers were based on work by Lu *et al.*²⁰ Following coupling with 5,5'-thiodipentanoic acid and the corresponding reactive group, the thioether was dissolved with methyl iodide (2.2 eq.) in dry acetonitrile (1.5 mL) and allowed to react while stirring at room temperature for 4 days. The solvent was then removed *in vacuo* yielding a yellow oil which was used for cross-linking reactions without further purification.

Cross-linking of AcAAKA. Each linker was dissolved in DMSO to 10 mM with AcAAKA in a 1:1 linker:peptide ratio with 1 eq. of DIEA. The mixture was incubated at room temperature for 1 hour. The mixture was then diluted to 10 μM in 50:50 water:acetonitrile for analysis by negative ion mode tandem MS and with 0.1 % aqueous formic acid for analysis by positive ion mode tandem MS.

Cross-linking of hen egg lysozyme. Lysozyme was dissolved in PBS to a concentration of 1 mg/mL. The relevant cross-linker was dissolved in DMSO (10 mM) and added to lysozyme solution at a ratio of 20:1 linker:lysozyme ratio. Protein was used for SDS-PAGE immediately or was buffer exchanged into ammonium acetate (100 mM) using a 10 kDa MWCO Amicon Ultra-0.5 Centrifugal Filter Device (Millipore, Jaffrey, NH, USA).

Copper click reaction. The general procedure for CuAAC reactions was performed using a method adapted from a protocol available from Broadpharm.³¹ Peptide/protein conjugated with azide/alkyne linker (1mg/mL, 20 μL) was diluted with PBS (90 μL), followed by addition of aqueous solutions of 2.5 mM biotin-azide (20 μL) or 2.5 mM Cy3-azide, 100 mM tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (BTAA, 10 μL; Click Chemistry Tools, Scottsdale, AZ, USA), 20 mM copper sulfate (10 μL) and 300 mM sodium ascorbate (10 μL). The reaction was incubated at room temperature for 30 min, then purified and concentrated into PBS (100 μL) using a 10 kDa MWCO Amicon Ultra-0.5 Centrifugal Filter Device.

SDS-polyacrylamide gel electrophoresis. Lysozyme was cross-linked and examined by SDS-PAGE (12% gels) (Bio-rad, Hercules, CA, USA). Bands were visualised using Coomassie Brilliant Blue stain (ThermoFisher, Waltham, MA, USA) according to the manufacturer's instructions. Fluorescence imaging was performed on a ChemiDoc gel image system (Bio-rad) using an excitation wavelength of 553 nm and emission was measured at 556 nm.

Acknowledgements

HMS and KGS were supported by a Faculty of Sciences Divisional Scholarship from the University of Adelaide. This work was supported in part by a Discovery Project Grant awarded to TLP by the Australian Research Council (DP170102033).

Competing Interests

The authors declare that they have no competing interests associated with the contents of this manuscript.

Author Contributions

HMS, KGS, ADA, and TLP designed the research; HMS, ERB, KGS and KMD performed the experiments; HMS and TLP analysed experimental data; HMS, ADA and TLP prepared the manuscript.

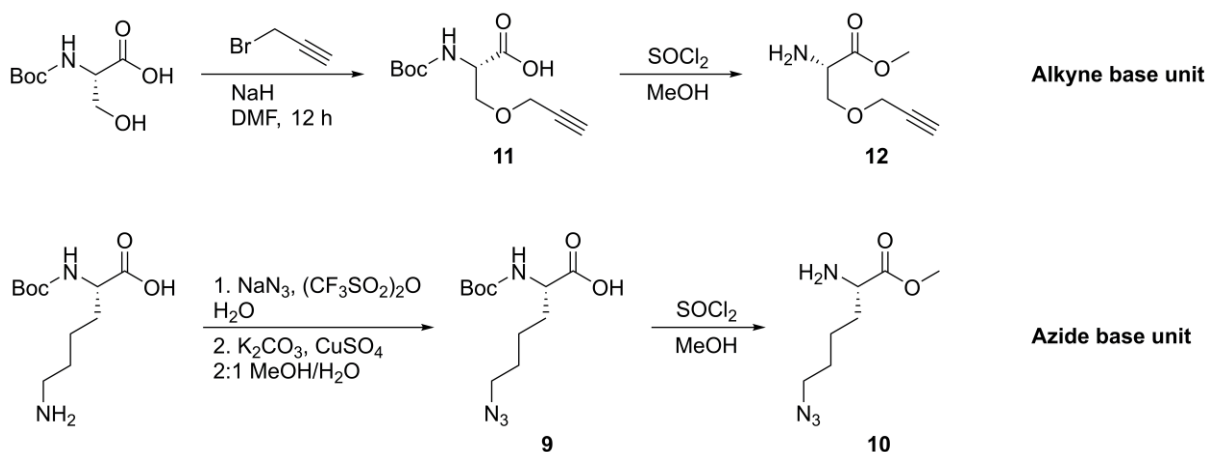
References

1. Politis, A.; Stengel, F.; Hall, Z.; Hernández, H.; Leitner, A.; Walzthoeni, T.; Robinson, C. V.; Aebersold, R. A Mass Spectrometry–Based Hybrid Method for Structural Modeling of Protein Complexes. *Nat. Methods* 2014, 11 (4), 403–406. <https://doi.org/10.1038/nmeth.2841>.
2. Schmidt, C.; Zhou, M.; Marriott, H.; Morgner, N.; Politis, A.; Robinson, C. V. Comparative Cross-Linking and Mass Spectrometry of an Intact F-Type ATPase Suggest a Role for Phosphorylation. *Nat. Commun.* 2013, 4. <https://doi.org/10.1038/ncomms2985>.
3. O'Reilly, F. J.; Rappsilber, J. Cross-Linking Mass Spectrometry: Methods and Applications in Structural, Molecular and Systems Biology. *Nat. Struct. Mol. Biol.* 2018, 25 (11), 1000–1008. <https://doi.org/10.1038/s41594-018-0147-0>.
4. Nguyen, V. Q.; Ranjan, A.; Stengel, F.; Wei, D.; Aebersold, R.; Wu, C.; Leschziner, A. E. Molecular Architecture of the ATP-Dependent Chromatin-Remodeling Complex SWR1. *Cell* 2013, 154 (6), 1220–1231. <https://doi.org/10.1016/j.cell.2013.08.018>.
5. Liu, F.; Rijkers, D. T. S.; Post, H.; Heck, A. J. R. Proteome-Wide Profiling of Protein Assemblies by Cross-Linking Mass Spectrometry. *Nat. Methods* 2015, 12 (12), 1179–1184. <https://doi.org/10.1038/nmeth.3603>.
6. Chavez, J. D.; Weisbrod, C. R.; Zheng, C.; Eng, J. K.; Bruce, J. E. Protein Interactions, Post-Translational Modifications and Topologies in Human Cells. *Mol. Cell. Proteomics* 2013, 12 (5), 1451–1467. <https://doi.org/10.1074/mcp.M112.024497>.
7. Gutierrez, C. B.; Block, S. A.; Yu, C.; Soohoo, S. M.; Huszagh, A. S.; Rychnovsky, S. D.; Huang, L. Development of a Novel Sulfoxide-Containing MS-Cleavable Homobifunctional Cysteine-Reactive Cross-Linker for Studying Protein-Protein Interactions. *Anal. Chem.* 2018, 90 (12), 7600–7607. <https://doi.org/10.1021/acs.analchem.8b01287>.
8. Leitner, A.; Faini, M.; Stengel, F.; Aebersold, R. Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem. Sci.* 2016, 41 (1), 20–32. <https://doi.org/10.1016/j.tibs.2015.10.008>.
9. Steigenberger, B.; Pieters, R. J.; Heck, A. J. R.; Scheltema, R. A. PhoX: An IMAC-Enrichable Cross-Linking Reagent. *ACS Cent. Sci.* 2019, 5 (9), 1514–1522. <https://doi.org/10.1021/acscentsci.9b00416>.
10. Petrotchenko, E. V.; Serpa, J. J.; Borchers, C. H. An Isotopically Coded CID-Cleavable Biotinylated Cross-Linker for Structural Proteomics. *Mol. Cell. Proteomics* 2011, 10 (2), 1–8. <https://doi.org/10.1074/mcp.M110.001420>.
11. Vellucci, D.; Kao, A.; Kaake, R. M.; Rychnovsky, S. D.; Huang, L. Selective Enrichment and Identification of Azide-Tagged Cross-Linked Peptides Using Chemical Ligation and Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 2010, 21 (8), 1432–1445. <https://doi.org/10.1016/j.jasms.2010.04.004>.
12. Sibbersen, C.; Lykke, L.; Gregersen, N.; Jørgensen, K. A.; Johannsen, M. A Cleavable Azide Resin for Direct Click Chemistry Mediated Enrichment of Alkyne-Labeled Proteins. *Chem. Commun.* 2014, 50 (81), 12098–12100. <https://doi.org/10.1039/c4cc05246c>.
13. Müller, M. Q.; Dreiocker, F.; Ihling, C. H.; Schäfer, M.; Sinz, A. Cleavable Cross-Linker for Protein Structure Analysis: Reliable Identification of Cross-Linking Products by Tandem MS. *Anal. Chem.* 2010, 82 (16), 6958–6968. <https://doi.org/10.1021/ac101241t>.
14. Kao, A.; Chiu, C.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L. Development of a Novel Cross-Linking Strategy for Fast and Accurate Identification of Cross-Linked Peptides of Protein Complexes. *Mol. Cell. Proteomics* 2011, 10 (1), M110.002212. <https://doi.org/10.1074/mcp.m110.002212>.

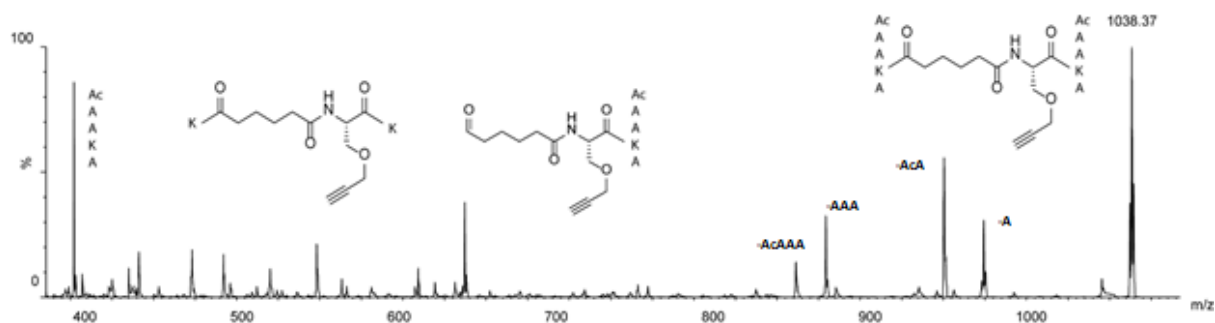
15. Iacobucci, C.; Götze, M.; Ihling, C. H.; Piotrowski, C.; Arlt, C.; Schäfer, M.; Hage, C.; Schmidt, R.; Sinz, A. A Cross-Linking/Mass Spectrometry Workflow Based on MS-Cleavable Cross-Linkers and the MeroX Software for Studying Protein Structures and Protein–Protein Interactions. *Nat. Protoc.* 2018, 13 (12), 2864–2889. <https://doi.org/10.1038/s41596-018-0068-8>.
16. Ihling, C. H.; Springorum, P.; Iacobucci, C.; Hage, C.; Götze, M.; Schäfer, M.; Sinz, A. The Isotope-Labeled, MS-Cleavable Cross-Linker Disuccinimidyl Dibutyric Urea for Improved Cross-Linking/Mass Spectrometry Studies. *J. Am. Soc. Mass Spectrom.* 2020, 31 (2), 183–189. <https://doi.org/10.1021/jasms.9b00008>.
17. Jones, A. X.; Cao, Y.; Tang, Y. L.; Wang, J. H.; Ding, Y. H.; Tan, H.; Chen, Z. L.; Fang, R. Q.; Yin, J.; Chen, R. C.; et al. Improving Mass Spectrometry Analysis of Protein Structures with Arginine-Selective Chemical Cross-Linkers. *Nat. Commun.* 2019, 10 (1), 3911. <https://doi.org/10.1038/s41467-019-11917-z>.
18. Gutierrez, C. B.; Yu, C.; Novitsky, E. J.; Huszagh, A. S.; Rychnovsky, S. D.; Huang, L. Developing an Acidic Residue Reactive and Sulfoxide-Containing MS-Cleavable Homobifunctional Cross-Linker for Probing Protein–Protein Interactions. *Anal. Chem.* 2016, 88 (16), 8315–8322. <https://doi.org/10.1021/acs.analchem.6b02240>.
19. Hage, C.; Iacobucci, C.; Rehkamp, A.; Arlt, C.; Sinz, A. The First Zero-Length Mass Spectrometry-Cleavable Cross-Linker for Protein Structure Analysis. *Angew. Chemie Int. Ed.* 2017, 56 (46), 14551–14555. <https://doi.org/10.1002/anie.201708273>.
20. Lu, Y.; Tanasova, M.; Borhan, B.; Reid, G. E. Ionic Reagent for Controlling the Gas-Phase Fragmentation Reactions of Cross-Linked Peptides. *Anal. Chem.* 2008, 80 (23), 9279–9287. <https://doi.org/10.1021/ac801625e>.
21. Lomant, A. J.; Fairbanks, G. Chemical Probes of Extended Biological Structures: Synthesis and Properties of the Cleavable Protein Cross-Linking Reagent [35S]Dithiobis(Succinimidyl Propionate). *J. Mol. Biol.* 1976, 104 (1), 243–261. [https://doi.org/10.1016/0022-2836\(76\)90011-5](https://doi.org/10.1016/0022-2836(76)90011-5).
22. Liu, F.; Lössl, P.; Scheltema, R.; Viner, R.; Heck, A. J. R. Optimized Fragmentation Schemes and Data Analysis Strategies for Proteome-Wide Cross-Link Identification. *Nat. Commun.* 2017, 8 (1), 15473. <https://doi.org/10.1038/ncomms15473>.
23. Calabrese, A. N.; Wang, T.; Bowie, J. H.; Pukala, T. L. Negative Ion Fragmentations of Disulfide-Containing Crosslinking Reagents Are Competitive with Aspartic Acid Side-Chain-Induced Cleavages. *Rapid Commun. Mass Spectrom.* 2013, 27 (1), 238–248. <https://doi.org/10.1002/rcm.6445>.
24. Calabrese, A. N.; Good, N. J.; Wang, T.; He, J.; Bowie, J. H.; Pukala, T. L. A Negative Ion Mass Spectrometry Approach to Identify Cross-Linked Peptides Utilizing Characteristic Disulfide Fragmentations. *J. Am. Soc. Mass Spectrom.* 2012, 23 (8), 1364–1375. <https://doi.org/10.1007/s13361-012-0407-x>.
25. Maroufi, B.; Ranjbar, B.; Khajeh, K.; Naderi-Manesh, H.; Yaghoubi, H. Structural Studies of Hen Egg-White Lysozyme Dimer: Comparison with Monomer. *Biochim. Biophys. Acta - Proteins Proteomics* 2008, 1784 (7–8), 1043–1049. <https://doi.org/10.1016/j.bbapap.2008.03.010>.
26. Kaake, R. M.; Wang, X.; Burke, A.; Yu, C.; Kandur, W.; Yang, Y.; Novitsky, E. J.; Second, T.; Duan, J.; Kao, A.; et al. A New in Vivo Cross-Linking Mass Spectrometry Platform to Define Protein–Protein Interactions in Living Cells. *Mol. Cell. Proteomics* 2014, 13 (12), 3533–3543. <https://doi.org/10.1074/mcp.M114.042630>.
27. Matzinger, M.; Kandjoller, W.; Doppler, P.; Heiss, E. H.; Mechtler, K. Fast and Highly Efficient Affinity Enrichment of Azide-A-DSBSO Cross-Linked Peptides. *J. Proteome Res.* 2020, 19 (5), 2071–2079. <https://doi.org/10.1021/acs.jproteome.0c00003>.
28. Peheré, A. D.; Sumbly, C. J.; Abell, A. D. New Cylindrical Peptide Assemblies Defined by Extended Parallel β -Sheets. *Org. Biomol. Chem.* 2013, 11 (3), 425–429. <https://doi.org/10.1039/C2OB26637G>.

29. Link, A. J.; Vink, M. K. S.; Tirrell, D. A., Preparation of the functionalizable methionine surrogate azidohomoalanine via copper-catalyzed diazo transfer. *Nat. Prot.* 2007, 2 (8), 1879-1883.
30. Barman, A. K.; Gour, N.; Verma, S. Morphological Transition Triggered by Mannose Conjugation to a Cyclic Hexapeptide. *Arkivoc* 2012, 2013 (2), 82–99. <https://doi.org/10.3998/ark.5550190.0014.208>.
31. Broadpharm. Click Chemistry Protocols [https://broadpharm.com/public/uploads/protocol_files/Click Chemistry Protocol.pdf](https://broadpharm.com/public/uploads/protocol_files/Click_Chemistry_Protocol.pdf) (accessed Jun 25, 2020).

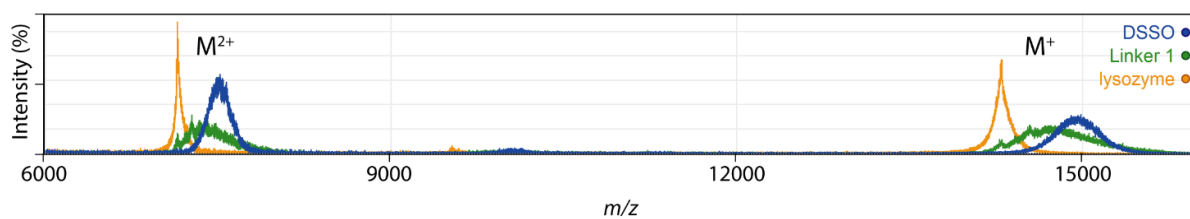
Supplementary Information



Supplementary Scheme 1

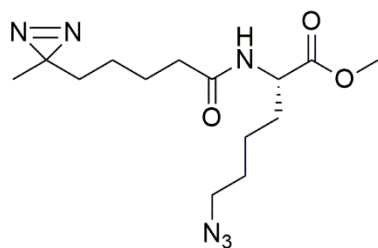


Supplementary Figure 1: MS/MS analysis of AcAAKA peptide cross-linked by compound **1** (m/z $[M+H]^+ = 1038$). Expected fragmentation patterns were used to confirm the presence of the peptide-linker complex.

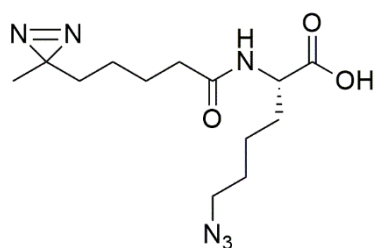


Supplementary Figure 2. MALDI spectrum shows doubly charged (approximately 7100 m/z) and singly charged (14,200 m/z) lysozyme. The protein is modified, increasing its mass, upon the covalent linkage with multiple cross-linkers.

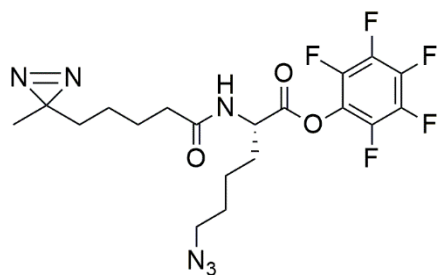
Characterisation Data



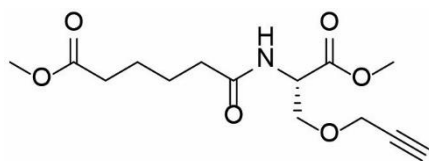
(*S*)-methyl 6-azido-2-(5-(3-methyl-3H-diazirin-3-yl)pentanamido)hexanoate. ^1H NMR (500 MHz, CDCl_3): δ 6.39 (d, $J = 8.0$ Hz, 1H), 4.46 (dt, $J = 8.0, 5.4$ Hz, 1H), 3.62 (s, 3H), 3.16 (dt, $J = 6.7, 1.2$ Hz, 2H), 2.08 (t, $J = 7.5$ Hz, 2H), 1.78-1.69 (m, 1H), 1.65-1.51 (m, 2H), 1.48 (m, 4H), 1.36-1.26 (m, 2H), 1.26-1.21 (m, 2H), 1.09-1.02 (m, 2H), 0.87 (s, 3H) ppm. ^{13}C NMR: (125 MHz, CDCl_3): δ 175.6, 174.9, 55.1, 54.4, 53.7, 38.8, 36.6, 34.7, 31.0, 28.3, 27.6, 26.3, 25.0, 22.5 ppm.



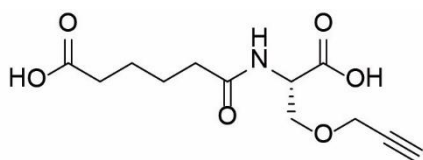
(*S*)-6-azido-2-(5-(3-methyl-3H-diazirin-3-yl)pentanamido)hexanoic acid. ^1H NMR (500 MHz, CDCl_3): δ 9.95 (br s, 1H), 6.49 (d, $J = 7.4$ Hz, 1H), 4.54 (m, 1H), 3.24 (t, $J = 6.8$ Hz, 2H), 2.19 (t, $J = 7.5$ Hz, 2H), 1.90-1.84 (m, 1H), 1.72-1.69 (m, 1H), 1.62-1.54 (m, 4H), 1.43-1.38 (m, 2H), 1.33-1.30 (m, 2H), 1.16-1.10 (m, 2H), 0.94 (s, 3H) ppm. ^{13}C NMR: (125 MHz, CDCl_3): δ 177.8, 176.5, 54.7, 53.7, 38.7, 36.5, 34.2, 31.0, 28.4, 27.7, 26.1, 25.0, 22.4 ppm.



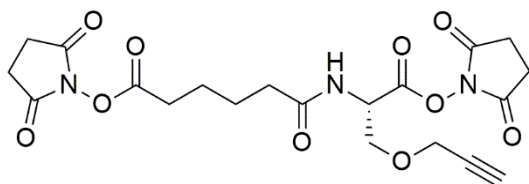
(*S*)-perfluorophenyl 6-azido-2-(5-(3-methyl-3H-diazirin-3-yl)pentanamido)hexanoate (**8**). HRMS (ESI+) calculated for $\text{C}_{19}\text{H}_{21}\text{F}_5\text{N}_6\text{O}_3$ $[\text{M}+\text{H}]^+ = 476.1595$, found $[\text{M}+\text{H}]^+$: 477.1667.



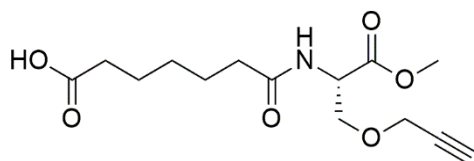
(*S*)-methyl 6-((1-methoxy-1-oxo-3-(prop-2-yn-1-yloxy)propan-2-yl)amino)-6-oxohexanoate. ^1H NMR (500 MHz, CDCl_3 , δ): 6.29 (d, $J = 7.9$ Hz, 1H), 4.79 (dt, $J = 8.1, 3.1$ Hz, 1H), 4.15 (dd, 2H), 3.97 (dd, $J = 9.4, 3.1$ Hz, 1H), 3.82 – 3.74 (m, 4H), 3.67 (s, 3H), 2.46 (t, $J = 2.4$ Hz, 1H), 2.34 (t, $J = 6.9$ Hz, 2H), 2.28 (t, $J = 6.4$ Hz, 2H), 1.75 – 1.61 (m, 4H). ^{13}C NMR (125 MHz, CDCl_3 , δ): 174.0, 172.5, 170.7, 79.0, 75.3, 69.6, 58.7, 52.8, 52.4, 51.7, 36.1, 33.9, 25.1, 24.5.



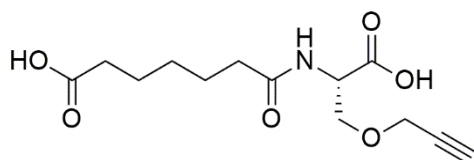
(*S*)-6-((1-carboxy-2-(prop-2-yn-1-yloxy)ethyl)amino)-6-oxohexanoic acid. ^1H NMR (500 MHz, CD_3OD , δ): 4.63 (t, $J = 4.2$ Hz, 1H), 4.18 (d, $J = 2.4$ Hz, 2H), 3.91 (dd, $J = 9.6, 5.0$ Hz, 1H), 3.79 (dd, $J = 9.6, 3.6$ Hz, 1H), 2.86 (t, $J = 2.4$ Hz, 1H), 2.31 (q, $J = 7.0$ Hz, 4H), 1.71 – 1.59 (m, 4H). ^{13}C NMR (125 MHz, CDCl_3 , δ): 174.0, 172.5, 170.7, 79.0, 75.3, 69.6, 58.7, 52.8, 52.4, 51.7, 36.1, 33.9.



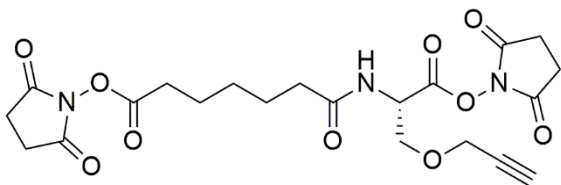
(*S*)-2,5-dioxopyrrolidin-1-yl 6-((1-((2,5-dioxopyrrolidin-1-yl)oxy)-1-oxo-3-(prop-2-yn-1-yloxy)propan-2-yl)amino)-6-oxohexanoate (**1**). HRMS (ESI+) calculated for $\text{C}_{20}\text{H}_{23}\text{N}_3\text{O}_{10}$ $[\text{M}+\text{Na}]^+ = 488.1261$, found $[\text{M}+\text{Na}]^+$: 488.1271.



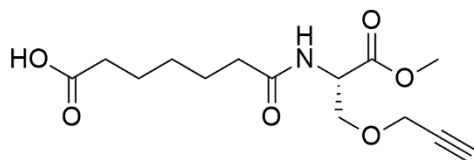
(*S*)-7-((1-methoxy-1-oxo-3-(prop-2-yn-1-yloxy)propan-2-yl)amino)-7-oxoheptanoic acid. ^1H NMR (500 MHz, CD_3OD , δ): 4.71 (dt, $J = 5.1, 3.7$ Hz, 1H), 4.22 (dd, $J = 2.4, 0.8$ Hz, 2H), 3.95 (dd, $J = 9.7, 5.1$ Hz, 1H), 3.81 (dd, $J = 9.7, 3.8$ Hz, 1H), 3.78 (s, 3H), 2.92 (t, $J = 2.4$ Hz, 1H), 2.33 (q, $J = 7.4$ Hz, 4H), 1.74 – 1.62 (m, 4H), 1.48 – 1.35 (m, 2H). ^{13}C NMR (125 MHz, CD_3OD , δ): 178.7, 177.5, 173.1, 81.3, 77.6, 71.3, 60.4, 55.2, 54.1, 37.6, 36.0, 30.9, 27.8, 27.0.



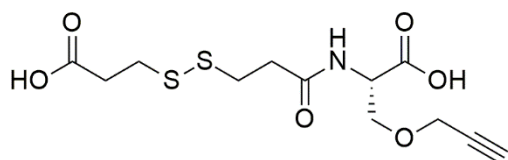
(*S*)-7-((1-carboxy-2-(prop-2-yn-1-yloxy)ethyl)amino)-7-oxoheptanoic acid. ^1H NMR (500 MHz, CD_3OD , δ): 4.64 (dt, $J = 5.0, 3.6$ Hz, 1H), 4.19 (d, $J = 2.4$ Hz, 2H), 3.92 (dd, $J = 9.6, 5.0$ Hz, 1H), 3.80 (dd, $J = 9.6, 3.7$ Hz, 1H), 2.87 (t, $J = 2.4$ Hz, 1H), 2.30 (td, $J = 7.5, 3.7$ Hz, 4H), 1.92 – 1.83 (m, 4H), 1.64 (dp, $J = 11.8, 7.5$ Hz, 2H). ^{13}C NMR (125 MHz, CD_3OD , δ): 178.8, 177.4, 174.3, 81.4, 77.6, 71.5, 60.5, 55.1, 37.7, 36.0, 31.0, 27.8, 27.0.



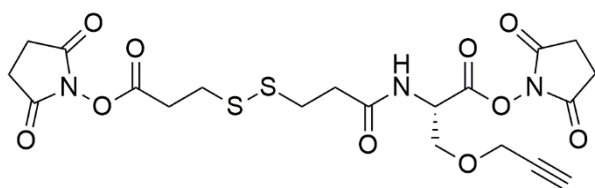
(*S*)-2,5-dioxopyrrolidin-1-yl 7-((1-((2,5-dioxopyrrolidin-1-yl)oxy)-1-oxo-3-(prop-2-yn-1-yloxy)propan-2-yl)amino)-7-oxoheptanoate (**2**). HRMS (ESI+) calculated for $\text{C}_{21}\text{H}_{25}\text{N}_3\text{O}_{10}$ $[\text{M}+\text{Na}]^+ = 502.1438$, found $[\text{M}+\text{Na}]^+$: 502.1421.



(*S*)-3-((3-((1-methoxy-1-oxo-3-(prop-2-yn-1-yloxy)propan-2-yl)amino)-3-oxopropyl)disulfanyl)propanoic acid. ^1H NMR (500 MHz, CD_3OD , δ): 4.71 (dt, $J = 5.1, 3.7$ Hz, 1H), 4.22 (dd, $J = 2.4, 0.8$ Hz, 2H), 3.95 (dd, $J = 9.7, 5.1$ Hz, 1H), 3.81 (dd, $J = 9.7, 3.8$ Hz, 1H), 3.78 (s, 3H), 2.92 (t, $J = 2.4$ Hz, 1H), 2.33 (q, $J = 7.4$ Hz, 4H), 1.74 – 1.62 (m, 4H), 1.48 – 1.35 (m, 2H). ^{13}C NMR (125 MHz, CD_3OD , δ): 178.7, 177.5, 173.1, 81.3, 77.6, 71.3, 60.4, 55.2, 54.1, 37.6, 36.0, 30.9, 27.8, 27.0.

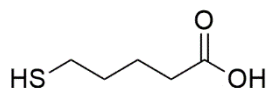


(S)-2-(3-((2-carboxyethyl)disulfanyl)propanamido)-3-(prop-2-yn-1-yloxy)propanoic acid. ^1H NMR (500 MHz, CD_3OD , δ): 4.66 (t, $J = 4.2, 4.2$ Hz, 1H), 4.19 (d, $J = 2.4$ Hz, 2H), 3.94 (dd, $J = 9.7, 4.8$ Hz, 1H), 3.80 (dt, $J = 9.8, 2.8$ Hz, 1H), 3.01 – 2.90 (m, 4H), 2.72 (tt, $J = 7.0, 3.2$ Hz, 4H). ^{13}C NMR (125 MHz, CD_3OD , δ): 176.7, 176.7, 175.1, 81.4, 77.7, 71.5, 60.5, 55.2, 37.5, 36.2, 36.1, 35.6.



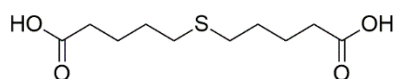
(S)-2,5-dioxopyrrolidin-1-yl 2-(3-((3-((2,5-dioxopyrrolidin-1-yl)oxy)-3-oxopropyl)disulfanyl)propanamido)-3-(prop-2-yn-1-yloxy)propanoate (**3**). HRMS (ESI+) calculated for $\text{C}_{20}\text{H}_{23}\text{N}_3\text{O}_{10}\text{S}_2$ $[\text{M}+\text{Na}]^+ = 552.0723$, found $[\text{M}+\text{Na}]^+$: 552.0727

5-mercaptopentanoic acid

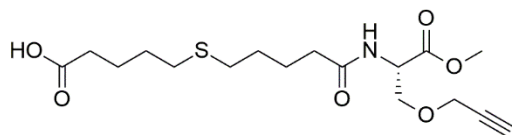


^1H NMR (500 MHz, CDCl_3 , δ): 11.29 (s br, 1H), 2.59 – 2.51 (m, 2H), 2.38 (t, $J = 7.1$ Hz, 2H), 1.80 – 1.62 (m, 4H), 1.36 (td, $J = 7.9, 1.1$ Hz, 1H). All NMR data is consistent with literature.²⁰

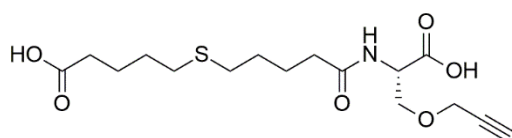
5,5'-thiodipentanoic acid



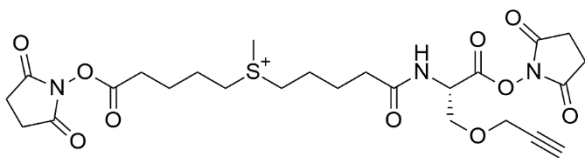
^1H NMR (500 MHz, CDCl_3 , δ): 2.54 (t, $J = 7.2$ Hz, 4H), 2.40 (t, $J = 7.2$ Hz, 4H), 1.80 – 1.61 (m, 8H). ^{13}C NMR (125 MHz, CDCl_3 , δ): 177.46, 36.02, 30.88, 27.00, 22.09. All NMR data as per literature²⁰



(S)-5-((5-((1-methoxy-1-oxo-3-(prop-2-yn-1-yloxy)propan-2-yl)amino)-5-oxopentyl)thio)pentanoic acid. ^1H NMR (500 MHz, CD_3OD , δ): 4.67 (dt, $J = 3.60, 1.37$ Hz, 1H), 4.18 (d, $J = 2.4$ Hz, 2H), 3.92 (dd, $J = 9.7, 5.0$ Hz, 1H), 3.77 (dd, $J = 9.7, 3.8$ Hz, 1H), 3.74 (s, 3H), 2.89 (t, $J = 2.4$ Hz, 1H), 2.59 – 2.51 (m, 4H), 2.31 (q, $J = 7.3$ Hz, 4H), 1.78 – 1.57 (m, 8H). ^{13}C NMR (125 MHz, CD_3OD , δ): 178.3, 176.3, 175.9, 83.1, 81.5, 72.6, 61.7, 55.9, 55.9, 38.2, 37.1, 34.6, 34.6, 32.5, 32.4, 28.3, 27.7.



(S)-5-((5-((1-carboxy-2-(prop-2-yn-1-yloxy)ethyl)amino)-5-oxopentyl)thio)pentanoic acid. ^1H NMR (500 MHz, CD_3OD , δ): 4.63 (t, $J = 4.3$ Hz, 1H), 4.19 (d, $J = 2.4$ Hz, 2H), 3.91 (dd, $J = 9.6, 5.0$ Hz, 1H), 3.80 (dd, $J = 9.6, 3.6$ Hz, 1H), 2.87 (t, $J = 2.5$ Hz, 1H), 2.54 (t, $J = 7.1$ Hz, 4H), 2.31 (q, $J = 6.8$ Hz, 4H), 1.78 – 1.66 (m, 4H), 1.63 (ddt, $J = 12.1, 7.7, 5.0$ Hz, 4H). ^{13}C NMR (125 MHz, CD_3OD , δ): 178.3, 176.1, 175.4, 83.8, 81.5, 73.0, 61.7, 55.9, 38.3, 37.1, 34.6, 34.6, 32.5, 32.5, 28.4, 27.6.



(S)-1-((2,5-dioxopyrrolidin-1-yl)oxy)-1-oxo-3-(prop-2-yn-1-yloxy)propan-2-yl)amino)-5-oxopentyl)(5-((2,5-dioxopyrrolidin-1-yl)oxy)-5-oxopentyl)(methyl)sulfonium iodide (4). HRMS (ES⁺) calculated for $\text{C}_{25}\text{H}_{34}\text{N}_3\text{O}_{10}\text{S}$ $[\text{M}]^+$ = 568.1959, found $[\text{M}]^+$: 568.2037

Conclusions and future directions

This thesis is based on the premise that there are diverse and often underutilised ways in which mass spectrometry (MS) can be used to study the human proteome. Furthermore, the bioanalytical capabilities of MS can be further enhanced when used in combination with the chemical modification of biomolecules, or 'bioconjugation.' The development of these methods into viable research tools or clinical platforms requires multidisciplinary expertise, access to specialised instrumentation and significant investments in time and funding. Firstly, however, the design and synthesis of novel reagents, optimisation of efficient conjugation reactions and development of effective data analysis workflows is required. The projects described here illustrate some of the ways in which our research group has achieved these early steps towards establishing novel strategies for the detection and analysis of pathologically significant proteins using various MS techniques.

An in-depth analysis of existing literature related to clinical MS revealed that immunoglobulins pose a significant analytical challenge, owing to their high intact molecular weight and heterogeneity.¹⁻² MS-cleavable 'mass tags' were introduced as a way of overcoming this problem, by substituting the direct detection of biomarkers antibody-antigen complexes for smaller reporter ions, which can be tailored to produce ions within optimum m/z ranges for standard MS instrumentation.³⁻⁵ We sought to further this approach by incorporating more versatile chemistries into the design of mass tags for detecting antibodies using matrix-assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI)-MS. A logical approach to this task was to include functional groups that can be used for 'click chemistry', which we hypothesised would provide a simple and fast way of modifying the mass tag m/z with a synthetic peptide. We first demonstrated this strategy using the archetypal click chemistry reaction, the copper-catalysed azide-alkyne cycloaddition (CuAAC, 'copper click').⁶ Using CuAAC, we were able to demonstrate the effective synthesis of mass tag bioconjugates and, importantly, the detection of a cleaved peptide reporter ion from a biomarker-specific antibody using MALDI-MS.

Next, we decided to explore the application of a somewhat novel click reaction, the Diels-Alder cycloaddition reaction between maleimides and pentafulvenes.⁷ Unlike copper click, this reaction requires no catalyst and leads to the formation of thermally reversible adducts, which we discovered also fragment efficiently during both MALDI and ESI-MS analyses.⁸ The concept of applying this type of chemistry for MS immunoassays is unique, and we aim to further develop it into a platform for detecting biomarkers in clinical samples. For both CuAAC and Diels-Alder conjugates, conjugation to a synthetic peptide modified with an azide or maleimide-modified amino acid, both of which can be easily incorporated into solid-phase peptide syntheses, will enable easy modification of the cleaved peptide m/z .⁹⁻¹⁰ This feature is important because it means that multiple labelled antibodies, with similar molecular weights but different mass tag m/z ratios, can be detected simultaneously, and therefore addresses what we view as one of the major limitations of existing MS immunoassay platforms. The next step is to develop a workflow for using labelled antibodies to detect antigens in more complex samples and ultimately apply these reagents in more clinically relevant studies, which focus on

the multiplexed detection of biomarkers in clinical samples. We also aim to further develop the design of these linkers to include additional features, such as affinity groups for signal amplification.¹¹

One of the main advantages of MS over other biomarker detection strategies is its ability to differentiate between different protein isoforms. Bottom-up proteomics is a well-established and effective method for achieving this; however, we hypothesised that this technique would be useful not only for identifying naturally occurring modifications, such as oxidation, but also for investigating the reactivity of a novel fluorophore developed by our collaborators (DPANzine) towards proteins.¹² We employed a bottom-up proteomics approach for identifying chemically modified amino acids in samples of myoglobin derivatised using combinations of metal-catalysed oxidation, chemical labelling with DPANzine and chemical reduction using sodium cyanoborohydride. The wide range of potential reaction products, coupled with a protein that is already prone to various *in vivo* modifications, created a complex data set for which we were required to design a data analysis workflow involving both targeted and open searches for multiple post-translational modifications. Using this search strategy, we were able to assess the labelling specificity of DPANzine for carbonylated residues in oxidised myoglobin. However, a major conclusion from this analysis was that multiple dynamic modifications can lead to misassigned peptides and missed peptide spectra matches altogether.¹³⁻¹⁴ To better understand these results, we now aim to conduct MS analysis of purified peptides labelled with DPANzine so that we can investigate the fragmentation mechanisms of this fluorophore, which might be preventing more efficient detection for protein digests.¹⁵

Lastly, the concept of chemical crosslinking, which involves the utilisation of a combination of bioconjugation and bottom-up proteomics to investigate quaternary protein structures and protein-protein interactions, is explored in a study aimed at developing a rational approach to reagent synthesis.¹⁶ We report the synthesis of several novel reagents with various attributes designed to improve MS analysis of crosslinked proteins, including MS-cleavable groups, affinity handles and several different combinations of biomolecule-reactive moieties.¹⁷ Preliminary model studies demonstrated the utility of these compounds, although future studies involving more complex systems, such as larger protein complexes or biological media will be the true measure of their capabilities. Regardless, fundamental studies such as these are important for validating crosslinking workflows.¹⁸⁻¹⁹ This modular synthetic strategy will hopefully inform the design of new crosslinkers that address some of the present limitations of crosslinking-MS, which are now centred around effective data analysis methods for cross-linking data.¹⁸

In summary, this thesis comprises a range of different methods for studying the complex chemistry and interactions of proteins, by utilising chemical derivatisation and MS. A variety of different chemistries and MS instrumentation were used to develop workflows for successfully analysing proteins at various levels of complexity, from enabling targeted study of intact heterogeneous proteins through labelling with MS-cleavable reporters, to identifying unique chemical modifications of individual amino acids in complex protein mixtures.

References

1. Bich, C.; Scott, M.; Panagiotidis, A.; Wenzel, R. J.; Nazabal, A.; Zenobi, R., Characterization of antibody–antigen interactions: comparison between surface plasmon resonance measurements and high-mass matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Biochem.* 2008, *375* (1), 35-45.
2. Zhang, Z.; Pan, H.; Chen, X., Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom. Rev.* 2009, *28* (1), 147-176.
3. Nedelkov, D., Mass spectrometry-based immunoassays for the next phase of clinical applications. *Expert Rev. Proteomic.* 2006, *3* (6), 631-640.
4. Bandura, D. R.; Baranov, V. I.; Ornatsky, O. I.; Antonov, A.; Kinach, R.; Lou, X.; Pavlov, S.; Vorobiev, S.; Dick, J. E.; Tanner, S. D., Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal. Chem.* 2009, *81* (16), 6813-6822.
5. Kang, N.; Lee, J.-M.; Jeon, A.; Oh, H. B.; Moon, B., Design and synthesis of new mass tags for matrix-free laser desorption ionization mass spectrometry (LDI-MS) based on 6,11-dihydrothiochromeno[4,3-b]indole. *Tetrahedron* 2016, *72* (36), 5612-5619.
6. Hou, J.; Liu, X.; Shen, J.; Zhao, G.; Wang, P. G., The impact of click chemistry in medicinal chemistry. *Expert Opin. Drug Discov.* 2012, *7* (6), 489-501.
7. Swan, E.; Platts, K.; Blencowe, A., An overview of the cycloaddition chemistry of fulvenes and emerging applications. *Beilstein J. Org. Chem.* 2019, *15* (1), 2113-2132.
8. Boutelle, R. C.; Northrop, B. H., Substituent effects on the reversibility of furan–maleimide cycloadditions. *J. Org. Chem.* 2011, *76* (19), 7994-8002.
9. Koehler, K. C.; Alge, D. L.; Anseth, K. S.; Bowman, C. N., Development of a maleimide amino acid for use as a tool for peptide conjugation and modification. *Int. J. Pept. Res. Ther.* 2013, *19* (3), 265-274.
10. Jagasia, R.; Holub, J. M.; Bollinger, M.; Kirshenbaum, K.; Finn, M. G., Peptide Cyclization and Cyclodimerization by CuI-Mediated Azide–Alkyne Cycloaddition. *J. Org. Chem.* 2009, *74* (8), 2964-2974.
11. Thiery, G.; Mernaugh, R. L.; Yan, H.; Spraggins, J. M.; Yang, J.; Parl, F. F.; Caprioli, R. M., Targeted multiplex imaging mass spectrometry with single chain fragment variable (SCFV) recombinant antibodies. *J. Am. Soc. Mass Spectrom.* 2012, *23* (10), 1689-1696.
12. Lisitsa, A.; Moshkovskii, S.; Chernobrovkin, A.; Ponomarenko, E.; Archakov, A., Profiling proteoforms: Promising follow-up of proteomics for biomarker discovery. *Expert Rev. Proteomic.* 2014, *11* (1), 121-129.
13. Verrastro, I.; Pasha, S.; Jensen, K. T.; Pitt, A. R.; Spickett, C. M., Mass spectrometry-based methods for identifying oxidized proteins in disease: advances and challenges. *Biomolecules* 2015, *5* (2), 378-411.
14. Rykær, M.; Svensson, B.; Davies, M. J.; Hagglund, P., Unrestricted mass spectrometric data analysis for identification, localization, and quantification of oxidative protein modifications. *J. Proteome Res.* 2017, *16* (11), 3978-3988.
15. Fedorova, M.; Bollineni, R. C.; Hoffmann, R., Protein carbonylation as a major hallmark of oxidative damage: Update of analytical strategies. *Mass Spectrom. Rev.* 2014, *33* (2), 79-97.
16. Sinz, A., The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert Rev. Proteomic.* 2014, *11* (6), 733-743.
17. Petrotchenko, E. V.; Serpa, J. J.; Borchers, C. H., An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol. Cell. Proteom.* 2011, *10* (2).

18. Beveridge, R.; Stadlmann, J.; Penninger, J. M.; Mechtler, K., A synthetic peptide library for benchmarking crosslinking-mass spectrometry search engines for proteins and protein complexes. *Nat. Comm.* 2020, *11* (1), 742.

19. Iacobucci, C.; Piotrowski, C.; Aebersold, R.; Amaral, B. C.; Andrews, P.; Bernfur, K.; Borchers, C.; Brodie, N. I.; Bruce, J. E.; Cao, Y.; Chaignepain, S.; Chavez, J. D.; Claverol, S.; Cox, J.; Davis, T.; Degliesposti, G.; Dong, M.-Q.; Edinger, N.; Emanuelsson, C.; Gay, M.; Götze, M.; Gomes-Neto, F.; Gozzo, F. C.; Gutierrez, C.; Haupt, C.; Heck, A. J. R.; Herzog, F.; Huang, L.; Hoopmann, M. R.; Kalisman, N.; Klykov, O.; Kukačka, Z.; Liu, F.; MacCoss, M. J.; Mechtler, K.; Mesika, R.; Moritz, R. L.; Nagaraj, N.; Nesati, V.; Neves-Ferreira, A. G. C.; Ninnis, R.; Novák, P.; O'Reilly, F. J.; Pelzing, M.; Petrotchenko, E.; Piersimoni, L.; Plasencia, M.; Pukala, T.; Rand, K. D.; Rappsilber, J.; Reichmann, D.; Sailer, C.; Sarnowski, C. P.; Scheltema, R. A.; Schmidt, C.; Schriemer, D. C.; Shi, Y.; Skehel, J. M.; Slavin, M.; Sobott, F.; Solis-Mezarino, V.; Stephanowitz, H.; Stengel, F.; Stieger, C. E.; Trabjerg, E.; Trnka, M.; Vilaseca, M.; Viner, R.; Xiang, Y.; Yilmaz, S.; Zelter, A.; Ziemianowicz, D.; Leitner, A.; Sinz, A., First community-wide, comparative cross-linking mass spectrometry study. *Anal. Chem.* 2019, *91* (11), 6953-6961.

Appendix

Table A1. Carbonylated peptide spectrum matches identified from Sequest HT targeted search of unmodified myoglobin.

<i>Annotated Sequence</i>	<i>Modifications</i>	<i>Charge</i>	<i>m/z [Da]</i>	<i>MH+ [Da]</i>	<i>RT (min)</i>	<i>XCorr</i>
[K].HPGDFGADAQGAMtK.[A]	T14(Thr->2-A3-KBA)	3	500.8892	1500.653	27.9416	3.74
[K].HPGDFGADAQGAMtK.[A]	T14(Thr->2-A3-KBA)	3	500.8892	1500.653	28.4119	3.83
[K].HPGDFGADAQGAMtK.[A]	T14(Thr->2-A3-KBA)	2	750.83	1500.653	28.4193	3.55
[K].HPGDFGADAQGAMtK.[A]	T14(Thr->2-A3-KBA)	3	500.8889	1500.652	28.4626	3.38
[K].HPGDFGADAQGAMtK.[A]	T14(Thr->2-A3-KBA)	2	750.8296	1500.652	28.5932	2.1
[K].HPGDFGADAQGAMtK.[A]	T14(Thr->2-A3-KBA)	3	500.8896	1500.654	28.8619	2.4
[K].HPGDFGADAQGAMtK.[A]	T14(Thr->2-A3-KBA)	2	750.8304	1500.653	29.1787	2.78
[K].HPGDFGADAQGAMtK.[A]	T14(Thr->2-A3-KBA)	2	750.8301	1500.653	29.7698	2.46
[K].HPGDFGADAqGAMtK.[A]	Q10(Deamidated); T14(Thr->2-A3-KBA)	3	501.2167	1501.636	28.1873	3.39
[K].HPGDFGADAqGAMtK.[A]	Q10(Deamidated); T14(Thr->2-A3-KBA)	3	501.2167	1501.636	28.712	3.83
[K].HPGDFGADAqGAMtK.[A]	Q10(Deamidated); T14(Thr->2-A3-KBA)	2	751.3234	1501.639	28.7228	2.24
[K].HPGDFGADAQGAmtK.[A]	M13(Oxidation); T14(Thr->2-A3-KBA)	3	506.2209	1516.648	28.6834	2.77
[K].HPGDFGADAQGAmtK.[A]	M13(Oxidation); T14(Thr->2-A3-KBA)	3	506.221	1516.649	29.2557	3.86
[K].HPGDFGADAQGAmtK.[A]	M13(Oxidation); T14(Thr->2-A3-KBA)	2	758.8277	1516.648	29.2791	2.59
[K].HpGDFGADAQGAmtK.[A]	P2(Oxidation); M13(Oxidation)	2	767.8326	1534.658	28.3539	2.11
[K].HPGDFGADAQGAMTK.[A]	K15(Lys->Alllysine)	2	751.3234	1501.639	28.7228	2.32
[K].HPGDFGADAQGAMTK.[A]	K15(Lys->Alllysine)	2	751.3234	1501.639	31.103	2.65

m/z, mass-to-charge ratio; Obs, molecular weight; Xcorr, cross correlation.

Table A2. Carbonylated peptide spectrum matches identified from Sequest HT targeted search of oxidised myoglobin.

<i>Annotated Sequence</i>	<i>Modifications</i>	<i>Charge</i>	<i>m/z [Da]</i>	<i>MH+ [Da]</i>	<i>RT (min)</i>	<i>XCorr</i>
[K].HGTVVLTALGGILKK.[K]	T7(Thr->2-A3-KBA)	4	376.9865	1504.924	32.5559	4
[K].HGTVVLTALGGILKK.[K]	T7(Thr->2-A3-KBA)	3	502.3127	1504.924	33.2549	2.98
[K].HGTVVLTALGGILKK.[K]	T7(Thr->2-A3-KBA)	3	502.3142	1504.928	33.6908	3.07
[K].HGTVVLTALGGILK.[K]	T7(Thr->2-A3-KBA)	2	688.9171	1376.827	34.0027	2.8
[K].HGtVVLTALGGILKK.[K]	T3(Thr->2-A3-KBA); T7(Thr->2-A3-KBA)	3	501.6408	1502.908	31.8075	3.03
[K].HGtVVLTALGGILKK.[K]	T3(Thr->2-A3-KBA)	4	376.9865	1504.924	32.5559	3.81
[K].HGtVVLTALGGILKK.[K]	T3(Thr->2-A3-KBA)	3	502.3128	1504.924	32.811	4.2
[K].KGGHAEELKPLAQSHATK.[H]	T17(Thr->2-A3-KBA)	3	660.681	1980.029	32.1092	2.43
[R].LFTGHpETLEK.[F]	P6(Oxidation); T8(Thr->2-A3-KBA)	2	643.3261	1285.645	30.9579	2.09
[R].LFTGHpETLEK.[F]	P6(Oxidation)	2	644.3334	1287.66	29.3343	3.21
[R].LFTGHpETLEK.[F]	P6(Oxidation)	3	429.8905	1287.657	29.4181	2.41
[R].LFTGHpETLEK.[F]	P6(Oxidation)	2	644.3329	1287.658	29.561	1.94
[R].LFTGHpETLEK.[F]	P6(Oxidation)	3	429.8911	1287.659	29.6725	3
[R].LFTGHpETLEK.[F]	P6(Oxidation)	2	644.3326	1287.658	29.7483	3.01
[R].LFTGHpETLEK.[F]	P6(Oxidation)	2	644.3329	1287.658	31.1526	2.06
[K].HKIpIK.[Y]	P4(Oxidation)	2	376.2445	751.4816	26.5697	1.95
[K].KGGHAEELKpLAQSHATK.[H]	P10(Oxidation)	3	666.6877	1998.049	27.2328	4.85
[K].KGGHAEELKpLAQSHATK.[H]	P10(Oxidation)	5	400.4157	1998.049	27.2491	4.55
[K].KGGHAEELKPLAQSHATK.[H]	K9(Lys->Alllysine)	3	661.0169	1981.036	27.3109	3.39
[K].HLkTEAEmK.[A]	K3(Lys->Alllysine); M8(Oxidation)	2	551.2658	1101.524	27.1392	2.46
[K].HLkTEAEmK.[A]	K3(Lys->Alllysine)	2	543.2672	1085.527	28.1715	2.9
[K].VEADIAGHGQEVLRILFTGHpETLEK.[F]	K26(Lys->Alllysine)	3	953.4966	2858.475	33.1486	2.73
[K].ALELFRNDIAAKYK.[E]	K12(Lys->Alllysine)	2	825.9458	1650.884	33.1492	3.5
[R].LFTGHpETLEKFDK.[F]	K11(Lys->Alllysine)	2	830.9168	1660.826	32.0946	3.88
[K].kHGtVVLTALGGILKK.[K]	K1(Lys->Alllysine); T4(Thr->2-A3-KBA)	3	544.6665	1631.985	33.1249	3.78

m/z, mass-to-charge ratio; Obs, molecular weight; Xcorr, cross correlation.

Table B1. DPANzine-labelled peptide spectrum matches identified from Sequest HT targeted search of oxidised myoglobin incubated with fluorophore.

<i>Annotated Sequence</i>	<i>Modifications</i>	<i>Charge</i>	<i>m/z [Da]</i>	<i>MH+ [Da]</i>	<i>RT (min)</i>	<i>XCorr</i>
HGTVVLTALGGILkk	K14(Lys-DPANzine); K15(Lys->Alllysine)	3	589.0069	1765.006	34.3079	2.38

m/z, mass-to-charge ratio; Obs, molecular weight; Xcorr, cross correlation.

Table B2. Peptide spectrum matches identified from Sequest HT targeted search of oxidised myoglobin incubated with fluorophore.

<i>Sample and treatment</i>	<i>Annotated Sequence</i>	<i>Modifications</i>	<i>Charge</i>	<i>m/z [Da]</i>	<i>MH+ [Da]</i>	<i>RT (min)</i>	<i>XCorr</i>
Oxidised myoglobin	kHGTVVLTALGGILKK	K1 (+259 Da)	3	632.0474	1894.128	34.4944	3.26
DPANzine-labelled unoxidised myoglobin	kHGTVVLTALGGILKK	K1 (+259 Da)	3	632.0476	1894.128	34.4988	2.49
DPANzine-labelled oxidised myoglobin	kHGTVVLTALGGILKK	K1 (+259 Da))	3	632.0472	1894.127	34.5031	3.07

m/z, mass-to-charge ratio; Obs, molecular weight; Xcorr, cross correlation.

Table C1. Mass shifts identified from Sequest HT error-tolerant search of unmodified myoglobin.

<i>Sequence</i>	<i>dM (corr)</i>	<i>dM (raw)</i>	<i>z</i>	<i>m/z (Obs)</i>	<i>M (Obs, z=0)</i>	<i>M (Calc, z=0)</i>	<i>RT (s)</i>	<i>Xcorr</i>
VEADIAGHGQEVLR	252.0091984	252.0105236	3	620.2932731	1857.85799	1605.84747	1910.365	4.013017
HGTVVLTALGGILK	252.0124079	252.0137331	2	815.9313365	1629.84812	1377.83438	2127.691	3.14632
VEADIAGHGQEVLR	252.0125078	252.013833	2	929.9379265	1857.8613	1605.84747	2010.179	3.091216
HPGDFGADAQGAMTK	252.0130638	252.014389	2	877.8454565	1753.67636	1501.66198	1875.806	2.583178
HGTVVLTALGGILKK	252.0132567	252.0145819	3	586.9885865	1757.94393	1505.92935	2105.28	3.035545
LFTGHPETLEK	252.0211191	252.0224443	2	762.3463765	1522.6782	1270.65575	2001.989	2.442958
VEADIAGHGQEVLR	252.0322697	252.0335949	3	620.3009631	1857.88106	1605.84747	1887.316	4.382184
VEADIAGHGQEVLR	324.1055119	324.1068371	3	644.3253798	1929.95431	1605.84747	1854.272	3.650994
LFTGHPETLEK	324.1058224	324.1071476	3	532.5949098	1594.7629	1270.65575	1777.776	2.65054
LFTGHPETLEK	324.1066904	324.1080156	2	798.3891615	1594.76377	1270.65575	1779.033	2.48584
VEADIAGHGQEVLR	324.1074785	324.1088037	2	965.9854115	1929.95627	1605.84747	1857.866	2.797411
VEADIAGHGQEVLR	324.1771062	324.1784314	3	644.3492431	1930.0259	1605.84747	1869.812	3.558236

dM, delta mass; corr, corrected; z, charge; m/z, mass-to-charge ratio; Obs, observed; M, molecular weight; calc, calculated; RT, retention time; Xcorr, cross correlation.

Table C2. Mass shifts identified from Sequest HT error-tolerant search of DPANzine-labelled unoxidised myoglobin.

<i>Sequence</i>	<i>dM (corr)</i>	<i>dM (raw)</i>	<i>z</i>	<i>m/z (Obs)</i>	<i>M (Obs, z=0)</i>	<i>M (Calc, z=0)</i>	<i>RT (s)</i>	<i>Xcorr</i>
VEADIAGHGQEVLR	206.123592	206.1249682	5	363.4017645	1811.97244	1605.84747	1751.68738	3.364408493
VEADIAGHGQEVLR	206.1254804	206.1268566	3	604.9987165	1811.97432	1605.84747	1752.088849	2.356070995
VEADIAGHGQEVLR	206.1260774	206.1274536	4	454.0010065	1811.97492	1605.84747	1754.625735	3.213564873
VEADIAGHGQEVLR	206.1263215	206.1276977	4	454.001069	1811.97517	1605.84747	1743.920136	3.874533415
VEADIAGHGQEVLR	206.127176	206.1285522	4	454.0012815	1811.97602	1605.84747	1759.429451	3.26649642
VEADIAGHGQEVLR	252.0098799	252.0112561	3	620.2935165	1857.85872	1605.84747	1910.814606	3.798293114
HPGDFGADAQGAMTK	252.0141114	252.0154876	2	877.8460065	1753.67746	1501.66198	1878.58941	2.277663231
LFTGHPETLEK	252.0145983	252.0159745	2	762.3431415	1522.67173	1270.65575	2001.344833	2.45062089
VEADIAGHGQEVLR	252.016363	252.0177392	2	929.9398815	1857.86521	1605.84747	2010.295501	2.956699133
LFTGHPETLEK	252.1191381	252.1205143	3	508.5993665	1522.77627	1270.65575	1791.60621	3.089152575
ALELFNRDIAAK	252.1533574	252.1547336	3	538.3092031	1611.90578	1359.75105	1932.253884	3.936362505
HGTVVLTALGGILKK	270.0978004	270.0991766	3	593.0167831	1776.02852	1505.92935	2067.947073	3.219680548
VEADIAGHGQEVLR	270.1261518	270.127528	3	626.3322765	1875.975	1605.84747	2134.853388	3.686632872
LFTGHPETLEK	270.1265368	270.127913	2	771.3991065	1540.78366	1270.65575	2145.490391	2.072377443
LFTGHPETLEK	270.1275134	270.1288896	2	771.3995965	1540.78464	1270.65575	2128.668986	2.13303256
VEADIAGHGQEVLR	270.1281794	270.1295556	2	938.9957865	1875.97702	1605.84747	2134.440924	2.758950949

dM, delta mass; corr, corrected; z, charge; m/z, mass-to-charge ratio; Obs, observed; M, molecular weight; calc, calculated; RT, retention time; Xcorr, cross correlation.

Table C3. Mass shifts identified from Sequest HT error-tolerant search of DPANzine-labelled unoxidised myoglobin after treatment with sodium cyanoborohydride.

<i>Sequence</i>	<i>dM (corr)</i>	<i>dM (raw)</i>	<i>z</i>	<i>m/z (Obs)</i>	<i>M (Obs, z=0)</i>	<i>M (Calc, z=0)</i>	<i>RT (s)</i>	<i>Xcorr</i>
VEADIAGHGQEVLR	215.073798	215.0765026	3	607.9819331	1820.92397	1605.84747	1965.499717	2.448675394
VEADIAGHGQEVLR	215.0792912	215.0819958	3	607.9837631	1820.92946	1605.84747	1950.241187	2.341320038
VEADIAGHGQEVLR	215.0809392	215.0836438	3	607.9843131	1820.93111	1605.84747	1937.156232	3.179379225
VEADIAGHGQEVLR	215.0921086	215.0948132	3	607.9880365	1820.94228	1605.84747	1895.466451	4.046961308
HPGDFGADAQGAMTK	215.1258201	215.1285247	2	859.4025265	1716.7905	1501.66198	1732.204447	2.884992838
VEADIAGHGQEVLR	220.0850285	220.0877331	3	609.6523431	1825.9352	1605.84747	1845.779839	3.07248497
HPGDFGADAQGAMTK	220.0862559	220.0889605	3	574.9242565	1721.75094	1501.66198	1765.086639	2.838758945
VEADIAGHGQEVLR	220.087592	220.0902966	3	609.6531965	1825.93776	1605.84747	1857.399474	2.745932102
VEADIAGHGQEVLR	220.089423	220.0921276	3	609.6538098	1825.9396	1605.84747	1834.627064	3.544759274
ALELFRNDIAAK	220.1296291	220.1323337	3	527.6350698	1579.88338	1359.75105	1931.032055	3.325678587
KGHHEAELKPLAQSHATK	323.9747413	323.9774459	4	577.2639765	2305.0268	1981.04936	1765.929801	5.197162628
LFTGHPETLEK	323.975184	323.9778886	2	798.3240965	1594.63364	1270.65575	2116.411214	2.345133781
LFTGHPETLEK	323.9760385	323.9787431	2	798.3245215	1594.63449	1270.65575	2026.256166	2.646345139
VEADIAGHGQEVLR	323.9766911	323.9793957	3	644.2828965	1929.82686	1605.84747	2115.917495	3.958654165
VEADIAGHGQEVLR	324.1034001	324.1061047	3	644.3251331	1929.95357	1605.84747	1847.599192	2.929350853

dM, delta mass; corr, corrected; z, charge; m/z, mass-to-charge ratio; Obs, observed; M, molecular weight; calc, calculated; RT, retention time; Xcorr, cross correlation.

Table C4. Mass shifts identified from Sequest HT error-tolerant search of oxidised myoglobin detected using DeltaMass.

<i>Sequence</i>	<i>dM (corr)</i>	<i>dM (raw)</i>	<i>z</i>	<i>m/z (Obs)</i>	<i>M (Obs, z=0)</i>	<i>M (Calc, z=0)</i>	<i>RT (s)</i>	<i>Xcorr</i>
HGTVVLTALGGILKK	210.072392	210.073725	3	573.0082998	1716.00307	1505.92935	2057.672	4.162907
HGTVVLTALGGILKK	210.074223	210.075556	3	573.0089098	1716.0049	1505.92935	2068.607	3.181268
VEADIAGHGQEVLR	210.0752917	210.0766247	3	606.3153065	1815.92409	1605.84747	1905.145	3.182397
VEADIAGHGQEVLR	210.0760241	210.0773571	3	606.3155531	1815.92483	1605.84747	1875.825	2.921164
VEADIAGHGQEVLR	210.07932	210.080653	3	606.3166498	1815.92812	1605.84747	1915.726	4.362521
GHHEAELKPLAQSHATK	214.9958805	214.9972135	4	517.995179	2067.95161	1852.95439	1713.876	3.099591
KHGTVVLTALGGILKK	215.0042845	215.0056175	3	617.3505865	1849.02993	1634.02431	1985.749	3.96959
HGTVVLTALGGILKK	215.0048871	215.0062201	3	574.6524665	1720.93557	1505.92935	2043.697	3.788384
VEADIAGHGQEVLR	215.0813952	215.0827282	3	607.9840098	1820.9302	1605.84747	1961.794	2.337988
VEADIAGHGQEVLR	215.0841418	215.0854748	3	607.9849231	1820.93294	1605.84747	1934.458	4.448572
HPGDFGADAQGAMTK	215.1254827	215.1268157	2	859.4016715	1716.78879	1501.66198	1731.243	3.532392
HPGDFGADAQGAMTK	252.0058404	252.0071734	3	585.5636598	1753.66915	1501.66198	1877.492	2.30283
VEADIAGHGQEVLR	252.0108386	252.0121716	3	620.2938231	1857.85964	1605.84747	2009.771	3.88346
HGTVVLTALGGILKK	252.0110516	252.0123846	3	586.9878531	1757.94173	1505.92935	2093.323	3.136762
HGTVVLTALGGILK	252.0124001	252.0137331	2	815.9313365	1629.84812	1377.83438	2128.941	2.647179
LFTGHPETLEK	252.0124443	252.0137773	2	762.3420415	1522.66953	1270.65575	2000.472	2.385062
VEADIAGHGQEVLR	252.0130359	252.0143689	3	620.2945565	1857.86184	1605.84747	1909.673	3.844671
VEADIAGHGQEVLR	252.0135986	252.0149316	2	929.9384765	1857.8624	1605.84747	2010.055	2.916556
HGTVVLTALGGILKK	252.0143475	252.0156805	3	586.9889531	1757.94503	1505.92935	2022.397	3.768145
HPGDFGADAQGAMTK	252.0164739	252.0178069	2	877.8471665	1753.67978	1501.66198	1876.547	2.754563
LFTGHPETLEK	252.118632	252.119965	3	508.5991831	1522.77572	1270.65575	1792.127	2.340642

dM, delta mass; corr, corrected; z, charge; m/z, mass-to-charge ratio; Obs, observed; M, molecular weight; calc, calculated; RT, retention time; Xcorr, cross correlation.

Table C5. Mass shifts identified from Sequest HT error-tolerant search of DPANzine-labelled oxidised myoglobin detected using DeltaMass.

<i>Sequence</i>	<i>dM (corr)</i>	<i>dM (raw)</i>	<i>z</i>	<i>m/z (Obs)</i>	<i>M (Obs, z=0)</i>	<i>M (Calc, z=0)</i>	<i>RT (s)</i>	<i>Xcorr</i>
KHGTVVLTALGGILKK	207.8818903	207.8831809	3	614.9764398	1841.90749	1634.02431	2082.017	3.316782
KHGTVVLTALGGILK	207.8836661	207.8849567	2	857.9144265	1713.8143	1505.92935	2126.68	2.113276
VEADIAGHGQEVLR	207.9828048	207.9840954	3	605.6177965	1813.83156	1605.84747	1926.929	3.575413
VEADIAGHGQEVLR	207.98665	207.9879406	3	605.6190798	1813.83541	1605.84747	1916.304	3.759245
LFTGHPETLEK	207.9879505	207.9892411	2	740.3297715	1478.64499	1270.65575	1852.967	1.975107
HGTVVLTALGGILKK	213.1141824	213.115473	3	574.0222165	1719.04482	1505.92935	2005.228	3.041471
VEADIAGHGQEVLR	213.1152511	213.1165417	3	607.3286131	1818.96401	1605.84747	1893.835	4.219291
VEADIAGHGQEVLR	213.1152511	213.1165417	3	607.3286131	1818.96401	1605.84747	1893.909	4.27797
GHHEAELKPLAQSHATK	213.1233643	213.1246549	4	517.527039	2066.07905	1852.95439	1659.96	2.890634
HGTVVLTALGGILK	213.1479271	213.1492177	3	531.3351431	1590.9836	1377.83438	1945.603	3.172232
ALELFRNDIAAK	214.0819708	214.0832614	3	525.6187131	1573.83431	1359.75105	1931.156	2.707265
LFTGHPETLEK	214.0938026	214.0950932	3	495.9242231	1484.75084	1270.65575	1792.055	2.820765
GHHEAELKPLAQSHATK	214.1013917	214.1026823	4	517.7715465	2067.05708	1852.95439	1690.606	2.972081
HGTVVLTALGGILKK	214.1020364	214.103327	3	574.3514998	1720.03267	1505.92935	2064.118	4.144279
GHHEAELKPLAQSHATK	214.1101807	214.1114713	4	517.7737415	2067.06586	1852.95439	1710.619	3.187497
VEADIAGHGQEVLR	220.0858932	220.0871838	3	609.6521598	1825.93465	1605.84747	1842.793	3.74226
VEADIAGHGQEVLR	220.0868563	220.0881469	4	457.4911815	1825.93562	1605.84747	1844.349	2.721872
HPGDFGADAQGAMTK	220.0880361	220.0893267	3	574.9243765	1721.7513	1501.66198	1763.485	2.534413
VEADIAGHGQEVLR	220.0904708	220.0917614	3	609.6536865	1825.93923	1605.84747	1832.284	3.183724
HGTVVLTALGGILKK	220.1183328	220.1196234	3	576.3569331	1726.04897	1505.92935	1964.467	4.032542
LFTGHPETLEK	241.9861195	241.9874101	2	757.3288565	1512.64316	1270.65575	1977.905	2.698889
HGTVVLTALGGILKK	241.9932107	241.9945013	3	583.6485598	1747.92385	1505.92935	2102.869	5.299643
HGTVVLTALGGILK	241.9940099	241.9953005	2	810.9221215	1619.82969	1377.83438	2132.992	3.825875
NDIAAKYKELGFQG	242.0233585	242.0246491	2	898.4138815	1794.81321	1552.78856	1935.536	2.857787
HGTVVLTALGGILKK	242.0512552	242.0525458	3	583.6679065	1747.98189	1505.92935	2035.482	3.745485
LFTGHPETLEK	251.9311878	251.9324784	2	762.3013915	1522.58823	1270.65575	2112.591	2.416169
HGTVVLTALGGILKK	252.0112771	252.0125677	3	586.9879165	1757.94192	1505.92935	2093.662	3.664014
LFTGHPETLEK	252.0128529	252.0141435	2	762.3422215	1522.66989	1270.65575	2000.203	2.124929

VEADIAGHGQEVLR	252.0132614	252.014552	3	620.2946165	1857.86202	1605.84747	1911.187	3.797593
HGTVVLTALGGILK	252.0145177	252.0158083	2	815.9323715	1629.85019	1377.83438	2129.399	3.972513
VEADIAGHGQEVLR	252.01535	252.0166406	2	929.9393315	1857.86411	1605.84747	2010.979	2.08727
HPGDFGADAQGAMTK	252.0155398	252.0168304	2	877.8466815	1753.67881	1501.66198	1878.604	2.68511
ALELFRNDIAAK	252.1541754	252.155466	3	538.3094465	1611.90651	1359.75105	1932.312	3.520358
HGTVVLTALGGILKK	266.0560159	266.0573065	3	591.6694931	1771.98665	1505.92935	2116.015	4.833212
LFTGHPETLEK	266.0561878	266.0574784	2	769.3638915	1536.71323	1270.65575	1918.151	3.437925
GHHEAELKPLAQSHATK	266.0568495	266.0581401	3	707.3447865	2119.01253	1852.95439	1717.11	2.335826
KGHHEAELKPLAQSHATK	266.0586748	266.0599654	4	562.7846065	2247.10932	1981.04936	1701.211	4.967408
VEADIAGHGQEVLR	266.0590988	266.0603894	3	624.9765631	1871.90786	1605.84747	1960.886	4.739177
KGHHEAELKPLAQSHATK	266.0654021	266.0666927	3	750.0459598	2247.11605	1981.04936	1702.05	5.648095
HGTVVLTALGGILKK	266.0754251	266.0767157	3	591.6759631	1772.00606	1505.92935	2062.321	3.830939
GHHEAELKPLAQSHATK	270.0962647	270.0975553	4	531.770264	2123.05195	1852.95439	1717.344	3.827969
HPGDFGADAQGAMTK	270.1234499	270.1247405	2	886.9006365	1771.78672	1501.66198	2107.652	2.095015
VEADIAGHGQEVLR	270.1251388	270.1264294	3	626.3319098	1875.9739	1605.84747	2136.107	4.367565
VEADIAGHGQEVLR	270.1272885	270.1285791	2	938.9953015	1875.97605	1605.84747	2135.447	2.44599
ALELFRNDIAAK	270.128371	270.1296616	2	815.9476315	1629.88071	1359.75105	2151.147	2.819663
LFTGHPETLEK	323.9740345	323.9753251	2	798.3228165	1594.63108	1270.65575	2116.235	2.201968
LFTGHPETLEK	323.974889	323.9761796	2	798.3232415	1594.63193	1270.65575	2026.009	2.717435
GHHEAELKPLAQSHATK	323.9754151	323.9767057	4	545.2400515	2176.9311	1852.95439	1809.946	2.65208
VEADIAGHGQEVLR	323.9764572	323.9777478	3	644.2823498	1929.82522	1605.84747	2115.585	4.471684
KGHHEAELKPLAQSHATK	323.9793291	323.9806197	4	577.2647715	2305.02998	1981.04936	1762.252	6.453737
KGHHEAELKPLAQSHATK	323.9833098	323.9846004	3	769.3519298	2305.03396	1981.04936	1762.829	4.083819
KGHHEAELKPLAQSHATK	324.100667	324.1019576	4	577.295104	2305.15131	1981.04936	1640.011	3.800978
VEADIAGHGQEVLR	324.101152	324.1024426	3	644.3239131	1929.94991	1605.84747	1859.638	3.68951
KGHHEAELKPLAQSHATK	324.1043426	324.1056332	3	769.3922731	2305.15499	1981.04936	1639.716	2.55405
HGTVVLTALGGILK	324.1067284	324.108019	3	568.3214098	1701.9424	1377.83438	2081.447	2.804563
LFTGHPETLEK	324.1078712	324.1091618	3	532.5955798	1594.76491	1270.65575	1783.086	2.658306
HGTVVLTALGGILKK	324.108323	324.1096136	3	611.0202631	1830.03896	1505.92935	1962.974	2.799631
LFTGHPETLEK	324.1090443	324.1103349	2	798.3903215	1594.76609	1270.65575	1783.196	2.829444
HPGDFGADAQGAMTK	324.1116091	324.1128997	2	913.8947165	1825.77488	1501.66198	1781.635	2.79812

dM, delta mass; corr, corrected; z, charge; m/z , mass-to-charge ratio; Obs, observed; M, molecular weight; calc, calculated; RT, retention time; Xcorr, cross correlation.