LALDI tags: ionization enhancers for mass spectrometry–based glycomics

Jacob Richard Hauser

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Chemistry

May 2019

“There is scarcely any passion without struggle.”

—Albert Camus, The Myth of Sisyphus and Other Essays

Intellectual property and publication statements

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The chapter "*Scalable synthesis of 6-amino-2-cyanobenzothiazole*" refers to work from a research article that was published in September 2016 ('*Economical and scalable synthesis of 6-amino-2-cyanobenzothiazole*'. Hauser, J. R., Beard, H. A., Bayana, M. E., Jolley, K. E., Warriner, S. L., and Bon, R. S. *Beilstein J. Org. Chem.*, **2016**, 12, 2019-2025). JRH (the candidate) and HAB performed the initial synthetic studies, JRH synthesized the final compounds, JRH performed the calorimetry experiments with assistance from MEB and KEJ, SLW and RSB provided supervision, JRH prepared a draft which was edited into its final form by RSB.

Acknowledgements

I would firstly like to thank my primary supervisor Dr Robin Bon for his unwavering support throughout both my undergraduate and postgraduate studies at Leeds. I have really enjoyed working on this project and I honestly could not have asked for a better supervisor. I would also like to thank my co-supervisor Dr Stuart Warriner for being a constant source of help and advice throughout my PhD and being nothing short of a wonderful person. A special thanks to Prof. Jane Thomas-Oates and Dr Ed Bergström at the University of York for all of their time and patience and affording me the opportunity to come visit York and use their awesome FT-ICR MALDI spectrometer.

The work described in this thesis could not have been completed without the support of the technical staff in the School of Chemistry at Leeds. In particular, I would like to express my thanks to Dr Mark Howard for help with NMR experiments, Martin Huscroft for the HPLC analysis, and Dr Alex Kulak for the SEM imaging.

I've had a fantastic experience doing my PhD at Leeds, but it would have been nothing without the people who shared that time with me. I'd especially like to thank my colleagues and office mates in the Bon group (Hester, Aisling, and Izzy) who've put up with me and my music tastes for the past four years, the Fishwick Boyz next door (Ryan G, Lewis, Aaron, and Ryan H) for ensuring there was never a dull moment in the lab, and my best friend Molly for all of her unconditional support and friendship.

I would like to also like to acknowledge this brief list of the people and things that have helped me retain my sanity and gave my life purpose and meaning on a daily basis: The Sad Goth Club, Wharf Chambers, the Leeds HxC/Punk/DIY scene, Harsh Noise™, good vibes, beer, more beer, Leeds Fixed Gear Cycling Club, "the mosh life", every member of Leeds Snowriders past and present, caffeine, methylphenidate, fluoxetine, all 420 members of The Discerning Vegan FB group, Nong Shim Shin Ramyun Instant Noodles, falafel wraps, Skramz, dank memes, and anybody who ever let me incoherently ramble on at them while stressed out over this project.

Finally, I'd like to give a shout out to my maternal parental unit, Rachel Hauser, for letting me crash at her house rent-free for the last year of my PhD. If you're reading this, stay out of my room.

Cheers.

Abstract

Carbohydrates play an important role in almost every structural and functional aspect of biology. One of the most important analytical tools for studying the structural and functional profile of native carbohydrates is mass spectrometry. However, the isolation and analysis of native glycans from their biological source is not trivial. Recently, label-assisted laser desorption/ionization mass spectrometry (LALDI-MS) has emerged as a powerful analytical tool that facilitates the selective detection of labelled substrates against a complex background of other reagents.

This thesis explores the capabilities of LALDI-MS as a novel analytical tool for MS-based glycomics. *Chapter 1* provides an overview of the different approaches to laser desorption/ionization mass spectrometry, the field of glycomics, and the challenges associated with analyzing glycans by mass spectrometry. Current applications of LALDI-MS are discussed in Section 1.2.2. *Chapter 2* describes the development of a series of pyrene-based water-soluble ionization enhancers and investigations into their LALDI-MS sensitivity and stability. The results of these investigations highlight two novel ionization enhancing labels, which exhibit high LDI sensitivity and good MS stability, respectively. *Chapter 3* describes the incorporation of the two chosen ionization enhancers from *Chapter 2* into the design of novel chemical reagents for analyzing native glycans by LALDI-MS. Using these reagents, selective labelling and analysis of simple glycans was achieved using LALDI-MS, including the in situ labelling and direct analysis of lactose from cow's milk. *Chapter 4* describes the development of a straightforward and readily scalable synthesis of 6-amino-2-cyanobenzothiazole via a 1,4-diazabicyclo[2.2.2]octane-catalyzed cyanation. This procedure allows the synthesis of the final product to be achieved in high purity at a multigram scale, requiring only filtrations and recrystallization for purification of each intermediate. Overall, this thesis demonstrates the application of LALDI-MS as a tool for studying native carbohydrates.

Table of contents

Abbreviations

2-AA	2-aminobenzoic acid
2-AB	2-aminobenzamide
2D	Two-dimensional
Ac	Acyl
ACBT	6-amino-2-cyano-1,3-benzothiazole
ANTS	2-aminonaphthalene trisulfonic acid
ApA-PaP	Aplyronine A photoaffinity amidopyrene
ApA-PaP-OSu	Aplyronine A photoaffinity amidopyrene *N*-hydroxysuccinimide ester
ApA-PP	Aplyronine A photoaffinity pyrene
AP-OEG-2AB	1-amidopyrene-labelled 2-aminobenzamide LALDI tag
AP-OEG-Hz	1-amidopyrene-labelled hydrazide LALDI tag
AP-OEG-ONH_2	1-amidopyrene-labelled hydroxylamine LALDI tag
Appel's salt	4,5-dichloro-1,2,3-dithiazolium chloride
APTS	1-aminopyrene-3,6,8-trisulfonic acid
Asn	Asparagine
Boc	*tertiary*-butoxycarbonyl
CBT	2-cyano-1,3-benzothiazole
CID	Collision-induced dissociation
COSY	Correlation spectroscopy
DABCO	1,4-diazabicyclo[2.2.2]octane
DC	Direct current
DCM	Dichloromethane
DEPT	Distortionless enhancement by polarization
DIAD	Diisopropyl azodicarboxylate
DIOS	Desorption/ionization on porous silicon
DMF	*N*,*N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
EDC•HCl	*N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride
endo-α-GalNAc-ase D	Endo-α-*N*-acetylgalactosaminidases D
ESI	Electrospray ionization
Et	Ethyl
FALDI	Fluorophore-assisted laser desorption/ionization
Fmoc	Fluorenylmethyloxycarbonyl
FT-ICR	Fourier-transform ion cyclotron resonance

Gal	Galactose
GalNAc	*N*-acetylgalactosamine
GC	Gas chromatography
GlcNAc	*N*-acetylglucosamine
HBSS	Hanks' balanced salt solution
HILIC	Hydrophilic interaction chromatography
HMBC	Heteronuclear Multiple Bond Coherence
HMQC	Heteronuclear Multiple Quantum Coherence
HOBt	Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
HR-ESI-MS	High-resolution electrospray ionization mass spectrometry
IM	Ion mobility
IR	Infrared
IRMPD	Infrared multiphoton dissociation
KE	Kinetic energy
LALDI	Label-assisted laser desorption/ionization
LC-MS	Liquid chromatography mass spectrometry
LDI	laser desorption/ionization
m.p.	Melting point
m/z	Mass-to-charge ratio
$M^{\bullet+}$	Molecular ion
MALDI	Matrix-assisted laser desorption/ionization
Man	Mannose
Me	Methyl
MOE	Metabolic oligosaccharide engineering
MS	Mass spectrometry
MS^2	Tandem mass spectrometry
MS^n	Sequential mass spectrometry - to the n^{th} degree
NALDI™	Nano-assisted laser desorption/ionization
NBD	Nitrobenzoxadiazole
^{n}Bu	Butyl
NHS	*N*-hydroxysuccinimide
NIMS	Nanostructure-initiator mass spectrometry
NMP	1-(2-naphthyl)-3-methyl-5-pyrazolone
NMR	Nuclear magnetic resonance

OEG	Oligoethylene glycol
PA	2-aminopyridine
PBA	Pyrene-1-boronic acid
PBH	1-pyrenebutyric hydrazide
PCA	1-Pyrenecarboxaldehyde
PGC	Porous graphitized carbon
Ph	Phenyl
PMF	Peptide mass fingerprinting
PMP	1-phenyl-3-methyl-5-pyrazolone
PMPMP	1-(4-methoxy)-phenyl-3-methyl-5-pyrazolone
PNGase	Peptide:*N*-glycosidase
PU-OEG-2AB	Pyrene urea-labelled 2-aminobenzamide LALDI tag
PU-OEG-Hz	Pyrene urea-labelled hydrazide LALDI tag
PU-OEG-ONH_2	Pyrene urea-labelled hydroxylamine LALDI tag
QIT	Quadrupole ion trap
RF	Radio frequency
R_f	Retardation factor
R_t	Retention time
SALDI	Surface-assisted laser desorption/ionization
SAM	Self-assembled monolayer
SAMDI	Analysis of self-assembled monolayers by MALDI
SEM	Scanning electron microscope
Ser	Serine
S_NAr	Nucleophilic aromatic substitution
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Thr	Threonine
TLC	Thin layer chromatography
TOF	Time-of-flight
UV	Ultraviolet
UV-Vis	Ultraviolet–visible
Val	Valine

Chapter 1 – LALDI-MS as a tool for biochemical analysis

Label-assisted laser desorption/ionization mass spectrometry has recently emerged as a novel and powerful analytical tool that facilitates the selective detection of labelled substrates against a complex background of reagents.[1] The aim of this project has been to develop novel chemical tools that can be used to analyze native carbohydrates directly from a biological background using label-assisted laser desorption/ionization mass spectrometry. Therefore, to provide the appropriate context for the scope of work discussed in this thesis, an overview of the different approaches to laser desorption/ionization mass spectrometry (Section 1.1 – Section 1.2), the field of glycomics (Section 1.3), and the challenges associated with analyzing glycans by mass spectrometry (Section 1.4) is addressed in this introductory chapter.

1.1 Matrix-assisted laser desorption/ionization mass spectrometry

Matrix-assisted laser desorption/ionization (MALDI) is a soft ionization technique that is commonly used to vaporize and ionize both large and small (bio)molecules in a non-destructive manner for analysis by mass spectrometry (MS).[2,3] The first reported example of MALDI using an organic matrix was first described by Hillenkamp, Karas, and colleagues in their study on the effects of varying laser wavelengths on the desorption and ionization of amino acids.[4] It was found that the alanine could be detected at a much lower irradiance when mixed with tryptophan and irradiated at λ (226 nm) than by itself. They surmised that tryptophan, which readily ionized at λ (226 nm), was acting as an absorbing matrix and facilitated the ionization of the non-absorbing alanine.

Analysis by MALDI-MS can be broken down into three phases (*Figure 1.1*): firstly, the analyte is prepared as an analyte–matrix mixture. The sample is then loaded into the MALDI-MS and ablated via laser irradiation. Finally, the analyte is ionized by the matrix and subjected to mass-analysis.

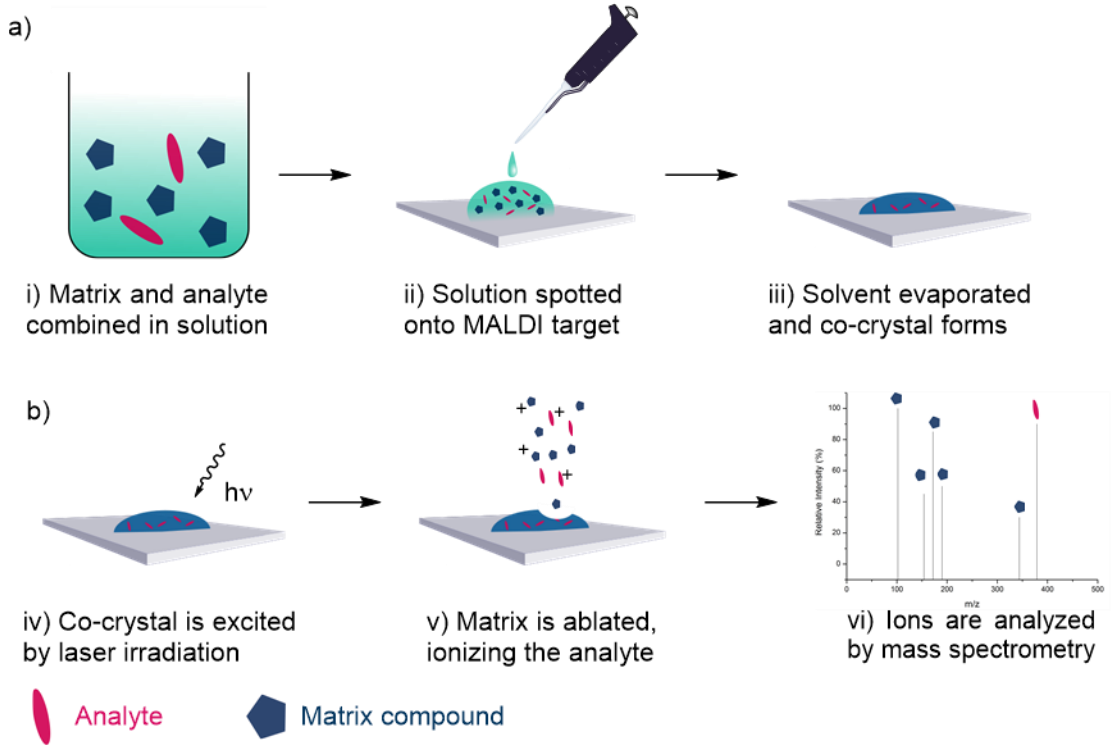


Figure 1.1: Matrix-assisted laser desorption/ionization mass spectrometry. a) Preparation of the matrix-analyte co-crystal b) General workflow for the analysis showing excitation of the matrix-analyte co-crystal followed by the subsequent desorption/ionization and mass analysis. Multiple signals are often observed for the matrix component in MALDI-MS analysis due to fragmentation, adduct formation, and clustering of the matrix ions in the mass spectrometer.[5]

1.1.1 Sample preparation

To prepare a sample for analysis by MALDI-MS, the analyte must first be combined in a solution with a molar excess of a suitable matrix compound. This is normally done using a mixture of highly purified water, an organic solvent, and trifluoracetic acid (TFA) to facilitate dissolution and co-crystallization of the matrix and analyte and enhance the relative intensity of the analyte ions.[6,7] The mixture is then spotted onto a 'target plate' and co-crystallized to give a solid-state mixture which can be subjected to analysis by MALDI-MS (*Figure 1.1a*). The matrix plays a key role in the MALDI process by absorbing the energy of the laser and facilitating the vaporization and ionization of the analyte. Therefore, to ensure that the best desorption/ionization of the analyte is achieved when performing MALDI, it is essential that the appropriate matrix compound for an analyte is chosen.

Matrix compounds usually meet certain physical and chemical properties: the matrix is usually a weak organic acid to facilitate ionization of the analyte through proton transfer,[8] although basic matrices have been reported.[9] Typically, the matrix is a crystalline solid with a low enough molecular weight that it can vaporized easily following excitation, but not so low that it evaporates during sample preparation.[2] The matrix also generally contains a chromophore or has extended conjugation to allow a strong absorption at the wavelength of excitation and promote efficient laser desorption/ionization.[5] While a number of established matrix compounds and commonly followed protocols are available,[10,11] identifying the most suitable matrix-analyte pairing can still prove difficult and can often require extensive trial-and-error testing due to the uncertainty surrounding the laser desorption/ionization mechanism.

1.1.2 Laser desorption/ionization

Ionization of the analyte is carried out by irradiating the analyte–matrix mixture with a laser pulse (*Figure 1.1b*), typically using wavelengths of 337 nm or 355 nm.[12] Upon excitement, the matrix is ablated and desorbed, vaporizing the analyte and inducing ionization. Ionization can occur in both positive and negative ionization modes, depending on the analyte and matrix used. Once ionized, analysis can be carried out by accelerating the charged species into a mass spectrometer. The exact mechanism of ionization in MALDI is not fully understood; however a number of theoretical explanations have been postulated including the photoionization and photochemical model,[13] the thermal ionization of a photo-excited matrix,[14,15] the energy pooling model,[16] the cluster ionization model,[6] the thermal proton transfer model,[17,18] the polar fluid model,[19,20] and the solid-state thermodynamic model.[21] In general, the ionization of the analyte is thought to occur via a proton transfer, with the matrix acting as either a proton donor or

acceptor.[22] However, the actual mechanism is suspected to be more complicated than a simple protonation/deprotonation, as the ionization has been shown to depend significantly on the matrix-analyte pairing, but not on the concentration of the analyte[23] or the number of acidic or basic groups on the analyte.[24]

Soft modes of ionization, like MALDI, are advantageous for working with large biomolecules and biopolymers due to the low degree of fragmentation that occurs, thus facilitating the identification of the analyte. A number of other soft ionization techniques are available, the most common of which is electrospray ionization (ESI). However, there are benefits to using MALDI that justify its application over the other available modes.[25] Firstly, unlike ESI, MALDI almost exclusively imparts a single charge onto the target, avoiding the generation of multiply charged species which are more prone to fragmentation.[26] Secondly, MALDI is a much faster method of mass analysis, being able to run one sample per second. This makes MALDI ideal for high-throughput screening.[26] MALDI is also reported to have a relatively high tolerance to ionization suppression by salts compared ESI,[27] although purification is still required to prevent contaminants obscuring the spectra.[28] Overall, the benefits of MALDI outweigh the complications and effort associated with sample preparation, making it an important and valuable tool for the analysis of biological species.

1.1.3 Mass analysis

Following desorption/ionization, a mass analyzer is used to separate the ions generated by MALDI according to their respective mass-to-charge ratio (*m/z*) before outputting them to a detector which converts the counted ions to a digital signal for interpretation. Consequently, both the sensitivity and resolution of the obtained data are related to the ion separation achieved by the mass analyzer.[26] As each mass analyzer operates differently with its own set of advantages and limitation, to achieve the optimal results the choice of mass analyzer should be founded on the requirements of that experiment.[29] Currently, the most frequently employed mass analyzers for modern mass spectrometry are: time-of-flight (TOF), quadrupole analyzer, quadrupole ion trap (QIT), and Fourier-transform ion cyclotron resonance (FT-ICR).

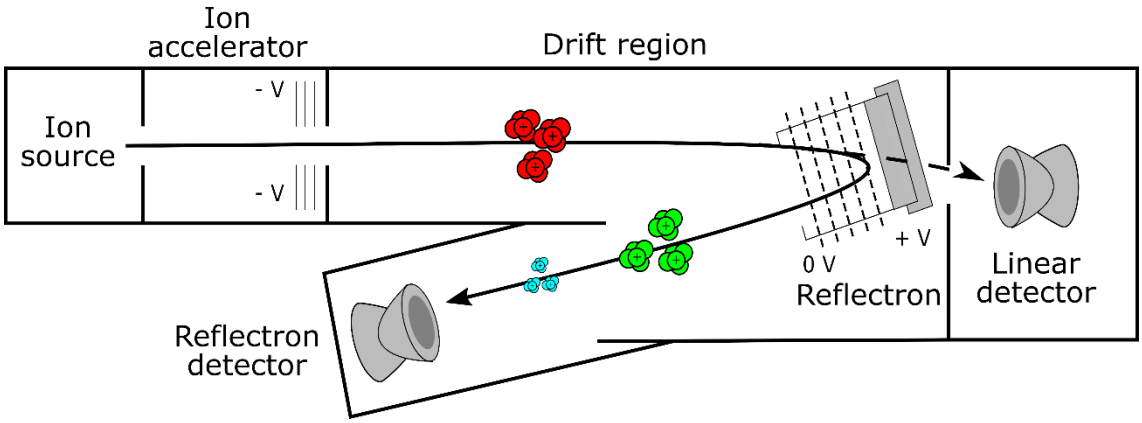


Figure 1.2: A schematic representation of a time-of-flight mass analyzer. Ions are separated in the TOF spectrometer by their relative m/z. The same kinetic energy is imparted to all the ions by the accelerator, however, heavier ions with the same charge have a slower velocity and take longer to reach the detector. As the imparted kinetic energy and distance of the drift region are known, the m/z ratio of the ion can be calculated from the time taken to reach the detector. In reflectron time-of-flight, an electrostatic field is used to reflect the ions, increasing the distance travelled and reducing temporal distribution, improving the resolution of the mass spectrometer.

TOF analyzers operate by calculating the *m/z* of each ion using the time taken to reach a detector (*Figure 1.2*).[8] Following ionization, an applied electric field is used to accelerate the ions and fire them across a field-free drift region towards a detector. Because the applied voltage in the accelerator is constant, each ion will have the same kinetic energy (KE). Since each ion has the same KE, equivalently charged ions with different mass will travel at different velocities through the drift region, with lighter ions travelling faster. Since both the applied voltage in the accelerator, and the distance of the drift region are known the *m/z* can be calculated from the time taken for an ion to reach the detector.[30] The advantages of TOF are that it has a theoretically unlimited mass range, and pairs well with MALDI.[31] However, slight variations in the kinetic energy of equivalent ions directly influence distribution of time taken for an ionic species to reach the detector. Since mass and time are proportional in TOF-MS, this creates an apparent distribution of masses, Δm, and negatively impacts on the resolution, which is defined as being the intensity of an *m/z* peak divided by its distribution, $\frac{m}{\Delta m}$. One solution to this problem is to couple the TOF with a reflectron, which extends the distance the ions travel by reflecting them back towards the detector using an electrostatic field, increasing the time, t, taken for an ion to reach the detector. In the reflectron, ions with higher KE penetrate deeper into the electrostatic field than ions with lower KE, which reduces the KE distribution of the ion and the temporal distribution, Δt. Therefore, by increasing t and decreasing Δt, the resolution of the mass spectrometer is improved.[8] However, this increased resolution does come at a cost to the sensitivity as the reflectron can result in a loss of ions.[26]

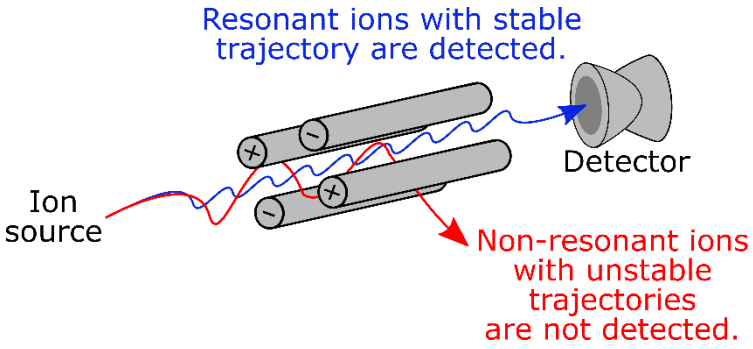


Figure 1.3: A schematic representation of a quadrupole mass analyzer. Ions are pulsed through a quadrupole and filtered using an oscillating electromagnetic field. Depending on the applied voltage only ions of a certain m/z range will have a stable enough trajectory to reach the detector. By varying the voltages applied to the quadrupole it is possible to scan across a broader range of m/z and obtain a complete mass spectrum.

Quadrupole mass analyzers utilize an oscillating electronic field to filter ions according to their *m/z* (*Figure 1.3*). In the instrument, the quadrupole comprises four parallel metal rods, with each opposing pair of rods linked electronically. When an alternating radio frequency (RF) voltage is applied to one pair and a direct current (DC) offset voltage is applied to the other, an oscillating field is generated between the quadrupole. This alternating polarity causes any ions travelling through the quadrupole to oscillate in turn. Depending on the applied DC and RF voltages, only ions within certain *m/z* range will experience a trajectory stable enough to reach the detector. The path of any ions outside of this range will become destabilized and fail to reach the detector. By varying the applied DC/RF voltages it is possible to scan across a broader range of *m/z* and obtain a mass spectrum.[32] The benefits of using a quadrupole mass analyzer are their low cost, simple maintenance, relatively small size, and the fact that they can function without requiring a high vacuum (>10^{-7} Torr).[26,33] However, they do not pair well with MALDI and are limited by their small mass range (ca. 4000 Da), single unit resolution, and an inability to perform sequential MS (MS^n) experiments alone.[31] However, this issue can be circumvented by using instruments such as triple quadrupole MS,[34] or quadrupole TOF-MS,[35] where the quadrupole has been coupled with another mass analyzer.

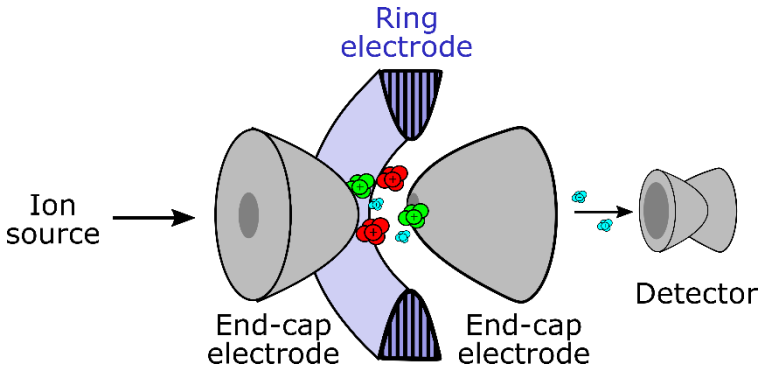


Figure 1.4: A schematic representation of a quadrupole ion trap. Ions are trapped in an oscillating electromagnetic field between three hyperbolic electrodes then sequentially ejected in order of increasing m/z for detection.

Quadrupole ion traps (QIT) function under the same principals as the quadrupole mass analyzer, however instead of allowing continual transmission of the ions through the analyzer, the ions are trapped and stored by an electromagnetic field before being ejected for detection (*Figure 1.4*).[32] The QIT consists of two hyperbolic end-cap electrodes and a ring electrode which sits between them. DC and RF voltages are applied to the electrodes which trap the ions in a circular flight path which, in the absence of collisions, can be stored in this state indefinitely.[36] Mass analysis can be achieved by amplifying the RF and DC voltages applied to the electrodes. This destabilizes the flight path of increasingly larger ions, ejecting them sequentially from the ion trap. Through this technique, the *m/z* of each ion can be determined based on the applied voltage required to eject and detect them.[37] The ability of these analyzers to trap and accumulate ions allows analysis to be performed on minute quantities of analytes.[38] Despite this high sensitivity, QITs are limited by their poor dynamic range and mass accuracy. Examples of quantitative analysis using QITs have been reported,[39–41] although their applications are generally reserved for qualitative experiments such as clinical screenings[42] or metabolite identification.[43] In addition to the QIT described here, other available ion traps include: the miniaturized 'cylindrical ion trap',[44] the two-dimensional (2D) 'linear quadrupole ion trap',[45] and the highly accurate 'orbitrap' which operates more closely to FT-ICR.[46]

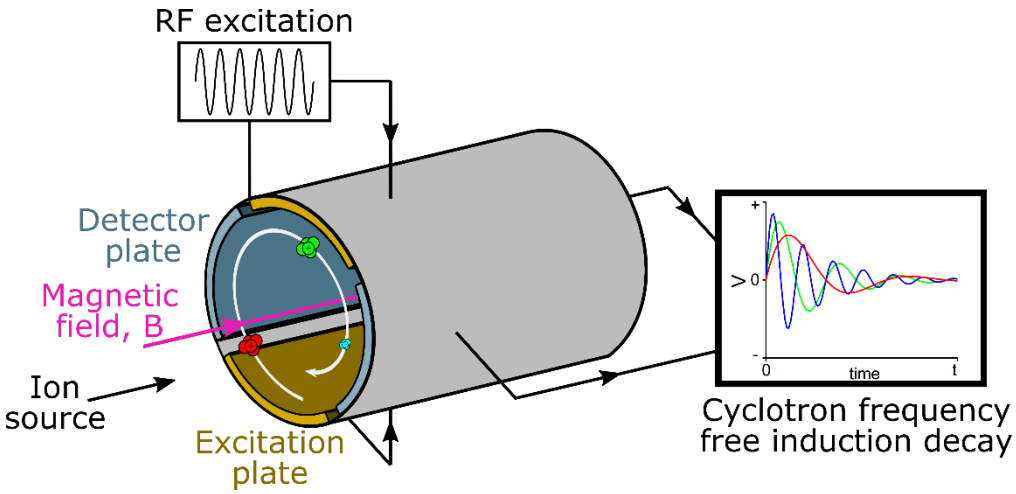


Figure 1.5: A schematic representation of a Fourier-transform ion cyclotron resonance mass analyzer. Ions trapped between the perpendicular electric and magnetic fields are excited into increasingly larger cyclotron orbits around the magnetic field using a radiofrequency and a detector records their cyclotron frequency. Analysis of the ions by FT-ICR occurs simultaneously and a Fourier transformation is used to resolve the signal and determine the m/z of each ion inside.

Fourier-transform ion cyclotron resonance (FT-ICR) is a highly accurate mass analysis technique which identifies the *m/z* of ions by measuring their cyclotron frequency around a static magnetic field.[47] Following ionization, the ions are injected into an analyzer cell and become trapped between a magnetic field and a perpendicular electric field. The trapped ions begin to resonate at their respective cyclotron frequencies which are inversely proportional to their *m/z*. A RF voltage is applied to the excitation electrodes, which excites the ions and increases the size of their cyclotron orbit. As the ion travels past the detector it induces a charge and changes the potential of the detector. Each ion is detected multiple times until they eventually relax back to their original orbit at the center of the analyzer cell. These changes in potential are recorded by the detector as a function of time to give the ion cyclotron frequency. A Fourier transformation is then used to deconvolute this signal and provide a complete mass spectrum of all the charged species present in the analyzer.[48] FT-ICR is unique amongst other mass analyzers as the ions are not detected individually by striking a detector, but counted simultaneously as they pass by the detector plate.[48] The biggest advantage of FT-ICR is that it can offer the highest resolution of any mass analyzer, delivering sub parts-per-million mass accuracy with attomole sensitivity.[49] This unparalleled resolution has established FT-ICR as an invaluable tool for various -omics studies which involve the characterization of complex biological mixtures.[50,51] However, the superconducting magnet and high vacuum (~10^{-8} Torr) required to achieve this extraordinary level of mass accuracy makes FT-ICR-MS a very large and expensive instrument to set up and run.[47,48] Therefore, unless considered essential, it may be overlooked in favor of a smaller and more cost effective instrument.

1.1.4 Applications

The applications of MALDI as an analytical tool extend beyond its ability to determine the mass of an analyte from the mass-to-charge ratio. The fast sample run time, high accuracy, and efficient ionization of bio-macromolecules have made MALDI a valuable tool for investigating complicated bio-chemical systems and carrying out both top-down and bottom-up -omics investigations.[52]

The high-throughput capabilities and non-destructive ionization of MALDI-MS makes it perfect for applications such as peptide mass fingerprinting (PMF)[53] and glycan linkage analysis,[54] while the high resolution and accuracy of instruments such as MALDI-FT-ICR allows entire biological systems to be profiled simultaneously. Examples of this include MALDI imaging, which can be used to map the localization of biomolecules in tissue samples[55] and whole-cell analysis, which can be used to identify pathogens, profile pathology and pathophysiology, and assess pathological immunity and antibacterial/antifungal resistance.[56,57]

The direct analysis of self-assembled monolayers (SAMs) by MALDI, referred to in the literature as “SAMDI”, has been used to rapidly and accurately assess the success of chemical and biochemical interactions that occur between an array of immobilized reagents and other substrates.[58] SAMDI-MS has been applied to a wide range of biochemical assays which include high-throughput compound screening,[59] multiplexing the activity and specificity of enzymes,[60,61] studying protein-protein and protein-ligand interactions,[62] and profiling drug metabolism kinetics.[63] SAMDI has also be used to achieve reaction discovery,[64] and to monitor SPOT synthesis of peptides and glycopeptides.[65]

Through a combination of these analytical techniques, MALDI-MS has great potential in aiding the detection and diagnosis of disease. A notable example of this has been use of MALDI-MS to identify membrane proteins that are associated with pancreatic cancer, a notoriously aggressive and difficult-to-diagnose disease.[66] Breakthroughs such as these are strong indicators that MALDI-MS has the potential to be a valuable tool for early detection of cancers and other diseases.

1.1.5 Limitations

While MALDI is a powerful and useful tool for mass analysis, it does have its drawbacks. Aside from the issues addressed on finding a successful matrix-analyte pairing, another limitation is its restricted detection window. Due to the requirement for the matrix compound to have a low molecular weight,[2] the resulting MS spectra are obscured by high intensity peaks between

0-1000 Da, corresponding to the ionized matrix material. These high intensity peaks subsequently hide any low molecular weight compounds, making that region of the mass spectra difficult to analyse.[67]

Small variations in how samples are prepared can lead to significant differences in the homogeneity of the resulting matrix-analyte co-crystals. These issues are often cited as one of the leading factors in the poor reproducibility of MALDI, with identical analytical protocols having been reported to yield different results.[67] Furthermore, ionization achieved through the matrix is indiscriminate, resulting in the ionization of the analyte and any impurities that are present in the co-crystal. As a result, most samples intended to be analyzed by MALDI-MS require rigorous purification to obtain a clean spectrum. These factors can make analysis of biological species especially difficult as reagents commonly used in biological studies, such as buffers, salts and detergents, can affect both the co-crystallization and ionization,[68] making analysis of samples directly from biological media difficult and impractical. If salt concentrations are too high, suppression of the analyte peaks can occur resulting in blank spectra. To prevent this, volatile buffers that can be removed under reduced pressure are recommended. When this is not possible, dialysis is often required to desalt samples prior to analysis.[69] Each of these problems presents a significant concern for the analytical accuracy of MALDI-MS. As a consequence, the time, effort, and costs associated with its use can make MALDI-MS less favored for applications that require frequent and quick analysis to be efficient.

1.1.6 Matrix-free laser desorption/ionization

Due to the limitations imposed by the organic matrix on the mass range and types of molecules that can be successfully analyzed using MALDI-MS, a number of alternatives that do not require an organic matrix have been developed. Collectively these are often referred to as surface-assisted laser desorption/ionization (SALDI). In SALDI, the analyte is adsorbed onto a solid substrate, which is then excited using a pulsed laser. The laser heats-up the substrate, resulting in the desorption and ionization of the analyte (*Figure 1.6*).[70] Therefore, as the desorption/ionization of the analyte relies heavily on both the photothermal conversion and heat transfer efficiencies of the substrate used, typical properties for SALDI substrates include good photoabsorptivity and good thermal conductivity.[70] Substrates for SALDI can be metal-based, carbon-based, or made from silicon nanostructures.

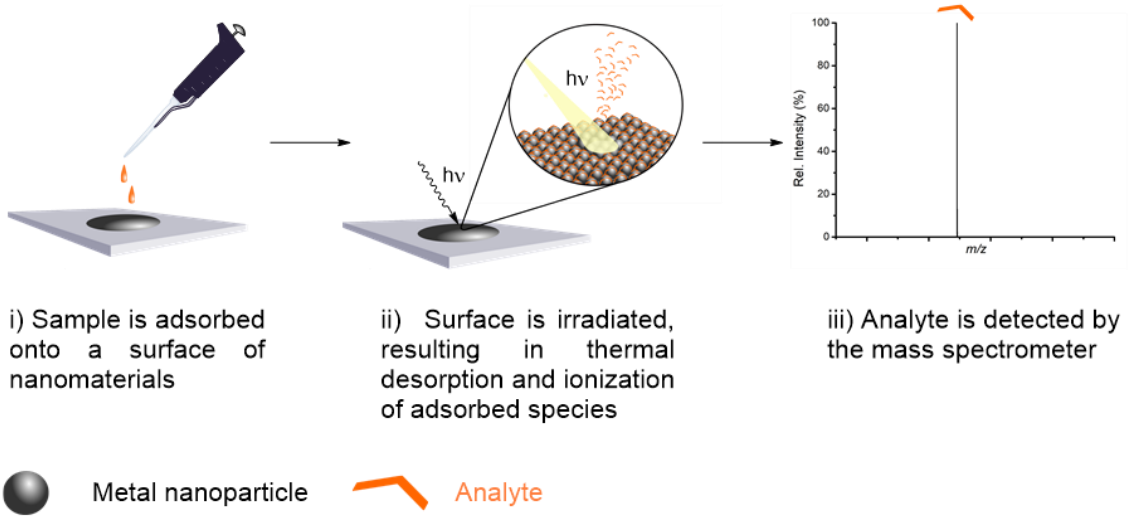


Figure 1.6: A schematic representation of SALDI-MS illustrated using metal nanoparticles as the surface material. i) A sample is adsorbed onto the surface material, in this example this is achieved by depositing a solution of the analyte onto the surface material, but this can also be achieved through other techniques such as premixing the two and depositing it onto the target. ii) The surface is then irradiated by a laser, resulting in the thermal desorption and ionization of the analyte. iii) Ionized species are detected by the mass spectrometer with little background interference from the surface material.[70]

The first example of metal-based SALDI was reported by Tanaka et al., who described the use of ultra-fine cobalt powders suspended in glycerol to achieve LDI-MS of proteins and polymers up to *m/z* 100000.[3] Au nanoparticles in particular are favored for SALDI as they exhibit excellent substrate-analyte heat transfer and can be functionalized easily using sulfur-containing molecules.[71] Nanostructured Au surfaces have also been reported as effective substrates to achieve matrix-free SAMDI.[72] Porous metal oxide nanoparticles, such as Fe_2O_3, are particularly useful for enriching trace amounts of small molecules from limited biological samples, facilitating their detection and identification.[73] However, metal-based substrates often require complicated functionalizations to achieve selective enrichment of the analyte.

Carbon-based substrates are a relatively cheap and versatile option for SALDI. Surfaces of the carbon substrates can be easily functionalized to tune the physical and electronic properties of the substrate and optimize UV absorption and the adsorption/desorption of different analytes.[74] The first reported example of carbon-based SALDI was reported by Sunner et al. and involved the use of graphite particles suspended in glycerol to analyze various peptides and proteins.[75] Other examples include carbon nanotubes, buckminsterfullerene, nanoporous graphitic carbon, highly oriented pyrolytic graphite, and nanodiamonds.[76] Graphene has been reported as a particularly useful substrate for SALDI for both polar and non-polar compounds,[77] with successful applications reported in both positive and negative ionization modes.[78] Graphene has also been shown to be advantageous for enriching aromatic analytes from biological samples

through π–π interactions.[79] However, precise control over the installation and modification of surface functional groups is difficult to achieve, and at high laser fluence carbon nanomaterials can produce carbon cluster ions that obscure the low molecular weight region of the mass spectrum.[75]

Porous silicon nanomaterial have also been found to be highly useful for analyte enrichment, and have been used extensively for the detection of many drugs and metabolites.[80,81] Various different names have been given to SALDI using nanostructured silicon including desorption/ionization on porous silicon (DIOS),[82] nanostructure-initiator mass spectrometry (NIMS),[83] and nano-assisted laser desorption/ionization (NALDI™).[84] Each case differs slightly in terms of structure, doping, and surface chemistry, although the desorption of the analyte occurs through the same mechanism. DIOS and NIMS both use porous silicon substrates etched from silicon wafers, while the commercially available NALDI™ uses a porous silicon surface made of silicon nanowires grown on a steel or silicon support.[85] However, the analyte desorption efficiency of the silicon substrate is strongly dependent on its structure, which can be affected by a large number of variables including the size and shape of the pores, the type and amount of dopant used, and the solvents/reagents used during surface preparation.[86,87] Furthermore, additional chemical modifications are often required to prevent oxidation and deactivation of the substrate over time.[88] These factors can make preparation of the silicon surfaces significantly more complicated than metal- or carbon-based nanomaterials.

Examples of silica and titania sol-gel surfaces have also been reported as substrates for SALDI.[89,90] Sol-gels were reported to have increased stability over porous silicon, with silica sol-gels still functional after one week and titania sol-gels functional for up to a month after formation.[91,92] However, only titania sol-gels are “matrix-free” as silica sol-gels require doping with 2,5-dihydroxybenzoic acid to achieve analyte desorption.[89]

In general, SALDI can offer a number of benefits, such as analyte enrichment and low substrate background, making it a useful technique for analyzing low molecular weight biomolecules. However, MALDI is still often favored over SALDI as many organic matrices are commercially available and can achieve a greater degree of versatility and sensitivity without the need for complicated or expensive substrate preparation.[93]

1.2 Label–assisted laser desorption/ionization mass spectrometry

Label-assisted laser desorption/ionization mass spectrometry (LALDI-MS) is a matrix-free approach to laser desorption/ionization mass spectrometry (*Figure 1.7*). First described by Kozmin et al.,[94] LALDI-MS was conceived following inspiration from reports of ionization-enhancing labels for ESI-MS that were reported to improve sensitivity[95,96] and permit high-throughput reaction evaluation.[97]

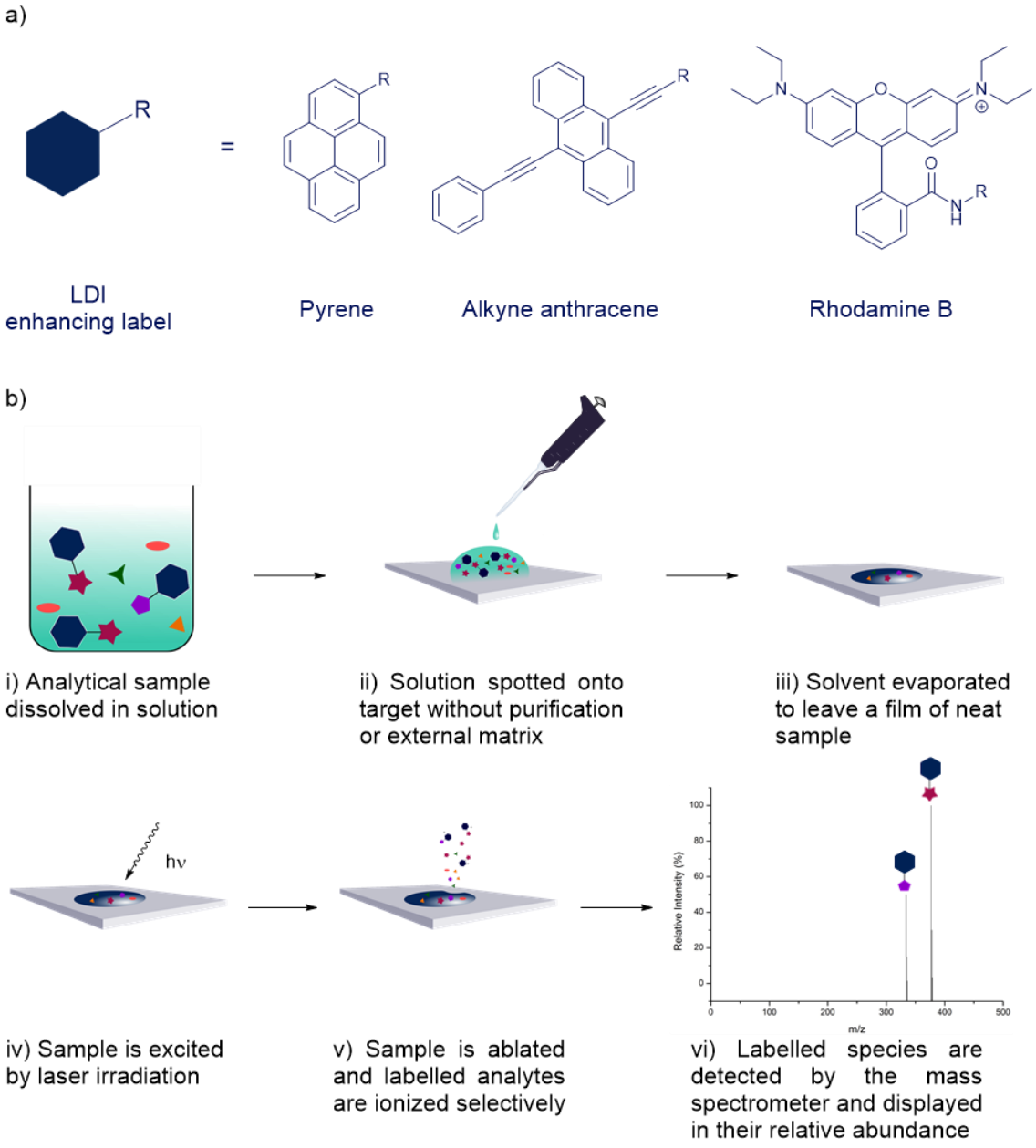


Figure 1.7: Label-assisted laser desorption/ionization mass spectrometry. a) Examples of laser desorption/ionization enhancers with their chemical structures shown. b) General workflow for the analysis of a complex mixture using LALDI-MS showing excitation of the sample followed by the subsequent desorption/ionization and mass analysis.

Analysis by LALDI-MS is performed in a similar manner to MALDI-MS, however, the incorporation of a polyaromatic group into the analyte allows the selective ionization and analysis of the labelled compound directly through the polyaromatic substituent. In doing so, the polyaromatic label eradicates the need for an external matrix. Polyaromatic compounds, such as pyrenes, were selected as the label component in LALDI due to their high molar absorptivity and capacity to undergo photoionization/desorption when excited by UV radiation, forming radical cations which can be measured by mass spectrometry.[98,99] These qualities make them ideal replacements for the matrix component.

1.2.1 Advantages over MALDI

In addition to retaining the benefits of MALDI, by covalently incorporating the matrix into the substrate, LALDI gains several additional advantages. Firstly, the absence of a matrix greatly facilitates the analysis of the resulting mass spectra. No molar excess of the matrix component means there are no large matrix peaks that obscure the lower *m/z* on the spectrum or drown out the other peaks, granting LALDI-MS with a wider detection range. Moreover, it was reported that LALDI-MS could produce spectra with better signal-to-noise and higher sensitivity than MALDI-MS analysis of the same amount of analyte.[94] As no matrix derived ion peaks were observed in the LALDI-MS spectra, the molecular ion peaks of the LALDI probes could be detected at lower concentrations than when combined with a matrix.[100]

Furthermore, concerns regarding the optimum matrix-analyte pairing and successful co-crystallization are now irrelevant, as preparation of a sample for analysis only requires the evaporation of the solvent. Additionally, the properties of high molar absorbance and facile ionization of the polyaromatic label ensure the efficient and selective ionization of the labelled analyte, thereby permitting analysis directly from a complex mixture without the subsequent ionization of any other reagents or impurities that are present.[94] As a result, the peaks for the labelled species have a greater relative intensity, making them easier to identify, and facilitating characterization of the analyte.

Initially, the development of LALDI as an analytical tool does demand more work than MALDI, requiring deliberation on which polyaromatic compound to use, how and where to attach the label, and whether activity or physicochemical properties of the analyte will be affected. However, once a suitable compound has been synthesized, it can be directly analyzed without any additional components or purification, significantly reducing the time and effort required to analyze each sample.

1.2.2 Current applications

Described below is an overview of the known approaches to LALDI-MS that have been published in the literature and their respective applications. A summary of the respective applications, probes used, and conditions of each study discussed in this section is represented in *Table 1.1*.

1.2.2.1 High-throughput reaction discovery

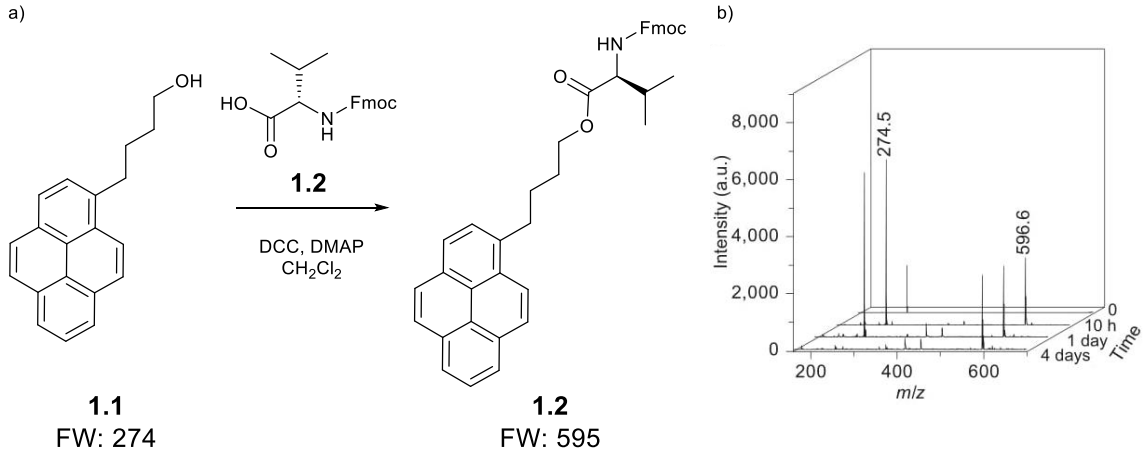
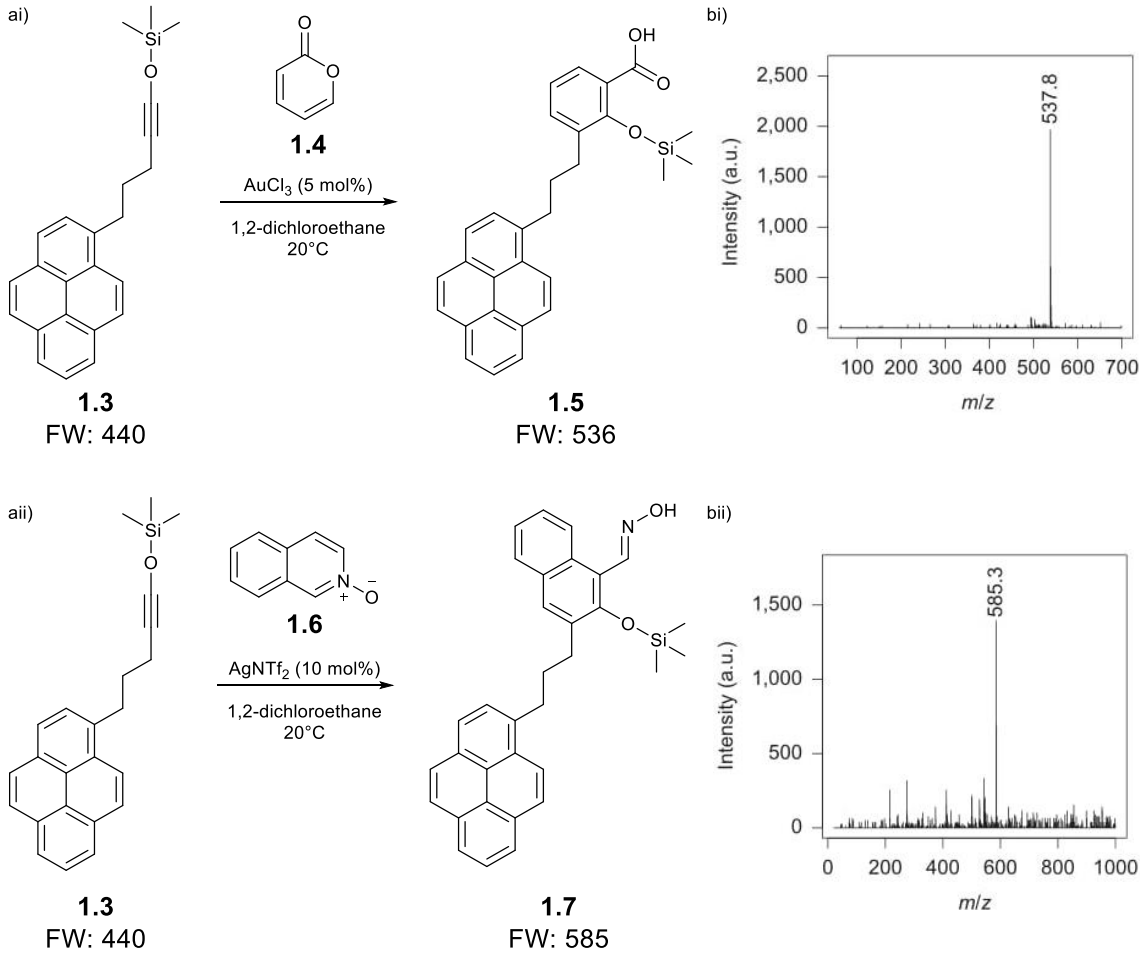


*Figure 1.8: Use of LALDI-MS to monitor the progress of a known esterification reaction. a) Reactants and conditions of the esterification reaction followed by LALDI-MS as a proof of concept experiment. b) LALDI-MS spectra for conversion of alcohol **1.1** (m/z 274.5) to ester **1.2** (m/z 596.6) showing the change in abundance of starting material and product over the course of the reaction. Reproduced with permission from Nature Publishing group: Nature Chemistry, **5**, 423–427, "Label-assisted mass spectrometry for the acceleration of reaction discovery and optimization", J. R. Cabrera-Pardo, D. I. Chai, S. Liu, M. Mrksich and S. A. Kozmin, Copyright (2013).*[94]

Since its inception, LALDI-MS has already been shown to have a number of analytical applications. The first reported example was described by Kozmin et al.[94] and employed pyrenes to carry out high-throughput reaction discovery directly from a complex mixture, requiring only nanomoles of the analyte. Initial studies were carried out to find the most suitable polyaromatic group. Pyrene was deemed most suitable for selective ionization due to its chemical inactivity and readiness to photoionize following irradiation.

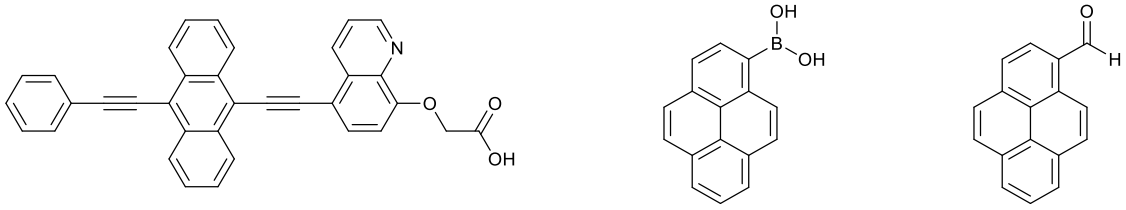
To prove the concept of LALDI-MS, laser desorption/ionization (LDI) MS was used to measure the progress of an ester formation reaction of alcohol **1.1** with Fmoc-Val-OH **1.2** directly from the reaction mixture. Resulting analysis showed that only the pyrene labelled compounds were visible in the mass spectra, allowing the turnover of the reaction to be measured by comparing the relative intensities of signals for the starting material **1.1** and product **1.2** (*Figure 1.8b*).

Following success with these initial studies, siloxyalkyne **1.3** was used to carry out high-throughput reaction discovery using a range of reactants and reagents that resulted in a total of 696 possible reactions. Using LALDI-MS, the group was able to analyze all reactions in two hours and proceeded to discover two previously unreported benzannulation reactions (*Scheme 1.1*).[94] Although the experiment proved to be successful, the probe they used would only be suitable for reactions that involve siloxyl alkynes. Therefore, to carry out high-throughput reaction discovery for any other reaction, a new probe would need to be synthesized. Additionally, this study does not investigate whether the proximity of the large pyrene moiety on the reactant could interfere with the reaction sterically, or even electronically through π-stacking, affecting the results of the high-throughput screening. This is an issue that could prove to be more prevalent in more complicated interactions, such as an enzyme-substrate complexation, where the specific conformation of the reactants is vital to ensure a successful reaction.



Scheme 1.1: LALDI-MS for reaction screening. a) Reactants and conditions of two previously unknown benzannulation reactions discovered through the LALDI-MS high-throughput reaction screening. b) LALDI-MS spectra of the reaction mixtures of each benzannulation reaction.[94] *Reproduced with permission from Nature Publishing group: Nature Chemistry, **5**, 423–427, "Label-assisted mass spectrometry for the acceleration of reaction discovery and optimization", J. R. Cabrera-Pardo, D. I. Chai, S. Liu, M. Mrksich and S. A. Kozmin, Copyright (2013).*[94]

1.2.2.2 Selective detection of chemical functionality



1.8 **1.9** **1.10**

*Figure 1.9: Structures of probes used by Basak et al. for selective detection of Zn^{2+} ions (**1.8**),[101] cis-1,2-diols (**1.9**),[102] and primary amines (**1.10**)[103] by LALDI-MS.*

Inspired by the work of Kozmin et al., Basak and colleagues developed a series of probes to selectively detect a number of specific chemical functionalities from a complex background. These included metal (II) ions,[101,104] cis-1,2-diols,[102] and primary amines.[103]

In the first study, reported by Addy et al., alkyne anthracene oxinyl acetic acid **1.8** (oxine = 8-hydroxyquinoline) was used to selectively detect zinc ions in a mixture of other transition metal ions.[101] The alkyne anthracene moiety was chosen as the polyaromatic group due to its wide UV absorption window and its capacity as a fluorophore, presenting an opportunity for a secondary analytical handle. Additionally, oxinyl acetic acid was chosen as the chelating ligand following reported strong affinity for Zn^{2+} ions. Addy et al. reported that anthracene **1.8** exhibited an affinity for Zn^{2+} ions and could be detected using LDI-MS without any assistance from an external matrix (*Figure 1.10*). However, fragmentation of the probe was observed and when tests were carried out in a mixture of metal ions, some chelation to the solvent and Ni^{2+} also took place. Therefore, the probe cannot be claimed to be specific. While this could be a useful tool for the selective detection of transition metal ions, the observed probe fragmentation and potential difficulties surrounding finding other suitable ligand-ion pairings may limit its application.

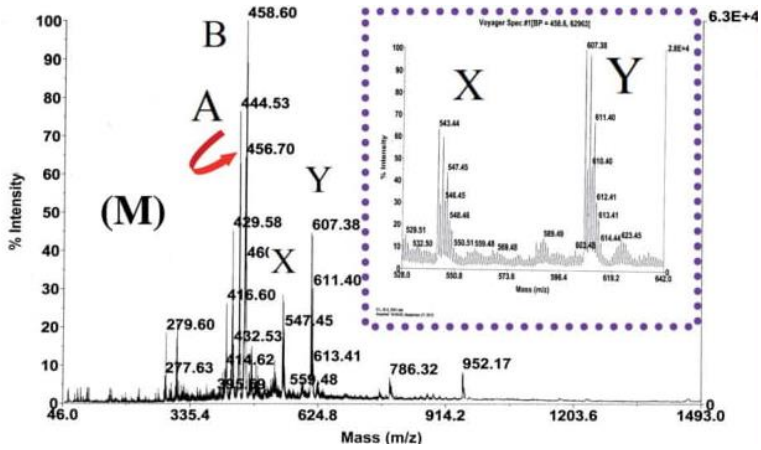


*Figure 1.10: LALDI-MS spectrum of an anthracene **1.8**-Zn^{2+} complexes in the presence of excess $Zn(ClO_4)_2$. peaks at m/z 543.5 (**X**) and 607.38 (**Y**) were correspond to a 2 : 1 and 1 : 1 complex between anthracene **1.8** and Zn^{2+} ions, respectively. Peaks labelled **A** and **B** correspond to fragments of anthracene **1.8**. Reproduced from P. S. Addy, S. Basu Roy, S. M. Mandal and A. Basak, "Polyaromatic label-assisted laser desorption ionization mass spectrometry (LA-LDI MS): a new analytical technique for selective detection of zinc ion", RSC Advances, **4**, 23314–23318, (2014) with permission from The Royal Society of Chemistry.*[101]

Addy et al. then went on to use pyrene-1-boronic acid (PBA) **1.9** to selectively detect cis-1,2-diol functionality.[102] These studies showed that PBA could be used to selectively detect nanomole quantities of cis-1,2-diols from a mixture of other species. However, a number of issues arose around the PBA probe. Following photoionization, PBA **1.9** was observed to both dimerize and act as an external matrix to unlabelled diols. Both of these problems resulted in additional complexity of the mass spectra. Detection of monosaccharides D-glucose and D-ribose was also demonstrated using PBA **1.9**. The LALDI-MS spectrum of D-ribose labelled with PBA **1.19** reported by Addy et al. has been reproduced in *Figure 1.11*. However, multiple labelling of the monosaccharides was observed due to them encompassing more than one cis-1,2-diol functionality. This effectively excludes PBA **1.9** from any viable applications within glycomics. Multiple labelling would over-complicate analysis of glycans by generating multiple signals per analyte, while inconsistencies in how efficiently different species undergo additional labelling could also affect the intensities of those signals and lead to misinformation about the relative abundances of analytes. Hence, for any analysis or quantification of complicated mixtures, labelling of analytes with a known and consistent stoichiometry would be preferred. Finally, while it was reported to be useful for biologically relevant molecules, no information was provided regarding the solubility of the PBA probe under aqueous conditions or if PBA **1.9** displayed selectivity to cis-1,2-diols over other biological species such as amines or alcohols, which would be essential to know before developing its application further. Considering the highly lipophilic nature of pyrene compounds, it is unlikely that this probe would be water-soluble without further modification.[105]

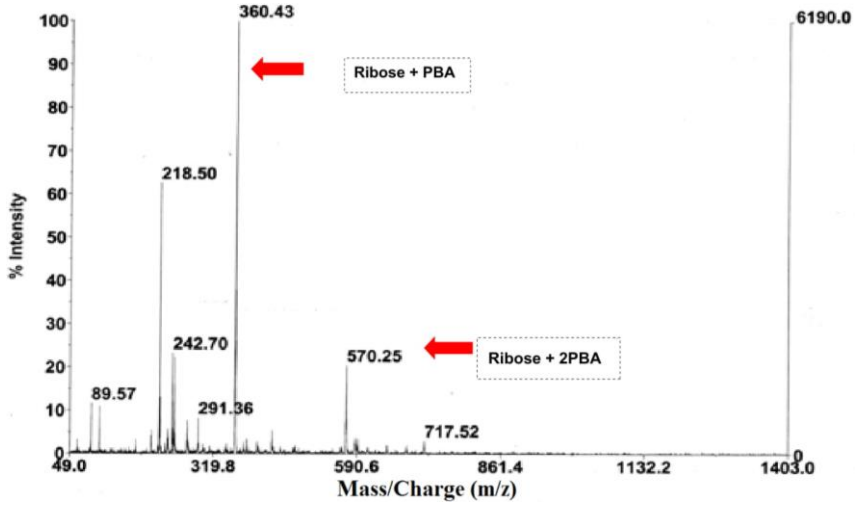


*Figure 1.11: LALDI-MS spectrum of D-ribose labelled with PBA **1.9**. Reproduced from P. S. Addy, A. Bhattacharya, S. M. Mandal and A. Basak, "Label-assisted laser desorption/ionization mass spectrometry (LA-LDI-MS): an emerging technique for rapid detection of ubiquitous cis-1,2-diol functionality", RSC Advances, **4**, 46555–46560, (2014) with permission from The Royal Society of Chemistry.*[102]

In a third study reported by Mandal et al., 1-pyrenecarboxaldehyde (PCA) **1.10** was employed for the selective detection of biogenic amines, amino acids, and peptides, broadening the scope of what could be detected by LALDI-MS.[103] The LALDI-MS spectra of alanine labelled with PCA **1.10** reported by Mandal et al. has been reproduced in *Figure 1.12*. Aldehyde **1.10** was shown to be able to selectively detect peptides, amines and biogenic amines from a mixture of other reagents, potentially providing a new method for small peptide profiling of a mixture. However, as with PBA **1.9**, multiple labelling was also observed for polyamines and no information reported on the aqueous solubility of PCA **3.9**.

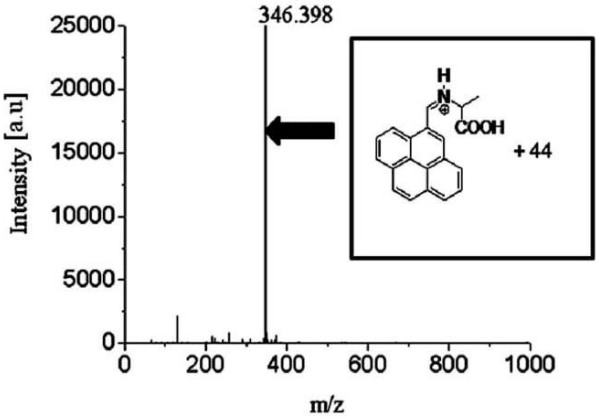


*Figure 1.12: LALDI-MS spectrum of alanine labelled with PCA **1.10**. Reproduced from A. Mandal, A. K. Das and A. Basak, "Label-assisted laser desorption/ionization mass spectrometry (LA-LDI-MS): use of pyrene aldehyde for detection of biogenic amines, amino acids and peptides", RSC Advances, **5**, 106912–106917, (2015) with permission from The Royal Society of Chemistry.*[103]

Most recently, Mandal et al. reported the development of a bifunctional tool capable of detecting metal(II) ions by both LALDI-MS and fluorescence.[104] The duel-purpose probe **1.11** consisted of a pyrene LDI enhancer conjugated to a tridentate terpyridine ligand that is known to complex with various transition metal ions (*Figure 1.13a*).

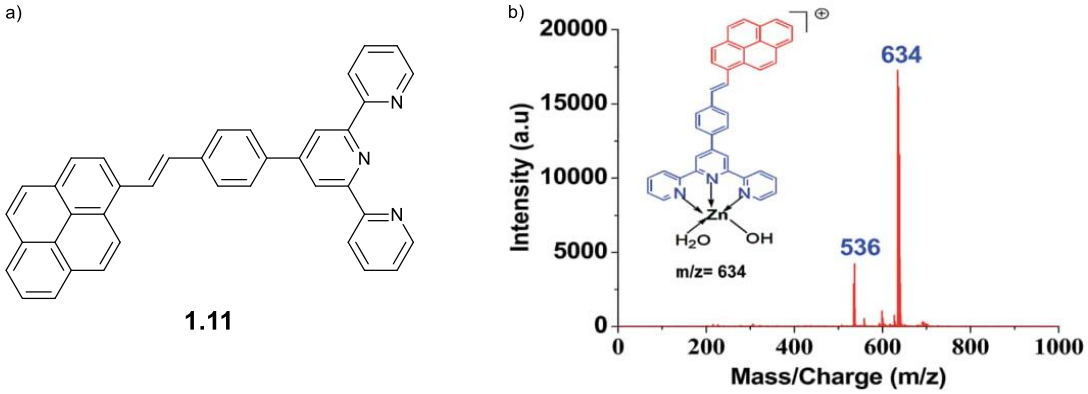


*Figure 1.13: a) Chemical structure of the bifunctional probe **1.11** capable of detecting Zn^{2+}, Co^{2+}, and Ni^{2+} ions by LALDI-MS and detect the presence of Zn^{2+} ions by monitoring changes in fluorescence emission. b) LALDI-MS spectrum of a pyrene **1.11**-Zn^{2+} complex in the presence of excess $Zn(ClO_4)_2$. Reproduced from A. Mandal, A. Maity, S. Bag, P. Bhattacharya, A. K. Das and A. Basak," Design and synthesis of dual probes for detection of metal ions by LALDI MS and fluorescence: Application in Zn(II) imaging in cells ", RSC Advances, **7**, 7163–7169, (2017). under Creative Commons Attribution 3.0 Unported License - published by The Royal Society of Chemistry.*[104]

Pyrene **1.11** was shown to have a good tolerance to both organic and ionic impurities, with clear detection of **1.11**-Zn^{2+}, -Co^{2+}, and -Ni^{2+} complexes possible even in the presence of tenfold molar excess of the perchlorate salts, and detection of **1.11**-Co^{2+} achieved in the presence of other common organic biological impurities. The LALDI-MS spectra of the **1.11**-Zn^{2+} complex reported by Mandal et al. has been reproduced in *Figure 1.13b.* Mandal et al. reported that the fluorescence emission of free **1.11** was quenched when complexed to Co^{2+}, Zn^{2+}, and Ni^{2+}. However, the **1.11**-Zn^{2+} complex also exhibited a bathochromic shift in the fluorescence emission from blue (487 nm) to red (620 nm) as the concentration of Zn^{2+} increased. It was proposed that the observed change in fluorescence could be exploited for sensing Zn^{2+} ions within live cells. To test this theory HaCaT cell were incubated with **1.11** then analyzed by fluorescence microscopy. A weak red fluorescence was observed in the cells, suggesting the presence of Zn^{2+}ions. Basak et al. proposed that this fluorescence was generated by **1.11** complexing to zinc-containing enzymes like metalloproteases. However, no additional evidence or experiments, such as LALDI-MS analysis of the cell lysates, were reported to support this claim and confirm the complexation of **1.11** to Zn^{2+} within the cells or determine the source of the zinc. Although no LALDI-MS analysis was carried out on analytes labelled from the cell, this work presented another clear example of a LALDI reagent being applied in cellulo. If cell permeable

LALDI reagents could be detected directly by LALDI-MS from within cells or tissue, it could be possible to achieve matrix-free MALDI imaging.

1.2.2.3 Identification of protein-ligand binding sites

A series of investigations undertaken by the Kigoshi group have been directed towards developing reagents that can be used to identify protein-ligand interactions in a complex mixture by LALDI-MS.[100,106,107]

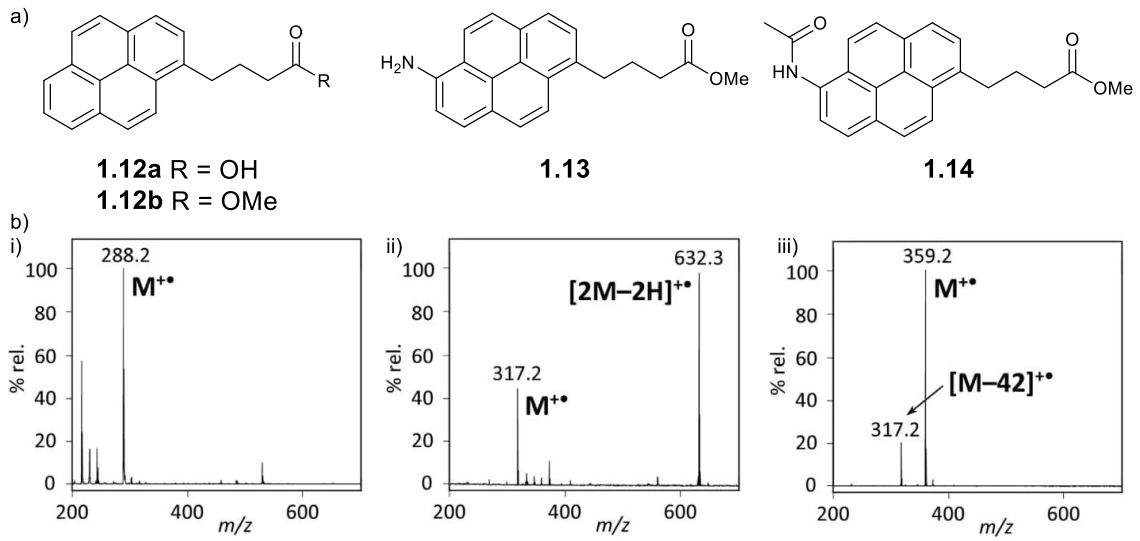


*Figure 1.14: LDI-enhancing labels investigated by Yoneda et al. in the pursuit of a more sensitive LALDI reagents. a) Structures LDI-enhancing labels investigated by Yoneda et al. b) LALDI-MS spectra for i) **1.12a**, ii) **1.13b**, iii) **1.14** recorded by Yoneda et al. Reproduced from K. Yoneda, Y. Hu, M. Kita and H. Kigoshi,"6-Amidopyrene as a label-assisted laser desorption/ionization (LA-LDI) enhancing tag: development of photoaffinity pyrene derivative", Scientific Reports, **5**, 17853, (2015) under Creative Commons Attribution 4.0 International License - published by Nature Publishing Group.*[100]

It was proposed that pyrene-functionalized probes could be used to label proteins in situ, then peptide mass fingerprinting (PMF) could be performed by LALDI-MS to elucidate the location of the binding site.[100] As a proof of concept, Yoneda et al. sought to use LALDI-MS to identify the aplyronine binding site on actin.[100] The initial investigation by Yoneda et al. was primarily focused on improving the LALDI-MS sensitivity of LDI enhancers used in their chemical probes.[100] This was to ensure that the probes would be able to achieve the sub-picomolar sensitivity required to perform PMF accurately.

Preliminary findings suggested that the LALDI-MS sensitivity was directly related to the absorbance of an LDI enhancer at the wavelength of excitation applied in the LDI mass spectrometer (355 nm). This was supported by their observations that aminopyrene **1.13** and amidopyrene **1.14** both absorbed strongly at 355 nm and exhibited a ten thousand-fold higher

sensitivity than pyrenebutyric acid **1.12a**, which absorbed poorly at 355 nm. However, a comparison of methyl pyrene-1-butanoate **1.12b**, amine **1.13**, and amide **1.14** found that all three LDI enhancers had an equivalent detection limit of 10 fmol, suggesting that the sensitivity of the LDI enhancers was influenced more by the presence of functional groups that can be easily protonated or deprotonated than by their UV absorbance.

Various degradations were observed during LALDI-MS analysis of amine **1.13** and amide **1.14**. Amine **1.13** (*Figure 1.14bii*) exhibited oxidative dimerization following photoionization, resulting in unpredictable fragmentation and complicated mass spectra. Amide **1.14** (*Figure 1.14biii*) experienced a single and consistent fragmentation at the terminal amide resulting in the loss of ketene.

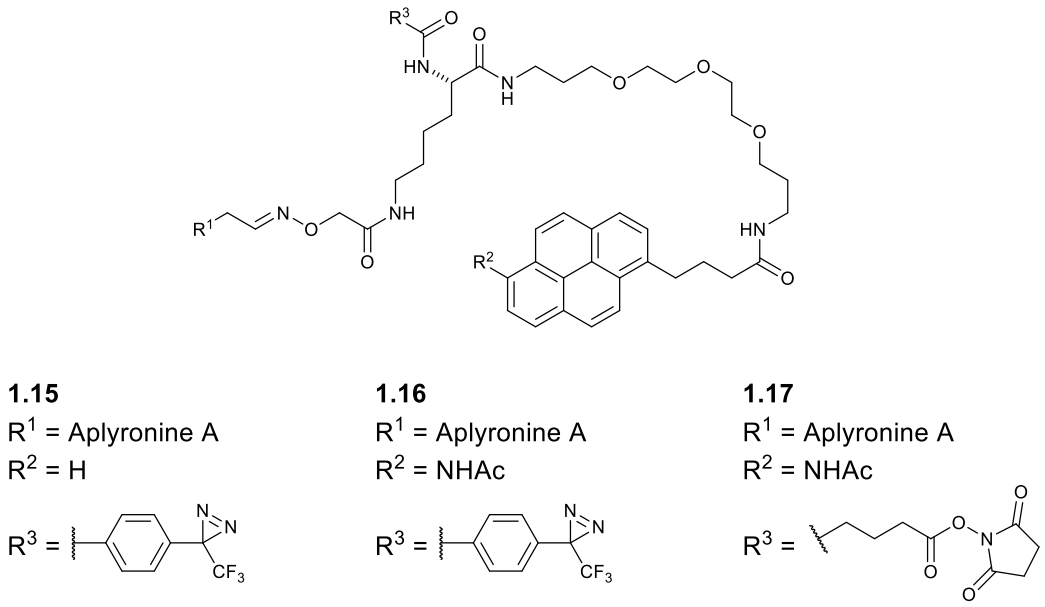


Figure 1.15: Structures of pyrene-functionalized aplyronine A affinity probes designed by Yoneda et al. for identifying the binding site of aplyronine A on actin. ***1.15*** *and* ***1.16*** *were designed to label actin through photo-reactive crosslinking,* ***1.17*** *was designed to label actin through amine-reactive crosslinking.*[100,106]

Yoneda et al. reported developing two photoaffinity probes for analyzing the binding site of actin by LALDI-MS, aplyronine A photoaffinity pyrene (ApA-PP) **1.15**, and aplyronine A photoaffinity amidopyrene (ApA-PaP) **1.16** (*Figure 1.15*).[100] Aplyronine A was chosen as the binding ligand due to its well-established interactions with actin and retention of its biological activity even when substituted with a large oligoethylene glycol (OEG) linker,[108–110] which was incorporated into the probe to improve its solubility in aqueous media. To compare the LALDI-MS stability and sensitivity of ApA-PP **1.15** and ApA-PAP **1.16** both probes were photocrosslinked to solvent and analysed by LALDI-MS (*Figure 1.16*). Both ApA-PP **1.15** and ApA-PaP **1.16** were detected by LALDI-MS as their fragment ions. The highest *m/z* observed in each spectrum correlated to a

proposed fragmentation at the N-O oxime linkage. ApA-PaP **1.16** was reported to have been observed as both $[M+Na]^+$ and $[M\text{-ketene}+Na]^+$ during the analysis of samples >100 pmol, although no data was provided to support this claim. Tandem MS (MS^2) analysis of ApA-PP **1.15** resulted in inconsistent and unidentifiable fragmentation of the probe that would make PMF analysis extremely difficult. It was postulated by Yoneda et al. that the complex fragmentation did not occur for ApA-PAP **1.16** due to the amido group providing a degree of stabilisation to the probe.

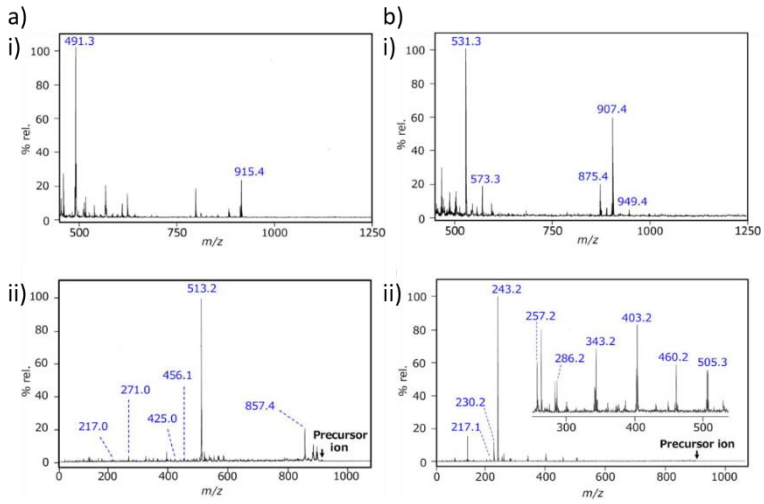


*Figure 1.16: LALDI-MS and MS^2 spectra of the photoaffinity probes developed by Yoneda et al. photocrosslinked to methanol. a) ApA-PP **1.15** (1 pmol) i) LALDI-MS spectrum – highlighted ion peaks were attributed to fragments cleaved at the N-O oxime linkage (m/z 915.1) and at the C-N amide bond (C1) of the lysine linker (m/z 491.3). ii) LALDI-MS^2 spectrum (precursor ion: m/z 915.4) – highlighted ion peaks were attributed to fragments cleaved at C-N amide bonds (m/z 217.1, 286.2, 403.2, 505.3), C-O bonds in the OEG chain (m/z 343.2, 460.2), C-C bonds of the butyramide linker (230.2, 257.2), and a McLafferty rearrangement of the butyramide linker (m/z 243.2). b) ApA-PaP **1.16** (1 pmol): i) LALDI-MS spectrum – two pairs of characteristic fragment ions with differences of m/z 42 (loss of ketene) were observed for fragments cleaved at the oxime N–O oxime linkage (m/z 949.4/907.4) and at the α C-C bond of the lysine linker (m/z 573.3/531.3). An additional peak was assigned to a fragment cleaved at the ε C-N bond of the lysine linker (m/z 875.4). ii) LALDI-MS^2 spectrum (precursor ion: 907.4) – ion peaks highlighted were attributed to fragments cleaved at C-N amide bonds (m/z 857.4, 513.2, 271.0, and 217.0) and a C-O bond in the OEG chain (m/z 456.1). Reproduced from K. Yoneda, Y. Hu, M. Kita and H. Kigoshi,"6-Amidopyrene as a label-assisted laser desorption/ionization (LA-LDI) enhancing tag: development of photoaffinity pyrene derivative", Scientific Reports, **5**, 17853, (2015) under Creative Commons Attribution 4.0 International License - published by Nature Publishing Group.*[100]

Attempts were made to label actin using ApA-PAP **1.16** via photocrosslinking. However, very little of the photocrosslinked product was observed.[106] To resolve this, the photo-reactive crosslinking group of **1.16** was substituted for an *N*-hydroxysuccinimide (NHS) ester to give the aplyronine A photoaffinity amidopyrene *N*- hydroxysuccinimide ester (ApA-AP-OSu) **1.17** (*Figure 1.15*), which was successfully used to label actin through amine-reactive crosslinking. The

labelled protein was subjected to enzymatic digestion with trypsin and Glu-C and analysed by both MALDI- and LALDI-MS (*Figure 1.17*). The amidopyrene labelled peptides could be easily detected by LALDI-MS and using MS^2 it was possible to confirm that ApA-AP-OSu **1.17** had reacted exclusively with the K113 residue of actin. However, the specificity of the amine-reactive crosslinking meant that it was not possible to attain any information about the ApA-actin binding site. Furthermore, the use of a long linker could also result in crosslinking at positions away from the binding site.

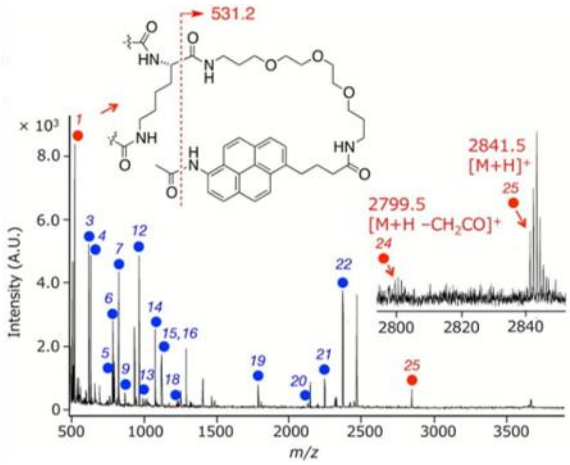
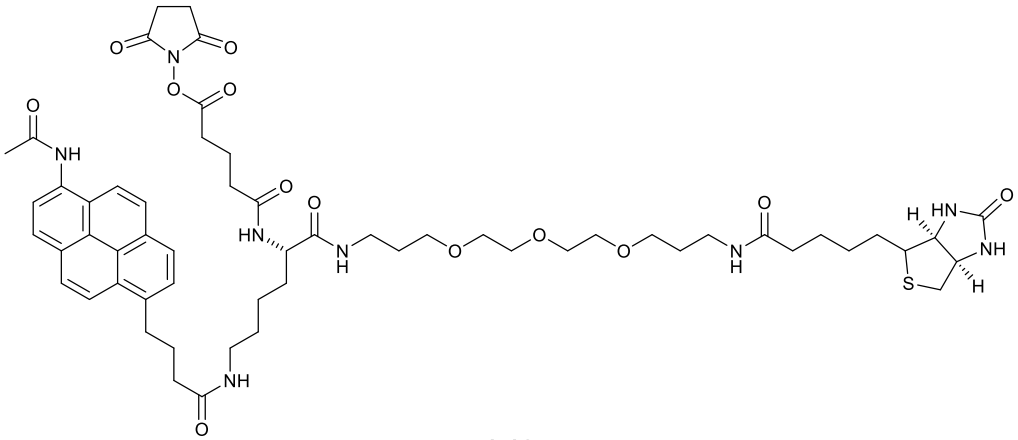


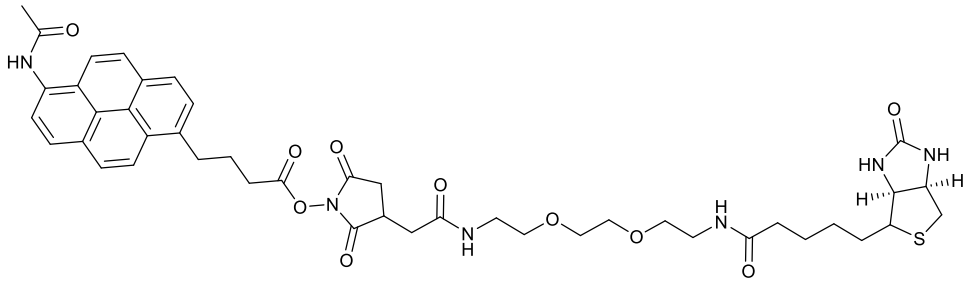
*Figure 1.17: LALDI-MS spectrum of the digested peptide fragments of ApA-AP-OSu **1.17**-labelled actin following treatment with trypsin and Glu-C protease enzymes. The signal at m/z 2841.5 corresponds to the precursor ion of the ApA-AP-OSu **1.17**-labelled peptide fragment. Red and blue circles indicate labelled and unlabelled peptides, respectively. Numerals indicate positions of digested peptides in actin arbitrarily defined by Yoneda et al. in their publication. Reproduced from K. Yoneda, Y. Hu, R. Watanabe, M. Kita and H. Kigoshi, "Binding position analysis of target proteins with the use of amidopyrene probes as LA-LDI enhancing tags", Organic & Biomolecular Chemistry, **14**, 8564–8569, (2016) with permission from The Royal Society of Chemistry.*[106]

A second study reported by Watanabe et al. focused on the identification of the binding position of biotin guided NHS esters on avidin.[107] Two amidopyrene-functionalized biotin affinity probes were created, a ligand-bound affinity probe Ap-biotin-OSu **1.18** and a ligand-dissociation affinity probe Ap-OSu-biotin **1.19** (*Figure 1.18*).

A sample of avidin was labelled using each probe, subjected to tryptic digestion, and then analysed by MALDI-MS. It was found that both **1.18** and **1.19** reacted specifically at the K135 residue of avidin. Detection of the peptide to which Ap-OSu-biotin **1.19** was bound could also be achieved using LALDI-MS, although no LALDI-MS data were provided for any samples labelled with Ap-biotin-OSu **1.18**. As with the previous study on ApA-actin it was not possible to identify the binding site of biotin on avidin using this technique.



1.18



1.19

*Figure 1.18: Chemical structures for the ligand-bound affinity probe Ap-biotin-OSu **1.18** and the ligand-dissociation affinity probe AP-OSu-biotin **1.19** used in the study by Watanabe* et al.[107]

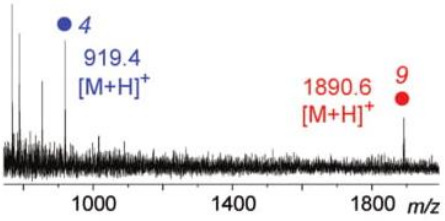
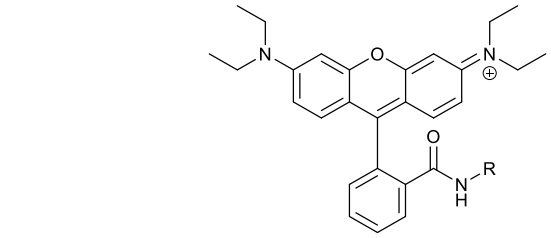


*Figure 1.19: LALDI-MS spectrum of the digested peptide fragments of Ap-OSu-biotin **1.19**-labelled avidin. The signal at m/z 1890.6 corresponds to the precursor ion of the Ap-OSu-biotin **1.19**-labelled peptide fragment. Red and blue circles indicate labelled and unlabelled peptides, respectively. Numerals indicate positions of digested peptides in actin arbitrarily defined by Watanabe et al. in their publication. Reproduced from R. Watanabe, Y. Hu, K. Iio, K. Yoneda, A. Hattori, A. Arai, H. Kigoshi and M. Kita, "Specific protein-labelling and ligand-binding position analysis with amidopyrene probes as LDI MS tags", Organic & Biomolecular Chemistry, **16**, 7883–7890, (2018) with permission from The Royal Society of Chemistry.*[107]

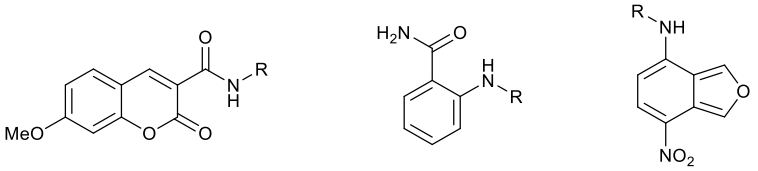
While neither study was able to identify the ligand-protein binding site, as originally intended, Kigoshi and co-workers were able to identify which residue each probe crosslinked to on both actin and avidin. This demonstrates that PMF could be achieved using LALDI-MS and indicates that ligand-protein binding site identification could be possible by LALDI-MS using alternative methods, such as an optimized photoreactive crosslinking probe. Additionally, the work described by the Kigoshi group has provided a valuable insight into aspects such as the MS stability and sensitivity of LDI enhancers, their applications in aqueous environments, and the

desorption and ionisation of larger biomolecules (2.5-3 KDa). All of this will be useful for any future work on analysing biomolecules by LALDI-MS. Despite achieving efficient and site-specific labelling for both actin and avidin with their amidopyrene LDI enhancers, there were no reports of any attempts to analyse the intact proteins by LALDI-MS.

1.2.2.4 Fluorophore-assisted laser desorption/ionization (FALDI)



1.20



1.21 **1.22** **1.23**

Figure 1.20: Structure of fluorophores used by Isailovic et al. to carry out FALDI-MS. Structure of HiLyte Fluor™ 555 was undisclosed by the manufacturer.[111,112]

Work carried out by Isailovic et al. showed that LALDI could be carried out at visible wavelengths of light as well as in the UV range using a fluorophore with extended conjugation.[111] Hence, in work by Isailovic and co-workers, they refer to this as fluorophore-assisted laser desorption/ionization (FALDI). In the first study by Isailovic et al., rhodamine B **1.20** and commercially available HiLyte Fluor™ 555 were chosen as the LDI enhancers. The solubility of the FALDI probes was found to be significantly better in aqueous solutions than other reported LDI enhancer, which would be important for any studies in biological systems. However, both fluorophores were found to have poor sensitivity and appeared to act as a weak matrix, ionizing other species and impurities that were not of interest present in the sample and complicating analysis of the mass spectrum (*Figure 1.21*).

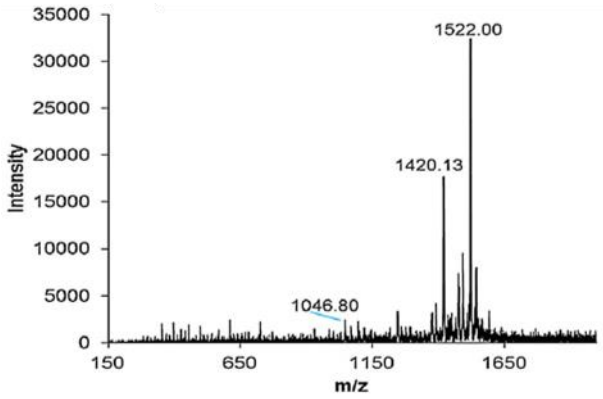


*Figure 1.21: FALDI-MS spectrum of a mixture of two distinct rhodamine B **1.20** labelled peptides, rhodamine B-TLLYVLFEV-COOH (m/z 1522.00) and rhodamine B-LLYVLFEV-COOH (m/z 1420.13), and angiotensin II (m/z 1046.80). The detection of the unlabelled angiotensin II suggests that the rhodamine B-labelled peptides also act as a weak matrix. Reproduced from International Journal of Mass Spectrometry, **353**, R. E. West, E. W. Findsen and D. Isailovic, "Fluorophore-assisted laser desorption/ionization-mass spectrometry (FALDI-MS)", pp 54–59, Copyright (2013) with permission from Elsevier.*[111]

Despite this, Isailovic and co-workers claimed that FALDI could be carried out with any fluorescently labelled compounds that can absorb at the wavelength of excitation. This claim was addressed in a second study by Isailovic and co-workers, which focused on the detection of other fluorescently labelled biomolecules such as proteins, oligosaccharides, and lipids.[112] Coumarin **1.21** labelled insulin and cytochrome C were both successfully detected using FALDI-MS, with each protein detected as a mixture of multiple protein-coumarin conjugates with masses that corresponded to the number of coumarin labels bound. In both cases, no signals were detected for either the free label or unlabelled protein. The FALDI-MS spectrum of cytochrome C labelled with coumarin **1.21** reported by West et al. has been reproduced in *Figure 1.22*.

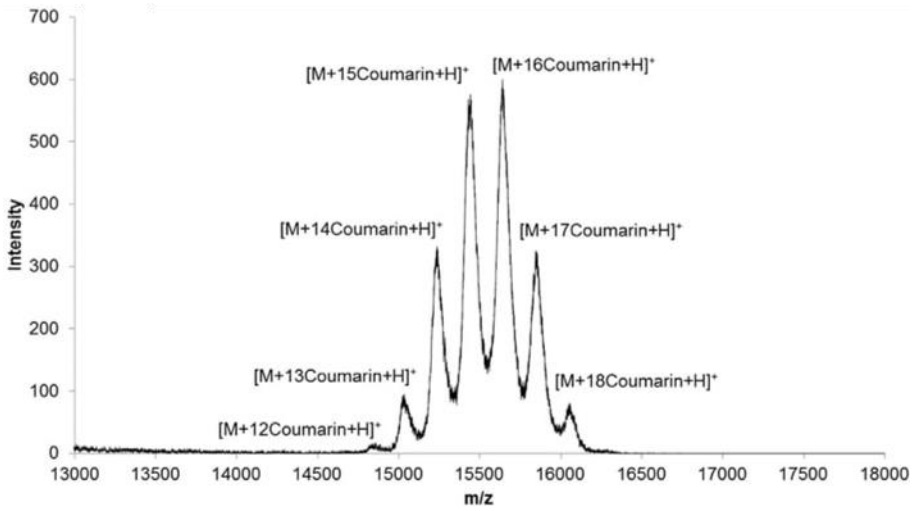


*Figure 1.22: FALDI-MS spectrum of cytochrome C labelled with coumarin **1.21** that shows the presence of multiple protein-coumarin conjugates. Reproduced from International Journal of Mass Spectrometry, **389**, R. E. West, J. B. Jacobs and D. Isailovic, "The analyses of fluorescently labelled biomolecules using ultraviolet-wavelength fluorophore-assisted laser desorption/ionization-mass spectrometry (UV FALDI-MS)", pp 39–46, Copyright (2015) with permission from Elsevier.*[112]

A brief investigation into FALDI-MS analysis of other fluorescently labelled biomolecules was also carried out, focusing on 2-aminobenzamide (2-AB)-labelled maltoheptaose **1.22** and nitrobenzoxadiazole (NBD)-labelled cholesterol **1.23**. Both fluorescently labelled biomolecules could be detected using FALDI-MS. However, 2-AB-labelled maltoheptaose **1.22** was found to have a number of unidentified higher *m/z* adducts (*Figure 1.23a*) and the spectrum for NBD labelled cholesterol **1.23** showed a number of high-abundance signals due to fragmentation (*Figure 1.23b*). In each case, the reported detection limits were relatively poor, with a detection limit of ~250 pmol for NBD labelled cholesterol, and ~150 nmol for 2-AB labelled maltoheptaose.

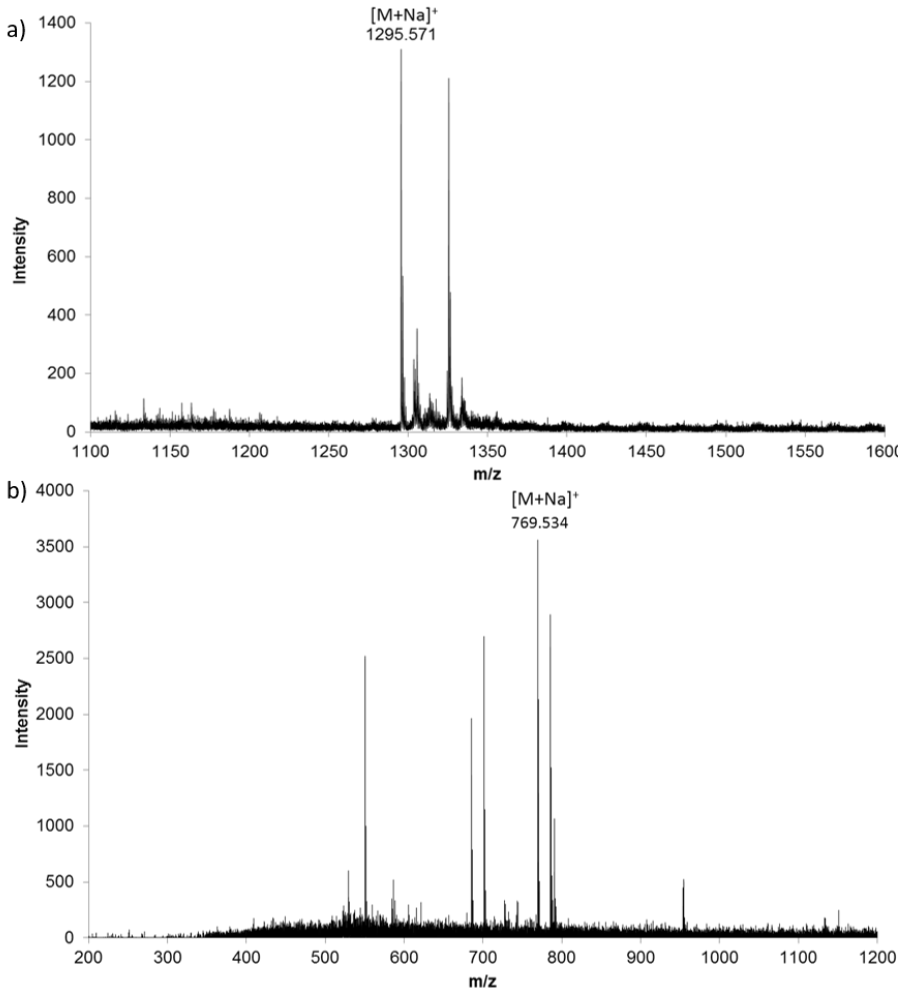


*Figure 1.23: FALDI-MS spectrum of a) NBD-cholesterol and b) 2-AB labelled maltoheptaose reported by West et al. Reproduced from International Journal of Mass Spectrometry, **389**, R. E. West, J. B. Jacobs and D. Isailovic, "The analyses of fluorescently labelled biomolecules using ultraviolet-wavelength fluorophore-assisted laser desorption/ionization-mass spectrometry (UV FALDI-MS)", pp 39–46, Copyright (2015) with permission from Elsevier.*[112]

The main advantage to FALDI is that there are numerous cheap and commercially available fluorophores that, in some cases, have already been optimised to work under biological conditions and target certain analytes. However, FALDI seems to be held back by the stability of its reagents and its poor detection limit, making it unsuitable for any applications that involve complex mixtures or require sub-picomole sensitivity, such as MS^n, glycomic analysis, or PMF.

Table 1.1: Summary of the applications, probes, and conditions of recent LALDI-MS studies

Application	Compound	Sample preparation solvent	Lowest reported detection (nmol)	Laser λ (nm)	MS instrument	Ionization Mode	Laser Source	Control Matrix Compound
High-throughput reaction discovery [94]	**1.3**	1,2-dichloro ethane	80.0	355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	N/A
Selective detection of zinc ions [101]	**1.8**	Dimethyl sulfoxide	2.0×10^{-1}	380	Applied Biosystems Voyager-DE PRO Time-of-Flight Mass Spectrometer	Positive Linear TOF mode	N_2 Laser	Sinapinic acid
Selective detection of cis -1,2-diols [102]	**1.9**	Na phosphate buffer (pH 7) : THF (1 : 1)	5.0	337	Applied Biosystems Voyager-DE PRO Time-of-Flight Mass Spectrometer	Positive Linear TOF mode	N_2 Laser	α -cyano-4-hydroxy cinnamic acid
Detection of biogenic amines [103]	**1.10**	MeOH	2.0×10^{-1}	355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	α -cyano-4-hydroxy cinnamic acid
Detection of metal ions by LALDI-MS and fluorescence [104]	**1.11**	MeCN	2.5×10^{-5}	355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	α -cyano-4-hydroxy cinnamic acid
Protein-ligand binding site identification [100,106,107]	**1.15**	50% Aq. MeOH or 1% TFA in MeCN	1.0×10^{-3}	355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	α -cyano-4-hydroxy cinnamic acid
	1.16		1.0×10^{-3}					
	1.17		0.2					
	1.18		3.5×10^{-1}					
	1.19		3.5×10^{-1}					
Detection and structural identification of peptides [111]	**1.20**	50% Aq. MeOH + 0.1% TFA	21.0	545	Thermo Finnigan LCQ Deca XPPlus ESI-IT-MS	Positive ion transmission geometry	Continuum Minilite I Nd:YAG laser	2-amino-3-nitrophenol
			21.6×10^{-3}	355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	Sinapinic acid
	HiLyte Fluor™ 555 labelled peptide		12.0	555	Thermo Finnigan LCQ Deca XPPlus ESI-IT-MS	Positive ion transmission geometry	Continuum Minilite I Nd:YAG laser	2-amino-3-nitrophenol
				355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	Sinapinic acid
Detection of proteins [112]	**1.21**	50% Aq. MeCN	160.0	355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	Sinapinic acid
Detection of maltoheptaose [112]	**1.22**		150.0	355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	N/A
Detection of cholesterol [112]	**1.23**		2.5×10^{-1}	355	Bruker Ultraflextreme MALDI–TOF	Positive ion reflector TOF mode	Bruker Smartbeam II (N_2, NdYag)	N/A

1.3 Glycomics

Carbohydrates are the most ubiquitous and structurally diverse group of molecules found in nature. Ranging from simple monosaccharides to large polymeric glycans and complicated glycoconjugates, these biomolecules are involved in almost every aspect of biology, both structural and functional.[113] In addition to the production and storage of energy, glycans are essential for the regulation of many biochemical pathways due to their role in processes such as cell signaling, cell adhesion, intercellular interactions, and molecular recognition.[114] Consequently, these factors make glycans excellent markers for disease that can be exploited for both diagnostics and therapeutics.[115] To fully understand the pathological and physiological significance of glycans within a biological system they must be studied collectively as a whole, and not just as individuals. The total complement of glycans and glycoconjugates that are present within a cell or organism is referred to collectively as 'the glycome'. Glycomics, therefore, describes investigations that focus on the identification of the structural and functional profile of the glycome.[116]

1.3.1 Challenges associated with glycomic analysis

Several analytical challenges can arise during glycomic investigations, one major barrier being that even the simplest glycans can be extensively isomeric. Due to the number of stereocenters present, monosaccharide can exist as multiple diastereomers. Larger monosaccharides, such as hexoses, can also cyclize via an intramolecular hemiacetal and exist as both ring-open and ring-closed forms. Depending on which hydroxyl group the hemiacetal forms with, cyclic monosaccharide can adopt different ring sizes, each of which can undergo mutarotation to give both an α and β anomeric form.[117] This isomerism increases exponentially as the oligosaccharides grows due to their potential to adopt multiple branched structures. A single reducing hexasaccharide made up of only D-hexose monosaccharides can have over a trillion possible linear and branched isomers.[118] In reality, very few isomers actually exist for each glycan due to the specific enzymatic nature of glycan biosynthesis.[119] However, identifying the structure and absolute stereochemistry of the isolated isomer can still prove laborious.

Similarities in size, polarity, and structure can make chromatographic separation difficult. Isolation can be achieved through a combination of techniques applied in succession.[120] However, detection of glycans by ultraviolet–visible (UV-Vis) spectroscopy during purification is fairly limited due to the absence of any chromo- or fluorophoric groups in native glycans. Structural analysis of both free glycans and glycoconjugates by X-ray crystallography is often

impractical due to the flexible and dynamic nature of oligosaccharides which makes crystallization difficult and results in conformational heterogeneity within the crystal.[121] Arguably, the most powerful technique for analyzing carbohydrates is nuclear magnetic resonance (NMR) spectrometry, which is capable of achieving both structural and stereochemical resolution.[122] However, its applications for glycomics are limited as it is not often possible to accrue enough material to perform the necessary NMR experiments. Unlike proteins and nucleic acids, the biosynthesis of carbohydrates is not template-driven or under transcriptional control. Hence, artificial amplification of specific glycans cannot be achieved, ultimately restricting the available material to the minute heterogeneous quantities obtained from natural sources.[123] Analysis by mass spectrometry, on the other hand, is far better equipped to deal with such small quantities. Especially when equipped with the means to trap and accumulate ions.[31] For this reasons, mass spectrometry coupled with chromatography is generally considered to be the most enabling tools for glycomics.[27,28]

1.4 Analysis of glycans by mass spectrometry

1.4.1 Ionization techniques and mass analysis

The two primary modes of ionization used to carry out MS analysis of glycans are ESI and MALDI.[124] The general advantages and disadvantages of MALDI have been discussed in Section 1.1.2. However, a specific problem encountered when analyzing native carbohydrates by MALDI is that, at the requisite low pressure, labile groups such as sialic acids, uronic acids, sulfates, and phosphates dissociate quite readily and fail to be accounted for. This dissociation is particularly problematic for monosaccharides.[125–127] However, chemical derivatizations have been developed to help stabilize these glycans and prevent this fragmentation from occurring (Section 1.4.5).

ESI involves the use of a strong voltage to vaporize and ionize a droplet of solubilized analyte as it is sprayed through a fine needle. While ESI is generally a gentler approach to ionization than MALDI, it can generate multiply charged species and complicated spectra.[26] However, ESI is generally more suitable for profiling intact native glycans as the labile acidic groups are reported to fragment less.[28] Furthermore, spectra produced by ESI are not obscured by the presence of signals relating to a matrix due to their absence from the sample preparation.[128] However, when using ESI, expensive commercial glycan standards are often required as an internal standard.[129] ESI is far more sensitive to the presence of salts than MALDI, so desalting is required prior to

analysis to prevent ion suppression and multiple adduct formation. However chromatographic separation can be integrated into the mass spectrometer to improve detection.[27,130]

The optimal mass instruments for analyzing the complex mixtures encountered in glycomics would be either FT-ICR, or orbitrap, both of which can be readily coupled with MALDI and ESI.[46,50] In both systems, the ions are trapped indefinitely in a cyclotron orbit and their *m/z* is determined from their ion cyclotron frequency. This allows analysis to be carried out on smaller analytical samples and provides a greater degree of accuracy and resolution than other methods such as TOF or QIT.[50]

1.4.2 Glycan release

Prior to MS analysis, glycans are usually cleaved from their respective glycoconjugates. This is done to help determine the number and variety of structurally different glycans within a sample, and is central to the characterization of the glycome.[115] The two most common types of glycosylation are *N*- and *O*-linked glycosylation (*Figure 1.24*).

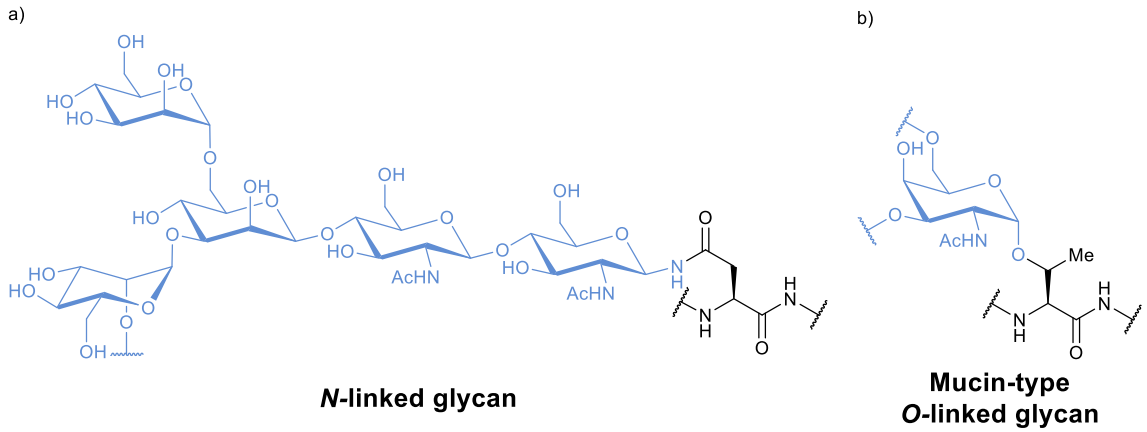


Figure 1.24: a) Example of an N-linked glycan showing the glycosidic linkage between the conserved β-linked $Man_3GlcNAc_2$ pentasaccharide core structure and an Asn residue. b) Example of a mucin-type O-linked glycan showing the glycosidic linkage between the conserved α-linked GalNAc and a Thr residue.[131]

All *N*-glycans begin with a common pentasaccharide core structure made up of two *N*-acetylglucosamine (GlcNAc) and three mannose (Man) residues. *N*-glycosidic linkages form between the amino side chain of asparagine (Asn) and the terminal β-linked *N*-acetylglucosamine (GlcNAc) (*Figure 1.24a*).[131] *O*-glycosidic linkages do not have a shared glycan core structure, affording them a much greater degree of structural diversity. Mucin-type *O*-glycosylations are the most commonly encountered and formed between an α-linked *N*-acetylgalactosamine (GalNAc) and the hydroxyl side chains of serine (Ser) or threonine (Thr)

(*Figure 1.24b*). Other *O*-glycosylations include linkages between Ser/Thr and glucose, xylose, mannose, or fucose residues.[131] The cleavage of these glycosidic linkages can be achieved through either enzymatic or chemical techniques.[132]

Enzymatic cleavage of *N*-linked glycans from their Asn linkage can be achieved using *N*-glycanase enzymes such as peptide:*N*-glycosidase F (PNGase F) (*Figure 1.25a*).[132] The universally shared core structure of *N*-glycans allows PNGase F to be applied across almost all *N*-linked glycoproteins in the glycome.[132] However, PNGase F cannot cleave *N*-linked glycans that have a α1-3 fucosylation at the Asn-bound GlcNAc.[133] Cleavage of those fucosylated *N*-linked glycans can be achieved using peptide:*N*-glycosidase A (PNGase A).[133] A limited number of *O*-glycanases have been reported, although only endo-α-*N*-acetylgalactosaminidases D (endo-α-GalNAc-ase D) is commercially available. The applications of this enzyme are also considerably restricted due to its high specificity for only one *O*-linked disaccharide, Gal(β1-3)GalNAc(α1-O)Ser/Thr (*Figure 1.25b*).[132]

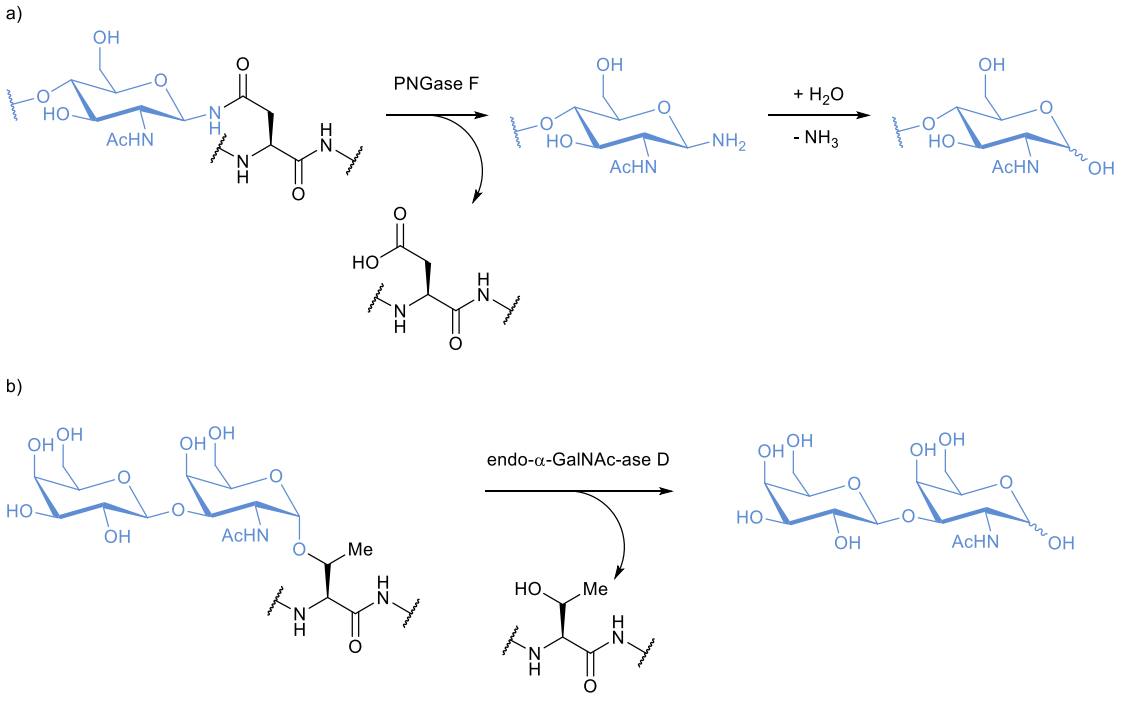


Figure 1.25: A schematic representation of enzymatic release of N- and O- glycans from their glycoconjugates. a) Enzymatic cleavage of N-linked glycoproteins at the GlcNAc(β1-N)Asp linkage using PNGase F. b) Enzymatic cleavage of O-linked glycoproteins at the Gal(β1-3)GalNAc(α1-O)Thr linkage using endo-α-GalNAc-ase D.[132]

O-glycans are typically released using chemical cleavage. Most commonly, base catalyzed β-eliminations using NaOH are used to selectively release *O*-glycans from their Ser/Thr linkages (*Figure 1.26*).[28] However, unwanted degradations known as "peeling" can occur, resulting in monosaccharides being cleaved from the reducing end.[134] The addition of sodium borohydride is often used to try and prevent this by reducing the glycan to the corresponding alditol. However, this removes the reducing end from the *O*-glycan, making further derivatizations more difficult.[28] Alternatively, β-eliminations of *O*-glycans can be achieved using NH_4OH, which can be easily removed following the reaction by evaporation due to the volatility of the reagent, eliminating the need for desalting steps and greatly simplifying sample clean-up.[135] Modification of the amino acid residues at glycosylation sites is known to occur during chemical cleavage with NH_4OH, allowing positive and unambiguous *O*-glycosylation site identification.[135] It was proposed that, following glycan release, NH_3 adds across double bond of the dehydroalanine cleavage product via a Michael-type addition reaction.[134]

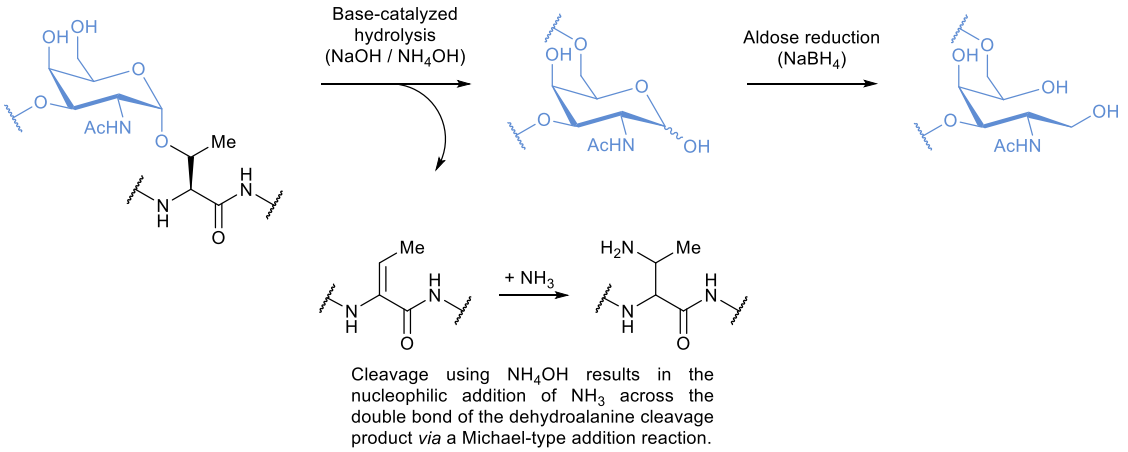


Figure 1.26: A schematic representation of chemical cleavage of O-linked glycoproteins via base-catalyzed hydrolysis. Release of the O-glycan chain from its glycoconjugate occurs via β-elimination under basic conditions. Released glycans can then be reduced using $NaBH_4$ to prevent degradation under the basic conditions.[132] *Glycolytic cleavage using NH_4OH, results in the formation of modified non-natural amino acid residues at hydrolyzed O-glycosylation site(s).*

Chemical release of *N*- and *O*-linked glycans can also be achieved using hydrazinolysis. Depending on the reaction conditions, the selectivity can be tuned to released *O*-glycans selectively, or both *N*- and *O*-glycans simultaneously.[134] While studies have deduced that hydrazinolysis of *N*-linked glycans occurs via nucleophile substitution of the Asn residue (*Figure 1.27a*), the exact reaction mechanism for hydrazinolysis of *O*-linked glycans has not yet been established.[132] It has been proposed that cleavage could occur via an initial β-elimination reaction, followed by the formation of hydrazone derivatives (*Figure 1.27b*).[136] Using hydrazinolysis, *O*-glycans have a greater chance of being recovered with the reducing end intact

compared to base-catalyzed β-elimination.[134,136] However, the associated safety risks of using anhydrous hydrazine make it less favourable.[134] Glycan release using hydrazine hydrate has been reported as a safer, but less efficient, alternative for large scale glycan production (1-1000 mg).[137]

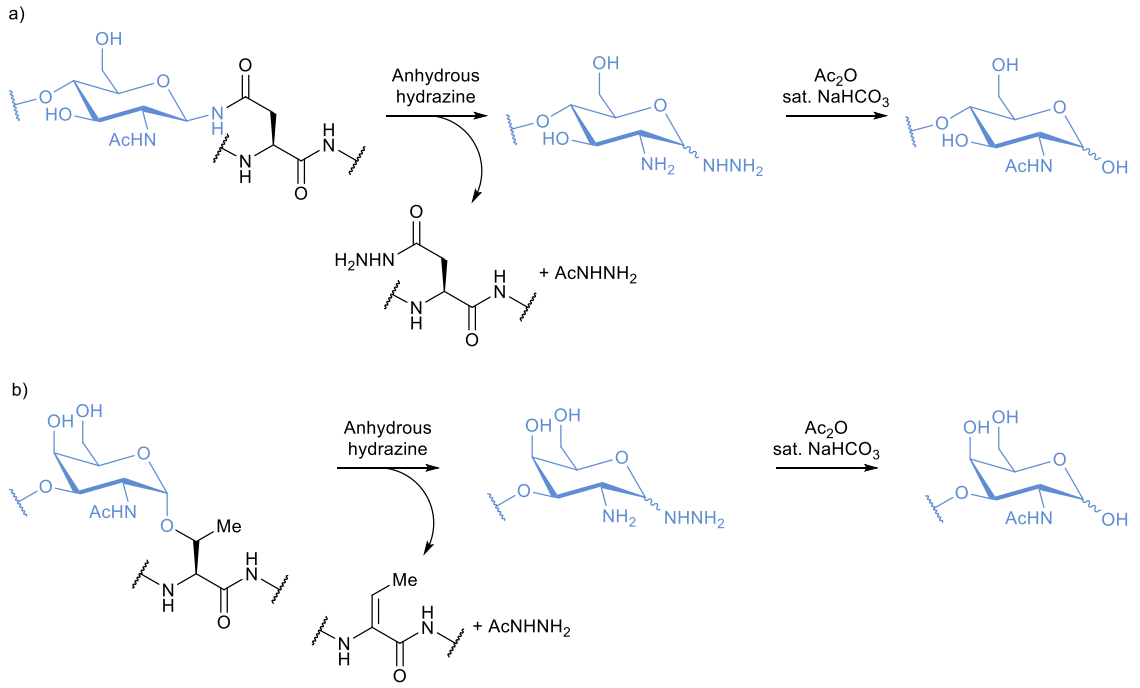


Figure 1.27: A schematic representation of chemical cleavage of a) N-linked and b) O-linked glycoproteins using the hydrazinolysis followed by re-N-acetylation using acetic anhydride.[132,136]

1.4.3 Separating mixtures of glycans

Separating mixtures of glycans prior to MS analysis is not always necessary but can help to reduce the complexity of a sample and deconvolute the analysis, especially when working with glycans derived from a biological source. However, due to the large variety of sterically similar and isomeric glycans that can be present in these mixtures, separation often demands high performance technique.[28] Commonly used chromatographic techniques include reversed-phase high-performance liquid chromatography (HPLC),[138] hydrophilic interaction chromatography (HILIC),[139] and porous graphitized carbon (PGC) chromatography.[140] As PGC separates glycans using electrostatic forces and the hydrophobic effect, it particularly useful for separating isomeric glycans.[140] Each of these can be integrated into ESI-MS to streamline the analytical workflow. Ion mobility (IM) can also be coupled to MS as an additional option for resolving mixtures of isomeric glycans.[141] IM-MS separates glycans prior to mass analysis based on their mobility through a drift tube filled with a buffer gas. Compact glycans travel more easily through the gas and leave drift tube in less time than their more branched isomers. The resulting data

are then displayed as a 2D plot of *m/z* vs drift time to provide a quick method of identifying the number of isomers in a sample.[141]

1.4.4 Structural elucidation techniques

Various MS techniques can be applied when studying glycans in order to discern their structural information. The simplest of these is just basic mass analysis, which can be used to identify the monoisotopic mass and subsequent isotopic distribution for each charged species. With a high enough accuracy, this information can be used to determine the molecular formula and serve as the first step towards structural elucidation. MS^n experiments can also be particularly useful for obtaining structural information about glycans through their fragmentation. Glycan sequence composition can be established using fragmentation at the glycosidic linkages, and cross-ring fragmentation can provide details about how the glycans are linked. Common examples of MS^n experiments are collision-induced dissociation (CID),[142] infrared multiphoton dissociation (IRMPD),[143] and electron capture/excitation/detachment dissociation techniques.[144–146] To obtain specific information about the structure of glycans, MS and MS^n can be coupled with techniques such as enzymatic sequencing and permethylation linkage analysis.[28,147]

Enzymatic sequencing of purified oligosaccharides can be achieved using specific exoglycosidase enzymes.[147] These enzymes sequentially cleave the monosaccharides from the non-reducing end of the glycan, but do not affect the internal bonds of the oligosaccharide chain. By carrying out multiple rounds of digestions using an array of exoglycosidases that are specific to various monosaccharides, enantio- and anomeric configurations, and glycosidic linkages, it is sometimes possible to learn a great deal about the structure of the glycan.[148] However, not all structural features can be resolved using this technique.[147]

Permethylation linkage analysis is a technique for gathering information about the location and types of glycosidic linkages using a combination of chemical derivatization and gas chromatography-mass spectrometry (GC-MS).[28] In linkage analysis, native glycans are permethylated then chemically hydrolyzed to yield a collection of partially methylated monosaccharides with free hydroxyl groups at the positions involved in the glycosidic linkages. A reducing agent is used to fix the partially methylated monosaccharides in the ring-opened conformation. The use of a deuterated reducing agent at this stage can help identify the reducing end of the monosaccharide.[149] Finally, the remaining free hydroxyl groups are capped with a different group, and a combination of MS^n fragmentation patterns and GC retention times

are used to piece together the structural information. Linkage analysis is useful for identifying features such as the terminal residues, the position and number of substitutions, monosaccharide ring size, and glycosidic linkages and branching points. However, it cannot provide information about the anomeric conformations of the monosaccharides or the specific glycans sequence.[54]

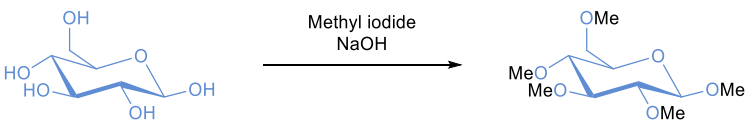
To achieve comprehensive glycomic analysis, multiple experiments using different analytical techniques should be used in parallel. Furthermore, additional resources such as computational prediction software,[150] and international glycomics databases[151] can be used to support investigations by comparing primary experimental data to both computational, and secondary experimental sources. This is especially useful when analyzing known glycans and glycoconjugates.

1.4.5 Improving analysis through chemical derivatization

Compared to proteins, the analysis and separation of native glycans is significantly more complicated. To address this, a wide range of chemical derivatizations have been developed. These derivatizations can be used to label glycans with permanent charges, or ionization enhancing groups to promote MS detection. Other techniques focus on improving the stability of glycans to prevent the loss of labile groups, enhancing sensitivity and supporting structural characterizations. Furthermore, derivatizations can also be used to facilitate purification by modifying the polarity of glycans to improve separation, introducing groups that enable UV-Vis detection, and providing ways to immobilize glycans onto solid surfaces.

A summary of the glycan derivatizations discussed in this section is represented in *Table 1.2.*

1.4.5.1 Permethylation

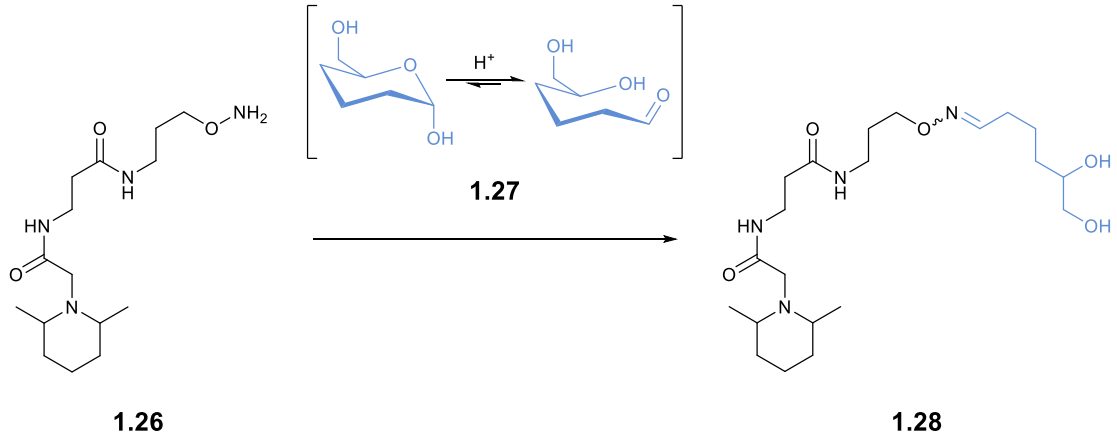


1.24 **1.25**

Scheme 1.2: An example of general reaction conditions used to modify glycans by permethylation, illustrated using β-D-glucopyranose. Under basic conditions, the hydrogen atoms of free hydroxy, amino, and carboxyl groups are substituted for methyl groups using methyl iodide.[152]

Permethylation (*Scheme 1.2*), a process of substituting the hydrogen atoms of free hydroxy, amino, and carboxyl groups with methyl groups, can also be used to achieve simultaneous analysis of both acidic and neutral oligosaccharides in positive mode,[153] stabilize labile acidic groups in MALDI,[154] facilitate purification by reducing the polarity of glycans by making separation more reproducible, and has even been shown to facilitate quantitative analysis.[155] As previously mentioned, permethylation is key to carrying out linkage analysis and identifying glycan structures and linkages.[54] However, incomplete derivatizations can result in heterogeneous mixtures with variable ionization efficiencies and fragmentation patterns, resulting in convoluted spectra and complicated analysis.[152]

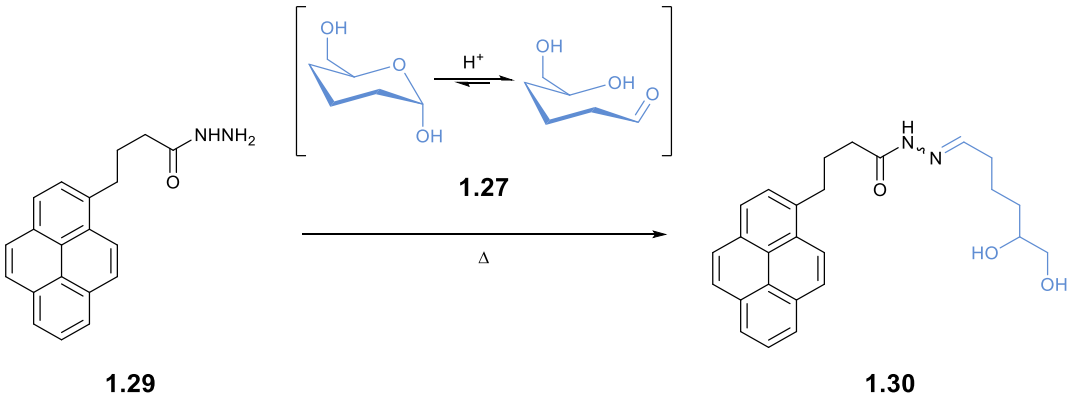
1.4.5.2 Oxime ligation



Scheme 1.3: An example of the general conditions required for labelling glycans by oxime ligation. Illustrated using a generic α-pyranose and an aminooxy tandem mass tag (TMT) used for isobaric labelling. Under acidic conditions, the hydroxylamine of the label reacts at the reducing end of the glycan through nucleophilic addition. The subsequent elimination of water from that position then gives the corresponding oxime.[156]

Oxime ligation is a chemoselective and bioorthogonal approach to labelling glycans at their reducing end using a hydroxylamine-functionalized probe under acidic conditions (*Scheme 1.3*).[157] The low pH is important as it shifts the conformational equilibrium of the reducing glycan into its open-ring conformation and generates the carbonyl at which the ligation occurs. Oxime ligations are typically high yielding, but reaction rates can be slow unless catalyzed. Aniline and anthranilic acid analogues have been reported as efficient catalysts for oxime ligations.[158,159] This type of derivatization has a broad range of applications including non-natural glycoconjugate synthesis,[160] solid-phase extraction,[161] immobilization on a solid surface for the development of glyconanoparticles[162] and microarrays,[163] and installing isobaric tags for multiplexing oligosaccharides.[156]

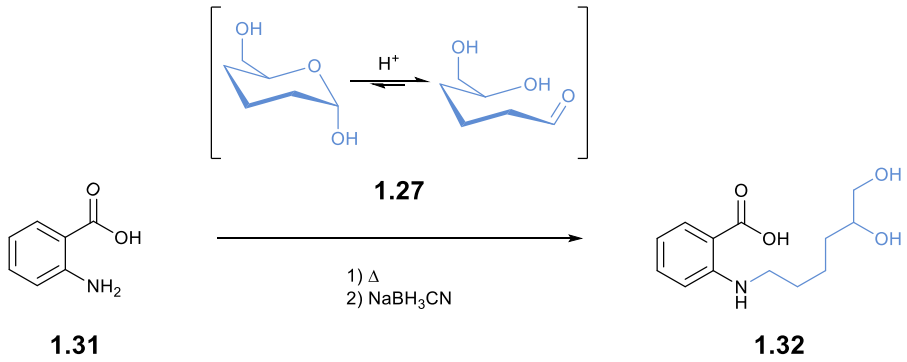
1.4.5.3 Hydrazone ligation



Scheme 1.4: An example of the general conditions required for labelling glycans by hydrazone ligation. Illustrated using a generic α-pyranose and pyrenebutyric hydrazide. Under acidic conditions and heating, the hydrazide of the label reacts at the reducing end of the glycan through nucleophilic addition. The subsequent elimination of water from that position then gives the corresponding hydrazone.[164]

Hydrazone ligation is another bioorthogonal labelling technique which involves labelling glycans at their reducing end using a hydrazide-functionalized probe under acidic conditions (*Scheme 1.4*).[157] Hydrazones are generally more labile than oximes, so present an increased risk of hydrolysis or hydrazone exchange in solution.[157] To prevent this, the hydrazone linkage can be optionally stabilized by reducing it to its corresponding hydrazide using sodium cyanoborohydride.[165] Reagents such as carboxymethyl trimethylammonium hydrazide (Girard's T reagent) provide a simple but effective method for improving MS detection by installing a permanent positive charge onto a glycan,[166] while biotin hydrazide can be used to immobilize glycans onto streptavidin-linked surfaces for carrying out biofunctional investigations.[167] Additionally, several studies by Amano et al. have found that glycans that were derivatized with 1-pyrenebutyric hydrazide (PBH) exhibited increased sensitivity and stability in MALDI-MS^n.[164,168,169] This was especially noticeable in negative mode MALDI-MS, which is normally completely unable to detect neutral oligosaccharides.[164] They concluded that the ionization was being enhanced by the pyrene, proposing that its strong aromatic nature facilitated the acceptance of energy from the matrix and stabilized the charge.[164] In addition to improving MS detection and stability, they also noted that pyrene derivatization resulted better shot-to-shot reproducibility of spectra due to better co-crystallization of the PBH-glycan and the 2,5-dihydroxybenzoic acid matrix.[170]

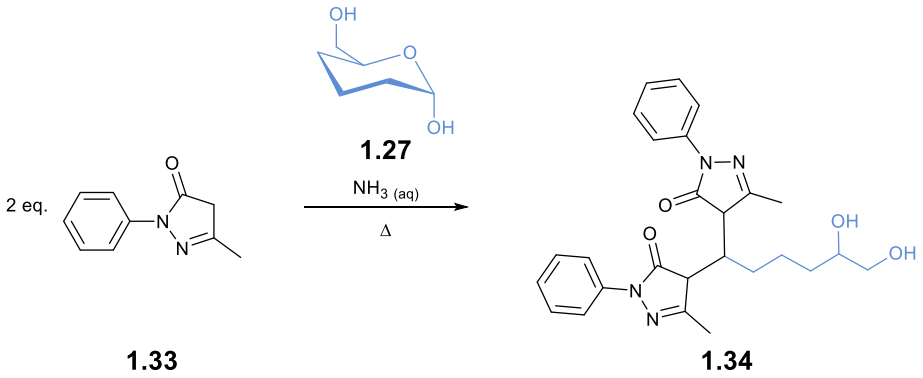
1.4.5.4 Reductive amination



Scheme 1.5: An example of the general conditions required for labelling glycans by reductive amination. Illustrated using a generic α-pyranose and 2-aminobenzamide. Under acidic conditions and heating, the primary amine of the label reacts at the reducing end of the glycan through nucleophilic addition. The subsequent elimination of a water from that position gives a secondary imine which is reduced to the secondary amine.[164]

Reductive amination can also be used to label glycans at their reducing end using a probe functionalized with a primary amine (*Scheme 1.5*). Under acidic conditions the reaction proceeds via an imine formation at the reducing end of a glycan which can then be reduced, typically using sodium cyanoborohydride, to yield a secondary amine.[171] Reductive amination is generally used to incorporate a fluorescent label onto the glycan to facilitate detection and separation,[172] although examples of quantitative glycomics have been reported through reductive amination using light and heavy stable isotope labels.[173,174] Labelling reagents such as 2-aminobenzamide (2-AB),[175] 2-aminobenzoic acid (2-AA),[175] 2-aminopyridine (PA)[176] are well known for permitting good separation and structural profiling by HILIC (2-AB, 2-AA) and HPLC (PA) using extensive databases of standardized elution times.[177,178] Labels that can confer fluorescence and multiple negative charges, such as 2-aminonaphthalene trisulfonic acid (ANTS),[179] and 1-aminopyrene-3,6,8-trisulfonic acid (APTS)[180], have excellent applications for separation by capillary electrophoresis,[181] and have been shown to improve sensitivity in negative mode MALDI sensitivity.[179,180]

1.4.5.5 Pyrazolone labelling



Scheme 1.6: An example of the general conditions required for labelling glycans by aldol condensation and Michael addition. Illustrated using a generic α-pyranose and 1-phenyl-3-methyl-5-pyrazolone. Under basic conditions the pyrazolone is deprotonated to form an enolate, which undergoes an aldol condensation with the reducing end of the glycan to form an α,β-unsaturated carbonyl compound. In the second step, a second deprotonated pyrazolone reacts with the α,β-unsaturated carbonyl by Michael addition to give the 1,4-addition product.[182]

A shared problem for oxime ligation, reductive amination, and hydrazine ligation is that the low pH required to efficiently label reducing glycans can result in the unwanted degradation of sialic acid groups and mischaracterization of the desialylated glycans.[171] Pyrazolones can be used to label glycans at their reducing end under basic conditions via a two-step process (*Scheme 1.6*).[182] Firstly, the pyrazolone is deprotonated to form an enolate, which attacks the reducing end of the glycan and undergoes aldol condensation to yield an α,β-unsaturated carbonyl. In the second step, a second deprotonated pyrazolone reacts with the α,β-unsaturated carbonyl to give the 1,4-addition product.[182] The benefit of this labelling technique is that the basic conditions significantly reduce the risk of desialylation. However, a drawback to this labelling technique is the need for two equivalents of label per glycan.[182] Typically, the applications of pyrazolone labelling have been limited to the introduction of a fluorophore,[182–184] although isotopic labelling for quantitative glycomics has also been reported.[185,186] Common pyrazolone labelling reagents include 1-phenyl-3-methyl-5-pyrazolone (PMP),[182] 1-(4-methoxy)-phenyl-3-methyl-5-pyrazolone (PMPMP),[187] and 1-(2-naphthyl)-3-methyl-5-pyrazolone (NMP).[182] A bifunctional amino pyrazolone moiety developed by Yu et al. has been shown to be capable of labelling glycans through both reductive amination and Michael addition depending on whether basic or acidic conditions are used.[188]

With the exception of permethylation, each of these derivatizations occur by modifying the reducing end of the glycan. However, as previously discussed, *O*-glycans released by reductive β-elimination are released as the alditol and cannot be derivatized through these techniques.[171]

1.4.5.6 Metabolic oligosaccharide engineering

First introduced by Bertozzi and co-workers,[189] metabolic oligosaccharide engineering (MOE) describes the metabolic labelling or chemoenzymatic tagging of glycans using an unnatural monosaccharide functionalized with a bioorthogonal chemical reporters.[131] These reactive handles can then be used to achieve a broad range of glycosylation-specific derivatization for visualization, purification, and enrichment using appropriate secondary probes.[190]

While MOE has become recognized as an important complementary tool for glycobiology, the incorporation of unnatural monosaccharide residues places this technique outside of the scope of native glycan labelling strategies. A comprehensive overview of this subject has been recently discussed in the review by Krishnan and Bertozzi,[131] and a detailed summary of the chemical probes currently available in the MOE molecular toolbox and the bioconjugation techniques they enable can be found in the review by Sminia et al.[190]

Table 1.2: Summary of native glycan derivatization techniques, featuring examples of reactants and commonly applied reaction conditions used to achieve derivatization of native glycans isolated from a biological sample.

Derivatization	Reactants	Additional reagents	Solvent	Temperature (°C)	Reaction time (h)
Permethylation [155]	Methyl iodide	-	DMSO : NaOH	25	1.5
Oxime ligation [191]	AminooxyTMT ONH_2-functionalized surfaces ONH_2-functionalized nanoparticles Aminoxy biotin	Aniline/anthranilic acid (optional catalysts)	Sodium phosphate buffer (pH 3.0)	25	4 (catalyzed) >24 (uncatalyzed)
Hydrazone ligation [165]	Biotin hydrazide Girard's T reagent PBH	$NaBH_3CN$ (optional reductant)	DMSO : AcOH	60	2-3
Reductive amination [175]	2-AB 2-AA PA ANTS APTS	$NaBH_3CN$ (reductant)	DMSO : AcOH	60	2-3
Pyrazolone labelling [187]	PMP PMPMP NMP	-	MeOH : $NH_{3\ (aq)}$	70	0.5

1.5 Project aims

Advancements in the field of glycomics have been central to understanding the complex biological roles of glycans and their impact on human health and development. Mass spectrometry in particular has been recognized as one of the most important tools for characterizing the structural and functional profile of the glycome. However, problems relating to the purification and ionization of glycans makes their MS analysis far more complicated than for other biomolecules. Therefore, efforts directed towards improving the MS analysis of glycans hold great potential for profiling disease, characterizing pathophysiology, and establishing novel approaches to diagnostics and treatment. Recently, LALDI has emerged as a relatively simple approach to achieving "matrix-free" LDI of various biomolecules, demonstrating selective detection of labelled analytes in the presence of complex samples without requiring prior purification. These properties could be of great value for overcoming some of the problems faced in MS-based glycomics.

The aim of this project was to develop LDI enhancers that can be used to selectively detect reducing glycans from a complex biological background using LALDI-MS. To achieve this, the project involved the creation of chemical reagents, referred to as "LALDI tags", that comprise an LDI-enhancing label conjugated to a glycan-selective reactive handle through a solubilizing linker (*Figure 1.28*). *Chapter 2* describes the initial development of a series of water-soluble LDI enhancers and the optimization of their detection by LALDI-MS. *Chapter 3* details the design, synthesis, and application of a series of LALDI tag reagents used to detect reducing glycans by LALDI-MS. *Chapter 4* looks at the development of a scalable synthesis of 6-amino-2-cyanobenzothiazole, a reactive handle for bioorthogonal labelling reactions with 1,2-aminothiol groups.

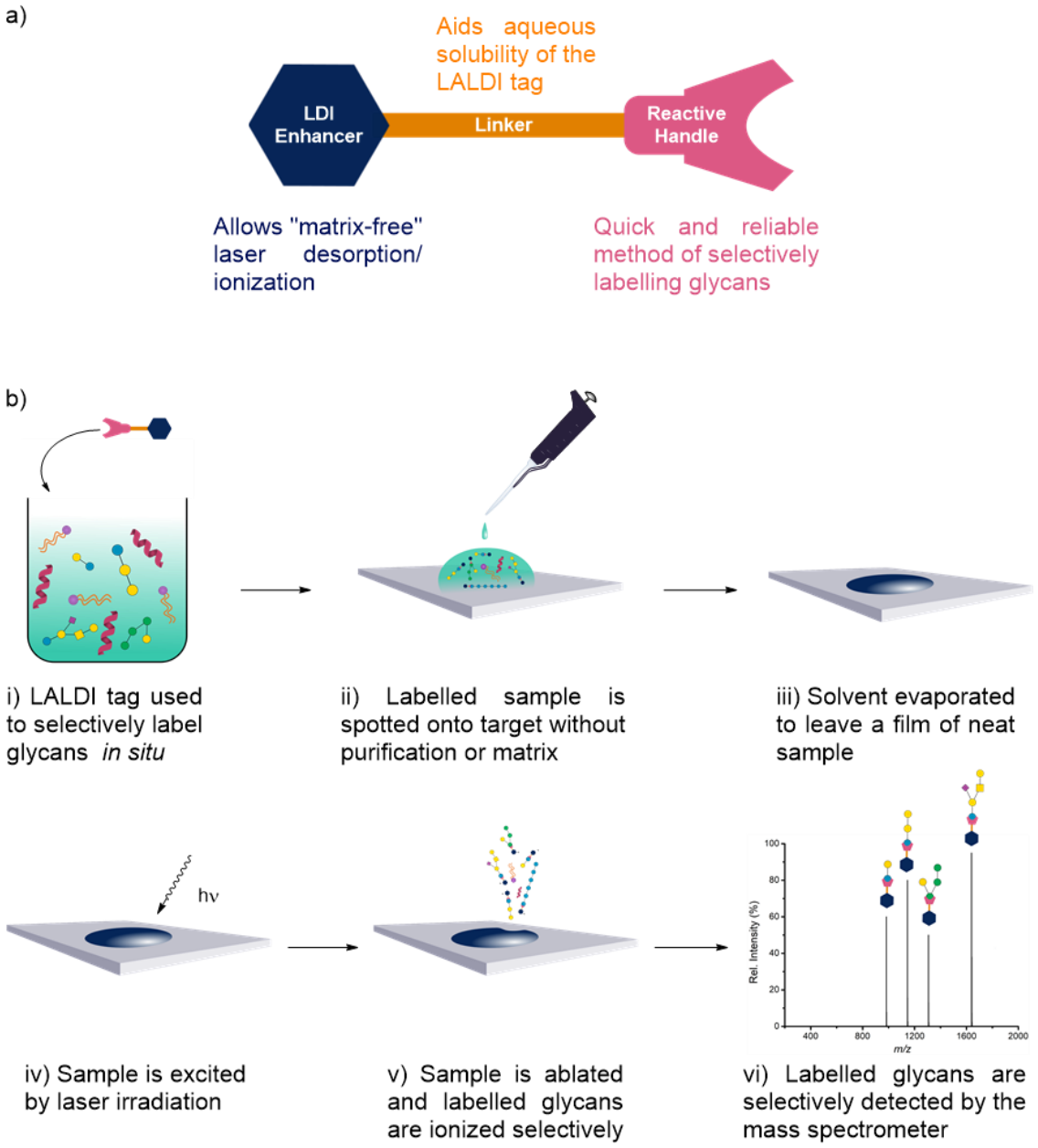


Figure 1.28: LALDI tags as a tool for MS-based glycomics a) General design of LALDI tag reagent incorporating an LDI-enhancing label (dark blue), a linker (orange), and a reactive handle (pink). b) Proposed general workflow for the analysis of glycans from biological samples using LALDI-MS.

Chapter 2 – Development of water-soluble LDI enhancers

2.1 Introduction

MALDI-MS is considered to be one of the most powerful analytical tools for studying high molecular weight bio-molecules, such as oligonucleotides, lipids, and glycoconjugates (Section 1.1).[26] The soft ionization and high throughput capabilities of MALDI-MS have made it particularly useful for a number of important clinical and bio-chemical applications (Section 1.1.4).[25,192,193] However, many of the problems and limitations encountered in MALDI-MS, have been attributed to matrix-related complications (Section 1.1.5). These limitations have prompted the development of several matrix-free approaches to LDI-MS (Section 1.1.6).

LALDI-MS has recently emerged as a novel approach to matrix-free LDI-MS, where desorption/ionization is facilitated by the incorporation of a polyaromatic label, or "LDI enhancer" into the analyte (Section 1.2).[1] Selective ionization and MS detection of the labelled analyte can then be achieved directly from complex samples, without requiring prior purification or an external matrix component (*Chapter 1 – Figure 1.7b*). Incorporation of the ionization enhancing label allows LALDI-MS to retain the benefits of MALDI-MS, such as soft ionization and fast run time, while simultaneously avoiding some of the matrix-related issues (Section 1.2.1).

Inspired by the initial LALDI-MS study by Cabrera-Pardo et al.,[94] which described the used of LALDI-MS to follow the progress of a chemical reaction, it was proposed that LALDI-MS could also be used to monitor bio-chemical processes directly from a complex biological environment. However, as very little research had been conducted on the direct applications of LALDI-MS for studying native biological samples, initial efforts in this project were directed towards establishing a reliable protocol for LALDI-MS analysis and the creation of a suitable LDI enhancer with the appropriate sensitivity and tolerance for aqueous biological environments.

A series of LALDI-MS investigations were undertaken to determine the optimum conditions for the detection and identification of labelled compounds without the need for an external matrix. Analysis was focused primarily on improving analytical parameters such as sensitivity, stability, and reproducibility, in order to optimize data and facilitate the analysis of more complicated labelled species.

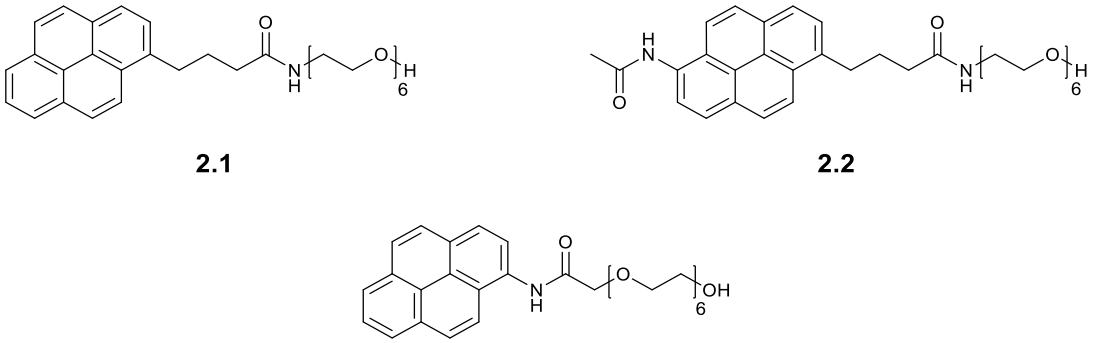
Pyrene-based LDI enhancers were selected for the preliminary experiments as their ability to promote laser desorption/ionization has been well established in the literature, with data available for the LALDI-MS sensitivity, stability, and expected mass spectra of various pyrene

derivatives (Section 1.2.2 – *Table 1.1*).[1,94,100,102–104,107] Calibration of the LDI mass spectrometer and optimization of analysis could then be facilitated through direct comparison of experimentally obtained data to results from the literature.

Finally, as the sensitivity and accuracy of a mass spectrometer are effectively defined by the mass analyzer used (Section 1.1.3), LALDI-MS analysis was conducted using a high-performance LDI-FT-ICR mass spectrometer (*Chapter 1 – Figure 1.5*). As FT-ICR offers the highest resolution of any other mass analyzer, it was considered ideal for optimizing the LALDI-MS analysis of biological samples and enabling the unambiguous identification of reagents even in the presence of a complex mixture.[194]

2.2 Optimization of LALDI-MS and calibration of LDI-MS instrument

Preliminary investigations into the optimization of LALDI-MS analysis were conducted using a series of reagents functionalized with three distinct pyrene-based LDI enhancers (*Figure 2.1*).



2.1 **2.2** **2.3**

*Figure 2.1: Chemical structure of pyrene **2.1**, 6-amidopyrene **2.2**, and 1-amidopyrene **2.3** used in the initial LALDI-MS optimization investigations.*

Two of these initial LALDI reagents, pyrene **2.1** and 6-amidopyrene **2.2**, featured LDI enhancers which had been previously studied by Yoneda et al.[100] Hence, **2.1** and **2.2** were chosen as references to help calibrate the available LDI mass spectrometer, highlight analytical parameters that could be optimized to improve data, and facilitate the analysis of more complicated labelled species.

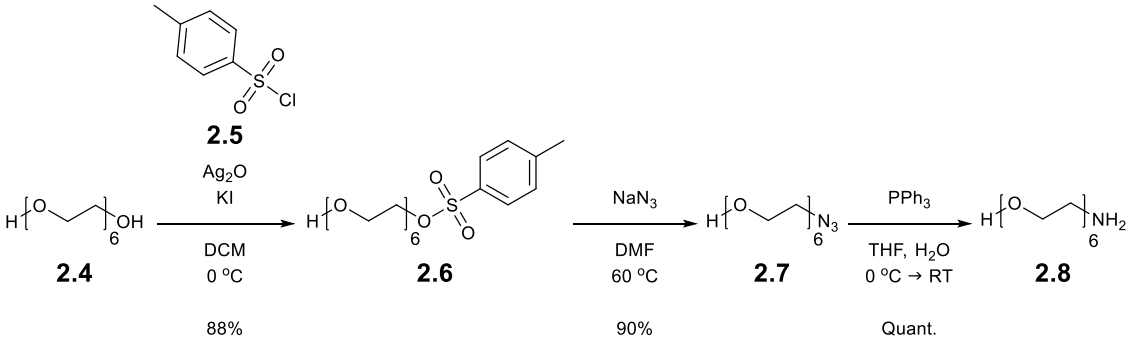
The third LALDI reagent **2.3**, featured a novel 1-amidopyrene LDI enhancer. 1-amidopyrene **2.3** was chosen to determine whether alternative amidopyrene LDI enhancers could be used to achieve the increased LALDI-MS sensitivity reported for 6-amidopyrenes, while avoiding the need for lengthy syntheses and MS-degradations from loss of ketene.[100]

Oligoethylene glycol (OEG) chains were incorporated into the design of each reagent to facilitate the solubility of the lipophilic pyrene labels in either aqueous or partially aqueous environments.

2.2.1 Synthesis of preliminary LALDI reagents

2.2.1.1 Pyrene butyramide **2.1**

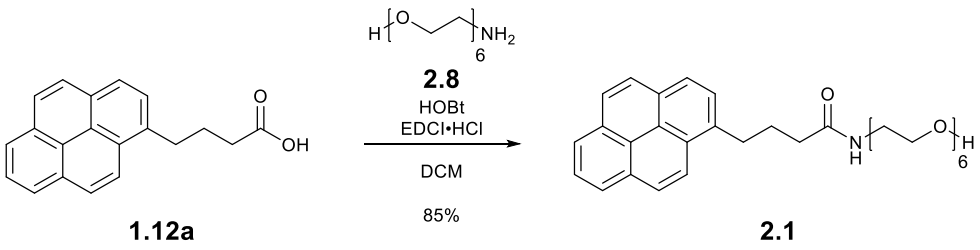
The OEG component of **2.1** and **2.2** was synthesized in three steps from hexaethyleneglycol **2.4** following literature procedures (*Scheme 2.1*).[195–198]



*Scheme 2.1: Synthetic route to 17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.8**.*[195–198]

Hexaethyleneglycol **2.4** was activated with silver(I) oxide for selective mono-substitution with tosylchloride **2.5** to afford tosylate **2.6** in 88% yield.[196] The tosyl group of **2.6** was then substituted via an S_N2 mechanism with sodium azide to afford azide **2.7** in 90% yield.[197] Finally, azide **2.7** was then reduced via a Staudinger reaction to afford amine **2.8** in quantitative yield.[198]

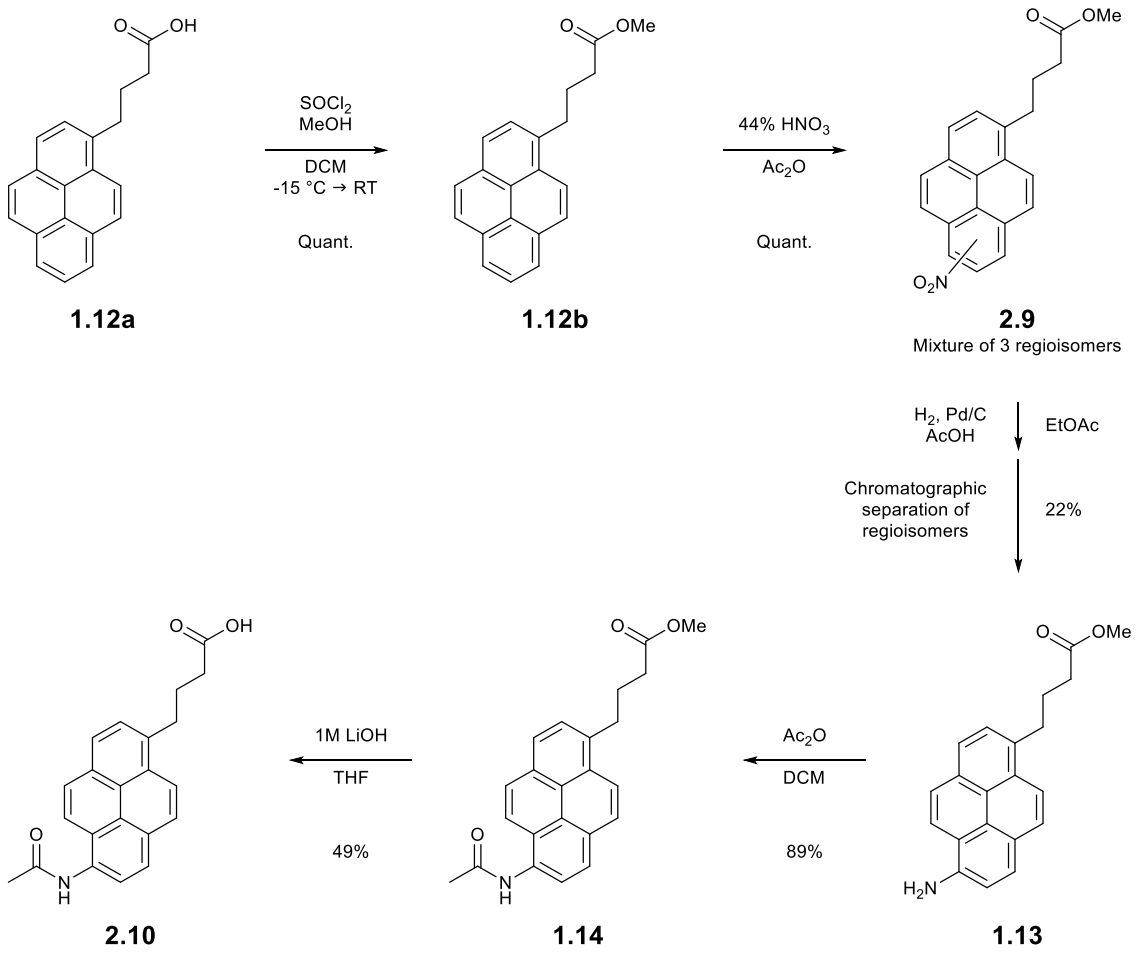
The synthesis of pyrene butyramide **2.1** was completed via an amide coupling reaction between pyrenebutyric acid **1.12a** and amine **2.8**, using hydroxybenzotriazole (HOBt) and *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC•HCl) as coupling reagents to afford alcohol **2.1** in 85% yield (*Scheme 2.2*).[199]



*Scheme 2.2: Reaction conditions for the synthesis of N-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-4-(pyren-1-yl)butanamide **2.1**.*[199]

2.2.1.2 6-amidopyrenebutyramide **2.2**

The 6-amidopyrene component of **2.2**, 4-(6-acetamidopyren-1-yl)butanoic acid **2.10**, was synthesized in 5 steps from pyrenebutyric acid **1.12a** following a procedure reported by Yoneda et al. (*Scheme 2.3*).[100]



*Scheme 2.3: Synthetic route to 4-(6-acetamidopyren-1-yl)butanoic acid **2.10**.*[100]

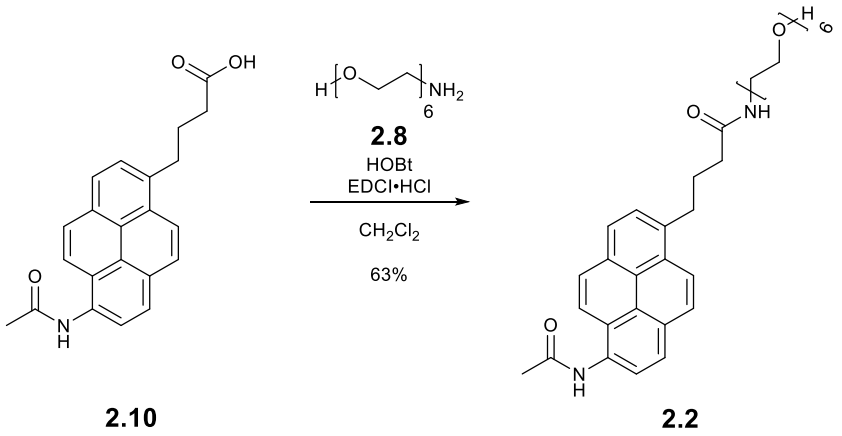
Firstly, the carboxylic acid group of **1.12a** was protected via esterification. This was accomplished by generating an acyl chloride intermediate in situ with thionyl chloride, followed by immediate substitution of the chloride group with methanol to afford methyl ester **1.12b** in quantitative yield.

Nitration of pyrene **1.12b** via electrophilic aromatic substitution was achieved using nitric acid and acetic anhydride to give nitro pyrene **2.9** as a 1:1:1 mixture of regioisomers in quantitative yield. The nitration reaction was confirmed to produce 3 regioisomers by ^{13}C NMR, which produced a spectrum with 3 distinct signals for each environment. The ratio of regioisomers was confirmed by measuring the integration of each isomer in the ^{1}H NMR spectrum.

Nitro pyrenes **2.9** were taken forward as the mixture of regioisomers and reduced via Pd-catalyzed hydrogenation to give amino pyrene **1.13**. Initial attempts at this reduction resulted in <10% yield and the formation of an insoluble precipitate. Yoneda et al. described amine **1.13** as having poor light stability and reported dimerization upon excitation with UV light.[100] Therefore, it was suspected that the insoluble precipitate was formed by light induced dimerization/degradation of the product. The reduction was repeated with all steps of the reaction and purification shielded from light to prevent decomposition of the product. Chromatographic separation of the regioisomers gave an isomerically pure sample of 6-aminopyrene **1.13** in 22% yield (66% isomeric yield) and a mixture of the other two regioisomers in 51% yield. 6-amino pyrene **1.13** was the only regioisomer to be taken forward in the synthesis as the other two regioisomers could not be separated. The structure of the 6-amidopyrene **1.13** was confirmed by comparing experimentally obtain spectra to examples reported in the literature.[100]

Acetylation of 6-amino pyrene **1.13** was achieved by stirring with acetic anhydride to afford 6-acetamido pyrene **1.14** in 83% yield. Methyl ester **1.14** was then deprotected by hydrolysis using LiOH to give 6-amidopyrene butyric acid **2.10** 49% yield.

The synthesis of 6-amidopyrene butyramide **2.2** was completed via an amide coupling reaction between 6-amidopyrenebutyric acid **2.10** and amine **2.8**, using HOBt and EDC•HCl as coupling reagents to afford alcohol **2.2** in 63% yield (*Scheme 2.4*).[199]



*Scheme 2.4: Reaction conditions for the synthesis of 4-(6-acetamidopyren-1-yl)-N-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)butanamide **2.2**.*[199]

2.2.1.3 Synthesis of 1-amidopyrene **2.3**

It was proposed that an alternative 1-amidopyrine LDI enhancer could be generated by conjugating 1-aminopyrene **2.12** to the OEG chain through an acyl linkage (*Figure 2.2*). The forward synthesis of 1-amidopyrene **2.10** is shown in *Scheme 2.5*.[200,201]

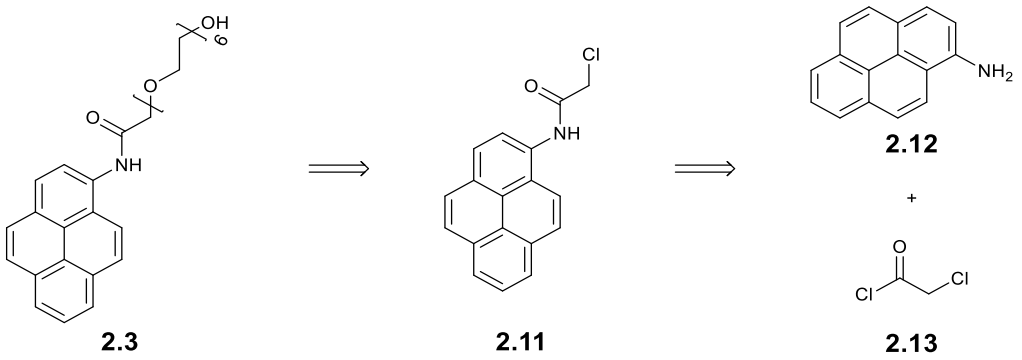
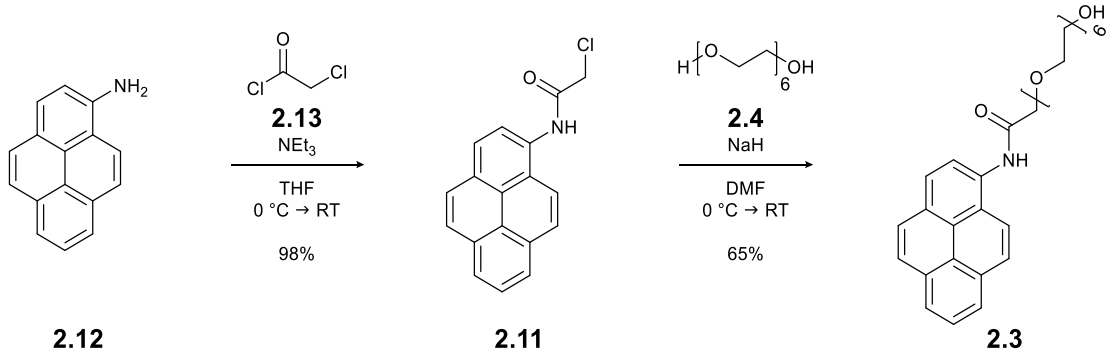


*Figure 2.2: Proposed retrosynthetic analysis of 20-hydroxy-N-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **2.3**.*

Following a procedure reported by Xu et al., 1-aminopyrene **2.12** was conjugated to chloroacetamide **2.13** via nucleophilic acyl substitution to afford chloroacetamide **2.11** in 98% yield.[200] Chloroacetamide **2.11** was then conjugated to hexaethylene glycol **2.4** under basic conditions via a Williamson ether synthesis to yield alcohol **2.3** in 65% yield.[201]



*Scheme 2.5: Synthetic route to 20-hydroxy-N-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **2.3**.*[200,201]

2.2.2 UV-Vis analysis of LALDI reagents

Following a precedent established by Yoneda et al.,[100] prior to LALDI-MS analysis each compound was first analyzed by UV-Vis spectroscopy. It was proposed by Yoneda et al. that an insight into the potential sensitivity of an LDI enhancer could be gained by measuring the UV-Vis absorption.[100]

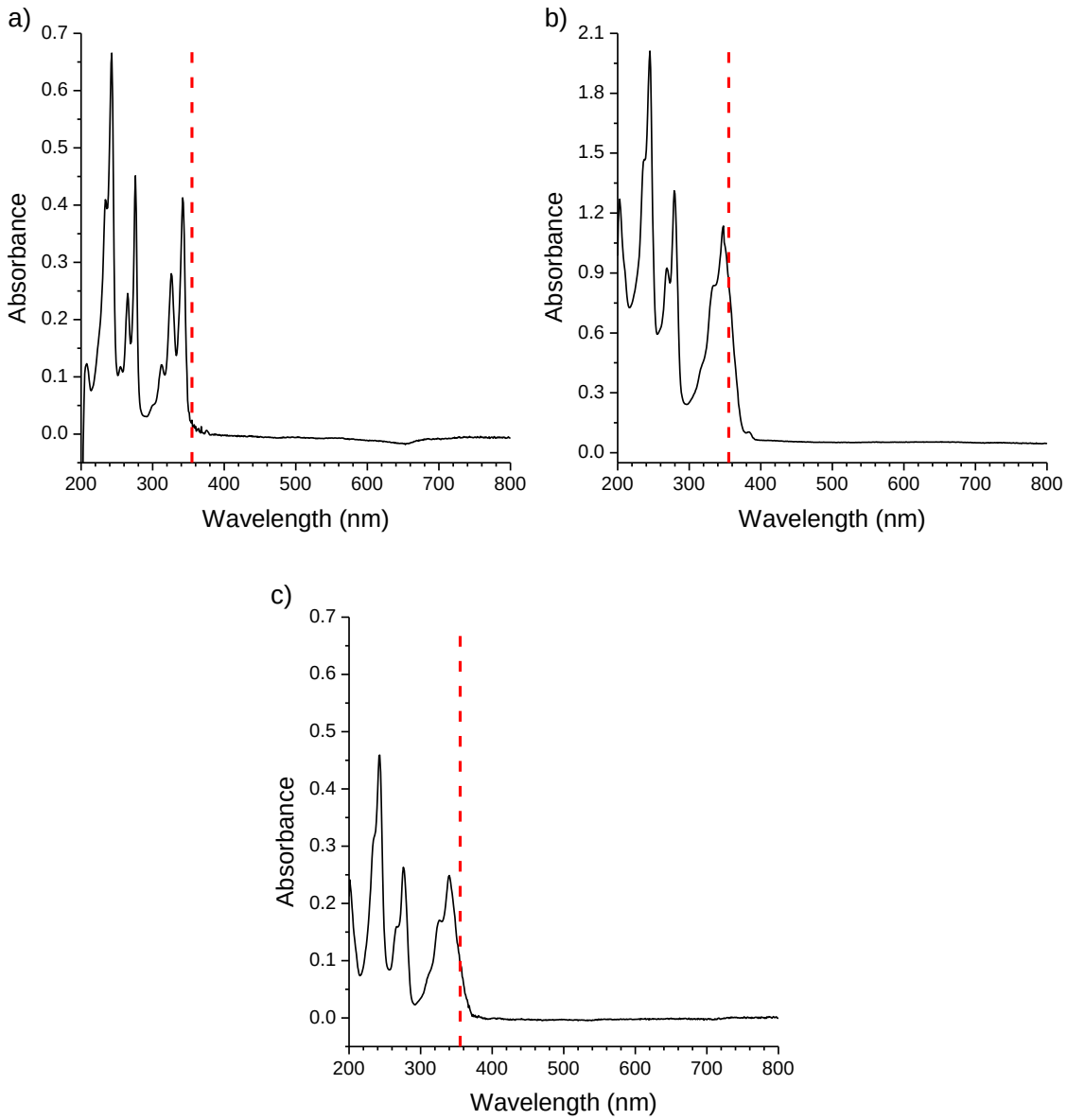


*Figure 2.3: UV-Vis absorption profile for 10 µM solution of a) pyrene **2.1**, b) 6-amidopyrene **2.2**, and c) 1-amidopyrene **2.3**. The wavelength of excitation used by the LDI-MS (355 nm) is highlighted by the dashed red line.*

The UV-Vis analysis of pyrene **2.1** (*Figure 2.3a*) and 6-amidopyrene **2.2** (*Figure 2.3b*) exhibited absorption profiles identical to the ones reported by Yoneda et al. for pyrene **1.12b** and 6-amidopyrene **1.14**.[100] Therefore, it was predicted that pyrene **2.1** and 6-amidopyrene **2.2** could also exhibit similar LALDI-MS sensitivity to those observed by Yoneda et al.[100]

Analysis of the novel 1-amidopyrene species **2.3** (*Figure 2.3b*) revealed an elevated absorbance at the wavelength of excitation used in the LDI-MS (355 nm). According to the work by Yoneda et al.,[100] these results suggested that **2.3** could also exhibit an increased LALDI-MS sensitivity. However, in LALDI-MS samples are analysed neat, not as solutions. Evidence in the literature has shown that solid state UV-Vis absorption spectra generally exhibit red shifted absorption

maxima compared to solution phase spectra.[202] A clear example of this for pyrenes has been provided by Sharma et al., which shows that in solution the UV absorption profile of pyrene is identical to that of **2.1**, while solid pyrene exhibits a broader and red shifted UV absorption profile with a much higher absorbance at 355 nm.[203] Sharma et al. proposed that the observed change in absorbance was due to the effect of increased π-stacking interactions that occurred during aggregation of pyrene molecules in the solid phase. Therefore, while the information gained on the solution phase UV absorbance of **2.3** was useful for comparison to other LDI enhancers in the literature, it is important to note that it should not be relied on for an accurate insight into the absorbance of an LDI enhancer during LALDI-MS.

2.2.3 Initial LALDI-MS analysis and determination of the lower detection limits

To determine the LALDI-MS lower detection limit of **2.1**, **2.2**, and **2.3**, each reagent was analyzed by LALDI-MS as a sequence of samples with the amount of analyte serially reduced from 100 pmol – 1 fmol. The lower detection limit was then defined as the lowest amount at which the molecular ion, protonated molecular ion, or metal adduct ion could be detected by LALDI-MS.

Samples were prepared by dissolving the LALDI tags in methanol, spotting a known amount of each compound onto a target plate, then allowing the solvent to evaporate. LALDI-MS analysis was performed on the neat samples without any additional chemical modifications, MALDI matrices, or SALDI surface materials with at least three individual repeats measured for each sample.

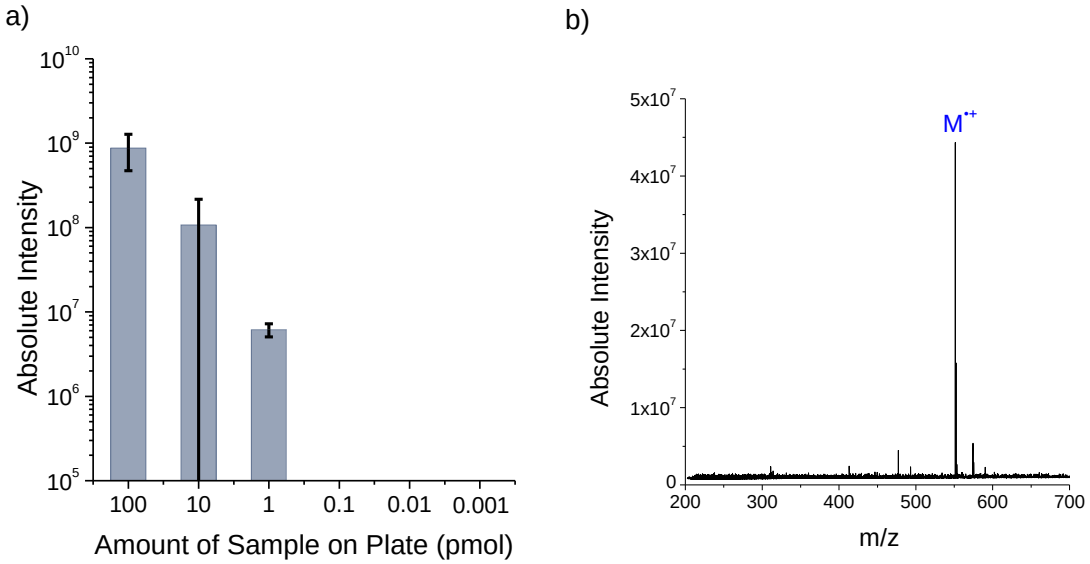


*Figure 2.4: LALDI-MS analysis of pyrene **2.1**. a) Analysis of the LALDI-MS lower detection limit for **2.1**. Bars represent the mean sum of all (+) ions related to of **2.1** ($M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$) that were detected by LALDI-MS at decreasing sample amounts. Error bars represent the standard deviation of the mean, where n=3 (individual samples on separate target spots). b) Representative LALDI-MS spectrum of **2.1** (10 pmol) with the corresponding signal for the $M^{\bullet+}$ (m/z 539.2515) highlighted.*

Pyrene **2.1** was observed to have a lower detection limit of 1 pmol and was detected primarily as the molecular ion ($M^{\bullet+}$), with minor detection observed for the sodium and potassium adduct ions. The experimentally obtained value for the lower detection limit was observed between the values reported by Yoneda et al. for **1.12a** (100 pmol) and **1.12b** (10 fmol), suggesting that the addition of the OEG chain could have affected the ionization efficiency of the LDI enhancers.

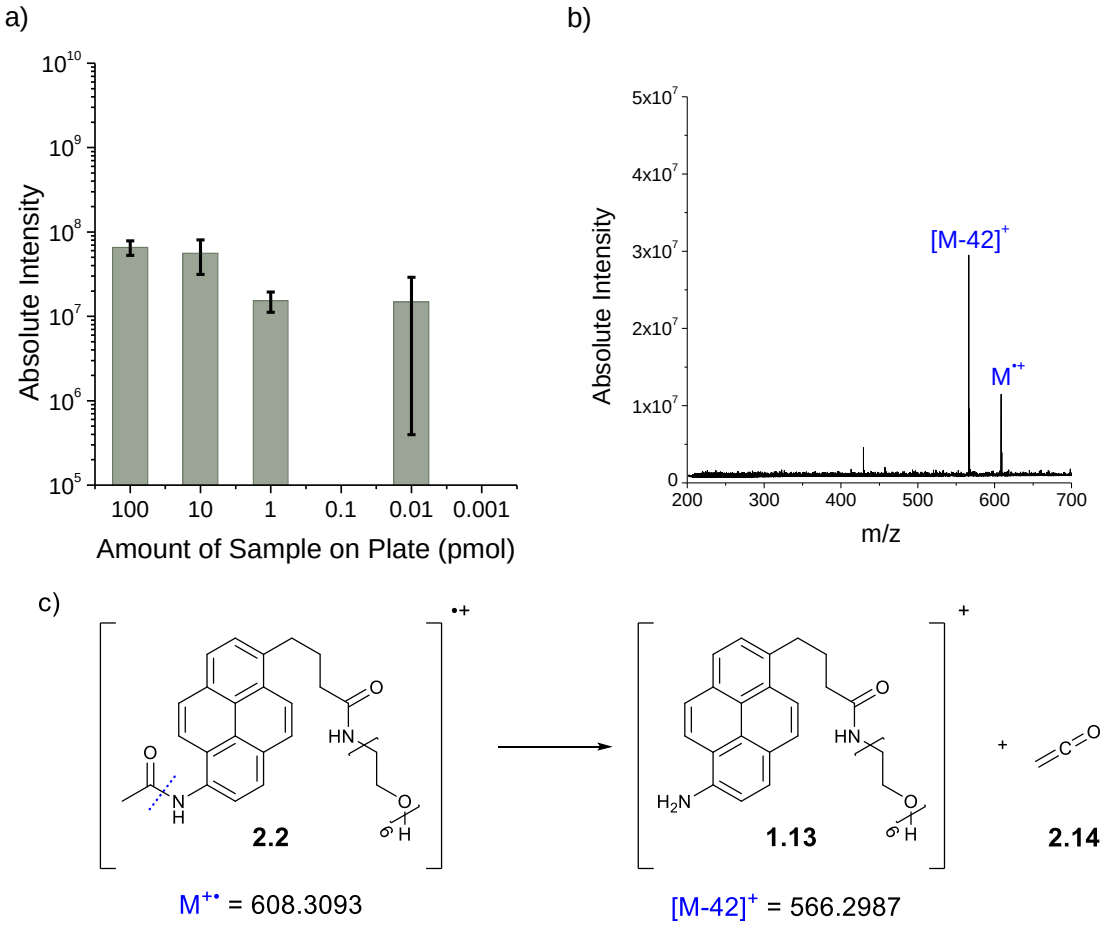


*Figure 2.5: LALDI-MS analysis of 6-amidopyrene **2.2**. a) Analysis of the LALDI-MS lower detection limit for **2.2**. Bars represent the mean sum of all (+) ions related to of **2.2** ($M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$, $[M-42]^+$) that were detected by LALDI-MS at decreasing sample amounts. Error bars represent the standard deviation of the mean, where n=3 (individual samples on separate target spots). b) Representative LALDI-MS spectrum of **2.2** (10 pmol) with the corresponding signals for the $M^{\bullet+}$ (m/z 608.3093) and $[M-42]^+$ (m/z 566.2987) ions highlighted. c) Fragmentation of 6-amidopyrene **2.2** proposed by Yoneda et al. to result in the formation of an $[M-42]^+$ ion **1.13** and the liberation of ketene **2.14**.*[100]

LALDI-MS analysis of amidopyrene **2.2** was found to be consistent with the data reported by Yoneda et al. for 6-amidopyrene **1.14**.[100] Amidopyrene **2.2** was observed to have a lower detection limit of 10 fmol (*Figure 2.5a*), and could be clearly detected as the $M^{\bullet+}$ (*Figure 2.5b*). An additional signal was also observed in the mass spectrum at $[M-42]^+$. Yoneda et al. concluded that the $[M-42]^+$ ion corresponded to the loss of ketene from 6-amidopyrene **2.2** as a result of fragmentation at the acetamide bond during LALDI-MS analysis (*Figure 2.5c*).[100] Accurate mass analysis of the sample confirmed that $[M-42]^+$ did indeed have an *m/z* value that corresponded exactly to the calculated value for $[M\text{-ketene}+H]^+$. While the formation of this fragment did not significantly complicate the spectra when analyzing a single reagent, analysis of a complex sample containing multiple species labelled with a 6-amidopyrene LDI enhancer could result in a more convoluted spectrum that would become difficult to resolve.

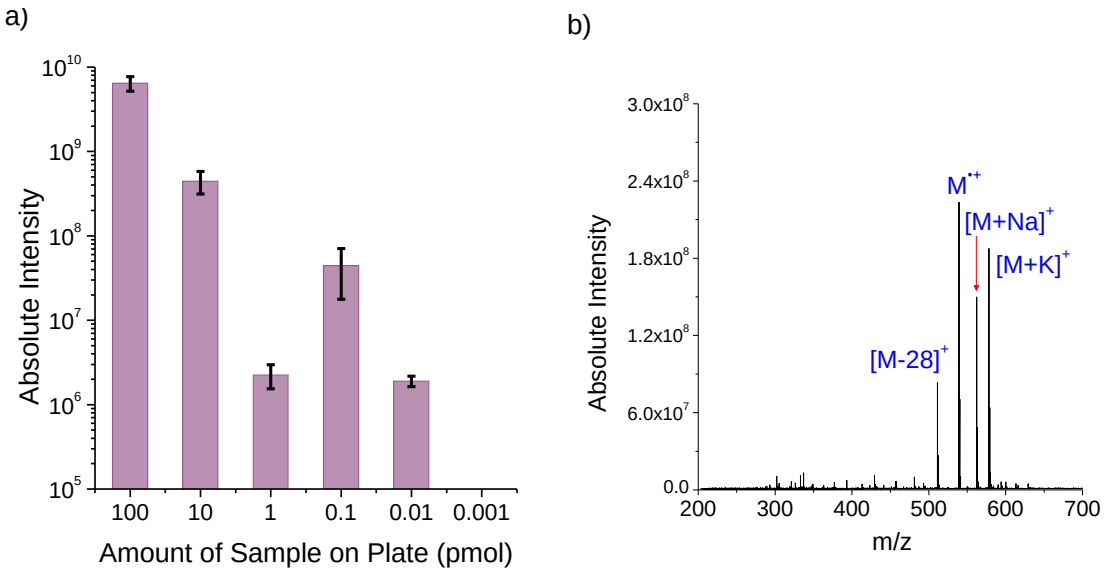


*Figure 2.6: LALDI-MS analysis of 1-amidopyrene **2.3**. a) Analysis of the LALDI-MS lower detection limit for **2.3**. Bars represent the mean sum of all (+) ions related to of **2.3** ($M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$, $[M-28]^+$) that were detected by LALDI-MS at decreasing sample amounts. Error bars represent the standard deviation of the mean, where n=3 (individual samples on separate target spots). b) Representative LALDI-MS spectrum of **2.3** (10 pmol) with the corresponding signals for the $M^{\bullet+}$ (m/z 539.2515) and $[M-28]^+$ (m/z 511.2566) ions highlighted.*

Novel 1-amidopyrene **2.3** was found to have comparable sensitivity to 6-amidopyrene **2.2**, with an observed lower detection limit of 10 fmol. These results supported the claims made by Yoneda et al. that amidopyrene LDI enhancers exhibited increased LALDI-MS sensitivity.[100] 1-amidopyrene **2.10** was detected as the $M^{\bullet+}$, and both $[M+Na]^+$and $[M+K]^+$ adduct ions in high abundance.

An additional peak correlating to $[M-28]^+$ was repeatedly observed in the LALDI-MS spectra for 1-amidopyrene **2.3**. Analytical high-pressure liquid chromatography (HPLC) was able to confirm that the analytical sample was pure, suggesting that the $[M-28]^+$ ion was a product of degradation during MS analysis and not an impurity (see *Figure 2.24* in Section 2.8.2). Through the use of high-resolution mass spectrometry, it was found that this peak had an *m/z* 28.9949 less than $M^{\bullet+}$, which correlated exactly to a loss of CO from **2.3**. However, it was not clear how this species could have formed through fragmentation, as **2.3** has no terminal CO that could be readily liberated by fragmentation. Furthermore, this signal at $[M-28]^+$ was only detected following ionization by LALDI and not with ESI. Therefore, it was hypothesised that $[M-28]^+$ was likely to be a product of LALDI excitation process. A further investigation into the source of this degradation is discussed in Section 2.4.

The results of these experiment showed that all three reagents pyrene **2.1** (*Figure 2.4*), 6-amidopyrene **2.2** (*Figure 2.5*), and 1-amidopyrene **2.3** (*Figure 2.6*) could achieve matrix-free ionization and subsequent MS detection following excitation of the sample via a laser pulse, complementing literature reports that pyrenes can be readily ionized by LDI-MS without an external matrix.[94,100,102,103] Detection of $M^{\bullet+}$ in the LALDI-MS spectra was expected, as the detection of the molecular ion of pyrene species has been reported in both MALDI and ESI-MS analysis.[204,205] Large polyaromatic ring structures, such as pyrenes, are known to stabilize the radical cations through delocalization of the charge across the extensive π-system.[206] The salt adduct ions observed in the spectra were attributed to the leaching of Na^+ and K^+ salts from the glassware used to store solvents and prepare samples. It was proposed that the OEG component of these compounds may have a strong affinity for sodium and potassium ions, like the strong affinity seen between 18-crown-6 and K^+ or 15-crown-5 and Na^+,[207] facilitating their MS detection. Salt adduct ions were included in the assessment of lower detection limits, as they were considered useful for detection and identification of the molecular ion of the LALDI reagents.

Several anomalies were observed in the experimental data obtained during these initial LALDI-MS investigations. Firstly, during analysis of a 0.1 pmol sample of **2.2** (*Figure 2.5*), no ions related to the parent species were detected, despite this being above the recorded lower detection limit. Furthermore, the abundance of molecular ions detected during the analysis of a 1 pmol sample of **2.3** (*Figure 2.6*) was significantly lower than for a 0.1 pmol sample of the same compound. To address these anomalies, it was decided that further independent repeats of these experiments were required to ensure that the results are accurate and robust. However, achieving good reproducibility of data between discrete samples proved difficult. Individual repeats of the same sample were found to give completely different measurements regarding the absolute ion count and signal-to-noise, resulting in the large error bars reported for the data (*Figure 2.5, Figure 2.6*). Multiple additional individual repeats were found to be necessary for each experiment as LALDI-MS analysis of samples above the limit of detection frequently resulted in blank spectra that were omitted from the quantification of the data.

Even technical repeats of the same sample target were observed to yield vastly different results. As good reproducibility was considered important for ensuring the validity of future analysis, efforts were directed towards improving the reproducibility of data obtained and is discussed further in Section 2.3.

Of the three reagents analyzed during this investigation, 1-amidopyrene **2.3** was considered to have the most favorable LDI-enhancing label for incorporation future LALDI-MS investigations. While both amidopyrenes were observed to have equivalent sensitivities, 1-amidopyrene **2.3** was favored over 6-amidopyrene **2.2**, as **2.3** had slightly better stability, and a much simpler synthesis.

2.3 Improving the reproducibility of LALDI-MS data

Poor reproducibility is common in MALDI-MS studies and can often be attributed to the poor homogeneity of the analyte across the sample during co-crystallization with the matrix component.[208] Therefore, it was proposed that the anomalies and inconstancies in the results reported in Section 2.2.3 could be attributed to the uneven distribution of the LALDI reagents on the target plate during evaporation of the sample solution.

2.3.1 Scanning electron microscope imaging of an LDI target plate

To better understand the issues behind the inconstancy and poor reproducibility of the data, an investigation was undertaken to examine how the LALDI reagents were deposited onto the surface of the ground-steel MALDI target plate used during the initial investigations. To achieve this, samples of 1-amidopyrene **2.3** were spotted onto a section of a ground steel LDI-MS target plate used for LALDI-MS experiments, then analyzed using a scanning electron microscope (SEM). SEM imaging was conducted with assistance from Dr Alexander Kulak. To understand why these inconstancies appeared across both large and small sample sizes, a 100 pmol and a 10 fmol sample of **2.3** were analyzed in this investigation (*Figure 2.7*).

Upon magnification of the target area, both samples exhibited concentrated areas of **2.3** at their outer edges. The heterogeneity of the film can be attributed to a phenomenon known colloquially as the “coffee-ring effect”.[209] As a droplet dries, the center of droplet evaporates slower than the edges. Capillary flow then causes the remaining liquid to move towards the edges. Therefore, when the liquid is a solution, this capillary flow results in a larger concentration of the solute at the edges of the droplet following evaporation.

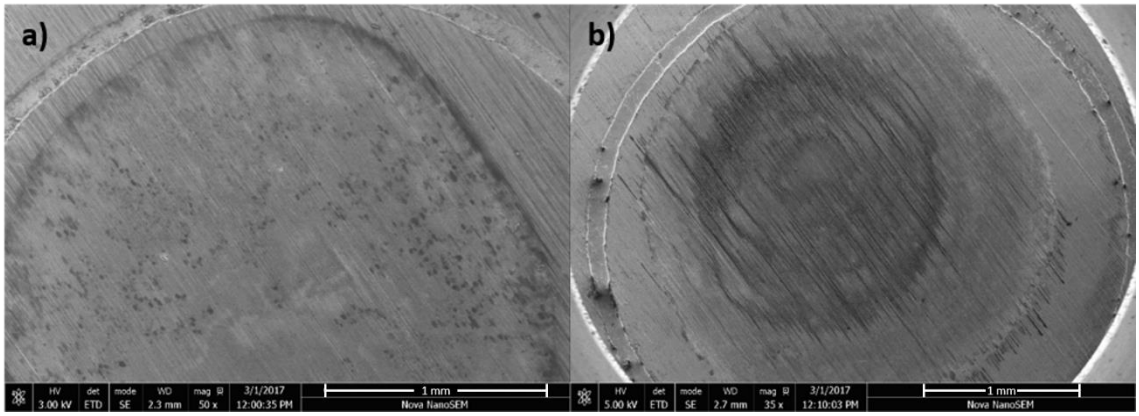


*Figure 2.7: SEM imaging of a Bruker MTP 384 ground steel target plate spotted with a) 10 fmol (50 × magnification) and b) 100 pmol (35 x magnification) sample of **2.3**. The dark grey areas indicate where **2.3** has been deposited. Scale bars represent 1 mm. SEM images were recorded with assistance from Dr Alexander Kulak.*

Magnification of the 10 fmol sample (*Figure 2.7a*), showed that large areas of the target remained uncoated by the sample, especially in the center. These observations supported the assumption that poor coverage of the target area could be responsible for the anomalies in analyte detection at very low quantities. To circumvent this problem and improve the reliability of data when analyzing minute quantities of analyte, it was proposed that targeting the laser towards the outer edges of the sample could be used to increase the chance of finding an area occupied by the analyte.

Elemental analysis of the 10 fmol sample of **2.3**, suggested that the sample was concentrating in the small grooves in the ground steel LDI target plate (*Figure 2.8*). Elemental analysis showed that there was an increased concentration of carbon (purple trace) and oxygen (green trace) in the grooves of the plate, indicating that the grooves were filled with an organic material, such as **2.3**, further highlighting the heterogeneity of the sample on the target plate.

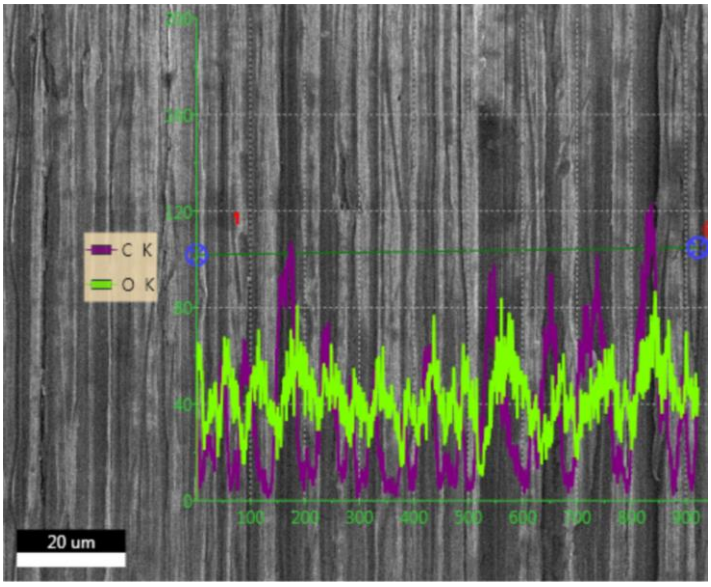


*Figure 2.8: Elemental analysis of the 10 fmol sample of **2.3**. Dark grey vertical lines in the image show the grooves in the ground steel target plate. The trace overlaid on the magnified image of the plates surface shows the levels of carbon (purple) and oxygen (green) detected between the start and end points of the elemental analysis (blue crosshairs). The trace shows that there are spikes in the detection of these two elements where there are grooves in the steel plate, indicating the presence of an organic material. SEM images were recorded with assistance from Dr Alexander Kulak.*

Magnification of the 100 pmol sample of **2.3** (*Figure 2.7b*) revealed that the overall coverage of the target plate was much better than for the 10 fmol sample. However, during evaporation of the solvent, several rings of sample were shown to have formed across the target area resulting in a pattern of high and low-density areas of the analyte. As with the 10 fmol sample, these images suggested that the problems encountered with the precision and poor reproducibility of data obtained from the larger samples was also a consequence of uneven coverage of sample on the target plate.

Knowing that samples dried in these 'coffee ring' patterns when deposited on the ground steel targets, the lower detection limit investigations were repeated with attempts made to improve the reproducibility by aiming the laser at the edges of each sample (*Figure 2.9*). While this did improve the results slightly, the reproducibility was still poor and analysis of samples above the limit of detection still resulted in occasional blank spectra that were omitted from the quantification of the data, resulting in the continued need for additional individual repeats to obtain the data.

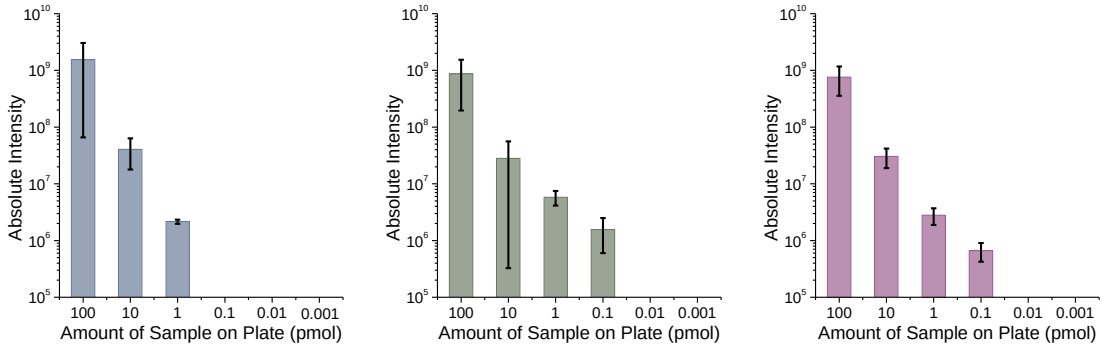


*Figure 2.9: Comparison of the LALDI-MS analysis of the lower detection limit for **2.1** (blue), **2.2** (green), and **2.3** (purple) using a ground-steel MALDI target plate where data were collected by aiming the laser specifically at the edges of the sample. Bars represent the mean sum of all (+) ions related to **2.1**, **2.2**, and **2.3** that were detected by LALDI-MS at decreasing sample amounts of each reagent. Error bars represent the standard deviation of the mean, where n=3 (individual samples on separate target spots).*

2.3.2 Improving reproducibility using a micro-focusing target plate

During previous LALDI-MS experiments, analysis was carried out using a standard Bruker MTP 384 ground steel target plate, which has targets with a diameter of ~3.5 mm. One of the most common problems faced during previous LALDI-MS analysis, especially at very low sample quantities, was trying to find an area of the target occupied by the analyte for MS analysis. This problem led to blank measurements, inconsistent results, and was suggested as the main cause of the poor reproducibility observed between both independent repeats and technical replicates. It was proposed that this problem could be addressed through the use of a micro-focusing target plate that contains a super-hydrophilic core surrounded by a hydrophobic ring designed to concentrate the sample into a spot 800 µm in diameter.[210] Therefore, it was hypothesized that the use of a micro-focusing target plate could help improve analyte detection and reproducibility of data by concentrating the sample into a smaller area.

The lower detection limit investigations were repeated using a Bruker AnchorChip micro-focusing target plate (*Figure 2.10*). Using the micro-focusing target plate it was possible to significantly improve the reproducibility of the LALDI-MS analysis. Detection of analytes could be achieved immediately above the reported detection limits without having to search on the target and without encountering any blank runs. This meant that both analysis and data processing could be carried out much more effectively, and that processing the data no longer required omission of blank spectra from the results. Furthermore, using the micro-focusing target plate, it was possible to achieve detection of pyrene **2.1** below the previous experimentally observed detection limit.

Following this investigation, it was concluded that the AnchorChip target plate would be used for all further LALDI-MS investigations.

LALDI-MS spectra obtained during the optimized lower detection limit analysis are reported in *Appendix I.*

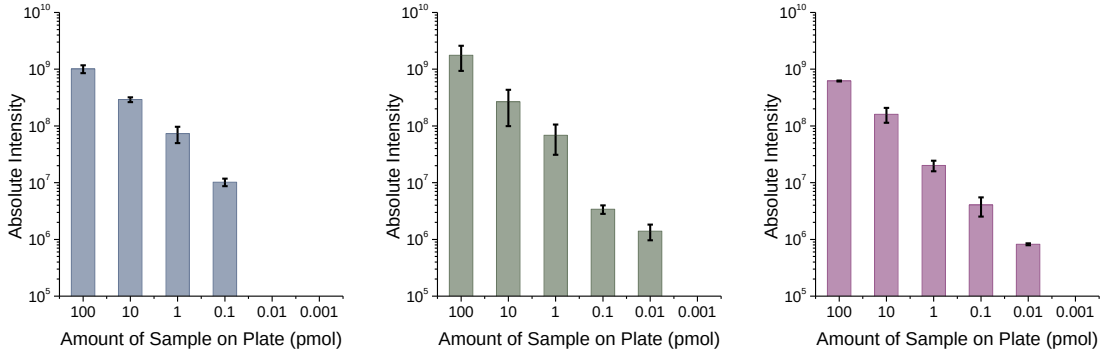


*Figure 2.10: Comparison of the LALDI-MS analysis of the lower detection limit for **2.1** (blue), **2.2** (green), and **2.3** (purple) using a Bruker AnchorChip™ micro-focusing MALDI target plate. Bars represent the mean sum of all (+) ions related to **2.1**, **2.2**, and **2.3** that were detected by LALDI-MS at decreasing sample amounts of each reagent. Error bars represent the standard deviation of the mean, where n=3 (individual samples on separate target spots).*

2.4 Investigating loss of CO from 1-amidopyrene LDI enhancers

An investigation was undertaken to try and identify the source of the $[M\text{-}28]^+$ ion generated during LALDI-MS analysis of 1-amidopyrene **2.3**. Previously, LALDI-MS analysis of **2.3** had indicated that $[M\text{-}28]^+$ was generated by a loss of CO, but it was unclear how this could have occurred via fragmentation (Section 2.2.3). To determine whether $[M\text{-}28]^+$ was generated by fragmentation of **2.3**, LALDI-MS^2 analysis was performed on the radical cation of 1-aminopyrene **2.3** ($M^{\bullet+}$(**2.3**)) (*Figure 2.11a*). While the LALDI-MS^2 analysis of $M^{\bullet+}$(**2.3**) resulted in a number of fragments (*Figure 2.11b*), $[M\text{-}28]^+$ was not observed in the mass spectrum. This suggested that $[M\text{-}28]^+$ was not a product of fragmentation, but formed through some other mechanism.

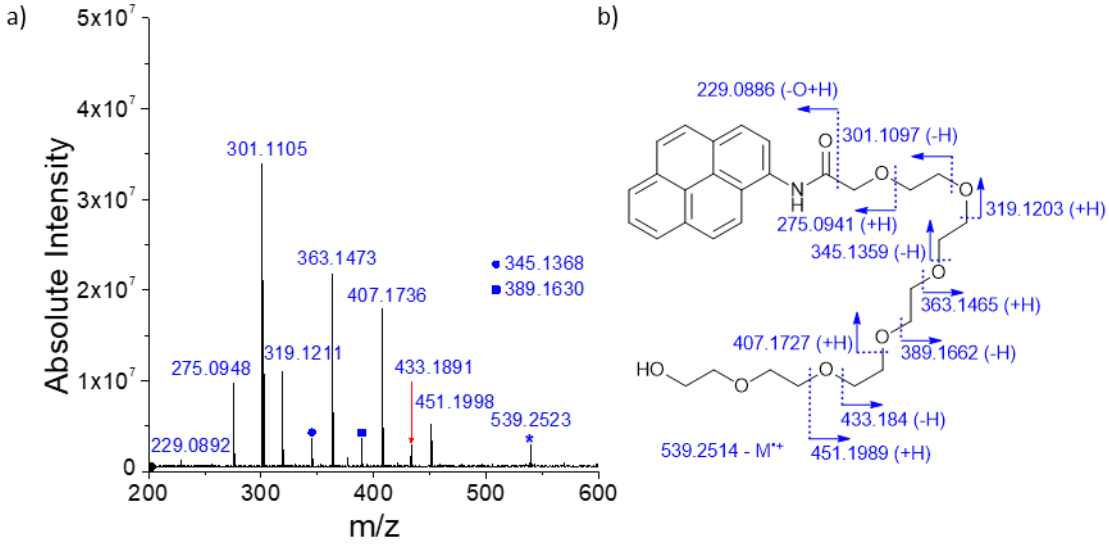
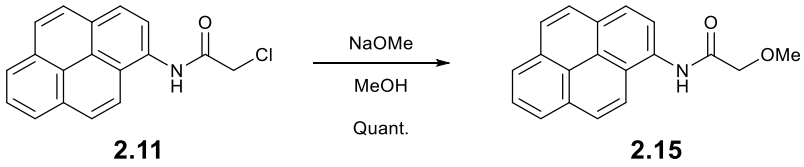


*Figure 2.11: LALDI-MS^2 analysis of M$^{\bullet+}$(**2.3**) (m/z 539.2523) (100 pmol). a) LALDI-MS^2 spectrum for M$^{\bullet+}$(**2.3**) with identified fragments annotated with their observed m/z. The parent ion (m/z 539.2523) has been annotated with an asterisk (*). b) Proposed assignments for signals observed in the LALDI-MS^2 spectrum. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

As [M-28]$^+$ had not been observed during the LALDI-MS analysis of **2.1** (Figure 2.4b) or **2.2** (*Figure 2.5*b), it was hypothesized that CO was being liberated from the 1-amidopyrene LDI enhancer rather than the OEG chain. To test this theory, a simple 1-amidopyrene analogue **2.15** was synthesised (*Scheme 2.6*) and analysed by both LALDI-MS (*Figure 2.12*) and MS^2 (*Figure 2.13, Figure 2.14*).



2.11 **2.15**

*Scheme 2.6: Synthesis of 2-methoxy-N-methyl-N-(pyren-1-yl)acetamide **2.15**.*

1-amidopyrene **2.15** was prepared in quantitative yield from chloroacetamide **2.11** via alkylchloride substitution using NaOMe/MeOH.

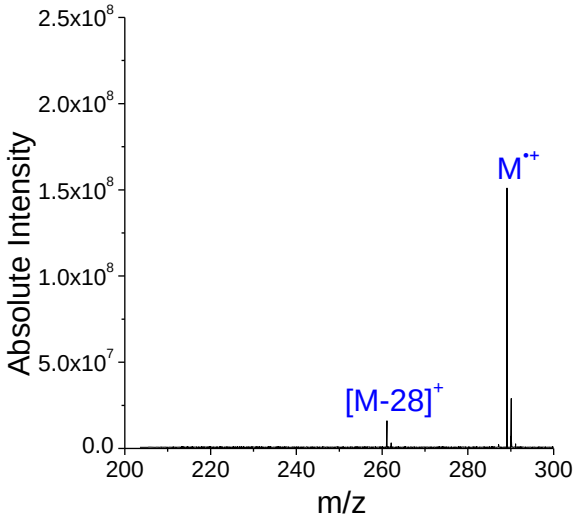


*Figure 2.12: LALDI-MS spectrum of 1-amidopyrene **2.15** (100 pmol) with the corresponding signals for the $M^{\bullet+}$ (m/z 289.1099) and $[M-28]^+$ (m/z 261.1150) ions highlighted.*

As observed for 1-amidopyrene **2.3**, $[M-28]^+$ was observed in the LALDI-MS spectrum of 1-amidopyrene **2.15** (*Figure 2.12*), but not in the LALDI-MS^2 spectrum of $M^{\bullet+}$(**2.15**) (*Figure 2.13*). These findings supported the proposed idea that $[M-28]^+$ was generated from the 1-amidopyrene LDI enhancer and not from the OEG chain.

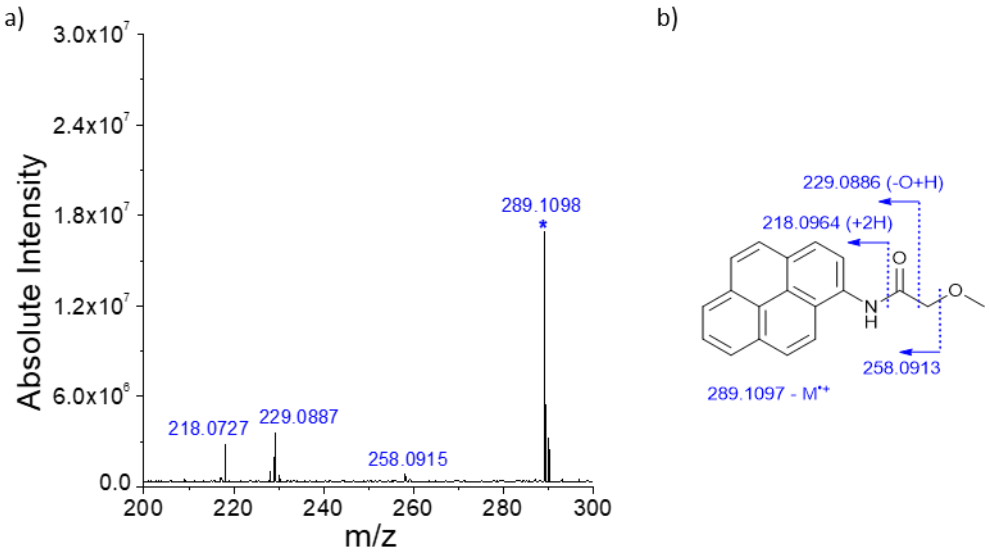


*Figure 2.13: LALDI-MS^2 analysis of $M^{\bullet+}$(**2.15**) (m/z 289.1097) (100 pmol). a) LALDI-MS^2 spectrum for $M^{\bullet+}$(**2.15**) with identified fragments annotated with their observed m/z. The parent ion (m/z 539.2523) has been annotated with an asterisk (*). b) Proposed assignments for signals observed in the LALDI-MS^2 spectrum of $M^{\bullet+}$(**2.15**). Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

LALDI-MS^2 analysis then was carried out on the [M-28]$^+$(**2.15**) to determine which of the two oxygens were being lost in the original degradation by looking at the fragmentation pattern of the [M-28]$^+$ ion (*Figure 2.14*). In the LALDI-MS^2 analysis of [M-28]$^+$(**2.15**) (*Figure 2.14*), a fragment was observed with an *m/z* 228.0810, which is consistent with the loss of a methoxy group. The observation that the terminal methoxy group can be lost by fragmentation of [M-28]$^+$, suggests that the original degradation was likely to be from the loss the carbonyl CO as there are no other available oxygens.

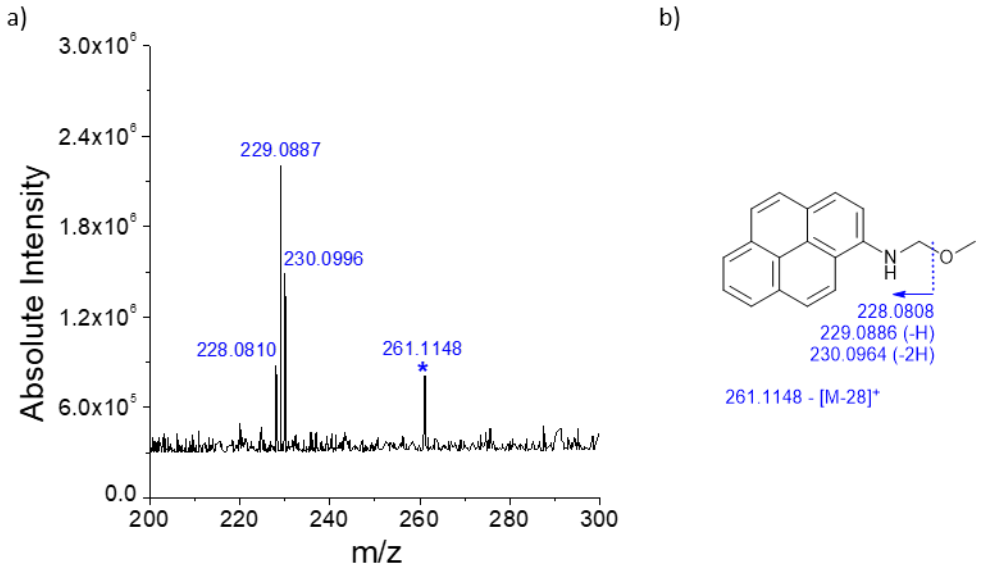


*Figure 2.14: LALDI-MS^2 analysis of [M-28]$^+$(**2.15**) (m/z 289.1097) (100 pmol). a) LALDI-MS^2 spectrum for [M-28]$^+$(**2.15**) with identified fragments annotated with their observed m/z. The parent ion (m/z 539.2523) has been annotated with an asterisk (*). b) Proposed structure for [M-28]$^+$ with proposed assignments for signals observed in the LALDI-MS^2 spectrum of [M-28]$^+$(**2.15**). Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

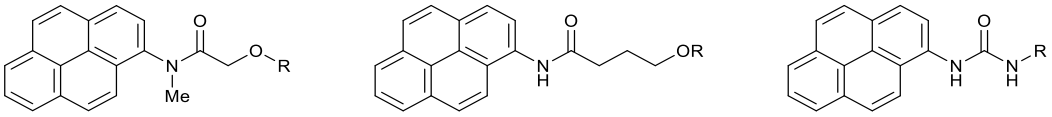
2.5 Improving stability

While previous LALDI-MS analysis demonstrated that amidopyrene LDI enhancers were capable of femtomole-quantity lower detection limits (*Figure 2.10*), both 6-amidopyrene **2.2** (*Figure 2.5b*) and 1-amidopyrene **2.3** (*Figure 2.6b*) also experienced degradation during LALDI-MS analysis. While the formation of the additional signals did not significantly complicate the spectra when analyzing a single reagent, analysis of a sample containing multiple species could result in a complicated spectrum that may be difficult to resolve.

Therefore, efforts were directed towards developing a novel LDI enhancer with good LALDI-MS sensitivity and increased stability. It was proposed that the best approach to this would be to

investigate various analogues of the 1-amidopyrene LDI enhancer and improve the stability by preventing the loss of CO.

Three amidopyrene derivatives were put forward as means of stabilizing the carbonyl group during LALDI excitation (*Figure 2.15*): *N*-methylamidopyrene was chosen to assess whether substituting the amide proton for a methyl group would improve resistant to MS-degradation; *N*-pyrenylbutyramide was suggested to determine whether extending the alkyl spacer and moving the OEG chain away the carbonyl group could prevent the observed loss of CO; and pyrene urea was selected as it was hypothesized that the additional resonance from a second nitrogen to the carbonyl group could increase the MS stability.[211]

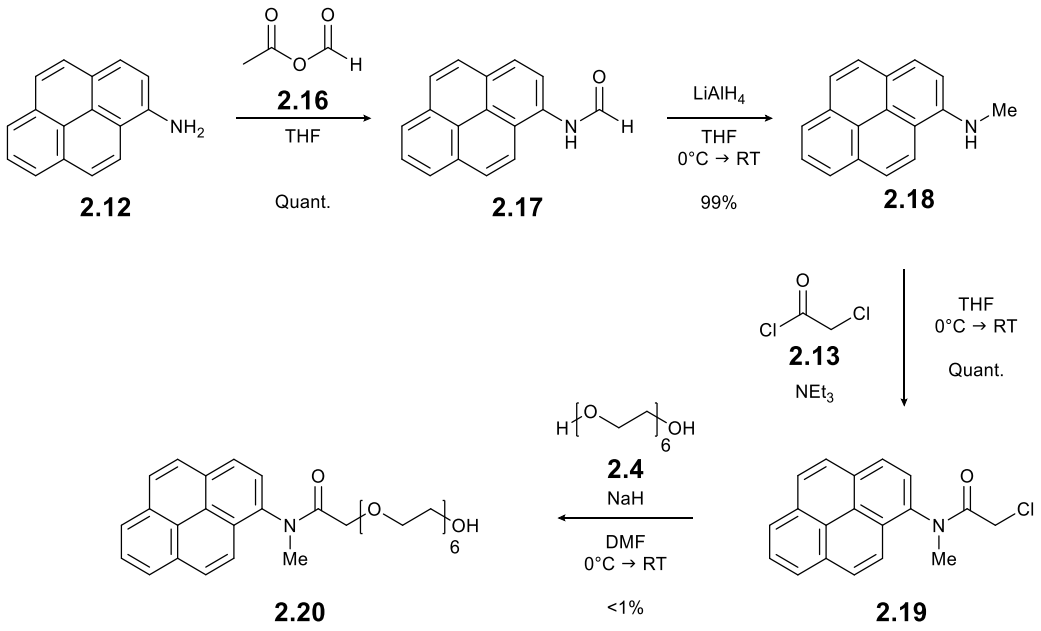


N-methylamidopyrene **N-pyrenylbutyramide** **Pyrene urea**

Figure 2.15: Generic structure for the three amidopyrene derivatives proposed for the investigation into novel LDI enhancers with improved MS-stability.

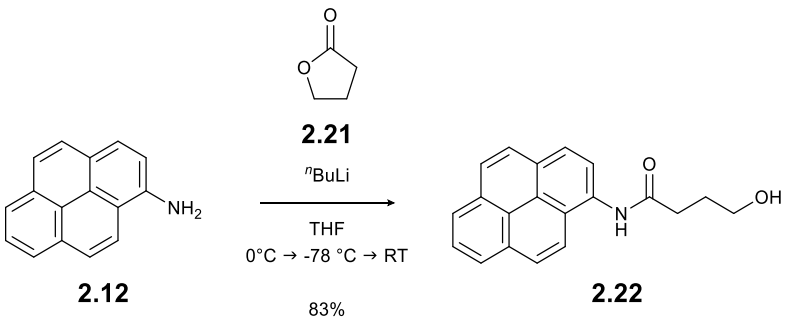
2.5.1 Synthesis of novel LDI enhancers

N-methylamidopyrene **2.20** was synthesized in 4 steps from amino pyrene (*Scheme 2.7*). Firstly, *N*-formylation of aminopyrene **2.12** was achieved using acetic formic anhydride **2.16** to afford formamide **2.17** in quantitative yield. Formamide **2.17** was then reduced using $LiAlH_4$ to afford *N*-methylamino pyrene **2.18** in 99% yield.[212] *N*-methylaminopyrene **2.18** was conjugated to chloroacetyl chloride **2.13** via nucleophilic acyl substitution to afford *N*-methyl chloroacetamide **2.19** in quantitative yield.[200] Finally, *N*-methylchloroacetamide **2.19** was then conjugated to hexaethylene glycol **2.4** under basic conditions via a Williamson ether synthesis to yield alcohol **2.20** in <1% yield.[201] LCMS analysis of the reaction mixture indicated that a large amount of *N*-methylaminopyrene **2.18** had formed during the reaction. Hence, the low recovered yield of **2.20** was attributed to hydrolysis of the amide linkage during the reaction.



*Scheme 2.7: Synthetic route to 20-hydroxy-*N*-methyl-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide* ***2.20***.[201]

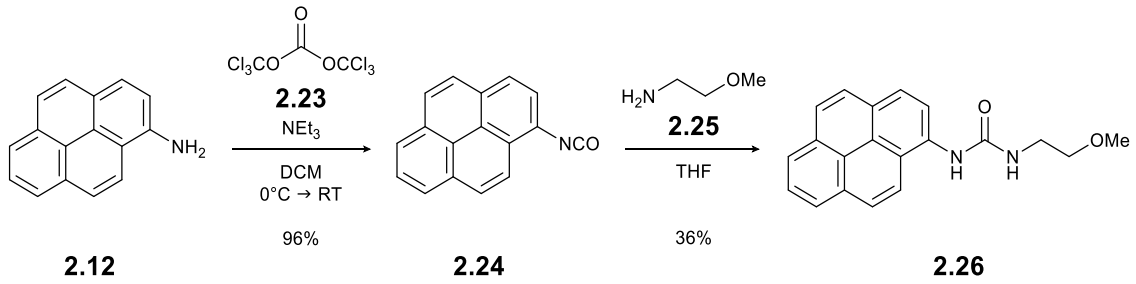
N-methylamidopyrene **2.22** was synthesized in a single step from aminopyrene **2.12** via nucleophilic acyl substitution of lactone **2.21** under strong basic conditions to afford *N*-pyrenylbutyramide **2.22** in 83% yield (*Scheme 2.8*).[213]



Scheme 2.8: Synthetic route to 4-hydroxy-N-(pyren-1-yl)butanamide ***2.22***.[213]

Pyrene urea **2.26** was synthesized in 2 steps from amino pyrene **2.12** (*Scheme 2.9*).[214,215] Following a procedure by Biedermann et al., aminopyrene **2.12** was treated with triphosgene **2.23** to give pyrene isocyanate **2.24** in 96% yield.[214] 2-methoxyethylamine **2.25** was then conjugated to isocyanate **2.24** via a urea formation reaction adapted from the literature to give pyrene urea **2.26** in 36% yield.[215] Conversion of isocyanate **2.24** to urea **2.26** was deemed almost quantitative by TLC and LCMS, but resulted in a reduced isolated yield due to problems encountered during chromatographic separation. However, as only a small purified sample of

pyrene urea **2.26** was required for LALDI-MS analysis, further optimization of the synthetic procedure was deemed unnecessary.



*Scheme 2.9: Synthetic route to 3-(2-methoxyethyl)-1-(pyren-1-yl)urea **2.26**.*[214,215]

2.5.2 Analysis of novel LDI enhancers

Following synthesis, purified samples of each of the novel LDI enhancers were analyzed by LALDI-MS (*Figure 2.16*) and UV-Vis spectroscopy (*Figure 2.18*). The stability and sensitivity of each reagent in the mass spectrometer was then assessed using this LALDI-MS analysis. A summary of the LALDI-MS analysis from this investigation can be found in *Table 2.1.*

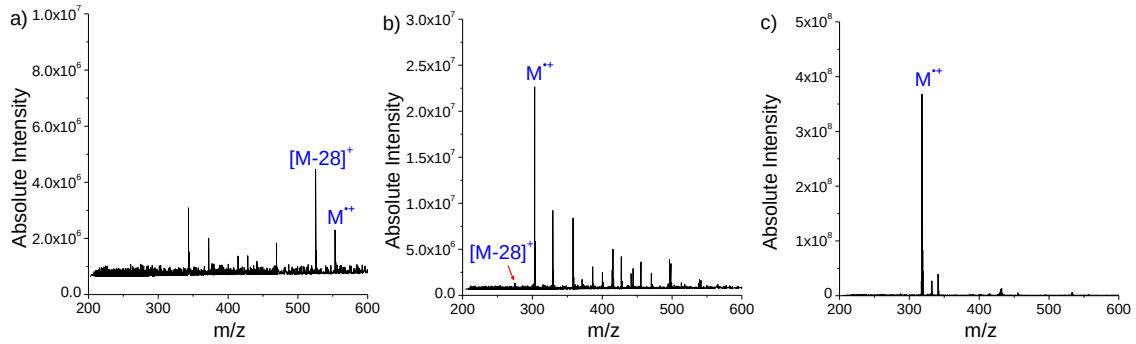


*Figure 2.16: LALDI-MS spectra for 10 pmol samples of a) **2.20** b) **2.22** c) **2.26** with signals correlating to the molecular radical cation ($M^{\bullet+}$) and loss of CO ($[M-28]^+$) highlighted on the spectra.*

LALDI-MS analysis showed that all three of the novel LDI enhancers were capable of achieving matrix-free ionization (*Figure 2.16*). *N*-methylamidopyrene **2.20**, *N*-pyrenylbutyramide **2.22**, and pyrene urea **2.26** could all be clearly detected as the $M^{\bullet+}$, but experienced varied levels of MS stability.

N-methylamidopyrene **2.20** was found to have reduced stability compared to the 1-amidopyrene LDI enhancer, with an increased relative abundance of $[M-28]^+$ and additional fragmentation observed in the spectrum (*Figure 2.16a)*. *N*-pyrenylbutyramide **2.22** showed a reduction in the relative abundance of $[M-28]^+$, but did produce a number of unidentifiable

higher *m/z* peaks (*Figure 2.16b*). Pyrene urea **2.26** was found to have the best MS stability of the three LDI enhancers tested, with no observed fragmentation or degradation. The results of the investigation suggested that pyrene urea LDI enhancers could be a viable option for developing LALDI reagents with high sensitivity and improved stability. Therefore, it was decided that pyrene urea LDI enhancers would be taken forward and their LALDI-MS performance investigated further.

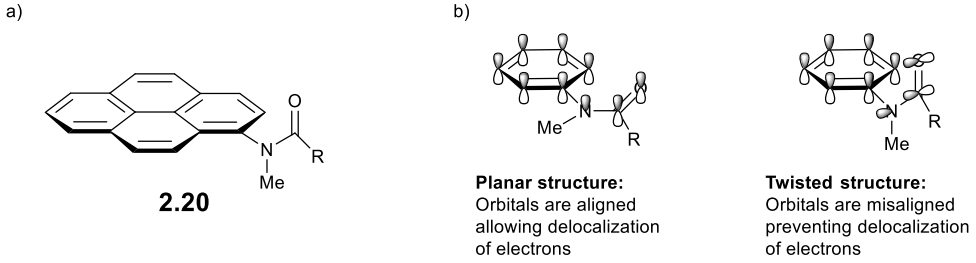


*Figure 2.17: a) Proposed conformational structure for N-methylamidopyrene **2.20**. The pyrene ring is twisted out of the plane of the amide group to prevent steric clashing between the N-methyl group/carbonyl oxygen and the pyrenyl protons. b) A schematic representation of the molecular orbitals in a planar and twisted conformation of a N-methyl-N-phenyl amide. The diagram shows that in the twisted conformation the molecular orbitals of the amide are orthogonal to the aromatic π-system, therefore delocalization cannot occur.*

It was proposed that the reduced LALDI-MS stability of *N*-methylamidopyrene **2.20** could be attributed to the conformational structure of the *N*-methylamidopyrene LDI enhancer. Crystal structures of similar *N,N*-disubstituted amidopyrenes reported in the literature suggested that steric clashing between the *N*-methyl group/carbonyl oxygen and the pyrene ring of **2.20** could result in twisting of the nitrogen-pyrene bond (*Figure 2.17a*). This torsion would prevent conjugation between the nitrogen lone pair and the pyrene π-system, affecting both the sensitivity and stability of **2.20** and may also explain why a significant percentage of **2.20** hydrolyzed during synthesis (*Scheme 2.7*).

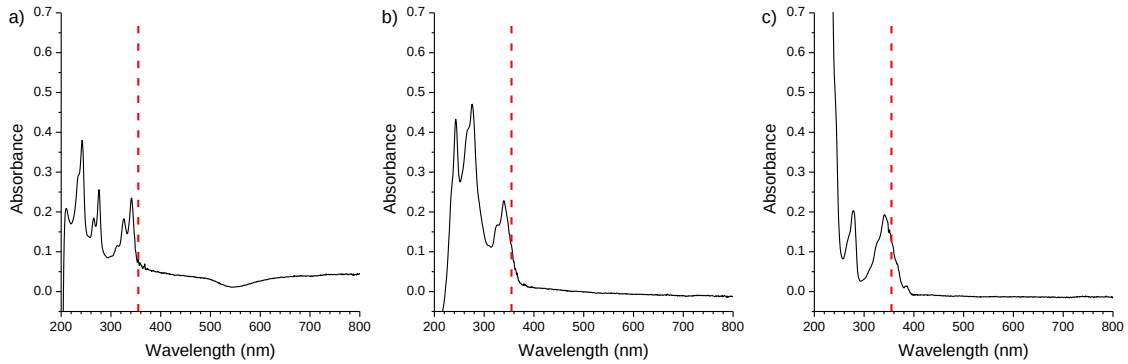


*Figure 2.18: UV-Vis absorption profile for 10 µM solutions of a) **2.20** b) **2.22** c) **2.26**. The wavelength of excitation used by the LDI-MS (355 nm) is highlighted by the dashed red line.*

This theory was further supported by the observed difference in the UV absorption profile of **2.20** (*Figure 2.18a*) compared to other amidopyrenes. Examples obtained both experimentally (*Figure 2.3b-c, Figure 2.18b-c*) and from the literature have generally shown that the UV absorption maxima of pyrene rings substituted with an amino or amido group experience a bathochromic shift.[100] The bathochromic shift is caused by delocalization of the nitrogen lone pair into the aromatic ring, which lowers the overall energy of the molecule and increases stability by increasing the available resonance forms. The greater the resonance, the smaller the gap between the highest occupied molecular orbital and the lowest unoccupied molecular orbital. Therefore, less energy is required to promote an electron between electronic levels, and the UV absorption maximum shifts to a longer wavelength. Therefore, as the absorption profile of *N*-methylamidopyrene **2.20** (*Figure 2.18a*) was found to be similar to that of pyrene **2.1** (Figure 2.3a), it was suggested that *N*-methylamidopyrene **2.20** was not experiencing any additional resonance as the nitrogen lone pair of was unable to delocalize into the pyrene ring due to the twisted structure (*Figure 2.17b*).

*Table 2.1: Summary of the LALDI-MS analysis of novel LDI enhancers compared to 1-amidopyrene **2.3**. MS stability was assessed on the relative abundance of the molecular ion ($M^{\bullet+}$) and metal adduct ions ($[M+Na]^+$, $[M+K]^+$) compared to the fragments, degree of fragmentation, and ease of fragment assignment then scored using an arbitrary scale from **++** to **– –**.*

LDI enhancer	MS stability	Limit of detection (pmol)
1 amidopyrene **2.3**	Degradation peak at $[M-28]^+$	0.01
N-methylamidopyrene **2.20**	Fragmentation and $[M-28]^+$ –	1 –
N-pyrenylbutyramide **2.22**	$[M-28]^+$ reduced +	1 –
Pyrene urea **2.26**	No $[M-28]^+$ observed ++	0.1 +

2.6 Exploring the potential of pyrene urea and pyrene thiourea LDI enhancers

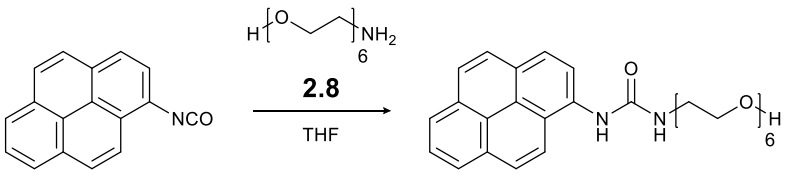
Further investigation into the LALDI-MS performance of the novel pyrene urea LDI enhancer were conducted using pyrene urea **2.27** (*Figure 2.19*). Pyrene urea **2.27** was designed to be structurally analogous to LALDI reagents **2.1**, **2.2** and **2.3** to limit the number of variables that could affect the analysis and facilitate a clearer comparison of the pyrene urea LDI enhancer to those previously tested in Section 2.2.3.

After observing improved stability for pyrene urea **2.27**, pyrene thioureas were also considered as potential LDI enhancer for LALDI-MS. Isothiocyanates are known to react with various nucleophiles present in a biological environment (e.g. amines, thiols, tyrosine phenolates, etc), but only form stable conjugate products with primary amines.[216] This makes isothiocyanates probes, such as fluorescein isothiocyanate, excellent tools for selectively modifying the *ε*- and *N*-terminal amines of peptides and proteins.[217] Therefore, it was considered that if pyrene thiourea **2.28** was found to have good LALDI sensitivity and good MS stability, then pyrene isothiocyanate **2.29** could potentially be developed as a novel probe for analyzing peptides and proteins by LALDI-MS.

2.27 **2.28**

*Figure 2.19: Chemical structure for 3-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **2.27**.*

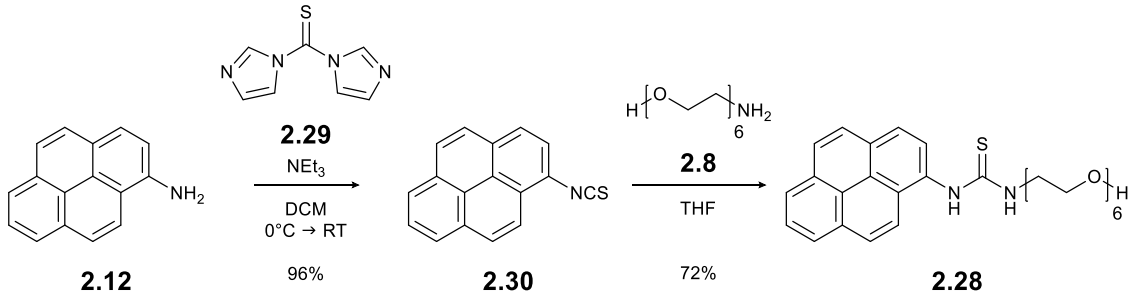
Pyrene urea **2.27** was synthesized in a single step via a urea formation reaction between pyrene isocyanate **2.24** and amine **2.8**, affording pyrene urea **2.28** in 56% yield *(Scheme 2.10)*.[215]



2.24 **2.27**

*Scheme 2.10: Synthesis of 3-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **2.27**.*[215]

Pyrene thiourea **2.28** was synthesized in 2 steps from aminopyrene **2.12** using conditions adapted from the literature (*Scheme 2.11*).[215,218] Pyrene isothiocyanate **2.30** was prepared in 96% yield following treatment of aminopyrene **2.12** with 1,1-thiocarbonyldiimidazole **2.29**.[218] Isothiocyanate **2.30** and amine **2.8** were then combined in a thiourea formation reaction to afford pyrene thourea **2.30** in 72% yield.[215]



*Scheme 2.11: Synthetic route to 3-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)thiourea **2.28**.*[215,218]

2.6.1 LALDI-MS analysis

LALDI-MS analysis of pyrene urea **2.27** and pyrene thiourea **2.28** was carried out under the optimized analytical protocol (Section 2.3.2) with analytical samples prepared via the previously described procedure for LALDI-MS (Section 2.2.3).

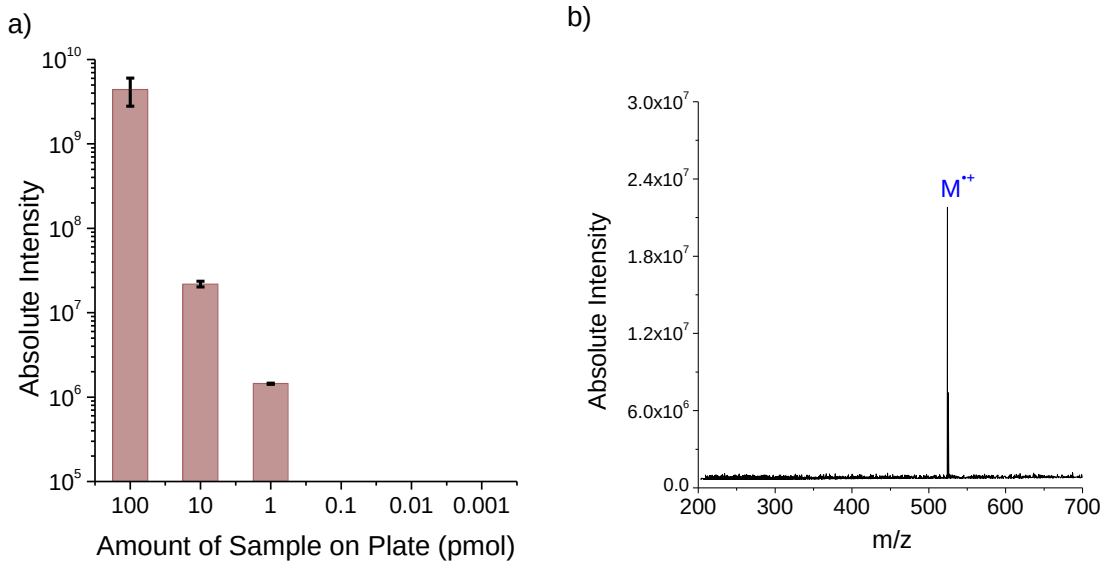


*Figure 2.20: LALDI-MS analysis of pyrene urea **2.27**. a) Analysis of the LALDI-MS lower detection limit for **2.27**. Bars represent the mean sum of all (+) ions related to of **2.27** ($M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$) that were detected by LALDI-MS at decreasing sample amounts. Error bars represent the standard deviation of the mean, where n=3 (individual samples on separate target spots). b) Representative LALDI-MS spectrum of **2.27** (10 pmol) with the corresponding signal for the $M^{\bullet+}$ (m/z 524.2518) highlighted.*

Pyrene urea **2.27** was detected exclusively as the $M^{\bullet+}$ with no additional fragmentations or degradations observed in the LALDI-MS spectrum (*Figure 2.20b*). The lower detection limit for **2.27** was found to be 1 pmol (*Figure 2.20a*). The LALDI-MS spectrum for pyrene urea **2.27** was found to be comparable to the one previously recorded for pyrene urea **2.26** (*Figure 2.16*), but the observed lower detection limit of **2.27** was noted as being ten-fold higher than the lowest recorded detection of pyrene urea **2.26** (*Table 2.1*). A similar differences in detection limits had previously been observed between pyrene **1.12b** (0.01 pmol – Yoneda et al.)[100] and pyrene **2.1** (0.1 pmol – *Figure 2.10a*).

LALDI-MS analysis of pyrene thiourea **2.30** showed a number of peaks relating to identifiable fragments but was unable to detect the molecular ion or any common metal adduct ions at any amount (*Figure 2.21*). Therefore, no further investigations were undertaken using pyrene thioureas as ionization enhancers for LALDI-MS.

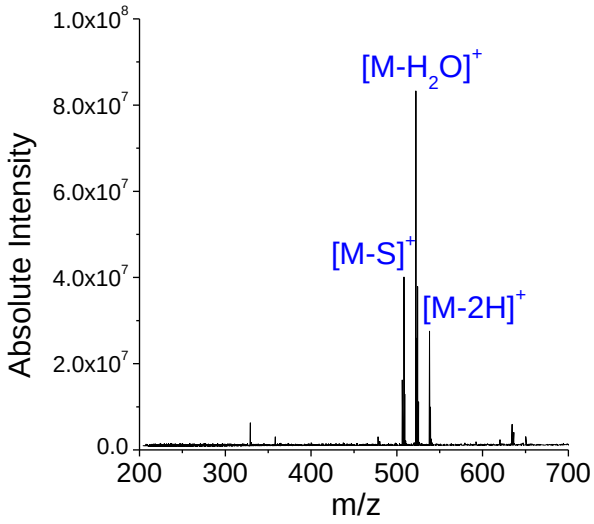


*Figure 2.21: LALDI-MS spectra for pyrene thiourea **2.28** (10 pmol) with the identified fragments and adducts annotated.*

The results of the LALDI-MS analysis confirmed that pyrene urea LDI enhancers could be used to develop LALDI reagents with improved MS stability compared to the amidopyrene labels but could probably not be used to achieve a comparable level of LALDI-MS sensitivity. Nevertheless, the improved stability exhibited by pyrene urea **2.27** was still considered a valuable property. Therefore, it was decided that both pyrene urea **2.27** and 1-amidopyrene **2.3** would be incorporated into the design of novel chemical reagents for future LALDI-MS investigations (see *Chapter 3*).

2.7 Conclusions

In summary, a series of investigations were undertaken to optimize the LALDI-MS detection of pyrene-based LDI enhancers and establish a reliable procedure for LALDI-MS analysis. Through these investigations, the novel 1-amidopyrene and pyrene urea LDI enhancers were developed, which exhibit high LDI sensitivity and good MS stability, respectively. Both novel LDI enhancers could be readily synthesized via robust synthetic routes and show promise for future applications as LDI enhancing labels in more complicated LALDI-MS studies.

Investigations into the preparation and excitation of samples in LDI-MS found that using a micro-focusing target plate was able to significantly improve the reproducibility of LALDI-MS data compared to the standard ground steel target.

Overall, these investigations have provided a better understanding of the how the structure and chemical functionality of an LDI enhancer affects the sensitivity and stability of reagents when analyzed by LALDI-MS.

2.8 Experimental

2.8.1 General experimental

All reactions were carried out under an atmosphere of dry nitrogen unless otherwise stated, using anhydrous solvents from a solvent purification system (Innovative Technology Inc. PureSolv), with the exception of anhydrous DMF, which was purchased from Acros Organics. Acetic formic anhydride **2.16** was prepared by combining acetic anhydride (1 equiv.) and formic acid (1.2 equiv.) with stirring at 55 °C for 2 h and used directly without purification. All other chemical starting reagents were purchased from commercial suppliers and used without further purification. Azeotropic distillation with toluene was used to remove persistent moisture from reagents with oligoethylene glycol chains before being used in a reaction. The identity and purity of known compounds was confirmed through the comparison of experimentally obtained data to values reported in the literature.

Thin layer chromatography (TLC) was carried out on Merck TLC Silica gel 60 F_{254} plates. Flash column chromatography was performed using Merck Geduran silica gel 60 (40–63 μm). All retention factors (R_f) are given to two decimal places along with the solvent system. Reversed-phase chromatography was performed using prepacked RediSep® Rf Reversed-phase C_{18} columns. Lyophilisation of compounds was performed using a Virtis Benchtop K freeze dryer.

The melting point of recrystallised solids was determined using a Griffin MFB-590 Melting Point Apparatus with a glass capillary melting point tube. Melting points (m.p.) are recorded as a range between the meniscus point and liquefaction with values given to the nearest degree in Celsius (°C).

^{1}H and ^{13}C NMR spectra were recorded in deuterated solvents on a Bruker AVANCE 500 Ultrashield or Bruker AVANCE 400 Ultrashield NMR Spectrometer. Chemical shifts are referenced to residual solvent peaks and are quoted in ppm. Coupling constants (*J*) are reported to the nearest 0.1 Hz. Assignment of spectra was based on expected chemical shifts and coupling constants, aided by DEPT, COSY, HMQC, and HMBC where appropriate.

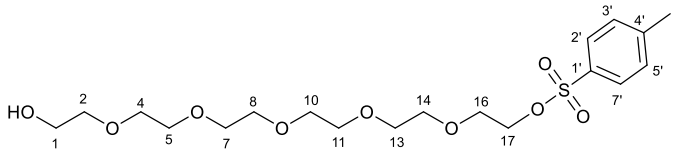
High-resolution electrospray mass spectrometry (HR-ESI-MS) was performed using a Bruker MaXis Impact spectrometer; *m/z* values are reported to four decimal places. LC-MS was recorded on an Agilent Technologies 1200 series HPLC combined with a Bruker HCT Ultra ion trap using 50 × 20 mm C_{18} reversed-phase columns with a solvent system of increasing acetonitrile (5 to 95%) in water, each containing 0.1% formic acid. A flow rate of 1.5 cm^3 min^{-1} was used and *m/z* values are given to one decimal place. Retention times (R_t) are provided in minutes to the nearest two decimal places. LALDI-MS was carried out on a solariX XR FTMS 9.4T MALDI mass spectrometer. All optimised analysis was performed on a Bruker Daltonics MTP AnchorChip Target 384. Ionisation/desorption of the samples was achieved using a fixed wavelength (355nm) Bruker Smartbeam II (N_2, NdYag) laser*; m/z* values are reported to four decimal places.

Infrared (IR) spectroscopy was carried out using a Bruker Alpha Platinum ATR. Samples were analysed neat and absorption maxima (ν_{max}) are given in wave numbers (cm^{-1}) to the nearest whole wavenumber. Signals are defined as either strong (s), medium (m), weak (w), or broad (br). UV-Vis absorption was measured using an Agilent Technologies Cary 100 UV-Vis Spectrophotometer with all samples were analysed as solutions in a 10 mm Hellma Analytics High Precision Quarts Suprasil cell. Absorption maxima (λ_{max}) are given in nanometers to the nearest whole nanometer with the corresponding molar extinction coefficient (ε) given in M^{-1} cm^{-1}.

Scanning electron microscopy (SEM) was carried out with assistance from Dr. Alexander Kulak using a Nova NanoSEM450 Scanning Electron Microscope. SEM elemental analysis was carried out using EDAX elemental mapping software. SEM analysis of the LDI-MS target plate was carried out on a Bruker Daltonics ground-steel MALDI MTP 384 target plate.

2.8.2 Experimental procedures

17-[(4-methylbenzenesulfonyl)oxy]-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.6** [195,196]



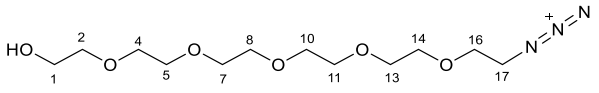
2.6

Silver (I) oxide (1.23 g, 5.31 mmol), and potassium iodide (118 mg, 1.71 mmol) were added to a stirred solution of hexaethylene glycol **2.4** (1.00 g, 3.54 mmol) in DCM (20 mL) at 0 °C (ice bath). A solution of 4-toluenesulfonyl chloride **2.5** (743 mg, 3.89 mmol) in DCM (20 mL) was then added drop-wise and the reaction mixture allowed to warm to room temperature and stir overnight (16 h). The reaction was filtered through a silica plug, eluted with EtOAc (250 mL), and concentrated in vacuo. The crude tosylate was subjected to flash chromatography (SiO_2; EtOAc–MeOH 95:5) to afford the product **2.6** as a colourless oil (1.37 g, 3.13 mmol, 88%). $\boldsymbol{R_f}$ = 0.18 (SiO_2; EtOAc–MeOH 95:5). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.78 (d, 2H, *J* 8.3 Hz, C*H*-2', C*H*-7'), 7.33 (d, 2H, *J* 8.3 Hz, C*H*-3', C*H*-5'), 4.15 (t, 2H, *J* 4.81, CH_2-17), 3.71 – 3.57 (m, 22H, OCH_2CH_2O), 2.44 (s, 3H, CH_3-4). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 144.9 (*C*CH_3-4'), 133.2 (*C*SO_2-1'), 129.9 (Aryl *C*H-3', *C*H-5'), 128.1 (Aryl *C*H-2', *C*H-7'), 72.7 (*C*H_2), 70.9 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.6 (*C*H_2), 70.4 (*C*H_2), 69.4 (*C*H_2-17), 68.8 (*C*H_2), 61.8 (*C*H_2), 21.8 (*C*H_3-4). **HR-ESI-MS** Calculated for $C_{19}H_{33}O_9S$: *m/z* 437.1840 $[M+H]^+$; found 437.1841. **LC-MS** Calculated for $C_{19}H_{33}O_9S$: *m/z* 437.18 $[M+H]^+$; found 437.6, $\boldsymbol{R_t}$ = 1.66.

1 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

Characterization is consistent with data reported in the literature.[1]

17-azido-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.7** [195,197]



2.7

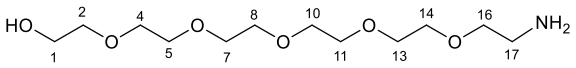
Sodium azide (726 mg, 11.2 mmol) was added to a stirred solution of 17-hydroxy-3,6,9,12,15-pentaoxaheptadecyl 4-methylbenzenesulfonate **2.6** (1.95 g, 4.47 mmol) in DMF (20 mL). The

mixture was then heated to 60 °C with stirring for 20 h, then left to cool to room temperature. The reaction was diluted with EtOAc (30 ml), filtered through a pad of SiO_2, and eluted with EtOAc (300 mL). The solvent was removed under reduced pressure to afford the product **2.7** as a clear yellow oil (1.23 g, 4.01 mmol, 90%). $\mathbf{R_f}$ = 0.24 (EtOAc–MeOH 10:1). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 3.70 – 3.69 (m, 2H, *CH*$_2$-1), 3.66 – 3.63 (m, 18H, OC*H*$_2$C*H*$_2$O), 3.59 – 3.57 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.36 (t, 2H, *J* 5.2 Hz, C*H*$_2$-17). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 72.7 (*C*H_2), 70.8 (*C*H_2), 70.8 (*C*H_2), 70.8 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.5 (*C*H_2), 70.2 (*C*H_2), 61.9 (*CH*$_2$OH-1), 50.8 (*C*H_2N_3-17). **HR-ESI-MS** Calculated for $C_{12}H_{26}N_3O_6$: *m/z* 308.1816 $[M+H]^+$; found 308.1823. **LC-MS** Calculated for $C_{12}H_{25}N_3O_6$: *m/z* 307.1738 $[M]^{+\bullet}$; found 307.8, $\boldsymbol{R_t}$ = 1.37. **IR** (Neat, ν_{max}/cm^{-1}) 3457 br s (OH hydroxy), 2099 s (N=N=N).

2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

Characterization is consistent with data reported in the literature.[56]

17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.8** [198]



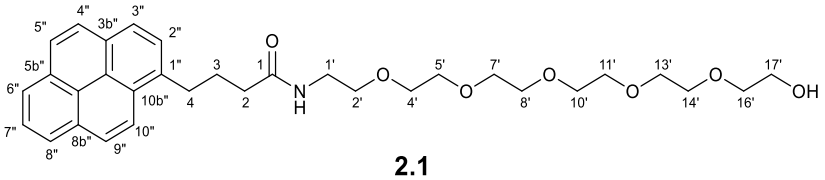
2.8

Triphenylphosphine (939 mg, 3.58 mmol) was added to a stirred solution of 17-azido-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.7** (1.00 g, 3.25 mmol) in THF (15 mL) at 0 °C (ice bath). The reaction mixture allowed to warm to room temperature and stir overnight (24 h). H_2O (1 mL) was added to the reaction mixture and left to stir (2 h) to hydrolyse the iminophosphorane intermediate. The reaction mixture was then concentrated under reduced pressure, diluted with H_2O (15 mL), and washed with toluene (3 × 15 mL). The aqueous layer was separated, and the water removed by lyophilization to afford the product **2.8** as a yellow oil (915 mg, 3.25 mmol, Quant.). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 3.70 – 3.69 (m, 2H, *CH*$_2$-1), 3.64 – 3.61 (m, 16H, OC*H*$_2$C*H*$_2$O), 3.59 – 3.57 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.51 (t, 2H, *J* 5.1 Hz, C*H*$_2$-16), 2.85 (t, 2H, *J* 5.1 Hz, C*H*$_2$-17). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 73.2 (*C*H_2), 73.0 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.6 (*C*H_2), 70.4 (*C*H_2), 70.3 (*C*H_2), 61.6 (*C*H_2OH-1), 41.8 (*C*H_2NH_3-17). **HR-ESI-MS** Calculated for $C_{12}H_{26}NO_6$: *m/z* 282.1911 $[M+H]^+$; found 282.1915. **LC-MS** found 282.0, $\boldsymbol{R_t}$ = 0.43. **IR** (Neat, ν_{max}/cm^{-1}) 3420 br m (OH hydroxy), 3361 m (NH_2 asym), 3309 m (NH_2 sym).

1 × CH_2 from the OEG chain was unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

Characterization is consistent with data reported in the literature.[198]

N-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-4-(pyren-1-yl)butanamide **2.1**



2.1

Hydroxybenzotriazole (112 mg, 0.832 mmol), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (119 mg, 0.624 mmol) and a solution of 17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.8** (107 mg, 0.624 mmol) in DCM (8.0 mL) were added sequentially to a stirred solution of pyrenebutyric acid **1.12a** (100 mg, 0.347 mmol) in dry DCM (8.0 mL). The reaction was stirred until determined complete by LC-MS (21 h). The reaction mixture was washed with sat. NH_4Cl (3 × 10 mL) and brine (10 mL). The organic layer was then dried (Na_2SO_4) and the solvent removed *in vacuo.* The crude residue was purified by flash chromatography (SiO_2; $CHCl_3$–MeOH 99:1 → 90:10) to afford the product **2.1** as a brown oil (163 mg, 0.295 mmol, 85%). $\boldsymbol{R_f}$ = 0.08 (SiO_2; $CHCl_3$–MeOH 95:5). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.32 (t, 1H, *J* 7.2 Hz, Aryl), 8.18 – 8.13 (m, 2H, Aryl), 8.12 – 8.08 (m, 2H, Aryl), 8.04 – 7.96 (m, 3H, Aryl), 7.88 (d, 1H, *J* 7.8 Hz, Aryl), 6.60 (br s, 1H, CON*H*-1), 3.64 (m, 2H, O*CH*$_2$*CH*$_2$O), 3.59 – 3.44 (m, 22H, O*CH*$_2$*CH*$_2$O), 3.40 (t, 2H, *J* 7.6 Hz, C*H*$_2$-4), 2.33 (t, 2H, *J* 7.2 Hz, C*H*$_2$-2), 2.26 – 2.19 (m, 2H, C*H*$_2$-3).**^{13}C NMR** (125 MHz, $CDCl_3$) δ 173.0 (*C*ONH-1), 136.3 (Aryl *C*), 131.6 (Aryl *C*), 131.1 (Aryl *C*), 130.0 (Aryl *C*), 129.0 (Aryl *C*), 127.6 (Aryl *C*H), 127.5 (Aryl *C*H), 127.5 (Aryl *C*H), 126.8 (Aryl *C*H), 126.0 (Aryl *C*H), 125.2 (Aryl *C*), 125.2 (Aryl *C*), 125.0 (Aryl *C*H), 124.9 (Aryl *C*H), 124.9 (Aryl *C*H), 123.7 (Aryl *C*H), 72.8 (*C*H$_2$), 70.7 (*C*H$_2$), 70.6 (*C*H$_2$), 70.6 (*C*H$_2$), 70.6 (*C*H$_2$), 70.5 (*C*H$_2$), 70.5 (*C*H$_2$), 70.3 (*C*H$_2$), 70.2 (*C*H$_2$), 70.1 (*C*H$_2$), 61.8 (*C*H$_2$), 39.4 (*C*H$_2$), 36.1 (*C*H$_2$-2), 33.0 (*C*H$_2$-4), 27.7 (*C*H$_2$-3). **HR-ESI-MS** Calculated for $C_{32}H_{42}NO_7$: *m/z* 552.2956 $[M+H]^+$; found 552.2961. **LC-MS** Calculated for $C_{32}H_{41}NNaO_7$: *m/z* 574.2775 $[M+Na]^+$; found 574.2, $\boldsymbol{R_t}$ = 1.89. **IR** (Neat, ν_{max}/cm^{-1}) 3325 br m (OH hydroxy), 3273 m (NH amide), 1642 s (C=O amide), 1525 – 1420 m (C=C aromatic). **UV-Vis** (MeOH) λ_{max} 234 nm (ε 25190 M^{-1}cm^{-1}), 243 nm (ε 41860 M^{-1}cm^{-1}), 265 nm (ε 15450 M^{-1}cm^{-1}), 276 nm (ε 28360 M^{-1}cm^{-1}), 326 nm (ε 17260 M^{-1}cm^{-1}), 342 nm (ε 25680 M^{-1}cm^{-1}), 355 nm (ε 730 M^{-1}cm^{-1}).

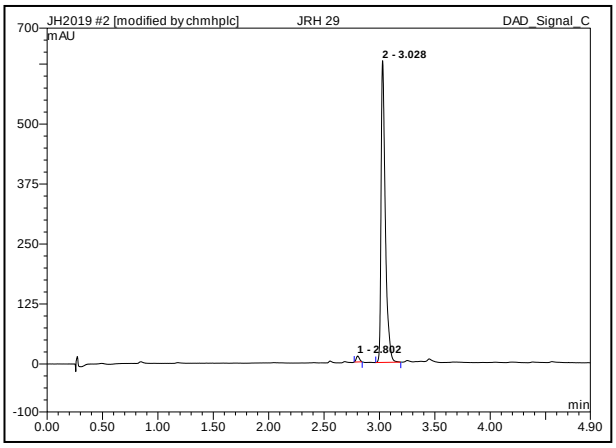
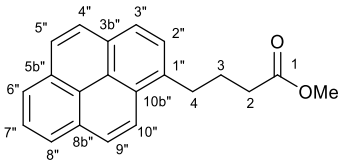


Figure 2.22: Analytical HPLC trace for N*-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-4-(pyren-1-yl)butanamide* ***2.1****.*

Methyl 4-(pyren-1-yl)butanoate **1.12b** [100]

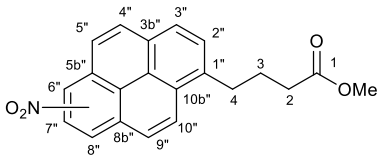


1.12b

Thionyl chloride (1.0 mL, 8.67 mmol) followed by dropwise addition of methanol (5 mL mmol) was added to a stirred suspension of pyrenebutyric acid **1.12a** (500 mg, 1.73 mmol) in DCM (7.5 mL) at -15 °C (NaCl/ice bath). The reaction mixture was then allowed to stir at room temperature until determined complete by TLC (3 h). The reaction mixture was then concentrated *in vacuo* and the crude residue was purified by flash chromatography (SiO_2; Hex–$CHCl_3$ 1:1 → 0:1) to afford **1.12b** as a yellow oil that solidified to yellow needles on cooling (523 mg, 1.73 mmol, Quant.). $\boldsymbol{R_f}$ = 0.19 (SiO_2; Hex–$CHCl_3$ 1:1). **m.p.** = 46–47 °C (Hex : $CHCl_3$). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.31 (d, 1H, *J* 9.2 Hz, Aryl), 8.20 – 8.15 (m, 2H, Aryl), 8.12 (dd, 2H, *J* 8.5, 3.2 Hz, Aryl), 8.05 – 7.96 (m, 3H, Aryl), 7.86 (d, 1H, *J* 7.8 Hz, Aryl), 3.70 (s, 3H, OC*H*$_3$), 3.40 (t, 2H, *J* 7.6 Hz, C*H*$_2$-4), 2.47 (t, 2H, *J* 7.3 Hz, C*H*$_2$-2), 2.21 (p, 2H, *J* 7.5 Hz, C*H*$_2$-3). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 174.1 (*C*OOMe-1), 135.8 (Aryl *C*), 131.6 (Aryl *C*), 131.0 (Aryl *C*), 130.1 (Aryl *C*), 128.9 (Aryl *C*), 127.6 (Aryl *C*H), 127.5 (Aryl *C*H), 127.5 (Aryl *C*H), 126.9 (Aryl *C*H), 126.0 (Aryl *C*H), 125.2 (Aryl *C*), 125.1 (Aryl *C*), 125.1 (Aryl *C*H), 124.9 (Aryl *C*H), 124.9 (Aryl *C*H), 123.4 (Aryl *C*H), 51.7 (O*C*H$_3$), 33.8 (*C*H$_2$-2), 32.9 (*C*H$_2$-4), 26.9 (*C*H$_2$-3). **HR-ESI-MS** Calculated for $C_{21}H_{19}O_2$: *m/z* 303.1379 $[M+H]^+$; found 303.1383. **LC-MS** Calculated for $C_{21}H_{18}NaO_2$: *m/z* 325.12 $[M+Na]^+$; found 325.2, $\boldsymbol{R_t}$ = 2.36.

Characterization is consistent with data reported in the literature.[100]

Methyl 4-(nitropyren-1-yl)butanoate **2.9** [100]



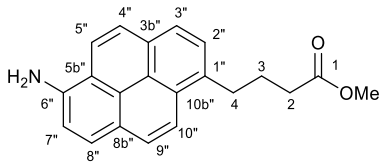
2.9

(mixture of 3 regioisomers)

44% nitric acid$_{(aq)}$ (110 µL, 1.00 mmol) was added to a stirred solution of methyl 4-(pyren-1-yl)butanoate **1.12b** (500 mg, 1.65 mmol) in acetic anhydride (11 mL) and stirred overnight. A second portion of 44% nitric acid (110 µL, 1.00 mmol) was added to the reaction and allowed to stir for another hour. The reaction mixture was then concentrated *in vacuo* and the crude residue was purified by flash chromatography (SiO_2; $CHCl_3$) to afford a 1:1:1 mixture of 3 regioisomers of **2.9** as an orange oil (573 mg, 1.65 mmol, Quant). $\boldsymbol{R_f}$ = 0.38–0.50 (SiO_2; $CHCl_3$). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.87 (d, 1H, *J* 9.7 Hz, Aryl), 8.76 (d, 2H, *J* 9.4 Hz, Aryl), 8.57 (t, 2H, *J* 8.1 Hz, Aryl), 8.50 (d, *J* = 9.8 Hz, 1H, Aryl), 8.45 – 8.41 (m, 2H, Aryl), 8.29 – 8.15 (m, 8H, Aryl), 8.11 (d, 1H, *J* 8.8 Hz, Aryl), 8.07 – 8.02 (m, 4H, Aryl), 7.96 (d, 1H, *J* 8.8 Hz, Aryl), 7.90 (d, 2H, *J* 7.8 Hz, Aryl), 3.65 (s, 3H, OC*H*$_3$), 3.64 (s, 3H+3H,2 × OC*H*$_3$), 3.38 – 3.30 (m, 2H+2H+2H, 3 × CH_2-4), 2.42 (m, 2H+2H+2H, 3 × CH_2-2), 2.17 – 2.09 (m, 2H+2H+2H, 3 × CH_2-3).**^{13}C NMR** (125 MHz, $CDCl_3$) δ 173.9 (*C*OOMe-1), 173.8 (*C*OOMe-1), 173.8 (*C*OOMe-1), 143.0 (2 × Aryl *C*) , 142.7 (Aryl *C*), 142.7 (Aryl *C*), 142.6 (Aryl *C*), 142.6 (Aryl *C*), 139.1 (Aryl *C*), 138.7 (Aryl *C*), 135.6 (Aryl *C*), 135.6 (Aryl *C*), 135.0 (Aryl *C*), 133.2 (Aryl *C*), 131.8(Aryl *C*H), 131.1 (Aryl *C*H), 131.0 (2 × Aryl *C*H), 130.9 (Aryl *C*), 130.5 (Aryl *C*), 129.8 (Aryl *C*), 129.1 (Aryl *C*), 128.8 (Aryl *C*), 128.7 (Aryl *C*H), 128.7 (Aryl *C*H), 128.1 (Aryl *C*), 127.8 (Aryl *C*H), 127.7 (Aryl *C*H), 127.6 (Aryl *C*H), 127.3 (Aryl *C*H), 127.2 (Aryl *C*H), 127.2 (Aryl *C*H), 127.0 (Aryl *C*H), 126.8 (Aryl *C*H), 126.4 (Aryl *C*H), 125.7 (Aryl *C*), 125.4 (Aryl *C*), 125.3 (Aryl *C*), 124.9 (Aryl *C*), 124.4 (Aryl *C*), 124.2(Aryl *C*), 124.2 (Aryl *C*H), 124.1 (Aryl *C*H), 123.6 (Aryl *C*), 123.6 (Aryl *C*H), 123.0 (Aryl *C*H), 122.9 (Aryl *C*H), 122.8 (Aryl *C*H), 121.9 (Aryl *C*H), 121.8 (Aryl *C*H), 121.1 (Aryl *C*H), 51.8 (O*C*H$_3$), 51.8 (2 × O*C*H$_3$), 33.7(2 × *C*H$_2$-2), 33.6 (*C*H$_2$-2), 33.1 (*C*H$_2$-4), 32.9 (*C*H$_2$-4), 32.9 (*C*H$_2$-4), 27.2 (2 × CH_2-3), 26.4 (*C*H$_2$-3). **HR-ESI-MS** Calculated for $C_{21}H_{18}NO_4$: *m/z* 348.1230 $[M+H]^+$; found 348.1233. **LC-MS** Calculated for $C_{21}H_{17}NNaO_2$: *m/z* 370.11 $[M+Na]^+$; found 370.1, $\boldsymbol{R_t}$ = 2.35. **IR** (Neat, ν_{max}/cm^{-1}) 1729 s (C=O ester), 1584–1435 w (C=C aromatic), 1506 s (NO_2 asym), 1307 s (NO_2 sym).

Characterization is consistent with data reported in the literature.[100]

Methyl 4-(6-aminopyren-1-yl)butanoate **1.13** [100]

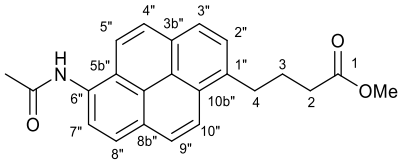


1.13

Throughout the reaction and work up the reaction and products were shielded from light. Under a N_2 atmosphere, acetic acid (1.00 mL, 18.0 mmol) and 10% Pd/C (180 mg) was added to a stirred solution of methyl 4-(nitropyren-1-yl)butanoate **2.9** (500 mg, 1.44 mmol) in EtOAc (20 mL). The inert atmosphere was then replaced with H_2 and stirred overnight (16h). The reaction was then filtered, and the residue washed with EtOAc then the filtrate was concentrated by azeotropic distillation (rotavap) with toluene. The crude residue was purified by flash chromatography (SiO_2; Toluene–EtOAc 99:1 → 97:3→ 95:5→ 93:7→ 9:1) to afford the product **1.13** as a yellow oil (101 mg, 0.318 mmol, 22%), $\boldsymbol{R_f}$ = 0.32 (SiO_2; Toluene–EtOAc 9:1) and a mixture of two regioisomers as an orange oil (232 mg, 0.731 mmol, 51%), $\boldsymbol{R_f}$ = 0.21 (SiO_2; Toluene–EtOAc 9:1). Characterisation data only given for 6-isomer of **1.13**. **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.04 – 7.87 (m, 6H, Aryl), 7.78 (d, 1H, *J* 7.7 Hz, Aryl), 7.38 (t, 1H, *J* 9.0 Hz, Aryl), 3.69 (s, 3H, OC*H*$_3$), 3.33 (t, 2H, *J* 7.7 Hz, C*H*$_2$-4), 2.46 (t, 2H, *J* 7.3 Hz, C*H*$_2$-2), 2.22 – 2.15 (m, 2H, C*H*$_2$-3). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 174.1 (*C*OOMe-1), 140.9 (Aryl *C*), 134.7 (Aryl *C*), 130.6 (Aryl *C*), 129.9 (Aryl *C*), 128.4 (Aryl *C*), 127.8 (Aryl *C*H), 127.6 (Aryl *C*H), 126.5 (Aryl *C*), 126.3 (Aryl *C*H), 126.0 (Aryl *CH*), 124.3 (Aryl *C*), 123.5 (Aryl *C*H), 119.8 (Aryl *C*H), 119.5 (Aryl *C*H), 117.3 (Aryl *C*), 114.2 (Aryl *C*H), 51.7 (O*C*H$_3$), 33.9 (*C*H$_2$-2), 33.0 (*C*H$_2$-4), 26.6 (*C*H$_2$-3). **HR-ESI-MS** Calculated for $C_{21}H_{20}NO_2$: *m/z* 318.1489 $[M+H]^+$; found 318.1482. **LC-MS** found 318.2, $\boldsymbol{R_t}$ = 2.39. **IR** (Neat, ν_{max}/cm^{-1}) 3465 m (NH_2 asym), 3373 m (NH_2 sym), 1720 s (C=O Ester), 1601–1498 m (C=C Aromatic).

Characterization is consistent with data reported in the literature.[100]

Methyl 4-(6-acetamidopyren-1-yl)butanoate **1.14** [100]

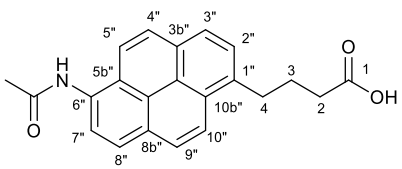


1.14

Acetic anhydride (36 µL, 0.38 mmol) was added to a stirred solution of methyl 4-(6-aminopyren-1-yl)butanoate **S11** (100 mg, 0.315 mmol) in DCM (3mL) and the mixture was allowed to stir overnight (16 h). The pH of the reaction mixture was then adjusted to 7 with sat. $NaHCO_{3\ (aq)}$. The layers were separated and the organic layer was washed with sat. $NaHCO_{3\ (aq)}$ (3 × 2 mL) and brine (5 mL). The organic layer was dried (Na_2SO_4) and solvent removed *in vacuo.* The crude residue was purified by flash chromatography (SiO_2; $CHCl_3$–Acetone 1:0 →10:1→ 5:1) to afford the product **1.14** as an amorphous off-white solid (94 mg, 0.27 mmol, 83%) $\boldsymbol{R_f}$ = 0.32 (SiO_2; $CHCl_3$–Acetone 9:1). **m.p.** = 196–197 °C (Amorphous). **^{1}H NMR** (500 MHz, DMSO) δ 10.27 (br s, 1H, CON*H*-1"), 8.32 (t, 1H, *J* 5.8 Hz, Aryl), 8.25 (d, 3H, *J* 10.2 Hz, Aryl), 8.22 – 8.13 (m, 3H, Aryl), 7.93 (d, 1H, *J* 7.8 Hz, Aryl), 3.61 (s, 3H, OCH_3), 3.33 (t, 2H, *J* 7.7 Hz, C*H*$_2$-4), 2.48 (t, 2H, *J* 7.2 Hz, C*H*$_2$-2), 2.27 (s, 3H, C*H*$_3$CONH-2"), 2.07 – 2.00 (m, 2H, C*H*$_2$-3). **^{13}C NMR** (125 MHz, DMSO) δ 173.2 (*C*OOMe-1), 169.0 (*C*ONH-1"), 136.2 (Aryl *C*), 131.8 (Aryl *C*), 129.2 (Aryl *C*), 128.3 (Aryl *C*), 127.9 (Aryl *C*), 127.7 (Aryl *C*H), 127.2 (Aryl *C*H), 127.1 (Aryl *C*H), 124.7 (Aryl *C*H), 124.7 (Aryl *C*), 124.6 (Aryl *C*H), 124.3 (Aryl *C*), 123.9 (Aryl *C*), 123.3 (Aryl *C*H), 122.5 (Aryl *C*H), 121.5 (Aryl *C*H), 51.2 (O*C*H$_3$), 32.9 (*C*H$_2$-2), 32.0 (*C*H$_2$-4), 26.6 (*C*H$_2$-3), 23.5 (*C*H$_3$CONH-2"). **HR-ESI-MS** Calculated for $C_{23}H_{21}NNaO_3$: *m/z* 382.1414 $[M+Na]^+$; found 382.1418. **LC-MS** found 382.3, $\boldsymbol{R_t}$ = 1.98. **IR** (Neat, $\nu_{max/}cm^{-1}$) 3265 m (NH), 1731 s (C=O Ester), 1682 (C=O Amide), 1602–1448 m (C=C Aromatic).

Characterization is consistent with data reported in the literature.[100]

4-(6-acetamidopyren-1-yl)butanoic acid **2.10** [100]



2.10

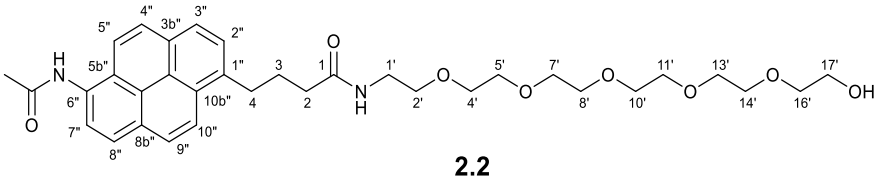
1 M $LiOH_{\ (aq)}$ (2.96 mL, 2.96 mmol) was added to a stirred solution of methyl 4-(6-acetamidopyren-1-yl)butanoate **1.14** (93 mg, 0.27 mmol) in dry THF (2 mL) and the mixture was allowed to stir overnight (24 h). The reaction mixture was then acidified to pH 3 using 1M HCl

$_{(aq)}$ and extracted with EtOAc (5 × 15 mL). The organic layers were combined, washed with brine (10 mL), dried (Na_2SO_4) and solvent removed *in vacuo.* The crude residue was then dry loaded onto ceilite (0.6 g) and purified by flash chromatography (SiO_2; $CHCl_3$–Acetone 2:1 → 1:1→ 1:2→ 0:1) to afford the product **2.10** as an amorphous off-white solid. (45.5 mg, 0.132 mmol, 49%). $\boldsymbol{R_f}$ = 0.49 (SiO_2; $CHCl_3$–Acetone 1:1). **m.p.** = 253–254 °C ($CHCl_3$: Acetone). **^{1}H NMR** (500 MHz, DMSO) δ 12.11 (br s, 1H, O*H*), 10.28 (br s, 1H, CON*H*-1"), 8.33 (d, 1H, *J* 9.2 Hz, Aryl), 8.28 – 8.22 (m, 3H, Aryl), 8.22 – 8.14 (m, 3H, Aryl), 7.92 (d, 1H, *J* 7.8 Hz, Aryl), 3.35 – 3.32 (m, 2H, CH_2-4), 2.39 (t, 2H, *J* 7.2 Hz, 2H, CH_2-2), 2.28 (s, 3H, CH_3CONH), 2.05 – 1.98 (m, 2H, CH_2-3).**^{13}C NMR** (125 MHz, DMSO) δ 174.4 (*C*OOH-1), 169.1 (CH_3*C*ONH-1"), 136.4 (Aryl *C*), 131.8 (Aryl *C*), 129.2 (Aryl *C*), 128.4 (Aryl *C*), 127.9 (Aryl *C*), 127.7 (Aryl *C*H), 127.2 (Aryl *C*H), 127.1 (Aryl *C*H), 124.7 (Aryl *C*H), 124.6 (Aryl *C*H), 124.4 (Aryl *C*), 124.0 (Aryl *C*), 123.3 (Aryl *C*H), 122.5 (Aryl *C*H), 121.5 (Aryl *C*H), 33.3 (*C*H_2-2), 32.1 (*C*H_2-4), 26.8 (*C*H_2-3), 23.5 (*C*H_3CONH-2"). **HR-ESI-MS** Calculated for $C_{23}H_{21}NO_3$: *m/z* 346.1438 $[M+H]^+$; found 346.1432. **LC-MS** found 346.5, $\boldsymbol{R_t}$ = 1.74. **IR** (Neat, ν_{max}/cm^{-1}) 3443 br m (OH acid), 3272 m (NH), 1713 s (C=O Acid), 1682 s (C=O Amide), 1553–1498 m (C=C Aromatic).

1 × aromatic C was unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

Characterization is consistent with data reported in the literature.[100]

4-(6-acetamidopyren-1-yl)-*N*-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)butanamide **2.2**



2.2

Hydroxybenzotriazole (25 mg, 0.19 mmol), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (25 mg, 0.14 mmol) and a solution of 17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.8** (22 mg, 0.077 mmol) in dry DCM (2.5 mL) were added sequentially to a stirred suspension of 4-(6-acetamidopyren-1-yl)butanoic acid **2.10** (24 mg, 0.069 mmol) in dry DCM (2.5 mL). The reaction was stirred until determined complete by LC-MS (21 h). The reaction mixture was diluted with DCM (10 mL) and filtered to remove precipitate. The filtrate was washed with sat. NH_4Cl (3 × 10 mL) and brine (10 mL). The organic layer was then dried (Na_2SO_4) and the solvent removed *in vacuo.* The crude residue was purified by flash chromatography (SiO_2; $CHCl_3$–MeOH 1:0 → 95:5) to afford the product **2.2** as a yellow waxy solid (27 mg, 0.044 mmol, 63%). $\boldsymbol{R_f}$ = 0.08 (SiO_2; $CHCl_3$–MeOH 95:5). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.25 (d, 1H, *J* 8.2 Hz, Aryl), 8.21

– 8.17 (m, 1H, Aryl), 8.06 (d, 1H, *J* 8.2 Hz, Aryl), 8.01 – 7.94 (m, 3H, Aryl), 7.80 (d, 1H, *J* 7.8 Hz, Aryl), 6.79 (br t, 1H, *J* 5.1 Hz, CON*H*-1), 3.63 (d, 2H, *J* 4.4 Hz, OC*H*$_2$C*H*$_2$O), 3.58 – 3.44 (m, 22H, OC*H*$_2$C*H*$_2$O), 3.33 (t, 2H, *J* 7.6 Hz, C*H*$_2$-4), 2.38 (s, 3H, C*H*$_3$CONH-2'''), 2.32 (t, 2H, *J* 7.2 Hz, C*H*$_2$-2), 2.22 – 2.14 (m, 2H, C*H*$_2$-3). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 173.1 (*C*ONH-1), 169.5 (CH_3*C*ONH-1'''), 136.7 (Aryl *C*), 130.3 (Aryl *C*), 129.6 (Aryl *C*), 129.1 (Aryl *C*), 129.0 (Aryl *C*), 128.0 (Aryl *C*H), 127.7 (Aryl *C*H), 127.3, (Aryl *C*H), 125.5 (Aryl *C*), 125.2 (Aryl *C*), 124.9 (Aryl *C*H), 124.4 (Aryl *C*), 123.1 (Aryl *C*H), 123.0 (Aryl *C*H), 119.9 (Aryl *C*H), 72.8 (*C*H_2), 70.6 (*C*H_2), 70.5 (*C*H_2), 70.4 (*C*H_2), 70.4 (*C*H_2), 70.4 (*C*H_2), 70.1, (*C*H_2), 70.0 (*C*H_2), 61.7 (*C*H_2), 39.4 (*C*H_2), 36.0 (*C*H_2-2), 33.0 (*C*H_2-4), 27.6 (*C*H_2-3), 24.4 (*C*H_3CONH-2'''). **HR-ESI-MS** Calculated for $C_{34}H_{44}N_2NaO_8$: *m/z* 631.2990 $[M+Na]^+$;: found 552.2961. **LC-MS** Calculated for $C_{34}H_{45}N_2O_8$: *m/z* 609.3170 $[M+Na]^+$; found 609.35, $\boldsymbol{R_t}$ = 0.52. **IR** (Neat, ν_{max}/cm^{-1}) 3313 br m (OH hydroxy), 3269 m (NH amide), 1639 s (C=O amide), 1550 – 1459 m (C=C aromatic), 1093 s (C-O aliphatic ether). **UV-Vis** (MeOH) λ_{max} 245 nm (*ε* 41029 $M^{-1}cm^{-1}$), 279 nm (*ε* 40023 $M^{-1}cm^{-1}$), 348 nm (*ε* 39430 $M^{-1}cm^{-1}$), 355 nm (*ε* 25090 $M^{-1}cm^{-1}$).

1 × aromatic CH and 2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

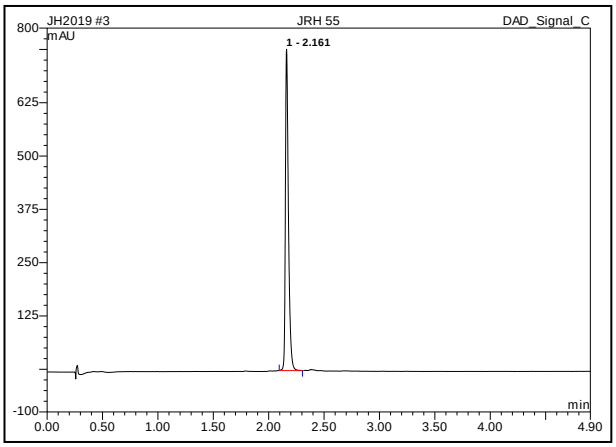
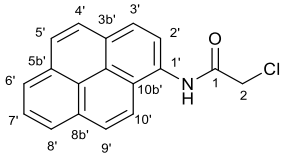


*Figure 2.23: Analytical HPLC trace for 4-(6-acetamidopyren-1-yl)-*N*-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)butanamide* ***2.2****.*

2-chloro-*N*-(pyren-1-yl) acetamide **2.11** [200]

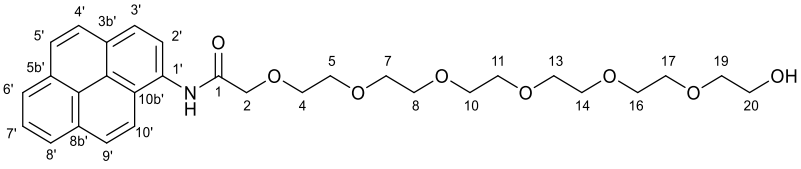


2.11

Under an N_2 atmosphere, triethylamine (1.27 mL, 9.12 mmol) and chloroacetyl chloride **2.13** (550 µL, 6.91 mmol) were added dropwise and sequentially to a stirred solution of 1-aminopyrene **2.12** (1.00 g, 4.56 mmol) in THF (20 mL) at 0 °C (ice bath). The reaction was then allowed to warm to room temperature and stir until determined complete by TLC (5 h). The solvent was removed under reduced pressure and the residue taken up in DCM (250 mL) and washed with sat. $NaHCO_3$ (2 × 50 mL), H_2O (50 mL), 1M HCl (2 × 50 mL), H_2O (2 × 50 mL), and brine (20 ml). The organic layer was dried (Na_2SO_4) and concentrated *in vacuo* to afford the product **2.11** as a brown solid (1.32 g, 4.49 mmol, 98%). $\boldsymbol{R_f}$ = 0.65 (SiO_2; DCM–EtOAc 90:10). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.99 (br s, 1H, CON*H*-1), 8.41 (d, 1H, *J* 8.3 Hz, Aryl), 8.22 – 8.14 (m, 3H, Aryl), 8.11 (d, 1H, *J* 9.2 Hz, Aryl), 8.06 – 7.98 (m, 4H, Aryl), 4.43 (s, 2H, C*H*$_2$-2). **^{13}C NMR** (125 MHz, $CDCl_3$) 164.8 (*C*ONH-1), 131.4 (Aryl *C*), 130.9 (Aryl *C*), 129.7 (Aryl *C*), 129.01 (Aryl *C*), 128.6 (Aryl *C*H), 127.4 (Aryl *C*H), 127.4 (Aryl *C*H), 126.4 (Aryl *C*H), 125.9 (Aryl *C*H), 125.5 (Aryl *C*H), 125.3 (Aryl *C*H), 125.2 (Aryl *C*), 124.8 (Aryl *C*), 123.8 (Aryl *C*), 121.9 (Aryl *C*H), 119.8 (Aryl *C*H), 43.5 (*C*H$_2$-2). **HR-ESI-MS** Calculated for $C_{18}H_{12}ClNNaO$: *m/z* 316.0500 $[M+Na]^+$; found 316.0494. **LC-MS** Calculated for $C_{18}H_{13}ClNO$: m/z 294.06802 $[M+H]^+$; found 294.33. **IR** (Neat, ν_{max}/cm^{-1}) 3233 m (NH amide), 1665 s (C=O amide), 1557–1409 m (C=C aromatic), 840 s (C-Cl alkyl).

Characterization is consistent with data reported in the literature.[200]

20-hydroxy-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **2.3**



2.3

A solution of *N*-pyren-1-yl chloroacetamide **2.11** (500 mg, 1.70 mmol) in DMF (2.5 mL) was added dropwise to an ice-cooled and stirred solution of hexaethylene glycol **2.4** (1.28 mL, 5.09 mmol) and 60% NaH in mineral oil (204 mg, 5.10 mmol) in DMF (3.5 mL) and stirred for 30

minutes. The reaction was then allowed to warm to room temperature and stir overnight until determined complete by LC-MS (16 h). The reaction was quenched with 1M HCl (10 mL) then extracted with $CHCl_3$ (3 × 10 mL). The org layers were collected and washed with 1M LiCl (5 × 10 mL), and brine (10 mL). The organic layer was then dried (Na_2SO_4) and the solvent removed *in vacuo* to give a crude brown oil. The crude product was then purified by flash chromatography (SiO_2; DCM–EtOAc–MeOH 9:9:2) to yield the product **2.3** as a brown oil (597 mg, 1.11 mmol, 65%). R_f = 0.35 (SiO_2; DCM–EtOAc–MeOH 9:9:2) **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.61 (br s, 1H, CON*H*-1), 8.48 (d, 1H, *J* 8.2 Hz, Aryl), 8.21 – 8.15 (m, 4H, Aryl), 8.11 (d, 1H, *J* 9.2 Hz, Aryl), 8.05 – 7.98 (m, 3H, Aryl), 4.36 (s, 2H, C*H*$_2$-2), 3.98 – 3.92 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.85 – 3.80 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.73 – 3.68 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.68 – 3.63 (m, 4H, OC*H*$_2$C*H*$_2$O), 3.56 – 3.47 (m, 8H, OC*H*$_2$C*H*$_2$O), 3.45 – 3.41 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.40 – 3.35 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.27 (s, 4H, OC*H*$_2$C*H*$_2$O). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 169.3 (*C*ONH-1), 131.5 (Aryl *C*), 131.0 (Aryl *C*), 130.3 (Aryl *C*), 129.3 (Aryl *C*), 128.0 (Aryl *C*H), 127.5 (Aryl *C*H), 127.0 (Aryl *C*H), 126.3 (Aryl *C*H), 125.5 (Aryl *C*H), 125.4 (Aryl *C*H), 125.3 (Aryl *C*), 125.1 (Aryl *C*H), 124.9 (Aryl *C*), 123.9 (Aryl *C*), 122.4 (Aryl *C*H), 121.0 (Aryl *C*H), 72.7 (*C*H$_2$), 71.7 (*C*H$_2$), 71.2 (*C*H$_2$), 70.8 (*C*H$_2$), 70.6 (*C*H$_2$), 70.5 (*C*H$_2$), 70.4 (*C*H$_2$), 70.4 (*C*H$_2$), 70.3 (*C*H$_2$), 70.3 (*C*H$_2$), 61.8 (*C*H$_2$). **HR-ESI-MS** Calculated for $C_{30}H_{38}N_2O_{10}$: *m/z* 540.2591 $[M+H]^+$; found 540.2600. **LC-MS** found 540.30, R_t = 0.65. **IR** (Neat, ν_{max}/cm^{-1}) 3417 br m (OH hydroxy), 3269 m (NH amide), 1639 s (C=O amide), 1550–1458 m (C=C aromatic), 1093 s (C-O aliphatic ether). **UV-Vis** (MeOH) λ_{max} 242 nm (*ε* 35512 $M^{-1}cm^{-1}$), 276 nm (*ε* 26195 $M^{-1}cm^{-1}$), 340 nm (*ε* 24603 $M^{-1}cm^{-1}$), 355 nm (*ε* 10870 $M^{-1}cm^{-1}$).

2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

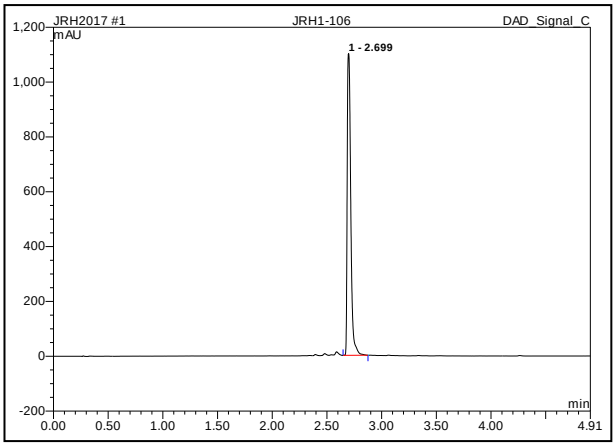
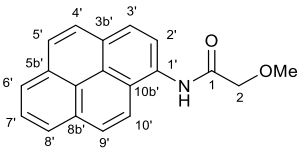


*Figure 2.24: Analytical HPLC trace for 20-hydroxy-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **2.3.***

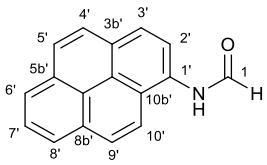
2-methoxy-*N*-(pyren-1-yl) acetamide **2.15**



2.15

2-chloro-*N*-(pyren-1-yl)acetamide **2.11** (128 mg, 0.436 mmol) was dissolved in a methanolic solution of NaOMe (25% w/w, 10 mL, 131 mmol) and allowed to stir at room temperature until determined complete by TLC (16 h). The pH of the reaction mixture was quenched with 2M HCl (40 ml) and extracted with DCM (5 × 30 mL). The organic layers were then combined, washed with brine (30 ml) and dried (Na_2SO_4). The solvent was then removed under reduced pressure to give the product **2.15** as a brown oil (126 mg, 0.436 mmol, Quant.) $\boldsymbol{R_f}$ = 0.09 (SiO_2; DCM). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.03 (s, 1H, CON*H*), 8.54 (d, 1H, *J* 8.3, Aryl), 8.17 – 8.11 (m, 3H, Aryl), 8.06 (d, 1 H, *J* 9.2, Aryl), 7.99 – 7.95 (m, 4H, Aryl), 4.23 (s, 2H, CH_2-2), 3.64 (3H, OCH_3). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 168.1 (*C*ONH), 131.3 (Aryl *C*), 130.8 (Aryl *C*), 129.6 (Aryl *C*), 129.0 (Aryl *C*), 128.0 (Aryl *C*H), 127.3 (Aryl *C*H), 126.8 (Aryl *C*H), 126.1 (Aryl *C*H), 125.5 (Aryl *C*H), 125.3 (Aryl *C*H), 125.1 (Aryl *C*), 125.0 (Aryl *C*H), 124.7 (Aryl *C*), 122.8 (Aryl *C*), 121.2 (Aryl *C*H), 119.7 (Aryl *C*H), 72.5 (CH_2-2), 59.5 (O*C*H$_3$). **HR-ESI-MS** Calculated for $C_{19}H_{16}NO_2$: *m/z* 290.1176 $[M+H]^+$; found 290.1171. **LC-MS** found, 290.31 $\boldsymbol{R_t}$ = 0.65. **IR** (Neat, ν_{max}/cm^{-1}) 3266 m (NH amide), 1668 s (C=O amide), 1515–1429 m (C=C aromatic).

N-(pyren-1-yl)formamide **2.17**



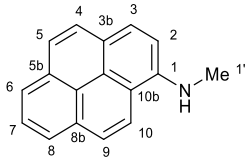
2.17

Acetic formic anhydride (55 µL, 0.69 mmol) was added to a stirred solution of 1-aminopyrene **2.12** (100 mg, 0.460 mmol) in THF (1 mL) and allowed to stir until determined complete by LC-MS (3 h). Solvent removed under reduced pressure to yield the product **2.17** as an off-white solid (113 mg, 0.460 mmol, Quant.). $\boldsymbol{R_f}$ = 0.19 (SiO_2; DCM–EtOAc 9:1). **^{1}H NMR** (400 MHz, DMSO) δ 10.85 (d, 2H, *J* 10.4 Hz, N*H*), 10.69 (br s, 5H, N*H*), 8.81 (d, 2H, *J* 10.5 Hz, C*H*O-1), 8.61 (s, 4H, C*H*O-1), 8.57 (d, 5H, *J* 8.3 Hz, Aryl), 8.38 (d, 7H, *J* 9.3 Hz, Aryl), 8.30 – 8.21 (m, 28H, Aryl), 8.17 – 8.02 (m, 24H, Aryl). **^{13}C NMR** (100 MHz, DMSO) δ 164.5 (*C*HO-1), 160.4 (*C*HO-1), 131.0 (Aryl *C*),

130.9 (Aryl *C*), 130.4 (Aryl *C*), 127.9 (Aryl *C*), 127.5 (Aryl *C*H), 127.3 (Aryl *C*H), 127.2 (Aryl *C*H), 126.6 (Aryl *C*H), 126.5 (Aryl *C*H), 125.6 (Aryl *C*H), 125.3 (Aryl *C*H), 125.1 (Aryl *C*H), 124.9 (Aryl *C*H), 124.4 (Aryl *C*), 123.9 (Aryl *C*), 122.1 (Aryl *C*), 121.8 (Aryl *C*H), 121.4 (Aryl *C*H), 121.3 (Aryl *C*H), 120.0 (Aryl *C*H). **HR-ESI-MS** Calculated for $C_{17}H_{12}NO$: 246.0913 *m/z* $[M+H]^+$; found 246.0908.2 **LC-MS** found 246.30, $\boldsymbol{R}_t$ = 0.70. **IR** (Neat, ν_{max}/cm^{-1}) 3269 m (NH amide), 1696 s (C=O amide), 1550–1458 m (C=C aromatic).

From the experimentally obtained data it was noted that **2.17** appears as a 2:5 mixture of rotamers in the ^{13}C and ^{1}H NMR spectra.

N-methylpyren-1-amine **2.18**

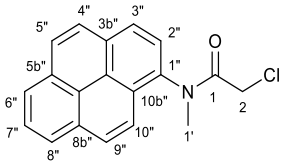


2.18

Under an N_2 atmosphere, 1M $LiALH_4$ (377 µL, 0.377 mmol) was added dropwise to a rapidly stirred suspension of *N*-(pyren-1-yl)formamide **2.17** (75 mg, 0.31 mmol) in THF (2.0 mL) at 0 °C (ice bath). The reaction was then allowed to warm to room temperature and stir until determined complete by LC-MS (2 h). The reaction was then cooled to 0 °C and quenched with EtOAc (10 mL). $Na_2SO_4 \bullet 10H_2O$ and Celite (1:1 w/w) (1.00 g) was added to the reaction mixture and allowed to stir (10 min) then filtered and washed thoroughly with EtOAc (100 mL). The solvent was removed under reduced pressure to yield the product **2.18** as a brown waxy solid (70 mg, 0.30 mmol, 99%). $\boldsymbol{R}_f$ = 0.18 (SiO_2; DCM–Hex 40:60). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.12 – 7.87 (m, 7H, Aryl), 7.77 (d, 1H, *J* 8.8 Hz, Aryl), 7.32 (d, 1H, *J* 8.4 Hz, Aryl), 4.85 (br s, 1H, N*H*), 3.16 (s, 3H, C*H*$_3$-1'). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 143.6 (Aryl *C*), 132.6 (Aryl *C*), 131.8 (Aryl *C*), 127.9 (Aryl *C*H), 126.6 (Aryl *C*H), 126.0 (Aryl *C*H), 125.9 (Aryl *C*H), 125.9 (Aryl *C*), 123.9 (Aryl *C*H), 123.3 (Aryl *C*H), 123.1 (Aryl *C*), 119.5 (Aryl *C*H), 116.6 (Aryl *C*), 108.4 (Aryl *C*H), 31.3 (*C*H$_3$-1'). **HR-ESI-MS** Calculated for $C_{17}H_{14}N$: 232.1121 *m/z* $[M+H]^+$; found 232.1115. **LC-MS** found 232.33, $\boldsymbol{R}_t$ = 0.85. **IR** (Neat, ν_{max}/cm^{-1}) 3267 m (NH amine), 1555 – 1455 m (C=C aromatic).

1 × aromatic C and 1 × aromatic CH were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

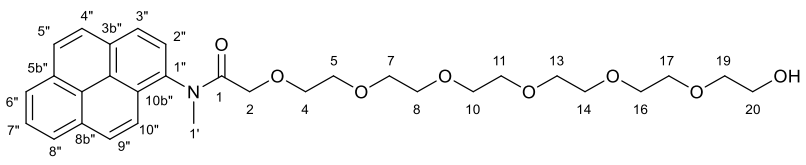
2-chloro-*N*-methyl-*N*-(pyren-1-yl)acetamide **2.19**



2.19

Under an N_2 atmosphere, triethylamine (255 µL, 1.83 mmol) and chloroacetyl chloride **2.13** (100 µL, 1.26 mmol) were added sequentially to a stirred solution of *N*-methylpyren-1-amine **2.18** (200 mg, 0.912 mmol) in THF (5.0 mL) at 0 °C (ice bath). The reaction was then allowed to warm to room temperature and stir until determined complete by TLC (1 h). The solvent was removed under reduced pressure and the residue taken up in DCM (30 mL) and washed with sat. $NaHCO_3$ until the aqueous layer remained basic (pH 8). The organic phase was washed with H_2O (30 mL × 2), and brine (20 ml), then dried (Na_2SO_4) and concentrated *in vacuo* to afford the product **2.19** as an amorphous yellow solid (280 mg, 0.911 mmol, Quant.). $\boldsymbol{R_f}$ = 0.65 (SiO_2; DCM–EtOAc 90:10). **^{1}H NMR** (400 MHz, $CDCl_3$) ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.33 – 8.20 (m, 4H, Aryl), 8.17 (d, 1H, *J* 8.8 Hz, Aryl), 8.14 – 8.06 (m, 2H, Aryl), 8.01 (d, 1H, *J* 9.2 Hz, Aryl), 7.91 (d, 1H, *J* 8.1 Hz, Aryl), 3.78 (d, 1H, *J* 13.4 Hz, C*H*-2), 3.73 (d, 1H, *J* 13.4 Hz, C*H*-2), 3.56 (s, 3H, CH_3-1'). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 167.4 (*C*$CONR_2$-1), 135.6 (Aryl *C*), 136.5 (Aryl *C*), 131.8 (Aryl *C*), 131.2 (Aryl *C*), 130.9 (Aryl *C*), 130.0 (Aryl *C*H), 128.8 (Aryl *C*H), 127.8 (Aryl *C*), 127.1 (Aryl *C*H), 126.9 (Aryl *C*H), 126.5 (Aryl *C*H), 126.3 (Aryl *C*H), 125.8 (Aryl *C*H), 125.6 (Aryl *C*H), 124.6 (Aryl *C*), 120.9 (Aryl *C*H), 42.1 (*C*H_2-2), 38.4 (*C*H_3-1"). **HR-ESI-MS** Calculated for $C_{19}H_{15}ClNO$: 308.0837 *m/z* $[M+H]^+$; found 308.0830. **LC-MS** found 308.33, $\boldsymbol{R_t}$ = 0.8. **IR** (Neat, ν_{max}/cm^{-1}) 1660 s (C=O amide), 1557–1409 m (C=C aromatic), 843 s (C-Cl alkyl).

20-hydroxy-*N*-methyl-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **2.20**



2.20

A solution of 2-chloro-*N*-methyl-*N*-(pyren-1-yl)acetamide **2.19** (133 mg, 0.432 mmol) in DMF (1 mL) was added dropwise to an ice-cooled and stirred solution of hexaethylene glycol **2.4** (325 µL, 1.30 mmol) and 60% NaH in mineral oil (57 mg, 1.4 mmol) in DMF (1 mL) and stirred for 30 minutes. The reaction was then allowed to warm to room temperature and stir overnight until

determined complete by LC-MS (16 h). The reaction was quenched with 1M HCl (1 mL) then extracted with $CHCl_3$ (3 × 2 mL). The org layers were collected and washed with 1M LiCl (5 × 2 mL), and brine (2 mL). The organic layer was then dried (Na_2SO_4) and the solvent removed *in vacuo* to give a crude brown oil. The crude product was then purified by flash chromatography (SiO_2; DCM–MeOH 95:5) to yield the product **2.20** as a brown oil (3 mg, 0.005 mmol, >1%). R_f = 0.17 (SiO_2; DCM–MeOH 95:5). **HR-ESI-MS** Calculated for $C_{31}H_{40}NO_8$: 554.2748 *m/z* $[M+H]^+$; found 554.2758. **LC-MS** found 554.65, R_t = 0.75. **UV-Vis** (MeOH) λ_{max} 242 nm (ε 20980 $M^{-1}cm^{-1}$), 265 nm (ε 12280 $M^{-1}cm^{-1}$), 276 nm (ε 17470 $M^{-1}cm^{-1}$), 312 nm (ε 6210 $M^{-1}cm^{-1}$), 326 nm (ε 12200 $M^{-1}cm^{-1}$), 341 nm (ε 16360 $M^{-1}cm^{-1}$), 355 nm (ε 3350 $M^{-1}cm^{-1}$).

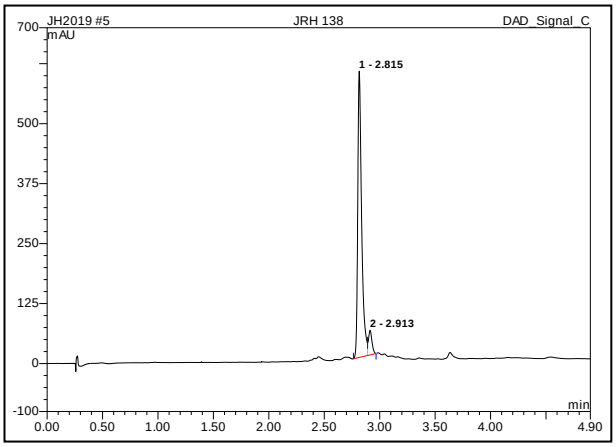
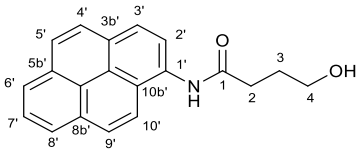


Figure 2.25: Analytical HPLC trace for 20-hydroxy-N-methyl-N-*(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **2.20***

4-hydroxy-*N*-(pyren-1-yl)butanamide **2.22**

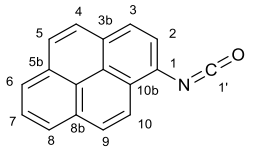


2.22

Under an N_2 atmosphere, 1.6 M ^{n}butyllithium in hexanes (633 µL, 1.01 mmol) was added dropwise to a stirred solution of 1-aminopyrene **2.12** (100 mg, 0.460 mmol) in anhydrous THF (2.0 mL) at 0 °C (ice bath) and allowed to stir (1 h). The reaction was then cooled to -78 °C (dry ice in acetone) and 4-butyrolactone **2.21** (72 µL, 0.47 mmol) was added dropwise with stirring. The reaction was then allowed to warm to room temperature and continue stirring until determined complete by TLC (21 h). The reaction was then quenched by addition of sat. NH_4Cl (20 mL) followed by extraction with EtOAc (3 × 20 mL). Organic extracts were then combined, washed with brine (20 mL), then dried (Na2SO4), and the solvent removed in vacuo. The crude

residue was then dry loaded onto celite (600 mg) and purified by flash chromatography (SiO_2; EtOAc–MeOH 1:0 → 95:5) to afford the product **2.22** as a brown amorphous solid (115 mg, 0.379 mmol, 83%). R_f = 0.44 (SiO_2; EtOAc–MeOH 90:10). **^{1}H NMR** (400 MHz, DMSO) δ 10.28 (s, 1H, CON*H*), 8.33 – 8.19 (m, 6H, Aryl), 8.18 – 8.11 (m, 2H, Aryl), 8.07 (t, 1H, *J* 7.6, Aryl), 4.61 (t, 1H, *J* 5.0, O*H*), 3.57 (q, 2H, *J* 6.0, CH_2-4), 2.64 (t, 2H, *J* 7.4 CH_2-2), 1.90 (p, 2H, *J* 7.0, CH_2-3). **^{13}C NMR** (100 MHz, DMSO) δ 172.2 (*C*ONH-1), 132.0 (Aryl *C*), 130.8 (Aryl *C*), 130.5 (Aryl *C*), 128.1 (Aryl *C*), 127.2 (Aryl *C*H), 127.0 (Aryl *C*H), 126.5 (Aryl *C*H), 126.4 (Aryl *C*H), 125.2 (Aryl *C*H), 124.9 (Aryl *C*H), 124.8 (Aryl *C*H), 124.4 (Aryl *C*), 123.9 (Aryl *C*), 123.8 (Aryl *C*), 123.5 (Aryl *C*H), 122.4 (Aryl *C*H), 60.4 (*C*H_2-4), (*C*H_2-2), 32.8, 28.7 (*C*H_2-3). **HR-ESI-MS** Calculated for $C_{20}H_{18}NO_2$: 304.1332 *m/z* $[M+H]^+$; found 304.1333. **LC-MS** found 304.33, R_t = 0.65. **IR** (Neat, ν_{max}/cm^{-1}) 3580 – 3142 br s (OH hydroxy), 3230 m (NH amide), 1545–1455 m (C=C aromatic). **UV-Vis** (MeOH) λ_{max} 243 nm (ε 6780 M^{-1}cm^{-1}), 276 nm (ε 7310 M^{-1}cm^{-1}), 340 nm (ε 7450 M^{-1}cm^{-1}), 355 nm (ε 7030 M^{-1}cm^{-1}).

1-isocyanatopyrene **2.24** [214]



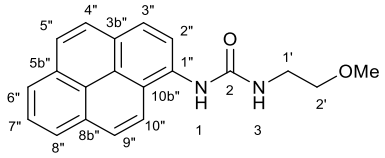
2.24

Triphosgene **2.23** (298 mg, 1.00 mmol) was added portion-wise slowly to a stirred solution of 1-aminopyrene **2.12** (200 mg, 0.912 mmol) in DCM (10 mL) at 0 °C (ice bath) and allowed to stir for 1hr. The reaction was then allowed to warm to room temperature with continued stirring (24 h). The reaction mixture was then dry loaded onto celite (200 mg) and purified by filtration through a short silica plug, eluting with DCM. Solvent was removed *in vacuo* to afford the product **2.24** as a yellow solid (214 mg, 0.880 mmol, 96%). R_f = 0.83 (SiO_2; DCM). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.29 (d, 1H, *J* 9.1 Hz, Aryl), 8.20 (d, 2H, *J* 7.6 Hz, Aryl), 8.15 (d, 1H, *J* 9.1 Hz, Aryl), 8.10 (d, 1H, *J* 8.1 Hz, Aryl), 8.07 – 8.01 (m, 3H, Aryl). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 131.5 (Aryl *C*), 131.2 (Aryl *C*), 129.4 (Aryl *C*), 128.6 (Aryl *C*H), 127.5 (Aryl *C*H), 127.17 (Aryl *C*H), 127.0 (N*C*O-1'), 126.7 (Aryl *C*H), 125.6 (Aryl *C*H), 125.5 (Aryl *C*H), 125.4 (Aryl *C*), 125.5 (Aryl *C*H), 123.4 (Aryl *C*), 123.2 (Aryl *C*H), 122.0 (Aryl *C*H). **IR** (Neat, ν_{max}/cm^{-1}) 2255 s (N=C=O).

2 × aromatic Cs were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

Characterization is consistent with data reported in the literature.[214]

3-(2-methoxyethyl)-1-(pyren-1-yl)urea **2.26** [215]

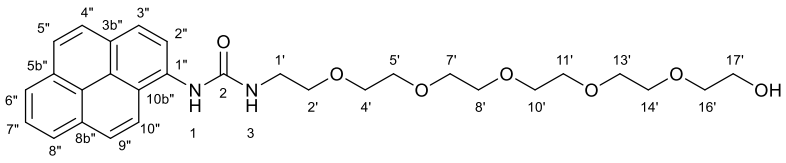


2.26

Under an N_2 atmosphere, 2-methoxyethylamine **2.25** (94 µL, 1.1 mmol) was added dropwise to a stirred solution of 1-isocyantopyrene **2.24** (175 mg, 0.719 mmol) in anhydrous THF (10 mL). The reaction was then allowed to stir until determined complete by TLC (16 h). The solvent was then removed in vacuo and the crude product dry loaded onto celite (250 mg) and purified by flash chromatography (SiO_2; EtOAc) to afford the product **2.26** as a brown waxy solid (83 mg, 0.26 mmol, 36%). $\boldsymbol{R_f}$ = 0.27 (SiO_2; EtOAc). **^{1}H NMR** (400 MHz, DMSO) δ 8.91 (s, 1H, N*H*-1), 8.64 (d, 1H, *J* 8.5 Hz, Aryl), 8.30 (d, 1H, *J* 9.3 Hz, Aryl), 8.13 – 8.04 (m, 4H, Aryl), 7.99 – 7.90 (m, 3H, Aryl), 6.77 (t, 1H, *J* 5.3 Hz, N*H*-3), 3.49 (t, 2H, *J* 5.0 Hz, CH_2-2'), 3.41 (q, *J* 5.0 Hz, 2H, CH_2-1'), 3.36 (s, 3H, OCH_3). **^{13}C NMR** (100 MHz, DMSO) δ 155.5 (NH*C*ONH-2), 133.8 (Aryl *C*), 131.0 (Aryl *C*), 130.5 (Aryl *C*), 127.0 (Aryl *C*H), 126.1 (Aryl *C*H), 125.7 (Aryl *C*H), 125.0 (Aryl *C*H), 124.8 (Aryl *C*H), 124.5 (Aryl *C*H), 124.3 (Aryl *C*H), 123.7 (Aryl *C*H), 120.8 (Aryl *C*H), 120.1 (Aryl *C*H), 119.0 (Aryl *C*H), 71.4 (*C*H_2-2'), 57.9 (O*C*H_3), 39.0 (CH_2-1'). **HR-ESI-MS** Calculated for $C_{20}H_{19}N_2O_2$: *m/z* 319.1441 $[M+H]^+$; found 319.1437. **LC-MS** found 319.3, $\boldsymbol{R_t}$ = 0.60. **IR** (Neat, ν_{max}/cm^{-1}) 3270 m (NH urea), 1626 s (C=O urea), 1521–1415 m (C=C aromatic). **UV-Vis** (MeOH) λ_{max} 223 nm (ε 133386 M^{-1}cm^{-1}), 278 nm (ε 23040 M^{-1}cm^{-1}), 341 nm (ε 21640 M^{-1}cm^{-1}), 355 nm (ε 17180 M^{-1}cm^{-1}).

Characterization is consistent with data reported in the literature.[215]

3-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **2.27**



2.27

A solution of 17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.8** (180 mg, 0.740 mmol) in THF (1.0 mL) was added dropwise to a stirred solution of 1-isocyantopyrene **2.24** (228 mg, 0.812 mmol) in THF (10 mL) at 0 °C (ice bath) and allowed to stir for 30 min. The reaction was then allowed to warm to room temperature and stir overnight (21 h). Solvent was then removed under reduced pressure and the crude residue purified by flash chromatography (SiO_2; DCM–

MeOH 95:5) to afford the product **2.27** as a brown waxy solid (217 mg, 0.414 mmol, 56%). R_f = 0.17 (SiO_2; DCM–MeOH 95:5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.35 (d, 1H, *J* 8.3 Hz, Aryl), 8.18 (d, 1H *J* 9.2 Hz, Aryl), 8.10 – 8.00 (m, 4H, Aryl), 7.94 – 7.88 (m, 4H, Aryl + N*H*-1), 6.25 (br s, 1H, N*H*-3), 3.66 – 3.42 (m, 24H, O*C*H_2*C*H_2O), 3.15 (br s, 1H, O*H*). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 157.2 (NH*C*ONH-2), 132.7 (Aryl *C*), 131.5 (Aryl *C*), 131.1 (Aryl *C*), 128.0 (Aryl *C*), 127.4 (Aryl *C*H), 127.1 (Aryl *C*H), 126.1 (Aryl *C*H), 126.0 (Aryl *C*H), 125.4 (Aryl *C*H), 125.3 (Aryl *C*), 124.9 (Aryl *C*H), 124.9 (Aryl *C*), 124.5 (Aryl *C*H), 123.5 (Aryl *C*), 122.0 (Aryl *C*H), 121.4 (Aryl *C*H), 72.6 (*C*H_2), 70.6 (*C*H_2), 70.6 (*C*H_2), 70.5 (*C*H_2), 70.4 (*C*H_2), 70.2 (*C*H_2), 70.2 (*C*H_2), 61.6 (*C*H_2), 40.4 (*C*H_2). **HR-ESI-MS** Calculated for $C_{29}H_{37}N_2O_7$: *m/z* 525.2595 $[M+H]^+$; found 525.2591. **LC-MS** found 525.54, R_t = 0.55. **IR** (Neat, ν_{max}/cm^{-1}) 3359 br m (OH hydroxy), 3285 m (NH urea), 1626 s (C=O urea), 1525–1416 m (C=C aromatic).

3 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

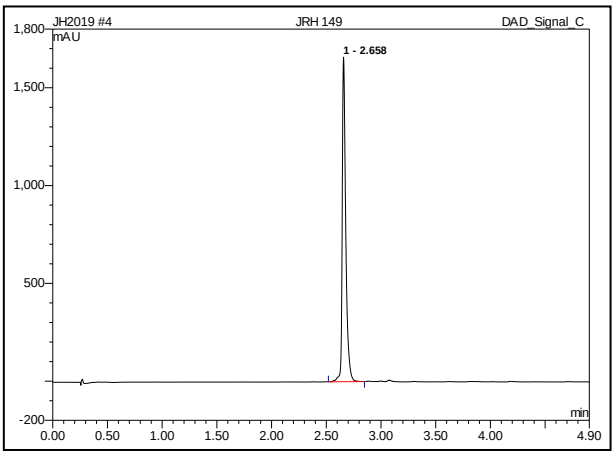
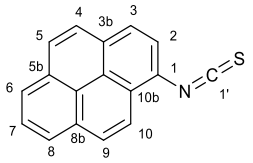


*Figure 2.26: Analytical HPLC trace for 3-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **2.27**.*

1-isothiocyanatopyrene **2.30** [218]

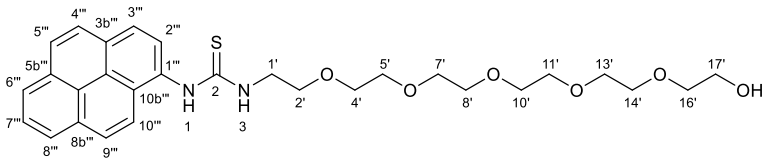


2.30

A solution of 1-aminopyrene **2.12** (500 mg, 2.30 mmol) in anhydrous DCM (5 mL) was added drop wise to a stirred solution of 1,1-thiocarbonyldiimidazole **2.29** (635 mg, 3.57 mmol) in anhydrous DCM (5.0 mL) at 0 °C (ice bath) and allowed to stir for 1hr. The reaction was then allowed to warm to room temperature with continued stirring (24 h). The reaction mixture was then concentrated and purified by filtration through a short silica plug, eluting with DCM. Solvent was removed *in vacuo* to afford the product **2.30** as a yellow solid (575 mg, 2.22 mmol, 96%). $\boldsymbol{R_f}$ = 0.79 (SiO_2; DCM). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.20 – 8.17 (3H, m, Aryl), 8.09 (d, 1H, *J* 9.2 Hz, Aryl), 8.06 – 7.99 (m, 3H, Aryl), 7.96 (d, 1H, *J* 8.9 Hz, Aryl), 7.80 (d, 1H, *J* 8.2 Hz, Aryl). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 136.4 (N=*C*=S), 131.3 (Aryl *C*), 131.0 (Aryl *C*), 130.3 (Aryl *C*), 128.2 (Aryl *C*H), 127.0 (Aryl *C*H), 126.9 (Aryl *C*), 126.8 (Aryl *C*H), 126.1 (Aryl *C*H), 126.0 (Aryl *C*H), 125.1 (Aryl *C*H), 125.0 (Aryl *C*), 124.2 (Aryl *C*), 124.1 (Aryl *C*), 123.9 (Aryl *C*H), 121.7 (Aryl *C*H). **HR-ESI-MS** Calculated for $C_{17}H_{10}NS$: *m/z* 260.0528 $[M+H]^+$; found 260.0527. **IR** (Neat, ν_{max}/cm^{-1}) 2098 s, 2026 s (N=C=S).

Characterization is consistent with data reported in the literature.[218]

3-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)thiourea **2.28**



2.28

A solution of 17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.8** (119 mg, 0.423 mmol) in THF (0.5 mL) was added dropwise to a stirred solution of 1-isothiocyanatopyrene **2.30** (100 mg, 0.386 mmol) in THF (1 mL) at 0 °C (ice bath) and allowed to stir for 30 min. The reaction was then allowed to warm to room temperature and stir overnight (21 h). Solvent was then removed under reduced pressure and the crude residue purified by flash chromatography (SiO_2; DCM–MeOH 95:5) to afford the product **2.28** as a yellow oil (143 mg, 0.264 mmol, 72%). $\boldsymbol{R_f}$ = 0.22

(SiO_2; DCM–MeOH 95:5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.46 (br s, 1H, NH-1), 8.22 – 8.16 (m, 4H, Aryl), 8.12 (d, 1H, *J* 9.3 Hz, Aryl), 8.09 (d, 1H, *J* 9.0 Hz, Aryl), 8.07 – 7.98 (m, 3H, Aryl), 6.86 (br s, 1H, NH-3), 3.86 (dd, 2H, *J* 10.2, 5.1 Hz, OC*H*$_2$C*H*$_2$O), 3.66 – 3.54 (m, 4H, OC*H*$_2$C*H*$_2$O), 3.50 – 3.46 (m, 6H, OC*H*$_2$C*H*$_2$O), 3.46 – 3.43 (m, 6H, OC*H*$_2$C*H*$_2$O), 3.32 (dd, 2H, *J* 6.0, 3.0 Hz, OC*H*$_2$C*H*$_2$O), 3.29 – 3.20 (m, 4H, OC*H*$_2$C*H*$_2$O). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 182.4 (NH*C*SNH-2), 131.2 (Aryl *C*), 131.0 (Aryl *C*), 130.9 (Aryl *C*), 130.1 (Aryl *C*), 128.7 (Aryl *C*H), 128.1 (Aryl *C*H), 127.6 (Aryl *C*), 127.2 (Aryl *C*H), 126.6 (Aryl *C*H), 125.9 (Aryl *C*H), 125.7 (Aryl *C*H), 125.6 (Aryl *C*), 125.4 (Aryl *C*H), 124.6 (Aryl *C*), 121.9 (Aryl *C*H), 72.7 (*C*H$_2$), 70.5 (*C*H$_2$), 70.4 (*C*H$_2$), 70.4 (*C*H$_2$), 70.3 (*C*H$_2$), 70.3 (*C*H$_2$), 70.1 (*C*H$_2$), 70.1 (*C*H$_2$), 69.4 (*C*H$_2$), 61.6 (*C*H$_2$), 45.3 (*C*H$_2$). δ **HR-ESI-MS** Calculated for $C_{29}H_{37}N_2O_6S$: *m/z* 541.2367 $[M+H]^+$; found 541.2373. **LC-MS** found 541.12, R_t = 0.60. **IR** (Neat, ν_{max}/cm^{-1}) 3267 m br (NH thiourea), 1528 s (C=S). **UV-Vis** (MeOH) λ_{max} 207 nm (*ε* 24160 $M^{-1}cm^{-1}$), 244 nm (*ε* 34650 $M^{-1}cm^{-1}$), 266 nm (*ε* 25490 $M^{-1}cm^{-1}$), 276 nm (*ε* 31530 $M^{-1}cm^{-1}$), 328 nm (*ε* 23110 $M^{-1}cm^{-1}$), 342 nm (*ε* 31710 $M^{-1}cm^{-1}$), 355 nm (*ε* 14830 $M^{-1}cm^{-1}$).

1 × aromatic CH and 1 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

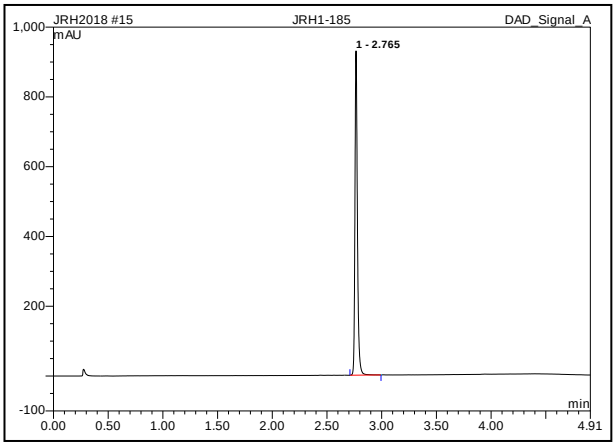


Figure 2.27: Analytical HPLC trace for 3-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)thiourea ***2.28****.*

Chapter 3 – LALDI tags for labelling and identifying carbohydrates

3.1 Introduction

The ubiquitous biological role of carbohydrates has made them excellent targets for developing improved diagnostic and therapeutic techniques (Section 1.3).[115] Mass spectrometry is recognized as one of the most important analytical tools for studying the structural and functional profile of native carbohydrates (Section 1.3.1).[27,28] However, the need for pure analytical samples can result in long and complicated workflows (Section 1.4.3 – 1.4.5).[28]

Recent developments in LALDI-MS have demonstrated its ability to selectively detect labelled analytes in the presence of complex samples, without requiring prior purification (Section 1.2.2).[1,107,112] These properties make LALDI-MS an attractive option for applications within glycomic analysis. Examples of reagents capable of labelling glycans and promoting "matrix free" ionization in LDI-MS have been reported in the literature (Section 1.2.2.2, Section 1.2.2.4).[112] However, in each case the analysis resulted in complex spectra due to problems with degradation or over-labelling.[102,112] While these results showed that labelled carbohydrates could be detected by LALDI-MS, issues regarding the sensitivity, stability, and labelling efficiency limit their applications in studying native glycans. It is proposed that by creating reagents with increased LALDI-MS sensitivity, stability, and glycan selectivity, LALDI-MS could be developed into a viable tool for glycomic analysis.

To address these challenges, a series of chemical tools referred to as "LALDI tags" were proposed. The design of these LALDI tags comprises an LDI-enhancing label conjugated to a glycan-selective reactive handle through a solubilizing linker (*Chapter 1 – Figure 1.28*). Glycans isolated from natural sources are often present in minute quantities and complex mixtures. Hence, effective detection of glycans using LALDI tags requires both an LDI enhancer with suitable sensitivity and a reactive handle that could selectively label glycans directly from crude, aqueous biological samples. The modular design of these reagents would allow optimization of the sensitivity, stability, selectivity, and solubility through modification of each constituent part.

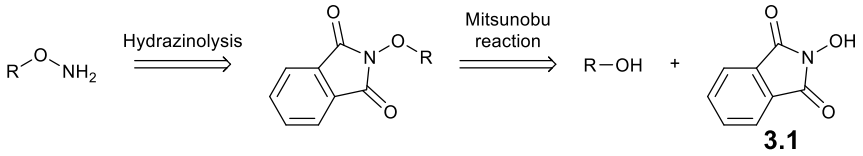
Following the results of Chapter 2, two LDI enhancers were selected for use in the LALDI tag reagents, 1-amidopyrene **2.3** for its high sensitivity (Section 2.3.2), and pyrene urea **2.27** for its improved MS stability (Section 2.6.1). Developing LALDI tags using these two different LDI enhancers would allow investigations with reagents displaying either increased stability or higher sensitivity.

Inspired by common glycan labelling strategies reported in the literature (Section 1.4.5),[171] three reactive handles were chosen for the LALDI tags: a hydroxylamine,[219–221] a hydrazide,[129,222] and 2-aminobenzamide .[223–225] Each of these handles was selected for its labelling efficiency, glycan selectivity, and ability to form a stable conjugate product in the presence of a complicated biological sample. Multiple reactive handles were chosen to determine which would be the best for both glycan labelling and LALDI-MS analysis, taking into account the labelling efficiency and the LALDI-MS sensitivity and stability of the labelled product.

3.2 Synthesis of LALDI tags for carbohydrate labelling

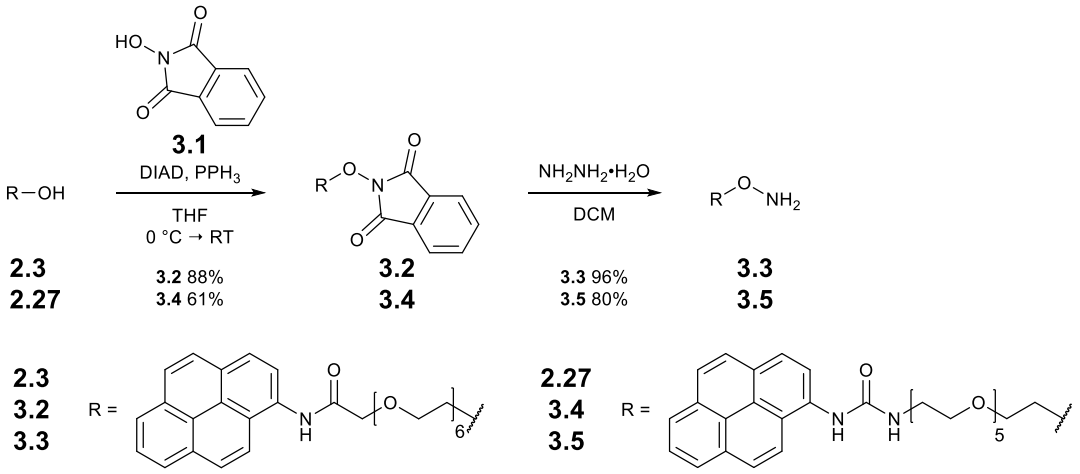
3.2.1 Synthesis of LALDI tags for labelling by glycans by oxime ligation

In order to label glycans via oxime ligation, a LALDI reagent functionalized with a hydroxylamine reactive handle was required. A simple method for installing an aminooxy group onto a molecule is through the hydrazinolysis of an *N*-hydroxyphthalimide moiety, which in turn can be generated through conjugation of an alcohol to *N*-hydroxyphthalimide **3.1**.



Scheme 3.1: Retrosynthetic analysis of a hydroxylamine-functionalized LALDI tag.

Both 1-amidopyrene-labelled hydroxylamine LALDI tag (AP-OEG-ONH_2) **3.1** and pyrene urea-labelled hydroxylamine LALDI tag (PU-OEG-ONH_2) **3.5** were synthesized in two steps from their respective water soluble LDI enhancers **2.3** and **2.27** via a Gabriel-like synthesis using conditions adapted from a literature procedure by Novoa-Carballal et al. (*Scheme 3.2*).[226]



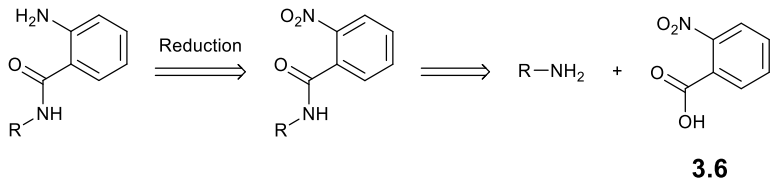
*Scheme 3.2: Synthetic route to 20-hydroxylamino-N-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.3** and 3-[17-(aminooxy)-3,6,9,12,15-pentaoxaheptadecan-1-yl]-1-(pyren-1-yl)urea **3.5**.*[226]

Alcohols **2.3** and **2.27** were converted to phthalimides **3.2** and **3.4** via a Mitsunobu reaction with *N*-hydroxyphthalimide **3.1** using diisopropylazodicarboxylate (DIAD) and triphenylphosphine to give **3.2** in 88% yield and **3.4** in 61% yield.[226] Phthalimides **3.2** and **3.4** were then subjected to hydrazinolysis to afford hydroxyl amine **3.3** in 96% yield and **3.5** in 80% yield.[226]

3.2.2 Synthesis of LALDI tags for labelling glycans by reductive amination

3.2.2.1 2-aminobenzamide reactive handle

It was reported by Locke et al. that secondary amide derivatives of 2-aminobenzamide were capable of derivatizing glycans as efficiently as 2-AB.[224] Encouraged by these results, a synthetic route towards 2-AB-functionalized LALDI tags was proposed which involved conjugating the reactive handle to the LDI enhancer though an amide bond formation between the commercially available 2-nitrobenzoic acid **3.6** and an amine-functionalized LDI enhancer (*Scheme 3.3*).



Scheme 3.3: Proposed retrosynthetic analysis of a 2-AB functionalized LALDI tag.

An amide linker was favored over an ester as amides linkages are more stable than esters and are less prone to hydrolysis. Furthermore, in a mechanism reported in the literature for the reductive amination of glycans with 2-AA **1.31** (*Figure 3.1*), it was proposed that the carboxylic

acid of 2-AA **1.31** plays an important role in the formation of the iminium intermediate, which in turn is stabilized by the carboxylate anion to aid in the reduction to the secondary amine.[227] Therefore, it was considered that having an available proton proximal to the amino group might be important for the efficiency of the reductive amination reaction.

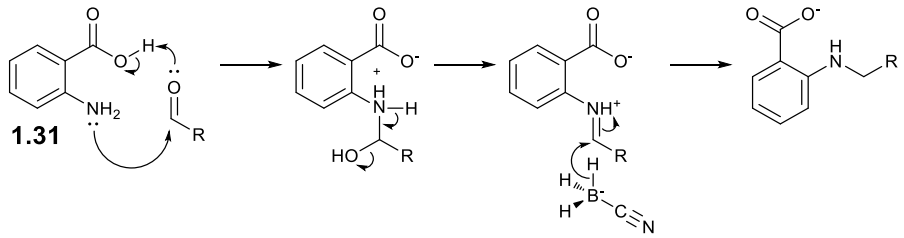
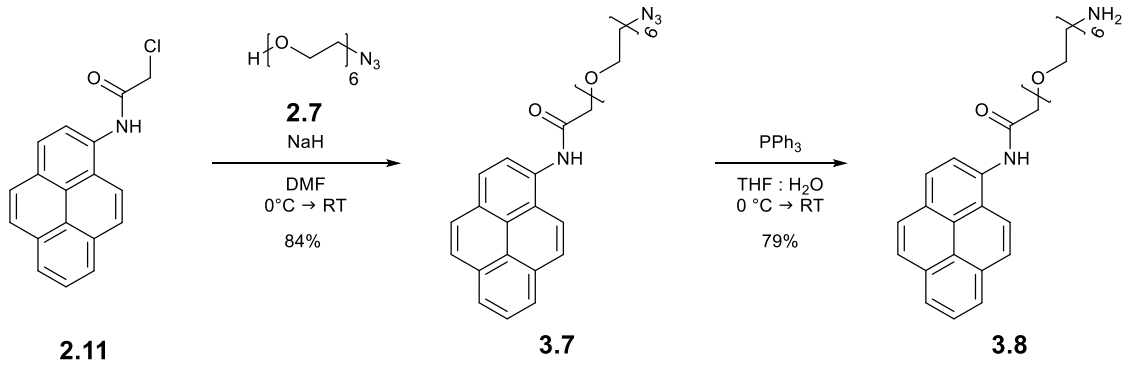


*Figure 3.1 :Proposed mechanism for the reductive amination of glycans with 2-AA **1.31** reported in Anumula's review on glycan labelling reagents.*[227]

Amine **3.8** was synthesized in two steps from chloride **2.11** using conditions adapted from literature procedures (*Scheme 3.4*).[201] Chloroacetamide **2.11** was conjugated to alcohol **2.7** under basic conditions via a Williamson ether synthesis to afford alcohol **3.7** in 84% yield.[201] Azide **3.7** was then reduced via a Staudinger reaction to afford amine **3.8** in 79% yield.[198]

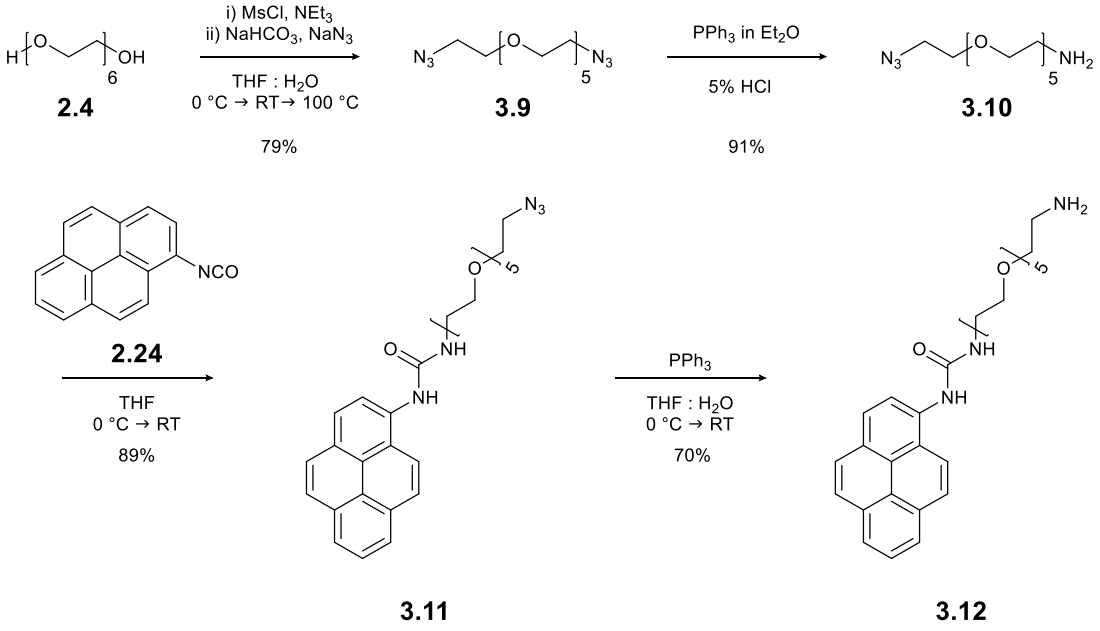


*Scheme 3.4: Synthetic route to 20-amino-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.8**.*[198,201]

Amine **3.12** was synthesized in four steps from hexaethylene glycol **2.4** (*Scheme 3.5*).[198,215,228] Following a procedure reported by Williams et al., hexaethylene glycol **2.4** was activated using mesyl chloride and triethylamine to give a dimesylated intermediate. Substitution of the mesyl groups was achieved by refluxing the dimesylated intermediate under basic aqueous conditions with sodium azide to afford diazide **3.9** in 79% yield.[228]

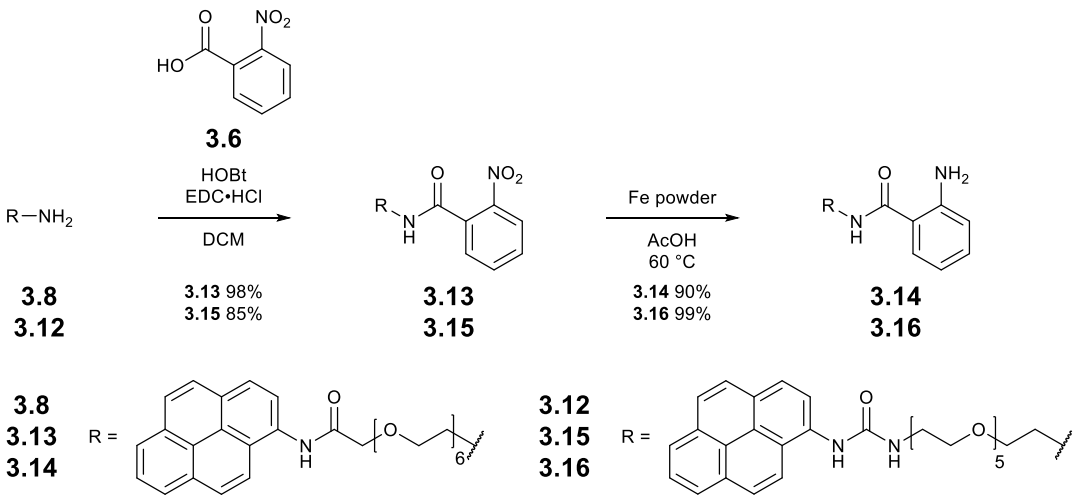
A Staudinger reaction, involving the dropwise addition of an ethereal solution of PPh_3 into a vigorously stirred suspension of diazide **3.9** in 5% $HCl_{(aq)}$, was used to give amine **3.10** in 91% yield.[228] It was observed that diazide **3.9** was insoluble in the acidic aqueous phase but would dissolve readily an organic solution. However, once a reduction had occurred, the mono amino product **3.10** could be readily protonated, allowing it to dissolve favorably in the acidic aqueous phase. Therefore, by utilizing phase separation, <1 equivalent of PPh_3, and the different solubilities of the reactants and products it was possible to selectively carry out a single Staudinger reaction with diazide **3.9** to exclusively obtain the mono reduced product **3.10**.

Using conditions adapted from a literature procedure, amine **3.10** was then conjugated to isocyanatopyrene **2.24** via a urea formation to give azide **3.11** in 89% yield.[215] Finally, azide **3.11** was reduced via a Staudinger reaction to afford amine **3.12** in 70% yield.[198]



*Scheme 3.5: Synthetic route to 3-(17-amino-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **3.12**.*[198,215,228]

1-amidopyrene-labelled 2-AB LALDI tag (AP-OEG-2AB) **3.14** and pyrene urea-labelled 2-AB LALDI tag (PU-OEG-2AB) **3.16** were synthesized in two steps from amines **3.8** and **3.12** respectively, using conditions adapted from various literature procedures (*Scheme 3.6*).[199,224]



*Scheme 3.6: Synthetic route to 20-[(2-aminophenyl)formamido]-N-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.14** and 2-amino-N-(17-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15-pentaoxaheptadecan-1-yl)benzamide **3.16**.*[198,199,229]

Amines **3.8** and **3.12** were conjugated to 2-nitrobenzoic acid **3.6** via an amide coupling reaction using EDC•HCl and HOBt to generate nitro compound **3.13** in 98% yield and **3.15** in 85% yield.[199]

Following a procedure by Locke et al., reduction of the nitro groups of **3.13** and **3.15** was initially attempted through Pd-catalyzed hydrogenation.[224] However, following purification by flash column chromatography, NMR analysis showed that both **3.14** and **3.16** were still contaminated by a small impurity. By comparing the chemical shifts of the impurities to those of **3.13** and **3.15** it was possible to confirm that the impurities were not starting material. Subjecting the impure samples of **3.14** and **3.16** to the hydrogenation conditions resulted in increased formation of the impurity. This indicated that the impurity was formed during the reaction and was not due to unreacted starting material or an impurity present in the starting material. Multiple rounds of purification were performed using both forward and reversed-phase chromatography and reversed-phase HPLC in an attempt to remove the impurity. However, it appeared that isolation of the product, when formed under these reaction conditions, was not possible due to the identical retention times of the product and impurity formed during the reaction.

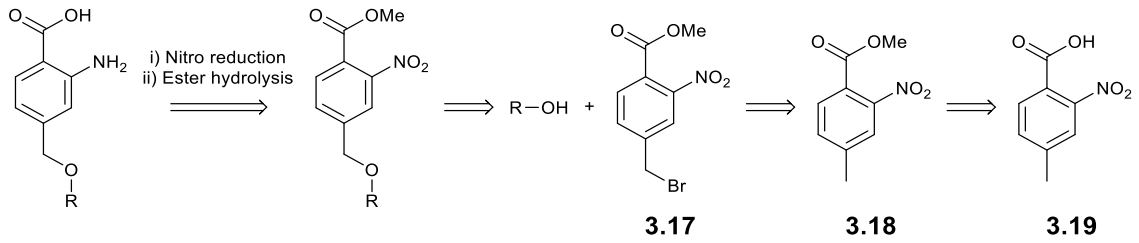
An alternative reduction of the nitro group on **3.13** and **3.15** was carried out using Fe powder in acetic acid (AcOH) at 60 °C (*Scheme 3.6*).[229] Reduction of **3.13** and **3.15** using Fe/AcOH proved to be much more effective than the Pd-catalyzed hydrogenation, giving higher yields in less time,

without any side product formation or need for chromatographic purification. Under these conditions AP-OEG-2AB **3.14** and PU-OEG-2AB **3.16** were obtained in 90% and 96% yield, respectively.

3.2.2.2 Towards 2-aminobenzoic acid reactive handle

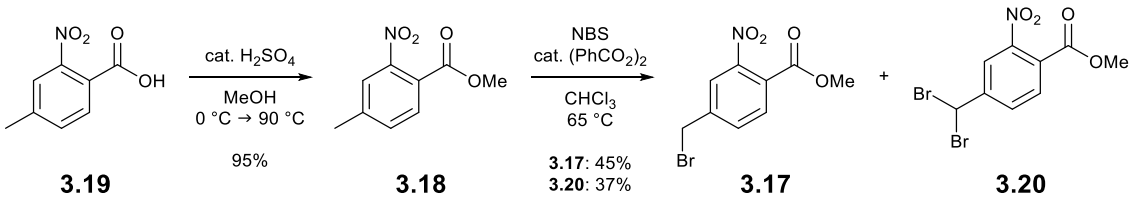
Following reports in the literature that 2-AA was more efficient at glycan derivatization than 2-AB,[227] attempts were also made to develop a LALDI tag functionalized with a 2-AA reactive handle.

In the mechanism reported in the literature for the reductive amination of glycans with 2-AA **1.31** (*Figure 3.1*), it can be seen that both the nitrogen lone pair and carboxylic proton are important for the glycan labelling efficiency of 2-AA.[227] Therefore, to avoid significantly altering the electronics of the aromatic ring, the reactive substituents, and the glycan labelling efficiency of the 2-AA reactive handle, the proposed design for the 2-AA-functionalized LALDI tag involved conjugation of the LDI enhancer to the benzene ring through a methylene bridge (*Scheme 3.7*). Conjugation of electron lone pair from an electron donating group such as oxygen or nitrogen to the aromatic ring would decrease the electrophilicity of the benzoic acid group and raise the pKa of the acidic proton. This could reduce the rate of glycan labelling by reducing the rate of imine formation. Conversely, the use of an electron withdrawing acyl linkage could also hinder the glycan labelling reaction by drawing electron density away from the aniline lone pair and reducing its nucleophilicity. While alkyl substituents are slightly electron donating through induction, the electronic donating effect would be lower than for an *N*- or *O*- substituent and would affect the glycan labelling efficiency of the 2-AA reactive handle less.



Scheme 3.7: Retrosynthetic analysis of a 2-AA-functionalized LALDI tag.

Methyl 5-bromomethyl-2-nitrobenzoate **3.17** was synthesized in two steps from 5-methyl-2-nitrobenzoic acid **3.19** following literature procedures (*Scheme 3.8*).[230,231]



*Scheme 3.8: Synthetic route to methyl 5-bromomethyl-2-nitrobenzoate **3.17**.*[230,231]

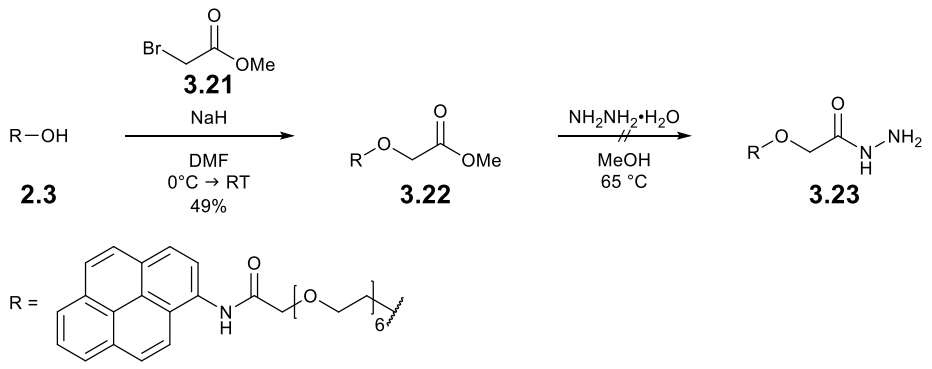
Protection of carboxylic acid **3.19** was achieved by refluxing **3.19** in methanol with a catalytic amount of sulfuric acid to give methyl ester **3.18** in 95% yield.[230] Benzylic bromination of **3.18** was then achieved by refluxing with *N*-bromosuccinimide with a catalytic amount of benzoyl peroxide acting as a radical initiator, affording bromide **17** in 43% yield.[231]

Bromination of **3.18** also resulted in the formation of an unwanted dibromo side product **3.20**, reducing the yield of the desired product **3.17** Formation of dibromide **3.20** was confirmed by comparison of experimentally obtained ^{1}H NMR to examples reported in the literature.[232]

It became clear that, while possible, the remaining three steps of the proposed synthesis (conjugation of **3.17** to an LDI enhancer, ester hydrolysis, and nitro reduction) could require extensive optimization to find suitable reaction conditions and synthetic order. Therefore, as the synthesis of 2-AB-functionalized LALDI tags had been achieved, in view of time constraints no further work was carried out on the synthesis of a 2-AA functionalized LALDI tag.

3.2.3 Synthesis of LALDI tags for labelling by glycans by hydrazone ligation

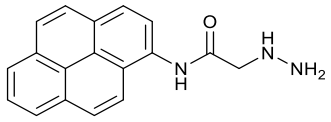
Following the standard procedure for hydrazide formation, originally reported by Yale et al.,[233] the synthesis of 1-amidopyrene-labelled hydrazide LALDI tag (AP-OEG-Hz) **3.23** was initially attempted through the hydrazinolysis of an ester (*Scheme 3.9*).[201,234]



*Scheme 3.9: Attempted synthesis of 1-(hydrazinecarbonyl)-N-(pyren-1-yl)-2,5,8,11,14,17,20-heptaoxadocosan-22-amide **3.23** by hydrazinolysis of methyl ester **3.22**.*[201,234]

Under basic conditions, alcohol **2.3** was conjugated to bromide **3.21** via a Williamson ether synthesis to give methyl ester **3.22** in 49% yield.[201]

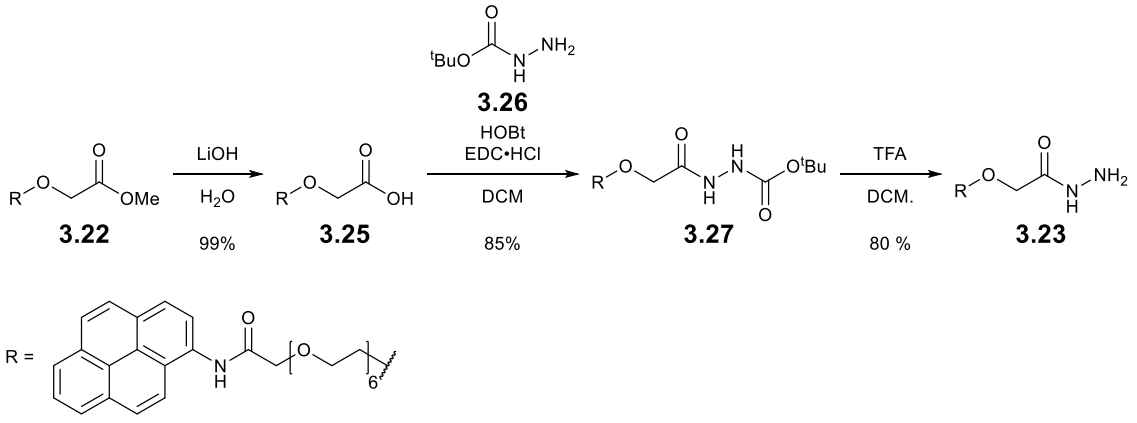
Methyl ester **3.22** was then heated at reflux with hydrazine hydrate in methanol.[234] However, the reaction failed to yield any AP-OEG-Hz **3.23**. NMR analysis of the major product of the reaction showed that the OEG chain had been cleaved from methyl ester **3.22**, but the methylene bridge α to the amide carbonyl was still present. A structure for the isolated product **3.24** was proposed using a combination of experimentally obtained ^{1}H-NMR, LC-MS data (*Figure 3.2*) The proposed structure for **3.24** was supported further by comparison of experimentally obtained data to examples reported in the literature.[200] It is proposed that **3.24** was formed by nucleophilic substitution of the OEG chain of **3.24** by hydrazine. While substitutions of *N*-aryl halido acetamides, such as **2.11**, by hydrazine have been reported in the literature,[200] strongly acid or basic conditions are usually required to cleave ethers.[235] It is proposed that substitution could have been driven by the large excess of hydrazine and heat applied to the reaction.



3.24

*Figure 3.2: Proposed structure for hydrazine **3.24**.*

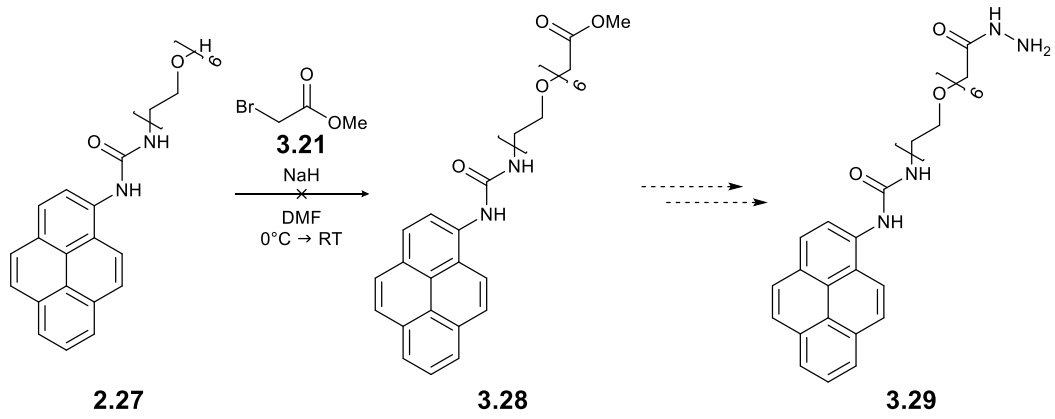
Synthesis of AP-OEG-Hz **3.23** was achieved via an alternative synthesis involving the formation of a *N*-substituted tert-butyl carbazate followed by *tert*-butoxycarbonyl (Boc) deprotection (*Scheme 3.10*).[236]



*Scheme 3.10: Synthesis of 1-(hydrazinecarbonyl)-N-(pyren-1-yl)-2,5,8,11,14,17,20-heptaoxadocosan-22-amide **3.23** through the formation and subsequent Boc deprotection of carbazate **3.27**.*[236]

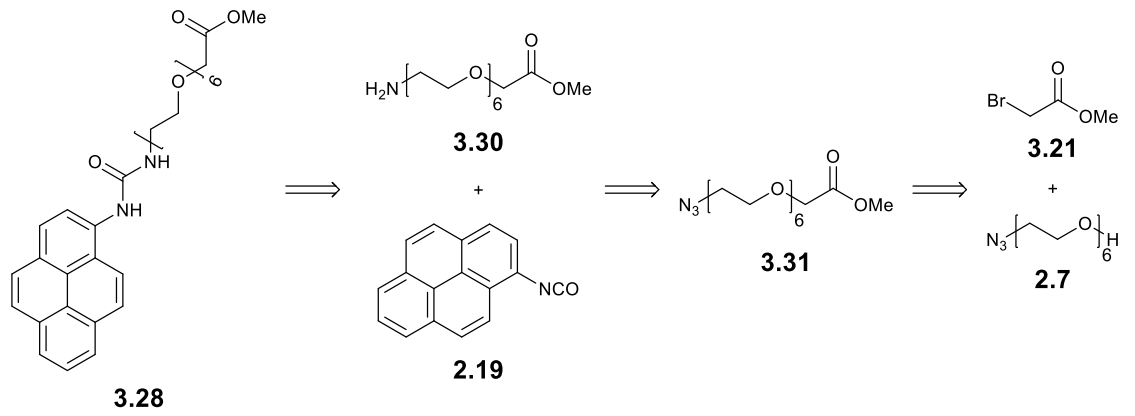
Methyl ester **3.22** was first hydrolyzed using LiOH to afford acid **3.25** in 99% yield. Acid **3.25** was then conjugated to tert-butyl carbazate **3.26** via an amide coupling reaction using EDC and HOBt to generate carbazate **3.27** in 85% yield.[236] Finally, trifluoracetic acid (TFA) was used to cleave the Boc group from carbazate **3.27** to give AP-OEG-Hz **3.23** in 80% yield.[236]

Synthesis of pyrene urea-labelled hydrazide LALDI tag (PU-OEG-Hz) **3.29** was initially attempted via the same synthetic route (*Scheme 3.11*). However, under the conditions described, attempts to form methyl ester **3.28** were unsuccessful. TLC and LCMS analysis of the reaction mixture suggested the reaction had stalled. Additional equivalents of both NaH and bromide **3.21** were added in an attempt to drive the reaction towards completion. Although, this also failed to result in the formation of multiple unidentifiable side products.



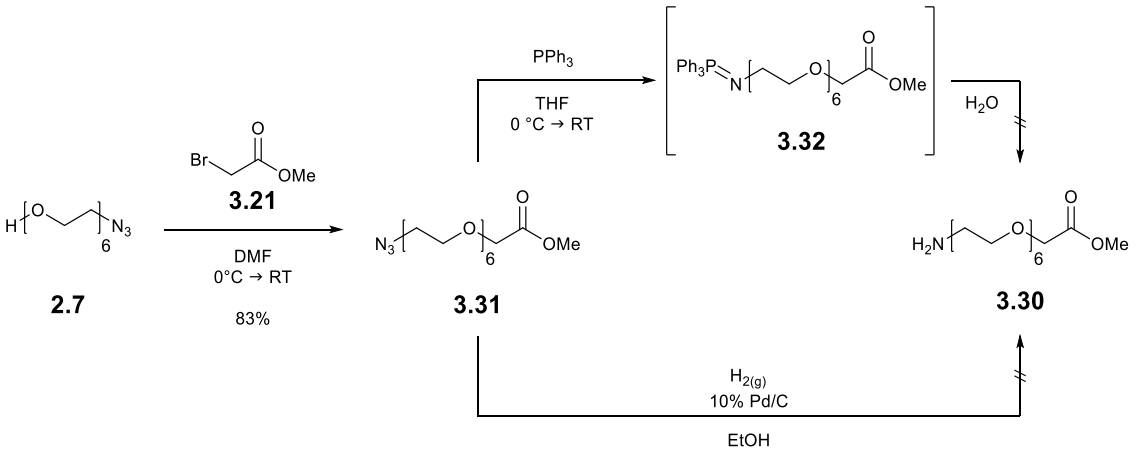
*Scheme 3.11: Attempted synthesis of methyl 20-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15,18-hexaoxaicosanoate **3.29**.*[201]

It was proposed that a urea formation reaction between isocyanatopyrene **2.24** and amine **3.30** could be used as an alternative way of synthesizing methyl ester **3.28** (*Scheme 3.12*).



*Scheme 3.12: Alternative retrosynthetic analysis of 3-[1-(hydrazinecarbonyl)-2,5,8,11,14,17-hexaoxanonadecan-19-yl]-1-(pyren-1-yl)urea **3.28** via urea formation reaction.*

Synthesis of amine **3.30** was attempted using conditions adapted from literature procedures (*Scheme 3.13*).[198,201,237]



*Scheme 3.13: A scheme detailing the reaction conditions for the synthesis of methyl 20-azido-3,6,9,12,15,18-hexaoxaicosanoate **3.31** and attempted reactions conditions for the synthesis of 20-amino-3,6,9,12,15,18-hexaoxaicosanoate **3.30**.*[198,201,237]

Azide **2.7** was conjugated to bromide **3.21** under basic conditions via a Williamson ether synthesis to give methyl ester **3.31** in 83% yield.[201]

Reduction of azide **3.31** was initially attempted through a Staudinger reaction.[198] LCMS analysis of the reaction indicated that total conversion of azide **3.31** to the iminophosphorane intermediate **3.32** had occurred. Hydrolysis of an iminophosphorane to the desired amine is normally achieved by the addition of water to the reaction mixture and allowing it to stir for

1-2 hours. After this time had elapsed, very little of the amine **3.30** was observed. The reaction was allowed to stir for an additional 24 hours and monitored by LC-MS. LC-MS analysis showed that instead of hydrolyzing the iminophosphorane of **3.32** hydrolysis of the methyl ester was occurring instead.

Following a literature procedure by Winther et al.,[237] reduction of azide **3.31** was attempted via Pd-catalyzed hydrogenation. Under hydrogenation conditions formation of amine **3.30** was observed. However, analysis of the crude product by MS and NMR indicated that a number of impurities with m/z higher than **3.30** had formed, suggesting that the Pd catalyst could have catalyzed a number of side reactions. Efforts to purify the crude sample using a combination of forward and reversed-phase chromatography were unable to isolate amine **3.30** from the impurities.

After encountering problems with impurities generated during Pd-catalyzed hydrogenation in both the nitro reduction of **3.13** and **3.15**, and the azide reduction of **3.31**, it is suspected that the cause of this problem could be related to contamination or degradation of the Pd/C catalyst used in the reaction.

In the interest of moving forward with the project, it was decided that having AP-OEG-Hz **3.23** would be sufficient to assess whether a hydrazide-functionalized LALDI tag is able to label a glycan and then promote its detection by LALDI-MS. Therefore, no further attempts were made to synthesize PU-OEG-Hz **3.29**. If future work requires the synthesis of PU-OEG-Hz **3.29**, the synthetic route proposed in *Scheme 3.12* could still be possible if repeated using fresh catalyst.

3.3 LALDI-MS analysis of unreacted LALDI tags

LALDI-MS analysis in *Chapter 2* focused on the MS performance and stability of the LDI enhancers conjugated to the solubilizing OEG linkers. In this study, samples of LALDI tags **3.5**, **3.16**, and **3.23** were analyzed by LALDI-MS to see how the addition of the reactive handles would affect their sensitivity and stability in the MS. Analysis of the unreacted label would also serve as a reference for the analysis of in situ labelling reactions, allowing distinctions to be made between signals relating to the unreacted label and other signals in the spectra. Due to the loss of CO observed during LALDI-MS analysis of 1-amidopyrene species (Section 2.5), where possible, LALDI tags featuring pyrene urea LDI enhancers were chosen for this investigation to simplify the analysis.

Samples were prepared by dissolving the LALDI tags in methanol, spotting a known amount of each compound onto a target plate, then allowing the solvent to evaporate. LALDI-MS analysis was performed on the neat samples without any additional chemical modifications, MALDI matrices, or SALDI surface materials with at least three individual repeats measured for each sample.

The LALDI-MS analysis of each LALDI tag reagent was assessed on both stability and sensitivity in the mass spectrometer. LALDI-MS spectra obtained from the analysis of the LALDI tag-labelled glycans are presented in *Figure 3.3.* A summary of the LALDI-MS analysis of **3.5**, **3.16** and **3.23** can be found in *Table 3.1*.

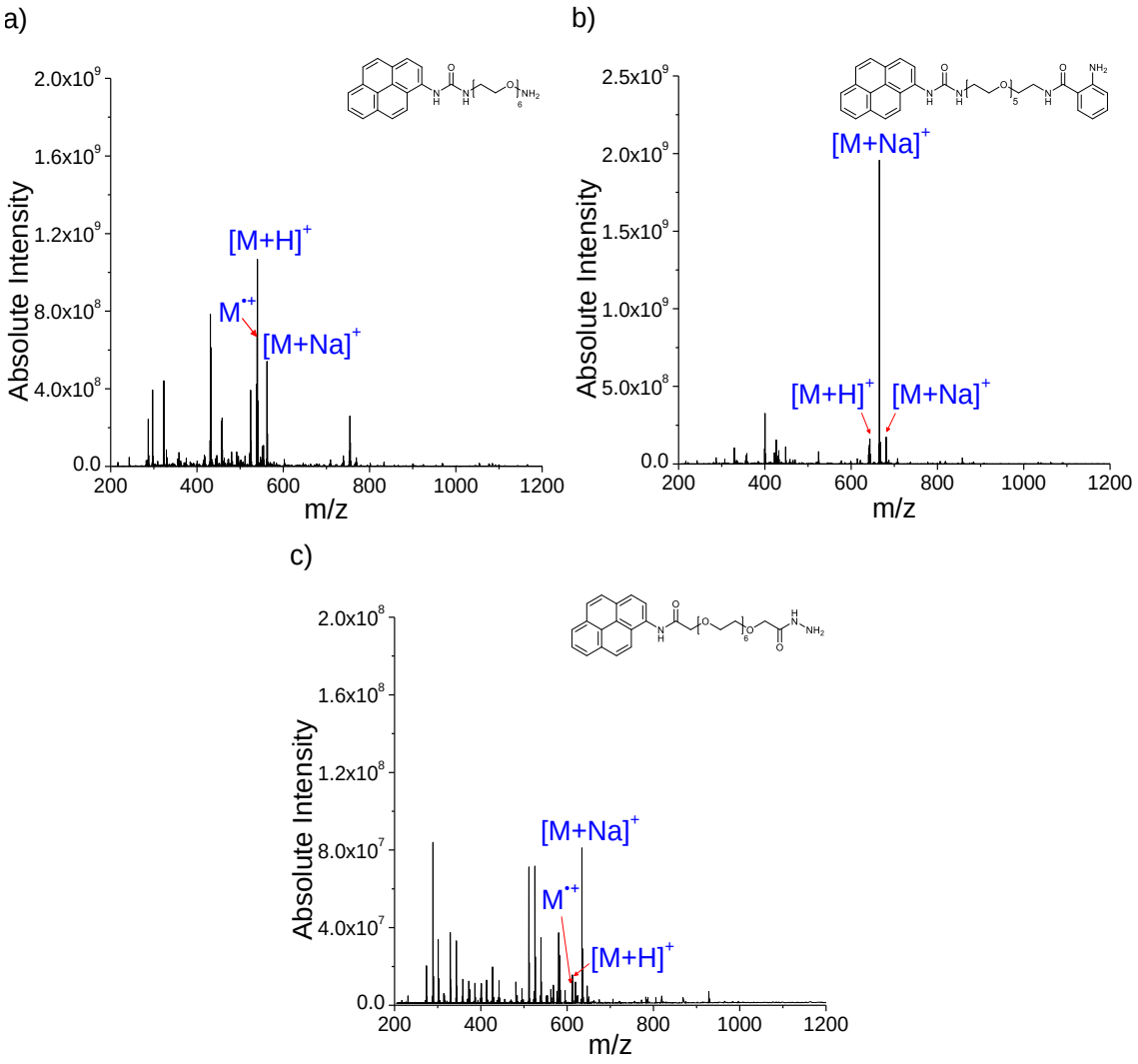


*Figure 3.3: LALDI-MS spectrum for 100 pmol samples of LALDI tags a) **3.35**, b) **3.16**, c) **3.23** with the corresponding signal for the $M^{\bullet+}$, protonated ion, and metal adduct ions highlighted.*

To understand how the reactive handles affected the stability of the LALDI tag reagents in the mass spectrometer, detailed analysis of the fragmentation patterns was carried out. An example of the fragmentation pattern analysis is shown in *Figure 3.4*. Analogous data for LALDI tags **3.16** and **3.23** can be found in *Appendix II.*

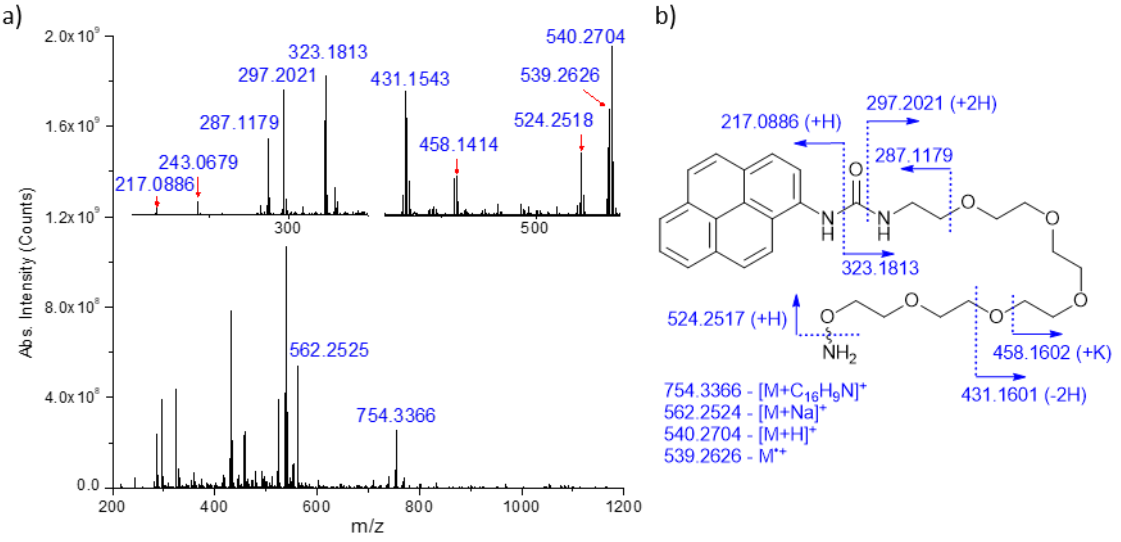


*Figure 3.4: LALDI-MS analysis of PU-OEG-ONH$_2$ **3.5**. a) LALDI-MS spectra for PU-OEG-ONH$_2$ **3.5** (100 pmol) with identified fragments and adducts annotated with their observed m/z. Inlaid graphs show expansion of the full spectra at m/z 200-350 and m/z 520-545. b) Proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

LALDI-MS analysis of PU-OEG-ONH$_2$ **3.5** showed clear detection of the molecular ion as the radical cation ($M^{\bullet+}$), along with the $[M+H]^+$, $[M+Na]^+$, and $[M+K]^+$ adduct ions (*Figure 3.4a*). The base ion peak in the spectrum (*m/z* 540.2704) was identified as $[M+H]^+$.

Several signals with *m/z* lower than **3.5** were attributed to fragmentation of **3.5** during LALDI-MS analysis. Fragmentation of **3.5** was mostly observed around the urea linkage, with other cleavages occurring along the OEG chain and at aminooxy group (*Figure 3.4b)*. These data indicated that glycans labelled by **3.3** and **3.5** could likely experience fragmentation at their oxime linkages. While most of the monoisotopic masses for the proposed fragments correlated well with the observed *m/z*, with errors <±0.001, the two signals at *m/z* 431.1543 and 458.1414 could not be accurately assigned to any obvious fragments, therefore the closest possible approximations were given.

An unexpected signal with a *m/z* higher than the molecular ion was observed at *m/z* 754.3361. Analysis of the sample by HR-ESI-MS, NMR, and analytical HPLC indicated that there were no

impurities present in the sample. During LALDI-MS analysis of a 10 pmol sample of **3.5**, no signal at *m/z* 754.3361 was detected, suggesting that the formation of this adduct may be concentration dependent. Using high-resolution mass spectrometry, it was possible to determine that the adduct was *m/z* 215.0734 greater than the parent ion. This value corresponded exactly to the monoisotopic mass of a species with the molecular formula $C_{16}H_9N$. The second most abundant ions in the spectrum (*m/z* 323.1813), corresponds exactly to a fragment generated through loss of $C_{16}H_{10}N$. Additionally, a minor fragment can be seen at *m/z* 217.0886 correlating a fragment with the molecular formula $C_{16}H_{11}N$. The presence of these signals indicates that fragmentation and loss of $C_{16}H_9N$ can occur during LALDI-MS analysis of **3.5**. In the investigation by Addy et al. on using pyrene boronic acid to achieve LALDI-MS analysis of cis-1,2-diols, it was reported that singly charged dimeric and trimeric species were generated through self-condensation of the PBA analytes during LALDI-MS analysis.[102] While a more in-depth investigation would be required to confidently explain the origin of [**3.5**+215]$^+$, one explanation could be that [**3.5**+215]$^+$ is the product of fragment dimerization in the mass spectrometer.

3.3.1 Summary of LALDI-MS analysis of unreacted LALDI tags

In line with the results from Chapter 2, the LALDI-MS analysis showed that all three LALDI tag reagents **3.5**, **3.16** and **3.23** could achieve matrix-free ionization and subsequent MS detection following excitation of the sample in the LDI mass spectrometer (*Figure 3.3*). The LALDI tag reagents did experience more degradation compared to the simple LALDI reagents comprising just the LDI enhancer and the solubilizing linker. However, all major peaks and the majority of other signals observed in the LALDI-MS spectra could be easily assigned with the monoisotopic masses for the proposed fragments being within ±0.001 of the observed *m/z*.

2-AB-functionalized LALDI tag **3.16** (*Appendix II– Figure A.5*) was found to be the most stable LALDI tag reagent, giving a stronger molecular ion signal and less degradation than hydroxylamine **3.5** (*Figure 3.4*) and Hz-functionalized LALDI tag **3.23** (*Appendix II – Figure A.6*) was found to be the least stable LALDI tag reagent, giving a high degree of fragmentation, with several unidentifiable signals present in the mass spectrum. The high degree of degradation and lower sensitivity observed during LALDI-MS analysis of **3.23** suggested that hydrazide-functionalized LALDI tags may not be suitable for analyzing glycans by LALDI-MS, especially if the observed instability is carried over to the conjugate product.

*Table 3.1: Summary of LALDI-MS analysis of LALDI tag reagents **3.5**, **3.16** and **3.23**. MS stability was assessed on the relative abundance of the molecular ion ($M^{\bullet+}$) and metal adduct ions ($[M+Na]^+$, $[M+K]^+$) compared to the fragments, degree of fragmentation, and ease of fragment assignment then scored using an arbitrary scale from **++** to **– –**.*

LALDI tag reagent	Detected as molecular ion	MS stability	Limit of detection (pmol)	Analysis location
PU-OEG-ONH_2 **3.5**	Yes	+/–	10	Section 3.3 – *Figure 3.4*
PU-OEG-2AB **3.16**	Yes	+ +	10	*Appendix II* – Figure A.5
AP-OEG-Hz **3.23**	Yes	–	100	*Appendix II – Figure A.6*

3.4 Optimization of glycan labelling and LALDI-MS analysis.

To label and detect glycans from a complex biological system, both the glycan labelling and LALDI-MS analysis first needed to be established in a controlled environment. Therefore, the purpose of these experiments was to determine whether LALDI tags could be used to label simple glycans using established glycan labelling reactions cited in the literature, and then promote their detection by LALDI-MS. These experiments would also provide information on the labelling efficiencies of each LALDI tag and the MS-detection and stability of their conjugate products.

The initial glycan labelling experiments were primarily performed using lactose. Lactose was considered an ideal proof-of-concept model for investigating the labelling and subsequent detection of carbohydrates by LALDI-MS. Lactose is a cheap and readily available reducing disaccharide, but can often be difficult to detect by MALDI-MS without further chemical modification,[238] additives,[239,240] or specialized nanoparticle supports.[241–243] The low cost and availability of lactose would enable the labelling experiments to be carried out on a scale which permits isolation of the labelled products, while the poor ionization of underivatized lactose would allow for a clear demonstration of the ionization-promoting capabilities of the LALDI tags.

In addition to lactose, labelling and analysis of *N*-acetyl glucosamine (GlcNAc) was also performed to investigate whether the LALDI tag reagents would be able to label and analyze simple *N*-acetylated glycans.

3.4.1 Glycan labelling reactions

A sample of lactose was labelled with each of the five synthesized LALDI tags **3.3**, **3.5**, **3.14**, **3.16**, and **3.16** using conditions taken from established glycan labelling procedures cited within the literature. Isolation of the labelled lactose was accomplished by preparative HPLC. Purity of the isolated products was assessed using analytical HPLC, and the identity of each labelled product was confirmed by HR-MS, under the assumption that each glycan labelling reaction occurred through the mechanisms described in the literature.[171]

The conditions and results of each labelling experiment are summarized in *Table 3.2*.

3.4.1.1 Oxime ligation

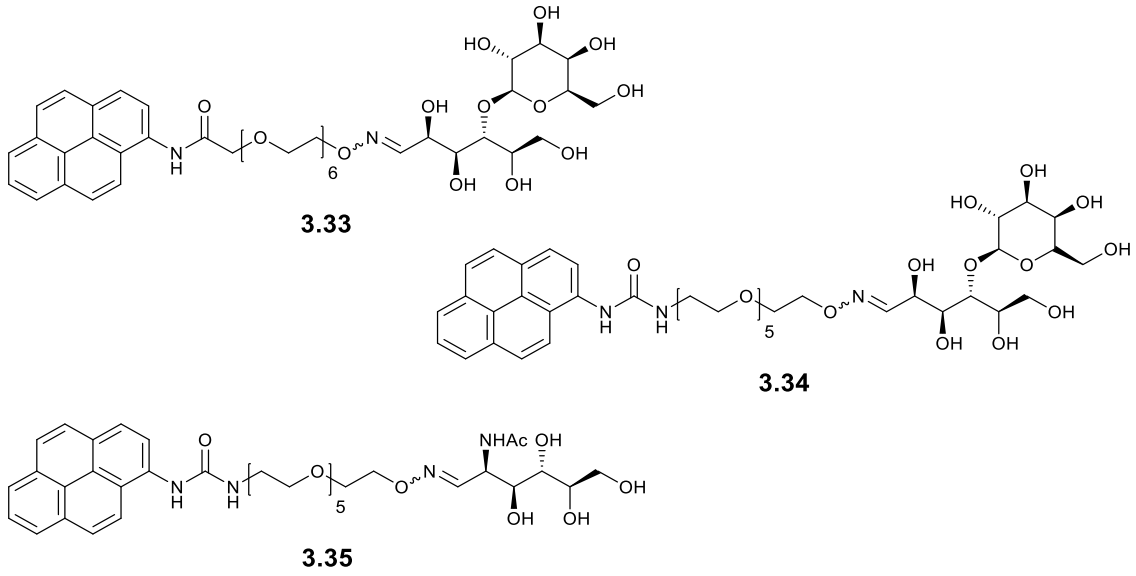
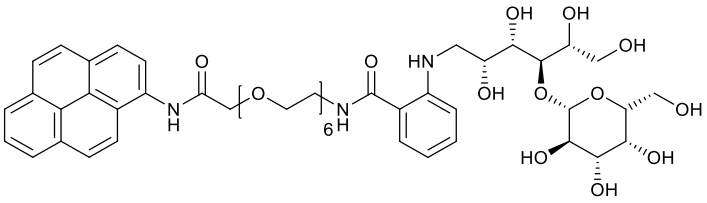


*Figure 3.5: Proposed chemical structures for AP-OEG-ONH$_2$-labelled lactose **3.33**, PU-OEG-ONH$_2$-labelled lactose **3.34**, PU-OEG-ONH$_2$-labelled GlcNAc **3.35** generated via the oxime ligation reactions described in Table 3.2.*[158]

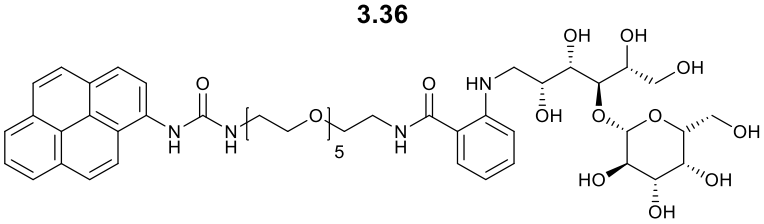
Labelling of lactose and GlcNAc with hydroxylamine-functionalized LALDI tags **3.3** and **3.5** was achieved via a 3,5-diaminobenzoic acid catalyzed oxime ligation reported by Crisalli et al.[158]

LC-MS analysis of the reactions after 24 h indicated that no side reactions had occurred and that all of the labelling reagent had been consumed during the reaction. Following purification by preparative HPLC hydroxylamine labelled glycans **3.33**, **3.34**, and **3.35** were obtained in 19%, 5%, and 34% respectively. The low isolated yield for **3.35** and **3.36** was due to troubleshooting issues during HPLC purification.

3.4.1.2 Reductive amination



3.36

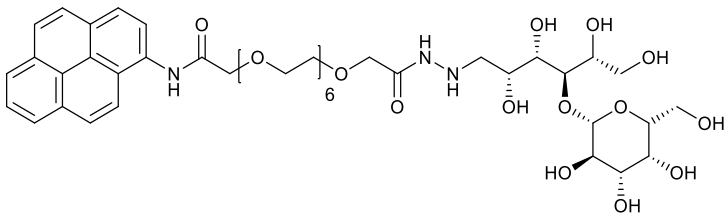


3.37

*Figure 3.6: Proposed chemical structures for AP-OEG-2AB-labelled lactose **3.36** and PU-OEG-2AB-labelled lactose **3.37** generated via the reductive amination labelling reactions described in Table 3.2.*[175]

Labelling of lactose with 2-AB-functionalized LALDI tags **3.14** and **3.16** was achieved via the optimized reductive amination conditions for 2-AB reported by Bigge et al.[175] The reaction was stopped after 2 h to prevent degradation of the labelled product. LC-MS analysis of the reaction mixtures indicated that no side reactions had occurred, and that 89-92% of labelling reagent had been consumed during the reaction. Purification of the reactions by preparative HPLC afforded the labelled 2-AB labelled glycans **3.36** and **3.37** in 34% and 68% yield respectively.

3.4.1.3 Reductive hydrazination



3.38

*Figure 3.7: Proposed chemical structures for AP-OEG-Hz labelled lactose **3.38** generated via the reductive hydrazination labelling reactions described in Table 3.2.*[244]

Conjugation of hydrazide probes to reducing glycans is typically achieved through hydrazide ligation via a mechanism analogous to oxime ligation. However, hydrazones are much less stable than their oxime counterparts.[245] Therefore, to reduce the chance of hydrolysis during purification and fragmentation during analysis, glycan labelling was pursued via reductive hydrazination. Glycan labelling with AP-OEG-Hz **3.23** was performed following a procedure

reported by Grün et al. which cited the reductive amination conditions optimized by Bigge et al. as a successful technique for reduction of the hydrazone conjugate product to a more stable hydrazide.[175,244] Grün et al. demonstrated that glycans derivatized through this procedure were stable during MALDI-MS analysis, predominantly appearing as a single peak in the mass spectra.[244]

Under the attempted reaction conditions, LC-MS analysis indicated that the labelling reaction had resulted in a complicated mixture of product, starting material, and side products. Purification of the reaction was attempted via preparative HPLC. However, absolute isolation of the hydrazide labelled lactose **3.38** from the reaction mixture could not be achieved. Analytical HPLC analysis of the isolated sample indicated that the product had co-eluted with an impurity. However, the identity of the impurity was unclear as it could not be detected by ESI-MS. Due to the limited quantity of **3.38**, no further attempts at purification were pursued and LALDI-MS analysis was carried out on the impure sample.

Table 3.2: Glycan labelling conditions applied LALDI tag reagents. † Conversion calculated by comparing integrations of UV absorbance peaks for the starting material and product in the reaction mixture.

Label	Reaction	Glycan	Reaction conditions	Product	% Label converted†	Isolated yield
3.3	Oxime ligation [158]	Lactose	3,5-diaminobenzoic acid, MeCN : Na citrate buffer (0.1 M, pH 3) (1:1), RT, 24 h.	**3.33**	100%	19%
3.5	Oxime ligation [158]	Lactose	3,5-diaminobenzoic acid, MeCN : Na citrate buffer (0.1 M, pH 3) (1:1), RT, 24 h.	**3.34**	100%	5%
3.5	Oxime ligation [158]	GlcNAc	3,5-diaminobenzoic acid, MeCN : Na citrate buffer (0.1 M, pH 3) (1:1), RT, 24 h.	**3.35**	100%	34%
3.14	Reductive amination [175]	Lactose	$NaBH_3CN$, DMSO : AcOH (7:3), 60 °C, 2 h.	**3.36**	92%	34%
3.16	Reductive amination [175]	Lactose	$NaBH_3CN$, DMSO : AcOH (7:3), 60 °C, 2 h.	**3.37**	89%	68%
3.23	Reductive hydrazination [244]	Lactose	$NaBH_3CN$, DMSO : AcOH (7:3), 60 °C, 2 h.	**3.38**	64%	N/A

3.4.2 LALDI-MS analysis of labelled glycans

Following the glycan labelling reactions (*Table 3.2*), isolated products **3.33**, **3.34**, **3.35**, **3.36**, **3.37**, and **3.38** were analyzed by LALDI-MS. Samples were prepared by dissolving the labelled glycans in methanol, spotting a known amount of each labelled glycan onto the target plate, then allowing the solvent to evaporate. Analysis was performed on the neat samples without any additional chemical modifications, MALDI matrices, or SALDI surface material with at least three individual repeats measured for each sample.

The LALDI-MS analysis of each labelled glycan was assessed on both stability and sensitivity in the mass spectrometer. LALDI-MS spectra obtained from the analysis of the LALDI tag-labelled glycans are presented in *Figure 3.8.* Detailed analysis of the fragmentation pattern observed in each LALDI-MS spectra was carried out to evaluate the LALDI-MS stability of the glycan conjugations. Annotated and assigned LALDI-MS spectra for the labelled glycans can be found in *Appendix III.* A summary of the LALDI-MS analysis from this investigation can be found in *Table 3.3*.

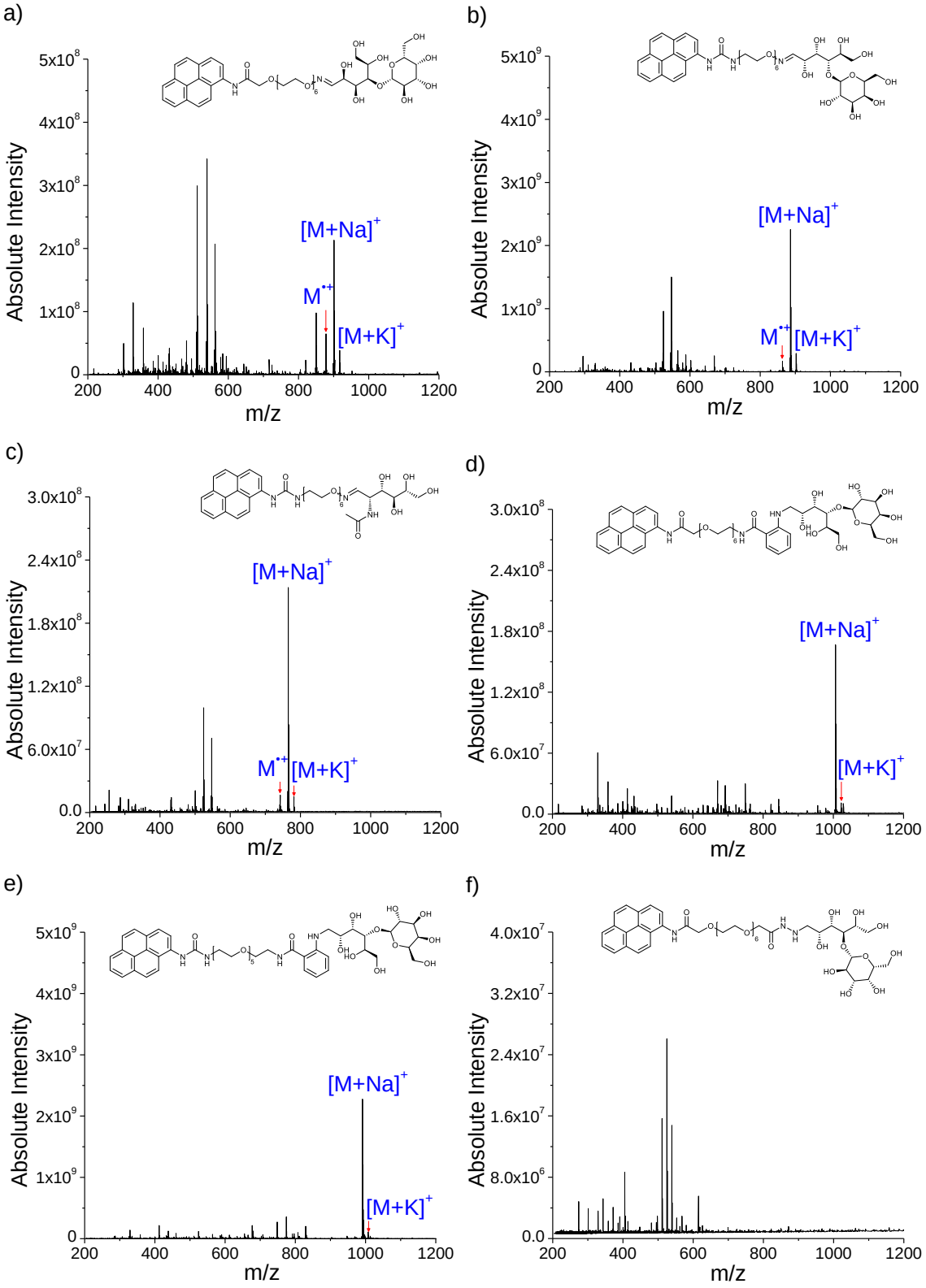


*Figure 3.8: LALDI-MS spectrum for 100 pmol samples of LALDI tag-labelled glycans a) **3.33**, b) **3.34**, c) **3.35**, d) **3.36**, e) **3.37**, and f) **3.38** with the corresponding signal for the $M^{\bullet+}$ and metal adduct ions highlighted.*

Analysis revealed that the hydroxyl amine labelled glycans **3.33**, **3.34**, and **3.35**, and the 2-AB-labelled lactose **3.36** and **3.37** could be clearly detected without the need for an external matrix, additives, or further chemical modification. The majority of signals observed in the LALDI-MS spectra could be easily assigned with the monoisotopic masses for the proposed fragments being within ±0.001 of the observed *m/z*. However, analysis of the hydrazide labelled lactose **3.38** resulted in significant MS-degradation with no peaks observed related to the molecular ion. Additionally, a large number of signals could not be identified in the LALDI-MS analysis of **3.38**. This could be due to the unidentified impurity that was present in the sample during analysis.

All three samples labelled with a 1-amidopyrene LDI enhancer **3.33**, **3.36**, and **3.38** exhibited degradation relating to a loss of CO from both the parent ion and major fragments. All three hydroxylamine labelled lactose products **3.33**, **3.34** and **3.35** displayed cleavage at the oxime linkage following excitation in the spectrometer. Both 2-AB labelled lactose **3.36** and **3.37** exhibited fragmentation around the benzamide and secondary amine, however both samples could be clearly detected as the $[M+Na]^+$ ion.

Due to the complications experienced during synthesis, purification and analysis of **3.38**, it was concluded that hydrazide-functionalized LALDI tags would not be suitable for glycomic analysis. Therefore, **3.38** was not included in any further LALDI-MS or glycan labelling investigations.

LALDI-MS analysis of the labelled glycans using the AnchorChip plate was generally found to have good precision, with the data recorded for labelled glycans **3.34**, **3.35**, **3.36**, **3.37**, and **3.38** having minimal standard deviation of data from the mean (*Figure 3.9*). This data supports previous assertions that the use of a micro-focusing target plate, such as an AnchorChip, could be beneficial for improving the reproducibility of data collected by LALDI-MS (Section 2.3.2). However, the LALDI-MS analysis of labelled lactose **3.33** was observed to have an greater error compared to the other analytes, indicating that additional independent repeats would still be required to ensure confidence in the data obtained using LALDI-MS; especially when investigating more complicated systems, such as those derived from biological samples.

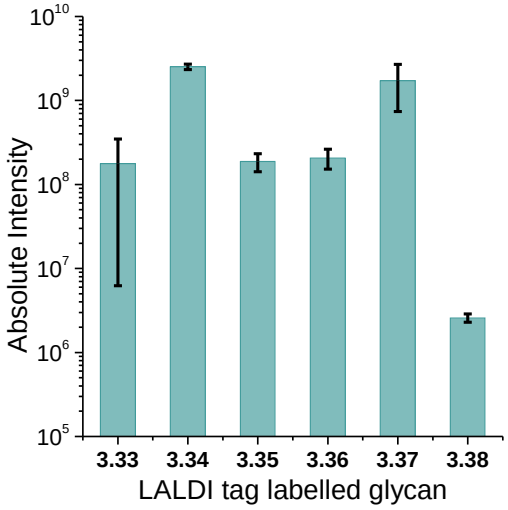


*Figure 3.9: Comparison of the LALDI-MS analysis of each of the LALDI tag labelled glycans **3.33**, **3.34**, **3.35**, **3.36**, **3.37**, and **3.38** analyzed as pure samples (100 pmol) illustrating the difference in sensitivity of each LALDI tag labelled glycan and the precision of the data collected using an AnchorChip target plate. Bars represent the mean sum of all (+) ions related to **3.33** ($[M\text{-}CO]^+$, $M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$), **3.34** ($M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$), **3.35** ($M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$), **3.36** ($[M\text{-}CO]^+$, $[M\text{-}CO+Na]^+$, $[M+Na]^+$, $[M+K]^+$, $[M\text{-}H+2Na]^+$), **3.37** ($M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$), and **3.38** ($[M\text{-}C_{12}H_{23}NO_{10}]^+$) that were detected by LALDI-MS at 100 pmol of each reagent. As no signals were detected for the molecular ion or adducts of **3.38** during LALDI-MS analysis, the reproducibility of the data for **3.38** was assessed using the measured abundance of the fragment ion generated through cleavage at the N-O oxime linkage (m/z 568.2776 – $[M\text{-}C_{12}H_{23}NO_{10}]^+$). Error bars represent the standard deviation of the mean, where n=3 (individual samples on separate target spots).*

*Table 3.3: Summary of LALDI-MS analysis of LALDI tag labelled glycans. MS stability was assessed on the relative abundance of the molecular ion ($M^{\bullet+}$) and metal adduct ions ($[M+Na]^+$, $[M+K]^+$) compared to the fragments, degree of fragmentation, and ease of fragment assignment then scored using an arbitrary scale from **++** to **– –**.*

LALDI tag-labelled glycan	Detected as molecular ion	MS Stability	Limit of detection (pmol)	Analysis location
3.33	Yes	**+/–**	1	*Appendix III – Figure A.7*
3.34	Yes	**+**	10	*Appendix III – Figure A.8*
3.35	Yes	**+**	10	*Appendix III – Figure A.9*
3.36	Yes	**+ +**	10	*Appendix III – Figure A.10*
3.37	Yes	**+ +**	10	*Appendix III – Figure A.11*
3.38	No	**– –**	100	*Appendix III – Figure A.12*

3.4.3 Analysis of labelled glycans from complicated background

Following glycoconjugate cleavage, the free glycans are normally contaminated by a mixture of salts, detergents, and various other biological and cellular material. Hence, before any analysis can be undertaken, rigorous purification of the free glycans from these contaminants is usually required.[246]

As one of the primary aims of LALDI tags is to reduce the amount of work required to achieve MS analysis of glycans by allowing direct detection of labelled glycans from a complex backgrounds without purification. Therefore, to challenge the sensitivity of the labels, analysis of the labelled products **3.33**, **3.34**, **3.35**, **3.36**, and **3.37** was repeated with the purified sample dissolved in Hanks' balanced salt solution (HBSS) then analyzed by LALDI-MS.

HBSS is a cell culture buffer composed of various buffer salts (Na^+, K^+, Mg^{2+}, Ca^{2+}), glucose, and a phenol red indicator. The purpose of this experiment was to provide an insight into whether detection of the LALDI tag-labelled lactose and GlcNAc could be achieved from a complex background similar to those encountered in current glycan derivatization protocols.

Samples were prepared by dissolving the labelled glycans in HBSS, spotting a known amount of each labelled glycan onto the target plate and then removing the solvent under reduced pressure. The abundance of salts and other contaminants present in each sample could be clearly observed following evaporation of the analyte/HBSS mixture (*Figure 3.10a*). Analysis was performed on the samples without any additional chemical modifications, MALDI matrices, or SALDI surface materials with at least three individual repeats measured for each sample.

As in the previous experiment, the LALDI-MS analysis of each labelled glycan was assessed on both their stability and sensitivity in the mass spectrometer. LALDI-MS spectra obtained from the analysis of the labelled glycans from HBSS are presented in *Figure 3.10.* A summary of the LALDI-MS analysis of the LALDI tag-labelled glycans from this investigation can be found in Table 3.4.

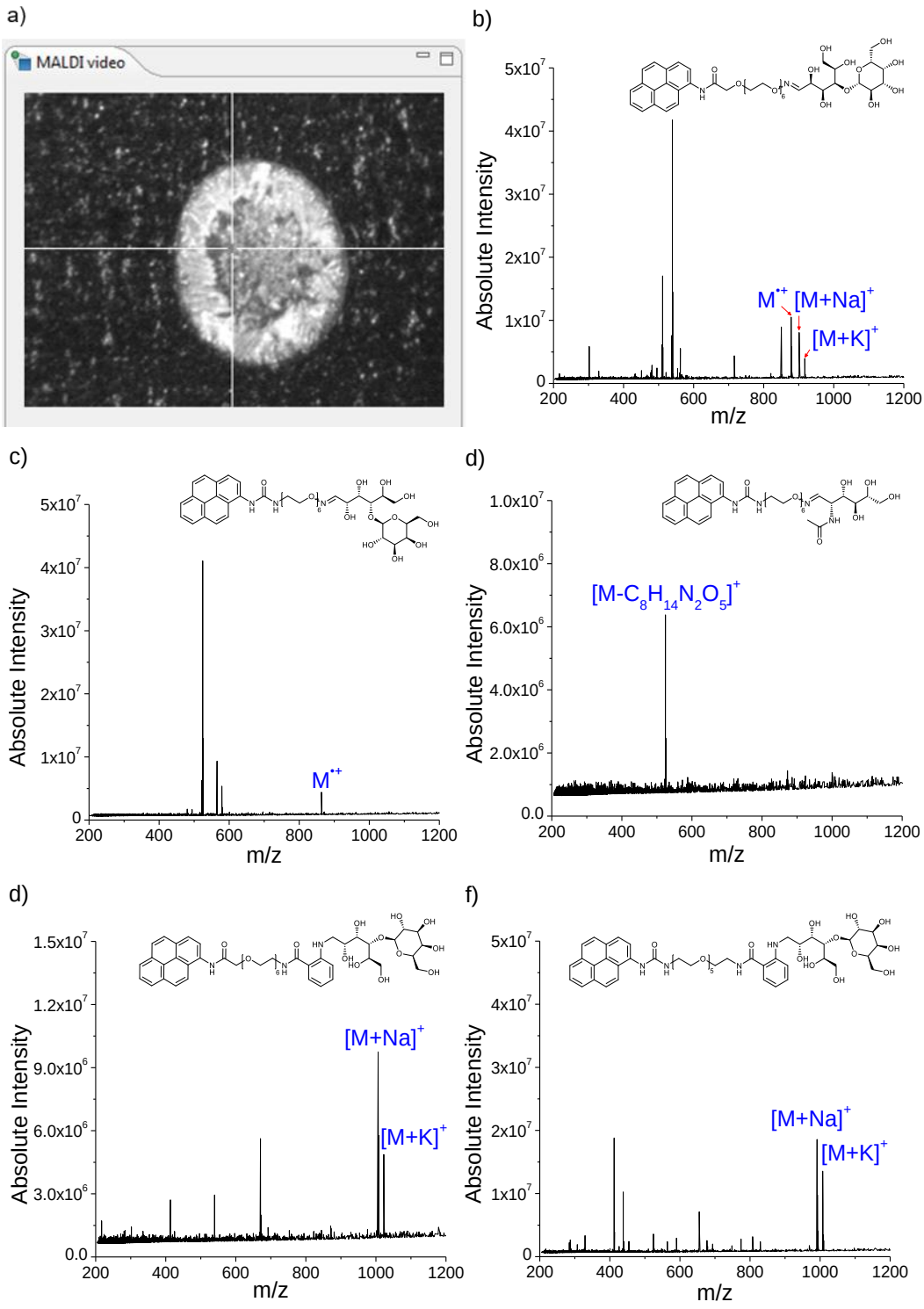


Figure 3.10: LALDI-MS analysis of labelled glycans from HBSS. a) An image showing an evaporated sample of LALDI-tag-labelled glycan dissolved in HBSS captured by the LDI-MS laser target camera. The crystalline appearance of the sample shows the high abundance of buffer salts present in the analytical samples analyzed during these experiments. LALDI-MS spectrum for 100 pmol samples of LALDI tag-labelled glycans b) ***3.33****, c)* ***3.34****, d)* ***3.35****, e)* ***3.36****, and f)* ***3.37*** *from a background of HBSS with the corresponding signal for the $M^{\bullet+}$ and metal adduct ions highlighted. For* ***3.35****, the fragment ion generated by cleavage at the N-O oxime linkage has been highlighted ($[M-C_8H_{14}N_2O_5]^+$).*

While the LALDI-MS sensitivity was affected, each of the LALDI tag-labelled glycans could be clearly detected from HBBS, with recorded spectra for the labelled lactose **3.33**, **3.34**, **3.36**, and **3.37** being comparable to the previous measurements taken on pure samples. No signals relating to the molecular ion or metal ion adducts were detected during the LALDI-MS analysis of labelled GlcNAc **3.35**, which was detected exclusively as a single peak correlating to fragmentation at the oxime linkage (*Figure 3.10d*).

LALDI-MS analysis of the labelled glycans from HBSS using the AnchorChip plate was found to have good precision, with the data recorded for each of the labelled glycans **3.33**, **3.34**, **3.35**, **3.36**, and **3.37** having minimal standard deviation of data from the mean (*Figure 3.11*).

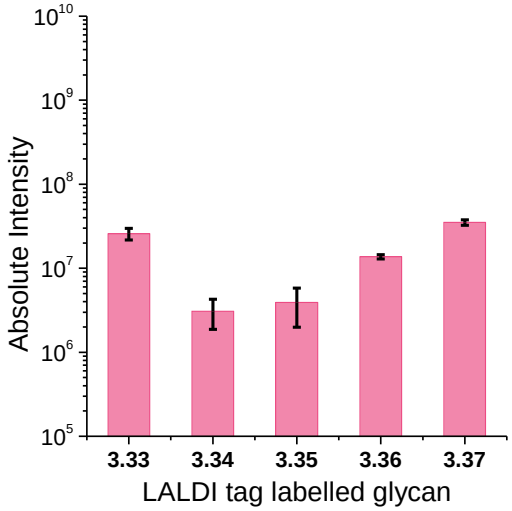


*Figure 3.11: Comparison of the LALDI-MS analysis of each of the LALDI tag labelled glycans **3.33**, **3.34**, **3.35**, **3.36**, and **3.37** analyzed from HBSS (100 pmol) illustrating the difference in sensitivity of each LALDI tag labelled glycan and the precision of the data collected using an AnchorChip target plate. Bars represent the mean sum of all (+) ions related to **3.33** ($[M\text{-}CO]^+$, $M^{\bullet+}$, $[M+Na]^+$, $[M+K]^+$), **3.34** ($M^{\bullet+}$), **3.35** ($[M\text{-}C_8H_{14}N_2O_5]^+$), **3.36** ($[M+Na]^+$, $[M+K]^+$), and **3.37** ($[M+H]^+$, $[M+Na]^+$, $[M+K]^+$)that were detected by LALDI-MS at 100 pmol of each reagent. As no signals were detected for the molecular ion or adducts of **3.35** during LALDI-MS analysis from HBSS, the reproducibility of the data for **3.35** was assessed using the measured abundance of the fragment ion generated through cleavage at the N-O oxime linkage (m/z 524.2518 – $[M\text{-}C_8H_{14}N_2O_5]^+$). Error bars represent the standard deviation of the mean, where n=3 (individual samples on separate target spots).*

These results highlight the ability of the LDI enhancing label to selectively promote ionization and MS detection of only the labelled species in the presence of buffer and contaminants. Assigned LALDI-MS spectra obtained from this experiment can be found in *Appendix IV.*

*Table 3.4: Summary of LALDI-MS analysis of LALDI tag-labelled glycans **3.33**-**3.37** from a background of HBSS. MS stability was assessed on the relative abundance of the molecular ion ($M^{\bullet+}$) and metal adduct ions ($[M+Na]^+$, $[M+K]^+$) compared to the fragments, degree of fragmentation, and ease of fragment assignment then scored using an arbitrary scale from **++** to **– –**.*

LALDI tag-labelled glycan	Detected as molecular ion	MS stability	Limit of detection (pmol)	Analysis location
3.33	Yes	**+/–**	1	*Appendix IV – Figure A.13*
3.34	Yes	**+/–**	10	*Appendix IV – Figure A.14*
3.35	No	**–**	10	*Appendix IV – Figure A.15*
3.36	Yes	**+**	10	*Appendix IV – Figure A.16*
3.37	Yes	**+**	10	*Appendix IV – Figure A.17*

3.5 In *situ* labelling and analysis of lactose from cow's milk

To determine if the LALDI tag reagents could be used to label reducing carbohydrates in situ, LALDI tag PU-OEG-NH_2 **3.5** and PU-OEG-2AB **3.16** were selected to label lactose directly in cow's milk. Cow's milk was chosen as a proof of concept because of its high lactose concentration (4 – 5% w/w).[247] Following each labelling reaction, samples were diluted ten-fold, subjected to centrifugation and the supernatant analyzed directly by LALDI-MS without any additional purification or modifications (*Figure 3.12*).

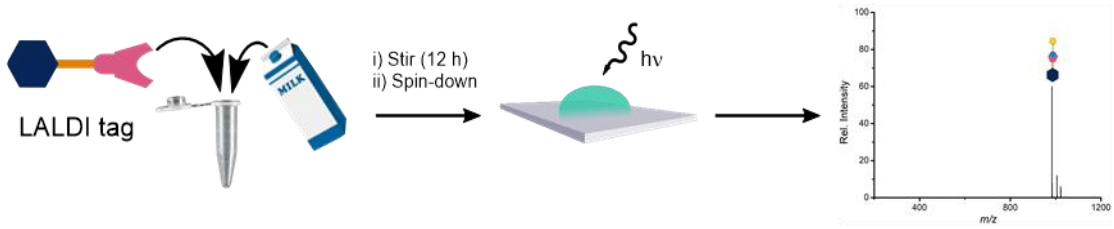


Figure 3.12: A schematic representation of the general workflow for in situ labelling of lactose from milk and subsequent LALDI-MS analysis.

3.5.1 In situ labelling and analysis with PU-OEG-NH_2 **3.5**

In situ labelling of lactose from milk using PU-OEG-NH_2 **3.5** was performed following the conditions described in *Table 3.2.* It was noted that PU-OEG-NH_2 **3.5** dissolved very readily under aqueous conditions.

Following the labelling reaction, samples were prepared by 10-fold dilution, followed by centrifugation. The supernatant was then spotted onto the target plate and the solvent evaporated under reduced pressure. The samples were then analyzed by LALDI-MS without any additional purification, chemical modifications, MALDI matrices, or SALDI surface materials with at least three individual repeats measured for each sample.

LALDI-MS analysis of the PU-OEG-NH_2 **3.5** labelled milk gave a clean spectrum with signals that could be clearly identified as the $[M+Na]^+$ and $[M+K]^+$ adduct ions of LALDI tag-labelled lactose **3.34** (*Figure 3.13*).

Only a single fragmentation was observed in the LALDI-MS analysis, which could be easily identified as fragmentation at the oxime linkage (*m/z* 525.2518). However, as the signal at *m/z* 525.2518 is also known to be generated by fragmentation of PU-OEG-NH_2 **3.5**, the presence of unreacted **3.5** in the sample could not be ruled out.

No data could be provided for the precision of this experiment as detection of LALDI tag-labelled lactose **3.34** was only achieved in one of the three individual repeats carried out during this experiment. This result emphasizes the need to continue optimizing LALDI-MS detection of the LALDI tag reagents and their labelled products.

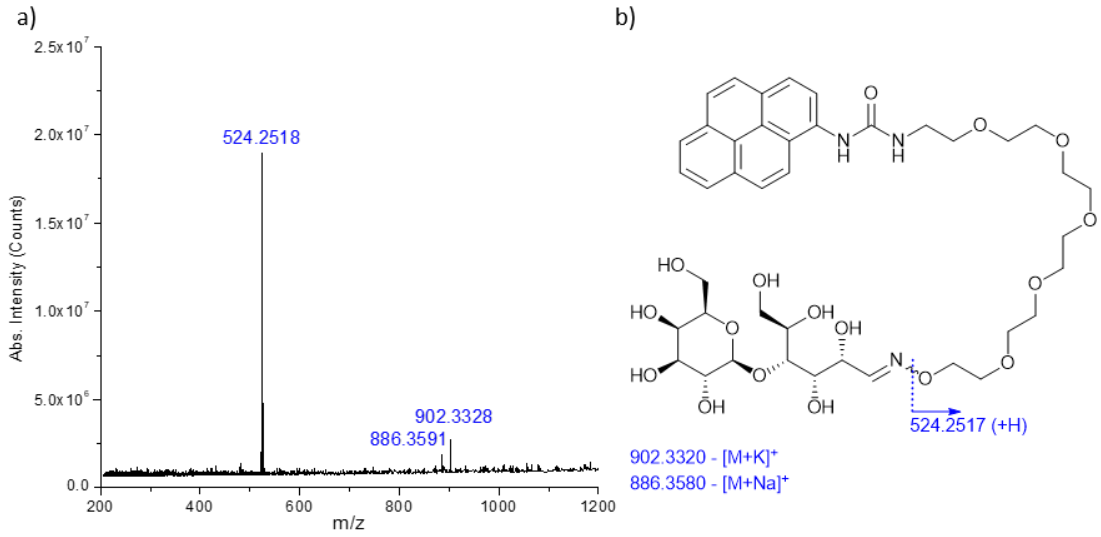


*Figure 3.13: LALDI-MS analysis of milk following glycan labelling reaction with LALDI tag **3.5**. a) LALDI-MS spectra obtained following direct LALDI-MS analysis of the labelling reaction, identified fragments and adducts are annotated with their observed m/z values. b) Chemical structure for LALDI tag-labelled lactose **3.34** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

3.5.2 In situ labelling and analysis with PU-OEG-2AB **3.16**

To determine whether glycan labelling with PU-OEG-2AB **3.16** could be achieved under aqueous conditions, the in situ labelling of lactose from milk using PU-OEG-2AB **3.16** was performed under two different set of conditions. One following the conditions described in *Table 3.2* (Rxn-1), and another where the reaction solvent was changed from 30% AcOH in dimethyl sulfoxide (DMSO) to 30% AcOH in H_2O (Rxn-2).

LC-MS analysis of each reaction indicated that PU-OEG-2AB **3.16** had successfully been able to label lactose from milk under both sets of conditions (*Figure 3.14*). LC-MS analysis found that Rxn-1 gave a cleaner mass chromatogram than Rxn-2. Although the UV chromatogram indicated that under both conditions only two of the compounds detected by the mass chromatogram were UV active, both of which could be identified as the starting material **3.16** and labelled product **3.37**.

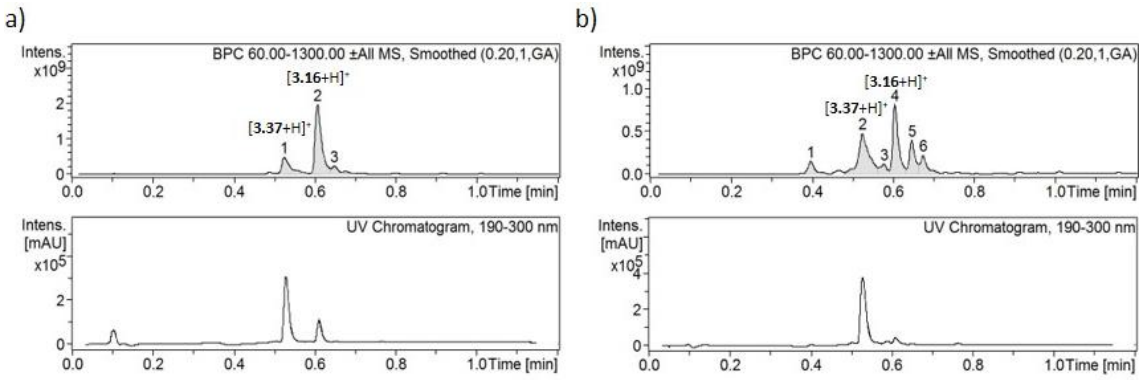


*Figure 3.14: A comparison of the mass and UV absorption chromatograms obtained through LC-MS analysis of the two in situ lactose labelling experiments performed in milk using PU-OEG-2AB **3.16**. a) Rxn-1 (Milk labelling reaction in AcOH:DMSO): Peak 1(R_t ~0.55 min) = Labelled lactose **3.37**. Peak 2 (R_t ~0.6 min) = unreacted PU-OEG-2AB **3.37**. The small peak in the UV chromatogram at R_t ~0.1 min corresponds to DMSO. b) Rxn-2 (Milk labelling reaction in AcOH:H_2O): Peak 2 (R_t ~0.55 min) = Labelled lactose **3.37**. Peak 4 (R_t ~0.6 min) = unreacted PU-OEG-2AB **3.16**.*

Following the labelling reactions, samples were prepared by 10-fold dilution followed by centrifugation. The supernatant was then spotted onto the target plate and the solvent evaporated under reduced pressure. The samples were then analyzed by LALDI-MS without any additional purification, chemical modifications, MALDI matrices, or SALDI surface materials with at least three individual repeats measured for each sample.

LALDI-MS analysis of Rxn-1 showed clear detection of PU-OEG-2AB **3.16** as its $M^{\bullet+}$ and $[M+Na]^+$ ions, but was unable to detect any PU-OEG-2AB-labelled lactose **3.37** (*Appendix V – Figure A.18*).

It is unclear why none of labelled lactose **3.37** was observed during the LALDI-MS analysis of Rxn-1, although non-volatile solvents, such as DMSO, are known to be deleterious for MALDI-MS analysis as they can cause signal suppression by preventing volatilization of the analyte.[248] Therefore, it was suspected that incomplete evaporation of DMSO from the sample could have prevented the desorption and ionization of the labelled lactose during LDI-MS analysis of Rxn-1.

LALDI-MS analysis of Rxn-2 gave a complicated spectrum with a large number of high-intensity signals between *m/z* 200-800 (*Figure 3.15a*). However, it was possible to confirm that labelling of lactose from milk had been achieved as PU-OEG-2AB-labelled lactose **3.37** could be clearly identified as the $[M+Na]^+$ ion. While a number of other signals were identified as either known fragments of **3.37** or unreacted PU-OEG-2AB **3.16**, several other peaks could not be identified. One explanation is that the unassigned species could have formed from either unexpected side reactions between PU-OEG-2AB **3.16** and other species in the milk, or through degradation of **3.16** and **3.37** in the reaction mixture due to the high concentration of acid and reducing agent.

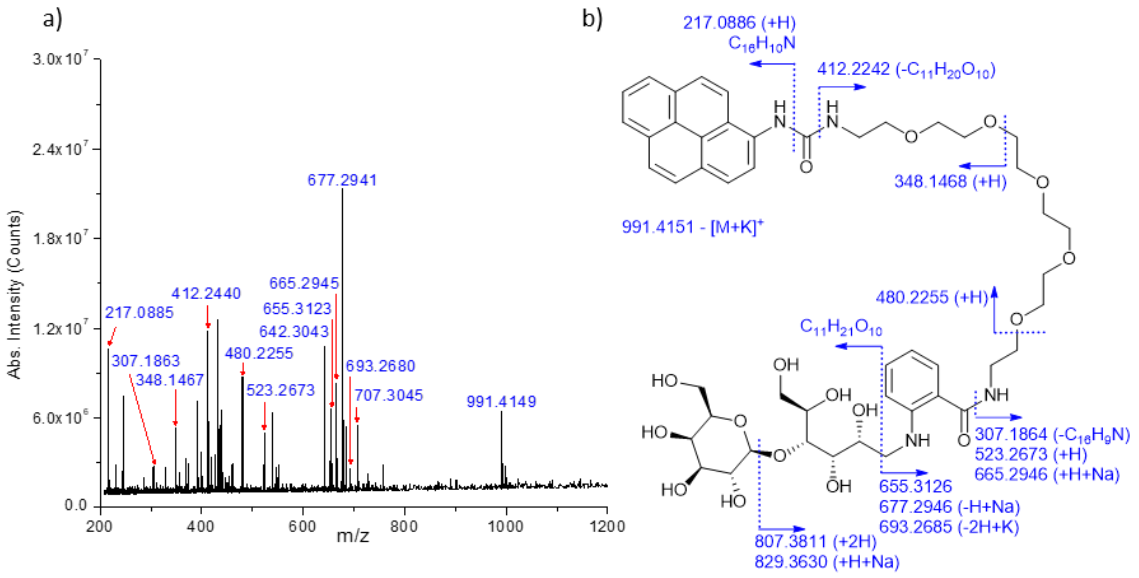


*Figure 3.15: LALDI-MS analysis of Rxn-2 following glycan labelling reaction with LALDI tag **3.16** under aqueous conditions. a) LALDI-MS spectra obtained following direct LALDI-MS analysis of Rxn-2, identified fragments and adducts are annotated with their observed m/z values. b) Chemical structure for LALDI tag-labelled lactose **3.37** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

The results shown in *Figure 3.13* and *Figure 3.15* demonstrate that both labelling and analysis of reducing carbohydrates with our LALDI tags can be performed directly from complex biological samples. However, the presence of unidentifiable peaks in *Figure 3.15* suggests that the LALDI

tag could be labelling other species in the sample or ionizing them by acting as a matrix. Therefore, additional independent repeats of the labelling reaction and analysis will be required to confirm this. If the repeats are found to be consistent with the data shown in *Figure 3.15*, additional experimentation will be required to determine the cause of the additional peaks, followed by further optimization of both the labelling reaction and sample analysis to circumvent this issue and improve the analysis.

No data could be provided for the precision of this experiment as detection of LALDI tag-labelled lactose **3.37** was only achieved in one of the three individual repeats carried out during this experiment.

3.6 Attempts at labelling more complicated glycans

Following the successful labelling and detection of simple mono- and disaccharides, an investigation was undertaken to determine if LALDI tags could be applied to larger and more complicated glycans. Naturally occurring glycans are often large and branched structures, are usually present in complex mixtures of other glycans, and can often featuring labile groups such as sialic acids. Therefore, to ensure that LALDI tags are suitable for carrying out analysis on such biomolecules, it was important to make sure that large, branched, and sialylated glycans can be labelled efficiently and easily detected, not just in isolation but as mixtures as well. To address these challenges, four glycans were selected (*Figure 3.16*): 3-α-sialyl-*N*-acetyllactosamine **3.39**, α-D-galactosyl-β-D-mannotriose **3.40**, α-maltohexaose **3.41**, and Gm1 pentasaccharide **3.42**. Glycan labelling was attempted using PU-OEG-2AB **3.16** via the aqueous labelling conditions described for Rxn-2 in Section 3.5.2.

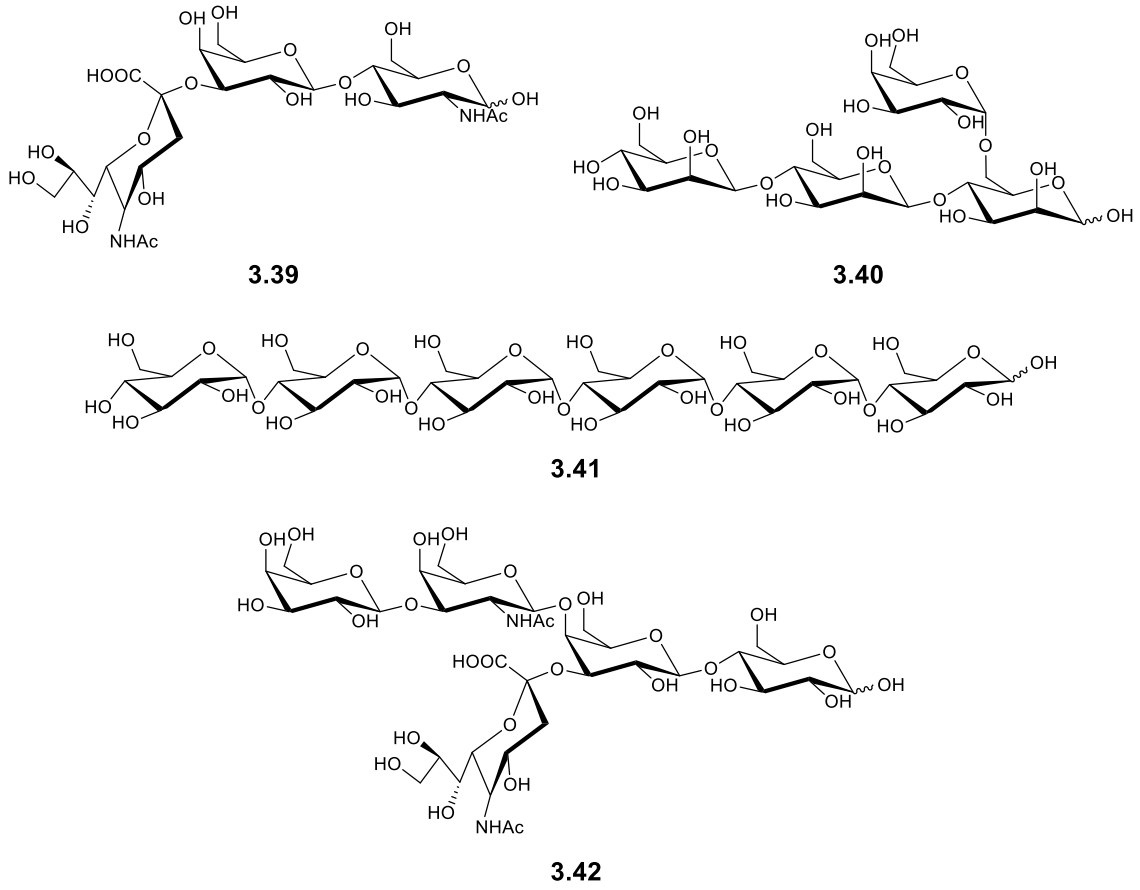


*Figure 3.16: Structures of glycans used in labelling experiments: 3-α-sialyl-N-acetyllactosamine **3.39**, α-D-galactosyl-β-D-mannotriose **3.40**, α-maltohexaose **3.41**, and Gm1 pentasaccharide **3.42**.*

Due to the limited availability of the glycans, the labelling reactions were carried out using a reduced concentration of label **3.16**. This also allowed the volume of the reaction mixture to be increased to a more manageable scale. To determine if the labelling was possible under the adapted reaction conditions a test reaction using maltohexaose **3.43** was carried out. Analysis of the test labelling reactions using HR-ESI-MS showed that a trace amount of **3.16**-labelled maltohexaose was present in the reaction mixture. Labelling was then performed on each of the four glycans, followed by LALDI-MS analysis of the crude reaction mixtures. However, no signals relating to any of the labelled glycans were detected by LALDI-MS, only signals relating to the label **3.16** observed in LALDI-MS spectra. These results suggested that glycan labelling had occurred in the reaction, but that the amount of labelled glycan present in the reaction was too low for detection by LALDI-MS.

It was suspected that the poor labelling efficiency of the reaction was due to the reduced concentration of **3.16** during the labelling reaction. The simplest solution would be to try improve the labelling efficiency by repeating the labelling reactions at the concentrations

optimized reported by Bigge et al.[175] However, due to limited time and material, no further work was carried out on this study.

3.7 Conclusions

In summary, a series of five unique LALDI tag reagents were synthesized which incorporated the two selected LDI enhancers **2.3** and **2.27** and three reactive handles. Labelling of simple glycans with LALDI tags was possible via established glycan labelling procedures derived from the literature.[158,175,244] Analysis of the labelled lactose and GlcNAc by LALDI-MS showed that detection of the labelled glycans could be achieved without the need for additional chemical modifications, MALDI matrices, or SALDI surface materials.

Detection of the labelled lactose and GlcNAc was also shown to be possible from a complex background of Hanks' balanced salt solutions, demonstrating the capabilities of the LALDI tag reagents to achieve ionization even in the presence of an excess of salts and other contaminants.

In situ labelling and detection of lactose directly from cow's milk was also demonstrated using hydroxylamine LALDI tag **3.5** and 2-AB LALDI tag **3.16**, highlighting the potential of LALDI tags for labelling and analyzing glycans directly from a biological source. However, further applications of this chemistry are currently limited by a number of unknowns regarding the stability of labelled glycans during different LALDI-MS experiments. To address these issues, further investigations into the labelling of native glycans and the analysis of the labelled products will need to be conducted.

Attempts were made to analyze larger and more complicated glycans using the LALDI tag reagents. While it was not possible to achieve detection using LALDI-MS under the conditions used, evidence from HR-ESI-MS indicated that glycan labelling had occurred but concentrations of the labelled glycans were below the limit of detection for LALDI-MS. Therefore, it was proposed that detection of the larger glycans could be achieved by improving the efficiency of the labelling reaction by increasing the concentration of label used.

3.8 Experimental

3.8.1 General experimental

All reactions were carried out under an atmosphere of dry nitrogen unless otherwise stated, using anhydrous solvents from a solvent purification system (Innovative Technology Inc. PureSolv), with the exception of anhydrous DMF, which was purchased from Acros Organics. Commercial oligosaccharides were purchased from Carbosynth and used directly without any additional purification. Gm1 pentasaccharide **3.42** was synthesized and purified by Ryan McBerney of the Turnbull group at the University of Leeds. All other chemical starting reagents were purchased from commercial suppliers and used without further purification. Azeotropic distillation with toluene was used to remove persistent moisture from reagents with oligoethylene glycol chains before being used in a reaction. The identity and purity of known compounds was confirmed through the comparison of experimentally obtained data to values reported in the literature.

Thin layer chromatography (TLC) was carried out on Merck TLC Silica gel 60 F_{254} plates. Flash column chromatography was performed using Merck Geduran silica gel 60 (40–63 μm). All retention factors (R_f) are given to two decimal places along with the solvent system. Reversed-phase chromatography was performed using prepacked RediSep® Rf Reversed-phase C_{18} columns. Lyophilisation of compounds was performed using a Virtis Benchtop K freeze dryer.

The melting point of recrystallised solids was determined using a Griffin MFB-590 Melting Point Apparatus with a glass capillary melting point tube. Melting points (m.p.) are recorded as a range between the meniscus point and liquefaction with values given to the nearest degree in Celsius (°C).

^{1}H and ^{13}C NMR spectra were recorded in deuterated solvents on a Bruker AVANCE 500 Ultrashield or Bruker AVANCE 400 Ultrashield NMR Spectrometer. Chemical shifts are referenced to residual solvent peaks and are quoted in ppm. Coupling constants (*J*) are reported to the nearest 0.1 Hz. Assignment of spectra was based on expected chemical shifts and coupling constants, aided by DEPT, COSY, HMQC, and HMBC where appropriate.

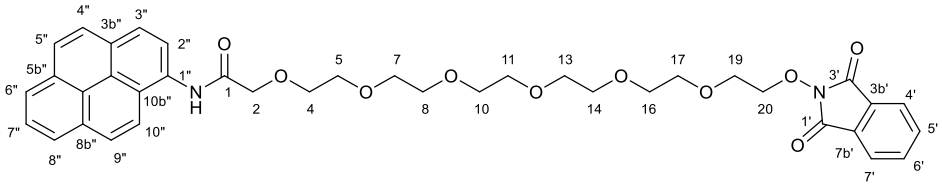
High-resolution electrospray mass spectrometry (HR-ESI-MS) was performed using a Bruker MaXis Impact spectrometer; *m/z* values are reported to four decimal places. LC-MS was recorded on an Agilent Technologies 1200 series HPLC combined with a Bruker HCT Ultra ion trap using 50 × 20 mm C_{18} reversed-phase columns with a solvent system of increasing

acetonitrile (5 to 95%) in water, each containing 0.1% formic acid. A flow rate of 1.5 cm^3 min^{-1} was used and *m/z* values are given to one decimal place. Retention times (R_t) are provided in minutes to the nearest two decimal places. LALDI-MS was carried out on a solariX XR FTMS 9.4T MALDI mass spectrometer. All optimised analysis was performed on a Bruker Daltonics MTP AnchorChip Target 384. Ionisation/desorption of the samples was achieved using a fixed wavelength (355nm) Bruker Smartbeam II (N_2, NdYag) laser*; m/z* values are reported to four decimal places. LALDI-MS fragmentation assignment was carried out with the assistance of a web application developed by ChemCalc that provides a list of possible molecular formulas from a given monoisotopic mass.[249,250]

Infrared (IR) spectroscopy was carried out using a Bruker Alpha Platinum ATR. Samples were analysed neat and absorption maxima (ν_{max}) are given in wave numbers (cm^{-1}) to the nearest whole wavenumber. Signals are defined as either strong (s), medium (m), weak (w), or broad (br). UV-Vis absorption was measured using an Agilent Technologies Cary 100 UV-Vis Spectrophotometer with all samples were analysed as solutions in a 10 mm Hellma Analytics High Precision Quarts Suprasil cell. Absorption maxima (λ_{max}) are given in nanometers to the nearest whole nanometer with the corresponding molar extinction coefficient (ε) given in $M^{-1}\,cm^{-1}$.

3.8.2 Experimental procedures

20-[(1,3-dioxo-2,3-dihydro-1H-isoindol-2-yl)oxy]-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.2**

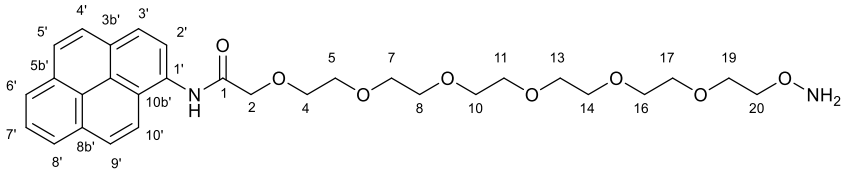


3.2

A solution of 20-hydroxy-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **2.3** (210 mg, 0.389 mmol) in THF (3 mL), and *N*-hydroxyphthalimide **3.1** (190 mg, 1.17 mmol) were added sequentially to an ice-cooled solution of DIAD (230 µL, 1.17 mmol) and triphenylphosphine (306 mg, 1.17 mmol) in THF (5 mL) and stirred for 10 min. The reaction was then allowed to warm to room temperature and stir overnight (20 h). The reaction mixture was then concentrated *in vacuo.* The crude residue was purified by flash chromatography (SiO_2; EtOAc–MeOH 100:0 → 90:10) to afford the product **3.2** as a clear orange oil (221 mg, 0.341 mmol, 88%). $\boldsymbol{R_f}$ = 0.23 (SiO_2; EtOAc–MeOH 90:10). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.57 (br s, 1H, CON*H*-1), 8.49 (d, 1H, *J* 8.3 Hz, Aryl), 8.21 – 8.15 (m, 5H, Aryl), 8.12 (d, 1H, *J* 9.2 Hz, Aryl), 8.06 – 7.98 (m, 4H, Aryl), 7.80 – 7.76 (m, 2H, C*H*-4', C*H*-7'), 7.72 – 7.65 (m, 2H, C*H*-5', C*H*-6'), 4.36 (s, 2H, CH_2-2), 4.34 – 4.32 (m, 2H, OCH_2CH_2O), 3.98 – 3.95 (m, 2H, OCH_2CH_2O), 3.86 – 3.82 (m, 2H, OCH_2CH_2O), 3.82 – 3.79 (m, 2H, OCH_2CH_2O), 3.73 – 3.70 (m, 2H, OCH_2CH_2O), 3.60 – 3.56 (m, 2H, OCH_2CH_2O), 3.52 – 3.50 (m, 2H, OCH_2CH_2O), 3.47 – 3.44 (m, 2H, OCH_2CH_2O), 3.41 – 3.37 (m, 2H, OCH_2CH_2O), 3.36 – 3.33 (m, 2H, OCH_2CH_2O), 3.27 (s, 4H, OCH_2CH_2O).**^{13}C NMR** (100 MHz, $CDCl_3$) δ 169.2 (*C*ONH-1), 163.54 (*C*ONH*C*O-1', *C*ONH*C*O-3'), 134.5 (*C*H-5', *C*H-6'), 131.5 (Aryl *C*), 131.0 (Aryl *C*), 130.2 (Aryl *C*), 129.3 (Aryl *C*), 129.1 (*C*-3b', *C*-7b'), 128.0 (Aryl *C*H), 127.5 (Aryl *C*H), 127.0 (Aryl *C*H), 126.3 (Aryl *C*H), 125.5 (Aryl *C*H), 125.4 (Aryl *C*H), 125.3 (Aryl *C*), 125.2 (Aryl *C*H), 124.9 (Aryl *C*), 123.8 (Aryl *C*), 123.6 (*C*H-4', *C*H-7'), 122.3 (Aryl *C*H), 120.9 (Aryl *C*H), 71.8 (*C*H_2), 71.2 (*C*H_2), 70.9 (*C*H_2), 70.8 (*C*H_2), 70.6 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.4 (*C*H_2), 70.4 (*C*H_2), 69.3 (*C*H_2). **HR-ESI-MS** Calculated for $C_{38}H_{41}N_2O_{10}^+$: *m/z* 685.2756 $[M+H]^+$; found 685.2765. **LC-MS** found 685.58, $\boldsymbol{R_t}$ = 0.75.

3 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

20-hydroxylamino-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.3**

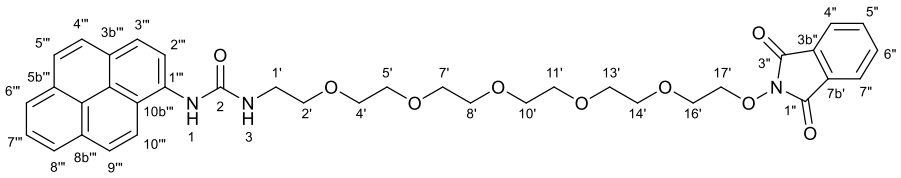


3.3

80% hydrazine hydrate (46 µL, 0.76 mmol) was added dropwise to a rapidly stirred solution of 20-[(1,3-dioxo-2,3-dihydro-1H-isoindol-2-yl)oxy]-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.2** (164 mg, 0.253 mmol) in THF (6 mL) until determined complete by LC-MS (3h). The reaction mixture was then filtered through celite and washed through with THF (20 mL). The solvent was then removed under reduced pressure to give the product **3.3** as a brown oil. (134 mg, 0.242 mmol, 96%) $\boldsymbol{R_f}$ = 0.16 (SiO_2; EtOAc–MeOH 80:20). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.60 (br s, 1H, CON*H*-1), 8.48 (d, 1H, *J* 8.2 Hz, Aryl), 8.21 – 8.15 (m, 4H, Aryl), 8.11 (d, 1H, *J* 9.2 Hz, Aryl), 8.05 – 7.98 (m, 3H, Aryl), 4.35 (s, 2H, C*H*$_2$-2), 3.97 – 3.92 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.85 – 3.81 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.80 – 3.76 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.72 – 3.68 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.61 – 3.58 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.56 – 3.53 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.52 – 3.48 (m, 4H, OC*H*$_2$C*H*$_2$O), 3.46 – 3.42 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.40 – 3.36 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.31 – 3.24 (m, 4H, OC*H*$_2$C*H*$_2$O). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 169.3 (*C*ONH-1), 131.4 (Aryl *C*), 131.0 (Aryl *C*), 130.2 (Aryl *C*), 129.3 (Aryl *C*), 127.9 (Aryl *C*H), 127.4 (Aryl *C*H), 127.0 (Aryl *C*H), 126.3 (Aryl *C*H), 125.5 (Aryl *C*H), 125.3 (Aryl *C*H), 125.3 (Aryl *C*), 125.1 (Aryl *C*H), 124.8 (Aryl *C*), 123.8 (Aryl *C*), 122.3 (Aryl *C*H), 120.9 (Aryl *C*H), 74.8 (*C*H$_2$), 71.7 (*C*H$_2$), 71.1 (*C*H$_2$), 70.8 (*C*H$_2$), 70.5 (*C*H$_2$), 70.4 (*C*H$_2$), 70.4 (*C*H$_2$), 70.4 (*C*H$_2$), 70.3 (*C*H$_2$). **HR-ESI-MS** Calculated for $C_{30}H_{39}N_2O_8^+$: *m/z* 555.2701 $[M+H]^+$; found 555.2698. **LC-MS** found 555.61, $\boldsymbol{R_t}$ **=** 0.55.

4 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

3-{17-[(1,3-dioxo-2,3-dihydro-1H-isoindol-2-yl)oxy]-3,6,9,12,15-pentaoxaheptadecan-1-yl}-1-(pyren-1-yl)urea **3.4**

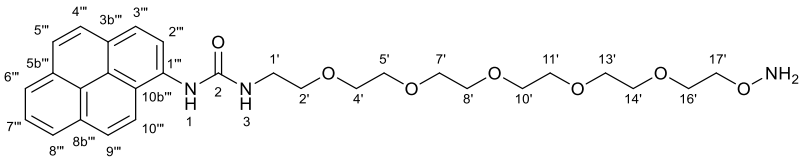


3.4

A solution of 3-(17-hydroxy-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **2.27** (500 mg, 0.953 mmol) in THF (3 mL), and *N*-hydroxyphthalimide **3.1** (466 mg, 2.86 mmol) were added sequentially to an ice-cooled solution of DIAD (563 µL, 2.86 mmol) and triphenylphosphine (750 mg, 2.86 mmol) in THF (5 mL) and stirred for 10 min. The reaction was then allowed to warm to room temperature and stir overnight (20 h). The reaction mixture was then concentrated *in vacuo.* The crude residue was purified by flash chromatography (SiO_2; EtOAc–MeOH 100:0 → 90:10) to afford the product **3.4** as a brown waxy solid (390 mg, 0.582 mmol, 61%). $\boldsymbol{R_f}$ = 0.38 (SiO_2; EtOAc–MeOH 90:10). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.28 (d, 1H, *J* 8.3 Hz, Aryl), 8.13 (d, 1H, *J* 9.2 Hz, Aryl), 8.02 (d, 2H, *J* 7.5 Hz, Aryl), 7.96 (d, *J* 8.5 Hz, Aryl), 7.89 – 7.80 (m, 4H, Aryl + N*H*-1), 7.71 – 7.65 (m, 2H, C*H*-4”, C*H*-7”), 7.65 – 7.55 (m, 2H, C*H*-5”, C*H*-6”), 6.20 (br s, 1H, N*H*-3), 4.22 – 4.14 (m, 2H, OCH_2CH_2O), 3.64 – 3.43 (m, 22H, OCH_2CH_2O). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 163.4 (*C*ONH*C*O-1”, *C*ONH*C*O-3”), 157.2 (NH*C*ONH-2), 134.4 (*C*H-5”, *C*H-6”), 132.5 (Aryl *C*), 131.4 (Aryl *C*), 130.9 (Aryl *C*), 128.8 (*C*-3b”, *C*-7b”), 128.0 (Aryl *C*), 127.3 (Aryl *C*H), 127.1 (Aryl *C*H), 126.1 (Aryl *C*H), 125.9 (Aryl *C*H), 125.3 (Aryl *C*H), 125.2 (Aryl *C*), 124.9 (Aryl *C*H), 124.7 (Aryl *C*), 124.5 (*C*H-4”, *C*H-7”), 123.6 (Aryl *C*), 123.4 (Aryl *C*H), 122.0 (Aryl *C*H), 121.3 (Aryl *C*H), 70.6 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.4 (*C*H_2), 70.3 (*C*H_2), 70.2 (*C*H_2), 69.1 (*C*H_2), 40.3 (*C*H_2). **HR-ESI-MS** Calculated for $C_{37}H_{40}N_3O_9$: 670.2759 *m/z* $[M+H]^+$; found 670.2765. **LC-MS** found 670.52, $\boldsymbol{R_t}$ **=** 0.65.

4 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

3-[17-(aminooxy)-3,6,9,12,15-pentaoxaheptadecan-1-yl]-1-(pyren-1-yl)urea **3.5**



3.5

80% hydrazine hydrate (45 µL, 0.75 mmol) was added dropwise to a rapidly stirred solution of 3-{17-[(1,3-dioxo-2,3-dihydro-1H-isoindol-2-yl)oxy]-3,6,9,12,15-pentaoxaheptadecan-1-yl}-1-(pyren-1-yl)urea **3.4** (100 mg, 0.149 mmol) in THF (1 mL) until determined complete by LC-MS (5 h). The solvent was then removed under reduced pressure and the crude residue purified by flash chromatography (SiO_2; DCM–MeOH 95:5) to give the product **3.5** as a brown waxy solid. (64 mg, 0.12 mmol, 80%) $\boldsymbol{R_f}$ = 0.24 (SiO_2; DCM–MeOH 95:5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.38 (d, 1H, *J* 8.3 Hz, Aryl), 8.23 (d, 1H, *J* 9.3 Hz, Aryl), 8.13 – 8.03 (m, 3H, Aryl), 8.01 – 7.87 (m, 5H, Aryl + N*H*-1), 6.29 (br s, 1H, N*H*-3), 3.72 – 3.67 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.66 – 3.43 (m, 22H, OC*H*$_2$C*H*$_2$O). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 157.1 (NH*C*ONH-2), 132.7 (Aryl *C*), 131.6 (Aryl *C*), 131.1 (Aryl *C*), 128.1 (Aryl *C*), 127.5 (Aryl *C*H), 127.2 (Aryl *C*H), 126.2 (Aryl *C*H), 126.1 (Aryl *C*H), 125.5 (Aryl *C*H), 125.4 (Aryl *C*), 125.0 (Aryl *C*H), 125.0 (Aryl *C*), 124.6 (Aryl *C*H), 123.6 (Aryl *C*), 122.0 (Aryl *C*H), 121.5 (Aryl *C*H), 74.7 (*C*H$_2$), 70.6 (*C*H$_2$), 70.5 (*C*H$_2$), 70.5 (*C*H$_2$), 70.4 (*C*H$_2$), 70.3 (*C*H$_2$), 70.2 (*C*H$_2$), 69.6 (*C*H$_2$), 40.4 (*C*H$_2$). **HR-ESI-MS** Calculated for $C_{29}H_{37}N_3NaO_7$: 540.2524 *m/z* $[M+Na]^+$; found 562.2526. **LC-MS** Calculated for $C_{29}H_{38}N_3O_7$: 540.27 *m/z* $[M+H]^+$; found 540.46, $\boldsymbol{R_t}$ = 0.60.

3 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

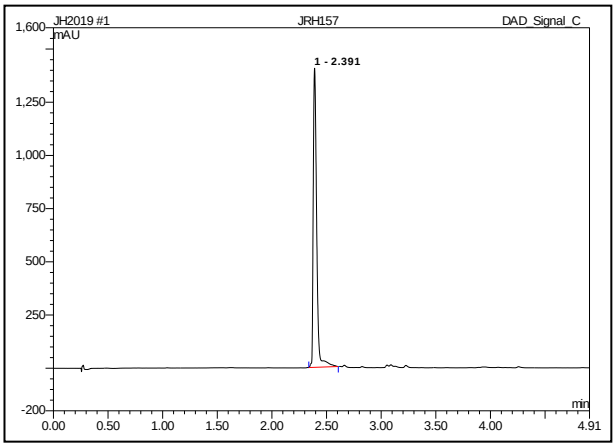
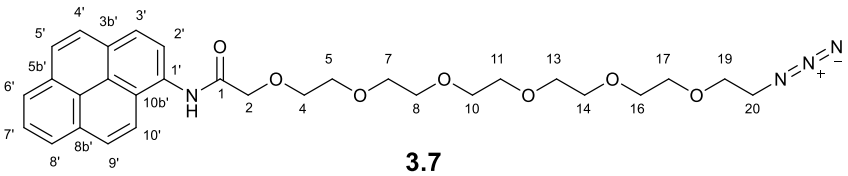


*Figure 3.17: Analytical HPLC trace for 3-[17-(aminooxy)-3,6,9,12,15-pentaoxaheptadecan-1-yl]-1-(pyren-1-yl)urea **3.5**.*

20-azido-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.7**

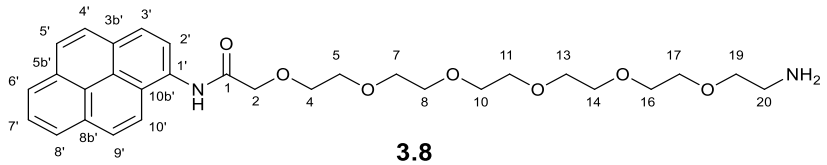


3.7

Under an N_2 atmosphere, 60% NaH in mineral oil (200 mg, 5.00 mmol) was added to an ice-cooled solution of 17-azido-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.7** (450 mg, 1.46 mmol) in DMF (5 mL), followed by the dropwise addition of a solution of 2-chloro-*N*-pyren-1-yl acetamide **2.11** (368 mg, 1.25 mmol) in DMF (5 mL). The reaction was then allowed to warm to room temperature and stirred until determined complete by LC-MS (21 h). The reaction was diluted by EtOAc (50 mL), and washed with 1M HCl (10 mL), 1M LiCl (5 × 25 mL). The aqueous layers were then combined, extracted with EtOAc (5 × 10 mL). Then the organic layers were combined, washed with brine (10 mL), and dried (Na_2SO_4). The solvent removed *in vacuo* to give a crude brown oil which was purified by flash chromatography (SiO_2; DCM–MeOH 1:0 → 95:5) to yield the product **3.7** as a brown oil (594 mg, 1.05 mmol, 84%). $\boldsymbol{R_f}$ = 0.42 (SiO_2; SiO_2; DCM–MeOH 95:5). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.56 (br s, 1H, CON*H*-1), 8.50 (d, 1H, *J* 8.3 Hz, Aryl), 8.21 – 8.16 (m, 4H, Aryl), 8.12 (d, 1H, *J* 9.2 Hz, Aryl), 8.07 – 7.97 (m, 3H, Aryl), 4.36 (s, 2H, C*H*$_2$-2), 3.97 – 3.94 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.85 – 3.82 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.72 – 3.69 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.61 – 3.58 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.57 – 3.54 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.54 – 3.50 (m, 4H, OC*H*$_2$C*H*$_2$O), 3.48 – 3.45 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.41 – 3.38 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.33 (t, 2H, *J* 5.1 Hz, CH_2-20), 3.29 (s, 4H, OC*H*$_2$C*H*$_2$O). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 169.2 (*C*ONH-1), 131.5 (Aryl *C*), 131.0 (Aryl *C*), 130.2 (Aryl *C*), 129.3 (Aryl *C*), 128.0 (Aryl *C*H), 127.5 (Aryl *C*H), 127.0 (Aryl *C*H), 126.3 (Aryl *C*H), 125.5 (Aryl *C*H), 125.3 (Aryl *C*H), 125.3 (Aryl *C*), 125.1 (Aryl *C*H), 124.9 (Aryl *C*), 123.8 (Aryl *C*), 122.2 (Aryl *C*H), 120.9 (Aryl *C*H), 71.8 (*C*H$_2$), 71.2 (*C*H$_2$-2), 70.9 (*C*H$_2$), 70.7 (*C*H$_2$), 70.6 (*C*H$_2$), 70.6 (*C*H$_2$), 70.6 (*C*H$_2$), 70.5 (*C*H$_2$), 70.4 (*C*H$_2$), 70.4 (*C*H$_2$), 70.1 (*C*H$_2$), 50.8 (*C*H$_2$-20). **HR-ESI-MS** Calculated for $C_{30}H_{37}N_4O_7^+$: *m/z* 565.2657 $[M+H]^+$; found 565.2661. **LC-MS** found 565.62, $\boldsymbol{R_t}$ = 0.67. **IR** (Neat, ν_{max}/cm^{-1}) 3304 m (NH amide), 2098 m (N=N=N), 1689 s (C=O amide), 1518–1416 m (C=C aromatic).

1 × CH_2 from the OEG chain was unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

20-amino-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.8**

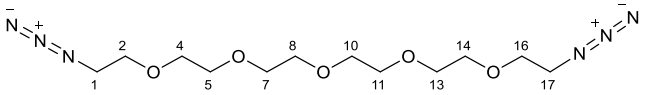


3.8

Under an N_2 atmosphere, a solution on triphenylphosphine (294 mg, 1.12 mmol) in THF (3.0 ml) was added to a solution of 20-azido-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.7** (575 mg, 1.019 mmol) in THF (7.0 mL) at 0 °C (ice bath). The reaction was then allowed to warm to room temperature and stirred until determined complete by TLC (24 h). H_2O was added (5.0 mL) and the reaction mixture allowed to stir for 1 h to hydrolyse the iminophosphorane intermediate. The solvent was then removed *in vacuo* and the crude produce purified by flash chromatography (SiO_2; DCM–sat. NH_3 in MeOH 1:0 → 95:5 → 90:10) to yield the product **3.8** as a yellow oil (435 mg, 0.808 mmol, 79%). $\boldsymbol{R_f}$ = 0.13 (SiO_2; DCM–sat. NH_3 in MeOH 95:5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.57 (br s, 1H, CON*H*-1), 8.49 (d, 1H, *J* 8.3 Hz, Aryl), 8.21 – 8.14 (m, 4H, Aryl), 8.11 (d, 1H, *J* 9.2 Hz, Aryl), 8.06 – 7.97 (m, 3H, Aryl), 4.35 (s, 2H, C*H*$_2$-2), 3.94 (dd, 2H, *J* 5.4, 3.2 Hz, OC*H*$_2$C*H*$_2$O), 3.82 (dd, 2H, *J* 5.4, 3.2 Hz, OC*H*$_2$C*H*$_2$O), 3.70 (dd, 2H, *J* 5.6, 3.8 Hz, OC*H*$_2$C*H*$_2$O), 3.54 – 3.48 (m, 6H, OC*H*$_2$C*H*$_2$O), 3.47 – 3.41 (m, 4H, OC*H*$_2$C*H*$_2$O), 3.39 (dd, 2H, *J* 5.9, 3.6 Hz, OC*H*$_2$C*H*$_2$O), 3.28 (s, 4H, OC*H*$_2$C*H*$_2$O), 2.80 (t, 2H, *J* 5.1 Hz, C*H*$_2$-20). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 169.2 (*C*ONH-1), 131.4 (Aryl *C*), 131.0 (Aryl *C*), 130.2 (Aryl *C*), 129.3 (Aryl *C*), 127.9 (Aryl *C*H), 127.5 (Aryl *C*H), 127.0 (Aryl *C*H), 126.3 (Aryl *C*H), 125.5 (Aryl *C*H), 125.3 (Aryl *C*H), 125.5 (Aryl *C*), 125.1 (Aryl *C*H), 124.9 (Aryl *C*), 123.8 (Aryl *C*), 122.3 (Aryl *C*H), 120.9 (Aryl *C*H), 73.3 (*C*H$_2$), 71.7 (*C*H$_2$), 71.1 (*C*H$_2$), 70.8 (*C*H$_2$), 70.5 (*C*H$_2$), 70.4 (*C*H$_2$), 70.4 (*C*H$_2$), 70.3 (*C*H$_2$), 70.3 (*C*H$_2$), 41.8 (*C*H$_2$-20). **HR-ESI-MS** Calculated for $C_{30}H_{39}N_2O_7^+$: *m/z* 539.2752 $[M+H]^+$; found 539.2749. **LC-MS** found 539.43, $\boldsymbol{R_t}$ **=** 0.55. **IR** (Neat, ν_{max}/cm^{-1}) 3377 m (NH_2 asym), 3303 m (NH_2 sym), 1688 s (C=O amide), 1518–1416 m (C=C aromatic).

3 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

1,17-diazido-3,6,9,12,15-pentaoxaheptadecane **3.9** [228]

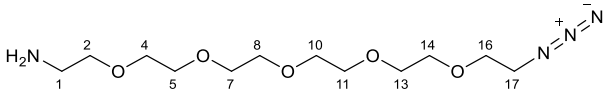


3.9

A solution of triethylamine (1090 µL, 7.820 mmol) in dry THF (1 mL) was added dropwise to a stirred solution of hexaethylene glycol **2.4** (1.00 g, 3.54 mmol) and methanesulfonyl chloride (610 µL, 7.89 mmol) in dry THF (4 mL) at 0 °C (ice bath) and then the mixture was allowed to stir for 1 h. The ice bath was then removed, and the reaction allowed to stir at room temperature for 4 h. The reaction mixture was then diluted with H_2O (5 mL) and the pH adjusted to 8 using $NaHCO_3$. NaN_3 (510 mg, 7.737 mmol) was then added to the reaction mixture and the resulting mixture was heated at reflux for 18 h. The reaction was allowed to cool and Et_2O (10 mL) was added and the layers separated. The aqueous layer was saturated with NaCl and extracted with Et_2O (5 × 10 mL). The organic layers were combined, washed with brine (10 mL) and dried ($MgSO_4$) and solvent removed *in vacuo* to yield the product **3.9** as a colourless oil (986 mg, 2.80 mmol, 79%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 3.67 – 3.64 (m, 20H, OC*H*$_2$C*H*$_2$O), 3.37 (t, 4H, *J* 5.0 Hz, C*H*$_2$-1,17). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 70.8 (*C*H_2), 70.8 (*C*H_2), 70.8 (*C*H_2), 70.8 (*C*H_2), 70.7 (*C*H_2), 70.1 (*C*H_2-2, 16), 50.8 (*C*H_2N_3-1, 17). **HR-ESI-MS** Calculated for $C_{12}H_{24}N_6NaO_5^+$: *m/z* 355.1700 $[M+Na]^+$; found 355.1717. **LC-MS** found 355.1, $\boldsymbol{R_t}$ **=** 1.68. **IR** (Neat, ν_{max}/cm^{-1}) 2120 s (N=N=N).

Characterization is consistent with data reported in the literature.[228]

17-azido-3,6,9,12,15-pentaoxaheptadecan-1-amine **3.10** [228]



3.10

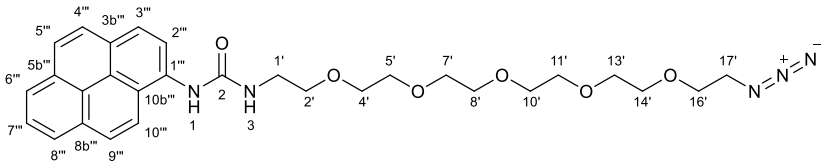
A solution of triphenylphosphine (719 mg, 2.74 mmol) in Et_2O (5 mL) was added dropwise to a vigorously stirred suspension of 1,17-diazido-3,6,9,12,15-pentaoxaheptadecane **3.9** (985 mg, 2.80 mmol) in 5% $HCl_{(aq)}$ (10 mL) then the mixture was allowed to stirred for 24 h. The reaction mixture was then washed with dichloromethane (3 × 40 mL) and the aqueous phase was cooled in an ice bath and the pH adjusted to 12 using KOH pellets. The water was removed by lyophilisation and product was taken up in dichloromethane (25 mL). The organic phase was dried (Na_2SO_4) and the solvent removed *in vacuo* to yield the product **3.10** as a yellow oil that was used without further purification (779 mg, 2.54 mmol, 91%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ

3.67 – 3.59 (m, 18H, OC*H*$_2$C*H*$_2$O), 3.49 (t, 2H, *J* 5.2 Hz, OC*H*$_2$C*H*$_2$O), 3.37 (t, 2H, *J* 5.2 Hz, C*H*$_2$-17), 2.84 (t, 2H, *J* 5.2 Hz, C*H*$_2$-1). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 73.6 (*C*H$_2$), 70.8 (*C*H$_2$), 70.8 (*C*H$_2$), 70.7 (*C*H$_2$), 70.7 (*C*H$_2$), 70.7 (*C*H$_2$), 70.7 (*C*H$_2$), 70.4 (*C*H$_2$), 70.1 (*C*H$_2$), 50.8 (*C*H$_2$N$_3$-17), 41.9 (*C*H$_2$NH$_3$-1). **HR-ESI-MS** Calculated for $C_{12}H_{26}N_4NaO_5^+$: *m/z* 329.1795 $[M+Na]^+$; found 329.1795. **LC-MS** Calculated for $C_{12}H_{27}N_4O_5^+$: *m/z* 307.20 $[M+H]^+$; found 307.20, R_t = 1.96. **IR** (Neat, ν_{max}/cm^{-1}) 3375 w (NH_2 asym), 3312 w (NH_2 sym), 2101 m (N=N=N).

1 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

Characterization is consistent with data reported in the literature.[228]

3-(17-azido-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **3.11**

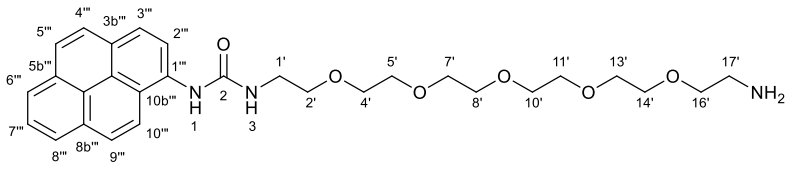


3.11

A solution of 17-azido-3,6,9,12,15-pentaoxaheptadecan-1-amine **3.10** (399 mg, 1.30 mmol) in THF (3.0 mL) was added dropwise to a stirred solution of 1-isocyantopyrene **2.24** (288 mg, 1.18 mmol) in THF (2.0 mL) at 0 °C (ice bath) and allowed to stir for 30 min. The reaction was then allowed to warm to room temperature and stir overnight (21 h). Solvent was then removed under reduced pressure and the crude residue purified by flash chromatography (SiO_2; DCM–MeOH 95:5) to afford the product **3.11** as a yellow waxy solid (560 mg, 1.057 mmol, 89%). R_f = 0.34 (SiO_2; DCM–MeOH 95:5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.37 (d, 1H, *J* 8.3 Hz, Aryl), 8.21 (d, 1H, *J* 9.2 Hz, Aryl), 8.13 – 8.04 (m, 3H, Aryl), 7.99 – 7.92 (m, 4H, Aryl), 7.90 (br s, 1H, N*H*-1), 6.15 (br t, 1H, *J* 4.8 Hz, N*H*-3), 3.63 – 3.46 (m, 18H, OC*H*$_2$C*H*$_2$O), 3.41 – 3.34 (m, 4H, OC*H*$_2$C*H*$_2$O), 3.17 – 3.11 (m, 2H, OC*H*$_2$C*H*$_2$O). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 157.1 (NH*C*ONH-2), 132.6 (Aryl *C*), 131.6 (Aryl *C*), 131.1 (Aryl *C*), 128.1 (Aryl *C*), 127.4 (Aryl *C*H), 127.3 (Aryl *C*H), 126.3 (Aryl *C*H), 126.1 (Aryl *C*H), 125.5 (Aryl *C*H), 125.4 (Aryl *C*), 125.1 (Aryl *C*H), 124.9 (Aryl *C*), 124.6 (Aryl *C*H), 123.6 (Aryl *C*), 122.0 (Aryl *C*H), 121.3 (Aryl *C*H), 70.6 (*C*H$_2$), 70.5 (*C*H$_2$), 70.4 (*C*H$_2$), 70.2 (*C*H$_2$), 69.9 (*C*H$_2$), 50.6 (*C*H$_2$), 40.4 (*C*H$_2$). **HR-ESI-MS** Calculated for $C_{29}H_{36}N_5O_6^+$: *m/z* 550.2660 $[M+H]^+$; found 550.2660. **LC-MS** found 550.40, R_t = 0.65. **IR** (Neat, ν_{max}/cm^{-1}) 3287 m (NH urea), 2098 m (N=N=N), 1625 s (C=O urea), 1523–1415 m (C=C aromatic).

5 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

3-(17-amino-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **3.12**

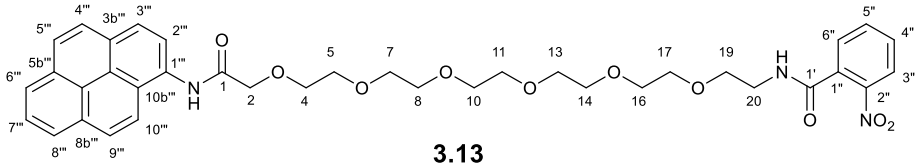


3.12

A solution on triphenylphosphine (490 mg, 0.675 mmol) in THF (3.0 ml) was added to a solution of 20-azido-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.11** (513 mg, 0.533 mmol) in THF (5.0 mL) at 0 °C (ice bath). The reaction was then allowed to warm to room temperature and stirred until determined complete by TLC (24 h). H_2O was added (2.0 mL) and the reaction mixture allowed to stir for 2 h to hydrolyse the iminophosphorane intermediate. The solvent was then removed *in vacuo* and the crude produce purified by flash chromatography (SiO_2; DCM–sat. NH_3 in MeOH 1:0 → 95:5 → 90:10) to yield the product **3.12** as a yellow waxy solid (342 mg, 0.653 mmol, 70%). $\boldsymbol{R_f}$ = 0.38 (SiO_2; DCM–sat. NH_3 in MeOH 9:1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.40 (d, 1H, *J* 8.3 Hz, Aryl), 8.24 (d, 1H, *J* 9.2 Hz, Aryl), 8.16 – 8.02 (m, 4H, Aryl), 8.00 – 7.90 (m, 4H, Aryl + N*H*-3), 6.33 (br t, 1H, *J* 4.5 Hz, N*H*-1), 3.64 – 3.48 (m, 18H, OC*H*$_2$C*H*$_2$O), 3.41 (dd, 2H, *J* 5.8, 3.4 Hz, OC*H*$_2$C*H*$_2$O), 3.29 (t, 2H, *J* 5.2 Hz, OC*H*$_2$C*H*$_2$O), 2.69 (t, 2H, *J* 5.2 Hz, C*H*$_2$-17). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 157.1 (NH*C*ONH-2), 132.8 (Aryl *C*), 131.6 (Aryl *C*), 131.1 (Aryl *C*), 128.0 (Aryl *C*), 127.5 (Aryl *C*H), 127.2 (Aryl *C*H), 126.2 (Aryl *C*H), 126.1 (Aryl *C*H), 125.5 (Aryl *C*H), 125.4 (Aryl *C*), 125.0 (Aryl *C*H), 124.5 (Aryl *C*H), 123.4 (Aryl *C*), 121.9 (Aryl *C*H), 121.4 (Aryl *C*H), 73.1 (*C*H_2), 70.6 (*C*H_2), 70.6 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.2 (*C*H_2), 70.1 (*C*H_2), 41.7 (*C*H_2-17), 40.4 (*C*H_2). **HR-ESI-MS** Calculated for $C_{29}H_{38}N_3O_6^+$: *m/z* 524.2755 $[M+H]^+$; found 524.2757. **LC-MS** found 524.49, $\boldsymbol{R_t}$ **=** 0.55. **IR** (Neat, ν_{max}/cm^{-1}) 3390 m (NH_2), 1625 s (C=O urea), 1524–1416 m (C=C aromatic).

1 × aromatic C and 2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

20-[(2-nitrophenyl)formamido]-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.13**

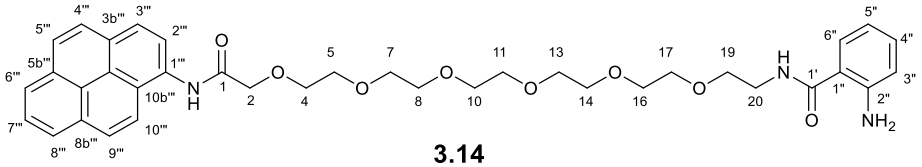


3.13

Under an N_2 atmosphere, hydroxybenzotriazole (176 mg, 1.30 mmol), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (199 mg, 1.04 mmol) and a solution of 20-amino-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.12** (280 mg, 0.520 mmol) in DCM (5.0 mL) were added sequentially to a stirred solution of 2-nitrobenzoic acid **3.6** (96 mg, 0.57 mmol) in DCM (5.0 mL). The reaction was stirred until determined complete by LC-MS (15 h). The solvent removed *in vacuo* and the crude residue was purified by flash chromatography (SiO_2; DCM–MeOH 1:0 → 95:5) to afford the product **3.13** as a brown oil (350 mg, 0.509 mmol, 98%). $\boldsymbol{R_f}$ = 0.32 (SiO_2; DCM–MeOH 95:5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.58 (br s, 1H, CON*H*-1), 8.48 (d, 1H, *J* 8.3 Hz, Aryl), 8.20 – 8.17 (m, 4H, Aryl), 8.12 (d, 1H, *J* 9.2 Hz, Aryl), 8.07 – 7.99 (m, 3H, Aryl), 7.93 (dd, 1H, *J* 8.3, 1.0 Hz, C*H*-3"), 7.50 (td, 1H, *J* 7.6, 1.0 Hz, C*H*-5"), 7.45 (dd, 1H, *J* 7.6, 1.6 Hz, C*H*-6"), 7.39 (ddd, 1H, *J* 8.3, 7.6, 1.6 Hz, C*H*-4"), 6.90 (br s, 1H, CON*H*-1'), 4.41 – 4.32 (m, 2H, OCH_2CH_2O), 3.98 – 3.93 (m, 2H, OCH_2CH_2O), 3.85 – 3.79 (m, 2H, OCH_2CH_2O), 3.67 (dd, 2H, *J* 5.5, 4.0 Hz, OCH_2CH_2O), 3.65 – 3.60 (m, 4H, OCH_2CH_2O), 3.59 – 3.55 (m, 2H, OCH_2CH_2O), 3.53 – 3.48 (m, 2H, OCH_2CH_2O), 3.46 (dd, 2H, *J* 5.5, 3.9 Hz, OCH_2CH_2O), 3.43 – 3.38 (m, 2H, OCH_2CH_2O), 3.30 (dd, 2H, *J* 5.7, 3.7 Hz, OCH_2CH_2O), 3.24 – 3.16 (m, 4H, OCH_2CH_2O). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 169.3 (*C*ONH-1), 166.6 (*C*ONH-1'), 146.6 (*C*-1"), 133.5 (*C*H-5"), 133.2 (*C*-2"), 131.4 (Aryl *C*), 130.9 (Aryl *C*), 130.2 (*C*H-4"), 129.2 (Aryl *C*), 128.9 (*C*H-6"), 127.9 (Aryl *C*H), 127.4 (Aryl *C*H), 127.0 (Aryl *C*H), 126.3 (Aryl *C*H), 125.5 (Aryl *C*H), 125.3 (Aryl *C*H), 125.3 (Aryl *C*), 125.1 (Aryl *C*H), 124.8 (Aryl *C*), 124.4 (*C*H-3"), 123.8 (Aryl *C*), 122.3 (Aryl *C*H), 120.9 (Aryl *C*H), 71.7 (*C*H_2), 71.1 (*C*H_2), 70.8 (*C*H_2), 70.4 (*C*H_2), 70.4 (*C*H_2), 70.4 (*C*H_2), 70.3 (*C*H_2), 70.3 (*C*H_2), 70.2 (*C*H_2), 70.2 (*C*H_2), 69.6 (*C*H_2), 40.1 (*C*H_2). **HR-ESI-MS** Calculated for $C_{37}H_{45}N_4O_{10}^+$: *m/z* 705.3130 $[M+NH_4]^+$; found 705.3135. **LC-MS** found 705.47, $\boldsymbol{R_t}$ **=** 0.65. **IR** (Neat, ν_{max}/cm^{-1}) 3291 m (NH amide), 1659 s (C=O amide), 1523 s (NO_2 Asym), 1347 s (NO_2 Sym), 1554–1438 m (C=C aromatic).

1 × aromatic C and 1 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

20-[(2-aminophenyl)formamido]-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.14**

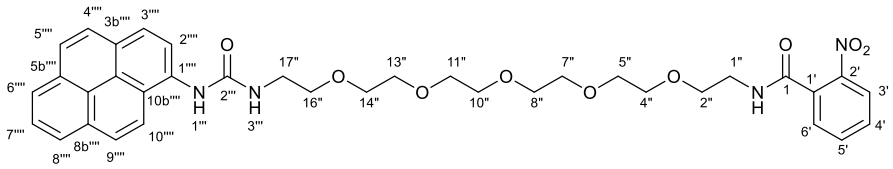


3.14

Fe powder (406 mg, 7.27 mmol) was added to a stirred solution of 20-[(2-nitrophenyl)formamido]-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.13** (100 mg, 0.145 mmol) in glacial acetic acid (5.0 mL) then heated to 60 °C and stirred (2 h). The reaction was then allowed to cool, and the reaction basified (pH 8) using sat. $NaHCO_3$. The reaction mixture was extracted with EtOAc (4 × 30 mL). The organic layers were combined, washed with sat. $NaHCO_3$ (20 mL), brine (10 mL), and dried ($Na2SO_4$). The solvent removed under reduced pressure to afford the product **3.14** as a brown oil (86 mg, 0.13 mmol, 90%). $\boldsymbol{R_f}$ = 0.17 (SiO_2; DCM–MeOH 96:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.61 (br s, 1H, CON*H*-1), 8.48 (d, 1H, *J* 8.2 Hz, Aryl), 8.21 – 8.16 (m, 4H, Aryl), 8.12 (d, 1H, *J* 9.2 Hz, Aryl), 8.04 – 7.99 (m, 3H, Aryl), 7.35 (dd, 1H, *J* 7.9, 1.4 Hz, C*H*-6”), 7.15 (ddd, 1H, *J* 8.2, 7.9, 1.4 Hz, C*H*-4”), 6.83 (br s, 1H, CON*H*-1’), 6.63 (dd, 1H, *J* 8.2, 1.0 Hz, C*H*-3”), 6.59 (td, 1H, *J* 7.9, 1.0 Hz, C*H*-5”), 5.51 (br s, 2H, NH_2), 4.35 (s, 2H, CH_2-2), 3.95 – 3.92 (m, 2H, OCH_2CH_2O), 3.82 – 3.78 (m, 2H, OCH_2CH_2O), 3.70 – 3.61 (m, 4H, OCH_2CH_2O), 3.59 – 3.54 (m, 4H, OCH_2CH_2O), 3.49 – 3.44 (m, 4H, OCH_2CH_2O), 3.42 – 3.40 (m, 2H, OCH_2CH_2O), 3.37 – 3.32 (m, 2H, OCH_2CH_2O), 3.26 – 3.21 (m, 4H, OCH_2CH_2O). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 169.5 (*C*ONH-1’), 169.3 (*C*ONH-1), 148.8 (*C*-1”), 132.2 (*C*H-4”), 131.5 (Aryl *C*), 131.0 (Aryl *C*), 130.2 (Aryl *C*), 129.3 (Aryl *C*), 128.0 (Aryl *C*H), 127.8 (*C*H-6”), 127.5 (Aryl *C*H), 127.0 (Aryl *C*H), 126.3 (Aryl *C*H), 125.5 (Aryl *C*H), 125.4 (Aryl *C*H), 125.3 (Aryl *C*), 125.2 (Aryl *C*H), 124.9 (Aryl *C*), 123.9 (Aryl *C*), 122.4 (Aryl *C*H), 121.0 (Aryl *C*H), 117.3 (*C*H-3”), 116.6 (*C*H-5”), 116.4 (*C*-2”), 71.7 (*C*H_2), 71.2 (*C*H_2), 70.8 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.4 (*C*H_2), 70.3 (*C*H_2), 70.2 (*C*H_2), 70.0 (*C*H_2), 39.5 (*C*H_2). **HR-ESI-MS** Calculated for $C_{37}H_{44}N_3O_8$: 658.3123 *m/z* $[M+H]^+$; found 658.3118. **LC-MS** found 658.47, $\boldsymbol{R_t}$ **=** 0.65. **IR** (Neat, ν_{max}/cm^{-1}) 3434 m (NH_2), 3322 m (NH amide), 1631 m (C=O amide), 1554–1415 m (C=C aromatic).

3 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

2-nitro-*N*-(17-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15-pentaoxaheptadecan-1-yl)benzamide **3.15**

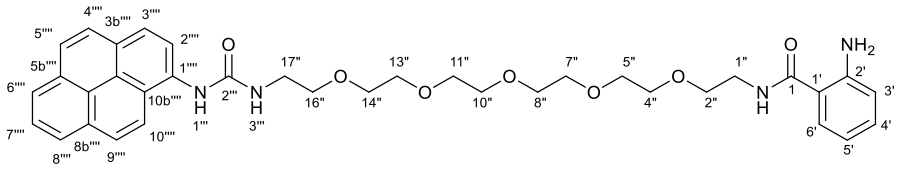


3.15

Hydroxybenzotriazole (73 mg, 0.54 mmol), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (83 mg, 0.43 mmol) and a solution of 3-(17-amino-3,6,9,12,15-pentaoxaheptadecan-1-yl)-1-(pyren-1-yl)urea **3.12** (114 mg, 0.216 mmol) in DCM (2.0 mL) were added sequentially to a stirred solution of 2-nitrobenzoic acid **3.6** (40 mg, 0.24 mmol) in DCM (3.0 mL). The reaction was stirred until determined complete by LC-MS (15 h). The solvent removed *in vacuo* and the crude residue was purified by flash chromatography (SiO_2; DCM–MeOH 95:5) to afford the product **3.15** as a yellow waxy solid (124 mg, 0.184 mmol, 85%). $\boldsymbol{R_f}$ = 0.24 (SiO_2; DCM–MeOH 95:5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.35 (d, 1H, *J* 8.3 Hz, Aryl), 8.18 (d, 1H, *J* 9.3 Hz, Aryl), 8.09 (dd, 1H, *J* 7.5, 0.8 Hz, Aryl), 8.06 – 8.01 (m, 3H, Aryl), 7.96 – 7.88 (m, 4H, Aryl + N*H*-1'"), 7.80 (dd, 1H, *J* 8.1, 0.8 Hz, C*H*-3'), 7.37 (td, 1H *J* 7.3, 0.8 Hz, C*H*-5'), 7.32 (dd, 1H, *J* 7.3, 1.8 Hz, C*H*-6'), 7.27 (ddd, 1H, *J* 8.1, 7.3, 1.8 Hz, C*H*-4'), 7.00 (br t, 1H, *J* 5.2 Hz, CONH-1), 6.28 (br t, 1H, *J* 5.2 Hz, N*H*-3'"), 3.60 – 3.45 (m, 24H, OCH_2CH_2O). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 166.9 (*C*ONH-1), 157.1 (NH*C*ONH-2'"), 146.4 (*C*-1'), 133.4 (*C*H-5'), 133.0 (*C*-2'), 132.8 (Aryl *C*), 131.6 (Aryl *C*), 131.1 (Aryl *C*), 130.0 (*C*H-4'), 128.9 (*C*H-6'), 127.8 (Aryl *C*), 127.4 (Aryl *C*H), 127.1 (Aryl *C*H), 126.1 (Aryl *C*H), 125.5 (Aryl *C*H), 125.3 (Aryl *C*), 125.0 (Aryl *C*H), 124.9 (Aryl *C*), 124.5 (Aryl *C*H), 124.2 (*C*H-3'), 123.1 (Aryl *C*), 121.5 (Aryl *C*H), 121.3 (Aryl *C*H), 70.6 (CH_2), 70.5 (CH_2), 70.5 (CH_2), 70.4 (CH_2), 70.3 (CH_2), 70.2 (CH_2), 70.2 (CH_2), 69.6 (CH_2), 40.4 (CH_2), 40.1 (CH_2). **HR-ESI-MS** Calculated for $C_{36}H_{41}N_4O_9^+$: *m/z* 673.2868 $[M+H]^+$; found 673.2866. **LC-MS** found 673.49, $\boldsymbol{R_t}$ = 0.7. **IR** (Neat, ν_{max}/cm^{-1}) 3279 m (NH urea), 1626 s (C=O urea/amide), 1562–1450 m (C=C aromatic).

1 × aromatic C and 2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

2-amino-*N*-(17-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15-pentaoxaheptadecan-1-yl)benzamide **3.16**



3.16

Fe powder (125 mg, 2.23 mmol) was added to a stirred solution of 2-nitro-*N*-(17-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15-pentaoxaheptadecan-1-yl)benzamide **3.15** (93 mg, 0.14 mmol) in glacial acetic acid (5.0 mL) then heated to 60 °C and stirred (2 h). The reaction was then allowed to cool, and the reaction basified (pH 8) using sat. $NaHCO_3$. The reaction mixture was extracted with EtOAc (4 × 30 mL). The organic layers were combined, washed with sat. $NaHCO_3$ (20 mL), brine (10 mL), and dried ($Na2SO_4$). The solvent removed under reduced pressure to afford the product (88 mg, 0.14 mmol, 99%) $\boldsymbol{R_f}$ = 0.28 (SiO_2; DCM–MeOH 94:6). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.42 (d, 1H, *J* 8.3 Hz, Aryl), 8.26 (d, 1H, *J* 9.2 Hz, Aryl), 8.16 – 8.06 (m, 3H, Aryl), 8.02 – 7.93 (m, 4H, Aryl), 7.93 (br s, 1H, N*H*-1'"), 7.30 (d, 1H, *J* 7.8 Hz, C*H*-6'), 7.14 (td, 1H, *J* 7.8, 1.0 Hz, C*H*-4'), 6.84 (br s, 1H, CON*H*-1), 6.62 (d, 1H, *J* 8.1 Hz, C*H*-3'), 6.56 (t, 1H, *J* 7.5 Hz, C*H*-5'), 6.31 (br s, 1H, NH-3'"), 5.47 (br s, 2H, NH_2), 3.64 – 3.56 (m, 15H, OCH_2CH_2O), 3.53 – 3.48 (m, 4H, OCH_2CH_2O), 3.46 – 3.39 (m, 6H, OCH_2CH_2O). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 169.6 (*C*ONH-1), 157.1 (NH*C*ONH-2'"), 148.7 (*C*-1"), 132.8 (Aryl *C*), 132.2 (*C*H-4"), 131.6 (Aryl *C*), 131.1 (Aryl *C*), 128.1 (Aryl *C*), 127.8 (*C*H-6"), 127.6 (Aryl *C*H), 127.3 (Aryl *C*H), 126.3 (Aryl *C*H), 126.2 (Aryl *C*H), 125.6 (Aryl *C*H), 125.5 (Aryl *C*), 125.1 (Aryl *C*H), 124.9 (Aryl *C*), 124.6 (Aryl *C*H), 123.5 (Aryl *C*), 121.9 (Aryl *C*H), 121.4 (Aryl *C*H), 117.3 (*C*H-3"), 116.7 (*C*H-5"), 116.4 (*C*-2"), 70.6 (*C*H_2), 70.6 (*C*H_2), 70.5 (*C*H_2), 70.4 (*C*H_2), 70.2 (*C*H_2), 70.1 (*C*H_2), 69.9 (*C*H_2), 40.4 (*C*H_2), 39.4 (*C*H_2). **HR-ESI-MS** Calculated for $C_{36}H_{42}N_4NaO_7$: 665.2946 *m/z* $[M+Na]^+$; found 665.2944. **LC-MS** found 665.47, $\mathbf{R_t}$ **=** 0.7. **IR** (Neat, ν_{max}/cm^{-1}) 3457 w (NH2) 3326 m (NH urea), 1626 s (C=O urea/amide), 1545–1459 m (C=C aromatic).

3 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

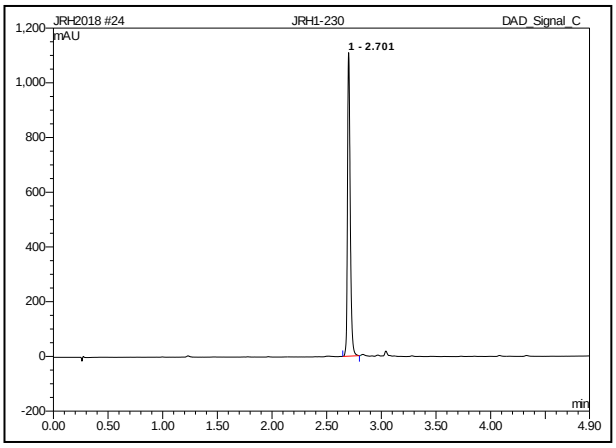
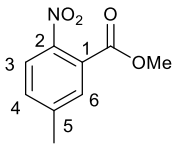


*Figure 3.18: Analytical HPLC trace for 2-amino-*N*-(17-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15-pentaoxaheptadecan-1-yl)benzamide* ***3.16****.*

Methyl 5-methyl-2-nitrobenzoate **3.18** [230]

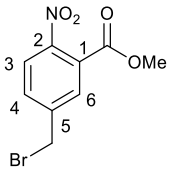


3.18

Conc. sulfuric acid (1.18 mL, 22.1 mmol) was added drop wise to a stirred solution of 5-methyl-2-nitrobenzoic acid **3.19** (2.00 g, 11.0 mmol) in methanol (20 mL) at 0°C (ice bath). The reaction was then heated to reflux and stirred until determined complete by LC-MS (16 h). The reaction was then allowed to cool to room temperature and concentrated *in vacuo.* The residue was then poured into water (40 mL) and extracted with EtOAc (2 × 40 mL). The combined organic layers were washed with sat. $NaHCO_3$ solution (2 × 20 mL), and brine (20 mL), then dried (Na_2SO_4). The solvent was then removed to give the product **3.18** as colorless microcrystals (2.038 g, 10.44 mmol, 95%). $\boldsymbol{R_f}$ = 0.25 (SiO_2; EtOAc– Hex – Toluene 85:10:5). **m.p.**= 78 – 79 °C (EtOAc). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.85 (d, 1H, *J* 8.3 Hz, C*H*-3), 7.47 (d, 1H, *J* 1.2 Hz, C*H*-6), 7.39 (dd, 1H, *J* 8.3, 1.2 Hz, C*H*-4), 3.91 (s, 3H, OCH_3), 2.46 (s, 3H, ArCH_3). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 166.5 (*C*OOMe), 145.8 (*C*-2), 144.7 (*C*-5), 132.0 (*C*H-4), 130.2 (*C*H-6), 128.2 (*C*-1), 124.2 (*C*H-3), 53.4 (O*C*H_3), 21.5 (Ar*C*H_3).

Characterization is consistent with data reported in the literature.[230,231]

Methyl 5-bromomethyl-2-nitrobenzoate **3.17** [231]

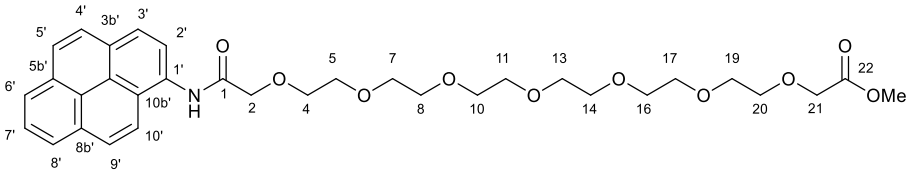


3.17

N-bromosuccinimide (3.65 g, 20.5 mmol) and 70% benzoylperoxide (354 mg, 1.46 mmol) was suspended in a stirred solution of methyl 5-methyl-2-nitrobenzoate **3.18** (2.00 g, 10.2 mmol) in $CHCl_3$ (140 mL). The reaction mixture was then heated to reflux and allowed to stir until determined complete by TLC (21 h). The reaction was cooled to room temperature and the reaction mixture washed with sat. $NaHCO_3$ (2 × 100 mL), H_2O (3 × 100 mL), and brine (50 mL). The organic layer was then dried (Na_2SO_4) and solvent removed *in vacuo*. The crude residue was then purified by flash chromatography (SiO_2; EtOAc–Hex–Toluene 85:10:5) to afford the product **3.17** as a yellow crystalline solid (1.20 g, 4.37 mmol, 43%) $\boldsymbol{R_f}$ = 0.19 (SiO_2; EtOAc– Hex – Toluene 85:10:5).**m.p.**= 59-60 °C ($CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.89 (d, 1H, *J* 8.3 Hz, C*H*-3), 7.74 (d, 1H *J* 2.0 Hz, C*H*-6), 7.63 (dd, 1H, *J* 8.3, 2.0 Hz, C*H*-4), 4.49 (s, 2H, CH_2Br), 3.93 (s, 3H, OCH_3). **^{13}C NMR** (100 MHz, $CDCl_3$) δ 166.5 (*C*OOMe), 147.7 (*C*-2), 143.5 (*C*-5), 132.2 (*C*H-4), 130.5 (*C*H-6), 128.3 (*C*-1), 124.7 (*C*H-3), 53.5 (COO*C*H_3), 30.2 (*C*H_2Br).

Characterization is consistent with data reported in the literature.[231]

Methyl 1-[(pyren-1-yl)carbamoyl]-2,5,8,11,14,17,20-heptaoxadocosan-22-oate **3.22**



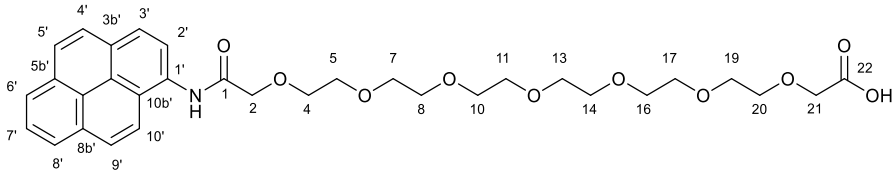
3.22

Methyl bromoacetate **3.21** (125 µL, 1.32 mmol) was added dropwise to a stirred solution of 20-hydroxy-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **2.3** (473 mg, 0.877 mmol) and 60% NaH in mineral oil (53 mg, 1.325 mmol) in DMF (2 mL) and stirred for 30 minutes. The reaction was then allowed to warm to room temperature and stir until determined complete (3 h). The reaction was diluted with EtOAc (20 mL) then washed with 1M LiCl (5 × 15 mL) and brine (10 mL). The organic layer was then dried (Na_2SO_4) and the solvent removed *in vacuo* to give a crude brown oil. The crude product was then purified by flash chromatography (SiO_2; DCM–EtOAc–

MeOH 6:3:1) to yield the product **3.22** as a brown oil (262 mg, 0.428 mmol, 49%). R_f = 0.07 (SiO_2; EtOAc–MeOH 8:2). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.28 (d, 1H, *J* 1.9 Hz, Aryl), 8.26 (d, 1H, *J* 2.1 Hz, Aryl), 8.23 – 8.19 (m, 2H, Aryl), 8.17 – 8.14 (m, 3H, Aryl), 8.11 – 8.07 (m, 2H, Aryl), 5.10 (d, 1H, *J* 17.1 Hz, C*H*$_2$-2), 4.11 (d, 1H, *J* 17.1 Hz, C*H*$_2$-2), 4.00 (d, 1H, *J* 15.7 Hz, C*H*$_2$-21), 3.79 (d, 4H, *J* 16.9 Hz, C*H*$_2$-21 + OC*H*$_3$), 3.68 (dd, 2H, *J* 5.1, 3.8 Hz, OC*H*$_2$C*H*$_2$O), 3.64 – 3.59 (m, 6H, OC*H*$_2$C*H*$_2$O), 3.58 – 3.54 (m, 6H, OC*H*$_2$C*H*$_2$O), 3.53 – 3.49 (m, 4H, OC*H*$_2$C*H*$_2$O), 3.49 – 3.46 (m, 4H, OC*H*$_2$C*H*$_2$O), 3.43 (dd, 2H, *J* 7.2, 3.9 Hz, OC*H*$_2$C*H*$_2$O). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 171.0 1(*C*ONH-1), 169.6 (*C*OOMe-22), 134.1 (Aryl *C*), 131.8 (Aryl *C*), 131.2 (Aryl *C*), 130.9 (Aryl *C*), 129.9 (Aryl *C*H), 128.8 (Aryl *C*H), 128.0 (Aryl *C*), 127.2 (Aryl *C*H), 126.8 (Aryl *C*H), 126.6 (Aryl *C*H), 126.4 (Aryl *C*H), 126.2 (Aryl *C*H), 125.7 (Aryl *C*), 125.6 (Aryl *C*H), 124.6 (Aryl *C*), 121.1 (Aryl *C*H), 72.7 (*C*H$_2$), 70.8 (*C*H$_2$), 70.7 (*C*H$_2$), 70.6 (*C*H$_2$), 70.5 (*C*H$_2$), 70.5 (*C*H$_2$), 70.5 (*C*H$_2$), 70.5 (*C*H$_2$), 70.3 (*C*H$_2$), 69.4 (*C*H$_2$-21), 61.8 (*C*H$_2$), 52.4 (O*C*H$_3$), 51.6 (*C*H$_2$-2). **HR-ESI-MS** Calculated for: 612.2803 *m/z* $[M+H]^+$; found 612.2810. **LC-MS** found, 612.67 R_t = 0.65. **IR** (Neat, ν_{max}/cm^{-1}) 3264 m (NH), 1748 m (C=O ester), 1672 m (C=O amide), 1506–1437 m (C=C aromatic).

2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

1-[(pyren-1-yl)carbamoyl]-2,5,8,11,14,17,20-heptaoxadocosan-22-oic acid **3.25**



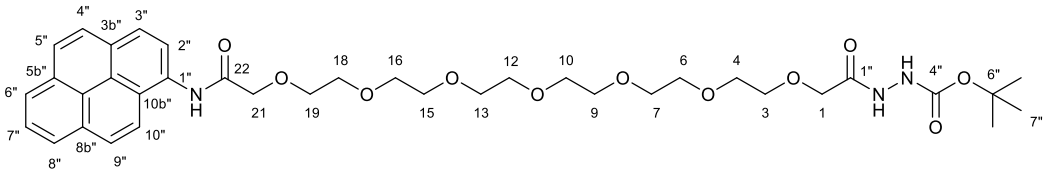
3.25

1 M LiOH (2.0 mL, 1.0 mmol) was added to a stirred solution of methyl 1-[(pyren-1-yl)carbamoyl]-2,5,8,11,14,17,20-heptaoxadocosan-22-oate **3.22** (109 mg, 0.371 mmol) in THF (2 mL) and allowed to stir until determined complete by TLC (2 h). The reaction was then acidified with 1M HCl (1 mL), organic solvent removed under reduced pressure, and the product extracted with EtOAc (5 × 3 mL). The organic layers were combined, washed with brine, and dried (Na_2SO_4). The solvent was removed to yield the product **3.25** as a brown oil (105 mg, 0.176 mmol, 99%). R_f = 0.19 (SiO_2; DCM–MeOH 95:5). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.26 (d, 1H, *J* 3.2 Hz, Aryl), 8.25 (d, 1H, *J* 3.4 Hz, Aryl), 8.23 – 8.18 (m, 3H, Aryl), 8.15 (d, 1H, *J* 9.0 Hz, Aryl), 8.13 (d, 1H, *J* 9.2 Hz, Aryl), 8.10 – 8.05 (m, 2H, Aryl), 5.12 (d, 1H, *J* 17.1 Hz, C*H*$_2$-2), 4.04 (d, 1H, *J* 17.1 Hz, C*H*$_2$-2), 3.99 (d, 1H, *J* 15.4 Hz, C*H*$_2$-21), 3.77 (d, 1H, *J* 15.4 Hz, C*H*$_2$-21), 3.74 (dd, 2H, *J* 5.0, 3.9 Hz, OC*H*$_2$C*H*$_2$O), 3.67 – 3.63 (m, 8H, OC*H*$_2$C*H*$_2$O), 3.63 – 3.57 (m, 8H, OC*H*$_2$C*H*$_2$O), 3.57 – 3.50 (m, 6H, OC*H*$_2$C*H*$_2$O).

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.2 (*C*OOH-22), 171.0 (*C*ONH-1), 134.4 (Aryl *C*), 131.8 (Aryl *C*), 131.2 (Aryl *C*), 130.9 (Aryl *C*), 129.9 (Aryl *C*H), 128.7 (Aryl *C*H), 127.9 (Aryl *C*), 127.2 (Aryl *C*H), 126.8 (Aryl *C*H), 126.6 (Aryl *C*H), 126.3 (Aryl *C*H), 126.2 (Aryl *C*H), 125.7 (Aryl *C*H), 125.6 (Aryl *C*), 124.6 (Aryl *C*), 121.1 (Aryl *C*H), 72.7 (*C*H_2), 70.6 (*C*H_2), 70.6 (*C*H_2), 70.6 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.3 (*C*H_2), 69.4 (*C*H_2-21), 61.6 (*C*H_2), 52.0 (*C*H_2-2). **HR-ESI-MS** Calculated for $C_{32}H_{40}NO_{10}$: 598.2647 *m/z* $[M+H]^+$; found 598.2640. **LC-MS** found 598.35., $\boldsymbol{R_t}$ **=** 0.55. **IR** (Neat, ν_{max}/cm^{-1}) 3439 br m (OH acid), 1734 m (C=O acid), 1670 m (C=O amide), 1507 – 1438 m (C=C aromatic).

3 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

1-{N'-[(tert-butoxy)carbonyl]hydrazinecarbonyl}-*N*-(pyren-1-yl)-2,5,8,11,14,17,20-heptaoxadocosan-22-amide **3.27**



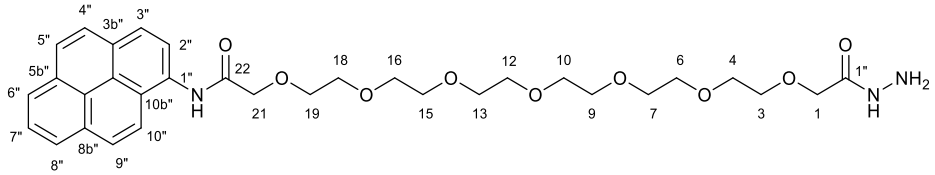
3.27

Hydroxybenzotriazole (11 mg, 0.081 mmol), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (13 mg, 0.068 mmol) and tert-butylcarbazate **3.26** (5 mg, 0.04 mmol) were added sequentially to a stirred solution of 1-[(pyren-1-yl)carbamoyl]-2,5,8,11,14,17,20-heptaoxadocosan-22-oic acid **3.25** (19 mg, 0.032 mmol) in DCM (1 mL). The reaction was stirred until determined complete by TLC (16 h). The reaction mixture was then concentrated *in vacuo* and the crude residue was purified by flash chromatography (SiO_2; DCM–MeOH 9:1) to afford the product **3.27** as a brown oil (19 mg, 0.027mmol, 85%). $\boldsymbol{R_f}$ = 0.43 (SiO_2; DCM–MeOH 9:1). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.71 (br s, 1H, CON*H*-22), 8.28 – 8.25 (m, 2H, Aryl), 8.22 (d, 1H, *J* 2.1 Hz, Aryl), 8.20 (d, 1H, *J* 3.4 Hz, Aryl), 8.17 – 8.15 (m, 2H, Aryl), 8.11 – 8.05 (m, 3H, Aryl), 6.73 (br s, 1H, CON*H*-1"), 5.05 (d, 1H, *J* 14.9 Hz, C*H*$_2$-21), 4.05 (d, 1H, *J* 14.9 Hz, C*H*$_2$-21), 3.99 (d, 1H, *J* 15.8 Hz, C*H*$_2$-1), 3.77 (d, 1H, *J* 15.8 Hz, C*H*$_2$-1), 3.70 – 3.67 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.65 – 3.40 (m, 22H, OC*H*$_2$C*H*$_2$O), 1.54 – 1.42 (m, 9H, OC(C*H*$_3$)$_3$-7").**^{13}C NMR** (125 MHz, $CDCl_3$) δ 172.1 (*C*ONH-1), 168.7(*C*ONHNH-1"), 155.4 (NH*C*OO^{t}Bu-4"), 133.8 (Aryl *C*), 132.0 (Aryl *C*), 131.2 (Aryl *C*), 130.9 (Aryl *C*), 130.1 (Aryl *C*H), 128.8 (Aryl *C*H), 127.6 (Aryl *C*), 127.2 (Aryl *C*H), 126.8 (Aryl *C*H), 126.6 (Aryl *C*H), 126.4 (Aryl *C*H), 126.2 (Aryl *C*H), 125.9 (Aryl *C*H), 125.7 (Aryl *C*), 124.6 (Aryl *C*), 120.8 (Aryl *C*H), 81.7 (O*C*(CH$_3$)$_3$-6"), 72.8 (*C*H_2), 70.8 (*C*H_2), 70.7 (*C*H_2), 70.6 (*C*H_2), 70.6 (*C*H_2), 70.6 (*C*H_2),

70.5 (*C*H_2), 70.5 (*C*H_2), 70.4 (*C*H_2), 69.4 (*C*H_2-1), 61.8 (*C*H_2), 53.1 (*C*H_2-21), 28.4 (OC(*C*H_3)$_3$-7”). **HR-ESI-MS** Calculated for $C_{37}H_{50}N_3O_{11}$: 712.3440 *m/z* $[M+H]^+$; found 712.3441. **LC-MS** Calculated for $C_{37}H_{53}N_4O_{11}$: 729.37 *m/z* $[M+NH_4]^+$; found 729.46, R_t = 0.55. **IR** (Neat, ν_{max}/cm^{-1}) 3439 br m (OH acid), 1734 m (C=O acid), 1670 m (C=O amide), 1507 – 1438 m (C=C aromatic).

2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

1-(hydrazinecarbonyl)-*N*-(pyren-1-yl)-2,5,8,11,14,17,20-heptaoxadocosan-22-amide **3.23**



3.23

Trifluoroacetic acid (300 mL, 60.0 mmol) was added to a stirred solution of 1-{N'-[(tert-butoxy)carbonyl]hydrazinecarbonyl}-*N*-(pyren-1-yl)-2,5,8,11,14,17,20-heptaoxadocosan-22-amide **3.27** (90 mg, 0.010 mmol) in DCM (1 mL) and allowed to stir until determined complete by TLC (2 h). The reaction was diluted with DCM (4 mL), washed with sat. $NaHCO_3$ (3 × 2 mL), brine (2 mL), and dried (Na_2SO_4). The solvent was removed under reduced pressure and the residue purified by reversed-phase chromatography (C_{18}; H_2O (+0.1% CH_3COOH)–MeCN (+0.1% CH_3COOH) 95:5→5:95) to give the product **3.23** as a brown oil (46 mg, 0.076 mmol, 60%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.20 (t, 2H, *J* 6.7 Hz, Aryl), 8.15 (d, 1H, *J* 4.0 Hz, Aryl), 8.13 (d, 1H, *J* 2.8 Hz, Aryl), 8.09 (d, 1H, *J* 8.9 Hz, Aryl), 8.05 – 7.98 (m, 3H, Aryl), 7.95 (d, 1H, *J* 9.1 Hz, Aryl), 4.92 (d, 1H, *J* 15.1 Hz, C*H*$_2$-21), 3.99 (d, 1H, *J* 15.1 Hz, C*H*$_2$-21), 3.91 (d, 1H, *J* 15.6 Hz, C*H*$_2$-1’), 3.68 (d, 1H, *J* 15.6 Hz, C*H*$_2$-1’), 3.64 – 3.59 (m, 2H, OC*H*$_2$C*H*$_2$O), 3.58 – 3.35 (m, 22H, OC*H*$_2$C*H*$_2$O). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 171.7 (*C*ONH-1), 169.1 (*C*ONHNH-1’), 134.0 (Aryl *C*), 132.0 (Aryl *C*), 131.2 (Aryl *C*), 130.8 (Aryl *C*), 130.1 (Aryl *C*H), 128.8 (Aryl *C*H), 127.7 (Aryl *C*), 127.2 (Aryl *C*H), 126.9 (Aryl *C*H), 126.5 (Aryl *C*H), 126.4 (Aryl *C*H), 126.3 (Aryl *C*H), 125.8 (Aryl *C*H), 125.7 (Aryl *C*), 124.6 (Aryl *C*), 120.8 (Aryl *C*H), 72.8 (*C*H_2), 70.8 (*C*H_2), 70.6 (*C*H_2), 70.6 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.5 (*C*H_2), 70.3 (*C*H_2), 69.5 (*C*H_2-1’), 61.7 (*C*H_2), 53.0 (*C*H_2-21). **HR-ESI-MS** Calculated for $C_{32}H_{42}N_3O_9$: 612.2916 *m/z* $[M+H]^+$; found 612.2908. **LC-MS** found 612.35, R_t = 0.50.

2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

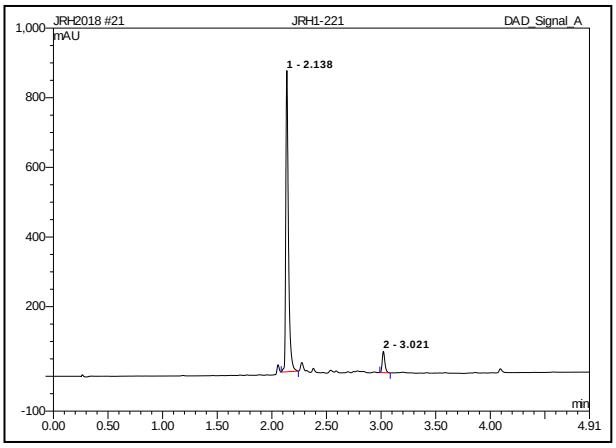
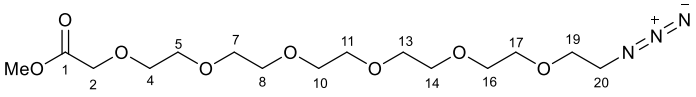


*Figure 3.19: Analytical HPLC trace for 1-(hydrazinecarbonyl)-N-(pyren-1-yl)-2,5,8,11,14,17,20-heptaoxadocosan-22-amide **3.23**.*

Methyl 20-azido-3,6,9,12,15,18-hexaoxaicosanoate **3.31**

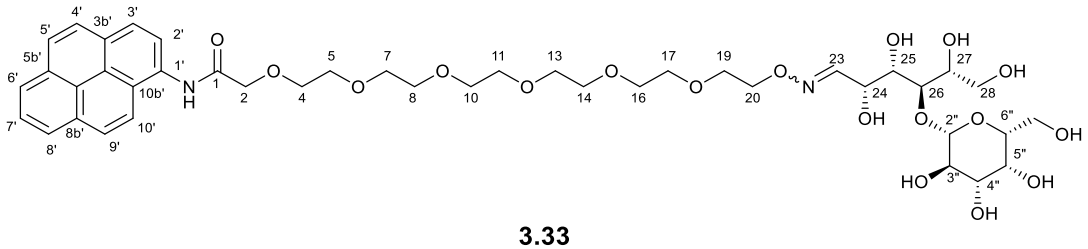


3.31

Methyl bromoacetate **3.21** (205 µL, 2.17 mmol) was added dropwise to a stirred solution of 17-azido-3,6,9,12,15-pentaoxaheptadecan-1-ol **2.7** (600 mg,1.65 mmol) and 60% NaH in mineral oil (150 mg, 3.75 mmol) in DMF (2 mL) and stirred for 10 minutes. The reaction was then allowed to warm to room temperature and stir until determined complete by TLC (18 h). The reaction was diluted with EtOAc (30 mL) and washed 1M LiCl (5 × 5 mL), brine (10 mL), then dried (Na_2SO_4) and the solvent removed *in vacuo* to yield Methyl 20-azido-3,6,9,12,15,18-hexaoxaicosanoate **3.31** as a yellow oil (613 mg, 1.62 mmol, 83%). $\boldsymbol{R_f}$ = 0.23 (SiO_2; EtOAc–MeOH 95:5). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 4.17 (s, 2H, CH_2-2), 3.75 (s, 3H, OCH_3), 3.74 – 3.71 (m, 2H, OCH_2CH_2O), 3.70 – 3.68 (m, 2H, OCH_2CH_2O), 3.67 – 3.64 (m, 18H, OCH_2CH_2O), 3.39 (t, 2H, *J* 5.1 Hz, CH_2-20). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 171.0 (*C*OOMe-1), 71.1 (*C*H_2), 70.8 (*C*H_2), 70.8 (*C*H_2), 70.8 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.7 (*C*H_2), 70.2 (*C*H_2), 68.8 (*C*H_2-2), 51.9 (O*C*H_3), 50.8 (*C*H_2N_3-20). **HR-ESI-MS** Calculated for $C_{15}H_{29}N_3O_8$: 379.1957 *m/z* $[M+Na]^+$; found 379.1957. **LC-MS** found 397.28, $\boldsymbol{R_t}$ = 0.50. **IR** (Neat, ν_{max}/cm^{-1}) 2098 m (N=N=N), 1730 s (C=O ester), 1518–1416 m (C=C aromatic).

2 × CH_2 from the OEG chain were unaccounted for in the ^{13}C NMR spectra due to overlapping signals.

24,25,27,28-tetrahydroxy-*N*-(pyren-1-yl)-26-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}-3,6,9,12,15,18,21-heptaoxa-22-azaoctacos-22-enamide **3.33**



3.33

3, 5-diaminobenzoic acid (31 mg, 0.20 mmol) was added to a stirred solution of lactose (349 mg, 1.02 mmol) and 20-hydroxylamino-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.3** (113 mg, 0.204 mmol) in a 1:1 mixture of MeCN : 0.1 M citrate buffer pH 3 (5 mL) and stirred until confirmed complete by LC-MS (72 h). The solvent was then removed under reduced pressure and purified by reversed-phase high-pressure liquid chromatography (C_{18}; H_2O (+0.1% CH_3COOH)–MeCN (+0.1% CH_3COOH) 95:5→50:50) to afford the product **3.33** as a brown oil (33 mg, 0.038 mmol, 19%) $\boldsymbol{R_f}$ = 0.05 (SiO_2; EtOAc–MeOH 80:20).^{1}**H NMR** (400 MHz, DMSO) δ 10.15 (br s, 1H, CON*H*-1), 8.48 (s, 1H, Aryl), 8.30 (d, 2H, *J* 7.9 Hz, Aryl), 8.26 – 8.19 (m, 3H, Aryl), 8.17 (d, 2H, *J* 0.9 Hz, Aryl), 8.09 (t, 1H, *J* 7.6 Hz, Aryl), 7.42 (d, 1H, *J* 6.9 Hz, RN=C*H*-1") 4.33 (s, 2H, CH_2-2), 4.07 – 4.02 (m, 2H), 3.86 – 3.82 (m, 2H), 3.76 – 3.21 (m, 36H). **HR-ESI-MS** Calculated for $C_{42}H_{58}N_2NaO_{18}^+$: *m/z* 901.3577 $[M+Na]^+$; found 901.3575. **LC-MS** Calculated for $C_{42}H_{59}N_2O_{18}^+$: *m/z* 879.38 $[M+H]^+$; found 879.43.

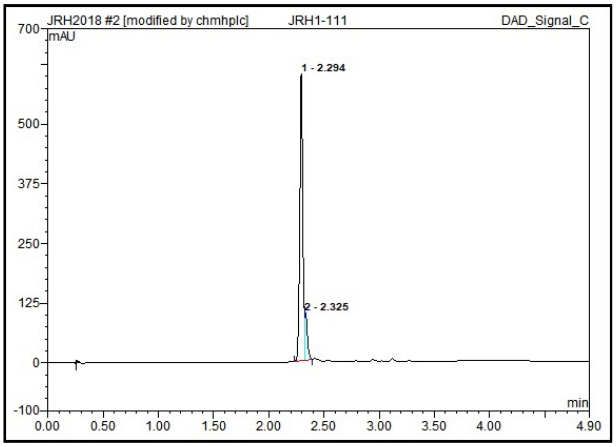
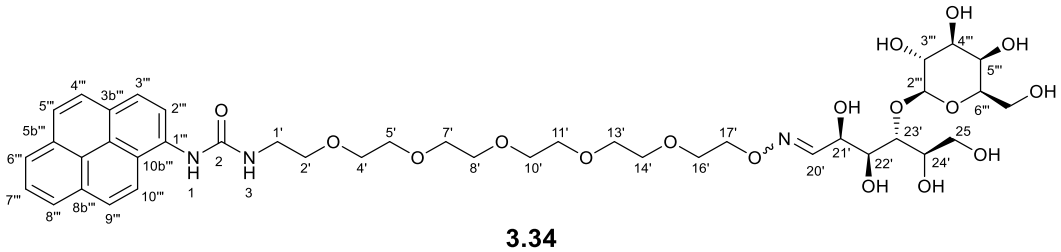


*Figure 3.20: Analytical HPLC trace for 24,25,27,28-tetrahydroxy-*N*-(pyren-1-yl)-26-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}-3,6,9,12,15,18,21-heptaoxa-22-azaoctacos-22-enamide **3.33**.*

1-(pyren-1-yl)-3-[(21,22,24,25-tetrahydroxy-23-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}-3,6,9,12,15,18-hexaoxa-19-azapentacos-19-en-1-yl]urea **3.34**



3.34

3, 5-diaminobenzoic acid (19 mg, 0.13 mmol) was added to a stirred solution of lactose (190 mg, 0.555 mmol) and 3-[17-(aminooxy)-3,6,9,12,15-pentaoxaheptadecan-1-yl]-1-(pyren-1-yl)urea **3.5** (60 mg, 0.11 mmol) in a 1:1 mixture of MeCN : 0.1 M citrate buffer pH 3 (0.5 mL) and stirred until confirmed complete by LC-MS (24 h). The solvent was then removed under reduced pressure and purified by reversed-phase high-pressure liquid chromatography (C_{18}; H_2O (+0.1% CH_3COOH)–MeCN (+0.1% CH_3COOH) 95:5→50:50) to afford the product **3.34** as brown waxy solid (4.6 mg, 0.05 mmol, 5%). **HR-ESI-MS** Calculated for $C_{41}H_{57}N_3NaO_{17}^+$: *m/z* 886.3580 $[M+Na]^+$; found 886.3580. **LC-MS** Calculated for $C_{41}H_{58}N_3O_{17}^+$: *m/z* 864.38 $[M+H]^+$; found 864.25.

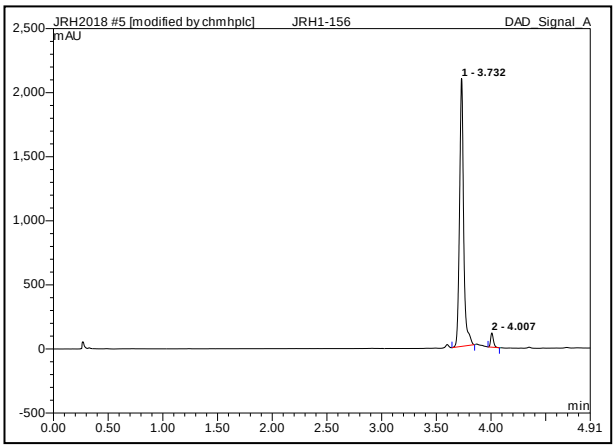
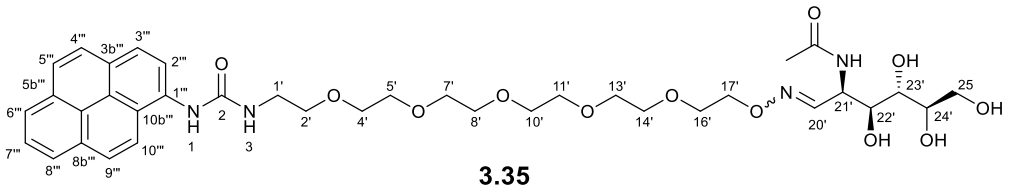


*Figure 3.21: Analytical HPLC trace for 1-(pyren-1-yl)-3-[(21,22,24,25-tetrahydroxy-23-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}-3,6,9,12,15,18-hexaoxa-19-azapentacos-19-en-1-yl]urea **3.34**.*

N-[22,23,24,25-tetrahydroxy-1-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15,18-hexaoxa-19-azapentacos-19-en-21-yl]acetamide **3.35**



3.35

3, 5-diaminobenzoic acid (31 mg, 0.20 mmol) was added to a stirred solution of *N*-acetylglucosamine (204 mg, 0.922 mmol) and 3-[17-(aminooxy)-3,6,9,12,15-pentaoxaheptadecan-1-yl]-1-(pyren-1-yl)urea **3.5** (100 mg, 0.185 mmol) in a 1:1 mixture of MeCN : 0.1 M citrate buffer pH 3 (0.5 mL) and stirred until confirmed complete by LC-MS (24 h). The solvent was then removed under reduced pressure and purified by reversed-phase high-pressure liquid chromatography (C_{18}; H_2O (+0.1% CH_3COOH)–MeCN (+0.1% CH_3COOH) 95:5→50:50) to afford the product **3.35** as colourless waxy solid (46.4 mg, 0.063 mmol, 34%). **HR-ESI-MS** Calculated for $C_{37}H_{51}N_4O_{12}^+$: *m/z* 743.3498 $[M+Na]^+$; found 743.3500. **LC-MS** found 743.36.

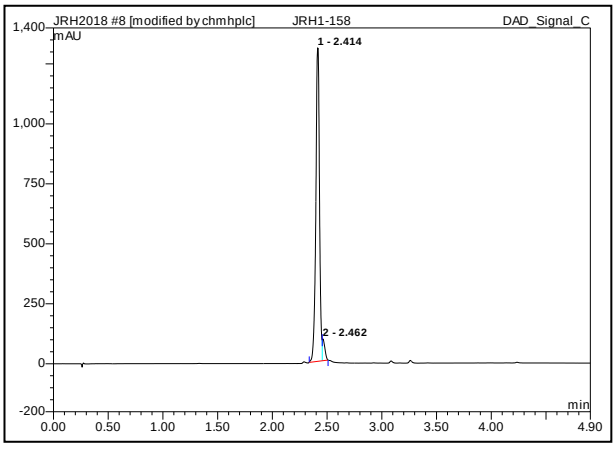
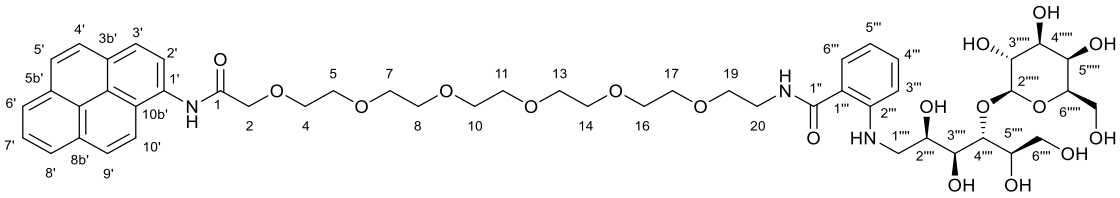


Figure 3.22: Analytical HPLC trace for N-*[22,23,24,25-tetrahydroxy-1-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15,18-hexaoxa-19-azapentacos-19-en-21-yl]acetamide* ***3.35****.*

N-(pyren-1-yl)-20-[(2-{[2,3,5,6-tetrahydroxy-4-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}hexyl]amino}phenyl)formamido]-3,6,9,12,15,18-hexaoxaicosanamide **3.36**



3.36

Lactose (21 mg, 0.061 mmol) was dissolved in a stirred solution of 20-[(2-aminophenyl)formamido]-*N*-(pyren-1-yl)-3,6,9,12,15,18-hexaoxaicosanamide **3.14** (20 mg, 0.030 mmol) and $NaBH_3CN$ (5.5 mg, 0.09 mmol) in 30% AcOH in DMSO (86.8 µL) then heated to 60 °C and stirred (2h). The reaction mixture was then allowed to cool and purified by reversed-phase high-pressure liquid chromatography (C_{18}; H_2O (+0.1% CH_3COOH)–MeCN (+0.1% CH_3COOH) 95:5→5:95) to afford the product **3.36** as a white solid (10.2 mg, 0.010 mmol, 34%). **HR-ESI-MS** Calculated for $C_{49}H_{66}N_3O_{18}^+$: *m/z* 984.4336 $[M+H]^+$; found 984.4335. **LC-MS** found 984.39.

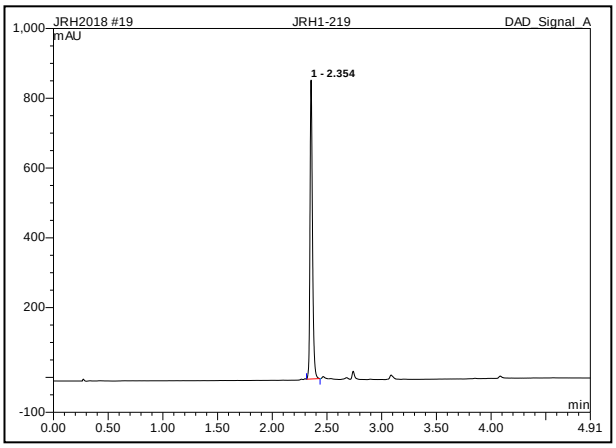
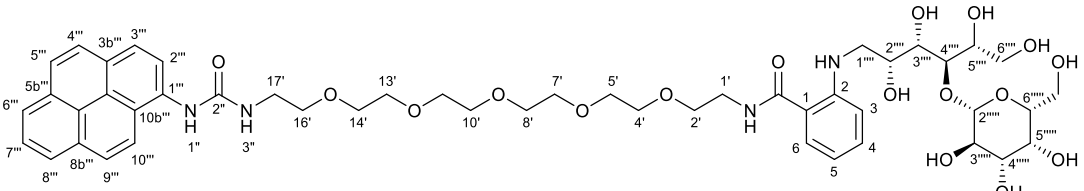


Figure 3.23: Analytical HPLC trace for N-*(pyren-1-yl)-20-[(2-{[2,3,5,6-tetrahydroxy-4-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}hexyl]amino}phenyl)formamido]-3,6,9,12,15,18-hexaoxaicosanamide* ***3.36****.*

N-(17-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15-pentaoxaheptadecan-1-yl)-2-{[2,3,5,6-tetrahydroxy-4-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}hexyl]amino}benzamide **3.37**



3.36

Lactose (7.2 mg, 0.02 mmol) was dissolved in a stirred solution of 2-amino-*N*-(17-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15-pentaoxaheptadecan-1-yl)benzamide **3.16** (6.7 mg, 0.01 mmol) and $NaBH_3CN$ (1.9 mg, 0.31 mmol) in 30% AcOH in DMSO (30 µL) then heated to 60 °C and stirred (2h). The reaction mixture was then allowed to cool and purified by reversed-phase high-pressure liquid chromatography (C_{18}; H_2O (+0.1% CH_3COOH)–MeCN (+0.1% CH_3COOH) 95:5→5:95) to afford the product **3.37** as a white solid (5.0 mg, 0.01 mmol, 68%). **HR-ESI-MS** Calculated for $C_{48}H_{65}N_4O_{17}^+$: *m/z* 969.4339 $[M+H]^+$; found 984.4335. **LC-MS** found 969.51.

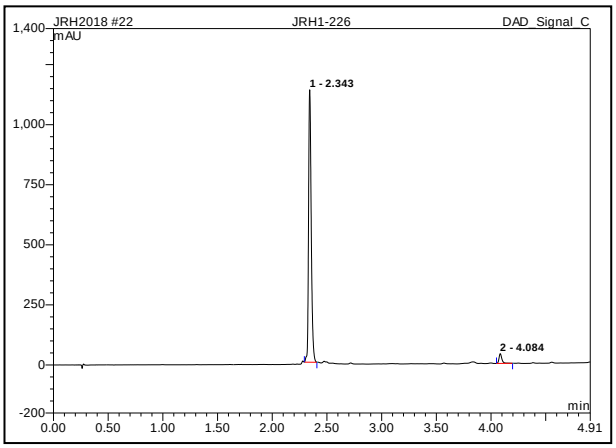
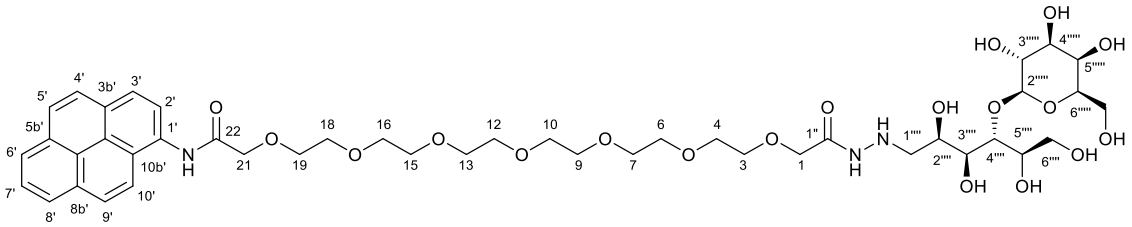


Figure 3.24: Analytical HPLC trace for N-*(17-{[(pyren-1-yl)carbamoyl]amino}-3,6,9,12,15-pentaoxaheptadecan-1-yl)-2-{[2,3,5,6-tetrahydroxy-4-{[3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl]oxy}hexyl]amino}benzamide **3.37**.*

N-(pyren-1-yl)-1-{N'-[2,3,5,6-tetrahydroxy-4-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}hexyl]hydrazinecarbonyl}-2,5,8,11,14,17,20-heptaoxadocosan-22-amide **3.38**



3.38

Lactose (25 mg, 0.07 mmol) was dissolved in a stirred solution of 1-(hydrazinecarbonyl)-*N*-(pyren-1-yl)-2,5,8,11,14,17,20-heptaoxadocosan-22-amide **3.23** (24 mg, 0.04 mmol) and $NaBH_3CN$ (5.9 mg, 0.09 mmol) in 30% AcOH in DMSO (47 µL) then heated to 60 °C and stirred (2h). The reaction mixture was then allowed to cool and subject to reversed-phase high-pressure liquid chromatography (C_{18}; H_2O (+0.1% CH_3COOH)–MeCN (+0.1% CH_3COOH) 95:5→5:95). Purification was incomplete and **3.38** was isolated as an impure compound (18.9 mg) and analysed without further purification. **HR-ESI-MS** Calculated for $C_{44}H_{64}N_3O_{19}^+$: *m/z* 938.4129 $[M+H]^+$; found 938.4134. **LC-MS** found 938.49.

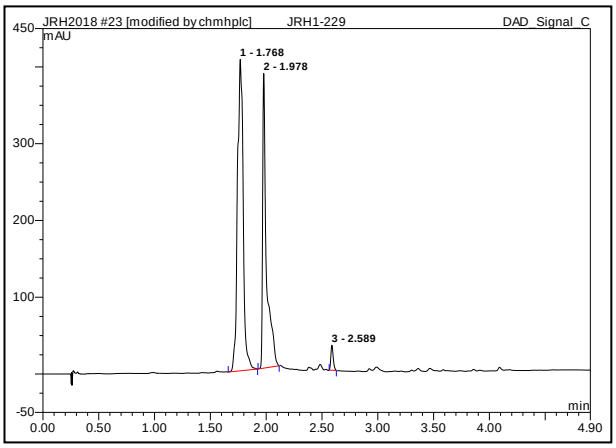


Figure 3.25: Analytical HPLC trace for N*-(pyren-1-yl)-1-{N'-[2,3,5,6-tetrahydroxy-4-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}hexyl]hydrazinecarbonyl}-2,5,8,11,14,17,20-heptaoxadocosan-22-amide* ***3.38****.*

3.8.3 *In situ* labelling procedures

3.8.3.1 In situ labelling of lactose from milk with LALDI tag **3.5** [158]

3, 5-diaminobenzoic acid (3 mg, 0.02 mmol) was added to a stirred solution of 3-[17-(aminooxy)-3,6,9,12,15-pentaoxaheptadecan-1-yl]-1-(pyren-1-yl)urea **3.5** (10 mg, 0.02 mmol) in a 1:1 mixture of cow's milk: 0.1 M citrate buffer pH 3 (0.5 mL) and stirred (12 h). The mixture was then diluted 10-fold, subjected to centrifugation (10000 g, 30 s) and the supernatant analysed directly by LALDI-MS.

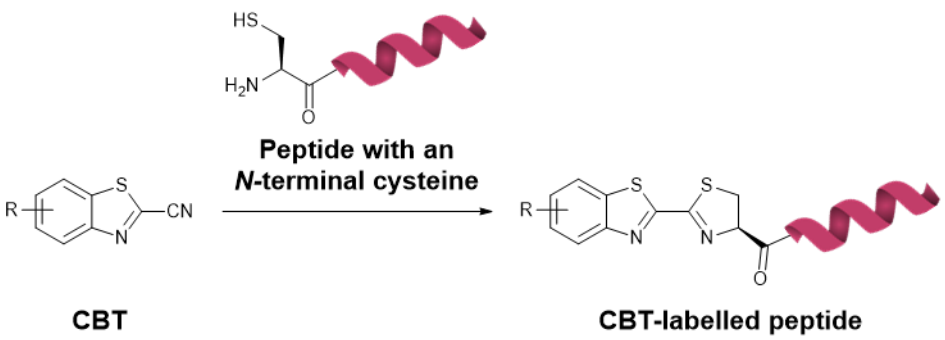
3.8.3.2 In situ labelling of lactose from milk with LALDI tag **3.16** [175]

Cow's milk and a solution of **3.16** (0.7 M) and $NaBH_3CN$ (2 M) in either 30% AcOH in DMSO or 30% AcOH in H_2O were combined in a 1:1 mixture then heated at 60 °C with stirring for 2 h. The mixture was then diluted 10-fold, subjected to centrifugation (10000 × g, 30 s), and the supernatant analysed directly by LALDI-MS.

Chapter 4 – Scalable synthesis of 6-amino-2-cyanobenzothiazole

4.1 Introduction

2-cyanobenzothiazoles (CBTs) are known to react rapidly and selectively with 1,2-aminothiols under physiological conditions, even in the presence of other thiols and amines.[251–255] This bioorthogonal selectivity has allowed the CBT reactivity to be exploited in the development of fast bioorthogonal reactions (k ~10 M^{-1} s^{-1}), known as CBT ligations. This chemistry has been demonstrated as especially useful for site-specific labelling or immobilization of peptides or proteins via a 1,2-aminothiol moiety, such as an *N*-terminal cysteine or non-natural amino acid residue (*Scheme 4.1*).[255,256]

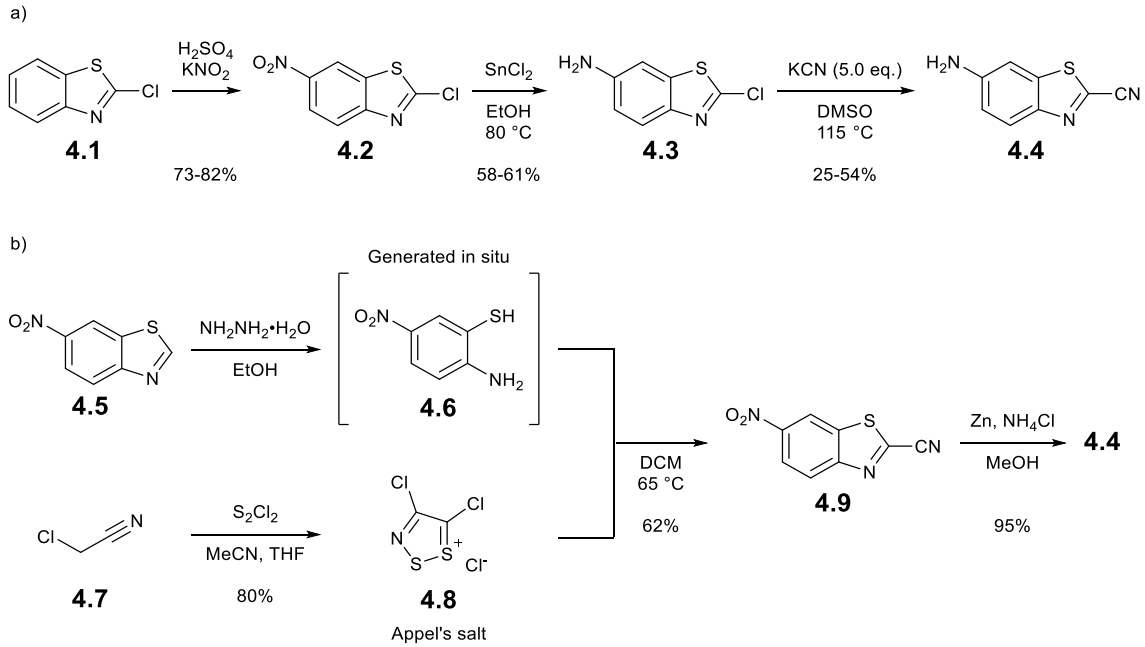


Scheme 4.1: A schematic representation of a bioorthogonal CBT ligation reaction used to achieve specific labelling/immobilization of peptides and proteins, illustrated using a functionalized CBT and a peptide with an N-terminal cysteine.

CBT-functionalized bioorthogonal probes have previously been explored within the Bon group as tools for both site-selective modification of proteins and immobilization of peptide.[257,258] As such, it was proposed that a LALDI tag functionalized with a CBT reactive handle could be used to selectively label and analyze 1,2-aminothiols by LALDI-MS directly from a biological background.

6-amino-2-cyanobenzothiazole (ACBT) **4.4** is considered a useful building block for generating reactive handles for CBT ligation, enabling straight forward incorporation of a CBT handle into a chemical probe via derivatization of the amino group.[253,256,259] Commercial sources of ACBT **4.4** are available, but the relatively high cost of this reagent can limit its applications in studies that require larger quantities. Therefore, a readily scalable synthetic route to ACBT **4.4** was considered to be advantageous. Synthesis of ACBT **4.4** had previously been achieved in the Bon group in 3 steps from 2-chlorobenzothiazole **4.1** by adapting literature procedures (*Scheme*

4.2a).[253,260–262] However, the final step of this route, cyanation of amine **4.3** in dimethylsulfoxide (DMSO), required elaborate extractive workups, produced low and variable yields, and required extensive safety measures as a consequence of the large amounts of KCN in the skin-permeable DMSO at high temperatures.[253,263] Therefore, this route was considered unsuitable for achieving a safe and straightforward scale-up under laboratory conditions.



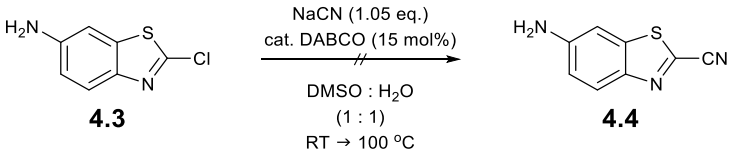
*Scheme 4.2: Previously established synthetic routes to ACBT **4.4**. a) Original route reported by Takakura et al.[263] and Wang* et al.[253] *used previously by the Bon group. b) Improved route reported by McCutcheon et al.*[264]

An alternative synthesis of ACBT **4.4** was reported by McCutcheon et al. that avoids the use of free cyanide by using 4,5-dichloro-1,2,3-dithiazolium chloride **4.8** (Appel's salt) (*Scheme 4.2b*).[264] Prescher and colleagues demonstrated the value and versatility of the synthetic route by using it to generate various different CBT derivatives, some of which on multi-gram scale.[264–267] However, Appel's salt **4.8** is a relatively expensive reagent that can only be obtained from a limited number of commercial suppliers. Appel's salt **4.8** can be generated synthetically, although this requires the use of sulfur monochloride, a reagent significantly more toxic than KCN.[268] These factors could potentially limit the scale-up of McCutcheon's procedure in a standard laboratory environment. These issues, therefore, prompted the search for an alternative synthetic route.

4.2 Synthesis of ACBT **4.4**

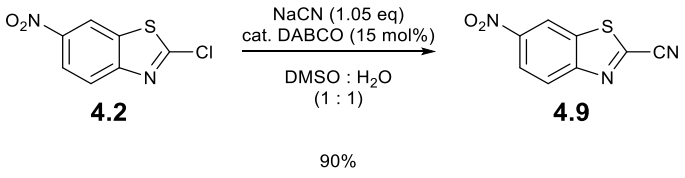
4.2.1 Alternative cyanation

Inspired by a report on 1,4-diazabicyclo[2.2.2]octane (DABCO)-catalyzed cyanations of heteroaryl halides by Svirdov et al.,[269] the applicability of DABCO as a catalyst for the cyanation of 2-chlorobenzothiazoles was investigated.



Scheme 4.3: *Attempted synthesis of 6-amino-1,3-benzothiazole-2-carbonitrile **4.4** using NaCN and DABCO.*[269]

Treatment of amine **4.3** with DABCO (15 mol%) and NaCN (1.05 equiv.) in DMSO : water (1:1) did not result in formation of ACBT **4.4**. Instead, only starting material was isolated. The reaction was repeated with the temperature incrementally increased from room temperature to 100 °C over time. As before the reaction failed to produce any of the product (*Scheme 4.3*). The failure of these reactions was postulated to be due to amino group of **4.3** deactivating the benzothiazole ring to nucleophilic aromatic substitution (S_NAr) at the chloride. This deactivation also explains why the high temperatures and excess of potassium cyanide were required for the cyanation of **4.3** in *Scheme 4.2*.



Scheme 4.4: *DABCO catalyzed cyanation of 2-chloro-6-nitro-1,3-benzothiazole **4.2** with NaCN in DMSO:H2O.*[269]

In an attempt to activate the benzothiazole ring for S_NAr at the chloride, the reaction was attempted using 6-nitro-2-chlorobenzothiazole **4.2** at room temperature under the same conditions.[269] The reaction was found to successfully convert chloride **4.2** to 6-nitro-2-cyanobenzothiazole **4.9**, a known precursor of ACBT **4.4**,[264] in 90% yield. Any unreacted cyanide in the reaction mixture was safely quenched by the addition of an iron (III) chloride solution (3 mM). It is thought that DABCO aids S_NAr through nucleophilic catalysis (*Figure 4.1*). The sterically unhindered amine centers of DABCO make it a strong nucleophile and allow it to

substitute the chloride more easily than the free cyanide ions. The resulting tertiary amine that is generated is a much more reactive intermediate that can be substituted more easily by the free cyanide ions, resulting in the simultaneous formation of the cyanated product **4.9** and regeneration of the DABCO catalyst.

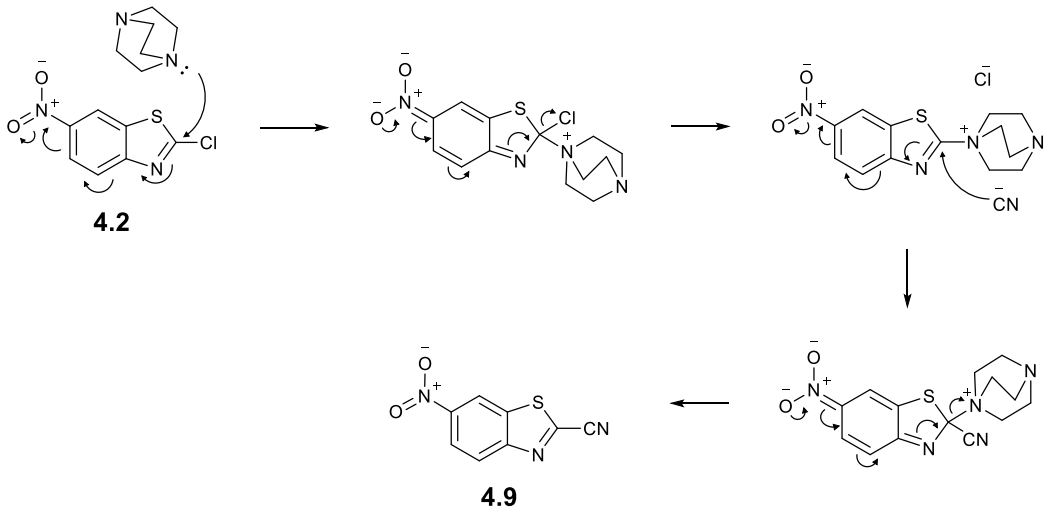
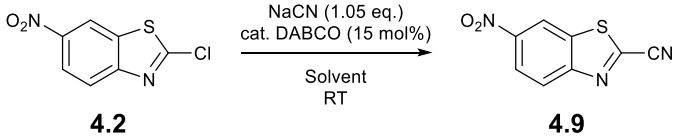


*Figure 4.1: Proposed mechanism for DABCO catalyzed S_NAr of chloride **4.2**.*

A control experiment was carried out without DABCO at room temperature. However, no conversion of the starting material **4.2** was observed. This validated that DABCO acts as a catalyst in the cyanation of **4.2** and was necessary for the reaction to proceed under these conditions.

4.2.1.1 Solvent optimization

As DMSO can quickly penetrate the skin, working with NaCN in DMSO required substantial safety measures. Furthermore, the high boiling point and aqueous solubility of DMSO lead to persistent complications during work up and purification. To address these issues, a solvent screen was carried out to find a suitable alternative and optimize the solvent system. Additionally, a parallel investigation was carried out to determine if the order of addition of reagents held any significance to the success of the reaction. Investigations were conducted via a series of 10 mg microscale reactions, and the progress of each reaction was monitored using LC-MS. The results from these experiments revealed that the DABCO-catalyzed cyanation proceeded under the same reaction conditions in a range of solvents and that the order of addition was inconsequential to the success of the reaction (*Table 4.1*).



*Scheme 4.5: Conditions for the microscale solvent screen of DABCO-catalyzed cyanation of 2-chloro-6-nitrobenzothiazole **4.2**.*

Table 4.1: Conditions and observed outcome of the DABCO-catalyzed cyanation solvent screen.

Reaction	Solvent System	Ratio	Order of Addition	Conversion of 4.2 to 4.9	Side Product
A	DMSO : H_2O	1 : 1	DABCO, NaCN	Yes	No
B	DMSO : H_2O	1 : 1	NaCN, DABCO	Yes	No
C	H_2O	-	DABCO, NaCN	Yes	Hydroxy adduct **4.10**
C	H_2O	-	DABCO, NaCN	Yes	Hydroxy adduct **4.10**
D	MeCN	-	DABCO, NaCN	Yes	No
E	MeCN	-	DABCO, NaCN	Yes	No
F	MeCN : H_2O	1 : 1	DABCO, NaCN	No	No
G	MeCN : H_2O	1 : 1	DABCO, NaCN	No	No
H	EtOH	-	DABCO, NaCN	Yes	Ethoxy adduct **4.11** Ethyl ester **4.12**
I	EtOH	-	DABCO, NaCN	Yes	Ethoxy adduct **4.11** Ethyl ester **4.12**
J	EtOH : H_2O	1 : 1	DABCO, NaCN	No	No
K	EtOH : H_2O	1 : 1	DABCO, NaCN	No	No

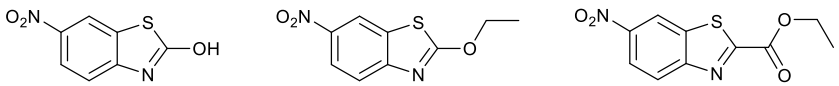
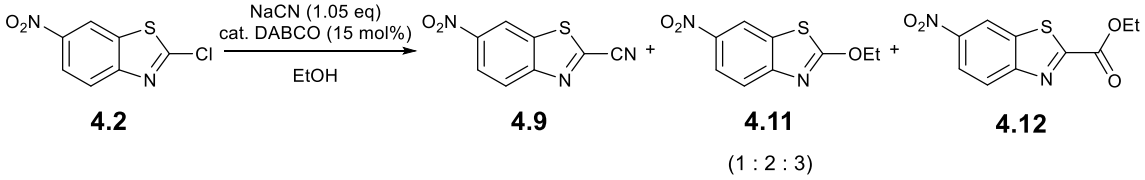


Figure 4.2: Structures of observed side products formed during microscale solvent screen

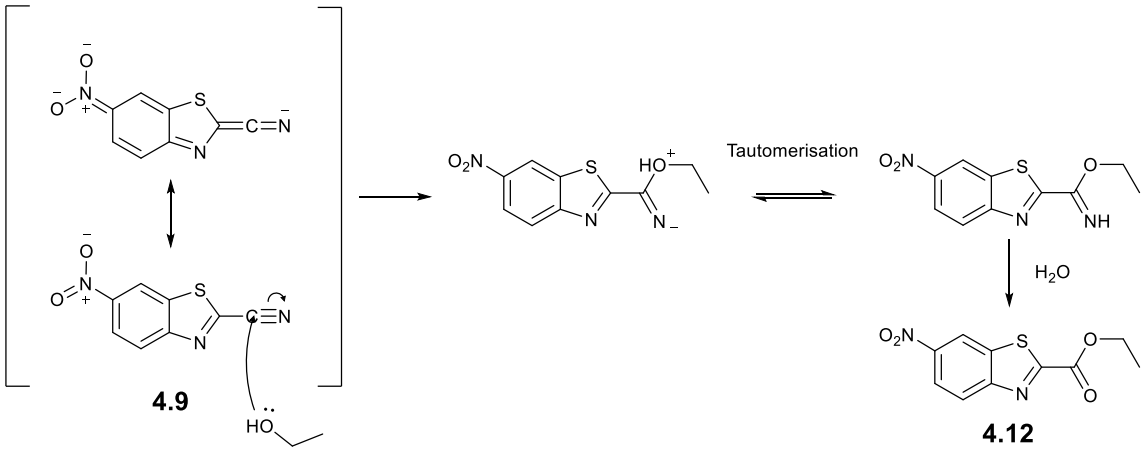
The results of the solvent screen highlighted water, ethanol and acetonitrile as potential candidates for a safer alternative to DMSO. However, due to the poor aqueous solubility of **4.2**, the reactions in water without an organic co-solvent resulted in a biphasic reaction mixture, poor conversion, and the formation of an unwanted hydroxyl adduct **4.10**. Therefore, further investigations were focused on reactions in organic solvents with water used as a co-solvent to assist dissolution of sodium cyanide.

The reactions in ethanol and acetonitrile were scaled up to a 200 mg scale to analyze their respective efficiency and ease of work up. In ethanolic solvents (*Scheme 4.6*), the reaction gave a 1 : 2 : 3 ratio of the desired product **4.9** and the unwanted side products **4.11** and **4.12**, respectively. Ethyl ester **4.12** was suspected to be the structure of the third compound.



*Scheme 4.6: DABCO catalyzed cyanation of 2-chloro-6-nitro-1,3-benzothiazole **4.2** with NaCN in ethanol.*

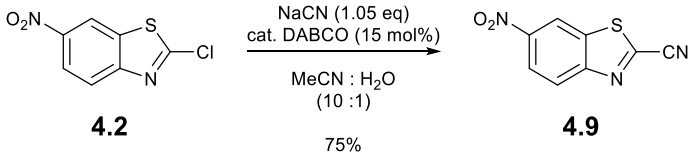
It was proposed that nitrile **4.9**, activated by the electron withdrawing nitro group, underwent a Pinner reaction with ethanol to give a carboximidate intermediate which then hydrolyzed to give ethyl ester **4.12** (*Scheme 4.7*).[270] The presence of the desired product and unwanted side products was confirmed by comparing the observed mass in the LC-MS and ^{1}H-NMR spectra of the crude product to the literature data of each compound.[269,271] The ratio of compounds present in the crude mixture was determined using two separate methods. Firstly, by integrating the absorptions for each compound in the LC-MS UV trace, then by integrating the equivalent protons for each compound in the ^{1}H-NMR spectra of the crude product. Both techniques indicated a 1 : 2 : 3 ratio of **4.9** : **4.11** : **4.12**.



*Scheme 4.7: Proposed mechanism of the Pinner reaction. Nucelophilic attack of the nitrile of **4.9** by ethanol leads to carboximidate intermediate, which is then hydrolysed to give ethyl ester **4.12**.*[270]

The most practical solvent system was found to be acetonitrile with a small volume of water (10%) to facilitate dissolution of sodium cyanide (*Scheme 4.8*). Once again, any unreacted cyanide was quenched by the addition of a 3 mM $FeCl_3$ solution. This method afforded the product in 75% yield and removed the need for DMSO. Immediate work up of the reaction upon completion, followed by the swift extraction of **4.9** from the ferric chloride solution was found to be important for preventing the hydrolysis of the cyano group of **4.9** under dilute aqueous conditions. Following the extractive work up, ^{1}H-NMR showed that the product was pure enough

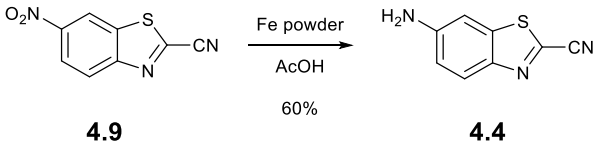
to be taken on without any need for further purification. However, the product could be subjected to filtration through a short plug (SiO_2; $CHCl_3$) to remove trace amounts of iron and gain an analytically pure sample without any loss of product. Recrystallisation of pure **4.9** could be achieved from methanol for melting point determination. While the reaction in acetonitrile/water afforded a lower yield than in DMSO/water, the purification of **4.9** no longer involved the resolution of complicated emulsions or extensive evaporation to remove persistent traces of solvent. Therefore, the benefits of the increased safety and simplified work up outweighed the reduced yield.



*Scheme 4.8: DABCO catalyzed cyanation of **4.2** with NaCN in MeCN.*

4.2.2 Nitro reduction

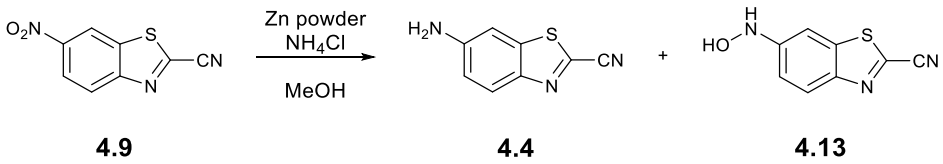
The synthesis of ACBT **4.4** was completed by the reduction of the nitro group of **4.9** (*Scheme 4.9*). The reduction of **4.9** was attempted through both a zinc- and an iron-mediated reduction. However, it was found that iron in acetic acid proved to be more effective than the literature reported Zn/NH_4Cl reduction. The iron reduction was performed on a 100 mg and 500 mg scale and consistently afforded ACBT **4.4** in 60% yield with no side the product formation or complications during work up. Purification of **4.4** was achieved by filtration of the crude product through a short plug (SiO_2; $CHCl_3$). Recrystallisation of pure **4.4** could be achieved from ethanol for melting point determination.



Scheme 4.9: Reduction of **4.9** using iron in acetic acid.[229]

Despite reports of quantitative yields and straightforward work up conditions by Prescher et al.,[264] success of the Zn/NH_4Cl reduction was hampered by the formation of a side products and persistent complications during work up. Correlations between the LC-MS analysis of the crude product and earlier reports by Prescher et al. indicated that reaction failed to go to completion and that the side product formed was an intermediate of the reduction, hydroxyl amine **4.13** (*Scheme 4.10*).[265] It was also postulated that the use of a higher-grade zinc powder or zinc

micropowder could have helped drive the reaction to completion. However, no information was given in the literature by Prescher et al. about the particle size, quality, or commercial source of the zinc reagent used in their reduction.[264,265] Therefore, following the success of the iron reduction of **4.9**, no further investigations into the zinc mediated reduction were pursued.



4.9 **4.4** **4.13**

Scheme 4.10: *Observed outcome by Prescher et al. of the Zn/NH_4Cl reduction of 6-nitro-1,3-benzothiazole-2-carbonitrile **4.9**.*[265]

4.2.3 Scale up of DABCO-catalyzed cyanation

4.2.3.1 Calorimetry

Before scale-up of the cyanation was carried out, calorimetry was performed to monitor the thermodynamics of the reaction. This precaution was taken to highlight any unexpected and potentially dangerous energy spikes that could occur during the scale-up. Calorimetry experiments were performed with the assistance of Dr Mary Bayana and Dr Kathrine Jolley. The calorimeter reactor temperature was maintained at a constant 21 °C through the use of a calorimeter jacket temperature set at 1 °C, and power compensation by an internal heater coil. In the experiment, slow addition of aqueous NaCN (49 mg in 2 mL) to a solution of **4.2** (200 mg) and DABCO (15 mol%) in acetonitrile (20 mL) resulted in an endotherm (*Figure 4.3*). It was noted that the dissolution of **4.2** in acetonitrile only occurred following the addition of DABCO.

Taking into account the temperature and specific heat capacity of the feed, it was calculated that an energy consumption of 0.42 kJ was determined for the addition of ~1 mmol of NaCN in 2 mL of water. A control experiment, in which only water was added to **4.2** and DABCO in acetonitrile, showed an endotherm of 0.51 kJ, corresponding to a calculated heat of mixing of water into the reaction mixture of ΔH = 4.57 kJ mol^{-1}. From these two experiments it was estimated that the energy output generated by the reaction was −0.09 KJ, presenting an enthalpy of −90 kJ mol^{-1} (*Equation 4.1*).

$$E_{(NaCN)} = E_{(NaCN\ +water)} - E_{(water)} = 0.42\ kJ - 0.51\ kJ = -0.09\ kJ$$
$$\Delta H_{(NaCN)} = E_{(NaCN)} \div n_{(NaCN)} = -0.09\ kJ \div 0.001\ mol = -90\ kJ\ mol^{-1}$$

Equation 4.1: Molar heat calculation for addition of NaCN to the reaction mixture. $E_{(NaCN)}$ = Total energy released from addition of NaCN (calculated). $E_{(NaCN\ +\ water)}$ = Total energy released from addition of NaCN in water (0.5 M) (measured). $E_{(water)}$ = Total energy released from addition of water (measured). $\Delta H_{(NaCN)}$ = Heat of reaction for the addition of 1 mole of NaCN (calculated). $n_{(NaCN)}$ = Moles of NaCN added to reaction (measured).

However, it was not possible to accurately quantify the heat of the reaction from these two experimentally obtained values due to complications caused by differences between the properties of the feeds in the experiment and the control. The difference in endotherms between the cyanation reaction (*Figure 4.3*) and the control experiment (*Figure 4.4*) indicated that the overall DABCO-catalyzed reaction of **4.2** and NaCN was exothermic, but under the experimental conditions, the endothermic addition of water to the acetonitrile solution outweighs this exotherm, making the overall reaction endothermic. Hence, these results suggested that as long as the concentration of the NaCN solution is sufficiently dilute to allow an overall reaction endotherm to be maintained, the slow addition of an aqueous NaCN solution was a safe method for scale-up of the cyanation reaction and would prevent potential thermal runaway upon scale-up.

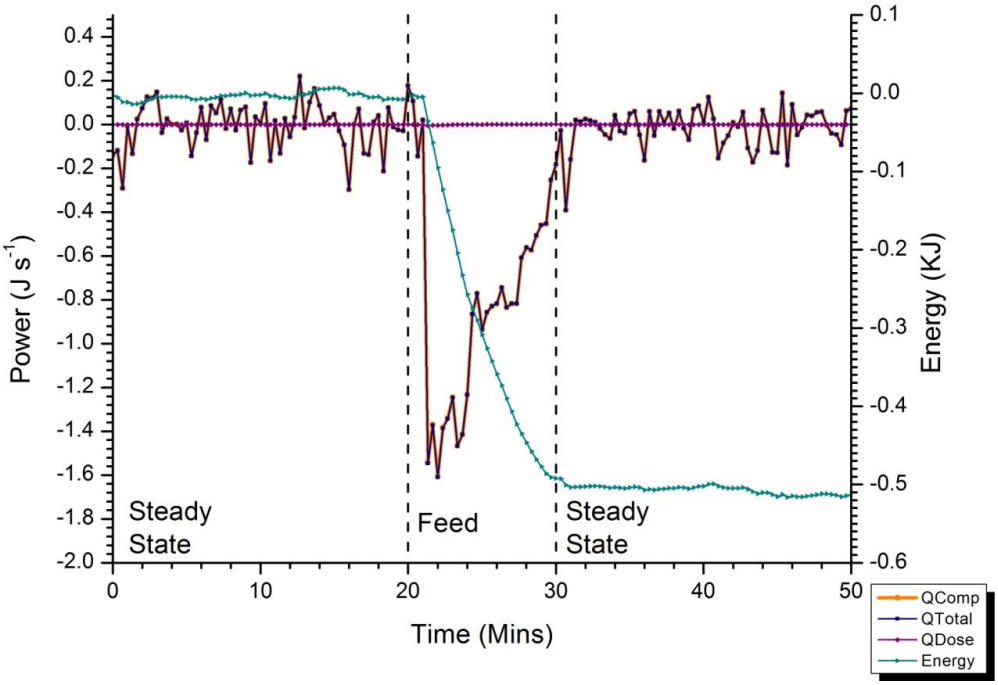


*Figure 4.3: Calorimeter traces for the addition of aqueous NaCN to a solution of **4.2** and DABCO in acetonitrile. Q Comp = compensatory power; Q Total = total power; Q Dose = power delivered by dosing of $NaCN_{(aq)}$ to the reaction solution. Energy = heat energy. Because Q Dose ~0 over the course of the experiment, traces for Q Comp (orange) and Q Total (blue) overlap.*

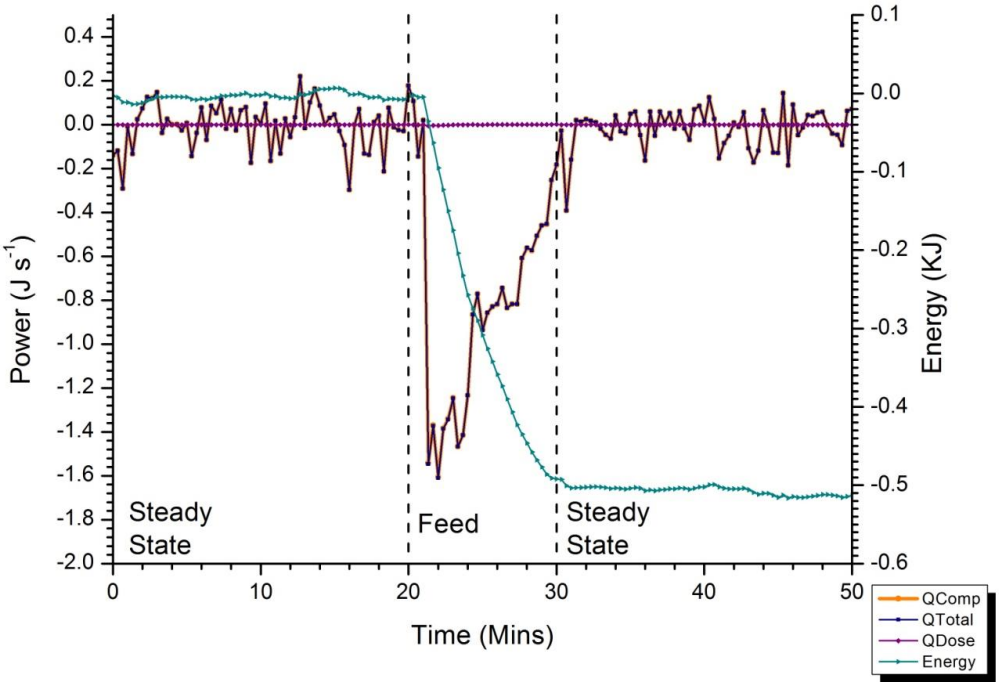


*Figure 4.4: Calorimeter traces for the addition of water to a solution of **4.2** and DABCO in acetonitrile. Q Comp = compensatory power; Q Total = total power; Q Dose = power delivered by dosing of water to the reaction solution. Energy = heat energy. Because Q Dose ~0 over the course of the experiment, traces for Q Comp (orange) and Q Total (blue) overlap.*

4.2.3.2 Multi-gram scale reactions

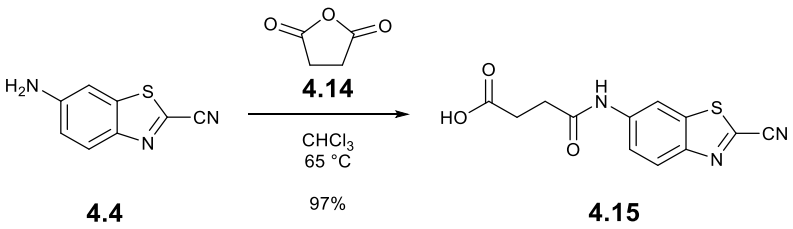
Nitration of chloride **4.1** was repeated on a 10 g scale in preparation for a large scale cyanation. Purification from acetonitrile afforded nitro **4.2** in 73% yield. Recrystallisation from acetonitrile requires only 25% of the volume of an ethanolic recrystallisation. However, acetonitrile is more expensive and less environmentally friendly than ethanol. Therefore, it may not be ideal for large scale applications, but still notable that recrystallisation can be performed in either solvent.

The cyanation of nitro **4.2** in acetonitrile and water was subsequently performed on a 2 g and 10 g scale, which were previously unfeasible in a standard research lab using the original cyanation procedure of chloride **4.3** to ACBT **4.4** in DMSO. Purification was achieved by rapid filtering of the crude product through a short plug (SiO_2; $CHCl_3$) to give the cyanated product **4.9** as pure material in 93% and 83% yield, respectively. This also demonstrated that the cyanation of nitro **4.2** is more efficient on scale.

Finally, the reduction of nitro CBT **4.9** using Fe in AcOH was repeated on a 1 g and 5 g scale. Purification was achieved by filtration of the crude through a short plug (SiO_2; $CHCl_3$) to yield ACBT **4.4** in a 60% and 70% yield, respectively.

4.2.3.3 Derivatization of ACBT **4.4**

To increase the applicability of the CBT handle to a wider range of chemical tools, a carboxylate functional group was installed on ACBT **4.4**, permitting the conjugation of a CBT handle onto amines and alcohols. Derivatization of ACBT **4.4** was performed according to literature procedure (*Scheme 4.11*).[272] ACBT **4.4** was heated at reflux with succinic anhydride **4.14** in chloroform to afford the succinylated product **4.15** in 97% yield. Additional recrystallisation of the pure product was achieved from ethanol for melting point determination.



*Scheme 4.11: Reaction conditions for synthesis of 3-[(2-cyano-1,3-benzothiazol-6-yl)carbamoyl]propanoic acid **4.15**.*[272]

4.3 Conclusion

In summary, a straight-forward, practical and readily scalable synthesis of ACBT **4.4** has been developed. Following calorimetric assessment of the cyanation reaction, synthesis of ACBT **4.4** was achieved safely in high purity and on a multigram scale, requiring only filtrations and recrystallization for purification of each intermediate. The endothermic nature of the controlled cyanation step and sole use of simple purification techniques would facilitate further scale up of the reaction in a safe and straightforward manner, if required.

Progress was made towards the synthesis of a CBT-functionalized LALDI tag through the synthesis of the succinylated CBT **4.15**, thereby facilitating the incorporation of a CBT handle onto amine-functionalized LALDI reagents via a simple amide coupling reaction.

Its noteworthy to highlight that sequential work by Sharma et al. has shown that the DABCO-catalyzed cyanation shown here for 6-nitro-substituted precursor **4.2** could also be applied to the 6-fluoro and 6-bromo analogues.[273] This reactivity was used to generate a wide variety of *N*- and *N,N*-substituted ACBT analogues via Buchwald–Hartwig amination using various primary and secondary amine. This successive work by Sharma et al. has significantly expanded the potential for derivatizing CBTs and it should be recognized as a valuable synthetic tool for developing novel chemical probes for CBT ligation.

4.4 Experimental

4.4.1 General experimental

All reactions were carried out under an atmosphere of dry nitrogen unless otherwise stated, using anhydrous solvents from a solvent purification system (Innovative Technology Inc. PureSolv), with the exception of anhydrous DMF, which was purchased from Acros Organics. All chemical starting reagents were purchased from commercial suppliers and used without further purification. Azeotropic distillation with toluene was used to remove persistent moisture from reagents with oligoethylene glycol chains before being used in a reaction. The identity and purity of known compounds was confirmed through the comparison of experimentally obtained data to values reported in the literature.

Thin layer chromatography (TLC) was carried out on Merck TLC Silica gel 60 F_{254} plates. Flash column chromatography was performed using Merck Geduran silica gel 60 (40–63 µm). All retention factors (R_f) are given to two decimal places along with the solvent system. Reversed-phase chromatography was performed using prepacked RediSep® Rf Reversed-phase C_{18} columns. Lyophilisation of compounds was performed using a Virtis Benchtop K freeze dryer.

The melting point of recrystallised solids was determined using a Griffin MFB-590 Melting Point Apparatus with a glass capillary melting point tube. Melting points (m.p.) are recorded as a range between the meniscus point and liquefaction with values given to the nearest degree in Celsius (°C).

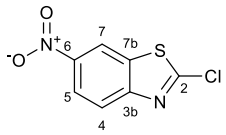
^{1}H and ^{13}C NMR spectra were recorded in deuterated solvents on a Bruker AVANCE 500 Ultrashield or Bruker AVANCE 400 Ultrashield NMR Spectrometer. Chemical shifts are referenced to residual solvent peaks and are quoted in ppm. Coupling constants (*J*) are reported to the nearest 0.1 Hz. Assignment of spectra was based on expected chemical shifts and coupling constants, aided by DEPT, COSY, HMQC, and HMBC where appropriate.

High-resolution electrospray mass spectrometry (HR-ESI-MS) was performed using a Bruker MaXis Impact spectrometer; *m/z* values are reported to four decimal places. LC-MS was recorded on an Agilent Technologies 1200 series HPLC combined with a Bruker HCT Ultra ion trap using 50 × 20 mm C_{18} reversed-phase columns with a solvent system of increasing acetonitrile (5 to 95%) in water, each containing 0.1% formic acid. A flow rate of 1.5 cm^3 min^{-1} was used and *m/z* values are given to one decimal place. Retention times (R_t) are provided in minutes to the nearest two decimal places.

Infrared (IR) spectroscopy was carried out using a Bruker Alpha Platinum ATR. Samples were analysed neat and absorption maxima (ν_{max}) are given in wave numbers (cm^{-1}) to the nearest whole wavenumber. Signals are defined as either strong (s), medium (m), weak (w), or broad (br). UV-Vis absorption was measured using an Agilent Technologies Cary 100 UV-Vis Spectrophotometer with all samples were analysed as solutions in a 10 mm Hellma Analytics High Precision Quarts Suprasil cell. Absorption maxima (λ_{max}) are given in nanometers to the nearest whole nanometer with the corresponding molar extinction coefficient (ε) given in $M^{-1}\,cm^{-1}$.

4.4.2 Experimental procedures

2-chloro-6-nitro-1,3-benzothiazole **4.2** [253,260]

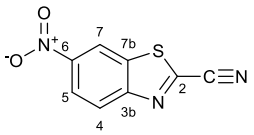


4.2

2-chloro-1,3-benzothiazole **4.1** (10.00 g, 59.0 mmol) was added dropwise to concentrated H_2SO_4 (60 mL) in a cooled round bottom flask (ice bath). Potassium nitrate (6.56 g, 64.9 mmol) was added portion wise, and the resulting reaction mixture was stirred at 0 °C (ice bath) for 30 mins, and then at room temperature for 18 h. The solution was poured onto ice and precipitate collected by filtration. The collected solid was washed with ice cold water until the flow through was acid free, and then dried under reduced pressure. Crude product was purified by recrystallisation from MeCN (~100 mL) or EtOH (~425 mL) to yield the product **4.2** as fine off-white needles; from EtOH (10.37g, 48.3 mmol, 82%), from MeCN (9.21g, 42.9 mmol, 73%). **m.p.** = 193–194 °C (EtOH). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.75 (d, 1H, *J* 2.3 Hz, C*H*-7), 8.38 (dd, 1H, *J* 9.0, 2.3 Hz, C*H*-5), 8.07 (d, 1H, *J* 9.0 Hz, C*H*-4). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 158.9 (*C*Cl), 154.9 (*C*NO_2), 145.6 (*C*-7b), 136.6 (*C*-3b), 123.5 (*C*H-4), 122.4 (*C*H-5), 117.8 (*C*H-7). **HR-ESI-MS** Calculated for $C_7H_3ClN_2O_2S$: *m/z* 214.9677 $[M+H]^+$; found 214.9671. **LC-MS** found 214.8, $\boldsymbol{R_t}$ = 1.92. **IR** ν_{max} 1598-1403 m (C=C aromatic), 1510 s (NO_2 asym), 1327 s (NO_2 sym).

Characterization is consistent with data reported in the literature.

6-nitro-1,3-benzothiazole-2-carbonitrile **4.9** [269]

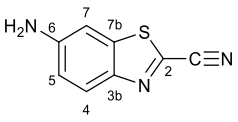


4.9

A solution of NaCN (2.40 g, 49.0 mmol) in H_2O (100 mL) was added slowly to a stirred solution of 2-chloro-6-nitro-1,3-benzothiazole **4.2** (10.0 g, 46.6 mmol) and DABCO (748 mg, 6.99 mmol) in MeCN (1000 mL). The reaction mixture was stirred at room temperature for 24 h. Excess cyanide was quenched by the addition of an aqueous $FeCl_3$ solution (30 mL, 0.3 M). The reaction mixture was diluted with H_2O (470 mL) and extracted with EtOAc (3 × 400 mL). The organic layers were combined, washed with brine (100 mL), dried with Na_2SO_4, filtered, and concentrated *in vacuo* to give a yellow solid. The crude product was dry loaded onto a short plug of silica gel, flushed through with $CHCl_3$ (ca. 1 L) and concentrated *in vacuo* to give the product **4.9** as a white solid (7.89g, 38.4 mmol, 83%). A sample of **4.9** was crystallized from MeOH for melting point analysis. $\boldsymbol{R_f}$ = 0.37 (SiO_2; $CHCl_3$). **m.p.** = 164–165 °C (MeOH). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.96 (d, 1H, *J* 2.1 Hz, C*H*-7), 8.52 (dd, 1H, *J* 9.1, 2.1 Hz, C*H*-5), 8.38 (d, 1H, *J* 9.1 Hz, C*H*-4). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 155.5 (*C*NO_2-6), 147.4 (*C*-7b), 141.9 (*C*-2), 135.7 (*C*-3b), 126.2 (*C*H-4), 123.2 (*C*H-5), 118.7 (*C*H-7), 112.1 (*C*≡N). **HR-ESI-MS** Calculated for $C_8H_4N_3O_2S$: *m/z* 206.0019 $[M+H]^+$; found 206.0010. **LC-MS** found 206.1, $\boldsymbol{R_t}$ **=** 1.83. **IR** ν_{max} 2235 w (C≡N), 1598-1405 m (C=C aromatic), 1518 s (NO_2 asym), 1342 s (NO_2 sym).

Characterization is consistent with data reported in the literature.[269]

6-amino-1,3-benzothiazole-2-carbonitrile **4.4** [229]



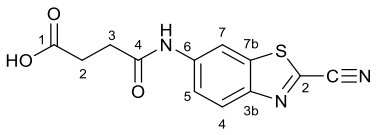
4.4

Fe powder (68.05 g, 1.218 mol) was added to a suspension of 6-nitro-1,3-benzothiazole-2-carbonitrile **4.9** (5.00 g, 24.4 mmol) in acetic acid (500 mL). The reaction mixture was stirred at room temperature for 24 h, diluted with water (1 L) and unreacted iron removed by filtration through celite. The aqueous solution was extracted with EtOAc (4 × 500 mL). The combined organic layers were washed with brine (2 × 300 mL), dried (Na_2SO_4) and concentrated *in vacuo*.

The crude product was dry loaded onto a short plug of silica gel, flushed through with $CHCl_3$ (ca. 1 L) and concentrated *in vacuo* to give the product **4.4** as yellow microcrystals (3.02 g, 17.2 mmol, 71%). A sample of **4.4** was crystallized from EtOH for melting point analysis. $\boldsymbol{R_f}$ = 0.2 (SiO_2; DCM). **m.p.** = 219–220 °C (EtOH). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.95 (d, 1H, J 8.9 Hz, C*H*-4), 7.08 (d, 1H, J 2.2 Hz, C*H-7*), 6.95 (dd, 1H, J 8.8, 2.2 Hz, C*H*-5), 4.14 (s, 2H, N*H*$_2$). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 147.8 (*C*-6), 145.7 (*C*-7b), 138.3 (*C*-3b), 131.2(*C*-2), 126.1 (*C*H-4), 117.8 (*C*H-5), 113.7 (*C*≡N), 104.0 (*C*H-7). **HR-ESI-MS** Calculated for $C_8H_6N_3S$: *m/z* 176.0277 $[M+H]^+$; found 176.0275. **LC-MS** found 176.1, $\boldsymbol{R_t}$ **=** 1.59. **IR** ν_{max} 3456 m (NH_2 asym), 3348 m (NH_2 sym), 2218 w (C≡N)**,** 1601-1455 m (C=C aromatic).

Characterization is consistent with data reported in the literature.[229]

3-[(2-cyano-1,3-benzothiazol-6-yl)carbamoyl]propanoic acid **4.15** [272]



4.15

Succinic anhydride **4.14** (171 mg, 1.71 mmol) was added to a solution of 6-amino-1,3-benzothiazole-2-carbonitrile **4.4** (200 mg, 1.14 mmol) in $CHCl_3$ (30 mL). The solution was heated to reflux and stirred overnight. The solvent was removed by rotary evaporation to yield a cream solid. The crude product subjected to flash chromatography (SiO_2; EtOAc–MeOH 95:5) to afford the product **4.15** as an off-white solid (304 mg, 1.11 mmol, 97%). A sample of **4.15** was crystallized from EtOH for melting point analysis. $\boldsymbol{R_f}$ = 0.2 (SiO_2; DCM). **m.p.** = 183-184 °C. **^{1}H NMR** (500 MHz, CD_3OD) δ 8.63 (d, 1H, *J* 1.9 Hz, C*H*-7'), 8.09 (d, 1H, *J* 9.0 Hz, C*H*-4'), 7.65 (dd, 1H, *J* 9.0, 2.0 Hz, C*H*-5'), 2.78 – 2.67 (m, 4H, C*H*$_2$-2, C*H*$_2$- 3). **^{13}C NMR** (125 MHz, CD_3OD) δ 176.3 (*C*OOH-1), 173.2 (*C*ONH-4), 149.6 (C-6'), 141.0 (C-7b'), 138.1 (C-3b'), 136.4 (C-2'), 125.9 (*C*H-7'), 122.0 (*C*H-5'), 114.10 (CN), 112.4 (*C*H-4'), 32.5 (*C*H_2-2), 29.8 (*C*H_2-3). **HR-ESI-MS** Calculated for $C_{12}H_{10}N_3O_3S$: *m/z* 276.04374 $[M+H]^+$; found 276.04375. **LC-MS** found 276.0, $\boldsymbol{R_t}$ **=** 1.58. **IR** ν_{max} 3370 m (NH), 2233 m (C≡N), 1702 m (C=O acid) 1675 s (C=O amide), 1567-1432 m (C=C aromatic).

Characterization is consistent with data reported in the literature.[272]

4.4.3 Calorimetry

Calorimetry experiments were carried out using HEL AutoMATE parallel reactors with HEL WinISO 2225 and HEL IQ 1.2.16 software. Temperature was controlled using a Julabo refrigerated/heating circulator (Model FP50-HD) and feed reactants introduced into the reaction using a Harvard syringe pump (model pump 11) connected to a 20 mL disposable syringe with an 8 inch, 16 gauge Luer fitting syringe needle.

4.4.3.1 Addition of $NaCN_{(aq)}$ to 6-nitro-1,3-benzothiazole-2-carbonitrile **4.2** and DABCO in MeCN

In a 50 mL vessel, 6-nitro-1,3-benzothiazole-2-carbonitrile **4.2** (200 mg, 0.932 mmol) and DABCO (15.7 mg, 0.14 mmol) were dissolved in MeCN (20.0 mL). 4.90 M $NaCN_{(aq)}$ was then fed into the reaction according to the required plan as detailed below. Following the reaction, any unreacted cyanide was quenched by the addition of 3 mM $FeCl_{3\ (aq)}$.

To determine the enthalpy of the reaction, power compensation calorimetry was used with the reaction plan as detailed below:

Steady State 1:

Stir	300 rpm
Reactor Temperature	21 °C
Circulator Temperature	1 °C
Time	60 min

Feed 1 :

Stir	300 rpm
Reactor Temperature	21 °C
Circulator Temperature	1 °C
Feed	4.90 M $NaCN_{(aq)}$ at 0.2 mL min^{-1} for 10 min (2.00 mL, 0.978 mmol NaCN)
Time	10 min

Steady State 2:

Stir	300 rpm
Reactor Temperature	21 °C
Circulator Temperature	1 °C
Time	1080 min

Feed 2 :

Stir	300 rpm
Reactor Temperature	21 °C
Circulator Temperature	1 °C
Feed	0.003 M $FeCl_{3\,(aq)}$ at 2 mL min^{-1} for 10 min (20.0 mL, 0.060 mmol $FeCl_3$)
Time	10 min

Steady State 3:

Stir	300 rpm
Reactor Temperature	21 °C
Circulator Temperature	1 °C
Time	60 min

The enthalpy (kJ mol^{-1}) of the reaction was determined using IQ analytical software. The software required input of the specific heat capacity of the feed for the calculation of compensatory power. For this, the specific heat capacity of water (4.18 J g^{-1} K^{-1}) was used.

4.4.3.2 Addition of water to 6-nitro-1,3-benzothiazole-2-carbonitrile **4.2** (200 mg) and DABCO in MeCN

In a 50 mL vessel, 6-nitro-1,3-benzothiazole-2-carbonitrile **4.2** (200 mg, 0.932 mmol) and DABCO (15.7 mg, 0.14 mmol) were dissolved in MeCN (20.0 mL). Water (2.0 mL) was then fed into the reaction according to the required plan as detailed below.

To determine the enthalpy of the reaction, power compensation calorimetry was used with the reaction plan as detailed below:

Steady State 1:

Stir	300 rpm
Reactor Temperature	21 °C
Circulator Temperature	1 °C
Time	90 min

Feed 1:

Stir	300 rpm
Reactor Temperature	21 °C
Circulator Temperature	1 °C
Feed	Water at 0.2 mL min^{-1} for 10 min (2.00 mL)
Time	10 min

Steady State 2:

Stir	300 rpm
Reactor Temperature	21 °C
Circulator Temperature	1 °C
Time	50 min

The enthalpy (kJ mol^{-1}) of the reaction was determined using IQ analytical software. The software required input of the specific heat capacity of the feed for the calculation of compensatory power. For this, the specific heat capacity of water (4.18 J g^{-1} K^{-1}) was used.

Chapter 5 – Thesis summary and future directions

This thesis explores the potential of LALDI-MS as a novel analytical tool for MS-based glycomics. Initial work on the optimization of LALD-MS analysis and development of improved LDI enhancers was described in *Chapter 2*, then applied towards the design, utilization, and analysis of glycan-selective LALDI tag reagents in *Chapter 3*.

Chapter 2 describes the development of novel water-soluble LDI enhancing labels through investigations into the LDI sensitivity and MS stability of various pyrene-based LDI enhancers. Various structural analogues of pyrene-based LDI enhancers were investigated in the development novel LDI enhancers that achieved high sensitivity and good MS-stability. These investigations confirmed previous reports that the structure of an LDI enhancers significantly affect the LALDI-MS performance. However, it was found that current methods quoted in the literature for rationalizing observed LALDI-MS sensitivity did not correlate as well as previously thought, noting that measuring the absorbance in the solution phase may not give an accurate insight into the UV absorbance and subsequent LDI sensitivity of a neat sample.

Future work will continue to focus on developing novel LDI enhancers with improved LALDI-MS performance. Future investigations could assess the absorption of LDI enhancers as neat samples, to evaluate whether a better correlation can be drawn between UV-absorption and LALDI-MS sensitivity. All investigations described within *Chapter 2* focused on the performance of the LDI enhancers in positive mode LALDI-MS. Pyrenes have also been shown to enhance negative mode MALDI-MS analysis,[164,168,169] especially for analysis of oligosaccharides.[164] However, negative mode LALDI-MS currently remains an unexplored area within the literature. Therefore, future research could also be directed towards assessing the performance of pyrenes as negative mode LDI enhancers.

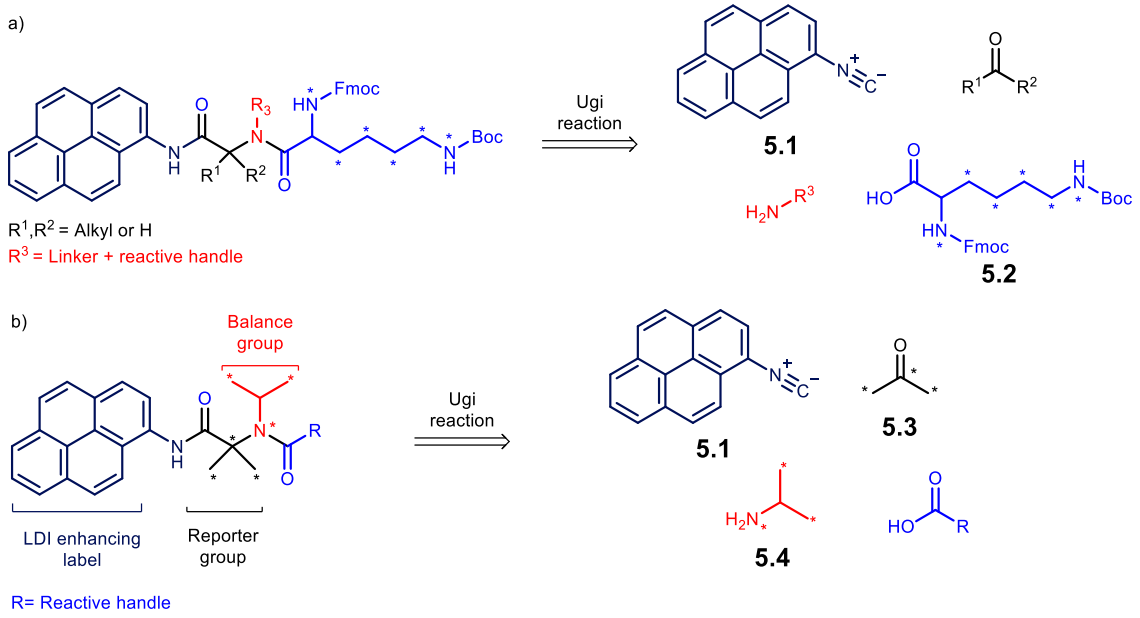
Chapter 3 details the development of a series of LALDI tag reagents for studying native glycans from complex biological samples by LALDI-MS. The LALDI tag design features an LDI-enhancing label conjugated to a glycan-selective reactive handle through a solubilizing linker. Labelling and analysis of simple glycans by LDI-MS was demonstrated using the LALDI tag reagents, without the need for additional chemical modifications, MALDI matrices, or SALDI surface materials, even in the presence of an excess of salts and other contaminants. To determine which LALDI tag reagent would be most suitable for carrying out glycan labelling and analysis directly from a biological sample, each of the LALDI tags was assessed on both labelling efficiency and LALDI-MS performance labelled glycans. This included an in-depth analysis of the fragmentation

patterns of both the LALDI tag reagents and labelled glycans. In situ labelling of lactose from cow's milk was achieved using two separate LALDI tag reagents, demonstrating that both labelling and analysis could be achieved directly from a native biological sample. These results suggested that LALDI tags could, indeed, have potential as tools for MS-based glycomics, although it was concluded that further optimizations of both the labelling and analysis would be required before being applied to more complicated systems. Analysis of larger and more complicated glycans was also attempted using the LALDI tag reagents. However, both glycan labelling and LALDI-MS analysis were hindered by due to the use sub-optimal reaction conditions.

Future work could be directed towards improving the efficiency of the glycan labelling reactions, optimizing the analysis of mixtures of labelled glycans, and establishing a suitable protocol for analyzing labelled glycans from the crude reaction mixtures. Once these aspects have been optimized, efforts could then be focused on establishing the applications of LALDI tags for the analysis of complex and biologically relevant glycans, the first step being the repetition of both the milk labelling experiments and analysis of larger glycans. As graphene has been shown to be advantageous for enriching aromatic analytes from biological samples through π–π interactions,[79] an investigation into the application of PGC solid phase extraction could be pursued as a quick and simple solution for improving the isolation and enrichment of LALDI tag labelled species. Additional work could also be directed towards the development of simpler LALDI tag reagents without the OEG linker. In glycomics studies, glycans isolated from native sources are generally lyophilized prior to derivatization rather than labelling directly from aqueous environment.[171] Consequently, this means that LALDI tags designed for glycomic investigations may not need to be soluble under aqueous conditions. Removing the OEG chain would simplify the structure of the LALDI tag reagents and potentially reduce the degree of fragmentation observed during LALDI-MS analysis. A simple test for this could be to use the commercially available pyrene hydrazide to label glycans, then analyze the conjugate products by LALDI-MS.

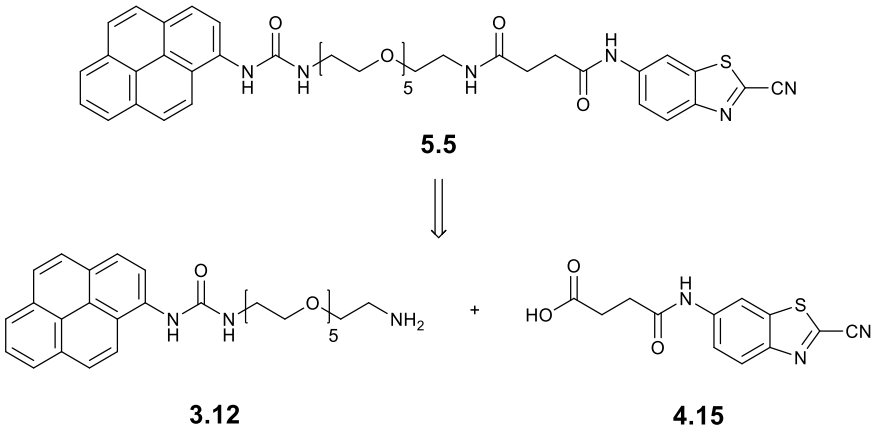
Furthermore, as it intended to use LALDI-MS to study mixtures of labelled analytes, efforts should also be made to investigate whether quantitative MS analysis can be achieved using LALDI-MS. It is proposed that multicomponent reactions, such as an Ugi reaction with pyrene isocyanide, generate LDI enhancers functionalized with heavy and light mass tags (*Scheme 5.1a*). Furthermore, if these pyrene-functionalized Ugi products exhibit consistent fragmentation patterns, isobaric mass tags could be generated by varying the distribution of heavy isotopes

around structure of the tag (*Scheme 5.1b*). Reagents such as these could have valuable applications for achieving quantitative -omics analysis by LALDI-MS.



*Scheme 5.1: Isotopically labelled LDI enhancers. a) Proposed synthesis of heavy/light LALDI mass tags using pyrene isocyanide **5.1** and different isotopically labelled lysines **5.2**. b) Proposed synthesis of isobaric LALDI mass tags using pyrene isocyanide **5.1** and isotopically labelled isopropylamine **5.3** and acetone **5.4**. Isotopically labelled atoms are marked by an asterisk (*).*

Chapter 4 described the development of a straightforward and readily scalable synthesis of ACBT **4.4** via a DABCO-catalyzed cyanation. Considerations towards aspects such as solvents, purification techniques, and reaction thermodynamics allowed scale up of the procedure to be accomplished safely, affording ACBT **4.4** in high purity on a multigram scale. Future work will involve the synthesis of a LALDI tag reagent functionalized with a CBT reactive handle (*Scheme 5.2*). Such reagents could then be used to selectively label and analyze peptides or proteins with *N*-terminal cysteine residues by LALDI-MS, thereby expanding the scope of LALDI-MS to include a potential for proteomics studies.



*Scheme 5.2: Retrosynthetic analysis of a CBT-functionalized LALDI tag **5.5** featuring a pyrene urea LDI enhancing label.*

Appendix I – LALDI-MS spectra recorded during lower detection limit investigations

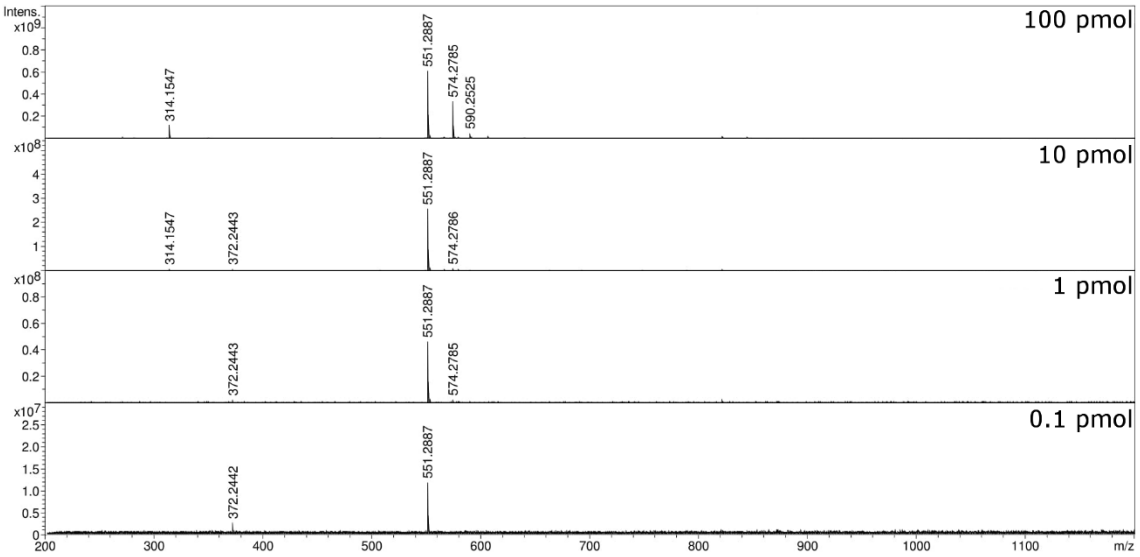


*Figure A.1: LALDI-MS spectra of pyrene **2.1** (100 pmol - 0.1 pmol) obtained during lower detection limit analysis using a Bruker AnchorChip™ micro-focusing MALDI target plate.*

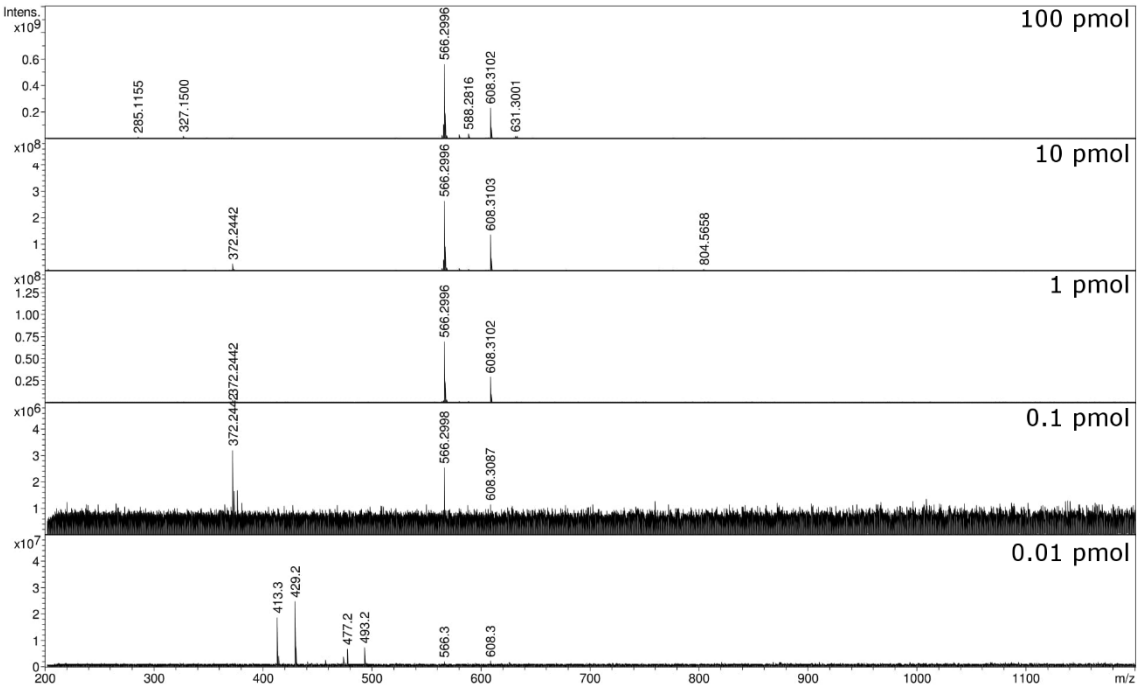


*Figure A.2: LALDI-MS spectra of 6-amidopyrene **2.2** (100 pmol - 0.01 pmol) obtained during lower detection limit analysis using a Bruker AnchorChip™ micro-focusing MALDI target plate.*

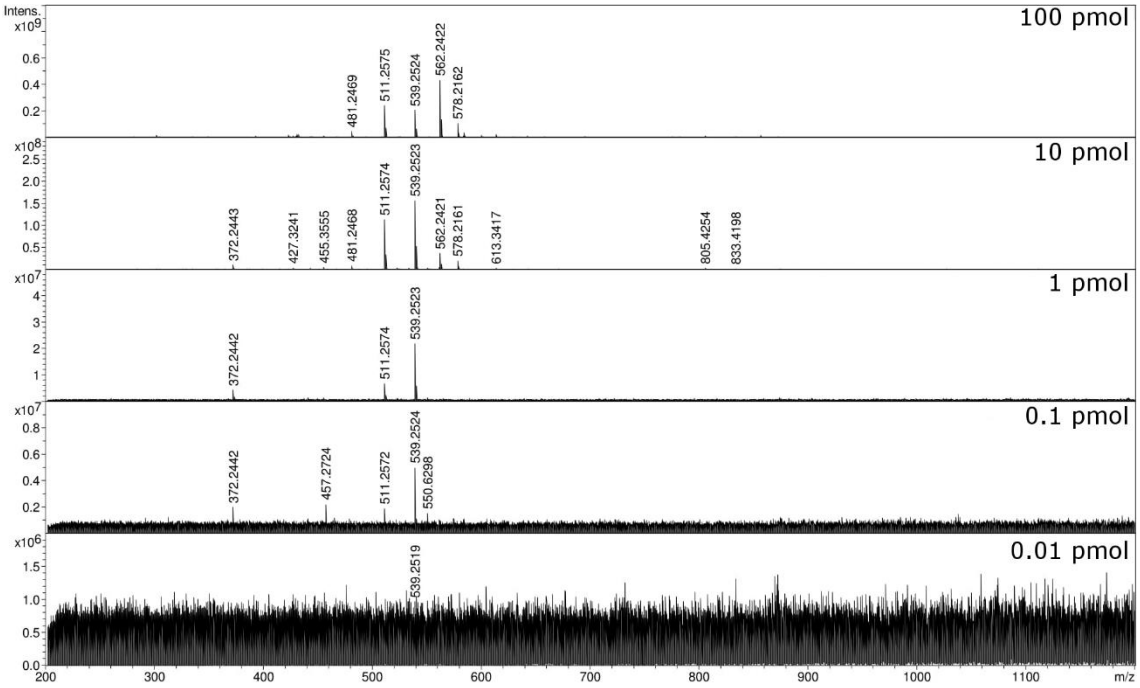


*Figure A.3: LALDI-MS spectra of 1-amidopyrene **2.3** (100 pmol - 0.01 pmol) obtained during lower detection limit analysis using a Bruker AnchorChip™ micro-focusing MALDI target plate.*

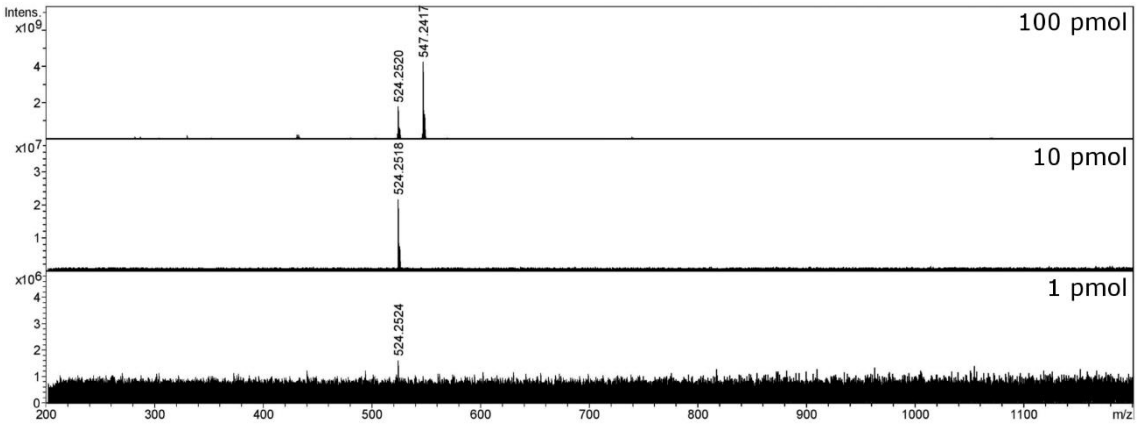


*Figure A.4: LALDI-MS spectra of 1-amidopyrene **2.27** (100 pmol - 1 pmol) obtained during lower detection limit analysis using a Bruker AnchorChip™ micro-focusing MALDI target plate.*

Appendix II – Detailed analysis of LALDI-MS spectra of unreacted LALDI tags

LALDI-MS analysis of PU-OEG-2AB **3.16**

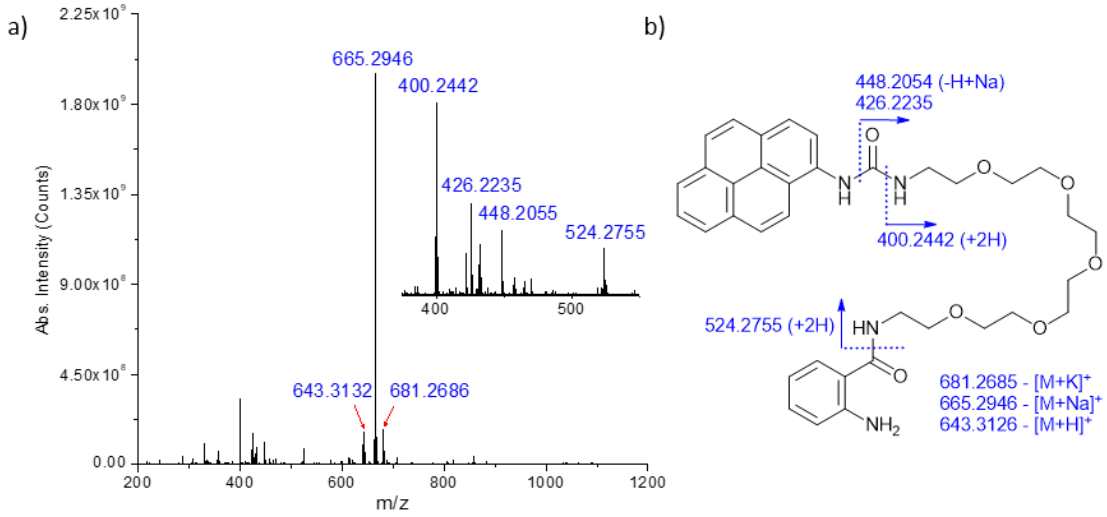


*Figure A.5: LALDI-MS analysis of PU-OEG-2AB **3.16**. a) LALDI-MS spectra for PU-OEG-2AB **3.16** (100 pmol) with identified fragments and adducts annotated with their observed m/z. Inlaid graph shows expansion of the full spectra at m/z 375-550.b) Proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

LALDI-MS analysis of PU-OEG-2AB **3.16** showed clear detection of the molecular ion as its $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$ adduct ions (*Figure A.5a*). The base ion peak in the spectrum (*m/z* 655.2946) was identified as $[M+Na]^+$.

LALDI-MS analysis of AP-OEG-Hz **3.23**

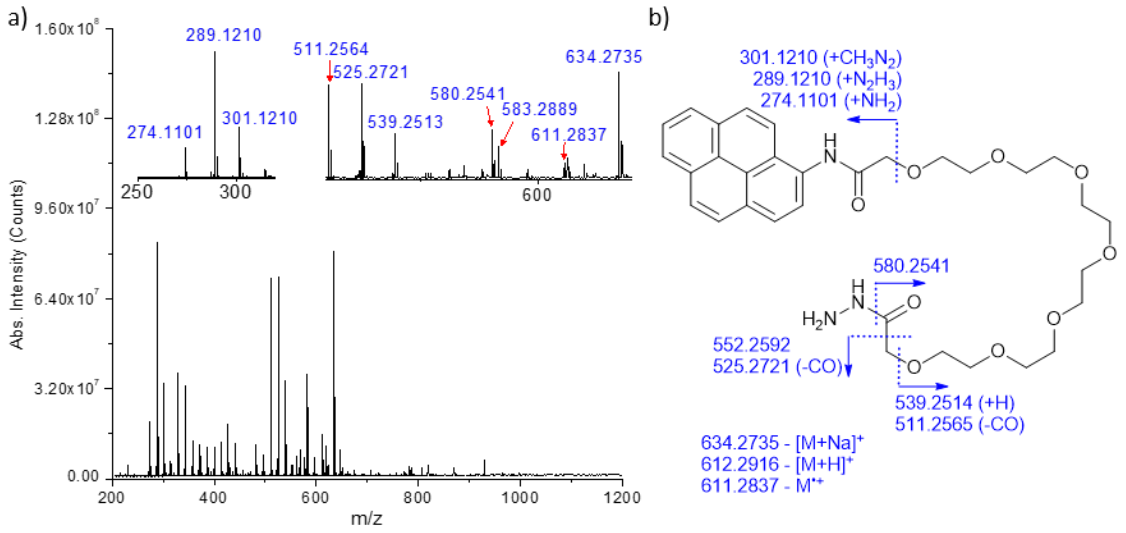


*Figure A.6: LALDI-MS analysis of AP-OEG-Hz **3.23**. a) LALDI-MS spectra for AP-OEG-Hz **3.23** (100 pmol) with identified fragments and adducts annotated with their observed m/z. Inlaid graphs show expansion of the full spectra at m/z 250-320 and m/z 510-640. b) Proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

LALDI-MS analysis of PU-OEG-2AB **3.23** showed clear detection of the $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$ adduct ions (*Figure A.6*). Significant degradation of AP-OEG-Hz **3.23** was observed in the LALDI-MS analysis, with a large number of signals corresponding to fragments and other degradation products. As expected, degradations resulting in the loss of CO were observed, complementing previous observations about the behavior of 1-amidopyrene reagents during LALDI-MS analysis. Most of the identifiable fragments of the molecular ion were observed close to the hydrazide reactive handle with the others occurring at the C-O bond adjacent to the aryl amide, including the base ion peak (*m/z* 289.1210) (*Figure A.6b*).

Proposed assignments for *m/z* 274.1101, 289.1210, and 301.1210 were generated by calculating a potential molecular formula from the respective *m/z* values. The calculated molecular formulas matched exactly but did not fit well into the structure of **3.23**. The closest explanation involved the addition of various N-containing adducts. It is proposed that the adducts could have been generated by degradation of the hydrazide group during LALDI-MS analysis. Several signals between *m/z* 329–480 could not be identified by any possible combination of fragmentation or dimerization that could have stemmed from the molecular ion.

It was noted that HR-ESI-MS analysis of **3.23** did not cause any fragmentation or degradation, appearing exclusively as the adducts of the molecular ion, suggesting that the degradations observed in *Figure A.6* were a product of the LDI process.

Appendix III – Detailed analysis of LALDI-MS spectra of pure samples of LALDI tag-labelled glycans

LALDI-MS analysis of LALDI tag-labelled lactose **3.33**, **3.34** and GlcNAc **3.35** isolated from the oxime ligation labelling reactions.

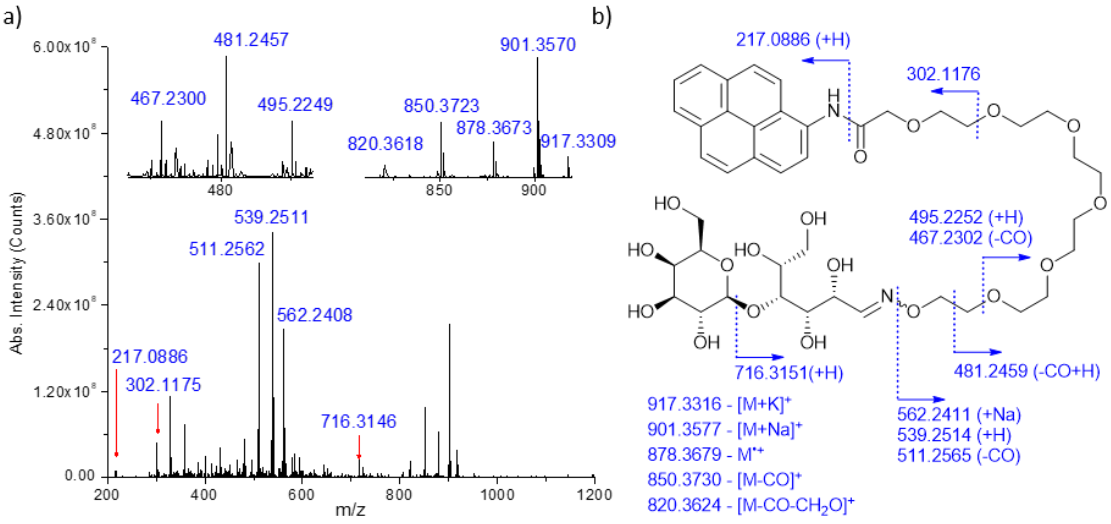


*Figure A.7: LALDI-MS analysis of LALDI tag-labelled lactose **3.33**. a) LALDI-MS spectra for LALDI tag-labelled lactose **3.33** (100 pmol) analyzed as a pure sample, identified fragments and adducts are annotated with their observed m/z values. Inlaid graphs show expansion of the full spectra at m/z 460-500 and m/z 810-920. b) Chemical structure for LALDI tag-labelled lactose **3.33** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

In the LALDI-MS analysis of LALDI tag-labelled lactose **3.33** the molecular ion was observed as the radical cation ($M^{\bullet+}$), along with the $[M+Na]^+$, and $[M+K]^+$ adduct ions (*Figure A.7a*). The base ion peak in the spectrum (*m/z* 539.2511) was identified as a fragment generated by cleavage at the oxime. (*Figure A.7b*). This degradation was expected following experimentally observed cleavage of the aminooxy group of **3.5** (*Chapter 3 – Figure 3.4*) and data published by Yoneda et al. showing cleavage at the oxime linkages ApA probes **1.5** and **1.6** during LALDI-MS analysis.[100] Other expected degradations that were observed included loss of CO from molecular and fragment ions, and cleavage of the aminopyrene group.

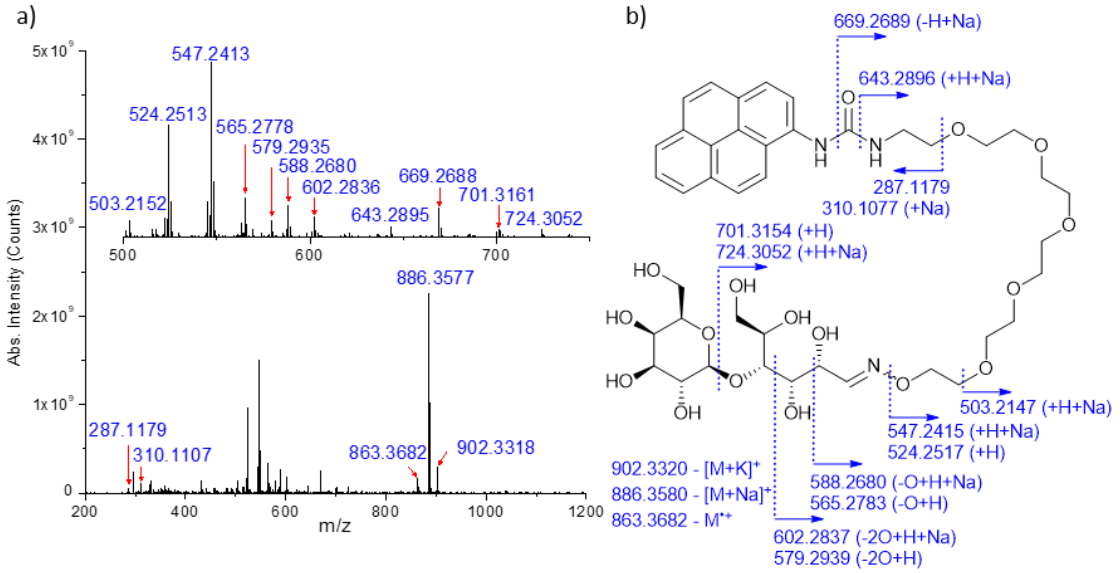


*Figure A.8: LALDI-MS analysis of LALDI tag-labelled lactose **3.34**. a) LALDI-MS spectra for LALDI tag-labelled lactose **3.34** (100 pmol) analyzed as a pure sample, identified fragments and adducts are annotated with their observed m/z values. Inlaid graph shows expansion of the full spectra at m/z 500-750.b) Chemical structure for LALDI tag-labelled lactose **3.34** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

In the LALDI-MS analysis of LALDI tag-labelled lactose **3.34** the molecular ion was detected as the radical cation ($M^{\bullet+}$), and $[M+K]^+$ adduct ions (*Figure A.8a*). The base ion peak in the spectrum (*m/z* 886.3577) was identified as the $[M+Na]^+$ adduct (*Figure A.8b*). Fragmentations were also observed at a number of bonds in the lactose conjugate, including loss of galactose (*m/z* 701.3161) and its sodiated adduct (*m/z* 724.3052).

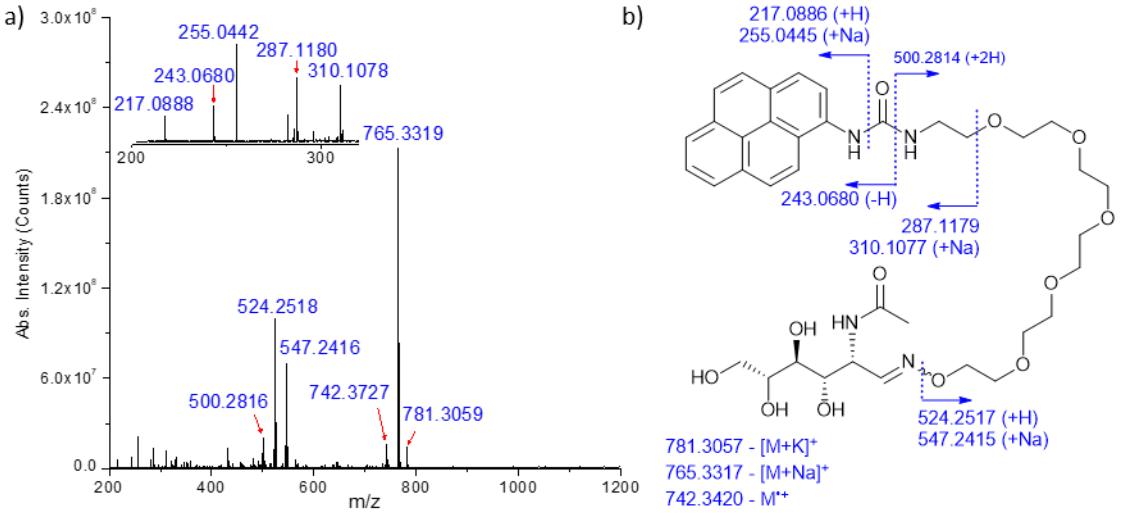


*Figure A.9: LALDI-MS analysis of LALDI tag-labelled GlcNAc **3.35**. a) LALDI-MS spectra for LALDI tag-labelled GlcNAc **3.35** (100 pmol) analyzed as a pure sample, identified fragments and adducts are annotated with their observed m/z values. Inlaid graph shows expansion of the full spectra at m/z 200-320. b) Chemical structure for LALDI tag-labelled GlcNAc **3.35** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

In the LALDI-MS analysis of LALDI tag-labelled lactose **3.35** the molecular ion could be observed as the radical cation ($M^{\bullet+}$), and $[M+K]^+$ adduct ions (*Figure A.9a*). The base ion peak in the spectrum (*m/z* 765.3319) was identified as the $[M+Na]^+$ adduct (*Figure A.9b*).

LALDI-MS analysis of LALDI tag-labelled lactose **3.36** and **3.37** isolated from the reductive amination labelling reactions.

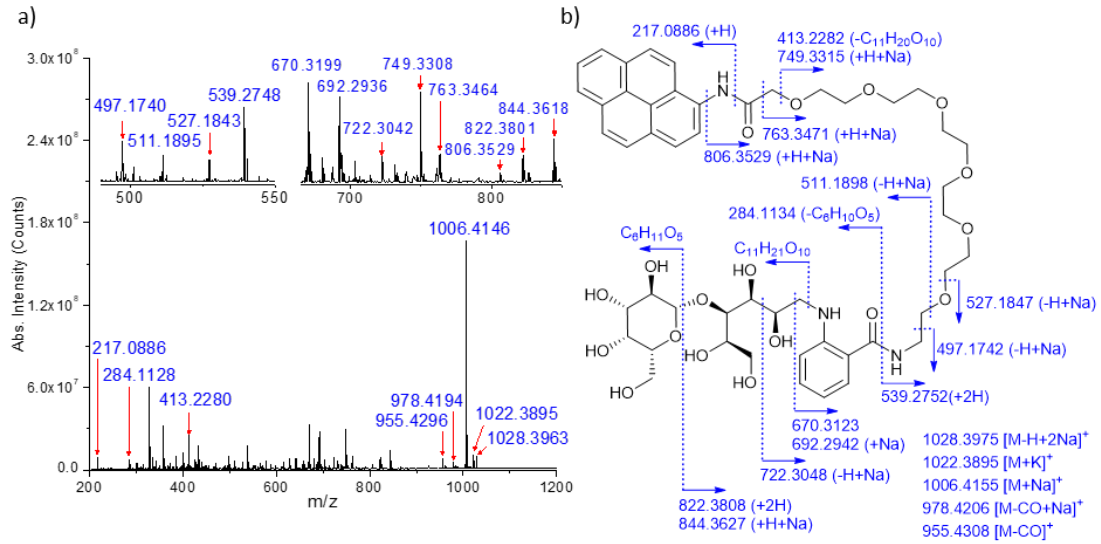


*Figure A.10: LALDI-MS analysis of LALDI tag-labelled lactose **3.36**. a) LALDI-MS spectra for LALDI tag-labelled lactose **3.36** (100 pmol) analyzed as a pure sample, identified fragments and adducts are annotated with their observed m/z values. Inlaid graphs show expansion of the full spectra at m/z 490-550 and m/z 665-850. b) Chemical structure for LALDI tag-labelled lactose **3.36** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

LALDI-MS analysis of LALDI tag-labelled lactose **3.36** was able to detect the molecular ion as the $[M+Na]^+$, $[M+K]^+$ and $[M-H+2Na]^+$ adduct ions (*Figure A.10a*). The base ion peak in the spectrum (*m/z* 1006.4146) was identified as the $[M+K]^+$ adduct ion. In addition to the cleavages around the benzamide linkage, fragmentations were also observed at the glycosidic linkage of lactose resulting in the loss of galactose (m/z 844.3627, 822.3808, 284.1134), and at the secondary amine (*m/z* 796.3082, 413.2282) resulting in the loss of lactose (*Figure A.10b*). Minor loss of CO from molecular and fragment ions was also observed in the LALDI-MS spectrum.

While a significantly higher degree of fragmentation was observed for **3.36** compared to any of the glycans labelled by oxime ligation, most signals were very low in abundance compared to the sodium adduct of the molecular ion.

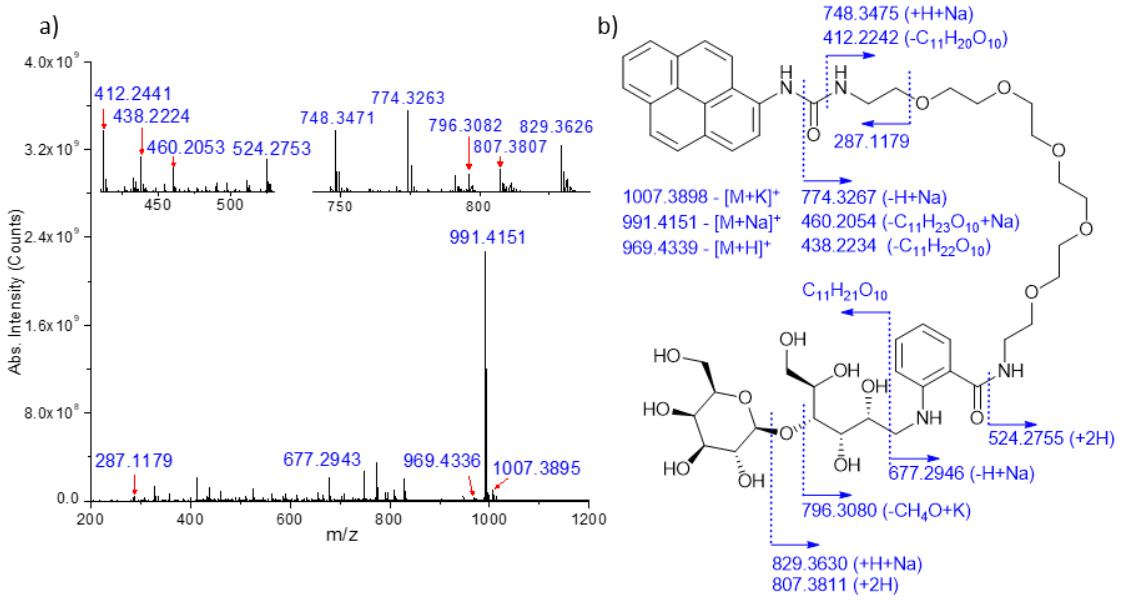


*Figure A.11: LALDI-MS analysis of LALDI tag-labelled lactose **3.37**. a) LALDI-MS spectra for LALDI tag-labelled lactose **3.37** (100 pmol) analyzed as a pure sample, identified fragments and adducts are annotated with their observed m/z values. Inlaid graphs show expansion of the full spectra at m/z 410-530 and m/z 740-840. b) Chemical structure for LALDI tag-labelled lactose **3.37** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

LALDI-MS analysis of LALDI tag-labelled lactose **3.37** was able to detect the molecular ion as the $[M+H]^+$, $[M+Na]^+$ and $[M+K]^+$ adduct ions (*Figure A.11a*). The base ion peak in the spectrum (*m/z* 991.4151) was identified as the $[M+Na]^+$ adduct ion. In addition to the cleavages around the benzamide linkage, fragmentations were also observed at the glycosidic linkage of lactose resulting in the loss of galactose (m/z 796.3082) (*Figure A.11b*).

LALDI-MS analysis of LALDI tag-labelled lactose **3.38** isolated from the reductive hydrazination labelling reaction

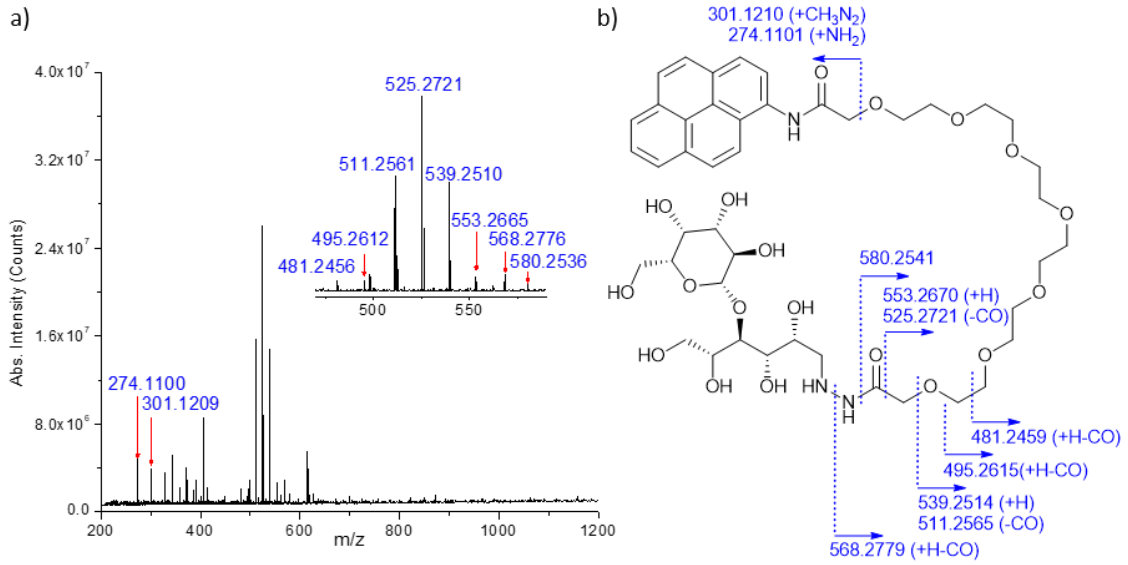


*Figure A.12: LALDI-MS analysis of LALDI tag-labelled lactose **3.38**. a) LALDI-MS spectra for a LALDI tag-labelled lactose **3.38** (100 pmol) contaminated by an unknown impurity according to analytical HPLC, identified fragments and adducts are annotated with their observed m/z values. Inlaid graph shows expansion of the full spectra at m/z 470-590. b) Chemical structure for LALDI tag-labelled lactose **3.38** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

No signals relating to a parent ion for hydrazide labelled lactose **3.38** were detected. A large number of fragments could not be identified. This could be due to the unidentified impurity that was present in the sample during analysis.

Appendix IV – Detailed analysis of LALDI-MS spectra of pure samples of LALDI tag-labelled glycans from HBSS

LALDI-MS analysis of LALDI tag-labelled lactose **3.33**, **3.34** and GlcNAc **3.35** from HBSS

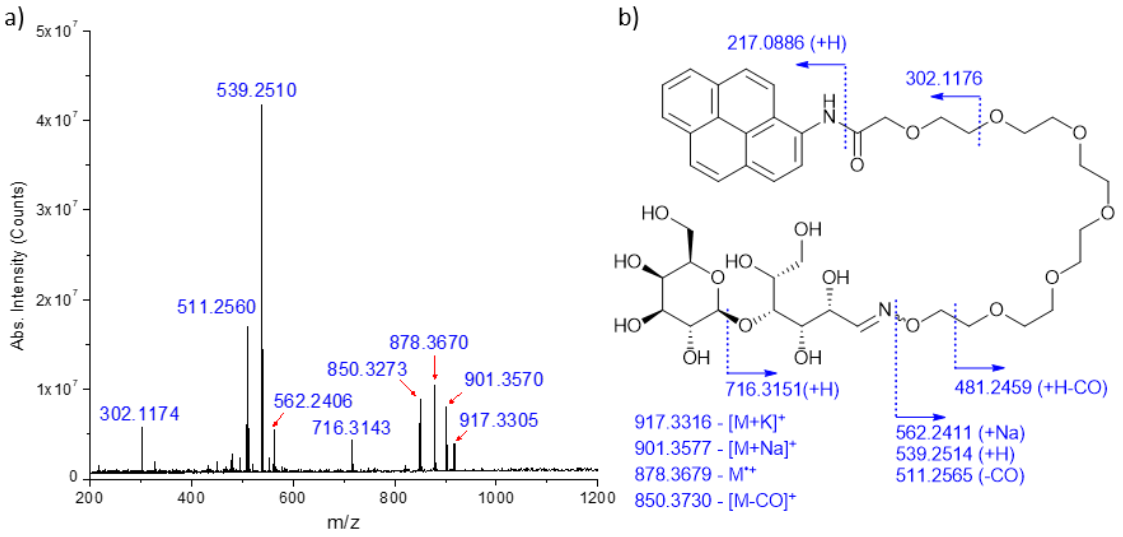


*Figure A.13: LALDI-MS analysis of LALDI tag-labelled lactose **3.33** dissolved in Hanks' buffered salt solution (HBBS). a) LALDI-MS spectra for LALDI tag-labelled lactose **3.33** (100 pmol) analyzed from a background of HBSS, identified fragments and adducts are annotated with their observed m/z values. b) Chemical structure for LALDI tag-labelled lactose **3.33** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

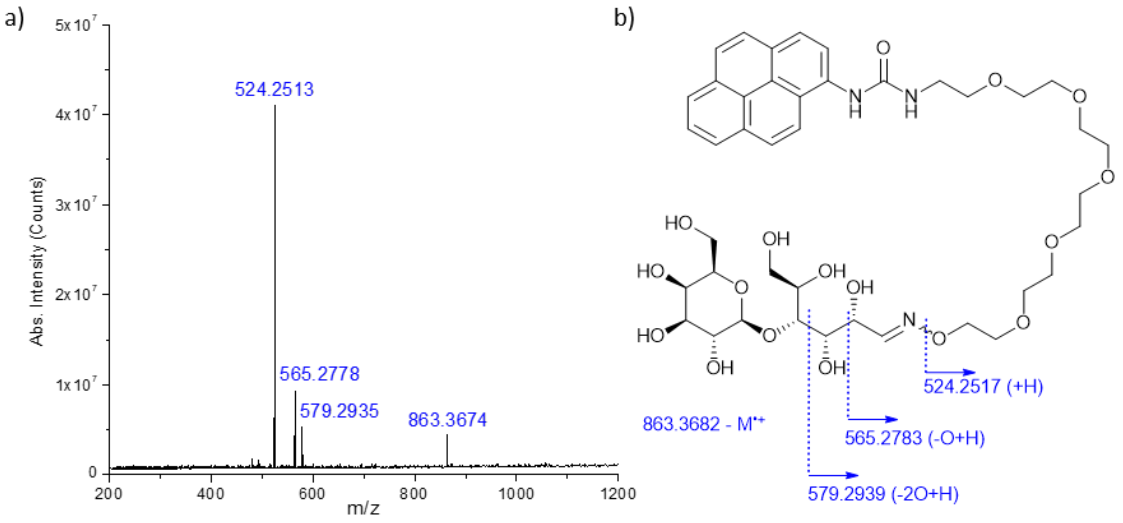


*Figure A.14: LALDI-MS analysis of LALDI tag-labelled lactose **3.34** dissolved in Hanks' buffered salt solution (HBBS). a) LALDI-MS spectra for LALDI tag-labelled lactose **3.34** (100 pmol) analyzed from a background of HBSS, identified fragments and adducts are annotated with their observed m/z values. b) Chemical structure for LALDI tag-labelled lactose **3.34** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

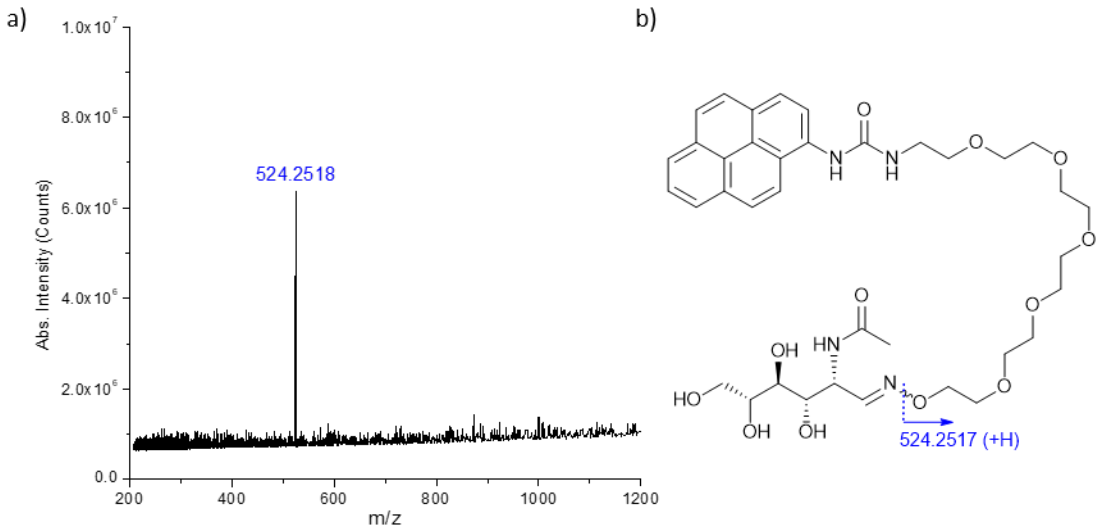


*Figure A.15: LALDI-MS analysis of LALDI tag-labelled GlcNAc **3.35** dissolved in Hanks' buffered salt solution (HBBS). a) LALDI-MS spectra for LALDI tag-labelled lactose **3.35** (100 pmol) analyzed from a background of HBSS., identified fragments and adducts are annotated with their observed m/z values. b) Chemical structure for LALDI tag-labelled GlcNAc **3.35** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

LALDI-MS analysis of LALDI tag-labelled lactose **3.36** and **3.37** from HBSS

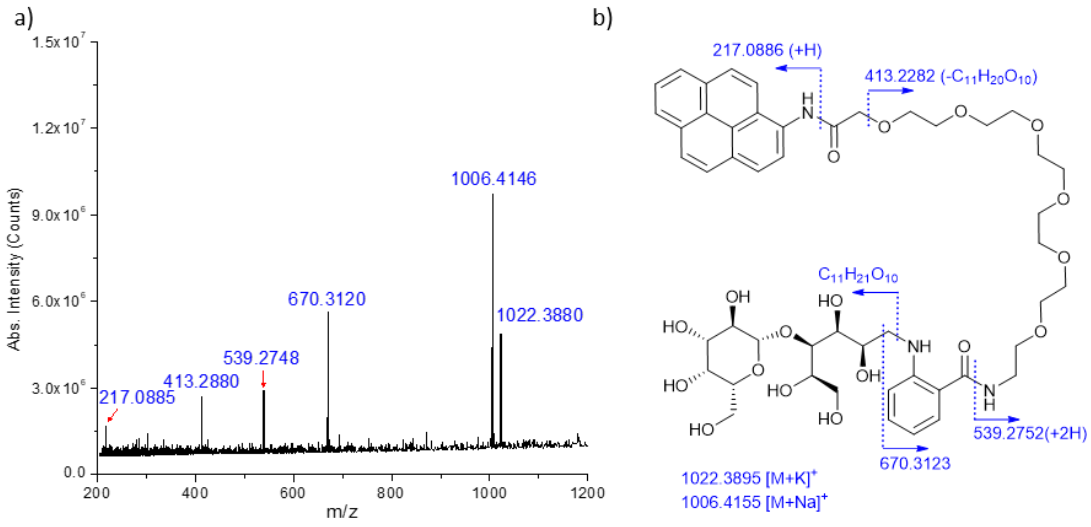


*Figure A.16: LALDI-MS analysis of LALDI tag-labelled lactose **3.36** dissolved in Hanks' buffered salt solution (HBBS). a) LALDI-MS spectra for LALDI tag-labelled lactose **3.36** (100 pmol) analyzed from a background of HBSS identified fragments and adducts are annotated with their observed m/z values. b) Chemical structure for LALDI tag-labelled lactose **3.36** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

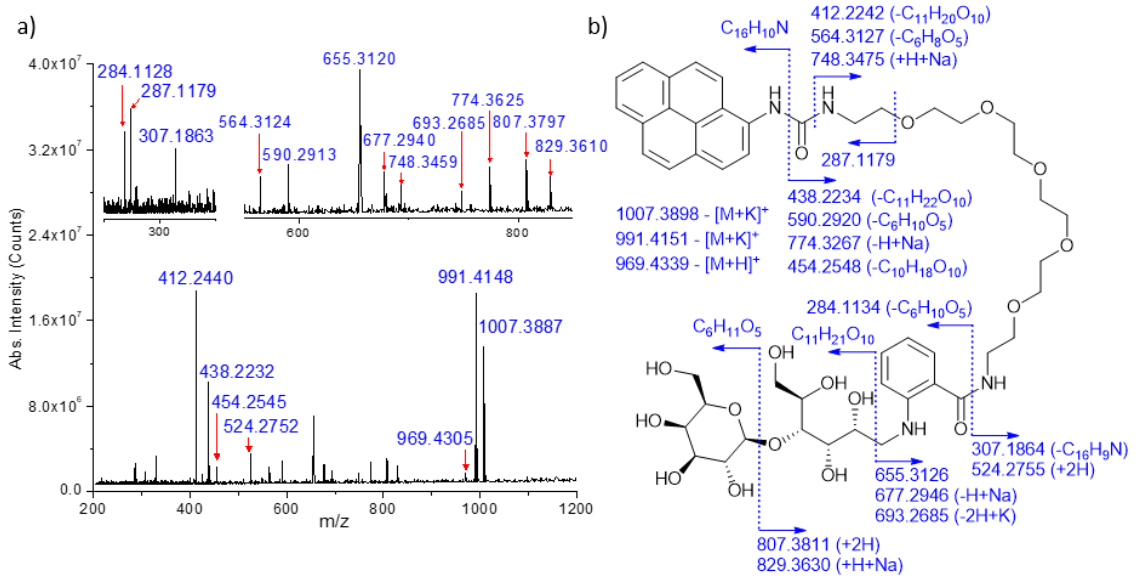


*Figure A.17: LALDI-MS analysis of LALDI tag-labelled lactose **3.37** dissolved in Hanks' buffered salt solution (HBBS). a) LALDI-MS spectra for LALDI tag-labelled lactose **3.37** (100 pmol) analyzed from a background of HBSS, identified fragments and adducts are annotated with their observed m/z values. Inlaid graphs show expansion of the full spectra at m/z 275-325 and m/z 550-850. b) Chemical structure for LALDI tag-labelled lactose **3.37** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

Appendix V – LALDI-MS spectra of milk labelled with **3.16** (Rxn-1)

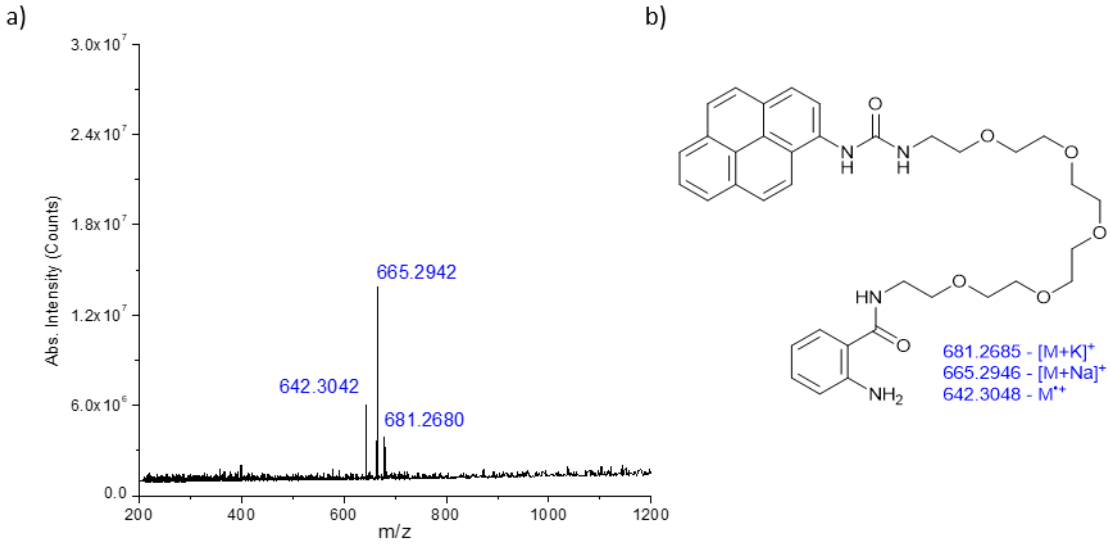


*Figure A.18: LALDI-MS analysis of Rxn-1 following glycan labelling reaction with LALDI tag **3.16**. a) LALDI-MS spectra obtained following direct LALDI-MS analysis of Rxn-1 which shows that only peaks related to **3.16** were observed following LALDI-MS analysis. The identified fragments and adducts are annotated with their observed m/z values. b) Chemical structure for LALDI tag-labelled lactose **3.16** with proposed assignments for fragments and adducts observed in the LALDI-MS spectra. Bonds at which proposed fragmentations occur are indicated by blue dashed lines and are accompanied by the corresponding calculated monoisotopic mass for the fragment.*

Bibliography

1 A. Mandal, M. Singha, P. S. Addy and A. Basak, *Mass Spectrom. Rev.*, 2019, **38**, 3–21.

2 M. Karas and F. Hillenkamp, *Anal. Chem.*, 1988, **60**, 2299–2301.

3 K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida and T. Matsuo, *Rapid Commun. Mass Spectrom.*, 1988, **2**, 151–153.

4 M. Karas, D. Bachmann and F. Hillenkamp, *Anal. Chem.*, 1985, **57**, 2935–2939.

5 R. Zenobi and R. Knochenmuss, *Mass Spectrom. Rev.*, 1998, **17**, 337–366.

6 M. Karas and R. Krüger, *Chem. Rev.*, 2003, **103**, 427–439.

7 L. Molin, R. Seraglia, Z. Czarnocki, J. K. Maurin, F. A. Pluciński and P. Traldi, *J. Anal. Methods Chem.*, 2012, **2012**, 1–8.

8 J. K. Lewis, J. Wei and G. Siuzdak, *Encycl. Anal. Chem.*, 2000, 5880–5894.

9 M. C. Fitzgerald, G. R. Parr and L. M. Smith, *Anal. Chem.*, 1993, **65**, 3204–3211.

10 W. A. Korfmacher, *Using Mass Spectrometry for Drug Metabolism Studies*, CRC Press, Bocta Raton, FL. United States, 2nd edn., 2009.

11 K. Wiangnon and R. Cramer, *Anal. Chem.*, 2015, **87**, 1485–1488.

12 I.-C. Lu, C. Lee, Y.-T. Lee and C.-K. Ni, *Annu. Rev. Anal. Chem.*, 2015, **8**, 21–39.

13 H. Ehring, M. Karas and F. Hillenkamp, *Org. Mass Spectrom.*, 1992, **27**, 472–480.

14 D. A. Allwood, P. E. Dyer, R. W. Dreyfus and I. K. Perera, *Appl. Surf. Sci.*, 1997, **109/110**, 616–620.

15 D. A. Allwood, P. E. Dyer and R. W. Dreyfus, *Rapid Commun. Mass Spectrom.*, 1997, **11**, 499–503.

16 R. Knochenmuss, *Anal. Chem.*, 2002, **75**, 2199–2207.

17 I. C. Lu, C. Lee, H. Y. Chen, H. Y. Lin, S. W. Hung, Y. A. Dyakov, K. T. Hsu, C. Y. Liao, Y. Y. Lee, C. M. Tseng, Y. T. Lee and C. K. Ni, *J. Phys. Chem. B*, 2014, **118**, 4132–4139.

18 K. Y. Chu, S. Lee, M. T. Tsai, I. C. Lu, Y. A. Dyakov, Y. H. Lai, Y. T. Lee and C. K. Ni, *J. Am. Soc. Mass Spectrom.*, 2014, **25**, 310–8.

19 S. Niu, W. Zhang and B. T. Chait, *J. Am. Soc. Mass Spectrom.*, 1997, **9**, 1–7.

20 X. Chen, J. A. Carroll and R. C. Beavis, *J. Am. Soc. Mass Spectrom.*, 1998, **9**, 885–891.

21 Y.-H. Lai, C.-C. Wang, S.-H. Lin, Y. T. Lee and Y.-S. Wang, *J. Phys. Chem. B*, 2010, **114**, 13847–52.

22 F. Hillenkamp, M. Karas, D. Holtkamp and P. Klusener, *Int. J. Mass Spectrom. Ion Process.*, 1986, **69**, 265–276.

23 S. H. Ahn, K. M. Park, Y. J. Bae and M. S. Kim, *J. Mass Spectrom.*, 2013, **48**, 299–305.

24 K. G. Standing and W. Ens, *Methods and mechanisms for producing ions from large*

molecules, New York, NY, 1991.

25 E. Dudley, in *Advancements of Mass Spectrometry in Biomedical Research*, eds. A. G. Woods and C. C. Darie, Springer International Publishing, Cham, 2014, pp. 33–58.

26 A. El-Aneed, A. Cohen and J. Banoub, *Appl. Spectrosc. Rev.*, 2009, **44**, 210–230.

27 J. Zaia, *Chem. Biol.*, 2008, **15**, 881–892.

28 L. Han and C. E. Costello, *Biochem.*, 2013, **78**, 710–720.

29 C. Brunnée, *Int. J. Mass Spectrom. Ion Process.*, 1987, **76**, 125–237.

30 U. Boesl, *Mass Spectrom. Rev.*, 2016, **36**, 86–109.

31 C. Dass, *Principles and Practice of Biological Mass Spectrometry*, Wiley-Interscience, New York City, NY. US, 1st edn., 2001.

32 R. E. March, *Mass Spectrom. Rev.*, 2009, **28**, 961–989.

33 Á. Somogyi, in *Medical Applications of Mass Spectrometry*, eds. K. Vekey, A. Telekes and A. Verte, Elsevier Science, Amsterdam, NL, 1st edn., 2008, pp. 93–140.

34 R. A. Yost and C. G. Enke, *Anal. Chem.*, 1979, **51**, 1251–1264.

35 I. V. Chernushevich, A. V. Loboda and B. A. Thomson, *J. Mass Spectrom.*, 2001, **36**, 849–865.

36 X. Zhang, S. V. B. Garimella, S. A. Prost, I. K. Webb, T. C. Chen, K. Tang, A. V. Tolmachev, R. V. Norheim, E. S. Baker, G. A. Anderson, Y. M. Ibrahim and R. D. Smith, *Anal. Chem.*, 2015, **87**, 6010–6016.

37 R. E. March, in *Encyclopedia of Spectroscopy and Spectrometry*, ed. J. C. Lindon, Elsevier Science, Amsterdam, NL, 1st edn., 1999, pp. 1000–1009.

38 D. Nolting, R. Malek and A. Makarov, *Mass Spectrom. Rev.*, 2017, **Epub ahead**, 1–19.

39 J. P. Stanoeva and M. Stefova, *J. Mass Spectrom.*, 2012, **47**, 1395–1406.

40 L. Bijlsma, E. Emke, F. Hernández and P. De Voogt, *Anal. Chim. Acta*, 2013, **768**, 102–110.

41 S. Vorce, J. Sklerov, K. Kalasinsky and E. McLeman, *J. Anal. Toxicol.*, 2000, **24**, 595–601.

42 F. L. Sauvage, F. Saint-Marcoux, B. Duretz, D. Deporte, G. Lachatre and P. Marquet, *Clin. Chem.*, 2006, **52**, 1735–1742.

43 F. Badalà, K. Nouri-mahdavi and D. A. Raoof, *Trends Anal. Chem*, 2008, **144**, 724–732.

44 G. E. Patterson, A. J. Guymon, L. S. Riter, M. Everly, J. Griep-Raming, B. C. Laughlin, Z. Ouyang and R. G. Cooks, *Anal. Chem.*, 2002, **74**, 6145–6153.

45 D. J. Douglas, A. J. Frank and D. Mao, *Mass Spectrom. Rev.*, 2005, **24**, 1–29.

46 Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman and R. G. Cooks, *J. Mass Spectrom.*, 2005, **40**, 430–443.

47 A. G. Marshall, C. L. Hendrickson and G. S. Jackson, *Mass Spectrom. Rev.*, 1998, **17**, 1–35.

48 E. N. Nikolaev, Y. I. Kostyukevich and G. N. Vladimirov, *Mass Spectrom. Rev.*, 2016, **35**,

219–258.

49 J. C. L. Erve, W. DeMaio and R. E. Talaat, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 3015–3026.

50 M. Ghaste, R. Mistrik and V. Shulaev, *Int. J. Mol. Sci.*, 2016, **17**, 816–838.

51 B. Bogdanov and R. D. Smith, *Mass Spectrom. Rev.*, 2005, **24**, 168–200.

52 F. Girolamo, I. Lante, M. Muraca and L. Putignani, *Curr. Org. Chem.*, 2013, **17**, 2891–2905.

53 J. Webster and D. Oxley, *Protein Identification by MALDI-TOF Mass Spectrometry - Chemical Genomics and Proteomics*, Humana Press, New York City, NY. US, 2012.

54 N. P. J. Price, *Appl. Biochem. Biotechnol.*, 2008, **148**, 271–276.

55 A. R. Buchberger, K. DeLaney, J. Johnson and L. Li, *Anal. Chem.*, 2018, **90**, 240–265.

56 G. Vrioni, C. Tsiamis, G. Oikonomidis, K. Theodoridou, V. Kapsimali and A. Tsakris, *Ann. Transl. Med.*, 2018, **6**, 240–240.

57 R. Ouedraogo, J. Textoris, A. Daumas, C. Capo and J.-L. Mege, in *Immunoproteomics: Methods and Protocols*, eds. K. M. Fulton and S. M. Twine, Humana Press, Totowa, NJ, 2013, pp. 197–209.

58 M. Mrksich, *ACS Nano*, 2008, **2**, 7–18.

59 Z. A. Gurard-Levin, M. D. Scholle, A. H. Eisenberg and M. Mrksich, *ACS Comb. Sci.*, 2011, **13**, 347–350.

60 J. M. Van Munster, B. Thomas, M. Riese, A. L. Davis, C. J. Gray, D. B. Archer and S. L. Flitsch, *Sci. Rep.*, 2017, **7**, 1–13.

61 N. Laurent, J. Voglmeir, A. Wright, J. Blackburn, N. T. Pham, S. C. C. Wong, S. J. Gaskell and S. L. Flitsch, *ChemBioChem*, 2008, **9**, 883–887.

62 P. T. O'Kane and M. Mrksich, *J. Am. Chem. Soc.*, 2017, **139**, 10320–10327.

63 L. L. Anderson, E. J. Berns, P. Bugga, A. L. George and M. Mrksich, *Anal. Chem.*, 2016, **88**, 8604–8609.

64 T. J. Montavon, J. Li, J. R. Cabrera-Pardo, M. Mrksich and S. A. Kozmin, *Nat. Chem.*, 2012, **4**, 45–51.

65 N. Laurent, R. Haddoub, J. Voglmeir, S. C. C. Wong, S. J. Gaskell and S. L. Flitsch, *Chembiochem*, 2008, **9**, 2592–2596.

66 N. Zhong, Y. Cui, X. Zhou, T. Li and J. Han, *Tumor Biol.*, 2014, **36**, 1221–1231.

67 J. Albrethsen, *Clin. Chem.*, 2007, **53**, 852–858.

68 K. O. Börnsen, in *Mass Spectrometry of Proteins and Peptides. Methods in Molecular Biology™, vol 146*, ed. J. R. Chapman, Humana Press, Totowa, NJ, 2000, pp. 387–404.

69 K. T. M. Myasein, J. S. Pulido, R. M. Hatfield, C. A. McCannel, R. F. Dundervill and S. A. Shippy, *Analyst*, 2007, **132**, 1046–1052.

70 H. W. Chu, B. Unnikrishnan, A. Anand, J. Y. Mao and C. C. Huang, *J. Food Drug Anal.*, 2018, **26**, 1215–1228.

71 B. Unnikrishnan, C. Y. Chang, H. W. Chu, A. Anand and C. C. Huang, *Anal. Methods*, 2016, **8**, 8123–8133.

72 R. Nayak and D. R. Knapp, *Anal. Chem.*, 2010, **82**, 7772–7778.

73 M. O. Amin, M. Madkour and E. Al-Hetlani, *Anal. Bioanal. Chem.*, 2018, **410**, 4815–4827.

74 N. Karousis, N. Tagmatarchis and D. Tasis, *Chem. Rev.*, 2010, **110**, 5366–5397.

75 J. Sunner, E. Dratz and Y. C. Chen, *Anal. Chem.*, 1995, **67**, 4335–4342.

76 H. W. Tang, K. M. Ng, W. Lu and C. M. Che, *Anal. Chem.*, 2009, **81**, 4720–4729.

77 X. Dong, J. Cheng, J. Li and Y. Wang, *Anal. Chem.*, 2010, **82**, 6208–6214.

78 M. Lu, Y. Lai, G. Chen and Z. Cai, *Anal. Chem.*, 2011, **83**, 3161–3169.

79 B. Gulbakan, E. Yasun, M. I. Shukoor, Z. Zhu, M. You, X. Tan, H. Sanchez, D. H. Powell, H. Dai and W. Tan, *J. Am. Chem. Soc.*, 2010, **132**, 17408–17410.

80 P. J. O'Brien, M. Lee, M. E. Spilker, C. C. Zhang, Z. Yan, T. C. Nichols, W. Li, C. H. Johnson, G. J. Patti and G. Siuzdak, *Cancer Metab.*, 2013, **1**, 4.

81 T. M. Guinan, D. Neldner, P. Stockham, H. Kobus, C. B. Della Vedova and N. H. Voelcker, *Drug Test. Anal.*, 2017, **9**, 769–777.

82 Z. Shen, J. J. Thomas, C. Averbuj, K. M. Broo, M. Engelhard, J. E. Crowell, M. G. Finn and G. Siuzdak, *Anal. Chem.*, 2001, **73**, 612–619.

83 T. R. Northen, O. Yanes, M. T. Northen, D. Marrinucci, W. Uritboonthai, J. Apon, S. L. Golledge, A. Nordström and G. Siuzdak, *Nature*, 2007, **449**, 1033–1036.

84 R. H. Daniels, S. Dikler, E. Li and C. Stacey, *J. Lab. Autom.*, 2008, **13**, 314–321.

85 T. Guinan, M. Ronci, R. Vasani, H. Kobus and N. H. Voelcker, *Talanta*, 2015, **132**, 494–502.

86 C.-S. Lee, E.-M. Kim, S.-H. Lee, M.-S. Kim, Y.-K. Kim and B.-G. Kim, *Biotechnol. Bioprocess Eng.*, 2005, **10**, 212.

87 G. Piret, H. Drobecq, Y. Coffmier, O. Melnyk and R. Boukherroub, *Langmuir*, 2010, **26**, 1354–1361.

88 J. J. Thomas, Z. Shen, R. Blackledge and G. Siuzdak, *Anal. Chim. Acta*, 2001, **442**, 183–190.

89 C. Y. Lo, J. Y. Lin, W. Y. Chen, C. T. Chen and Y. C. Chen, *J. Am. Soc. Mass Spectrom.*, 2008, **19**, 1014–1020.

90 Y. S. Lin and Y. C. Chen, *Anal. Chem.*, 2002, **74**, 5793–5798.

91 C. T. Chen and Y. C. Chen, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 1956–1964.

92 Y. S. Lin, C. H. Yang and Y. C. Chen, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 313–318.

93 V. L. Brown, Q. Liu and L. He, in *Mass Spectrometry Imaging of Small Molecules*, ed. L. He, Springer New York, New York, NY, 2015, pp. 175–184.

94 J. R. Cabrera-Pardo, D. I. Chai, S. Liu, M. Mrksich and S. A. Kozmin, *Nat. Chem.*, 2013, **5**, 423–427.

95 K. Yoshino, T. Takao, H. Murata and Y. Shimonishi, *Anal. Chem.*, 1995, **67**, 4028–4031.

96 Y. H. Ahn and J. S. Yoo, *Rapid Commun. Mass Spectrom.*, 1998, **12**, 2011–2015.

97 J. W. Szewczyk, R. L. Zuckerman, R. G. Bergman and J. A. Ellman, *Angew. Chemie Int. Ed.*, 2001, **40**, 216–219.

98 T. McCarley, R. McCarley and P. Limbach, *Anal. Chem.*, 1998, **70**, 4376–4379.

99 S. F. Macha, T. D. McCarley and P. A. Limbach, *Anal. Chim. Acta*, 1999, **397**, 235–245.

100 K. Yoneda, Y. Hu, M. Kita and H. Kigoshi, *Sci. Rep.*, 2015, **5**, 17853.

101 P. S. Addy, S. Basu Roy, S. M. Mandal and A. Basak, *RSC Adv.*, 2014, **4**, 23314–23318.

102 P. S. Addy, A. Bhattacharya, S. M. Mandal and A. Basak, *RSC Adv.*, 2014, **4**, 46555–46560.

103 A. Mandal, A. K. Das and A. Basak, *RSC Adv.*, 2015, **5**, 106912–106917.

104 A. Mandal, A. Maity, S. Bag, P. Bhattacharya, A. K. Das and A. Basak, *RSC Adv.*, 2017, **7**, 7163–7169.

105 T. A. Andersson, K. M. Hartonen and M.-L. Riekkola, *J. Chem. Eng. Data*, 2005, **50**, 1177–1183.

106 K. Yoneda, Y. Hu, R. Watanabe, M. Kita and H. Kigoshi, *Org. Biomol. Chem.*, 2016, **14**, 8564–8569.

107 R. Watanabe, Y. Hu, K. Iio, K. Yoneda, A. Hattori, A. Arai, H. Kigoshi and M. Kita, *Org. Biomol. Chem.*, 2018, **16**, 7883–7890.

108 M. Kita, O. Ohno, C. Han and D. Uemura, *Chem. Rec.*, 2010, **10**, 57–69.

109 M. Kita, Y. Hirayama, M. Sugiyama and H. Kigoshi, *Angew. Chemie Int. Ed.*, 2011, **50**, 9871–9874, S9871/1-S9871/15.

110 M. Kita, K. Yoneda, Y. Hirayama, K. Yamagishi, Y. Saito and Y. Sugiyama, *ChemBioChem*, 2012, **8522**, 1754–1758.

111 R. E. West, E. W. Findsen and D. Isailovic, *Int. J. Mass Spectrom.*, 2013, **353**, 54–59.

112 R. E. West, J. B. Jacobs and D. Isailovic, *Int. J. Mass Spectrom.*, 2015, **389**, 39–46.

113 K. F. Aoki-Kinoshita, *PLoS Comput. Biol.*, 2008, **4**, 1–7.

114 A. Varki and J. B. Lowe, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. US, 2nd edn., 2009.

115 H. J. An, S. R. Kronewitter, M. L. A. de Leoz and C. B. Lebrilla, *Curr. Opin. Chem. Biol.*, 2009, **13**, 601–607.

116 C. R. Bertozzi and R. Sasisekharan, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler,

Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. US, 2nd edn., 2009.

117 J.-P. Schermann, in *Spectroscopy and Modeling of Biomolecular Building Blocks*, ed. J.-P. Schermann, Elsevier Science, Amsterdam, NL, 1st edn., 2008, pp. 297–307.

118 R. A. Laine, *Glycobiology*, 1994, **4**, 759–767.

119 H. S. Overkleeft and P. H. Seeberger, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, T. Darvill, Alan G. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. US, 3rd edn., 2017.

120 R. J. Linhardt and H. G. Bazin, in *Glycoscience: Chemistry and Chemical Biology I--III*, eds. B. O. Fraser-Reid, K. Tatsuta and J. Thiem, Springer Berlin Heidelberg, Berlin, Heidelberg, 2001, pp. 63–74.

121 M. Nagae and Y. Yamaguchi, *Int. J. Mol. Sci.*, 2012, **13**, 8398–8429.

122 J. Duus, C. H. Gotfredsen and K. Bock, *Chem. Rev.*, 2000, **100**, 4589–4614.

123 M. J. Kailemia, L. R. Ruhaak, C. B. Lebrilla and I. J. Amster, *Anal. Chem.*, 2014, **86**, 196–212.

124 N. Leymarie and J. Zaia, *Anal. Chem.*, 2012, **84**, 3040–3048.

125 J. Zaia, *Mass Spectrom. Rev.*, 2004, **23**, 161–227.

126 Y. Wada, P. Azadi, C. E. Costello, A. Dell, R. A. Dwek, H. Geyer, R. Geyer, K. Kakehi, N. G. Karlsson, K. Kato, N. Kawasaki, K. H. Khoo, S. Kim, A. Kondo, E. Lattova, Y. Mechref, E. Miyoshi, K. Nakamura, H. Narimatsu, M. V. Novotny, N. H. Packer, H. Perreault, J. Peter-Katalinić, G. Pohlentz, V. N. Reinhold, P. M. Rudd, A. Suzuki and N. Taniguchi, *Glycobiology*, 2007, **17**, 411–422.

127 J. Zaia, in *The Encyclopedia of Mass Spectrometry.Volume 6: Ionization Methods*, eds. M. L. Gross and R. M. Caprioli, Elsevier Science, Amsterdam, NL, 1st edn., 2006, pp. 889–903.

128 M. B. Satterfield and M. J. Welch, *Clin. Biochem.*, 2005, **38**, 166–174.

129 S. H. Walker, A. D. Taylor and D. C. Muddiman, *J. Am. Soc. Mass Spectrom.*, 2013, **24**, 1376–1384.

130 J. Zaia, *Mass Spectrom. Rev.*, 2009, **28**, 254–272.

131 K. K. Palaniappan and C. R. Bertozzi, *Chem. Rev.*, 2016, **116**, 14277–14306.

132 J. P. Kamerling and G. J. Gerwig, in *Comprehensive Glycoscience: From Chemistry to system Biology.*, ed. H. Kamerling, Elsevier, Oxford, 2007, pp. 1–68.

133 V. Tretter, F. Altmann and L. März, *Eur. J. Biochem.*, 1991, **199**, 647–652.

134 A. H. Merry, D. C. A. Neville, L. Royle, B. Matthews, D. J. Harvey, R. A. Dwek and P. M. Rudd, *Anal. Biochem.*, 2002, **304**, 91–99.

135 G. J. Rademaker, S. A. Pergantis, L. Blok-Tip, J. I. Langridge, A. Kleen and J. E. Thomas-

Oates, *Anal. Biochem.*, 1998, **257**, 149–160.

136 K. R. Anumula, *Anal. Biochem.*, 2008, **37**, 104–111.

137 S. ichi Nakakita, W. Sumiyoshi, N. Miyanishi and J. Hirabayashi, *Biochem. Biophys. Res. Commun.*, 2007, **362**, 639–645.

138 G. C. M. Vreeker and M. Wuhrer, *Anal. Bioanal. Chem.*, 2017, **409**, 359–378.

139 M. Wuhrer, A. R. de Boer and A. M. Deelder, *Mass Spectrom. Rev.*, 2009, **28**, 192–206.

140 L. R. Ruhaak, A. M. Deelder and M. Wuhrer, *Anal. Bioanal. Chem.*, 2009, **394**, 163–174.

141 Y. Mu, B. L. Schulz and V. Ferro, *Molecules*, 2018, **23**, 1–17.

142 Y. Mechref, P. Kang and M. V. Novotny, *Rapid Commun. Mass Spectrom.*, 2006, **20**, 1381–1389.

143 K. S. Lancaster, H. J. An, B. Li and C. B. Lebrilla, *Anal. Chem.*, 2006, **78**, 4990–4997.

144 C. Zhao, B. Xie, S. Y. Chan, C. E. Costello and P. B. O'Connor, *J. Am. Soc. Mass Spectrom.*, 2008, **19**, 138–150.

145 X. Yu, Y. Huang, C. Lin and C. E. Costello, *Anal. Chem.*, 2012, **87**, 7487–7494.

146 J. J. Wolff, I. J. Amster, L. Chi and R. J. Linhardt, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 234–244.

147 B. Mulloy, A. Dell, P. Stanley and J. H. Prestegard, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, T. Darvill, Alan G. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. US, 3rd edn., 2017.

148 S. Prime and T. Merry, in *Glycoanalysis Protocols*, ed. E. F. Hounsell, Humana Press, Totowa, NJ, 1998, pp. 53–69.

149 A. Shajahan, C. Heiss, M. Ishihara and P. Azadi, *Anal. Bioanal. Chem.*, 2017, **409**, 4483–4505.

150 C.-Y. Yu, A. Mayampurath and H. Tang, in *Mass Spectrometry of Glycoproteins: Methods and Protocols*, eds. J. J. Kohler and S. M. Patrie, Humana Press, Totowa, NJ, 2013, pp. 269–276.

151 D. B. Hizal, D. Wolozny, J. Colao, E. Jacobson, Y. Tian, S. S. Krag, M. J. Betenbaugh and H. Zhang, *Clin. Proteomics*, 2014, **11**, 1–10.

152 D.-S. Lee, C. Wu and H. H. Hill Jr, *J. Chromatogr. A*, 1998, **822**, 1–9.

153 W. Morelle, V. Faid and J. C. Michalski, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 2451–2464.

154 I. Ciucanu and F. Kerek, *Carbohydr. Res.*, 1984, **131**, 209–217.

155 W. Morelle and J. Michalski, *Nat. Protoc.*, 2007, **2**, 1585–1602.

156 R. C. Robinson, N. A. Poulsen and D. Barile, *PLoS One*, 2018, **13**, 1–18.

157 D. K. Kölmel and E. T. Kool, *Chem. Rev.*, 2017, **117**, 10358–10376.

158 P. Crisalli and E. T. Kool, *J. Org. Chem.*, 2013, **78**, 1184–1189.

159 M. B. Thygesen, H. Munch, J. Sauer, E. Cló, M. R. Jørgensen, O. Hindsgaul and K. J. Jensen, *J. Org. Chem.*, 2010, **75**, 1752–1755.

160 J. J. Reina, A. Rioboo and J. Montenegro, *Synth.*, 2018, **50**, 831–845.

161 Y. Zhang, M. Yu, C. Zhang, W. Ma, Y. Zhang, C. Wang and H. Lu, *Anal. Chem.*, 2014, **86**, 7920–7924.

162 M. B. Thygesen, K. K. Sørensen, E. Cló and K. J. Jensen, *Chem. Commun.*, 2009, 6367–6369.

163 A. Hoang, E. Laigre, D. Goyard, E. Defrancq, F. Vinet, P. Dumy and O. Renaudet, *Org. Biomol. Chem.*, 2017, **15**, 5135–5139.

164 J. Amano, D. Sugahara, K. Osumi and K. K. Tanaka, *Glycobiology*, 2009, **19**, 592–600.

165 G. T. Hermanson, in *Bioconjugate Techniques*, ed. G. T. Hermanson, Academic Press, Boston, 3rd edn., 2013, pp. 229–258.

166 Y. Zhang, B. Wang, W. Jin, Y. Wen, L. Nan, M. Yang, R. Liu, Y. Zhu, C. Wang, L. Huang, X. Song and Z. Wang, *Anal. Chim. Acta*, 2018, **1048**, 105–114.

167 C. L. O'Neil, K. J. Stine and A. V. Demchenko, *J. Carbohydr. Chem.*, 2018, **37**, 225–249.

168 T. Nishikaze, T. Nakamura, H. Jinmei and J. Amano, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.*, 2011, **879**, 1419–1428.

169 J. Amano, M. Osanai, T. Orita, D. Sugahara and K. Osumi, *Glycobiology*, 2009, **19**, 601–614.

170 T. Nishikaze, H. Okumura, H. Jinmei and J. Amano, *Int. J. Mass Spectrom.*, 2013, **333**, 8–14.

171 L. R. Ruhaak, G. Zauner, C. Huhn, C. Bruggink, A. M. Deelder and M. Wuhrer, *Anal. Bioanal. Chem.*, 2010, **397**, 3457–3481.

172 K. R. Anumula, *Anal. Biochem.*, 2006, **350**, 1–23.

173 B. Xia, C. L. Feasley, G. P. Sachdev, D. F. Smith and R. D. Cummings, *Anal. Biochem.*, 2009, **387**, 162–170.

174 G. Ridlova, J. C. Mortimer, S. L. Maslen, P. Dupree and E. Stephens, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 2723–2730.

175 J. C. Bigge, T. P. Patel, J. A. Bruce, P. N. Goulding, S. M. Charles and R. B. Parekh, *Anal. Biochem.*, 1995, **230**, 229–238.

176 L. Veillon, Y. Huang, W. Peng, X. Dong, B. G. Cho and Y. Mechref, *Electrophoresis*, 2017, **38**, 2100–2114.

177 L. Royle, M. P. Campbell, C. M. Radcliffe, D. M. White, D. J. Harvey, J. L. Abrahams, Y. G.

Kim, G. W. Henry, N. A. Shadick, M. E. Weinblatt, D. M. Lee, P. M. Rudd and R. A. Dwek, *Anal. Biochem.*, 2008, **376**, 1–12.

178 N. Takahashi, K. B. Lee, H. Nakagawa, Y. Tsukamoto, Y. Kawamura, Y. T. Li and Y. C. Lee, *Anal. Biochem.*, 1995, **230**, 333–342.

179 J. Lemoine, M. Cabanes-Macheteau, M. Bardor, J. Michalski, L. Faye and P. Lerouge, *Rapid Commun. Mass Spectrom.*, 2000, **14**, 100–104.

180 H. Suzuki, O. Muller, A. Guttman and B. L. Karger, *Anal. Chem.*, 1997, **69**, 4554–4559.

181 G. Lu, C. L. Crihfield, S. Gattu, L. M. Veltri and L. A. Holland, *Chem. Rev.*, 2018, **118**, 7867–7885.

182 J. You, X. Sheng, C. Ding, Z. Sun, Y. Suo, H. Wang and Y. Li, *Anal. Chim. Acta*, 2008, **609**, 66–75.

183 G. Zauner, C. A. M. Koeleman, A. M. Deelder and M. Wuhrer, *Biochim. Biophys. Acta - Gen. Subj.*, 2012, **1820**, 1420–1428.

184 C. Wang, W. Fan, P. Zhang, Z. Wang and L. Huang, *Proteomics*, 2012, **11**, 4229–4242.

185 C. Wang, P. Zhang, W. Jin, L. Li, S. Qiang, Y. Zhang, L. Huang and Z. Wang, *J. Proteomics*, 2017, **150**, 18–30.

186 P. Zhang, Y. Zhang, X. Xue, C. Wang, Z. Wang and L. Huang, *Anal. Biochem.*, 2011, **418**, 1–9.

187 Z. Sun, C. Song, L. Xia, X. Wang, Y. Suo and J. You, *Chromatographia*, 2010, **71**, 789–797.

188 Y. Lu, C. Wang, R. Liu, W. Jin, Y. Wen, L. Huang and Z. Wang, *Anal. Biochem.*, 2018, **549**, 1–11.

189 L. K. Mahal, K. J. Yarema and C. R. Bertozzi, *Science*, 1997, **276**, 1125–1128.

190 T. J. Sminia, H. Zuilhof and T. Wennekes, *Carbohydr. Res.*, 2016, **435**, 121–141.

191 M. Wendeler, L. Grinberg, X. Wang, P. E. Dawson and M. Baca, *Bioconjug. Chem.*, 2014, **25**, 93–101.

192 M. W. Duncan, D. Nedelkov, R. Walsh and S. J. Hattan, *Clin. Chem.*, 2016, **62**, 134–143.

193 A. P. Kafka, T. Kleffmann, T. Rades and A. McDowell, *Int. J. Pharm.*, 2011, **417**, 70–82.

194 S. C. Brown, G. Kruppa and J. L. Dasseux, *Mass Spectrom. Rev.*, 2005, **24**, 223–231.

195 J. Murray, D. Nowak, L. Pukenas, R. Azhar, M. Guillorit, C. Wälti, K. Critchley, S. Johnson and R. S. Bon, *J. Mater. Chem. B*, 2014, **2**, 3741–3744.

196 A. L. Dif, F. Boulmedais, M. Pinot, V. Roullier, M. Baudy-Floc'h, F. M. Coquelle, S. Clarke, P. Neveu, F. Vignaux, R. Le Borgne, M. Dahan, Z. Gueroui and V. Marchi-Artzner, *J. Am. Chem. Soc.*, 2009, **131**, 14738–14746.

197 H. a Chokhawala, S. Huang, K. Lau, H. Yu, J. Cheng, V. Thon, N. Hurtado-Ziola, J. a Guerrero, A. Varki and X. Chen, *ACS Chem. Biol.*, 2008, **3**, 567–576.

198 S. Svedhem, C. a. Hollander, J. Shi, P. Konradsson, B. Liedberg and S. C. T. Svensson, *J. Org. Chem.*, 2001, **66**, 4494–4503.

199 A. Chiva, D. E. Williams, A. B. Tabor and H. C. Hailes, *Tetrahedron Lett.*, 2010, **51**, 2720–2723.

200 L. Xu, Y. C. Chen, J. Chong, A. Fin, L. S. McCoy, J. Xu, C. Zhang and D. Wang, *Angew. Chemie - Int. Ed.*, 2014, **53**, 11223–11227.

201 H. Ji, Y. Yang, X. Xu and G. Brown, *Org. Biomol. Chem.*, 2006, **4**, 770–772.

202 K. J. Smit, *J. Energ. Mater.*, 1991, **9**, 81–103.

203 K. K. Sharma, G. H. Kannikanti, T. R. R. Baggi and J. R. Vaidya, *Methods Appl. Fluoresc.*, 2018, **6**, 035004.

204 H. Moriwaki, *J. Mass Spectrom.*, 2016, **51**, 1096–1102.

205 S. F. Macha, P. a. Limbach and P. J. Savickas, *J. Am. Soc. Mass Spectrom.*, 2000, **11**, 731–737.

206 F. Tureček, *Mass Spectrom.*, 2013, **2**, S0003–S0003.

207 C. C. Liou and J. S. Brodbelt, *J. Am. Soc. Mass Spectrom.*, 1992, **3**, 543–548.

208 A. I. Gusev, W. R. Wilkinson, A. Proctor and D. M. Hercules, *Anal. Chem.*, 1995, **67**, 1034–1041.

209 R. D. Deegan, O. Bakajin, T. F. Dupont, G. Huber, S. R. Nagel and T. A. Witten, *Nature*, 1997, **389**, 827–829.

210 WO 2006083151 A1, 10th August, 2006.

211 V. C. Armstrong, D. W. Farlow and R. B. Moodie, *J. Chem. Soc. B*, 1968, 1099–1103.

212 R. S. Bon, C. Hong, M. J. Bouma, R. F. Schmitz, F. J. J. De Kanter, M. Lutz, A. L. Spek and R. V. A. Orru, *Org. Lett.*, 2003, **5**, 3759–3762.

213 X. Y. Qian, S. Q. Li, J. Song and H. C. Xu, *ACS Catal.*, 2017, **7**, 2730–2734.

214 F. Biedermann, E. A. Appel, T. Gruendling, C. Barner-kowollik and O. A. Scherman, *Macromolecules*, 2011, **44**, 4828–4835.

215 JP 2011059656 A, 2011.

216 D. Podhradský, Ľ. Drobnica and P. Kristian, *Experientia*, 1979, **35**, 154–155.

217 A. Jorbágy and K. Király, *Biochim. Biophys. Acta - Gen. Subj.*, 1966, **124**, 166–175.

218 K. M. Frazier and T. M. Swager, *Anal. Chem.*, 2013, **85**, 7154–7158.

219 S. Ulrich, D. Boturyn, A. Marra, O. Renaudet and P. Dumy, *Chem. - A Eur. J.*, 2014, **20**, 34–41.

220 H. S. Ewan, C. S. Muli, S. Touba, A. T. Bellinghiere, A. M. Veitschegger, T. B. Smith, W. L. Pistel, W. T. Jewell, R. K. Rowe, J. P. Hagen and H. Palandoken, *Tetrahedron Lett.*, 2014, **55**, 4962–4965.

221 K. D. McReynolds, D. Dimas and H. Le, *Tetrahedron Lett.*, 2014, **55**, 2270–2273.

222 S. R. King, E. S. Hecht and D. C. Muddiman, *Anal. Bioanal. Chem.*, 2018, **410**, 1409–1415.

223 S. Zhou, L. Veillon, X. Dong, Y. Huang and Y. Mechref, *Analyst*, 2017, **142**, 4446–4455.

224 D. Locke, C. G. Bevans, L. X. Wang, Y. Zhang, A. L. Harris and Y. C. Lee, *Carbohydr. Res.*, 2004, **339**, 221–231.

225 T. Keser, T. Pavić, G. Lauc and O. Gornik, *Front. Chem.*, 2018, **6**, 1–12.

226 R. Novoa-Carballal and A. H. E. Müller, *Chem. Commun.*, 2012, **48**, 3781.

227 K. R. Anumula, *Anal. Biochem.*, 2014, **457**, 31–37.

228 WO 2008072976 A2, 19th June, 2008.

229 WO 2013053657 A1, 18th April, 2013.

230 WO 2015038417 A1, 19th March, 2015.

231 J. Li, J. Chen, C. Gui, L. Zhang, Y. Qin, Q. Xu, J. Zhang, L. Liu, X. Shen and H. Jiang, *Bioorg. Med. Chem.*, 2006, **14**, 2209–2224.

232 T. N. Le, S. H. Yang, D. B. Khadka, S. H. Cho, C. Zhao and W. J. Cho, *Bull. Korean Chem. Soc.*, 2011, **32**, 4309–4315.

233 H. L. Yale, K. Losee, J. Martins, M. Holsing, F. M. Perry and J. Bernstein, *J. Am. Chem. Soc.*, 1953, **75**, 1933–1942.

234 A. J. Ruiz-Sanchez, P. L. Higgs, D. T. Peters, A. T. Turley, M. A. Dobson, A. J. North and D. A. Fulton, *ACS Macro Lett.*, 2017, **6**, 903–907.

235 R. L. Burwell, *Chem. Rev.*, 1954, **54**, 615–685.

236 A. A. Burke, E. S. Severson, S. Mool, M. J. Solares Bucaro, F. T. Greenaway and C. E. Jakobsche, *J. Enzyme Inhib. Med. Chem.*, 2017, **32**, 496–503.

237 A. M. L. Winther, H. Liu, Y. Sonntag, C. Olesen, M. Le Maire, H. Soehoel, C. E. Olsen, S. B. Christensen, P. Nissen and J. V. Møller, *J. Biol. Chem.*, 2010, **285**, 28883–28892.

238 D. J. Harvey, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.*, 2011, **879**, 1196–1225.

239 T. N. Laremore and R. J. Linhardt, *Rapid Commun. Mass Spectrom.*, 2007, **21**, 1315–1320.

240 T. Yamagaki, H. Suzuki and K. Tachibana, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 714–723.

241 C. L. Su and W. L. Tseng, *Anal. Chem.*, 2007, **79**, 1626–1633.

242 T. Watanabe, H. Kawasaki, T. Yonezawa and R. Arakawa, *J. Mass Spectrom.*, 2008, **43**, 1063–1071.

243 C. L. Wu, C. C. Wang, Y. H. Lai, H. Lee, J. Der Lin, Y. T. Lee and Y. S. Wang, *Anal. Chem.*, 2013, **85**, 3836–3841.

244 C. H. Grün, S. J. van Vliet, W. E. C. M. Schiphorst, C. M. C. Bank, S. Meyer, I. van Die and Y. van Kooyk, *Anal. Biochem.*, 2006, **354**, 54–63.

245 J. Kalia and R. T. Raines, *Angew. Chemie Int. Ed. English*, 2008, **47**, 7523–7526.

246 Z. Roth, G. Yehezkel and I. Khalaila, *Int. J. Carbohydr. Chem.*, 2012, **2012**, 1–10.

247 N. S. Scrimshaw and E. B. Murray, *Am. J. Clin. Nutr.*, 1988, **48**, 1099–1104.

248 A. M. Schaiberger and J. A. Moss, *J. Am. Soc. Mass Spectrom.*, 2008, **19**, 614–619.

249 L. Patiny, ChemCalc: Molecular formula from monisotopic mass, https://www.chemcalc.org/mf_finder/mfFinder_em_new, (accessed 22 July 2019).

250 L. Patiny and A. Borel, *J. Chem. Inf. Model.*, 2013, **53**, 1223–1228.

251 G. Liang, H. Ren and J. Rao, *Nat. Chem.*, 2010, **2**, 54–60.

252 A. Godinat, H. M. Park, S. C. Miller, K. Cheng, D. Hanahan, L. E. Sanman, M. Bogyo, A. Yu, G. F. Nikitin, A. Stahl and E. A. Dubikovskaya, *ACS Chem. Biol.*, 2013, **8**, 987–999.

253 P. Wang, C. J. Zhang, G. Chen, Z. Na, S. Q. Yao and H. Sun, *Chem. Commun.*, 2013, **49**, 8644–8646.

254 J. Jeon, B. Shen, L. Xiong, Z. Miao, K. H. Lee, J. Rao and F. T. Chin, *Bioconjug. Chem.*, 2012, **23**, 1902–1908.

255 D. P. Nguyen, T. Elliott, M. Holt, T. W. Muir and J. W. Chin, *J. Am. Chem. Soc.*, 2011, **133**, 11418–11421.

256 H. Ren, F. Xiao, K. Zhan, Y. P. Kim, H. Xie, Z. Xia and J. Rao, *Angew. Chemie Int. Ed.*, 2009, **48**, 9658–9662.

257 H. A. Beard, University of Leeds, 2018.

258 J. Murray, University of Leeds, 2014.

259 K. Lang and J. W. Chin, *ACS Chem. Biol.*, 2014, **9**, 16–20.

260 L. Katz, *J. Am. Chem. Soc.*, 1951, **73**, 4007–4010.

261 F. D. Bellamy and K. Ou, *Tetrahedron Lett.*, 1984, **25**, 839–842.

262 US 20110213124 A1, 1st September, 2011.

263 H. Takakura, R. Kojima, Y. Urano, T. Terai, K. Hanaoka and T. Nagano, *Chem. – An Asian J.*, 2011, **6**, 1800–1810.

264 D. C. McCutcheon, W. B. Porterfield and J. A. Prescher, *Org. Biomol. Chem.*, 2015, **13**, 2117–2121.

265 W. B. Porterfield, K. A. Jones, D. C. McCutcheon and J. A. Prescher, *J. Am. Chem. Soc.*, 2015, **137**, 8656–8659.

266 D. C. Mccutcheon, M. A. Paley, R. C. Steinhardt and J. A. Prescher, *J. Am. Chem. Soc.*, 2012, **134**, 7604–7607.

267 R. C. Steinhardt, J. M. O'Neill, C. M. Rathbun, D. C. McCutcheon, M. A. Paley and J. A. Prescher, *Chem. - A Eur. J.*, 2016, **22**, 3671–3675.

268 The National Institute for Occupational Safety and Health (NIOSH): 'Sulfur monochloride''. 4th December 2014.', https://www.cdc.gov/niosh/idlh/10025679.html,

(accessed 10 May 2019).

269 WO 2005011657 A2, 10th February, 2005.

270 A. Pinner and F. Klein, *Berichte der Dtsch. Chem. Gesellschaft*, 1877, **10**, 1889–1897.

271 G. B. Barlin and S. J. Ireland, *Aust. J. Chem.*, 1985, **38**, 1685–1691.

272 J. P. Collman, Y. L. Yan, J. Lei and P. H. Dinolfo, *Org. Lett.*, 2006, **8**, 923–926.

273 D. K. Sharma, S. T. Adams, K. L. Liebmann and S. C. Miller, *Org. Lett.*, 2017, **19**, 5836–5839.