

Fluorescent methods to detect and discover peptide and protein interactions *in vitro*

Joy B. Ghayche

Supervisor: Prof. Elizabeth J. New

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

The University of Sydney
School of Chemistry
(Faculty of Science)

2023

Abstract

Peptides and proteins are responsible for a diverse range of biological functions including catalysis, structure, transport, signalling and defence. The unique properties of peptide and protein interactions have been exploited in the field of medicine for uses like protein-targeted therapeutic, and diagnostic agents. However, such developments would not be possible without an understanding of the specific interactions of these biomolecules. As such, it is essential to develop methods that allow the study of these unique interactions. In this work, small fluorescent molecules were utilised for the development of such techniques.

A novel probe, **CyMT** was developed for multimodal mass spectrometry and fluorescence imaging within a single sample. The probe proved to be successful in the imaging of isolated insulin protein samples, first with fluorescence microscopy, followed immediately with mass spectrometry imaging. A second probe, **NpMT** was also designed and synthesised for future testing in this multimodal imaging setup, as well as assessment of a multiplexed multimodal setup.

Following from previous work on the development of a novel fluorescent redox sensor for protein tagging, the assessment of novel fluorescent scaffolds was performed. The fluorescent properties of **FCR1** and **NpFR1** were assessed computationally for compatibility with the HaloTag protein tagging system. This work revealed the suitability of **FCR1** for future incorporation into the HaloTag. Further computational studies on **NpFR1** helped to uncover the most suitable computation methods that would allow for further assessment and future developments of this sensor.

Finally, efforts were directed towards developing a novel peptide-receptor binding assay. The fluorescence probe, **BNp-COOH**, was designed with two functional moieties - one for solid-phase peptide synthesis attachment, and another as a pull-down partner. The fluorescence allows identification of receptor binding events, while the pull-down partner enables subsequent isolation and characterisation of the interaction.

The work outlined here, describes a novel set of tools for the study of peptide and protein interactions. These tools have the potential to enhance our understanding of these interactions, and aid in the development of novel treatments for the betterment of human health.

Statement of originality

This is to certify that, to the best of my knowledge, the content of this thesis is my own work undertaken at the School of Chemistry at The University of Sydney. This thesis has not been submitted for any other degree or purposes. This thesis contains work carried out by myself except where otherwise acknowledged.

Authorship attribution statement

Parts of Chapter 1 in this thesis are published as:

Ghrayche, J.; Graziotto, M. E.; Jeffcoat, P. I.; New, E. J. Labelling proteins and biomolecules with small fluorescent sensors. *Molecular Fluorescent Sensors for Cellular Studies 2022*, 201–239 DOI: 10.1002/9781119749844.ch7.

The sections reproduced in this thesis are exclusively my own work.

Joy Ghrayche

28th February 2023

As supervisor for the candidature upon which this thesis is based, I can confirm that the statements above are correct.

Elizabeth New

28th February 2023

COVID-19 impact statement

The COVID-19 pandemic significantly impacted the work presented in this thesis, as several government-enforced lock-downs occurred in Sydney during the research period. These lock-downs occurred on the following dates:

1. 18th March 2020 – 15th May 2020
2. 9th July 2020 – 31st July 2020
3. 26th June 2021 – 11th October 2021

During the lock-downs, closure of the university meant that I had no access to lab facilities. Given the hands-on nature of this research, no work could be completed during those times. Aside from these lock-downs, the university experienced isolated lock-downs due to positive cases in the building. This occurred on a number of occasions, causing major disruption to planned time-sensitive studies. Because of this, many overnight reactions could not be performed due to the safety concerns of leaving these reactions unattended in the event of an unforeseen shutdown.

Many of the compounds prepared for this work were found to be unstable, and therefore degraded during the course of the lock-downs. As a result, several months worth of synthesis had to be repeated to obtain these compounds and purify them for testing once again.

To mitigate these challenges, studies in chapter 3 focussed on computation chemistry work. I learnt computational chemistry with no background knowledge for this project. However, the original plan was to synthesise **FCR1-Tag** and **NpFR1-Tag** for testing with various HaloTag-modified cells that had already been obtained. **FCR1** and **NpFR1** with acid handles for this synthesis were already designed and synthesised, but due to time constraints, the final synthesis and testing could not be completed.

For the work in chapter 4, I had received multiple rounds of flow cytometry training in preparation for the cell studies. Cell lines that were positive and negative for the receptor of interest had been obtained and readied for cell culture. A collaboration with Prof. Joel Mackay was also prepared for the assessment of the pull-down feasibility. However, none of this work could be completed during my PhD candidature.

These unforeseen circumstances ultimately lead to the inability to perform the expected experiments within the required time-period.

Acknowledgements

Firstly, I would like to express my heartfelt gratitude to my wonderful supervisor, Professor Elizabeth New. Liz, thank you for your incredible support and understanding during the several years I have been working in your group. I am grateful for the opportunity to have witnessed your growth as a professor, and a mother in the short time that I have been around. Your journey has been truly inspiring.

I would like to thank Peter Gill for taking me on as a student with absolutely zero background knowledge in computational chemistry. Thank you so much for your kind and patient approach, which has been invaluable in helping me to develop my skills in this field. I also want to thank Andrew Gilbert for answering countless emails, and tireless efforts in answering my numerous novice questions about my computational problems. Both of you have been incredibly helpful in providing guidance, and I am truly grateful for your support in helping me to become (kind of) competent in a field of work that I never believed I would even comprehend.

To Nick Proschogo, thank you for taking the time to help me with my mass spectrometry studies, and for always being willing to fit me in at the last minute whenever I needed help. Apart from work, I had a blast getting to chat and laugh with you while running and analysing my experiments.

I would like to thank all group members, past and present, who have been there for me during my PhD journey. Your amazing support has fostered a wonderful work environment and made my research experience so much more enjoyable. To my cohort buddies, Veronica and Nat, I am beyond grateful for the unwavering support you both have provided me over the course of many years. Nian Kee, I had a blast going around Brisbane with you, and I am so grateful for your kindhearted nature. Haobo, thank you for introducing me to computational chemistry and for your incredible talent for art. To the "younger" PhDs who managed to submit before me, Amy and Paris, your bright and lively personalities and openness to listening to my struggles have been immensely helpful. Kylie, I thank you for guiding me during my first years as a researcher and helping me gain the necessary skills for my work. Hazel, I am beyond grateful for your friendship and kindness, and I want to thank you for all the coffees, lunches, and de-stress sessions we've shared over the years.

Vinitha, thank you for your kindness and for sharing your crazy stories that have made me both laugh and cry. Linda and Marcus, your presence in the office at the end of my candidature was greatly missed, and I thank you for being such a lively presence. Louis, it has been so amazing seeing you grow from my second-year lab student to a researcher in our group! Ivy, thank you for your sweet nature and kindness. Postdocs Liam and Tom, thank you so much for your infinite wisdom in synthetic chemistry and helping me get through all my chemistry woes. Jia Hao, thank you for spending so many days assisting me with my biological studies. To Leila, thank you for your amazing dry humour that always had me cracking up. Everyone's contributions have been invaluable, and I am so grateful for each and every one of you.

To my parents, thank you for your endless love and support. Without you, there is absolutely no way I would have come this far in my studies. You have provided me with more than I could ever ask for, and I am forever grateful for everything you have done for me.

To my second parents, Maria and Kennedy, thank you for all your kindness and care and for all the times you treated me like your own daughter. To your lovely children, Manuela and Patrick, thank you both for all the support and laughter you have provided me with over the years we grew up together.

To Michael, thank you for always being such a great listener to my endless complaints. Your kindness and support have been a lifeline for me, and I cannot thank you enough for being there for me whenever I needed you.

To the rest of my friends, both new and old, I want to thank you all for your unwavering support, and for always checking in on me. Without your presence in my life, I would never have been able to make it to where I am today. I am truly grateful for the wonderful network of friends that I have, and I feel blessed to have you all in my life.

Contents

1	Introduction	1
1.1	Fluorescence overview	1
1.1.1	Introduction to fluorescence	1
1.2	Small molecule fluorophores	3
1.3	Common fluorophores used in biology	4
1.3.1	Fluorescein	5
1.3.2	Rhodamine	5
1.3.3	Naphthalimide	6
1.3.4	Coumarin	6
1.3.5	Cyanine	7
1.3.6	BODIPY	7
1.4	Biological techniques to study fluorescence	7
1.4.1	Microscopy	8
1.4.2	Flow cytometry	9
1.5	Proteins and peptides in biology	9
1.6	Clinical significance of studying peptides and proteins	11
1.6.1	Therapeutics	11
1.6.2	Targeting agents	12
1.6.3	Diagnostics	13
1.6.4	Theranostics	14
1.7	Conventional methods to study peptides and proteins	14
1.7.1	Enzyme-linked immunosorbent assay (ELISA)	14
1.7.2	Western blotting	15

1.7.3	Immunohistochemistry	16
1.7.4	Display technologies	16
1.7.5	Radiolabelling	18
1.7.6	Mass spectrometry	19
1.8	Advanced fluorescent methods to study peptides and proteins	20
1.8.1	Labelling proteins with small molecule sensors <i>in vitro</i>	20
1.8.2	Fluorescence in multimodal imaging	28
1.9	Aims	32
2	Multimodal probe for fluorescence and mass spectrometry imaging	33
2.1	Overview	33
2.1.1	Principles of mass spectrometry imaging	33
2.1.2	Mass tags	35
2.1.3	Multimodal MSI	36
2.2	Aims	38
2.3	Design of a fluorescent mass-tag	39
2.3.1	Criteria for mass-tags	39
2.3.2	Probe design	40
2.4	Synthesis	43
2.4.1	Synthesis of photo-cleavable linker	43
2.4.2	Synthesis of CyMT	45
2.5	Fluorescence studies	49
2.5.1	Solution-phase fluorescence properties	49
2.5.2	Solid-state fluorescence properties	50
2.6	MALDI compatibility studies	54
2.6.1	MALDI low-resolution studies	54
2.6.2	Matrix sublimation compatibility	57
2.7	Design and implementation of a dual imaging protocol	60
2.7.1	Dual imaging protocol considerations	60
2.7.2	Dual imaging protocol proof of concept	61
2.8	Planned tissue studies	64
2.9	Design and synthesis of NpMT	64

2.10	Conclusions	69
3	Computational analysis of fluorophores to study the redox state of proteins <i>in vitro</i>	71
3.1	Overview	71
3.1.1	Redox sensing on the sub-cellular level	71
3.1.2	Protein tagging	79
3.2	Aims	80
3.3	Previous work (Honours)	81
3.4	Current work	83
3.5	Determining solvation environment for calculations	84
3.5.1	HaloTag docking	85
3.6	FRET efficiency studies on FCR1	86
3.6.1	FRET efficiency calculations	89
3.6.2	Calculating the spectral overlap integral, J	90
3.6.3	Calculating the inter-chromophore distance, R	91
3.6.4	Calculating the orientation factor, κ^2	92
3.6.5	Calculating the Förster radius, R_0	96
3.6.6	Calculating FRET efficiency, E	97
3.6.7	Designing future FRET systems	97
3.7	Benchmarking studies on a novel naphthalimide-flavin derivative	98
3.7.1	NpFR1	99
3.7.2	Choosing benchmarking methods	100
3.7.3	Geometry optimisation of the ground state	103
3.7.4	Assessment of configuration interaction singles (CIS)	105
3.7.5	Assessment of configuration interaction with single and double CIS(D)	107
3.7.6	Assessment of time-dependent density functional theory	108
3.7.7	Assessment of the maximum overlap method	111
3.8	Conclusions	113
4	Fluorescent method to determine peptide-receptor binding <i>in vitro</i>	115
4.1	Overview	115

4.1.1	Cell receptors in health and disease	115
4.1.2	Peptides as targeting agents	116
4.1.3	Novel fluorescent methods	117
4.2	Aims	118
4.3	Design and synthesis	119
4.3.1	Envisioned method of receptor discovery	119
4.3.2	Probe design	121
4.3.3	Synthesis	122
4.4	Proof of concept peptide	127
4.4.1	Targeting peptide	127
4.4.2	Synthesis of fluorescent analogues of Octreotide	128
4.5	Synthesis of a control	132
4.5.1	HPLC challenges	134
4.6	Planned biological studies	138
4.7	Molecular dynamics	138
4.7.1	MD binding simulations	138
4.8	Conclusion	143
5	Conclusions	144
5.1	Development of a fluorescent sensor for multimodal mass spectrometry and fluorescence imaging	144
5.2	Investigation of novel redox sensors for protein tagging	145
5.3	Studies towards a method for peptide-receptor binding analysis	147
5.4	Outlook	148
6	Experimental methods	149
6.1	Synthetic methods	149
6.1.1	Instrumentation	149
6.1.2	Materials	150
6.2	Synthesis	150
6.2.1	Photo-cleavable linker	150
6.2.2	CyMT	155

6.2.3	NpMT	160
6.2.4	HaloTag linker	164
6.2.5	FlavinTag	166
6.2.6	NaphTag	168
6.2.7	BNp-COOH	170
6.2.8	Modified amino acids	174
6.2.9	Np-COOH	175
6.2.10	General peptide synthesis procedure	176
6.2.11	Peptides	179
6.3	Fluorescence studies	180
6.3.1	Solution-phase fluorescence	180
6.3.2	Solid-state fluorescence	180
6.4	Electrochemical studies	181
6.5	Insulin preparation	181
6.5.1	CyMT tagged insulin fibril formation	181
6.5.2	CyMT tagged insulin fibril slide preparation	182
6.6	Confocal microscopy	182
6.6.1	CyMT tagged insulin microscopy	182
6.6.2	HaloTag microscopy	182
6.7	MALDI studies	183
6.7.1	Low resolution MALDI-MS	183
6.7.2	MALDI-MSI sample preparation	183
6.7.3	Matrix sublimation	183
6.7.4	MALDI-MSI	183
6.8	Computational chemistry	184
6.8.1	Geometry optimisations	184
6.8.2	Spectral overlap integral	185
6.8.3	Transition dipole moments	185
6.8.4	Orientation factor calculation	186
6.8.5	Excited state calculations	186
6.9	Computational Biology	187

6.9.1	Autodock 4	187
6.9.2	Molecular docking	187

Abbreviations

4CCA	α -Cyano-4-hydroxycinnamic acid
μ	Micro
λ_{ex}	Excitation wavelength
λ_{em}	Emission wavelength
\AA	Angstroms
ACE2	Angiotensin-converting enzyme 2
AGT	O ⁶ -alkylguanine-dna alkyltransferase
AIE	Aggregation induced emission
APCI	Atmospheric-pressure chemical ionization
BCS	Bathocuproine sulfonate
BG	O ⁶ -benzylguanine
Boc	<i>tert</i> -Butoxycarbonyl
BODIPY	Boron dipyrromethene
BSA	Bovine serum albumin
CC2	Coupled-cluster singles and doubles
CD ₃ OD	Methanol-d ₄
CDCl ₃	Chloroform-d
CFP	Cyan fluorescent protein
CIS	Configuration interaction with single excitations
COSY	Correlation spectroscopy
COVID	Coronavirus disease
CPCM	Conductor-like polarizable continuum model
CTR1	Copper transporter 1
CuAAC	copper catalysed azide-alkyne Huisgen cycloaddition

DCC	<i>N,N</i> -dicyclohexylcarbodiimide
DESI	Desorption electrospray ionization
DET	Dexter energy transfer
DEVD	Caspase-3 substrate peptide
DFT	Density functional theory
DIIS	Direct inversion in the iterative subspace
DIPEA	<i>N,N</i> -Diisopropylethylamine
DLC	Delocalised lipophilic cation
DMEM	Dulbecco's modified eagle's medium
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DMSO-d ₆	Dimethyl sulfoxide-d ₆
DNA	Deoxyribonucleic acid
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDT	Ethanedithiol
ELISA	Enzyme-linked immunosorbent assay
EMA	European medicines agency
ER	Endoplasmic reticulum
ESI	Electrospray ionization
FBS	Fetal bovine serum
FDA	Food and drug administration
FRET	Förster resonance energy transfer
GFP	Green fluorescent protein
GSH	Glutathione
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HER	Human epidermal growth factor receptor
HER2	Human epidermal growth factor receptor 2
HF	Hartree Fock
HMBC	Heteronuclear multiple bond correlation
HOBt	Hydroxybenzotriazole
HOMO	Highest occupied molecular orbital

HPLC	High performance liquid chromatography
HRMS	High-resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence spectroscopy
ICT	Intramolecular charge transfer
IMOM	Initial maximum overlap method
IR	Infrared spectroscopy
ITO	Indium tin oxide
LCMS	Liquid chromatography mass spectrometry
LDI	Laser desorption ionization
LRMS	Low-resolution mass spectrometry
LRP	Ligand-based receptor capture
LUMO	Lowest unoccupied molecular orbital
MALDI	Matrix-assisted laser desorption/ionization
MD	Molecular dynamics
MHz	Megahertz
MM	Molecular mechanics
MMGBSA	Molecular mechanics with generalized born and surface area solvation model
MOM	Molecular orbital method
MRCI	Multi-reference configuration interaction
MRI	Magnetic resonance imaging
MS	Mass spectrometry
MSI	Mass spectrometry imaging
MUC2	Mucin 2
MUC4	Mucin 4
NADPH	Nicotinamide adenine dinucleotide phosphate
NAPA	NAnopost array
NMR	Nuclear magnetic resonance
NOX2	Nadph oxidase 2
NOX4	Nadph oxidase 4
ONB	O-nitrobenzaldehyde
PAI	Photoacoustic imaging

PBS	Phosphate-buffered saline
PCM	Polarisable continuum model
PDA	Photo-diode array
PDB	Protein data bank
PDBID	Protein data bank ID
PES	Potential energy surface
PET	Positron emission tomography
PMMA	Poly(methyl methacrylate)
POI	Protein of interest
PSMA	Prostate-specific membrane antigen
PyBOP	Benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
R0	Forster distance
RMSD	Root mean square deviation
RNA	Ribonucleic acid
ROS	Reactive oxygen species
S_n AR	Nucleophilic aromatic substitution
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SCF	Self-consistent field
SEM	Scanning electron microscopy
SIMS	Secondary ion mass spectrometry
SMLM	Single molecule localization microscopy
SN_2	Nucleophilic substitution reaction 2
SPECT	Single-photon emission computed tomography
SPPS	Solid-phase peptide synthesis
SSTR	Somatostatin receptor
SSTR2	Somatostatin receptor type 2
SSTR5	Somatostatin receptor type 5
STED	Stimulated emission depletion microscopy
TA30	30% acetonitrile and 70% water with 0.1% TFA
TBAB	Tetrabutylammonium bromide
TD-DFT	Time-dependent density functional theory

TEA	Triethylamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIPS	Triisopropylsilane
TOF	Time-of-flight
UV-Vis	Ultraviolet-visible
VEGFR	Vascular endothelial growth factor receptor
XFM	X-ray fluorescence microscopy
XRF	X-ray fluorescence spectroscopy
YFP	Yellow fluorescent protein

List of Figures

1.1	A Jablonski diagram highlighting mechanism by which light absorption and fluorescence occurs.	2
1.2	Jablonski diagram representing possible absorbance and emission events that may occur, and their correlation to the possible absorbance and emission spectra of a molecule. Colours are indicative of probable wavelengths.	2
1.3	Representation of the difference between intensity-based and ratiometric fluorescent sensors.	3
1.4	Chemical structure of the fluorescein core.	5
1.5	Chemical structure of the Rhodamine core.	5
1.6	Chemical structure of the 4-amino-1,8-naphthalimide core.	6
1.7	Chemical structure of the coumarin core indicating the 7-position.	6
1.8	Chemical structure of the basic cyanine core.	7
1.9	Chemical structure of the BODIPY core.	7
1.10	Diagrammatic representation of the various forms of ELISA.	15
1.11	Various display technologies including: A) Phage display, where a phage with displayed peptides is used; B) Ribosome display, where the mRNA-ribosome-peptide complex is used and; C) mRNA display, where the mRNA-puromycin-mRNA complex is used.	18
1.12	Direct immunofluorescence vs indirect immunofluorescence.	20
1.13	Chemical structures of the biarsenic ligand, FIAsh-EDT ₂	22
1.14	Four cysteines arranged on one side of the alpha helix such that the thiol pairs are appropriately spaced for favourable binding to the biarsenic FIAsh ligand. One must also consider the isometric configuration in which the arsenic binds in a manner perpendicular to the alpha helix.	22

1.15	Conventional "always on" fluorescence labels vs fluorogenic labels for no wash imaging.	23
1.16	Chemical structures of various tetracysteine biarsenic ligands for use with the FAsH system including: ReAsH-EDT ₂ , a resorufin based biarsenic; SplAsH-Boc, a non fluorescent biarsenic scaffold used to tether various fluorophores; and CaGF, a FAsH compatible Ca ²⁺ sensor.	25
1.17	Protein of interest (POI) fused to a tag protein reacts with an exogenous label.	25
1.18	Schematic representation of the A) SNAP-tag, B) CLIP-tag, and C) HaloTag mechanisms. The core substrate can be appended to a small fluorescent sensor, shown as an orange star.	27
2.1	Schematic representation of the mass spectrometry imaging process where mass spectra are obtained at multiple points on a sample. The chosen mass is typically visualised as a false-colour heat map that represent the signal intensity.	34
2.2	Schematic representation of the mass tag containing a biomolecule point of attachment (red), linker (blue) and reporter molecule (purple) that is cleaved upon irradiation with the ionisation source.	35
2.3	Mass-tag design based on an ONB linker (blue) with an N-hydroxysuccinimide biomolecular point of attachment (red), and an alkyne moiety allowing for facile conjugation of an azide-functionalised fluorescent reporter using "click" chemistry.	41
2.4	Cleavage mechanism of the ONB moiety upon irradiation.	42
2.5	Design of the hemicyanine-based fluorescent mass-tag, CyMT	43
2.6	Normalised excitation (dashed) and emission (solid) spectra of CyMT taken in water ($\lambda_{ex} = 558$ nm, $\lambda_{em} = 591$ nm).	50
2.7	Normalised absorbance (left) and emission (right) spectra of CyMT taken in the solid-state ($\lambda_{abs} = 535, 572$ nm, $\lambda_{em} \approx 620$ nm).	51
2.8	Normalised absorbance (left) and emission (right) spectra of less concentrated CyMT taken in the solid-state ($\lambda_{abs} = 535, 572$ nm, $\lambda_{em} = 624$ nm).	52
2.9	Structure of the hemicyanine used for comparison of solid-state fluorescence properties.	52

2.10	Normalised absorbance (left) and emission (right) spectra of hemicyanine chromophore taken in the solid-state ($\lambda_{abs} = 527, 566$ nm, $\lambda_{em} = 641$ nm). . .	53
2.11	Emission spectra of CyMT dispersed in PMMA at 10, 5, 2.5, 1.25, and 0.625% taken in the solid-state ($\lambda_{ex} = 530$ nm).	54
2.12	Low resolution MALDI-MS spectrum of CyMT dried onto polished steel in 4CCA matrix.	56
2.13	MALDI-MS spectrum of CyMT on polished steel with no matrix	57
2.14	Visual representation of the matrix deposition process.	58
2.15	Visual image of A) the indium tin oxide coated glass slide smeared with CyMT , and MALDI-MSI images of the slide with a B) 515.31 (pink), C) 893.49 (blue), and D) an overlay of the two m/z spectral filters.	59
2.16	Point mass spectrum taken at an arbitrary point of the MSI segment.	59
2.17	Representation of the envisioned dual fluorescence and MALDI imaging protocol. An ITO-slide mounted sample is labelled with the mass-tag followed by fluorescence imaging. After which the same sample is used in MSI.	60
2.18	Amino acid sequence of the two-chain protein, human insulin, highlighting the free primary amines (red) available for conjugation with a succinimidyl ester.	61
2.19	MALDI-MS of CyMT -tagged insulin showing untagged, singly-tagged, doubly-tagged protein and their respective photo-cleaved adducts.	62
2.20	Representative images of bright field (left) and fluorescence (right) microscopy of CyMT -tagged insulin fibrils. Images were taken without the use of medium using a 40x objective (λ_{ex} 561 nm laser, λ_{em} 575 - 675 nm).	63
2.21	Visual image of the selected MSI region around the droplet of dried CyMT -tagged insulin fibrils (left) and MALDI-MSI of this droplet with a 515.31 m/z filter displayed in false colour based on intensity.	64
2.22	Design of the fluorescent mass-tag based on a naphthalimide core with a water soluble peptide to increase the mass, solubility, and ionisation potential.	65
3.1	Chemical structure of Pep3-NP1	73
3.2	Chemical structure of BODIPY-Se	74
3.3	Chemical structure of NucPE1	75

3.4	Chemical structure of TPAs-SCH3-2CN and TAN	76
3.5	Chemical structure of Mem-NO	77
3.6	Chemical structure of H₂RB-C₁₈	77
3.7	Chemical structure of CCA	79
3.8	Chemical structure of NAB	79
3.9	a) Normalised fluorescence response of FlavinTag to changes redox environment in HEPES buffer (20 mM). Reduction achieved using sodium ditionite (100 eq, red) and oxidation with hydrogen peroxide (100 eq, purple). Measurements were taken at $\lambda_{ex} = 443$ and $\lambda_{em} = 552$ nm. b) Cyclic voltammogram of FlavinTag (0.5 mM) in acetonitrile and tetrabutylammonium chloride (100 mM).	82
3.10	Fluorescence intensities of Halo-CTR1 DLD1 cells dosed for 20 min with a) NaphTag (0.5 μ M), and b) FlavinTag (175 μ M) after treatment with with CuSO ₄ (30 μ M) or BCS (50 μ) for 24 h. Measurements were taken at $\lambda_{ex} = 405$, $\lambda_{em} = 500 - 600$ nm. Values are expressed as mean \pm SEM (n = 4-7), * = p < 0.05.	83
3.11	Proposed chemical structures of NpFR1-Tag and FCR1-Tag	84
3.12	FCR1-Tag and NpFR1-Tag docked within the Halotag enzyme as determined using Autodock4.	86
3.13	Jablonski diagram showing the process of FRET occurring as a non-radiative transfer of energy from the excited donor to the acceptor, resulting in emission from the acceptor alone, and return of the donor to the ground state.	87
3.14	Simulated example of spectral overlap between two fluorophores. Donor absorbance and emission, and acceptor absorbance and emission are shown in purple, blue, green and orange respectively.	87
3.15	Chemical structure of FCR1 in the oxidised and reduced forms highlighting the effect of FRET between the donor, coumarin, and the acceptor, flavin, and the effect on its fluorescence output.	88
3.16	Normalised absorbance (dashed) and emission (solid) spectra of N-ethyl flavin (green) and 7-diethylamino coumarin (blue), respectively. Spectra were taken in MilliQ water. Red highlight indicates region of overlap.	90

3.17	Structure simplification for FCR1 calculations.	91
3.18	Inter-chromophore distance from optimised geometry of FCR1	92
3.19	Angles between the donor excited transition dipole moment (blue arrow), the acceptor transition dipole moment (green arrow) with respect to the axis of the inter-fluorophore distance (R)	93
3.20	Transition dipole moments (blue arrow) of the isolated coumarin- and flavin-based chromophores independently (separated at 1 carbon beyond the last atom involved in the chromophores' conjugation) and when bound by the cyclohexane linker as calculated using M11/6-31+G* level of theory.	95
3.21	Distance dependence of FRET efficiency, E.	96
3.22	Structure of potential FCR1B using a triazole linker.	97
3.23	Optimised structure of FCR1 with triazole linker.	98
3.24	Chemical structures of the naphthalimide and flavin cores, and the novel derivative NpFR1	99
3.25	Normalised absorbance spectrum of NpFR1	100
3.26	Optimised geometry of NpFR1 in its oxidised form calculated using B3LYP/6-31G* level of theory showing the front and side views.	104
3.27	HOMO and LUMO orbitals of NpFR1 calculated at B3LYP/6-31G* level of theory presented at ± 0.02 Å isovalues.	105
4.1	Varied receptor expression in healthy vs diseased cells	116
4.2	Size comparison of: Herceptin, an antibody-based drug; Octreotide, a peptide-based drug; and Ibuprofen, a small molecule drug (not shown to scale).	117
4.3	Design of the dual-functional fluorescent probe with the ability to be incorporated into the peptide of interest (blue chain) as well as containing a pull-down partner (red arrow).	119
4.4	After incubating with a cell line of interest, binding of the peptide of interest to a cell line expressing a relevant receptor can be identified by the presence of fluorescence.	120

4.5	After identification of binding by fluorescence, the cell contents can be incubated with pull-down-partner-decorated magnetic solid beads. The bound peptide can be isolated using external magnets, after which unbound contents can be washed away to yield the isolated peptide alone (not shown) or the peptide-receptor complex.	121
4.6	Chemical structure of the designed probe.	122
4.7	Normalised excitation (dashed) and emission (solid) spectra of BNp-COOH in 20 uM HEPES buffer ($\lambda_{ex} = 455$ nm, $\lambda_{em} = 547$ nm).	127
4.8	Chemical structure of octreotide.	128
4.9	Normalised excitation (dashed) and emission (solid) spectra of BNp-COOH -tagged octreotide in 20 uM HEPES buffer ($\lambda_{ex} = 452$ nm, $\lambda_{em} = 545$ nm).	132
4.10	Normalised excitation (dashed) and emission (solid) spectra of Np-COOH in 20 uM HEPES buffer ($\lambda_{ex} = 451$ nm, $\lambda_{em} = 548$ nm).	133
4.11	Normalised excitation (dashed) and emission (solid) spectra of NpCOOH labelled octreotide in HEPES buffer (20 mM).	134
4.12	HPLC trace of Np-COOH -conjugated octreotide assessing the same sample after A) 0 h in 20 - 60% B) 2 h in 30 - 60%, and C) 24 h at 20 - 100% acetonitrile in water with 0.1% TFA over 60 min.	135
4.13	HPLC trace of cyclised Np-COOH -conjugated octreotide assessing the same sample after A) 0 h in 0 - 100% B) 6 h in 0 - 100%, C) 18 h in 30 - 70%, and D) 24 h in 0 - 100% acetonitrile in water with 0.1% TFA over 60 min.	137
4.14	Changes in the RMSD of octreotide and the SSTR2 protein over time during MD binding simulations.	139
4.15	Binding interactions of octreotide with the various residues within the binding pocket of SSTR2 as determined by the MD simulations.	140
4.16	Changes in the RMSD of BNp-COOH conjugated octreotide and the SSTR2 protein over time during MD binding simulations.	141
4.17	Binding conformation of BNp-COOH conjugated octreotide within the SSTR2 binding pocket at 5 and 50 ns.	142

4.18	Binding interactions of BNp-COOH when bound to octreotide (not shown) with the various residues of the SSTR2 protein as determined by the MD simulations.	142
5.1	Suggested chemical structure of biotinylated fluorescent amino acid Fmoc-BNp-OH.	148
S1	Absorbance spectrum of <i>N</i> -ethyl flavin (0.100 mM) in water.	228
S2	Fitted Gaussians for each NpFR1 absorbance peak.	228
S3	HPLC trace of BNp-COOH -labelled octreotide in 0 - 100% acetonitrile in water with 0.1% TFA in water over 60 mins after using standard cleavage solution.	229
S4	HPLC trace of BNp-COOH -labelled octreotide in 0 - 100% acetonitrile in water with 0.1% TFA in water over 30 mins after using reagent K.	229
S5	HPLC trace of Np-COOH -labelled octreotide dissolved in DMSO after sitting for 12 h. Run in 15 - 100% acetonitrile in water with 0.1% TFA in water over 60 mins using C8.	230
S6	¹ HNMR spectrum of CyMT	231
S7	¹ HNMR spectrum of 21	232
S8	¹ HNMR spectrum of FlavinTag	233
S9	¹ HNMR spectrum of NaphTag	234
S10	¹ HNMR spectrum of BNp-COOH	235
S11	¹ HNMR spectrum of Np-COOH	236

List of Schemes

2.1	Attempted synthesis of the photo-cleavable linker.	44
2.2	Synthesis of the final photo-cleavable linker, 7	45
2.3	Attempted synthesis of an azide functionalised hemicyanine using a propyl linker.	46
2.4	Attempted synthesis of an azide functionalised hemicyanine using a hexane linker.	47
2.5	Ester hydrolysis of the provided hemicyanine.	47
2.6	Coupling of the amino azide to the acid functionalised hemicyanine.	48
2.7	Synthesis of the final fluorescent mass-tag, CyMT	49
2.8	Expected m/z of CyMT and its photo-cleaved adduct.	55
2.9	Synthetic scheme of protecting acid of existing naphthalimide	66
2.10	Synthetic scheme of protecting before substituting	66
2.11	Synthetic scheme of using amino azide directly on naphthalimide	67
2.12	Synthetic scheme of forming azide <i>via</i> mesylate intermediate	67
2.13	solid phase peptide synthesis of NpMT	69
3.1	Synthesis of the HaloTag substrate.	81
3.2	Synthesis of FlavinTag	81
3.3	Synthesis of NaphTag	82
4.1	Succinimidyl activation of biotin.	123
4.2	Attempted synthesis of the probe with the biotin at the imide position. . . .	123
4.3	Attempted S_NAR of the 4-position with ethylenediamine.	124
4.4	Synthesis of BNp-COOH	126

4.5	Attempted synthesis of acetal bound threonine for conjugation to rink amide resin.	129
4.6	Reduction of Fmoc-Thr-OH to Fmoc-Thr-ol.	130
4.7	SPPS of fluorescently labelled octreotide.	131
4.8	Synthesis of control probe, Np-COOH	133

List of Tables

2.1	Reaction conditions trialled for the synthesis of 5	45
3.1	Vertical excitation energies of various singlet excited states of coumarin. . . .	94
3.2	Transition dipole moment vectors of flavin and coumarin calculated with M11/6-31+G* level of theory	94
3.3	Calculated angles between the transition dipole moments of the specified angles.	96
3.4	Absorbance maxima of NpFR1 as determined by Gaussian fitting.	100
3.5	Vertical excitation parameters of NpFR1 calculated using CIS with 6-31G* or 6-31+G* basis sets in either the gas state, or solvated in water.	106
3.6	Vertical excitation parameters of NpFR1 calculated using CIS(D) with 6- 31+G* basis sets in either the gas state, or solvated in water.	108
3.7	Vertical excitation parameters of NpFR1 calculated using TD-DFT with the specified method and the 6-31+G* basis sets, solvated in water.	109
3.8	TD-DFT calculations using the B3LYP method and the specified triple zeta basis sets.	110
3.9	MOM with significant strengths in order of energy using B3LYP 6-31+G* in CPCM	112
3.10	Overlap of significant transitions calculated using MOM and the SG-3 grid. IMOM calculations yielded the same results.	113
4.1	Reaction conditions trialled for the synthesis of 30	125
S1	Overlap of significant transitions calculated using MOM and the SG-0 grid. .	229

Chapter 1

Introduction

1.1 Fluorescence overview

1.1.1 Introduction to fluorescence

Fluorescence is defined as the ability of a substance to emit visible light following absorption of light or other electromagnetic radiation. More specifically, fluorescence occurs when an electron in an excited state (S_{0+n}) relaxes back into the ground state (S_0), resulting in the emission of a photon. For each specific molecule, excitation of an electron into the excited state is quantised. That is, only a photon with energy equal to the energy difference between the ground and excited state is able to promote the electron to the excited state.

Within any electronic energy state (S_n) there are multiple vibrational states to which the electron may be excited. According to the Franck-Condon principle, electronic transitions back to the ground state can only occur from the lowest vibrational level of the excited state. As such, an electron will undergo non-radiative decay to reach the lowest vibrational level of the excited state before returning to the ground state. For this reason, emission always occurs at a lower energy and thus, a longer wavelength than the excitation wavelength.

Jablonski diagrams are often used to illustrate the various molecular processes that can occur within a molecule during excitation and emission. Vertical arrows depict electronic transitions resulting in light absorption or emission. Energy spacing between vibrational levels are represented by horizontal lines. Vertical wavy arrows depict non-radiative processes that occur when an electron relaxes to a lower energy state without the emission of light (Figure 1.1).

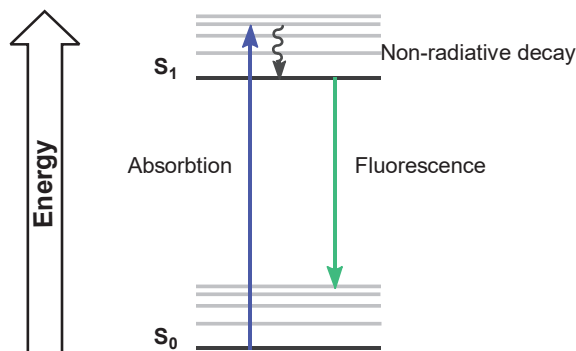


Figure 1.1: A Jablonski diagram highlighting mechanism by which light absorption and fluorescence occurs.

Excitation can occur to a variety of vibrational levels within an excited state. Similarly, emission can also occur to a variety of vibrational levels. This results in a bandwidth of possible absorbance and emission wavelengths. This phenomenon gives rise to the characteristic absorbance and emission spectra observed in fluorescence spectroscopy (Figure 1.2).¹ The difference in wavelength between the absorbance and emission maxima is known as the Stokes shift.

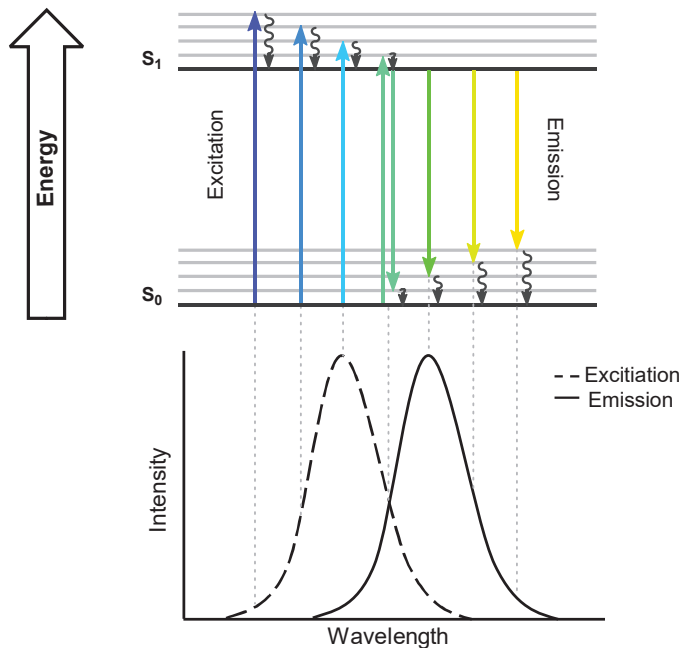


Figure 1.2: Jablonski diagram representing possible absorbance and emission events that may occur, and their correlation to the possible absorbance and emission spectra of a molecule. Colours are indicative of probable wavelengths.

1.2 Small molecule fluorophores

A small molecule fluorophore is a chemical capable of absorbing and re-emitting light. These molecules typically contain several aromatic groups, or planar structures with many π -bonds. These small molecules have become a staple in the field of biological studies. Because fluorescence is highly sensitive in nature, it lends itself to visualising biological systems making it suitable for detecting at very low concentrations. This makes it a very attractive tool for studying the microscopic phenomena of cellular biology.

In their most simple application, small molecule fluorophores can be used as stable dyes or stains.² But the small molecules can be modified in ways that allow them to respond to their environment. Such molecules are known as fluorescent sensors.

Fluorescent sensors can change in either colour or intensity in the presence of an analyte. In general, fluorescent sensors can be categorised as either intensity-based or ratiometric. Intensity-based fluorescent sensors can either increase (turn-on) or decrease (turn-off) their fluorescence emission, whereas ratiometric sensors exhibit a shift in their emission wavelength. Ratiometric sensors are more desirable as they are self-calibrating and are unaffected by instrumental parameters, however, they are more challenging to develop.

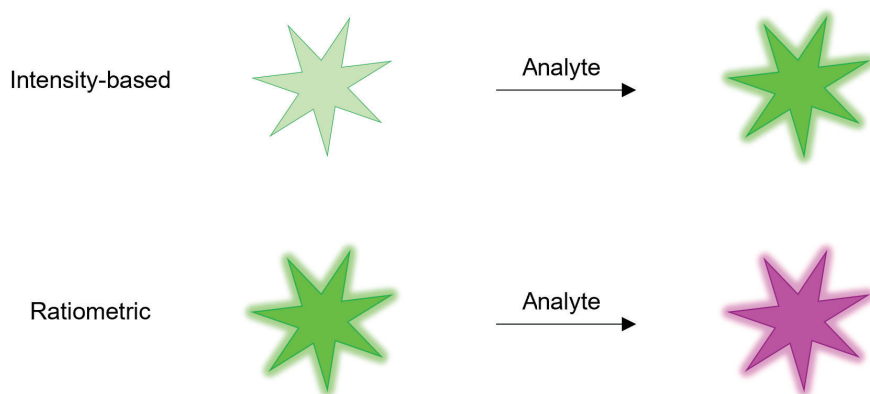


Figure 1.3: Representation of the difference between intensity-based and ratiometric fluorescent sensors.

Additionally, fluorescent sensors can be further categorised as being reversible or irreversible. Irreversible sensors interact with the analyte of interest by forming or breaking covalent bonds that cannot be restored. On the other hand, reversible sensors react to the

analyte of interest in a manner that allows it to be restored to its initial state. Given the transient nature of cellular processes, reversible sensors are advantageous as they can be used to monitor dynamic changes in the environment, which is not possible with irreversible sensors.

1.3 Common fluorophores used in biology

In order for a fluorophore to be compatible with live cells in biological studies, it must possess certain key characteristics. Typically, ideal characteristics include:

Red shifted excitation and emission maxima - This is because the shorter wavelengths are higher in energy and are known to cause considerable damage to DNA and proteins. In addition, shorter wavelengths may also induce auto-fluorescence of naturally occurring endogenous fluorophores.³

Sufficient brightness - This ensures that the fluorescence output is maximised such that the signal to noise ratio is sufficiently satisfactory to distinguish the desired signal from the background.⁴

Photostability - This means that the fluorophore is not prone to photo-degradation under standard imaging conditions, allowing for long-term experiments without fluorescence bleaching to occur.⁵

Biological compatibility - This includes chemical stability within the diverse cellular environment and high water solubility for use in aqueous media.⁶ Additionally, the fluorophore should also have low toxicity.

Synthetic compatibility - This enhances the versatility of the fluorescent molecule, allowing for tuning of its properties and the addition of various groups, such as targeting moieties.

The criteria for the most commonly used fluorescence molecules in biological applications vary based on the intended purpose. An overview of the commonly used fluorophores is presented below.

1.3.1 Fluorescein

Fluoresceins (Figure 1.4) were among the first fluorescent molecules to be synthesised.⁷ They have several properties that make them ideal for biological studies including excellent brightness, and good water solubility. Fluoresceins are extremely versatile dyes that have been used extensively in the field of biology. Some of these uses include pH sensing, copper sensing, and even fluorescent tumour tracer in brain tumour surgery.⁸⁻¹⁰ They are one of the most common fluorophores used in fluorescent-antibody conjugates and have seen vast use in techniques like flow cytometry and ELISA.^{11,12} However, fluoresceins are prone to photo-degradation, limiting their application in long-term experiments.¹³

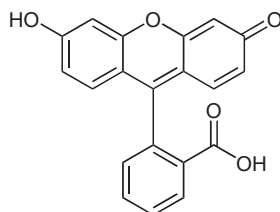


Figure 1.4: Chemical structure of the fluorescein core.

1.3.2 Rhodamine

Rhodamines are structurally related to fluoresceins (Figure 1.5), containing amino groups rather than the phenolic oxygens. They also exhibit excellent brightness, and good water solubility, but not to the same extent as fluoresceins. They can be readily modified synthetically for various applications including ion sensing, super resolution fluorescence imaging, and even wildlife tracing studies.¹⁴⁻¹⁷ Rhodamines are also less prone to photo-degradation when compared to fluorescein.¹⁸

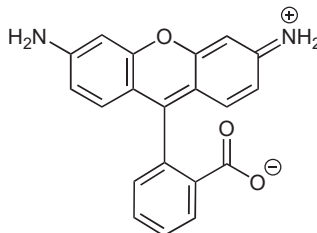


Figure 1.5: Chemical structure of the Rhodamine core.

1.3.3 Naphthalimide

1,8-Naphthalimide fluorophores are structurally comprised of a naphthalene ring fused to a dicarboxyl imide (Figure 1.6). Naphthalimides are generally substituted with an electron-donating group at the 4-position which enhances their fluorescence output. The mechanism of naphthalimide fluorescence is based on the ICT mechanism. They can be readily modified in at least 2 positions, and have highly tunable spectral properties depending on core substituents.¹⁹ Naphthalimides have found use in various applications including tumour hypoxia imaging, and gastrotransmitter sensing.^{20,21} However, they tend to suffer from poor water solubility, although this can be modulated through some structure modifications.²²

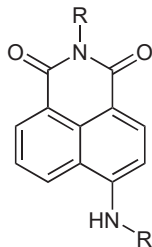


Figure 1.6: Chemical structure of the 4-amino-1,8-naphthalimide core.

1.3.4 Coumarin

Coumarins are based on the naturally occurring non-fluorescent parent compound (Figure 1.7). Substitution at the 6- or 7-position with an electron donating moiety such as an amine or hydroxyl group generates a highly fluorescent compound that possesses good photostability, and water solubility. They have found use as ion indicators, photoactivatable imaging agents, and fluorescent enzyme substrates.^{23–25} In general, they have higher energy excitation and emission profiles than other scaffolds although this can be modulated through synthetic modifications.²⁶

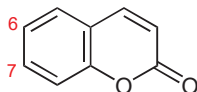


Figure 1.7: Chemical structure of the coumarin core indicating the 7-position.

1.3.5 Cyanine

Cyanines are a class of dyes that describe a family of molecules with an odd-numbered polymethine chain that links two one or more ring structures that generally contain tertiary amines (Figure 1.8). They are ideal for use in biological studies due to their orange-near IR emission profiles. Cyanines have been used in applications including drug delivery, metal sensing and live animal imaging.^{27–29} However, cyanines can have small Stokes shifts, which can cause self-quenching.

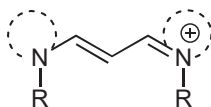


Figure 1.8: Chemical structure of the basic cyanine core.

1.3.6 BODIPY

BODIPY fluorophores are based on a boron difluoride and dipyrromethene core (Figure 1.9). They have excellent photostability, and are not sensitive to changes in solvent or pH environments. BODIPY derivatives have found use in several biological applications including commercial lysosome trackers and photo-cages.^{30,31} They are easy to functionalise but, suffer from small Stokes shifts making them susceptible to self-quenching.

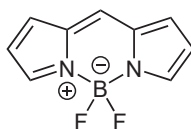


Figure 1.9: Chemical structure of the BODIPY core.

1.4 Biological techniques to study fluorescence

Fluorescence has become perhaps one of the most widely utilised technique for studying biological systems. Fluorophores and fluorescent sensors are highly tunable and can be adapted for use with many different techniques. Below are a few of the methods that utilise fluorescence as a means of studying biological systems.

1.4.1 Microscopy

Fluorescence microscopy is a technique used to visualise the distribution and localisation of fluorescently labelled cells or tissue. The most basic type of fluorescence microscopy involves the use of a widefield microscope. The sample is irradiated by a single light source which excites the fluorophore in a given sample. The emitted fluorescence passes through a wavelength specific filter, which separates the emitted light from the incident light that passes through to the detector. The advantage of widefield microscopy is the fast imaging, and low maintenance cost of the microscope itself. However, it suffers greatly with respect to image quality since it is prone to high background noise. In addition, it has low spatial resolution, which does not allow it to image cellular microstructures, such as organelles, in good detail.³²

Confocal microscopy has become standard in the field of fluorescence imaging. Confocal microscopes operate by emitting a laser which passes through a pinhole and is reflected towards the sample by a dichroic mirror. The emitted fluorescence is passed through a pinhole that blocks out of focus light, while the in-plane light passes through a photomultiplier tube before reaching the detector. Confocal microscopy offers significant advantages over widefield microscopy including improved contrast and higher spatial resolution. Confocal microscopes also offer greater flexibility in terms of excitation and emission parameters due to the large array of available lasers and filters.³³

Fluorescence lifetime imaging microscopy is a technique that creates an image based on the lifetime of a fluorophore: that is, the time that a molecule spends in the excited state before emitting a photon and returning to the ground state. The sample is irradiated with repeated laser pulses, which cause the fluorophore to emit a photon. Rather than intensity, the image is created based on the measurement of time taken between laser pulse and photon emission. One major disadvantage of this technique is that fluorescent molecules are prone to photobleaching caused by the repeated laser irradiation required for this technique. This limits its application in longer term experiments.³⁴

Super-resolution microscopy is a technique that produces high-resolution images of biological samples by overcoming the diffraction limit of conventional microscopy. There are several different techniques used to generate super-resolution images but the main forms include: structured illumination microscopy (SIM), where the sample is irradiated by a

patterned light source that is sampled multiple times to generate interference patterns that are deconvoluted into super-resolution images; stimulated emission depletion microscopy (STED), where fluorophores are selectively deactivated to minimise the focal point with the use of a depletion laser, in turn lowering the diffraction limit and increasing image resolution; and single-molecule localisation microscopy (SMLM), where only a small percentage of fluorophores is excited at one time, reducing the spatial overlap which increases image resolution.³⁵

1.4.2 Flow cytometry

Flow cytometry is a technique used to study populations of single cells on both a physical and chemical level. Cells are suspended in a buffered solution and are focussed into a single-cell stream that pass through a laser where the light gets scattered by the passing cell. The scattering is measured by multiple detectors, allowing the instrument to gather information about the size, granularity, fluorescence, and other properties of the cells. Flow cytometers are capable of analysing thousands of cells per second, making them a powerful tool for rapidly and accurately assessing large cell populations in real-time. By labelling cells with fluorescence-conjugated antibodies or other fluorophores, flow cytometers can detect and sort cells based on their unique fluorescence patterns, making it a useful tool for cell separation and analysis.³⁶

1.5 Proteins and peptides in biology

Peptides and proteins are biomolecules that play a pivotal role in cellular function. Structurally, both are made up of a series of amino acids connected by amide bonds. Traditionally, peptides are defined as consisting of between 2 and 50 amino acids, while proteins are defined as consisting of greater than 50 amino acids. Peptides and proteins are responsible for a diverse range of biological functions, and their properties have been exploited in the field of medicine. The main functions of peptides and proteins in biology are outlined below.

Catalysis

Proteins that catalyse reactions are known as enzymes. They are capable of performing both anabolic and catabolic reactions, which involve breaking down or building up molecules,

respectively. Each enzyme is highly specific for its reaction, meaning that it can only catalyse one specific set of reactions.³⁷

Structure

Structural proteins are one of the most abundant types of proteins found in nature, responsible for shaping structures in biology. Cell-based structural proteins, such as actin and microtubules, are crucial in forming the physical structure of cells and play a major role in organising organelle distribution and cell movement in processes such as cell division.³⁸

Transport

Transport proteins are involved in the movement of products in and out of the cell. Generally these proteins are known as channels or pumps. Perhaps the most well known of these transport proteins is the sodium/potassium pump that lies within the cell membrane and facilitates movement of sodium and potassium in and out of the cell to maintain osmotic equilibrium.³⁹

Other transport proteins play a role in the capture and movement of specific biomolecules. A well-known example of an important transport protein is haemoglobin. This protein is crucial for the transportation and delivery of oxygen and retrieval of CO₂ back to the lungs.⁴⁰

Defence

Defence peptide and proteins play a critical role in the body's defence against foreign invaders, including viruses, bacteria, and other pathogens. For example, antimicrobial peptides are a category of short peptides found in a variety of organisms, including humans. They serve as a component of the innate immune system and help to protect against pathogenic microorganisms. These peptides are produced by immune cells like neutrophils and macrophages, and they function by disrupting the membranes of pathogens, resulting in their death.⁴¹ These properties have led to the exploration of antimicrobial peptides as potential therapeutic agents for the treatment of various diseases including cystic fibrosis.^{42,43}

Antibodies are another class of protein that is involved in biological immunity. They help the body to identify and neutralise pathogens. Antibodies specifically recognise and bind

to pathogens, marking them for destruction by other immune cells. The specific structure of each antibody allows it to recognise and bind to a unique antigen, giving the immune system the ability to respond to a wide range of pathogens.⁴⁴

Signalling

Intracellular signalling peptides and proteins play a critical role in transmitting signals information between cells, tissues, and organs. An interplay between peptides, proteins and other small molecules responsible for driving such signal processes. They work by binding to specific receptors on the cell membrane or within cells, which initiates a series of intracellular signalling events.

Proteins and peptides are among the most common signalling molecules and include hormones, neurotransmitters, cytokines, and signalling proteins.⁴⁵ Hormones such as insulin and growth hormone regulate processes such as metabolism and growth.^{46,47} Neurotransmitters like dopamine and serotonin regulate communication between nerve cells.⁴⁸ Cytokines, such as interleukins, regulate communication between cells of the immune system.^{49,50} Signalling proteins, such as protein kinases, regulate a variety of cellular processes, including cell division and apoptosis.⁵¹

These signalling molecules bind to specific receptors to initiate a series of cascades within the cell. These receptors are specialised proteins that are capable of detecting, and responding to a specific signal. They are highly specific and will only bind with ligands of a particular structure. Receptor proteins can be classified into two main categories: transmembrane and intracellular receptors.⁵² The interplay between signalling molecules and receptors is essential for many biological processes, including growth, development, and response to changes in the environment.

1.6 Clinical significance of studying peptides and proteins

1.6.1 Therapeutics

The significance of understanding the function of proteins in biology was recently highlighted during the COVID-19 pandemic when it was discovered that the SARS-CoV-2 coronavirus enters cells through the angiotensin-converting enzyme 2 (ACE2). In the human body, ACE2

is an enzyme that cleaves angiotensin II into angiotensin (1-7), angiotensin I to angiotensin (1-9), and participates in the hydrolysis of other peptides.⁵³ It plays a crucial role in local and systemic haemodynamics as it primarily lowers blood pressure. The role of ACE2 in the uptake of SARS-CoV-2 into cells was discovered through the use of fluorescent antibody staining.⁵⁴

The discovery of this mechanism has led to the possibility of treatment options for those suffering from severe COVID-19 infection. For example, administering intravenous infusions of soluble recombinant ACE2 was found to reduce the severity of COVID infection. This is because the soluble ACE2 acts as a competing target, reducing the infection of host tissue by limiting the availability of the virus to bind to membrane-bound ACE2.⁵⁵ It was also found that the potential to neutralise ACE2 by targeting protein enzymes to block the reactive site and prevent COVID-19 entry.⁵⁶

Targeted therapeutics can act as direct modulators of the target receptor to either enhance or inhibit their function. For example, Herceptin is an approved, targeted therapeutic protein for the treatment of human epidermal growth factor receptor 2 (HER2) receptor positive breast cancer. HER2 is a transmembrane protein that facilitates cell growth, survival, proliferation and differentiation. Overexpression of the HER2 gene is seen in approximately 30% of breast cancers and is a significant indicator of low survival and high relapse rates in patients.^{57,58} Hereceptin is a humanised, monoclonal antibody with a high affinity for the HER2 receptor. The mechanism of action involves degradation of the HER2 receptor which in turn, inhibits proliferation and induces apoptosis.⁵⁹ Since HER2 is also involved in normal cell function, side effects are common. Most strikingly, a high incidence of cardiac toxicity has been noted.⁶⁰ For many patients, this side effect is partly reversible, but many patients discontinue treatment before completion.⁶¹ Many drugs, especially those used in chemotherapy, have a high incidence of systemic side effects.⁶²⁻⁶⁶

1.6.2 Targeting agents

Recent efforts to minimise off-target drug effects have focussed on the use of receptor targets to direct therapeutic drug delivery to a specified location. Actively targeted drug-delivery involves conjugating receptor-specific ligands to a drug, or drug delivery system.^{67,68} In turn, the drug is accumulated only when bound to the specific receptor. As such, drug

concentration is high only in the target tissue. The significance of targeted drug delivery lies in its ability to minimise systemic toxicity of drugs that act indiscriminately. For example, docetaxel is a chemotherapeutic drug used to treat prostate cancer, among others, that targets the G(2)M phase of the cell cycle, ultimately leading to apoptosis.⁶⁹ Its effect on the cell cycle affects all dividing cells indiscriminately, thus docetaxel is associated with many adverse side effects, including a 97% occurrence of neutropenia, and even death.⁷⁰ Research has been made to minimise these adverse events by using nano-encapsulated drug delivery systems. For example, BIND-014 is a docetaxel-carrying nanoparticle decorated with peptide ligands for the prostate specific membrane antigen (PSMA) - a receptor that is over expressed on prostate cancer cells.⁷¹ The targeting ligand utilised is a pseudomimetic dipeptide that was found to have an affinity and specificity similar to that of antibodies.⁷² BIND-014 was tested for use in metastatic castration-resistant prostate cancer and reached phase II clinical trials.⁷³

1.6.3 Diagnostics

Another way in which the interplay of peptides and proteins has been exploited is in the development of targeted diagnostic agents. For many progressive diseases, diagnostic imaging techniques are rarely able to detect early stage abnormalities due to a lack of sensitivity to small lesions. In such diseases, early detection is crucial in increasing survival rates. For example, pancreatic cancer has a high mortality rate of up to 91% within 5 years.⁷⁴ Currently, the only potential for long-term survival is complete tumour resection.⁷⁵ However, due to the highly metastatic and invasive nature of pancreatic cancer, the average size of a resectable tumour is <30 mm.⁷⁶ Only 10-15% of patients are diagnosed in early stages of disease, with the majority not being diagnosed until after significant metastasis, leading to a poor prognosis.^{77,78} As such, early detection is of great clinical importance. Many attempts have been made at designing targeted diagnostic agents for pancreatic cancer using strategies such as protein-based targeting to specific disease receptors.⁷⁹ For example, MUC4 receptors have been found to be overexpressed in pancreatic cancer.⁸⁰ One group conjugated anti-MUC4 antibody proteins onto MRI active nanoparticles in order to direct the imaging agent to pancreatic tumours.⁸¹ The preliminary studies *in vitro* and *in vivo* demonstrated a good sensitivity and specificity for pancreatic tumours.⁸¹

1.6.4 Theranostics

Theranostics are an emerging class of targeted medicines in which diagnostic imaging and therapeutic agents are integrated into a single system. The goal of a theranostic agent is to image disease and track drug delivery and pharmacokinetics in real time.⁸² This allows the evaluation of disease progression and efficacy of the therapeutic agent, eliminating the "trial and error" approach of disease treatment, and facilitating progress into personalised medicine. While theranostics are a new field in medicine, a few have been approved for use in clinical practice. ^{177}Lu -DOTATATE is a peptide-based radionuclide theranostic which was approved for the treatment of gastroenteropancreatic neuroendocrine tumours by the European Medicines Agency (EMA) and, the Food and Drug Administration (FDA) in September 2017 and January 2018 respectively.^{83,84} It is comprised of a modified form of the commercially available targeting peptide, octreotide, which is known to target somatostatin receptors. These receptors are expressed in 80-100% of neuroendocrine tumours and are generally the main target of therapy.⁸⁵ The radionuclide, ^{177}Lu , is a medium-energy β emitter making it an effective therapeutic. In addition, ^{177}Lu emits accompanying low-energy γ rays which allows the drug to be imaged through single-photon emission computed tomography (SPECT) imaging. In Phase III trials, ^{177}Lu -DOTATATE outperformed octreotide therapy in terms of tumour progression and quality of life benefits.⁸⁶

1.7 Conventional methods to study peptides and proteins

The study of proteins and their functions in the body is crucial for advancing our understanding of biological systems, and for developing new therapies and treatments for the betterment of human health. Conventional methods of studying peptides and proteins are detailed below.

1.7.1 Enzyme-linked immunosorbent assay (ELISA)

Enzyme-linked immunosorbent assay (ELISA) is a highly sensitive assay technique designed to detect and quantify biomolecules such as peptides, proteins, and antibodies that was first described in 1971.⁸⁷ In general, the assay works by immobilising the biomolecule of interest onto a solid surface which is then complexed with the corresponding antibody. This can

either occur directly onto the assay plate or *via* antibody capture (known as sandwich ELISA). Detection of the bound biomolecule is then detected directly, by labelling with a primary antibody, or indirectly by labelling the primary antibody with a secondary antibody (Figure 1.10). Detection can be analysed by either colourimetric, fluorescence or infrared output.

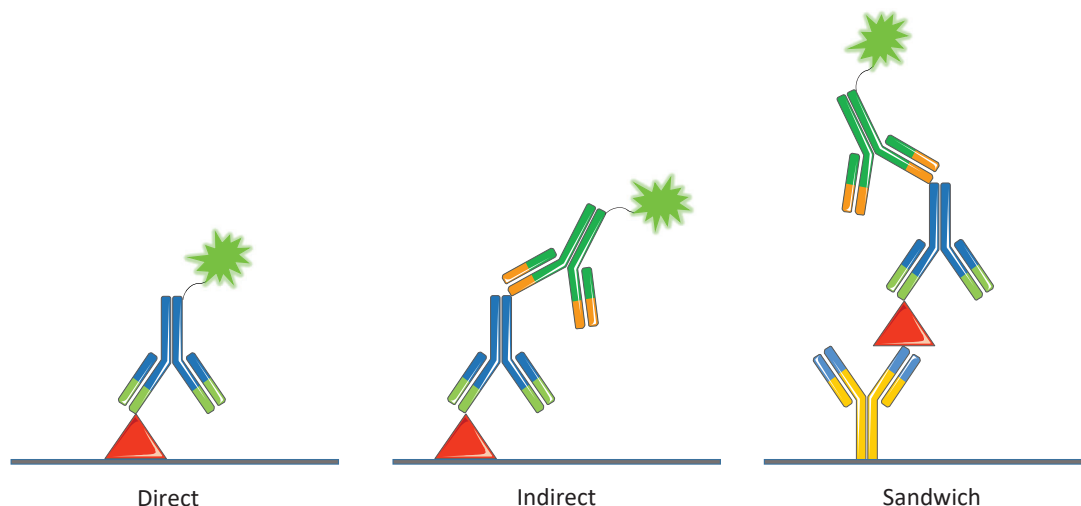


Figure 1.10: Diagrammatic representation of the various forms of ELISA.

ELISA is used extensively in the clinical field for the identification of various pathologies such as viruses, bacteria, and is also used in blood typing. Despite its widespread use, the ELISA technique suffers from several limitations, particularly in the research and development field. For example, biomolecule identification is limited by the commercial availability of the specific primary antibody. Additionally, there are many steps involved in this assay as well as the requirement for several reagents, making it a rather inefficient process. One of the major disadvantages of the ELISA technique is that the cells need to be lysed or fixed in order to quantify the biomolecule of interest, making it unsuitable for visualising live cellular processes.

1.7.2 Western blotting

Western blotting, or immunoblotting, is a technique used to detect the presence and quantity of a specific protein in a sample. It is applicable to various types of samples including cell lysates, cell homogenates, and body fluids. It involves the separation of a protein mixture

by electrophoresis onto a cell membrane. The protein of interest is identified with the use of fluorescent antibodies that allow visualisation by binding to the protein of interest, in a similar manner to that described for the ELISA technique. Detection of the antibody can be performed using fluorescence, colourimetric or IR.⁸⁸ Clinically, western blotting is a prominent tool used in antigen detection and diagnosis of infectious diseases.⁸⁹ Despite its widespread use, western blotting is time-consuming technique that requires specialist equipment and reagents, with multi-step procedures to complete.

1.7.3 Immunohistochemistry

Immunohistochemistry is a technique used to detect antigen proteins within a tissue sample. For example, in a clinical setting immunohistochemistry can be used for tumour diagnostics. Through the detection of a specific marker present on the tissue sample, one can determine the stage of tumourgenesis, as well as the cell type, and origin of metastasis. As with the above mentioned techniques, immunohistochemistry generally relies on the use of fluorescently antibodies that detect and bind to the protein of interest, and thus suffers from the same limitations.⁹⁰ Additionally, tissue preservation technique plays an important role in the quality of the results, and thus the technique is only semi-quantitative.⁹¹ Recent advances have allowed this technique to be used in a high-throughput format.⁹²

1.7.4 Display technologies

Phage display, a technique awarded the 2018 Nobel Prize in Chemistry, has been used extensively in the field of targeted drug discovery. The technique involves fusing the genetic sequence of a peptide or protein into the viral genome responsible for encoding surface capsid proteins of a bacteriophage. In turn, the protein of interest is displayed on the surface of the phage, in an functionally accessible manner.⁹³ Since its development, phage display has been used as a high-throughput approach for identifying peptides that bind to a specific target.^{94–96} Whole-cell applications of phage-displayed peptide libraries has lead to the discovery of various cell-specific peptides used in clinical applications.⁹⁷ Although effective, phage display in whole-cell based applications suffers from several limitations.⁹⁸ During a phage library screening, phage clones may not display with similar efficiency, resulting in certain peptides being under-represented, and thus not selected. Additionally, over-selection

during library refinement may lead to incorrectly selected peptides. This is because phage expression levels begin to drive the selection process rather than binding affinity. In addition, the use of a live bacteriophage host creates issues such as steric hindrance between the phage and target molecule, preventing proper binding from occurring. There is also the potential for the displayed peptide to exhibit toxicity toward the live host.^{95,99} In order to work around the limitations associated with using a live host, other display technologies such as ribosome display and mRNA display have been developed.^{100,101}

Ribosome display technology utilises a DNA library of randomised peptide sequences which are transcribed into mRNA strands. The generated mRNA is translated by the ribosomes into their corresponding peptides. The mRNAs are designed to lack a stop codon, as such the ribosome stalls onto the mRNA strand, generating a peptide-ribosome-mRNA complex library. The displayed peptides are then selected for their target molecule. mRNA display works in similar manner, where a DNA library is also transcribed into mRNA strands and translated by ribosomes into peptides. The mRNA contains puromycin which covalently links the mRNA to the translated peptide generating a peptide-puromycin-mRNA complex library from which the peptides can be selected (Figure 1.11). Ribosome display and mRNA display technologies both have the advantage of screening larger libraries, and can identify higher affinity targets than the phage display technique.^{102,103}

While display technology has been used to identify clinically relevant peptides, this technique is mostly inaccessible to mainstream laboratories due to its technically demanding nature.

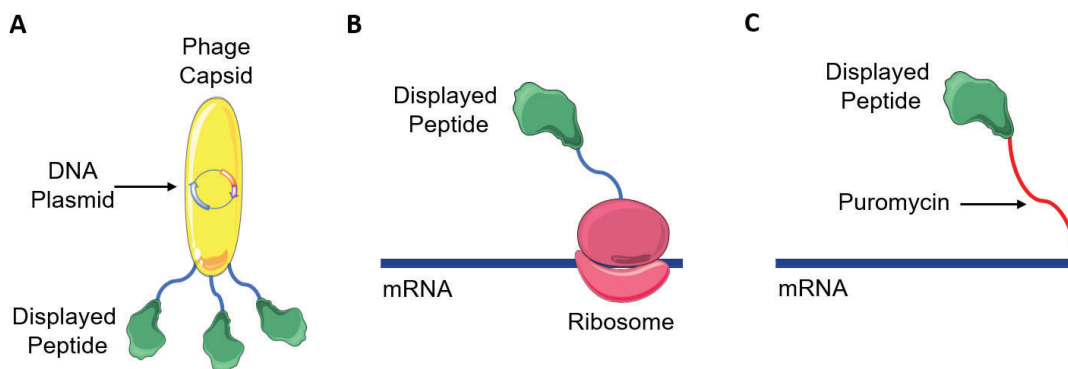


Figure 1.11: Various display technologies including: A) Phage display, where a phage with displayed peptides is used; B) Ribosome display, where the mRNA-ribosome-peptide complex is used and; C) mRNA display, where the mRNA-puromycin-mRNA complex is used.

1.7.5 Radiolabelling

Radiolabelling involves incorporating a radionuclide onto a compound of interest, and is commonly used to label peptides and proteins. The radio-emissive product can be detected which allows for the determination of binding. The most common radiolabelling strategy uses ^{125}I because of its high specific radioactivity (defined as the number of radioactive emissions per unit mass), and the convenience of detecting γ emission. Radiolabelling can also be performed using ^{131}I which has a higher specific radioactivity but has a shorter half life and higher γ emission, which means more shielding is required for handling. Radiolabelling procedures involve the use of an oxidising agent to convert Na^{125}I or Na^{131}I to its reactive form I^+ for electrophilic addition to the phenolic or imidazol side chains of tyrosine and histidine respectively. Preferential iodination at either the histidine or tyrosine can be achieved by altering the pH of the reaction.^{104,105} If iodination at histidine or tyrosine negatively impacts the activity of the peptide, amino groups (such as lysine residues or the peptide N-terminus) may be labelled with an iodinated acylating agent. However, this is higher costing than the previous method and often yields lower specific radioactivity. It is also possible to label peptides with ^{14}C or ^3H if it is crucial to obtain a radioactive product devoid of structural alteration. However, specific radioactivity of ^{14}C or ^3H are 35000 and 75 times lower than ^{125}I respectively. De novo synthesis using radiolabelled amino acids is also possible but this is much more costly than the above methods.

There are many drawbacks associated with the use of radiolabelled ligands. The amount of radioisotope used is not only governed by the cost of radioisotopes but is also limited by handling safety limits. This will compromise the specific radioactivity of the peptide since the amount of radionuclide is limiting relative to the amount of peptide. Additionally, some peptides may be unreactive to iodination or may be destroyed by the oxidative conditions required during the labelling reaction. This is especially true for oxidation-sensitive amino acids (particularly tryptophan, methionine and cysteine). The greatest risk for damage is with tryptophan residues since oxidation of the indole ring cannot be reversed with a reducing agent as it would with methionine or cysteine. Furthermore, the resulting radioactive peptides will decay over time depending on the half life of the radioisotope.¹⁰⁶ Most peptides have high affinity for their receptors therefore a high specific radioactivity is crucial.

1.7.6 Mass spectrometry

Mass spectrometry is a powerful technique that is frequently used to analyse peptides and proteins. Protein analysis using mass spectrometry can be performed using bottom-up or top-down methods.¹⁰⁷ In the bottom-up approach, proteins are enzymatically cleaved into peptides before analysis, providing information about peptide masses and fragmentation patterns. This approach is popular due to its simplicity and ease of peptide analysis, but results in loss of information such as protein modifications or splice variants. On the other hand, the top-down approach involves the direct analysis of intact proteins, allowing for a more comprehensive characterisation of the protein. However, this approach faces challenges such as low sensitivity and difficulties in analysing larger proteins. In both scenarios, the spectrum generates peptide or protein fragments that need to be matched to databases for identification, which requires prior knowledge of peptides and proteins for accurate analysis.¹⁰⁸ Moreover, the technique is destructive, leading to the loss of information regarding the native function and location of the peptide or protein.

1.8 Advanced fluorescent methods to study peptides and proteins

1.8.1 Labelling proteins with small molecule sensors *in vitro*

Traditionally, labelling of intracellular proteins was accomplished using immunolabelling techniques. This involves the use of an antibody, with high specificity to its target, as a means of directing a fluorescent molecule to a specific protein. There are two methods by which the fluorescent antibody can be targeted to the protein of interest: primary immunofluorescence, also known as direct immunofluorescence; and secondary immunofluorescence, also known as indirect immunofluorescence. Primary immunofluorescence utilises a single fluorescent antibody which binds directly to the target protein. With secondary immunofluorescence, one antibody is conjugated directly to the target, and a secondary fluorescently-conjugated antibody is targeted to the first antibody (Figure 1.12). One may select which of these methods to use depending on cost, sensitivity, signal strength, and ease of use. Direct immunofluorescence antibodies more expensive and give weaker signals, but they are more specific, and the protocol can be performed in fewer steps, when compared to indirect immunofluorescence.¹⁰⁹ A major limitation with immunofluorescence is that it is not possible to utilise this method for live cell imaging due to permeability issues. In order for the antibodies to reach the biomolecule of interest, the cell membrane must be made more permeable. This is done by cell fixation and permeabilisation.¹¹⁰ As such, this technique is unsuitable for imaging protein dynamics in real-time.

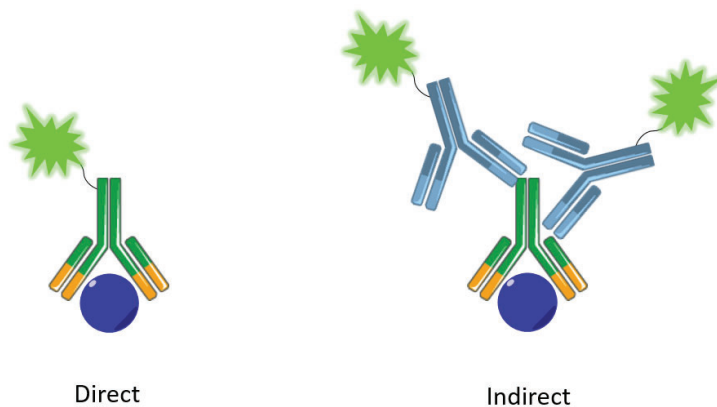


Figure 1.12: Direct immunofluorescence vs indirect immunofluorescence.

The discovery of green fluorescent protein (GFP) was of great importance to the field of biochemistry. GFP is an autofluorescent protein that was isolated from the *Aequorea Victoria* jellyfish in the 1960's.¹¹¹ However, its utility in biochemistry was not seen until the 1990's when the primary amino acid sequence was revealed, and the first examples of its expression in living organisms was demonstrated.¹¹²⁻¹¹⁴ GFP can be genetically fused to any protein of interest in live cells, allowing live visualisation of the fluorescently-labelled protein. The discovery and utility of GFP introduced the ability to study proteins in vitro without the need for cell fixation. Since its discovery, other fluorescent proteins of different colours have been isolated and used extensively to study real-time protein dynamics.¹¹⁵ However, fluorescent proteins suffer from several limitations including acid sensitivity, and the requirement of molecular oxygen to generate the chromophore.^{116,117} Most notably, is the large size of the fluorescent protein (approx. 30 kDa) which has been shown to disrupt the normal function of the protein of interest.¹¹⁸

To overcome the limitation of size, researchers have turned to the use of small fluorescent molecules. The advantage of using small fluorescent molecules is that they are small enough to enter the cell membrane without the need for permeabilisation, and cause minimal disruption to cellular function. There are several methods of introducing small fluorescent molecules into a protein of interest. These protein labelling methods are summarised below.

Self-labelling small peptides

FLAsH is the oldest known protein labelling system and is based on a short peptide sequence that can be readily appended to both in vitro and in cellulo proteins.¹¹⁹ The designed peptide consists of a tetracysteine motif in the sequence Cys-Cys-X-X-Cys-Cys (Where X is any amino acid except cysteine) that is genetically incorporated into the protein of interest. The short peptide motif adopts an alpha helical structure in which the four thiol groups are arranged on one side.¹²⁰

The FLAsH ligand, FLAsH-EDT₂ is a fluorescein-based, membrane-permeable small molecule containing two arsenic atoms complexed by 1,2-ethanedithiol (EDT) (Figure 1.13). Protein-specific labelling by this system is based on the reversible, high-affinity binding of arsenic to thiol pairs. In general, arsenic toxicity is attributed to its ability to bind tightly to proteins that contain closely spaced thiol pairs.¹²¹ However, this binding can be displaced by

small vicinal dithiols such as EDT as they have a higher binding affinity for arsenic than the cellular dithiols.¹²² Tsien *et al.* envisioned designing a short dithiol containing peptide domain with a higher affinity for the FAsH ligand than EDT ligands. The EDT ligands remain complexed to any unreacted probe, ensuring negligible arsenic toxicity when using the FAsH system.

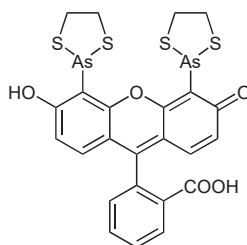


Figure 1.13: Chemical structures of the biarsenic ligand, FAsH-EDT₂.

The appropriately spaced arsenic centres of FAsH-EDT₂ bind to the tetracysteine motif in an entropically favourable manner (Figure 1.14). The advantage of using the FAsH system is the small size (<1 kDa) of the protein modification that minimises potential for interference with the function of the protein of interest when compared to fluorescent proteins (approx. 30 kDa).¹²³

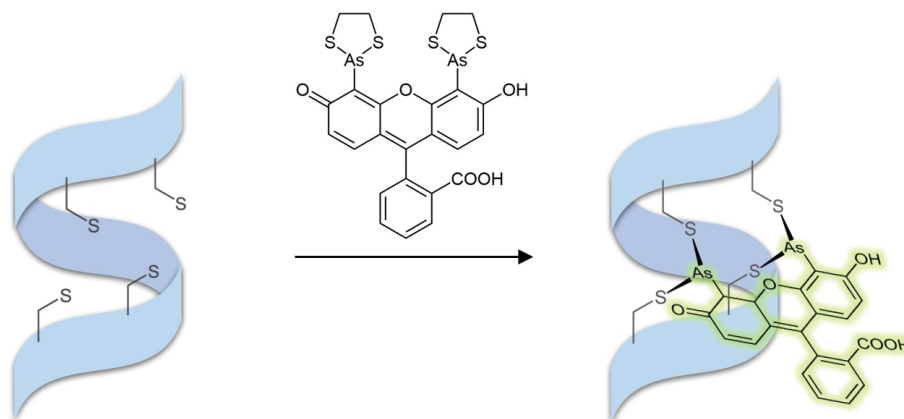


Figure 1.14: Four cysteines arranged on one side of the alpha helix such that the thiol pairs are appropriately spaced for favourable binding to the biarsenic FAsH ligand. One must also consider the isometric configuration in which the arsenic binds in a manner perpendicular to the alpha helix.

FlAsH-EDT₂ has been found to react with off-target endogenous cysteine-rich proteins, resulting in substantial background fluorescence.¹²⁴ However, the tetracysteine motif has since been optimised in attempt to maximise its affinity for FlAsH-EDT₂ thus decreasing the likelihood of non-specific background staining.¹²⁵

Of note, Flash-EDT₂ is non-fluorescent until bound to the tetracysteine motif, after which fluorescence emission can be observed at 528 nm. This property is known as fluorogenicity. It is proposed that the EDT ligands allow for the free rotation of the aryl-arsenic bonds, resulting in quenching of the excited state through photo-induced electron transfer (PET). PET is the process by which electron transfer occurs between the excited fluorophore to another redox active ligand. Once complexed to the tetracysteine, rotation is restricted which halts the quenching process and fluorescence is restored.¹²⁶ This fluorogenic property is advantageous in that it allows the protein of interest to be imaged without the need for extensive washing to remove unreacted dye before imaging (Figure 1.15).

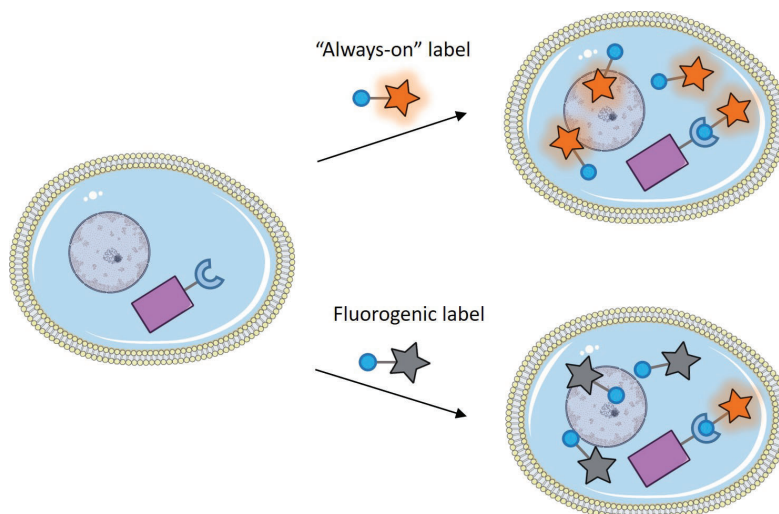


Figure 1.15: Conventional "always on" fluorescence labels vs fluorogenic labels for no wash imaging.

The success of the FlAsH protein labelling system led to the development of other FlAsH analogues such as resorufin-based, ReAsH (Figure 1.16, with a red fluorescence emission of 608 nm.¹²⁷ FlAsH and ReAsH are both commercially available, but can also be readily synthesised in two steps from inexpensive starting materials.¹²⁸

FlAsH and ReAsH can be utilised as FRET acceptors and have been combined with traditional fluorescent protein FRET donors such as CFP and YFP, allowing for ratiometric and quantitative imaging. For example, the CFP/FlAsH FRET pair was used to study the activation of G protein coupled receptors and was successfully used to visualise the activation of α -2A-adrenegic receptors in real-time.¹²⁹ Similarly, ReAsH used as a FRET acceptor to GFP was used for live-cell studies of protein folding and stability.¹³⁰

Intermolecular FRET between FlAsH and ReAsH has also been used in some protein studies. Jovin *et al.* used this FlAsH/ReAsH FRET pair to study the self-association of α -synuclein-C4, highlighting the potential for utilising this system to study the aggregation of amyloid proteins.¹³¹ While spectral overlap between FlAsH and ReAsH is not significant, increasing the amount of acceptor fluorophore to saturate the donor in a 4:1 ratio can be used to achieve more favourable FRET conditions.

Other biarsenic compounds have been developed for use with the tetracysteine motif. A modular targeting approach, SplAsH, was introduced as a non-fluorescent targeting scaffold, SplAsH-Boc, which would allow a variety of fluorophores to be used, by utilising various chemistries to substitute at the Boc-protected amine (Figure 1.16). For example, SplAsH-Alexa594 was shown to have high specificity and greater sensitivity for the tetracysteine motif than ReAsH.¹³² To the best of our knowledge, the utility of SplAsH based labels in live-cell imaging has not been extensively explored. However, there is great potential for its application in live cell studies.

Another interesting subset of biarsenic compounds comprises of sensor-based labels that change fluorescence in response to their environment, or a specific analyte. Tsien *et al.* developed a Ca^{2+} sensor based on the FlAsH-EDT₂ core tethered to a BAPTA-like chelator, CaGF (Figure 1.16). The sensor displayed a 10-fold fluorescence increase when bound to Ca^{2+} and was successfully used to image Ca^{2+} dynamics at the connexin 43 protein in live HeLa cells.¹³³

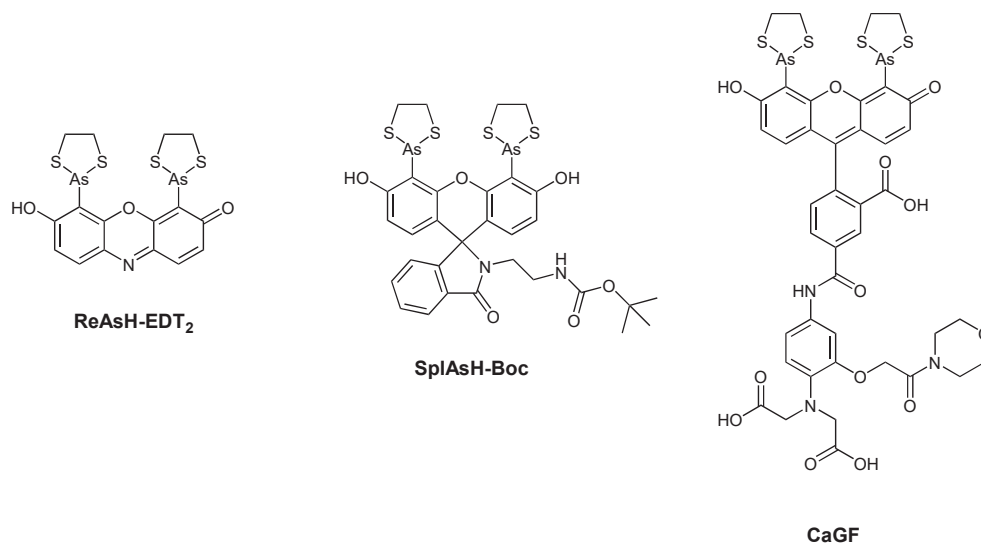


Figure 1.16: Chemical structures of various tetracysteine biarsenic ligands for use with the FlAsH system including: ReAsH-EDT₂, a resorufin based biarsenic; SplAsH-Boc, a non fluorescent biarsenic scaffold used to tether various fluorophores; and CaGF, a FlAsH compatible Ca²⁺ sensor.

Fusion proteins

Another protein-labelling strategy involves the use of fusion proteins. These systems work by genetically fusing an enzyme-tag to the protein of interest which are expressed in live cells. The fused enzyme reacts with an exogenous label to form an irreversible covalent bond, thus providing a method of highly specific and irreversible binding of various chemical functionalities, such as fluorescent sensors, to a protein of interest. This opens up the possibility for greater synthetic variation in labelling molecules (Figure 1.17).



Figure 1.17: Protein of interest (POI) fused to a tag protein reacts with an exogenous label.

There are two main steps involved with using these systems. Firstly, the protein of interest must be expressed as a fusion with the selected enzyme-tag. In order to do this,

a DNA vector with the genetic code for enzyme-tag can be obtained commercially and the genetic code of the protein of interest can be cloned at either the N- or C- terminus of the selected enzyme-tag. Selection of at which terminus to clone the protein of interest is largely determined by the proteins function and desired tagging location. For example, labelling a protein at the cell surface requires the tag to be oriented such that it is facing the extracellular surface of the cell. This POI-tag DNA vector can then be transfected into a cell-line of choice and the cell is instructed to make the protein of interest fused to the enzyme-tag. Secondly, the enzyme-tag is labelled with its specified substrate. The substrate itself can be modified to incorporate any small molecule such as a fluorescent sensor which acts as a label to the protein of interest (Figure 1.17). Since reaction of the core substrate to the enzyme is largely independent of the appended probe, the label can be virtually any molecule of choice. Details of the most prominently used enzyme tags are highlighted below.

SNAP-tag, CLIP-tag, and HaloTag

The SNAP-tag system is an self-labelling system based on a modified version of the human DNA repair protein, *O*⁶-alkylguanine-DNA alkyltransferase (hAGT) which reacts covalently with *O*⁶-benzylguanine (BG) derivatives.¹³⁴ The modified hAGT works by dealkylating the BG derivative through one of its cysteine residues to form a stable thioether bond between the enzyme and ligand, releasing a guanine in the process (Figure 1.18A). The SNAP-tag enzyme was modified from native hAGT using a directed evolution approach.^{135–137} Compared to wildtype hAGT, SNAP-tag possesses a 52-fold higher activity toward BG derivatives, lower susceptibility to proteolytic degradation, greater stability post-labelling, and does not bind to DNA.¹³⁸

A close relative to SNAP-tag, CLIP-tag utilises a different variant of hAGT which reacts irreversibly to *O*²-benzylcytosine (BC) derivatives. The mechanism occurs in a similar fashion to SNAP-tag, where a cysteine residue dealkylates the BC to form a thioether bond, releasing a cytosine (Figure 1.18B). CLIP-tag was designed to be orthogonal to SNAP-tag, allowing them both to be used for simultaneous labelling of two different proteins in the same cell without cross-reacting.¹³⁹

Another system known as the HaloTag system is based on a modified bacterial haloalkane dehalogenase enzyme, designed to link covalently to synthetic haloalkane-bearing ligands.

The enzyme cleaves the terminal halogen via an S_N2 type reaction, forming an ester bond between the alkane, and an aspartate residue located deep within the enzyme binding pocket. In the wild-type enzyme, this ester intermediate is readily cleaved through base-mediated hydrolysis, catalysed by a nearby histidine residue. The modified HaloTag enzyme is designed such that this histidine is replaced with a phenylalanine residue, which is ineffective as a base. This inhibits the ester hydrolysis, thereby trapping the intermediate as a stable covalent adduct (Figure 1.18C).¹⁴⁰

An array of fluorescent SNAP-tag, CLIP-tag and HaloTag substrates have been made commercially available. These fluorescent labels are based on a variety of "always-on" fluorophores which cover a range of the blue to near infrared spectrum.

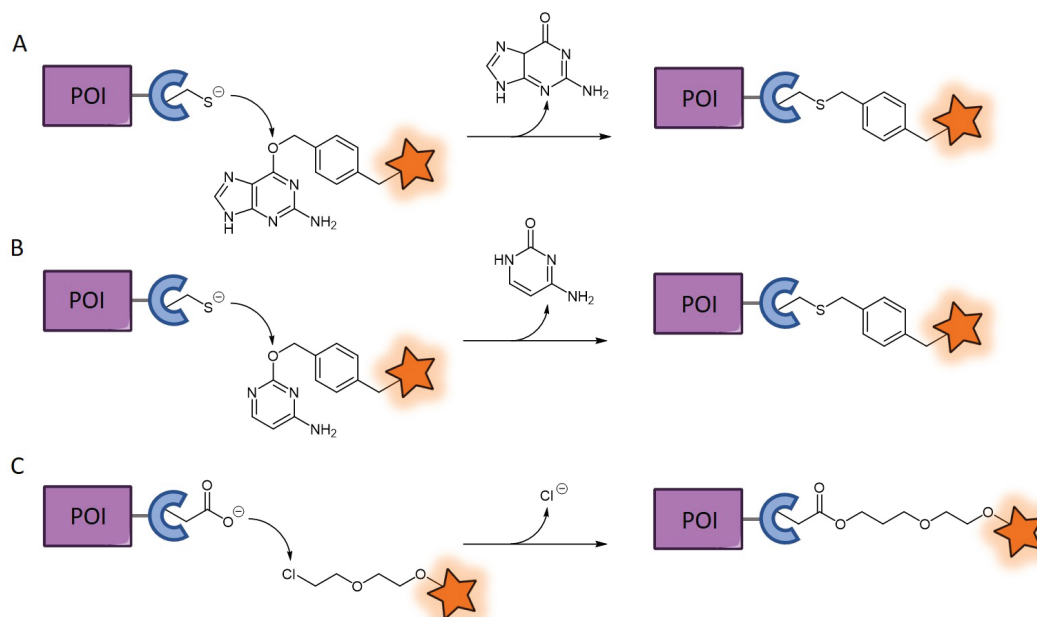


Figure 1.18: Schematic representation of the A) SNAP-tag, B) CLIP-tag, and C) HaloTag mechanisms. The core substrate can be appended to a small fluorescent sensor, shown as an orange star.

Synthesis of fluorescent derivatives of the core substrates is relatively straight forward. By synthesising a fluorescently attached derivative of the enzyme substrates, a protein of interest can be tagged. Fluorescent derivatives of the core substrate allows for the protein of interest to be imaged directly.

Ligand-directed labelling

Ligand-directed labelling is another method for conjugating fluorescent sensors to a POI. One widely utilised approach involves the use of tosylates, wherein a ligand specifically designed for the POI is synthesized with a tosyl ester attached to a fluorescent label. The labelled ligand selectively binds to the active binding site of the target protein. Subsequently, an S_N2 -type reaction takes place between the tosylate group and a nearby nucleophilic amino acid, resulting in the transfer of the fluorescent probe to the target protein through acyl transfer.¹⁴¹

Compared to other labelling methods, ligand-directed chemistry is considered less invasive since it doesn't require modification of the POI. This technique also enables labelling in close proximity to the active site of the protein.

1.8.2 Fluorescence in multimodal imaging

Fluorescence has been combined with various other analytical modalities for a multimodal output. Multimodal analysis is advantageous as it combines the strengths of different techniques to overcome the limitations of each individual technique. The combination of different techniques can provide complementary information that allows for a more comprehensive analysis, leading to a more holistic understanding of the system being studied. Some notable examples of multimodal analysis involving fluorescence are reviewed below.

Photoacoustic imaging

Photoacoustic imaging (PAI) is a technique that combines the strengths of both optical imaging and ultrasonic imaging. PAI, a short pulse of laser light is directed into the tissue and absorbed by biomolecules or exogenous contrast agents. The absorbed energy from the pulse quickly transforms into heat, resulting in a local increase in pressure due to thermal expansion which is detected as the photoacoustic signal. PAI is not limited by optical scattering, allowing for deep tissue imaging.^{142,143} This is in contrast to fluorescence, which is limited by optical scattering. However, penetration depth of fluorescence is largely dependent on wavelength, where longer wavelengths allow for deeper penetration.¹⁴⁴

PAI provides both anatomical and functional information, while fluorescence imaging provides enhanced temporal resolution (the duration of time between successive data mea-

surements), sensitivity, and the ability to image at the intracellular level. Together, PAI and fluorescence imaging can be a highly effective tool in biological imaging. As such, exogenous contrast agents with fluorescent properties have been used in a multimodal setup with PAI. Generally, the fluorophore needs to have a near infrared emission (I) window (650 - 900 nm) that matches the laser pulse of PAI instruments.¹⁴⁵

Wang *et al.* designed a new cyanine-based probe, CPH, that could target the Golgi apparatus and monitor its pH in real-time using both PAI and fluorescent imaging. The probe, CPH, was constructed with a phenylsulfonamide moiety to specifically target the Golgi apparatus, and a tricarbo-cyanine dye to produce both PAI and fluorescence signals. PAI imaging with CPH in whole mice revealed a Golgi pH decrease of 0.53 after LPS-stimulation. Fluorescence imaging of the mice was then used to confirm a Golgi pH of 5.72 as compared to 6.5 for normal mice.¹⁴⁶

Magnetic resonance imaging

Magnetic resonance imaging (MRI) is an imaging method that uses nuclear magnetic resonance (NMR) to produce images based on differences in NMR relaxation time in tissues. MRI is a versatile technique that is frequently used in research and diagnostic medicine to provide information about anatomy and physiological processes. However, MRI is a relatively insensitive technique, and often requires the use of contrast agents. These commonly include paramagnetic ions like gadolinium (III) or manganese (II).^{147,148} Other contrast agents include heavy isotopes of atoms with a non-zero net spin, such as ¹⁹F. A major drawback of MRI imaging is its low sensitivity to weakly magnetic structures and poor spatial resolution.¹⁴⁹ Dual functional MRI contrast agents that incorporate a fluorescent moiety have been developed for multimodal analysis which has allowed complimentary information to be gained using both techniques.

Mizukami *et al.* developed a novel dual-function probe based on a gadolinium complex, Gd-DOTA-DEVD-AFC, to detect caspase-3 enzyme activity. The probe consists of the Gd³⁺-DOTA complex, caspase-3 substrate peptide (DEVD), and ¹⁹F-containing fluorophore (AFC). When caspase-3 cleaves the C-terminus of the DEVD sequence, AFC is released, and both the ¹⁹F MRI signal and the fluorescence spectrum of AFC increase.¹⁵⁰ Such multimodal probes allow for the detection of complimentary macroscopic, and microscopic information.

Radioisotopic-imaging

Positron emission tomography (PET) and single-photon emission computed tomography (SPECT) are two commonly used imaging techniques used to visualise and study the distribution and behaviour of cells, tissues, and molecules in living organisms. PET uses radioisotopes that emit positrons as a tracer and can produce high-resolution images of molecular activity in the body. SPECT, on the other hand, uses gamma-emitting radiotracers, which has the ability to give correlated and quantified information that is not possible with fluorescence alone. Fluorescence imaging offers superior spatial resolution (~ 200 nm) in comparison to radioimaging (~ 5 mm). However, a drawback of fluorescence imaging is its limited tissue penetration depth, which varies based on the wavelength of the fluorophore employed. On the other hand, radioimaging has the advantage of deeper tissue penetration, enabling imaging of structures at greater depths. As such combining these techniques allows for a more comprehensive understanding of biomolecular samples.

Chen *et al.* developed a novel dual-modality imaging probe that effectively combines PET and fluorescence imaging to detect tumours *in vivo*. The probe was specifically designed to target protein binding, achieved by conjugating vascular endothelial growth factor receptor (VEGFR) to a fluorescent quantum dot, which was further functionalised with ^{64}Cu -labelled DOTA for PET detection. This dual-modal probe not only enabled fluorescent localisation of VEGFR-2 receptors in whole mice, but the addition of PET capability also allowed for accurate quantification of targeting efficacy.¹⁵¹

Rijpkema *et al.* developed a dual-labeled antibody, MN-14, tagged with both a radio-label (^{111}In) and a fluorophore (IRDye 800CW) for dual SPECT and fluorescence guided surgery. A target antigen, carcinoembryonic antigen, was also tagged as it has consistently high expression in both primary and metastatic colorectal cancer lesions. While useful at providing some level of localisation, SPECT imaging alone, does not allow for precise tumour delineation during surgery. By incorporating the fluorophore, it was found that greater delineation of the tumour could be achieved during surgery, which allowed for excellent results from the radical excision of lymph-node tumour tissue in mouse models.¹⁵²

Vibrational spectroscopy

Vibrational techniques such as infrared (IR) and Raman spectroscopy rely on the unique composition of chemical bonds within a structure. IR spectroscopy exploits the fact that molecular bonds absorb characteristic IR frequencies, while Raman spectroscopy measures the relative frequencies at which this radiation is scattered.¹⁵³ Vibrational spectroscopy imaging provides the capability to visualise critical macromolecular characteristics, including protein aggregation.¹⁵⁴ Despite its ability to provide important structural information, vibrational spectroscopy has limited spatial and temporal resolution.¹⁵⁵ This can be overcome by combining it with fluorescence, which provides complementary spatial and temporal information to the vibrational technique.

5-Cyanotryptophan (5CNTrp) was designed as a modified form of tryptophan with a nitrile group attached at the 5-position, and it possesses strong IR absorption and intrinsic fluorescence properties, making it a promising probe for exploring the structural and dynamic features of proteins. The IR-absorbing nitrile group provides information about the local structure of the protein because it is responsive to changes in electrostatics and hydrogen bond interactions, while fluorescence provides information on localisation.¹⁵⁶

X-ray fluorescence microscopy

X-ray fluorescence microscopy (XFM) is a useful imaging technique in biological cell and tissue studies as it gives information about the elemental distribution within cells and tissues.¹⁵⁷ Another commonly used X-ray based imaging method in this field is X-ray absorption spectroscopy (XAS), which provides insights into the oxidation state of metals present in cells and tissues.¹⁵⁸ These X-ray imaging techniques are particularly useful in the study of nanomaterials and their interactions with biological systems. XFM and XAS alone may not provide a complete understanding of the subcellular localisation of the element. In order to fully understand the function of the element within the cell, it is often helpful to combine XFM and XAS with other imaging such as fluorescence microscopy. Dual imaging agents for multimodal X-ray and fluorescence imaging have been developed, with the majority being based on nanoparticles.¹⁵⁹ However to the best of my knowledge, there are no XFM and XAS dual fluorescent sensors that have been developed for the study of protein interactions and this remains a niche to be filled.

1.9 Aims

Fluorescence has proven to be a powerful technique in the field of biology, providing a robust tool for investigating and understanding cellular function at the molecular level. In recent years, there have been significant advancements in the development of fluorescent probes and sensors, which have greatly expanded the range of applications. Recent developments in the field of fluorescence have led to the development of novel scaffolds for fluorescent probes and sensors, which have greatly enhanced their properties.

Proteins and peptides are essential biomolecules in biology, playing a fundamental role in many critical cellular processes. These biomolecules are responsible for an array of significant interactions like signalling, catalysis, and defence. The discovery of novel peptide and protein interactions has paved the way for the development of new and powerful therapeutic and diagnostic modalities.

The use of fluorescence as a tool for protein studies has been consistently used in more traditional techniques like ELISA. However, such techniques suffer in their utility by the fact that they do not have the capacity to assess the interactions of proteins and peptides in their native environment. This is because they require the cell to be lysed in order to access cellular contents.

As such, it is of great importance to develop novel methods that aid in our understanding of peptide and protein interactions. By utilising fluorescence for this goal, the potential for such interactions to be studied in their native state can be unlocked. This has the potential to provide unrivalled assessment of such interactions in their most pure biological state. As such, the aims of this thesis were to:

1. Develop a fluorescent probe for multimodal mass spectrometry imaging, and fluorescence spectroscopy.
2. Create and investigate new sensors designed to measure the redox state of particular proteins.
3. Work towards the development of a novel method for quick identification, and isolation of peptide-receptor binding.

Chapter 2

Multimodal probe for fluorescence and mass spectrometry imaging

2.1 Overview

Fluorescence is a powerful technique for studying cells and cellular processes. As outlined in section 1.8.2 the use of fluorescence imaging in a multimodal setup allows for the collection of orthogonal or complimentary information. This chapter will outline the development and implementation of a fluorescent probe that enables multimodal analysis by mass spectrometry imaging (MSI), and fluorescence imaging using a single sample.

2.1.1 Principles of mass spectrometry imaging

Mass spectrometry (MS) is a powerful analytical tool used to identify molecular masses present within a sample. This technique works by sample ionisation and detection based on their mass-to-charge ratios (m/z). The process of MS analysis begins with the ionisation of the molecular species present within the sample. Ionised compounds travel through a magnetic field, which deflects the samples based on their (m/z). The species are then detected by the mass spectrometer and are displayed as a spectrum of (m/z) vs intensity.

There are multiple ionisation methods that are used for a mass spectrometry experiment. These include desorption electrospray ionisation (DESI), secondary ion mass spectrometry (SIMS), matrix-assisted laser desorption-ionisation (MALDI) and laser desorption/ionisation (LDI). The most common of these techniques is MALDI, which involves the

application of a matrix onto the sample that absorbs energy from a laser and transfers it to the underlying sample molecules, causing them to ionise and desorb. MALDI is generally considered a soft ionisation method that causes minimal fragmentation of the ionised molecules. However, this also makes MALDI less sensitive than other MSI techniques. Additionally, the matrix can cause interference in the low molecular weight range, as the matrices ionise and fragment in this region of the mass spectrum. On the other hand, LDI uses the laser source to ionise and desorb the sample molecules directly, rather than through a matrix. This makes LDI more sensitive than MALDI but, the harder ionisation can cause fragmentation of the molecules, making interpretation more difficult.¹⁶⁰

With MSI, MS is used to record the molecular distribution at multiple individual sampling positions. The spectra are then combined to form a 2D raster, and an image is generated by selecting a single m/z to display in terms of relative signal intensity. The image is presented in false-colour, generally in the form of a heat-map Figure 2.1.

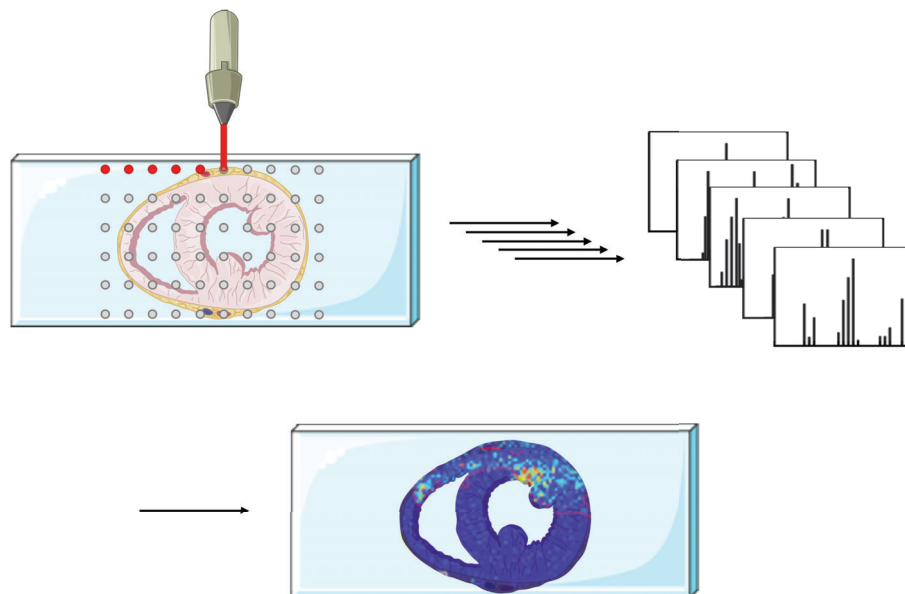


Figure 2.1: Schematic representation of the mass spectrometry imaging process where mass spectra are obtained at multiple points on a sample. The chosen mass is typically visualised as a false-colour heat map that represent the signal intensity.

When first developed, MSI was used for the analysis of protein expression in tissue samples.¹⁶¹ Since then, it has been utilised for a number of different biomolecules including drugs, proteins, peptides, ligands and many more.^{162,163} It has been successfully used to gain

information about their distribution, isoforms, post-translational modification, cleavage, and degradation.^{164–168} Despite this, some biomolecules remain inaccessible to MSI depending on the ionisation technique and sample preparation being used.¹⁶⁹ With MALDI-MSI, these mostly include hydrophobic, membrane-bound, high mass (>30 kD), or low-abundance proteins, as well as lipids, phospholipids, and oligonucleotides.

2.1.2 Mass tags

To overcome this limitation, mass-tags were designed as reporter molecules that cleave under MSI conditions to release a specific mass tracer. The mass-tag is comprised of three segments: the biomolecule point of attachment, cleavable linker, and reporter molecule. First, the mass-tag is bound to the biomolecule of interest. After this, irradiation with the ionisation source causes the linker to cleave and release the reporter molecule (Figure 2.2). The reporter is designed to be a molecule with a known mass that is easily detected by the mass spectrometer. This technique enables visualisation of otherwise inaccessible biomolecules of interest.

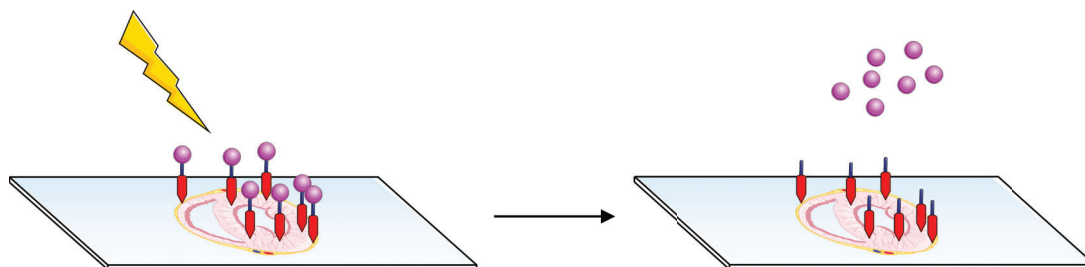


Figure 2.2: Schematic representation of the mass tag containing a biomolecule point of attachment (red), linker (blue) and reporter molecule (purple) that is cleaved upon irradiation with the ionisation source.

The concept of the mass tag was pioneered by Olejnik *et al.*, who designed a photo-cleavable linker that attached a peptide-based reporter to DNA sequences.^{170,171} This mass tag was used as a hybridisation target to detect particular DNA sequences within the sample by mass spectrometry. The first use of these photo-cleavable mass-tags in MSI was reported by Lemaire *et al.*, who similarly designed a photo-cleavable linker with a complementary mRNA sequence as the biomolecular point of attachment, and a peptide-based reporter. This mass tag was used to image mRNA sequences using MALDI-MSI.¹⁷² Since then, mass-

tags have been used frequently in the field of MSI.

The points of attachment of reported mass-tags have been very diverse. Many have been designed to be combined with other techniques like *in situ* hybridisation and immunohistochemistry. Other mass-tag biomolecule points of attachments have included lectin binders for glycan imaging, antibodies for immunohistochemistry-like MSI, and fluorophosphonates for analysis of serine hydrolases.¹⁷³⁻¹⁷⁷

Mass-tag reporter molecules are most commonly based on short peptide-sequences of known masses.¹⁷⁸ However, there are reports of mass-tag reporters that are based on small molecules and even nanoparticles.^{174,179} More recently, Han *et al.* designed mass-tags based on a Ru(II) polypyridine complex in which, upon UV irradiation, the polypyridyl undergoes a photosubstitution, releasing the ionised Ru(II) complex to provide a fingerprint signal. This mass-tag was used to image integrins $\alpha v \beta 3$ in human tumour tissue.¹⁸⁰

2.1.3 Multimodal MSI

While MALDI-MSI has become an integral technique spatial characterisation of biological samples, there are some limitations to this technique. Indeed, with MS in general, relative abundances cannot be determined due to differences in the ionisation potential of various compounds.¹⁸¹ This ionisation potential is also known to be influenced by the sample preparation method.¹⁸² As such, MSI is not a quantitative technique. Additionally, the resolution of MSI instruments ranges from 10 - 100 μm .¹⁸³ spatial resolutions of 1.4 μm have been made possible, but only with the use of specialised equipment which is not readily accessible to most labs.¹⁸⁴ Therefore, multimodal techniques can be advantageous in providing complimentary information to overcome the limitations of MSI.

There have been reports of a few techniques used with MSI in a multimodal setup. However, many of the multimodal analyses are performed on adjacent tissue slices, rather than the same sample, possibly due to interference between the two techniques.^{185,186} The use of adjacent tissue may not always be feasible, particularly when the sample being analysed is not tissue based. As such, the following section focusses only on techniques that have been applied to the same sample.

Multimodal approaches have been used as a means of identifying regions of interest within a tissue sample, prior to subsequent imaging with MSI. Dubbed "histology-directed

MSI", this has been achieved using autofluorescence microscopy as demonstrated by Patterson *et al.*, who used this technique to identify malaria infected hepatocytes and golmeruli in mouse liver and kidney tissue respectively.¹⁸⁷ More commonly, histology-directed MSI has been achieved using immunohistochemistry techniques. For example, Rujchanarong *et al.* utilised MSI to study the protein distribution of breast tumour tissue depending on whether immunohistochemical staining with PTEN occurred.¹⁸⁸ Similar techniques were employed to study the metabolic signatures alpha and beta cells within islets of Langerhans, and ganglioside monosialated ganglioside deposit distribution in brain tissue from Hunter's disease mouse model.^{189,190}

Multimodal MSI has also been used as a means of enhancing image sharpness and achieving a hybrid image that combines the desired properties of each technique. In 2015, Van de Plas *et al.* combined optical microscopy with MALDI-MSI which allowed the prediction of ion distribution in non-measured tissue down from 100 to 10 μm resolution.¹⁹¹ In 2018, Neumann *et al.* used IR spectroscopy to sharpen the MALDI-MSI image down to 5 μm resolution, and provided additional information not identifiable by the individual techniques.¹⁹²

MSI has also been used in a multimodal setup with other MSI techniques. For example MALDI-MSI has been used in combination with other ionisation techniques like DESI or SIMS.^{164,193} It has also been used in multimodal imaging with itself, by utilising multiple acquisition parameters to uncover more ionised species. For example, there are reports of the use of dual polarity acquisition (i.e. positive and negative ion modes) to reveal information about lipid microenvironments in lung cancer xenografts,¹⁹⁴ lipids involved with AB plaques,¹⁹⁵ and changes in peptide- and lipid-based pathologies in a mouse model of myocardial infarction.¹⁹⁶ multimodal MSI has also been achieved with the use of low, and high laser fluences to reveal MALDI, and NAPA spectra respectively, to study the varying lipid species in brain and lung mouse tissue.¹⁹⁷

Confocal raman microscopy has also been used in conjunction with MSI. Raman is a non-destructive technique that yields information about functional groups at higher resolutions than MSI. While MSI can identify characteristics of molecular masses with high specificity. Ahlf *et al.* used these two techniques in combination to characterise the composition of cell spheroid sections, by correlating the two techniques with principal component analysis.¹⁹⁸

X-ray fluorescence has also been used in a multimodal setup with MSI. This addresses the major limitation of MALDI-MSI - the varied ionisation potentials that lead to non-quantitative results. X-ray fluorescence is a technique based on core-electron excitation, and emission by X-ray irradiation, resulting in characteristic spectra unique to each individual element. Svirkova *et al.* combined MALDI-MSI for sphingomyelin, and phosphatidylcholines detection, in combination with μ X-ray fluorescence to give elemental information about bone structures though the imaging of elements like calcium and phosphorus .¹⁹⁹

Multimodal MSI and fluorescence

Fluorescence imaging has been used as a multimodal adjunct to MSI. The fluorescent component in all the cases presented are either a fluorescent protein, or an antibody labelled with a fluorescent tag.

For example, Chughtai *et al.* used a multimodal fluorescence and MSI setup. A fluorescent protein, dTomato, was genetically encoded into MDA-MB-231 breast cancer xenografts. This enabled fluorescence microscopy imaging of hypoxic regions within the xenograft. After this, MSI identified localisation of the dTomato protein by detecting a tryptic peptide that was ionised directly from the fluorescent protein. The tryptic peptide was only present in hypoxic but not normoxic tumour regions identified by fluorescence microscopy²⁰⁰

Hueber *et al.* used a multimodal fluorescence in situ hybridisation combined with MALDI for same sample analysis. MALDI-MSI was used on gastroesophageal adenocarcinoma which revealed AMP as a prognostic marker. After this, CEP17 in situ hybridisation with HER2 was performed to show an overexpression of this tumour marker. By utilising both techniques, the synergistic information was used to enhance patient prognosis.²⁰¹

More recently, Yagnik *et al.* combined fluorescent immunohistochemistry by using fluorescent antibodies tagged with a mass tag. This was done in up to 12-plex imaging to indentify multiple analytes of interest in tonsil and breast cancer tissue.¹⁷⁶

2.2 Aims

MSI is a powerful technique that has allowed for the analysis of specific analytes within a biological sample. However, depending on the ionisation source, there are limitations to the types of molecules that can be ionised within a sample. With MALDI-MSI, one of the biggest

limitations is its inability to assess lipid-based and large compounds (>30 kDa), which are inherent to almost all biological samples. Additionally, MSI is not a quantitative technique due to varying ionisation potentials of chemical species. To overcome the limitations of MSI, many techniques have been utilised. For example, MSI of inaccessible biomolecules of interest has been addressed with the development, and use of mass-tags. Further, multimodal analysis with multiple techniques such as fluorescence imaging has helped to overcome many other limitations including spatial resolution, and ionisation potentials.

To this end, this project was designed to combine the two powerful techniques of mass-tags and multimodality by creating a fluorescent mass-tag for direct fluorescence and MALDI-MSI analysis of a single sample. This would allow the analysis of any protein of interest. As such, the aims of this chapter were to:

1. Rationally design and synthesise a fluorescent mass-tag probe.
2. Assess the probes compatibility with fluorescence imaging and MALDI-MSI.
3. Validate the mass-tags utility in a protein- and tissue-based single-sample multimodal imaging setup.

2.3 Design of a fluorescent mass-tag

2.3.1 Criteria for mass-tags

In order for a molecule to be visible and compatible with biomolecular MALDI imaging, certain criteria need to be met. These are summarised below.

1. The photo-cleavable linker must be attached covalently to both the reporter molecule and the biomolecular point of attachment in a way that allows the reporter to cleave selectively.
2. The photo-cleavable linker must cleave at the wavelength of the MALDI laser (338 - 355 nm). This ensures that the reporter is released during the imaging sequence.
3. Cleavage of the reporter from the photo-cleavable linker must occur only during the imaging sequence. This ensures there is no migration of the reporter before imaging, allowing precise signal localisation.
4. The reporter molecule must be of known mass. This allows the user to identify the

reporter compound on the mass spectrum.

5. The mass of the reporter molecule must not lie within regions of high endogenous activity. This ensures the mass of the reporter is not lost within the dense signal, but also ensures mass is not misidentified as other common endogenous compounds.
6. The reporter molecule must be capable of ionising using the MALDI ionisation technique in order to be detected.
7. The reported molecule should be water soluble. This allows it to be compatible with both tissue preparation methods and MALDI sample preparation methods, both of which are generally carried out in aqueous media.

As such, the dual probe was designed to accommodate these requirements, while also ensuring it contained a fluorescent moiety. This fluorophore must not exhibit emission within the range of the photo-cleavable linker, so as to avoid self-induced cleavage.

2.3.2 Probe design

The chosen photo-cleavable linker was based on the ortho-nitrobenzyl (ONB) moiety. The development of this type of photo-cleavable linker was first described in 1998.¹⁷⁰ The ONB group is commonly used in the structure of mass-tags due to its cleavage absorbance of 290 - 340 nm, which corresponds to the wavelength of the MALDI ionisation laser.²⁰² In addition, these molecules can be designed with two points of attachment allowing the attachment of both the biomolecular binding moiety as well as the fluorescent reporter. The ONB derivative used in this work was designed by Katherine Stevens from the Pukala lab at the University of Adelaide (unpublished work) and features two points of attachment. The first attachment point is functionalised with a biomolecular point of attachment, a succinimidyl ester that reacts to primary amines, specifically, allowing it to be attached to any protein of interest containing a primary amine. Additionally, this molecule was designed to contain an alkyne moiety, which allows for easy attachment of an azide-functionalised fluorescent reporter molecule *via* a copper-catalysed azide-alkyne cycloaddition (CuAAC) or "click" reaction (Figure 2.3).

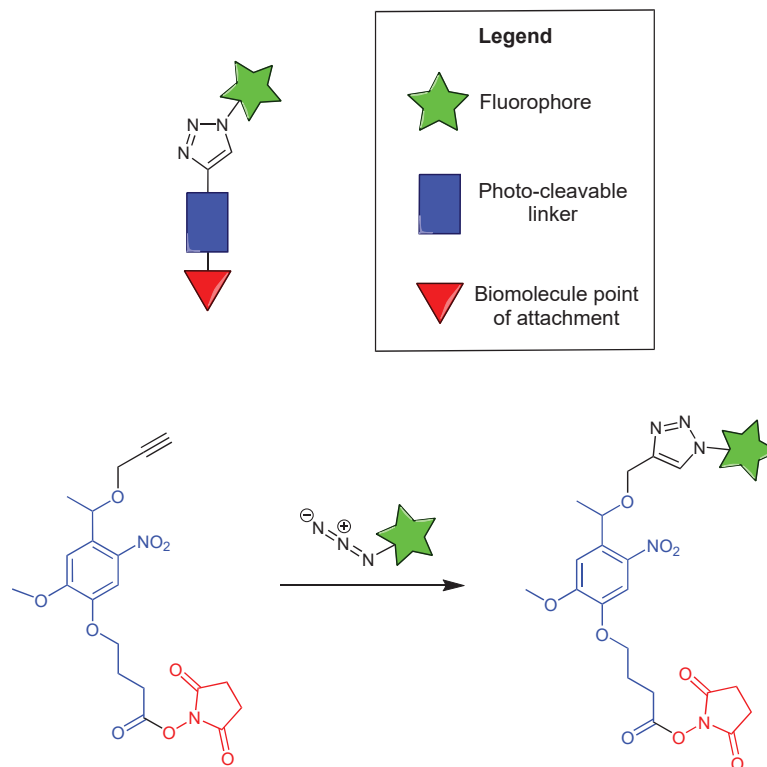


Figure 2.3: Mass-tag design based on an ONB linker (blue) with an N-hydroxysuccinimide biomolecular point of attachment (red), and an alkyne moiety allowing for facile conjugation of an azide-functionalised fluorescent reporter using "click" chemistry.

The fluorescent reporter is released from the ONB group through a radical mechanism. Upon irradiation, cyclisation occurs to generate a benzisoxazolidine intermediate, which subsequently undergoes ring opening. This results in the loss of the ether-bound reporter molecule, leaving behind a 2-nitrosobenzyl derivative (Figure 2.4)^{203,204}.

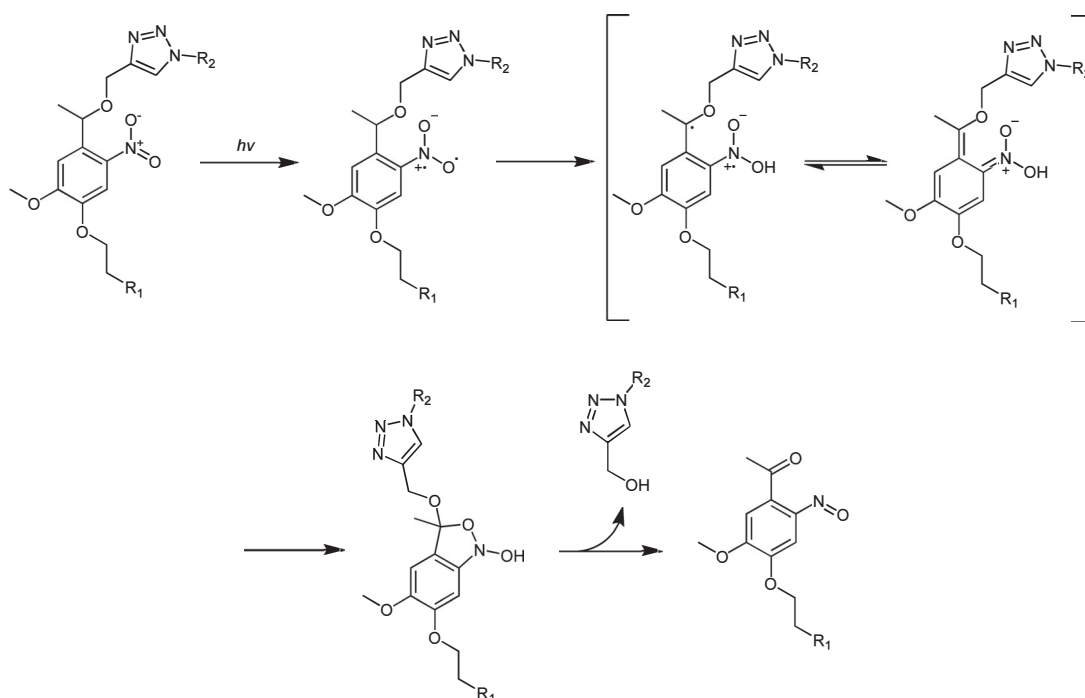


Figure 2.4: Cleavage mechanism of the ONB moiety upon irradiation.

With this in mind, the reporter must be designed to be ionisable, water-soluble, and fluorescent. For this reason, a hemicyanine fluorophore was chosen for the first design. Hemicyanines have an inherent positive charge that allows the tag to be readily detected by MS, eliminating the risk of poor ionisation. This charge also makes them water-soluble. Additionally, hemicyanines fluoresce in the orange region, making them ideal for deeper tissue imaging.¹⁴⁴ The final design is shown in Figure 2.5.

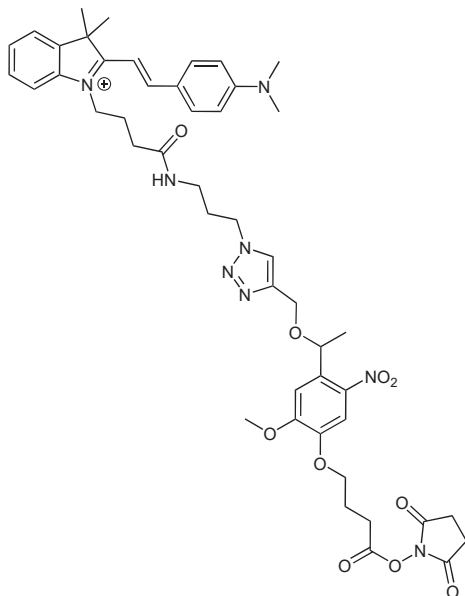
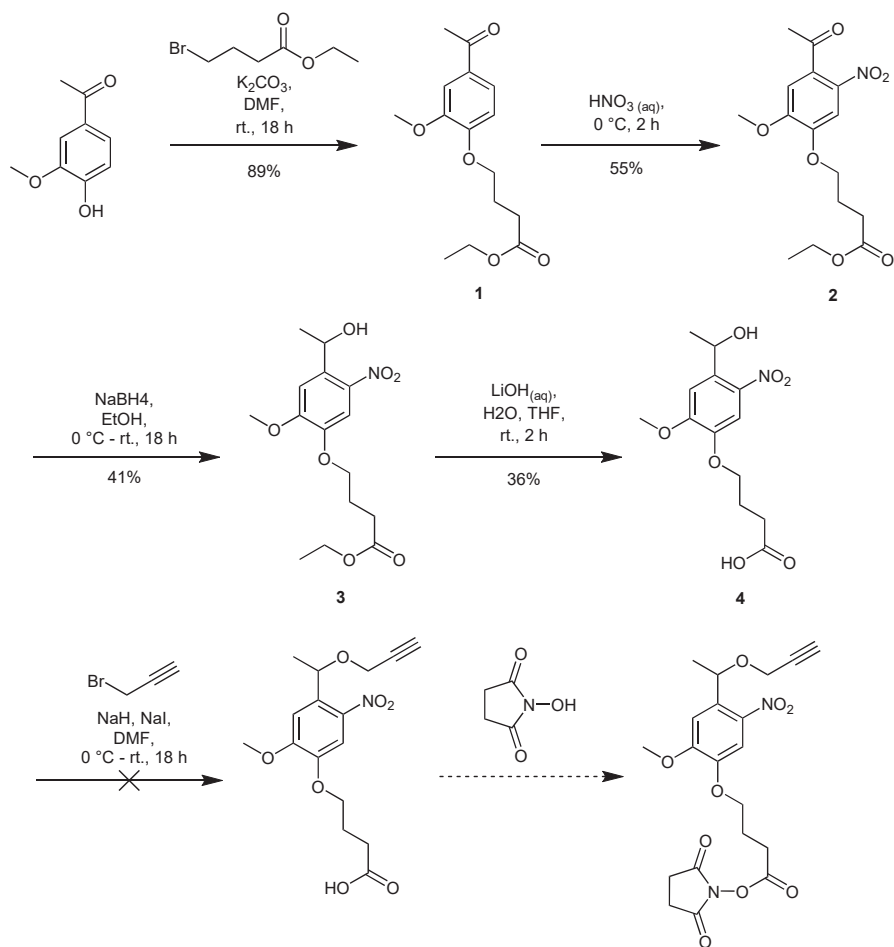


Figure 2.5: Design of the hemicyanine-based fluorescent mass-tag, **CyMT**.

2.4 Synthesis

2.4.1 Synthesis of photo-cleavable linker

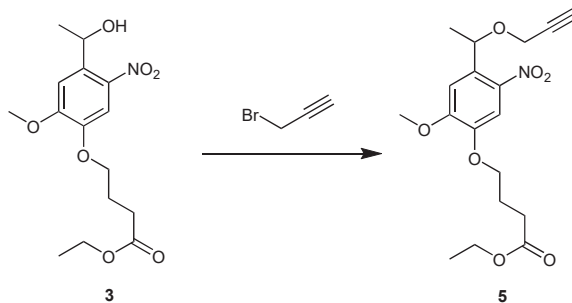
Synthesis of the photo-cleavable linker began with the alkylation of acetovanilone with an ester protected carboxylic acid using 4-ethylbromo butyrate to yield **1**. Nitration of the 2-position was performed in aqueous nitric acid to yield **2**. The ketone was then reduced to the secondary alcohol using sodium borohydride to give **3**. The ethyl ester was then hydrolysed in basic conditions to yield **4**. An attempt was made to install the alkyne at the free hydroxyl group with propargyl bromide using a Finkelstein-type reaction. However, this reaction resulted in loss of the starting material, but did not yield the desired product (Scheme 2.1). It was speculated that the free carboxylic acid may have been the reason for the formation of unwanted products. Therefore, this alkylation was reattempted while the carboxylic acid was still protected by the ester group by using **3**.



Scheme 2.1: Attempted synthesis of the photo-cleavable linker.

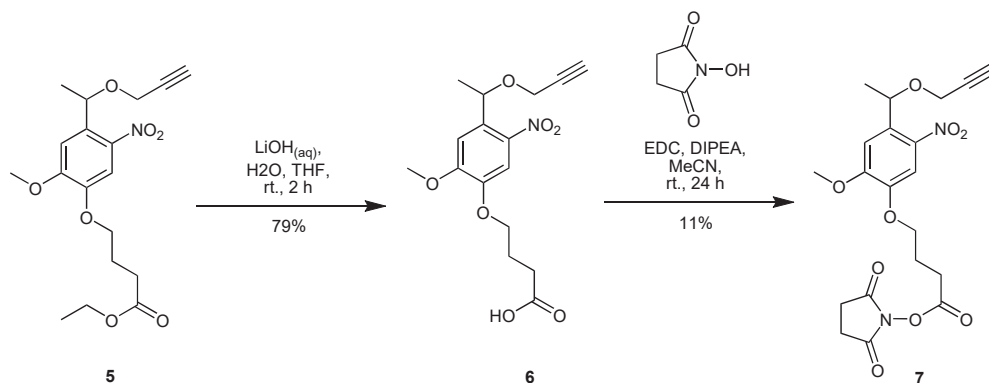
Initially, there were concerns that the use of sodium hydride could potentially result in hydrolysis of the ethyl ester, leading to the formation of multiple products. To avoid this, milder alkylating conditions were explored, using sodium hydroxide as a base and tetrabutylammonium bromide as a catalyst. However, this reaction did not yield the desired product. Subsequently, the Finkelstein-type reaction was attempted without the use of sodium hydride, but no reaction occurred even after leaving the reaction for 48 hours. Eventually, the original reaction conditions using sodium iodide and sodium hydride were used, which resulted in the formation of the desired product, **5** (Table 2.1).

Table 2.1: Reaction conditions trialled for the synthesis of **5**.



Reagent(s)	Conditions	Outcome
NaOH, TBAB	DMF, rt., 18 h	Unsuccessful
NaI	DMF 0 - 40 °C, 48h	No reaction
NaH, NaI	DMF 0 °C - rt., 18h	77.9% yield

From here, the ethyl ester was hydrolysed under basic conditions to yield **6**. the free carboxylic acid was coupled to *N*-Hydroxysuccinimide using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) as the coupling reagent, which yielded the final photo-cleavable linker, **7**.

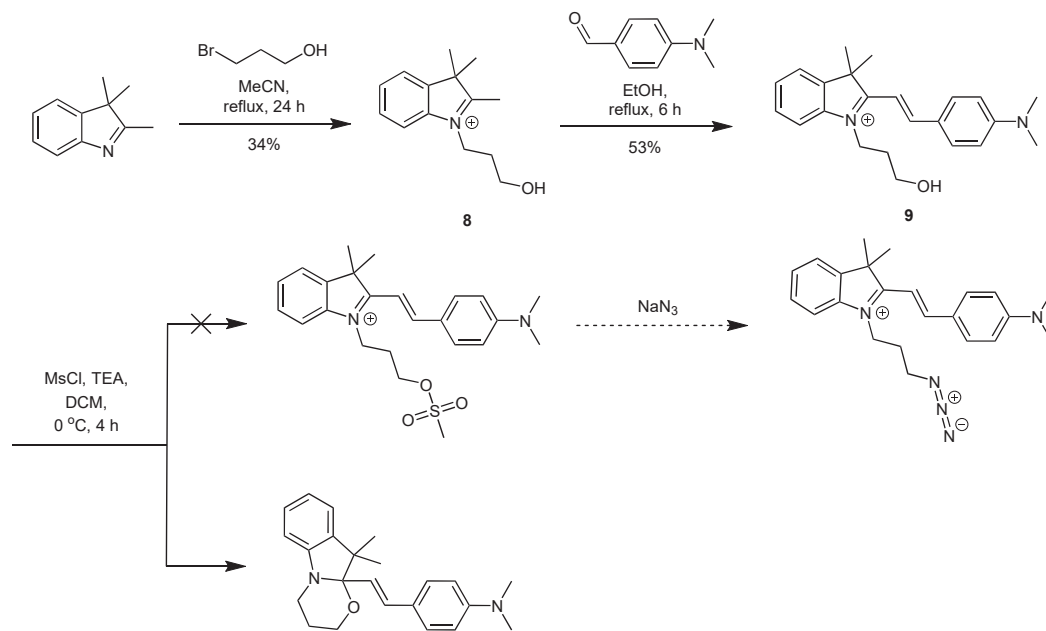


Scheme 2.2: Synthesis of the final photo-cleavable linker, **7**.

2.4.2 Synthesis of CyMT

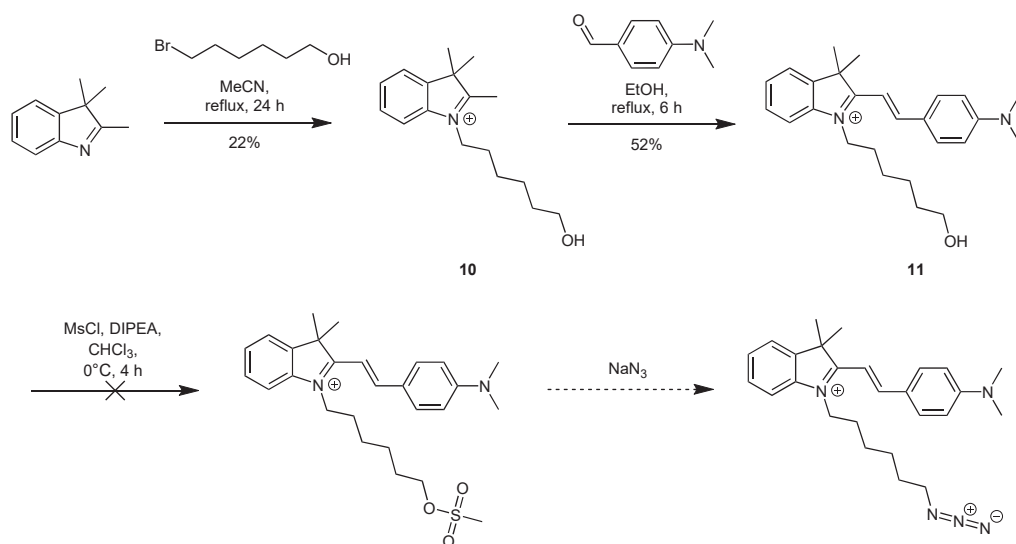
The first attempts at synthesising an azide-functionalised hemicyanine were performed by converting a tethered hydroxyl group to an azide *via* a mesylate intermediate. 2,3,3-Trimethylindolenine was alkylated at the imine position with 3-bromopropanol to yield **8**. The resulting indolinium was then subjected to a condensation with *N,N*-dimethylamino

benzaldehyde to form the hemicyanine **9**. After this, installation of the azide was attempted through a mesylate intermediate. However, this reaction resulted in formation of the cyclised product (Scheme 2.3). It is likely that activation of the hydroxyl group with the mesylate rendered it more likely to partake in nucleophilic attack of the iminium.



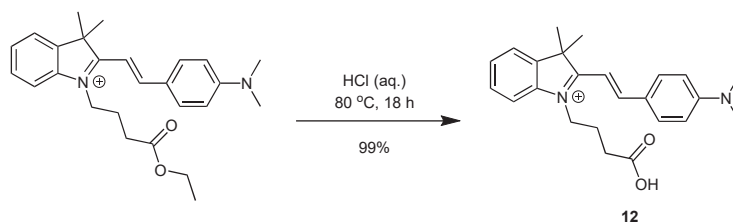
Scheme 2.3: Attempted synthesis of an azide functionalised hemicyanine using a propyl linker.

To reduce the likelihood of forming the cyclisation product, a longer linker was introduced between the indolenine and the hydroxyl group. A hexyl chain was chosen as the appropriate linker length since generating the cyclised product would involve generating a seven membered ring, which is highly unfavoured. With this, the synthesis was attempted according to Scheme 2.4. 2,3,3-Trimethylindolenine was alkylated at the imine position with 6-bromohexanol to yield **10**. The resulting indolinium was then subjected to a condensation with *N,N*-dimethylaminobenzaldehyde to form the hemicyanine, **11**. Mesylation of the alcohol was attempted, but the reaction was unsuccessful, resulting in multiple products that could not be isolated.



Scheme 2.4: Attempted synthesis of an azide functionalised hemicyanine using a hexane linker.

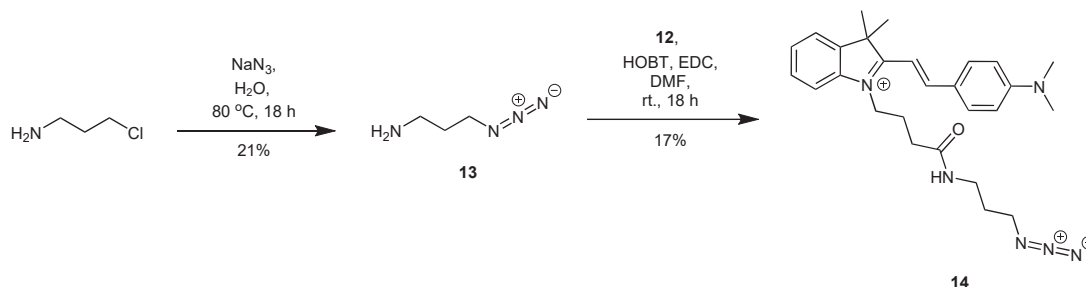
To avoid the issues encountered in attempting to install the azide directly on the hemicyanine, a convergent synthetic approach was pursued, where an acid-functionalised hemicyanine and an amine-functionalised alkyl azide were coupled *via* an amide bond. For this synthesis, an ethyl ester protected hemicyanine was provided by Dr. Kylie Yang. The ester was hydrolysed under acidic conditions to yield **12** (Scheme 2.5).



Scheme 2.5: Ester hydrolysis of the provided hemicyanine.

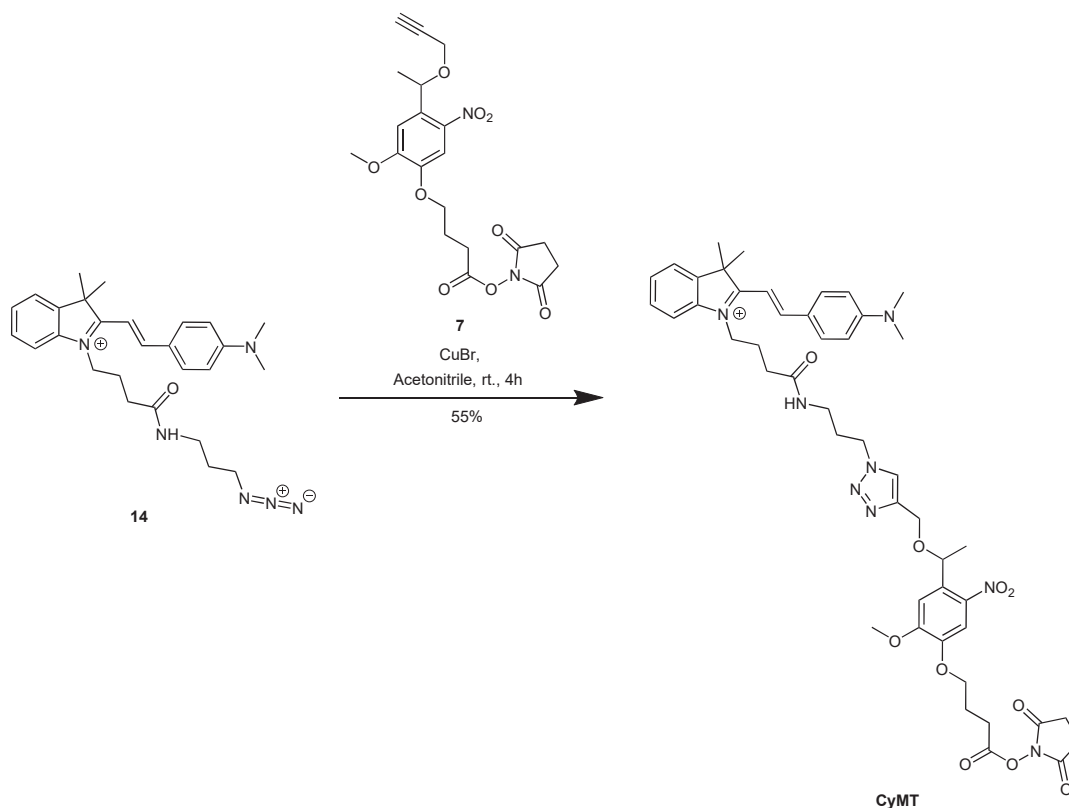
Separately, the amino azide was synthesised by halide displacement of 3-chloropropylamine with sodium azide to give **13** which was subsequently reacted with **12** using EDC and hydroxybenzotriazole (HOBt) as a catalyst to yield the desired product, **14**, as shown in Scheme 2.6. The lower than expected yield of this reaction were attributed to the use of an old batch of EDC.

SAFETY NOTE Both organic and inorganic azides are considered heat and shock sensitive and may decompose explosively with little external energy input. It is important to ensure that organic azides consisting of a C:N ratio ≤ 3 are not synthesised or stored in their pure form at quantities greater than 1 gram. Organic azides should not be distilled or sublimated, and purification should be limited to extraction or precipitation.



Scheme 2.6: Coupling of the amino azide to the acid functionalised hemicyanine.

At this point, the photo-cleavable linker, **7**, and the hemicyanine click partner, **14**, were subjected to a copper-catalysed cycloaddition. In general, click reactions are robust and not sensitive to water. As such, they can usually be performed in water mixtures to yield the click product in only a few hours.²⁰⁵ However, in this case there were two important factors that dissuaded the use of aqueous reaction conditions. Firstly, succinimidyl esters are prone to hydrolysis and this functional group is necessary for biomolecule conjugation. Secondly, the hemicyanine is inherently charged and, from experience, is very difficult to isolate from aqueous mixtures. For this reason, a water-free click reaction using copper(I) bromide in dry acetonitrile was pursued.²⁰⁶ The reaction successfully yielded the final probe, **CyMT** (Scheme 2.7).



Scheme 2.7: Synthesis of the final fluorescent mass-tag, **CyMT**

2.5 Fluorescence studies

2.5.1 Solution-phase fluorescence properties

In order to assess the compatibility of **CyMT** with fluorescence imaging, fluorescence spectra were examined. First, the excitation and emission spectra of the compound were measured in water to determine the compatibility of the probe with imaging media typically used for microscopy of biological samples. Any cleavage of the NHS ester in water should not affect fluorescent readings as it is not associated with the fluorescent core.

CyMT was dissolved in water and the emission and excitation spectra were measured. The probe exhibited an excitation and emission maximum of 558 and 591 nm, respectively (Figure 2.6). This result is consistent with what is expected of cyanine-based chromophores. As such, for fluorescence imaging studies that require imaging medium, the ideal microscope lasers would be either the 488 or 561 nm standard lasers.

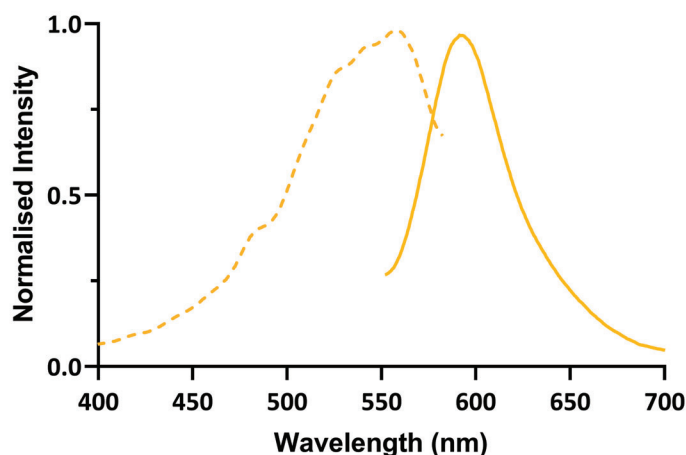


Figure 2.6: Normalised excitation (dashed) and emission (solid) spectra of **CyMT** taken in water ($\lambda_{ex} = 558$ nm, $\lambda_{em} = 591$ nm).

2.5.2 Solid-state fluorescence properties

Next, the compatibility of **CyMT** with medium-free imaging was assessed. Medium-free imaging eliminates the need for removal of imaging medium post fluorescence imaging, reducing the risk of sample disruption or contamination that can impact the accuracy of subsequent analysis. To assess this, the fluorescence properties of **CyMT** were measured in the solid state.

CyMT was dissolved in a small amount of dichloromethane, then smeared across decontaminated glass slides and left to dry before the spectra were taken. The absorbance spectrum reveals two peaks at 535 and 572 nm. However, the fluorescence spectrum shows a very broadened unresolved peak at around 620 nm (Figure 2.7). The broad, low-level fluorescence peak may be suggestive of aggregation-caused quenching. This phenomenon is commonly seen in highly conjugated, planar molecules. Aggregation caused quenching can occur from a variety of mechanisms include but are not limited to dexter energy transfer, Förster resonance energy transfer, or π - π stacking.²⁰⁷ Given the heavily aromatic and potentially planar structure of **CyMT**, this would be an expected possibility.

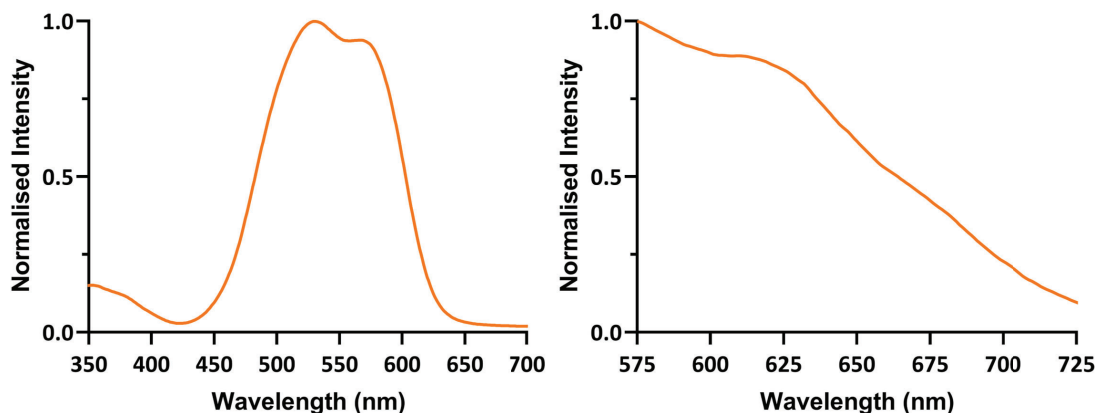


Figure 2.7: Normalised absorbance (left) and emission (right) spectra of **CyMT** taken in the solid-state ($\lambda_{abs} = 535, 572$ nm, $\lambda_{em} \approx 620$ nm).

To test this, a less concentrated sample of **CyMT** was prepared in the same manner described above. In this case, the fluorescence spectrum reveals a slightly more resolved emission peak at 624 nm. This does indeed support the speculation of aggregation-caused quenching, as decreasing concentration resulted in increasing fluorescence output. Interestingly, the absorbance spectrum also reveals two peaks at 535 and 572 nm. However, there is also a change in the relative absorbance intensities of each peak (Figure 2.8). This finding further supports the probability of aggregation occurring. It is well known that aggregation can cause a shift in the absorbance spectrum due to exciton coupling.^{208,209} That is, when there are interactions in the electronics of neighbouring molecules within the aggregate. Depending on the nature and direction of this coupling, the absorbance spectrum can shift bathochromically or hypsochromically. J-aggregates are characterised by the interaction of perpendicular transition dipole moments which leads to a blue-shift in the spectrum, while H-aggregates are characterised by the interaction of parallel transition dipole moments resulting in a red-shift in the absorbance spectrum.²¹⁰ It has been reported that cyanine-based molecules are capable of forming both J- and H-aggregates.^{211,212} From this data, it is not feasible to determine whether the secondary peak arises from J-aggregation or H-aggregation, as it is not possible to conclude which of the two peaks corresponds to the monomer absorbance. This cannot be elucidated from the solution-phase absorbance either, as the spectroscopic properties and behaviour of molecules in the solid state can differ

significantly from those in solution. To do so, further studies on the solid-state monomer absorbance spectrum would be needed, but this was beyond the scope of this project.

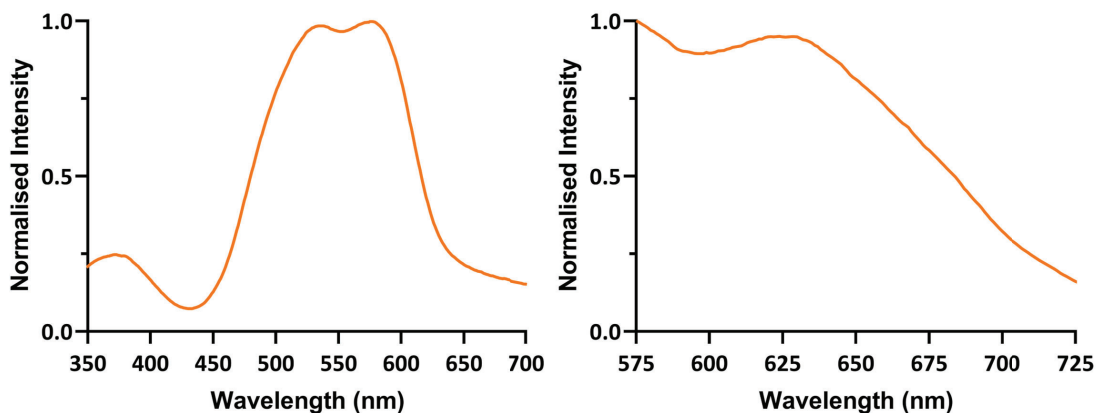


Figure 2.8: Normalised absorbance (left) and emission (right) spectra of less concentrated **CyMT** taken in the solid-state ($\lambda_{abs} = 535, 572$ nm, $\lambda_{em} = 624$ nm).

Regardless of the aggregate type, further testing of whether the low-intensity emission profile in the solid-state was inherent to the hemicyanine chromophore, or **CyMT** as a whole structure was ensued. To do this, the hemicyanine precursor provided by Dr. Kylie Yang, was selected to test the solid-state emission profile of the hemicyanine chromophore (Figure 2.9).

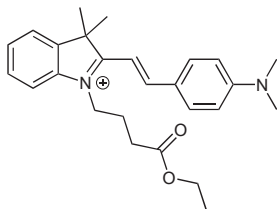


Figure 2.9: Structure of the hemicyanine used for comparison of solid-state fluorescence properties.

The sample was prepared in the same manner as above and the solid-state spectra were taken. The absorbance spectrum displayed a broadened double peak which may also be a result of aggregation, as explained above. However, in contrast to **CyMT**, the fluorescence spectrum of the hemicyanine chromophore shows a well resolved emission peak at 641 nm (Figure 2.10). This suggests that the photo-cleavable linker may be contributing to the

aggregation-caused quenching, leading to a low-intensity solid-state emission profile.

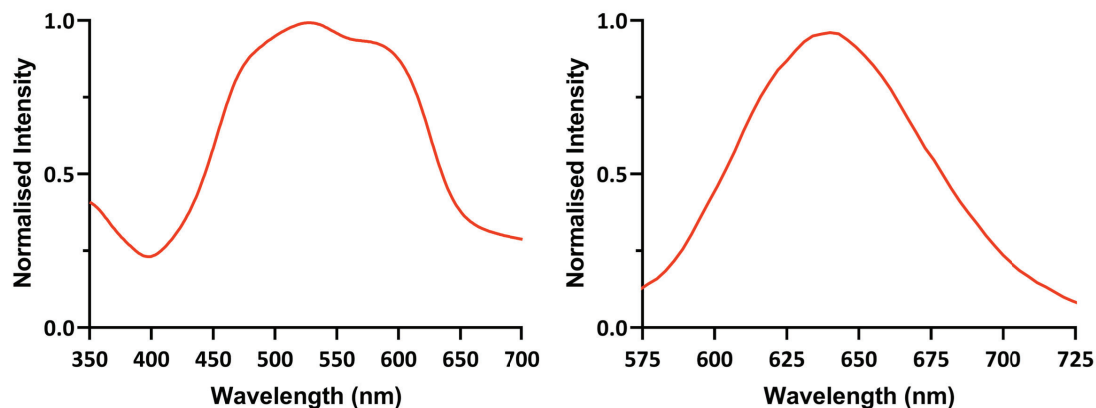


Figure 2.10: Normalised absorbance (left) and emission (right) spectra of hemicyanine chromophore taken in the solid-state ($\lambda_{abs} = 527, 566$ nm, $\lambda_{em} = 641$ nm).

In the case of solid-state fluorescence imaging of the tagged biomolecules, the distribution of **CyMT** will be limited by the number of free amines available on the biomolecule itself. On average, proteins contain a small number of primary amines, which comprise of the N-terminus and any lysine residues. As such, the distribution of amine-bound probes would likely be quite dispersed, which will inhibit aggregate formation. To emulate this, **CyMT** was dispersed into a matrix of Poly(methyl methacrylate) at 10, 5, 2.5, 1.25 and, 0.625% (w/v). This polymer was chosen due to its high refractive index and transparent optical properties.²¹³ The dispersed mixture was spin-coated to form a thin layer to minimise the formation of aggregates while drying. Spin-coating is a technique used to create thin, uniform film of material on a surface, which in this case was a glass slide. The surface is rotated at high speeds and a droplet of material is placed in the centre. This spreads the liquid uniformly across the surface through centrifugal force until the film is fully dried. One spin-coated, fluorescence spectra were taken (Figure 2.11). The spectra show that at <2.5% (w/v) aggregation caused quenching is alleviated. This shows that **CyMT** exhibits an emission maximum of about 680 nm in the solid state.

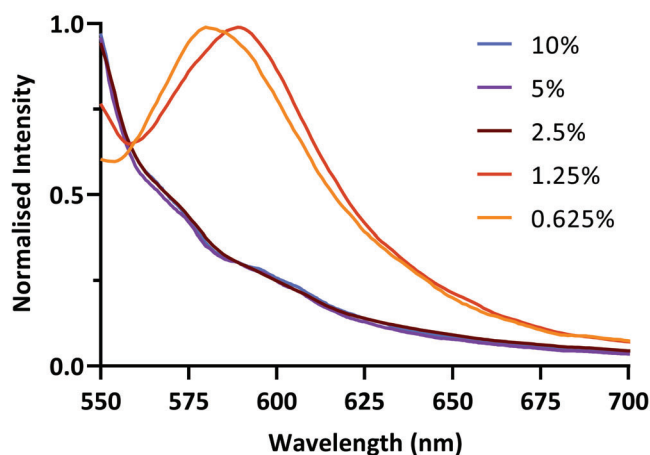


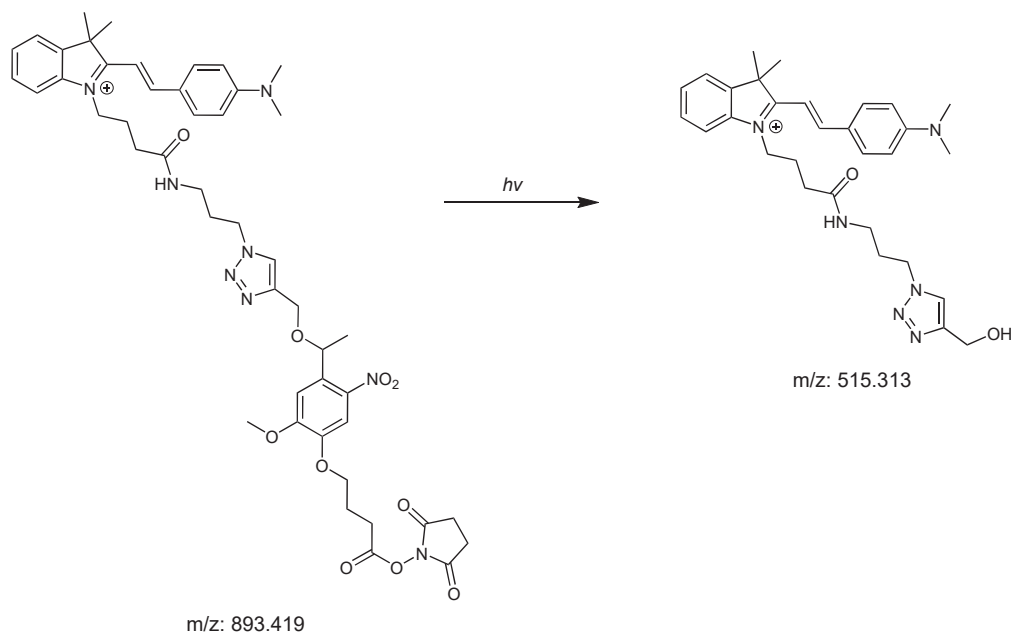
Figure 2.11: Emission spectra of **CyMT** dispersed in PMMA at 10, 5, 2.5, 1.25, and 0.625% taken in the solid-state ($\lambda_{ex} = 530$ nm).

From these studies, it is evident that **CyMT** exhibits aggregation-caused quenching. However, given the low number of free amines in biomolecules and the possibility of alleviating this quenching through a more dispersed arrangement, **CyMT** was still considered suitable for medium-free imaging tests.

2.6 MALDI compatibility studies

2.6.1 MALDI low-resolution studies

Having uncovered the suitable fluorescent properties of **CyMT**, attention was turned to assessing its compatibility with MALDI-MSI. Given the mechanism of cleavage explained in section 2.3.2, the expected mass to charge (m/z) ratio of the photo-cleaved fluorescent reporter should be 515.313 Da, with the intact molecule having a m/z of 893.419 Da (Scheme 2.8).



Scheme 2.8: Expected m/z of **CyMT** and its photo-cleaved adduct.

To ensure the compatibility of **CyMT** with MALDI, preliminary tests were carried out using a low-resolution instrument, with the assistance of Dr Leila Hill. This step was taken as it would allow for quick troubleshooting in case of any problems, before proceeding with the high-resolution instrument. The product was first tested in α -cyano-4-hydroxycinnamic acid (4CCA) matrix. This matrix was chosen for its compatibility with positive-ion mode analysis due to the inherent positive charge on the fluorescent reporter. Additionally, 4CCA is commonly used to measure m/z ratios under 5000 Da, making it appropriate for the analysis of cleaved **CyMT**.

The product was dissolved in 4CCA-saturated TA30 and spotted onto a polished steel plate for MALDI-MS analysis. To avoid cleavage by ambient light, the plate was dried in the dark. Low-resolution MALDI-MS analysis showed the expected m/z of both the intact (893.5 m/z) and cleaved (515.4 m/z) molecules (Figure 2.12). This confirms that **CyMT** cleaves in the expected manner under general MALDI-MS conditions.

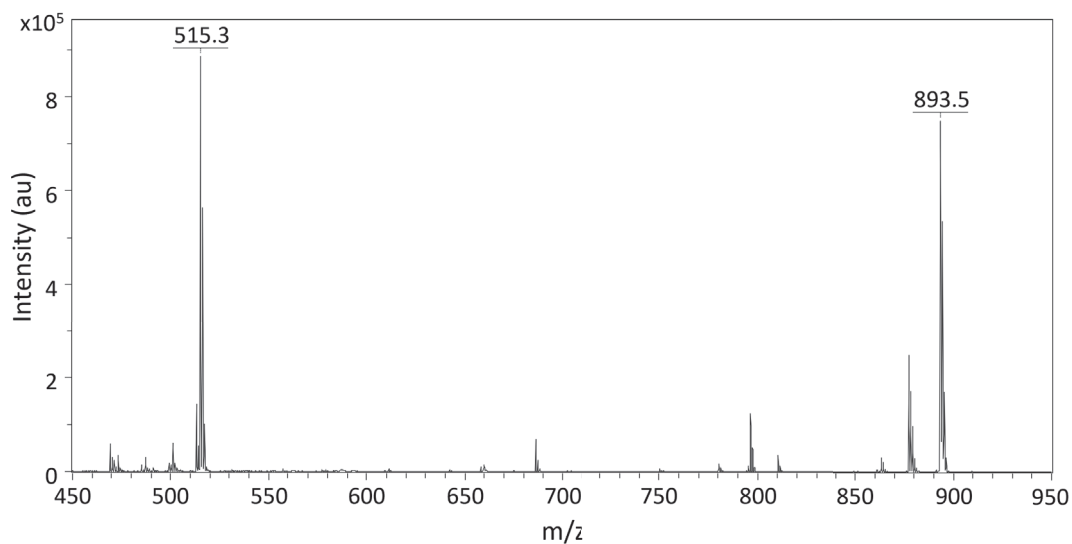


Figure 2.12: Low resolution MALDI-MS spectrum of **CyMT** dried onto polished steel in 4CCA matrix.

Next, the compatibility of **CyMT** with matrix-free mass spectrometry was examined. Matrix-free laser desorption ionisation mass spectrometry (LDI-MS) was designed as a potential means of reducing the interference of low-molecular weight compounds, but also minimising the extra steps required for matrix deposition prior to MALDI imaging, in general. As such, **CyMT** was dissolved in TA30 with no matrix, then spotted onto a polished steel plate. The mass spectrum reveals the mass of the photo-cleaved product (515.4 m/z) but no presence of the intact molecule. This is expected, given that the matrix absorbs a lot of the energy from the laser, resulting in milder conditions.

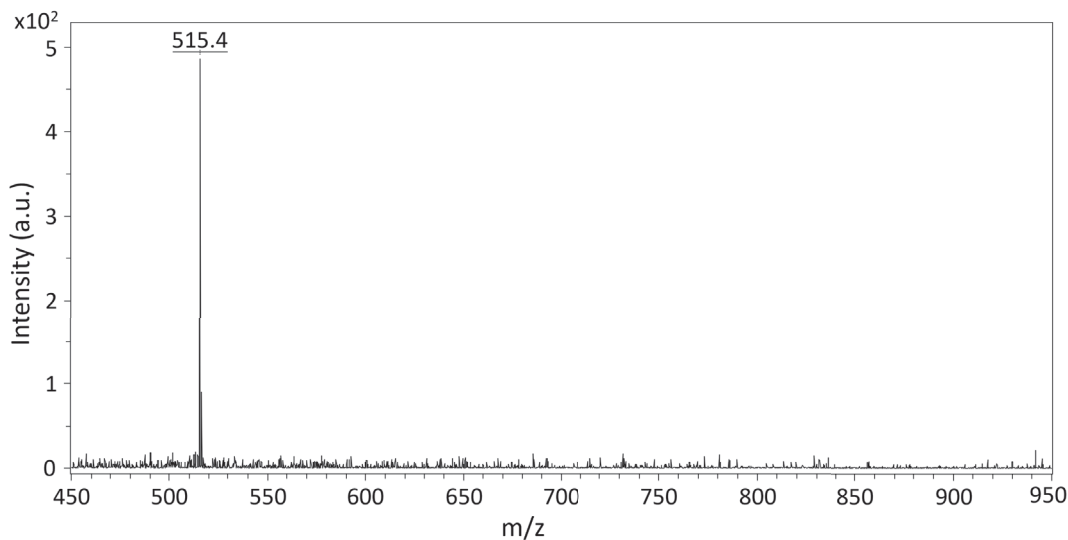


Figure 2.13: MALDI-MS spectrum of **CyMT** on polished steel with no matrix

Therefore, the results demonstrate that **CyMT** can be used in both matrix-assisted and matrix-free imaging, highlighting its utility and versatility for MS studies.

2.6.2 Matrix sublimation compatibility

In the case of matrix-assisted imaging, there are various techniques that can be used for matrix deposition onto a sample before MALDI experiments, with one of the most common being sublimation. This technique takes advantage of the fact that matrix compounds can sublime at high temperature and low pressure conditions, without decomposing. The matrix sublimation set-up involves an evacuated glass chamber containing a cooling chamber to which the glass-slide-mounted sample is secured. Below this is the heating element, onto which the matrix is placed in a petri dish. Upon heating, the matrix sublimates and deposits onto the cooled sample as a thin layer (Figure 2.14). The temperature and time for the sublimation process depend on the matrix being used.

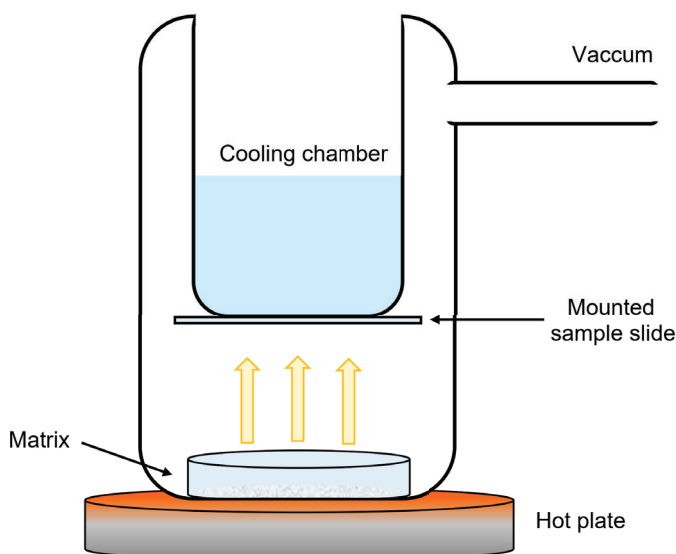


Figure 2.14: Visual representation of the matrix deposition process.

The stability of **CyMT** under the matrix sublimation process was investigated. To do this, a small amount of **CyMT** was dissolved in dichloromethane. The solution was smeared across an indium tin oxide coated glass slide and left to dry. The slide was then subjected to sublimation of 4CCA matrix at 140 - 170 °C. A small selection of this product was subjected to the MALDI imaging process as shown by the dashed border (Figure 2.15A). Two m/z filters were applied at 893.49 and 515.31 to visualise both the intact, and photo-cleaved probe. The 515.31 m/z filter shows a diffuse pattern with areas of higher intensities that generally correlate to the smeared plate (Figure 2.15A). This is similar to what is seen with the intact probe as visualised with the 893.49 m/z filter (Figure 2.15C) which is further seen in the overlay of both filters (Figure 2.15D).

A point spectrum from the image was extracted to show the full range of m/z present in the sample (Figure 2.16). The main m/z in the spectrum correspond to **CyMT** and its photo-cleaved product with little to no contamination of any other degradation products. These results show that **CyMT** is stable under matrix sublimation conditions and continues to cleave in the expected manner.

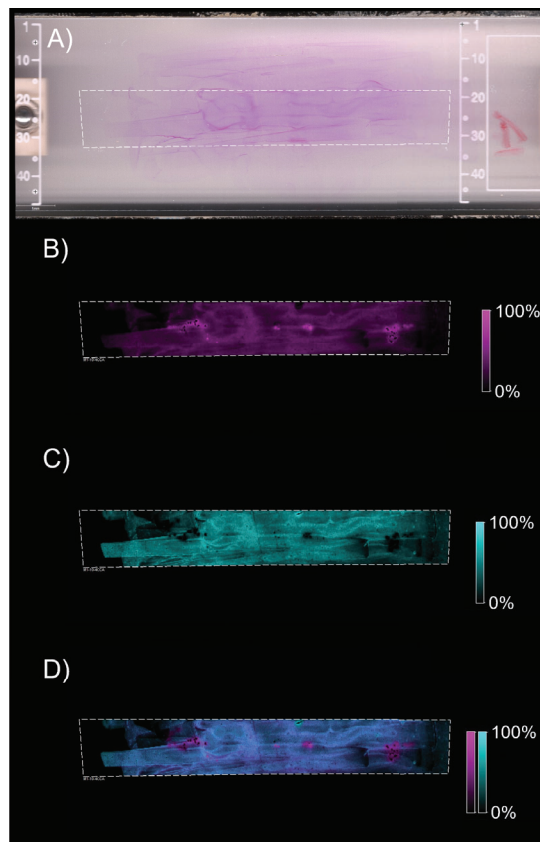


Figure 2.15: Visual image of A) the indium tin oxide coated glass slide smeared with **CyMT**, and MALDI-MSI images of the slide with a B) 515.31 (pink), C) 893.49 (blue), and D) an overlay of the two m/z spectral filters.

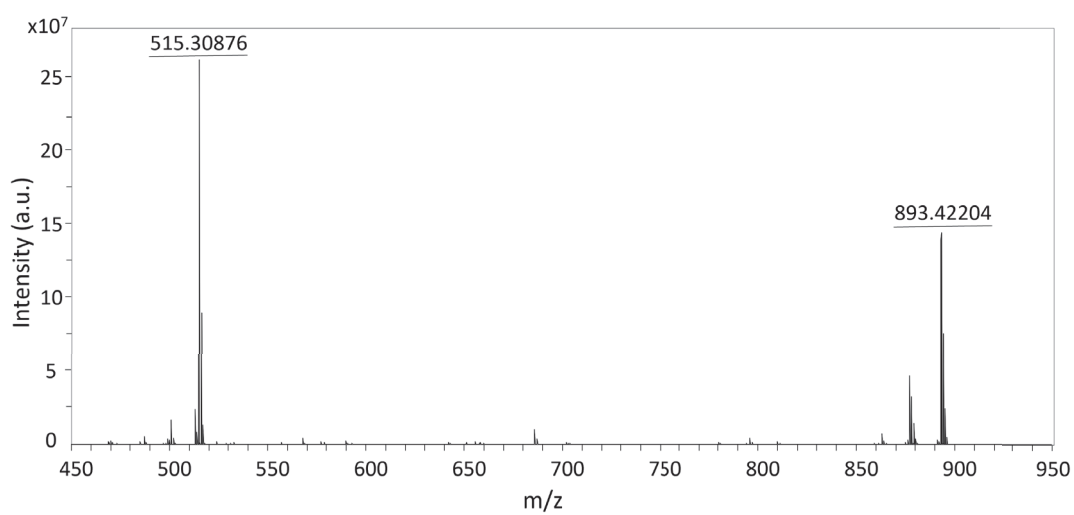


Figure 2.16: Point mass spectrum taken at an arbitrary point of the MSI segment.

2.7 Design and implementation of a dual imaging protocol

2.7.1 Dual imaging protocol considerations

Having demonstrated the fluorescence and MALDI compatible properties of **CyMT**, a protocol for the dual fluorescence and MALDI imaging on a single sample was conceptualised. Since MALDI imaging has optional sample preparation in the form of matrix deposition, it was determined that the order of techniques should begin with fluorescence imaging. Additionally with MALDI-MSI, the fluorescent reporter is released from the biomolecule with laser irradiation which would involve the risk of migration if performed before fluorescent imaging. This should not be a risk with laser irradiation from the microscope lasers, as the neither the 561 nor 488 nm standard microscopy lasers lie within the acceptable wavelength range that will induce cleavage of the ONB linker, by neither single-photon (310 or 340 nm) nor two-photon (710 or 750 nm) cleavage wavelengths.^{214–216}. Therefore, the protocol would proceed with sample mounting onto an ITO-coated glass slide, followed by fluorescence microscopy of the sample. After this, an optional step of matrix deposition onto the sample may be performed, before carrying out MALDI-MSI experiments (Figure 2.17).

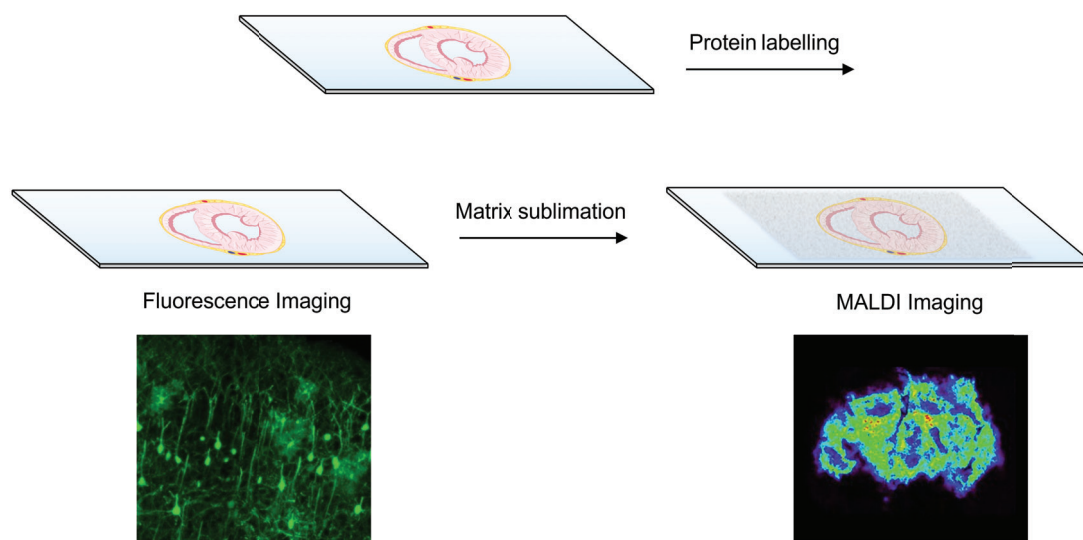


Figure 2.17: Representation of the envisioned dual fluorescence and MALDI imaging protocol. An ITO-slide mounted sample is labelled with the mass-tag followed by fluorescence imaging. After which the same sample is used in MSI.

2.7.2 Dual imaging protocol proof of concept

The viability of this protocol was tested using insulin as a biomolecule of interest. This protein is comprised of two chains, A and B, that are bound by disulfide bridges. There are three primary amines in monomeric insulin, namely the N-termini of each chain, and one lysine residue, which can serve as reactive sites for conjugation with **CyMT** Figure 2.18. Insulin monomers are capable of forming fibrils that exhibit many characteristics similar to many disease-forming amyloids.²¹⁷

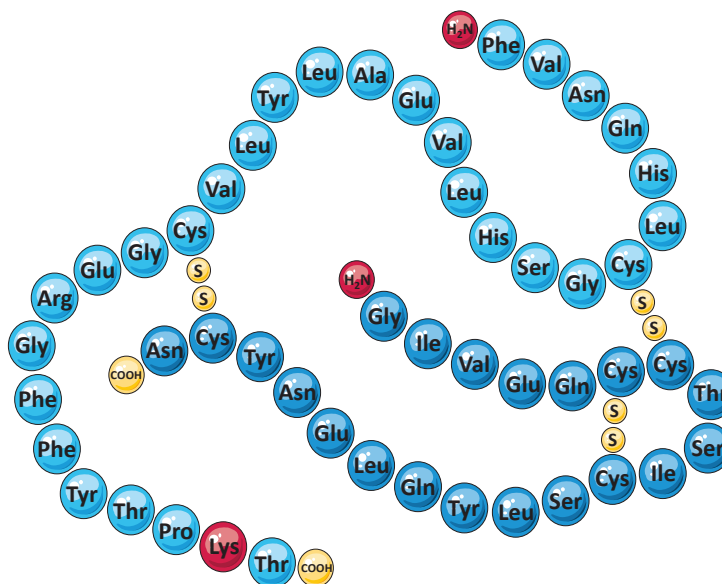


Figure 2.18: Amino acid sequence of the two-chain protein, human insulin, highlighting the free primary amines (red) available for conjugation with a succinimidyl ester.

Monomeric insulin was reacted with an excess of **CyMT** which was expected to tag one or more of the three primary amine sites. Coupling was confirmed by MALDI-MS. The mass spectrum shows un-tagged monomeric insulin ($5808.0\ m/z$), singly-tagged insulin ($6584.5\ m/z$) and its photo-cleaved ($6071.3\ m/z$) adduct, doubly-tagged insulin ($7361.1\ m/z$) and its singly- ($6848.2\ m/z$) and doubly-photo-cleaved ($6335.4\ m/z$) adducts. The spectrum did not reveal the mass of any triply-tagged insulin Figure 2.19.

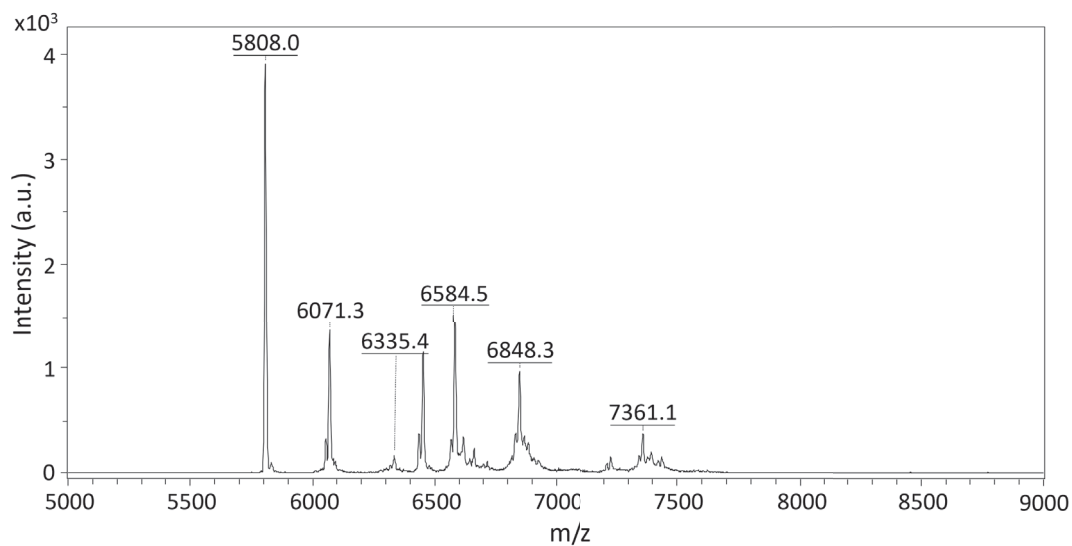


Figure 2.19: MALDI-MS of **CyMT**-tagged insulin showing untagged, singly-tagged, doubly-tagged protein and their respective photo-cleaved adducts.

The crude tagged insulin mixture was eluted from a size exclusion column to remove any remaining unbound probe. After this, tagged-insulin fibrils were formed by agitation under acidic conditions overnight. The fibrils were dried onto ITO-coated glass slides and subjected first to confocal fluorescence microscopy. Medium-free imaging was trialled.

Confocal images reveal a web-like morphology which is characteristic of insulin fibrils (Figure 2.20).²¹⁸ The fluorescent image demonstrates that biomolecular conjugation of **CyMT** diminishes the effect of aggregation induced quenching and confirms the feasibility of medium-free fluorescence imaging.

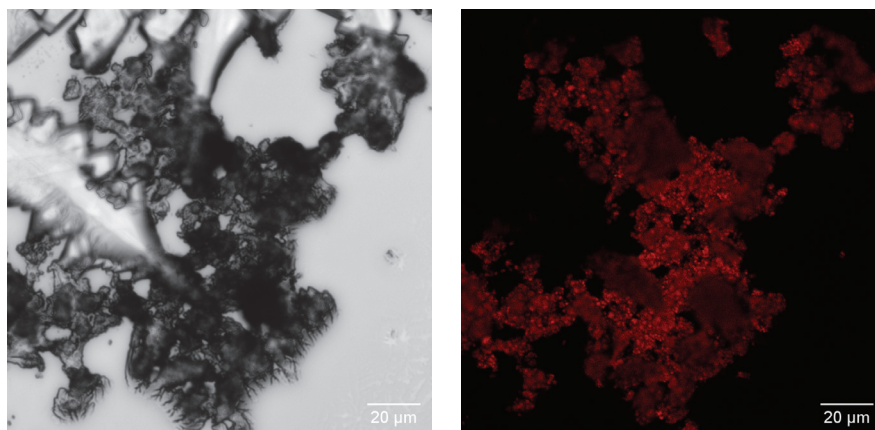


Figure 2.20: Representative images of bright field (left) and fluorescence (right) microscopy of **CyMT**-tagged insulin fibrils. Images were taken without the use of medium using a 40x objective (λ_{ex} 561 nm laser, λ_{em} 575 - 675 nm).

The benefit of medium-free imaging is that there is no need to remove or dry the medium before proceeding with the MALDI imaging protocol. This ensures there is minimal disruption to the sample in between techniques. As such, the dry slide was subjected to 4CCA matrix sublimation, followed by MALDI-MSI. A 515.31 m/z filter was applied to visualise the localisation of the cleaved reporter.

The MS images reveal a sparse distribution of reporter within the selected imaging area (Figure 2.21). This likely correlates to various clusters of insulin fibrils that formed during the slide-mounting process. This illustrates that protein-conjugated **CyMT** was capable of cleaving in the expected manner after being subjected to fluorescence microscopy.

This highlights the feasibility and effectiveness of the multimodal probe, **CyMT**, in the envisioned setup. The successful demonstration of **CyMT** using insulin as a biomolecule of interest provides a proof-of-concept for the probes potential applications in the multimodal imaging of other biomolecules.

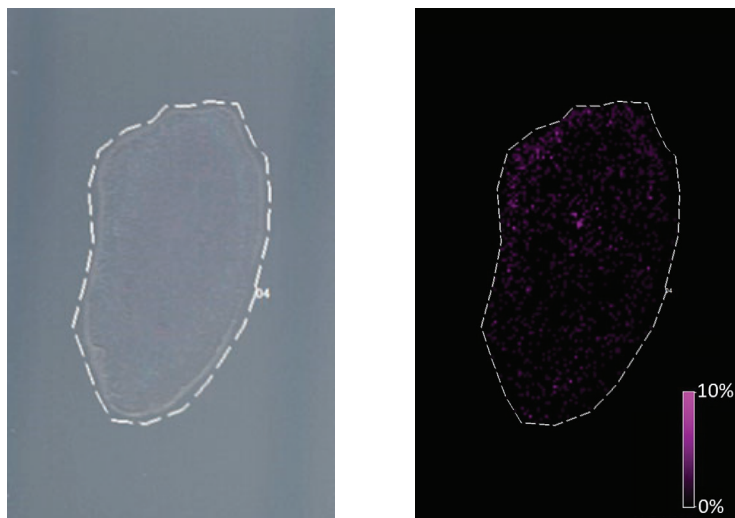


Figure 2.21: Visual image of the selected MSI region around the droplet of dried **CyMT**-tagged insulin fibrils (left) and MALDI-MSI of this droplet with a 515.31 m/z filter displayed in false colour based on intensity.

2.8 Planned tissue studies

Having demonstrated the successful implementation of **CyMT** in the multimodal imaging, the probe was sent for further testing in the Pukala Lab at the University of Adelaide. The intended studies aimed to demonstrate the conjugation of **CyMT** to an isolated protein using immunohistochemical techniques, allowing it to be assessed within a tissue sample for fluorescence imaging, followed by its assessment in matrix-free MSI. Unfortunately, the turbo pump on the MSI instrument in their lab broke down, and could not be repaired in time to acquire these images before the submission of this thesis.

2.9 Design and synthesis of NpMT

Given the successful implementation of **CyMT** in the proof of concept, a second probe was envisioned. To achieve this, a naphthalimide-based probe was considered. However, naphthalimides are not inherently charged and not very water-soluble. To address this issue, a water soluble peptide with good ionisation potential was selected to be conjugated to the naphthalimide dye. The amino acid sequence Ser-Tyr-Ser-Tyr was chosen for this

purpose. Additionally, by targeting a different region of the mass spectrum, this could allow this mass-tag to be used in multiplexed studies with **CyMT** to assess multiple proteins of interest within the same sample. Naphthalimide cores also contain at least two positions available for functionalisation.¹⁹ This would allow the photo-cleavable linker to be attached in addition to the peptide.

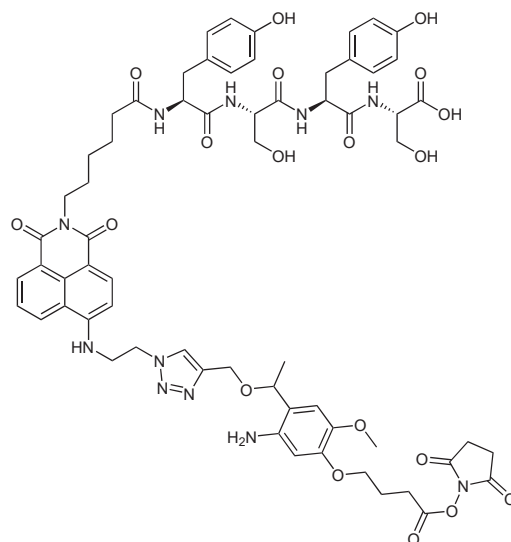
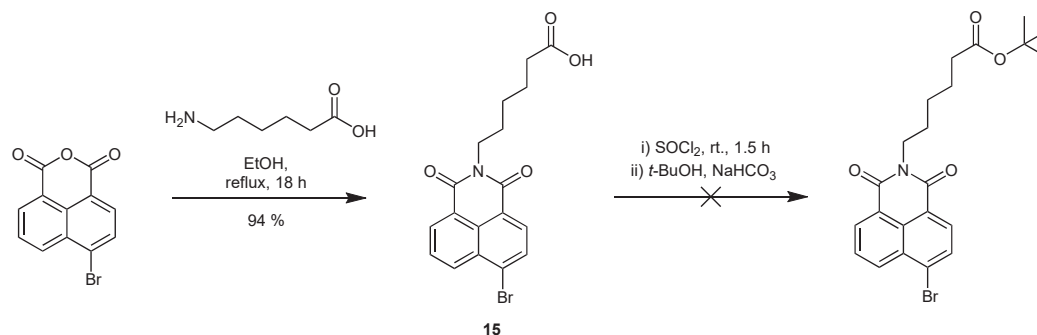


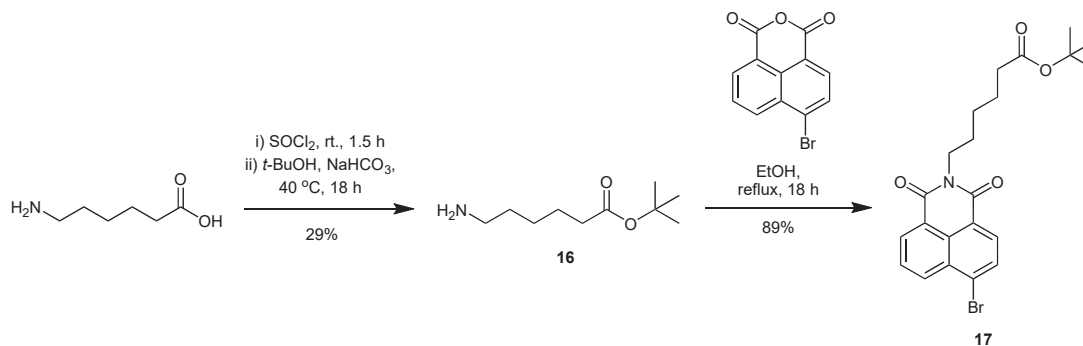
Figure 2.22: Design of the fluorescent mass-tag based on a naphthalimide core with a water soluble peptide to increase the mass, solubility, and ionisation potential.

The synthesis of the naphthalimide-based probe began with the protection of a carboxylic acid functionalised naphthalimide core, as illustrated in Scheme 2.9. 4-Bromo-1,8-naphthalic anhydride was substituted at the imide position with 6-aminohexanoic acid, yielding **15**. The next step involved protecting the free carboxylic acid with a *tert*-butyl ester using thionyl chloride followed by *tert*-butyl alcohol. Unfortunately, this reaction was unsuccessful, and the desired product could not be isolated.



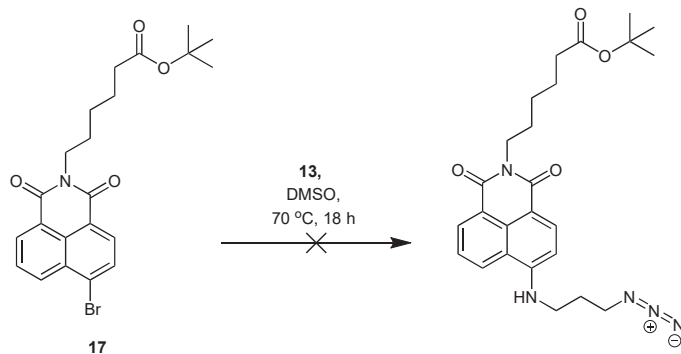
Scheme 2.9: Synthetic scheme of protecting acid of existing naphthalimide

Therefore, protection of the carboxylic acid was performed before substitution onto the naphthalimide core Scheme 2.10. First, 6-aminohexanoic acid was subjected to thionyl chloride, followed by reaction with tert-butyl alcohol to yield the tert-butyl-protected acid, **16**. The lower yields of this reaction may have been attributed to the unprotected amine forming side-products. **16** was then used to substitute the anhydride position of 4-bromo-1,8-naphthalic anhydride, resulting in the formation of **17**. With the desired product in hand, the next step was to introduce an azide at the 4-position of the naphthalimide scaffold.



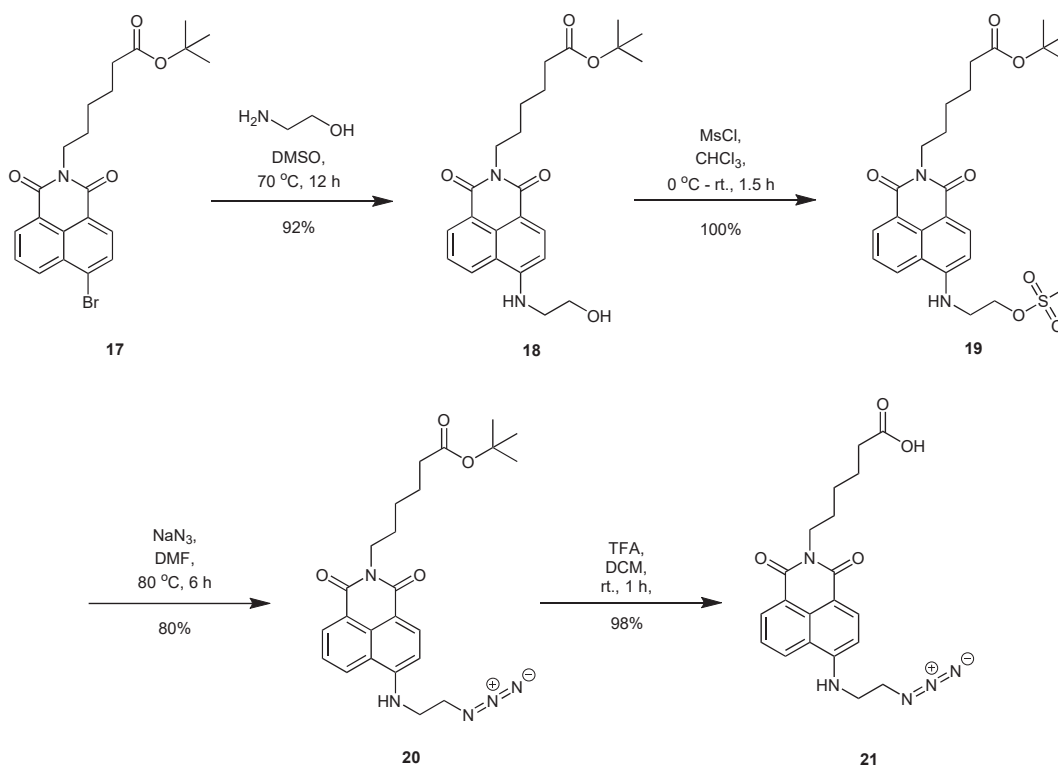
Scheme 2.10: Synthetic scheme of protecting before substituting

Initial attempts at this installation are given in Scheme 2.11. The previously synthesised aminopropyl azide, **13** was reacted with **17** in dimethylsulfoxide at 70 °C overnight. Unfortunately, the reaction was unsuccessful, and further troubleshooting was not possible due to safety concerns. **13** exceeded the C to N ratio for safe isolation and storage in larger amounts, so it was only synthesised in very small quantities immediately before use, making it inefficient for larger scale synthesis of the desired probe. Therefore, an alternative synthetic route was explored for the installation of the azide.



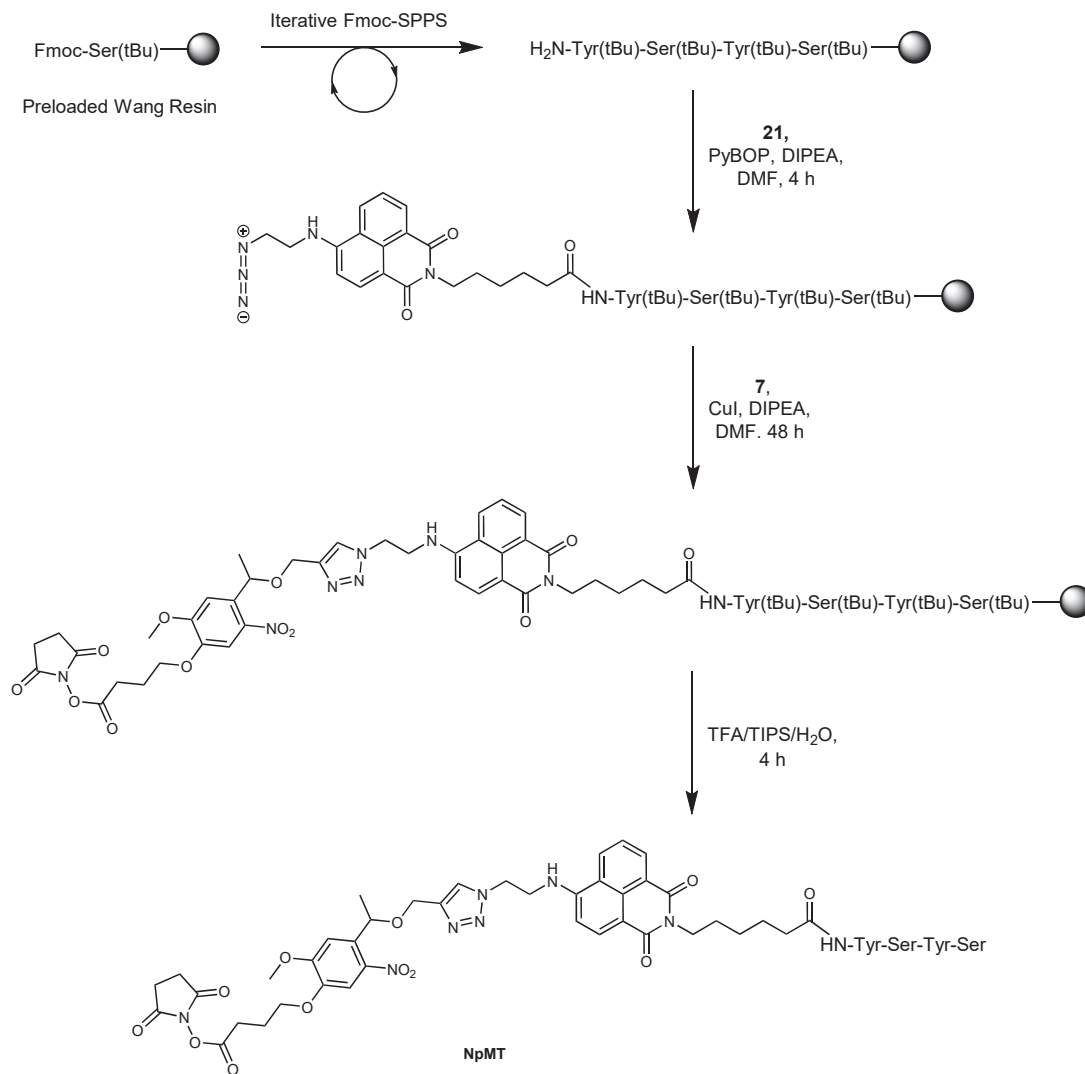
Scheme 2.11: Synthetic scheme of using amino azide directly on naphthalimide

The alternate route of choice is shown in Scheme 2.12. In this case, the azide was installed by conversion of a hydroxyl group attached to the naphthalimide core by an alkyl linker. To do this, ethanolamine was used to substitute of the 4-position of **17** to give **18**. The hydroxyl group was mesylated using mesyl chloride to give **19**, which was subsequently reacted with sodium azide to successfully yield the desired product, **20**. Finally, deprotection of the carboxylic acid was performed under acidic conditions to yield **21**.



Scheme 2.12: Synthetic scheme of forming azide *via* mesylate intermediate

With the desired naphthalide in hand, synthesis of the water soluble peptide was performed as shown in Scheme 2.13. Pre-loaded serine on Wang resin was used as the first amino acid. Peptide synthesis was performed according to standard protocols following an iterative method of Fmoc deprotection followed by amino acid coupling. The peptide was synthesised with the following sequence: H₂N-Tyr-Ser-Tyr-Ser. The carboxylic acid bearing probe, **21**, was then coupled onto the N-terminus using PyBOP as a coupling reagent in the presence of *N,N*-diisopropylamine. Following this, an on resin click reaction was performed using copper iodide as a catalyst in the presence of base for 48 h. After completion, the final peptide was cleaved from the resin using a peptide cleavage cocktail containing TFA/H₂O/triisopropyl silane (17:1:1 v/v/v). This yielded the desired final product, **NpMT**. The short exposure time to water during this reaction is assumed to have minimal effect on the NHS ester. However, this should be assessed more thoroughly in future syntheses. The product was characterised by low resolution mass spectrometry and analytical HPLC but due to time constraints purification and isolation of this compound was not performed.



Scheme 2.13: solid phase peptide synthesis of **NpMT**.

2.10 Conclusions

This work demonstrates the design, synthesis and implementation of a novel multimodal small-molecule probe for fluorescence and mass spectrometry imaging. The probe, **CyMT**, was designed based on the charged hemicyanine fluorophore, as the inherent charge allows detection without the risk of low ionisation, as well as matrix free imaging. This was tether to the photo-cleavable nitrobenzyl group functionalised with a succinimide for easy biomolecule conjugation.

CyMT was shown to be compatible with fluorescence spectroscopy in both aqueous

and solid-state form. Additionally, probe demonstrated compatibility with both matrix-assisted and matrix-free laser desorption ionisation mass spectrometry. In both cases, the probe cleaved to release the fluorescent reporter in the expected manner. The stability of **CyMT** to the matrix sublimation process was also confirmed. A single-sample dual-imaging protocol was devised, and tested on insulin fibrils in preliminary studies. **CyMT** was successfully conjugated to one, or two primary amines on insulin monomers which were then converted into fibrils. Despite its propensity towards aggregation-caused quenching, medium-free confocal fluorescence microscopy was used to successfully image the tagged insulin fibrils. The same sample was subjected to matrix sublimation and MALDI-MSI which confirmed compatibility with the dual imaging protocol.

Further to this, a secondary fluorescent multimodal probe, **NpMT** was designed and synthesised as a potential probe for future testing. The naphthlaimide moiety was selected for its biological compatibility and ability to be dual functionalised. A short, water-soluble peptide sequence was tether to increase its mass, water solubility and ionisation potential. Due to time constraints, **NpMT** was not tested, but is presented as a probe for which future testing should be performed.

Chapter 3

Computational analysis of fluorophores to study the redox state of proteins *in vitro*

3.1 Overview

3.1.1 Redox sensing on the sub-cellular level

The redox state of a cell is tightly regulated to balance the levels of oxidants and antioxidants. The main form of oxidants in the cell are known as reactive oxygen species (ROS). These are highly reactive oxygen-containing molecules such as hypochlorous acid (HOCl); hydrogen peroxide (H_2O_2); nitroxyl radical ($\text{NO}\bullet$); hydroxyl radical ($\text{OH}\bullet$); peroxynitrite (ONOO^-); singlet oxygen ($^1\text{O}_2$); and superoxide ($\text{O}_2^{\bullet-}$), that can cause significant damage to cellular macromolecules such as lipids, proteins, and DNA.²¹⁹ Chronic increases in ROS have been shown to be responsible for the development of various diseases including depression, Alzheimer's disease, and diabetes.^{220–223} However, despite their potential to cause damage, ROS are vital signalling molecules that are involved in significant signalling processes.^{224–226} On the subcellular level, ROS are both produced and processed by the individual organelles – each with a unique redox microenvironment, and specialised mechanisms for maintaining redox homeostasis.²²⁷ As such, the ability to study ROS at the subcellular level is crucial for our understanding of cellular signalling and the development of various pathologies.

Traditional methods of studying ROS involve indirect, ‘finger-printing’ measurements of markers like DNA oxidation, and lipid peroxidation. For example, oxidative damage can be measured by monitoring oxidised DNA adducts like deoxyguanosine-malondialdehydes.²²⁸ However, such techniques are destructive in nature, which means transient changes in live cells cannot be monitored. Additionally, the indirect nature of these measurements carries the potential for external artefacts to affect measurements.

In recent years, many have turned to the use of small fluorescent molecules to study the cellular environment. In the case of ROS, fluorescent sensors can either be designed for a single ROS, or they can be general redox sensors, where the individual ROS is not discriminated.

Over the years, a number of fluorescent sensors have been designed to respond to changes in the redox environment of a cell. However, as stated above, each organelle within a cell houses a unique redox microenvironment. As such, the development of targeted redox probes is of great importance in studying these distinctive environments.

Various methods can be used to target ROS sensors to specific cellular organelles. In general, targeting groups can be synthetically appended to the desired sensor, which allows it to accumulate in the organelle of choice. Targeting groups may be small synthetic molecules, or short peptide sequences designed to enter the organelle of choice. For a comprehensive overview of organelle targeting strategies, the reader is directed to the following review.²²⁹

Targeted sensors, in combination with powerful imaging tools such as confocal microscopy and flow cytometry, have proven to be an invaluable tool for studying the redox environment in cells. The following review will aim to summarise existing organelle targeted ROS sensors and uncover potential for the development of new sensors.

Mitochondria

Mitochondrial targeting is mainly achieved using delocalised lipophilic cations (DLCs). The positive charge on these molecules are drawn to the negatively charged inner membrane of the mitochondria allowing them to selectively accumulate in this organelle.²³⁰

The most common small molecule targeting groups used to localise redox sensors to the mitochondria are the aliphatic triphenylphosphonium group, and the pyridinium moiety. These groups have been used extensively to target redox sensors to the mitochondria^{231–236}

Other inherently charged functional groups have been incorporated into redox-sensor scaffolds that have also allowed these molecules to inherently target the mitochondria. Such chemical groups include benzothiazolium, and oxonium moieties.^{237–240}

Due to their intrinsic charge and lipophilic nature, some fluorophore scaffolds can inherently localise in mitochondria. In recent years, the most common DLC-type fluorescent scaffolds have been used for mitochondrial targeted redox sensing including rhodamine-, and indolenine-based scaffolds.^{241–246}

Peptide-based targeting groups have also been incorporated into small molecule sensors as a means of targeting them to the mitochondria. For example, **Pep3-NP1** is a ratiometric H_2O_2 sensor designed to incorporate a peptide sequence that promotes binding at the minor DNA grooves (Figure 3.1). Mitochondrial accumulation enhances fluorescence due to the low polarity of the DNA grooves. Reaction with H_2O_2 inhibits quenching and enhances green fluorescence to give a ratiometric response.²⁴⁷

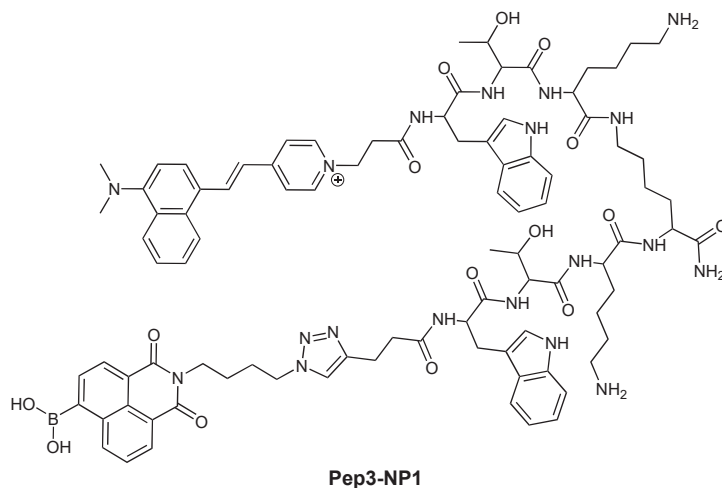


Figure 3.1: Chemical structure of **Pep3-NP1**.

Lysosomes

One of the most common methods of targeting small molecules to the lysosomes is by addition of weakly basic amine. Due to the acidic nature of the lysosome, the amine becomes protonated within the lumen of the lysosome. The positive charge inhibits the sensor from travelling back out of the membrane of the lysosome, resulting in accumulation, and retention of the probes.

The morpholine group is perhaps the most common lysosomal targeting group used to target ROS sensors to the lysosomes.^{248–252} However, there are examples of other basic amines that have been used for lysosomal targeting. For example Sun *et al.* developed a two-photon ratiometric sensor for lysosomal peroxynitrite based on a diethyl amine targeting moiety.²⁵³ Also, Wang *et al.* designed a pyridine-based sensor for hypochlorous acid sensing.²⁵⁴

Additionally, a unique selenamorpholine was used to incorporate a targeting group and a H₂O₂ sensor into a single moiety using a BODIPY core, **BODIPY-Se** (Figure 3.2). The change in the selenium’s electronic state upon contact with H₂O₂ induces a turn-on response of the BODIPY core.²⁵⁵ The basic amine of the selenamorpholine allows the sensor to accumulate within the lysosomes.

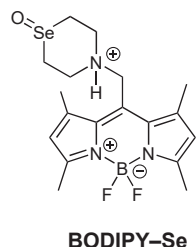


Figure 3.2: Chemical structure of **BODIPY-Se**.

Nucleus

Nuclear targeting is difficult to achieve due to the many cellular mechanisms that prevent mutagenic material from reaching nuclear DNA. As such, targeting the nucleus with fluorescent sensors is a difficult task to achieve. Most common molecular nuclear targeting groups comprise of crescent-shaped molecules that bind non-covalently in the minor grooves of DNA. For example, Hoechst dyes, that are themselves well-known nuclear stains, can be appended to small molecules to achieve DNA targeting. This tactic was utilised by Zhao *et al.* and Ikeda *et al.* to develop ratiometric, and reaction-based turn-on nitric oxide probes for the nucleus, respectively.^{256,257}

Protein tagging with the SNAP-Tag system has also been utilised to target redox sensors to the nucleus. This approach has enabled the development of a reaction-based H₂O₂ sensor, and nitric oxide sensor, which have both been successfully demonstrated in live cell

imaging.^{258,259} Additionally, nuclear targeted ratiometric H₂O₂ sensor has been designed using a nuclear localisation signal peptide.²⁶⁰

In an interesting example, Dickinson *et al.* designed a reaction-based hydrogen peroxide probe for the nucleus, **NucPE1** (Figure 3.3). The probe was used to successfully image nuclear hydrogen peroxide in HEK 293 cells and *C. elegans*. Intriguingly, the probe contained no nuclear targeting groups and the mechanism of nuclear targeting remains unclear.²⁶¹

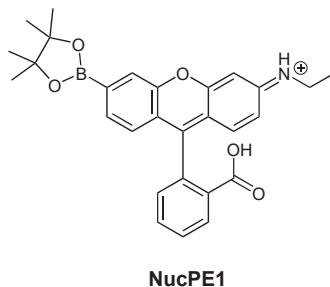


Figure 3.3: Chemical structure of **NucPE1**.

Peroxisomes

In general, short peptide based targeting groups are used to target the peroxisome. Small molecule targets have yet to be utilised for targeting redox sensors to the peroxisome.

To the best of our knowledge, only two fluorescent redox sensors have been designed for the peroxisome to date, and both target the peroxisome by utilising a variation on the SKL peptide. Ding *et al.* designed a superoxide sensor for the peroxisome by conjugating a QSKL peptide sequence to the a turn-on redox sensor.²⁶² Also, Zhou *et al.* utilised a novel peptide, HLKPLQSK, tethered to a two-photon fluorescent probe to successfully image peroxynitrite in SMMC-7721 cells and mouse livers.²⁶³

Lipid droplets

In general, targeting the lipid droplets does not require a specific targeting moiety. Instead, the types of molecules that are useful for imaging lipid droplets are highly conjugates fluorescent sensors that are significantly more emissive in non-polar, hydrophobic environments. This means that such probes may not necessarily be exclusively localised in the lipid droplets, but rather, are not visible in the more polar environment of the cell.

For example, **SX-1**, is a reaction-based peroxynitrite sensor that localises in the lipid

droplets.²⁶⁴ Additionally, **TPAs-SCH3-2CN** (Figure 3.4) and **XHZ** were designed as a reaction-based and ratiometric hypochlorous sensors, respectively.^{265,266} Another probe, **TAN** (Figure 3.4), was designed as a reaction-based nitric acid sensor for live cells.²⁶⁷ An extension of this probe, **SUN-TAN**, incorporated a triethylene glycol chain to further enhance anchoring of the probe to the lipid droplets due to its amphiphilic nature.²⁶⁸

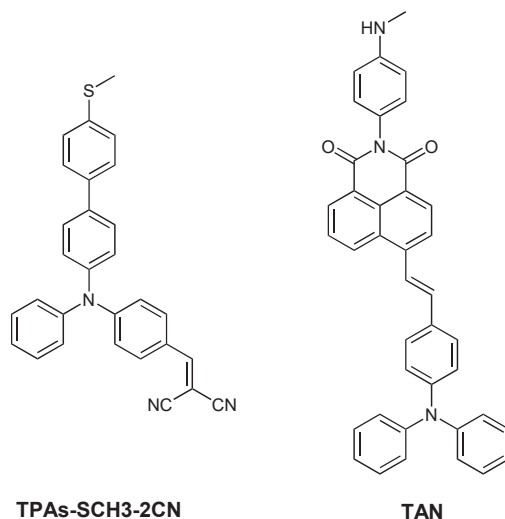


Figure 3.4: Chemical structure of **TPAs-SCH3-2CN** and **TAN**.

Cell membrane

The cell membrane is a semi-permeable lipid bilayer that regulates the movement of materials in and out of the cell. Most commonly, the cell membrane is targeted with fatty-acid or alkyl chain appendages that inset directly into the lipophilic centre of the lipid-bilayer. This anchors any appended fluorescent sensors to the cell membrane.

Ling He *et al.* developed a NO sensor with a structure mimicking the phospholipids of the cell membrane by appending an 18-carbon alkyl chain as the hydrophobic tail, and a quaternary ammonium group as the hydrophilic head, to the fluorescent core. The probe, **Mem-NO** (Figure 3.5), was successfully used to image NO in the membranes of HUVEC cells and neurons, as well as mouse brain tissue.²⁶⁹

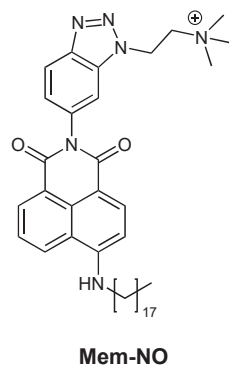


Figure 3.5: Chemical structure of **Mem-NO**.

Lipid membranes

Lipophilic probes that can cross the membrane may be able to tag internal lipid membranes in order to monitor internal membrane redox. For example, **H₂RB-C₁₈**, utilised an 18-carbon long alkyl chain to localised the ROS sensor to the cell membrane (Figure 3.6).²⁷⁰ **H₂RB-C₁₈** was used to image redox changes in pancreatic acinar cells induced by various compounds.

Another set of sensors utilise the structure of native fatty acids such as cis-parinaric acid and α -Tocopherol as the bases of membrane-targeted ROS sensors of lipid peroxidation – a consequence of ROS damage.²⁷¹

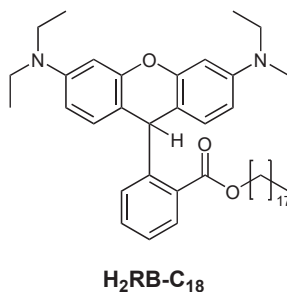


Figure 3.6: Chemical structure of **H₂RB-C₁₈**.

Phagosome

The SNAP-Tag system was utilised to target the phagosome, enabling imaging of phagosomal engulfment and the endogenous production of H₂O₂ during IgG ingestion in RAW264.7 cells. This was achieved through labeling a SNAP-Tag targeted hydrogen peroxide sensor

to SNAP-Tag tagged platelet-derived growth factor receptor (PDGFR).²⁷²

Zymosan particles have also been utilised as a targeting tool to send cargo directly to the phagosomes.²⁷³ These particles are inactivated fragments of *Saccharomyces cerevisiae* yeast cell walls that stimulates engulfment. Zymosan particles have been used to stimulate phagocytosis in cells pre-stained with a non-targeted ROS sensor to monitor redox changes during phagocytosis.^{274–277} However, Thekkan *et al.* designed a dual pH and HOCl fluorescent sensor covalently linked to zysoman particles by utilising a DNA-based linker.²⁷⁸ The probe was successful at imaging HOCl in the phagosomes of J774A.1 cells and macrophages.

Golgi apparatus

In general, phenylsulfonamides and aminoquinolines are used to target small molecules to the Golgi apparatus binding within the active site of the COX-2 enzyme.

Many phenylsulfonamide-based redox sensors have been designed to measure various ROS within the golgi apparatus. These include **Gol-ONOO[−]** and **MG-ONOO** are ratiometric probes that were used to image ONOO[−] at the golgi apparatus in live cells.^{279,280} Probe **Gogli-NO**, also based on phenylsulfonamide, was used to measure the relationship between golgi NO and amyloid- β induced Alzheimer in live cell models.²⁸¹ Furthermore, **Np-Golgi**, a two-photon ratiometric sensor, was developed for imaging H₂O₂ within the Golgi.²⁸²

Additionally, **FC-ONOO**, an aminoquinoline-based sensor, was developed for imaging peroxynitrite within the golgi apparatus in live cells and isolated mouse lung tissue in sepsis-induced acute lung injury.²⁸³ Aminoquinoline-based glutathione sensor, **GT-GSH**, exhibits a ratiometric fluorescence response to GSH which further reacts to form toxic by-products to selectively kill cells exhibiting golgi related oxidative stress.²⁸⁴

L-cysteine has been used as a golgi-targeting strategy to anchor to cysteine rich protein residues within the golgi apparatus. This strategy was used in the development of a reversible superoxide sensor, **CCA**, based on a caffeic acid scaffold (Figure 3.7). The probe revealed a hepatic ischemia-reperfusion signaling pathway that was mediated by superoxide generated from the golgi apparatus in hepatic cell lines and mouse models.²⁸⁵

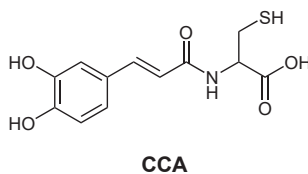


Figure 3.7: Chemical structure of **CCA**.

Endoplasmic reticulum

Small molecule targeting groups for the endoplasmic reticulum (ER) mainly includes sulfonylamides that bind tightly to ER-localised proteins. For example, ratiometric two-photon probe, **NAB**, was developed as a ONOO^- sensor for the endoplasmic reticulum using a sulfonylamide targeting group (Figure 3.8). The naphthalimide-based probe showed response to both endogenous and exogenous changes in ONOO^- . Additionally, by utilising **NAB** in depressive mice models, it revealed an increase in hippocampal ONOO^- .²⁸⁶

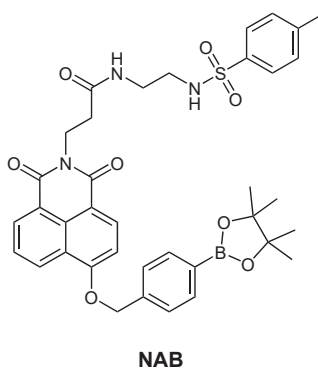


Figure 3.8: Chemical structure of **NAB**.

3.1.2 Protein tagging

The targeted redox probes discussed above have been an invaluable tool at revealing previously unknown redox mechanisms within the tightly regulated organelles. However, there are other locations inside and outside of the cell for which the redox state is significant. For example, the synaptic cleft is a unique extracellular environment in which redox mechanisms are suspected.²⁸⁷ Therefore, it would be of great value to develop redox sensors that are capable of being tagged to various proteins that would allow redox sensing to within structures that cannot be targeted with simple targeting groups. To do this, we turn to the

use of protein labelling systems. These systems are explored in detail in section 1.8.1. The advantage of using protein tagging systems is that there is a free choice of label which can be varied synthetically. Additionally, these systems can be used in a large range of cellular and even live animal studies. In this chapter, the HaloTag system will be utilised as the protein labelling system of choice.

3.2 Aims

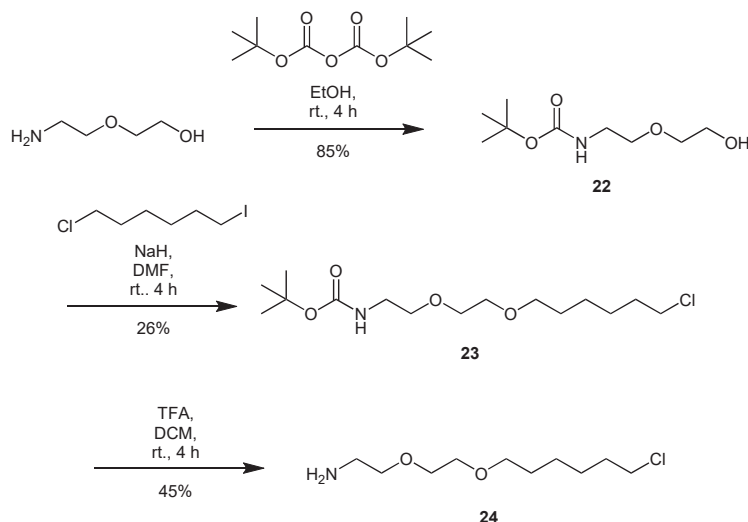
Excessive levels of ROS are known to cause damage to proteins, lipids and DNA in cells. However, ROS also play an important role as signaling molecules. Despite this, the specific ways in which ROS and cellular redox states contribute to various cellular processes are not yet well understood. Visualising redox changes in live cells is crucial for gaining insights into these processes. To date, this has typically been accomplished through the use of small molecule fluorescent sensors that respond to redox changes. These sensors can be used to measure the overall redox state of cells or can be targeted to specific cellular compartments. However, to fully understand the function of ROS in diseases and signaling pathways, the ability to measure the redox state at any sub-structure within the cell is crucial. Having a means of testing the redox state of any protein would be an invaluable tool to help further our understanding of the role of ROS in cellular processes.

To this end, this project was designed to combine the powerful techniques of protein-tagging and fluorescence sensing to develop novel protein-targeted redox sensors. As such, the aims of this chapter were to:

1. Develop novel redox-responsive fluorescent ligands for the HaloTag system.
2. Assess the molecular design of a ratiometric redox sensor using computational methods.
3. Perform computational benchmarking studies on a novel fluorescent redox scaffold to inform the design of future sensors.

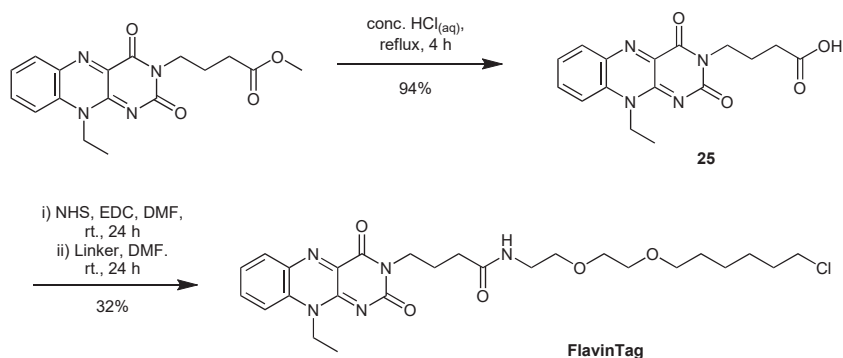
3.3 Previous work (Honours)

In previous work on this project, a novel fluorescent redox-responsive HaloTag substrate was designed, synthesised, and tested. To do this, the fluorescent sensor was tethered to the known HaloTag ligand that was synthesised as per Scheme 3.1.

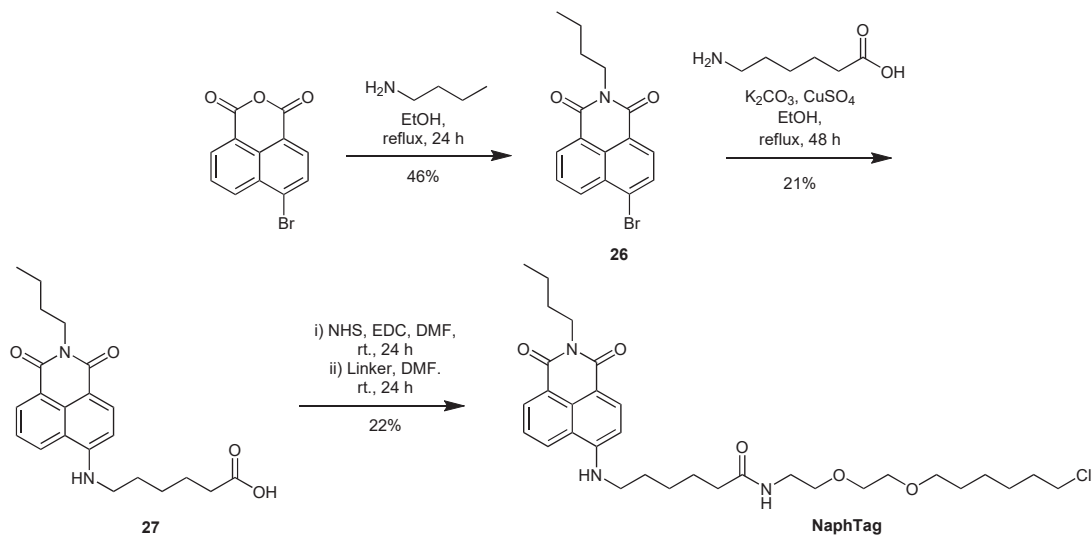


Scheme 3.1: Synthesis of the HaloTag substrate.

FlavinTag was designed to incorporate flavin, a redox sensitive fluorophore, *via* an amide bond onto the HaloTag substrate (Scheme 3.2). Further to this, a control substrate **NaphTag**, was synthesised in a similar manner (Scheme 3.2). **NaphTag** contains the naphthalimide fluorophore, but no redox-responsive group.



Scheme 3.2: Synthesis of **FlavinTag**.



Scheme 3.3: Synthesis of **NaphTag**.

FlavinTag was found to exhibit a reversible response to reducing and oxidising environments, and was stable to these changes over many cycles (Figure 3.9a). The reduction potential of this sensor lied within the biologically relevant range, making it suitable for testing in live cells (Figure 3.9b). Control substrate, **NaphTag** was found to have no fluorescent response to changing redox environments and was determined to be a suitable control (results not shown).

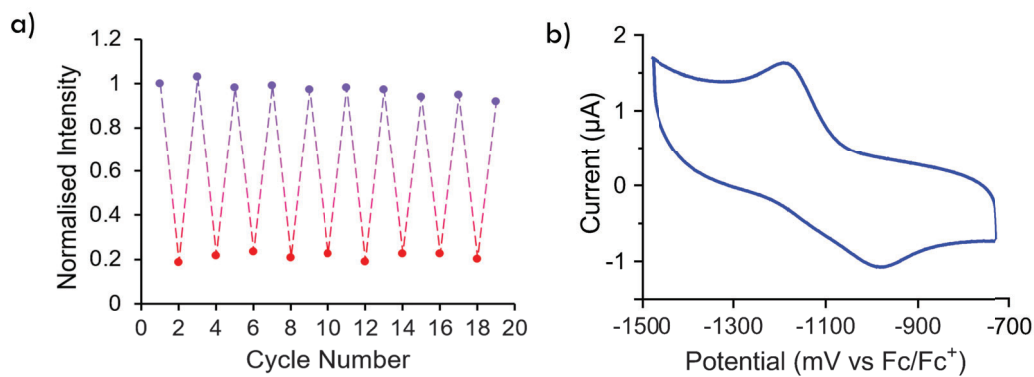


Figure 3.9: a) Normalised fluorescence response of **FlavinTag** to changes redox environment in HEPES buffer (20 mM). Reduction achieved using sodium ditionite (100 eq, red) and oxidation with hydrogen peroxide (100 eq, purple). Measurements were taken at $\lambda_{ex} = 443$ and $\lambda_{em} = 552$ nm. b) Cyclic voltammogram of **FlavinTag** (0.5 mM) in acetonitrile and tetrabutylammonium chloride (100 mM).

Copper transporter 1 (CTR1) was chosen as the protein of interest as the redox activity at this transporter was not known. Tagging this protein with **FlavinTag** revealed a previously unknown copper homeostasis mechanism. Under copper starvation, CTR1 was found to be in a reducing environment. On the other hand, when there was an excess of copper, CTR1 was found to be in an oxidising environment.

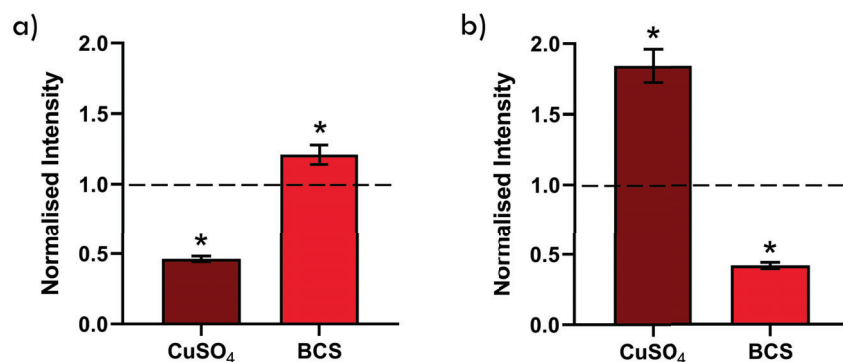


Figure 3.10: Fluorescence intensities of Halo-CTR1 DLD1 cells dosed for 20 min with a) NaphTag (0.5 μM), and b) **FlavinTag** (175 μM) after treatment with CuSO_4 (30 μM) or BCS (50 μM) for 24 h. Measurements were taken at $\lambda_{ex} = 405$, $\lambda_{em} = 500 - 600$ nm. Values are expressed as mean \pm SEM (n = 4-7), * = $p < 0.05$.

These results were validated with a non-targeted redox sensor, **NpFR1** that was designed to measure the overall redox state of the cell.²⁸⁸ Despite the redox changes occurring at the CTR1 protein, the global redox state of the cell was unchanging (results not shown here). This indicates that the redox changes are specific to the CTR1 protein, and not to the cell as a whole.

3.4 Current work

Having demonstrated the potential for **FlavinTag** to uncover previously undiscovered protein mechanisms, development of improved redox sensors was ensued. Two previously described redox sensors, flavin coumarin redox sensor 1 (**FCR1**) and naphthalimide flavin redox sensor 1 (**NpFR1**) were selected. Both of these fluorophores are derivatives of the flavin moiety, that improve on the properties of flavin alone. More information about the individual sensors are outlined in their respective sections below. The envisioned structures

of the HaloTag substrates, **FCR1-Tag** and **NpFR1-Tag** are shown in Figure 3.11. The design of these structures (i.e. the position and length of the alkyl chain binding the HaloTag substrate) was based on the developed acid functionalised **FCR1** and **NpFR1** that were synthesised by Dr Liam Adair. To assess the properties and suitability of these sensors, computational studies were employed.

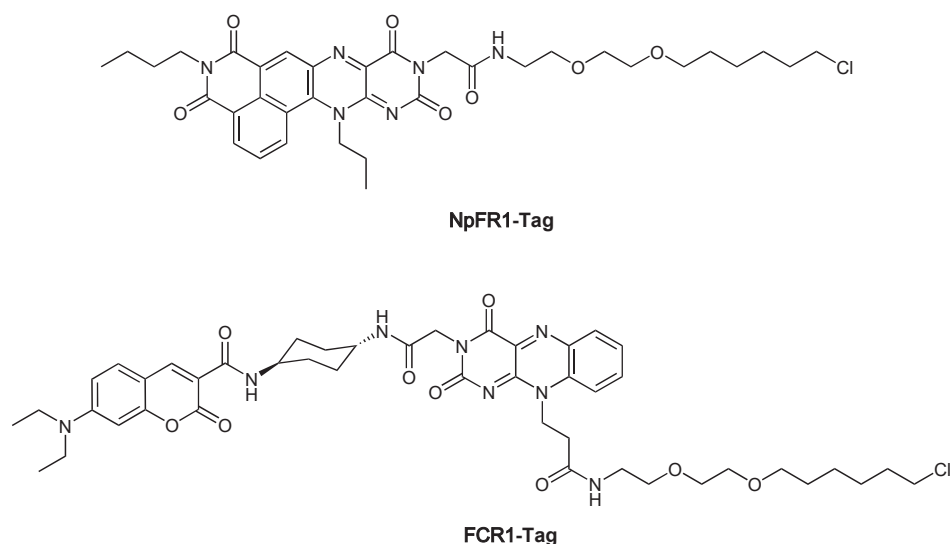


Figure 3.11: Proposed chemical structures of **NpFR1-Tag** and **FCR1-Tag**.

3.5 Determining solvation environment for calculations

Determining the appropriate dielectric constant for a given solvation environment is crucial to accurately model the properties of the redox sensors when bound to the HaloTag system. The dielectric constant of proteins is inconsistent depending on binding position and ranges from 4 to 80. However, when performing calculations on molecules within protein environments, dielectric constants ranging from 2 to 4 are typically used.²⁸⁹ It is important to note that the dielectric constant of the protein environment is not always constant and may vary depending on factors such as amino acid residue, and folding.²⁹⁰ For instance, regions of the protein that are more exposed to the environment may have a dielectric constant that is closer to that of pure water, while regions that are buried or surrounded by other residues may have a lower dielectric constant. Therefore, it is important to consider the relevant solvation environment when determining the appropriate dielectric constant for

computational studies involving sensors in a protein environment.

To begin the computational studies, the solvation environments of **NpFR1-Tag** and **FCR1-Tag** were first determined by performing molecular docking studies.

3.5.1 HaloTag docking

AutoDock4 was used to assess the docking conformation of **FCR1-Tag** and **NpFR1-Tag** when bound to the HaloTag protein.²⁹¹ The flexible side chain method was employed over the two-point attractor method due to its better performance accuracy, in the manner described by Bianco *et al.* In summary, protein crystal structure of ligand complexed HaloTag was obtained from the protein data bank (PDBID 6U32) was chosen based on its overall score. The protein structure was prepared by removing the complexed ligand and implicit water molecules, and adding hydrogen atoms. **FCR1-Tag** or **NpFR1-Tag** structure files were aligned within the binding pocket. AutoGrid4 software was used to generate calculation parameter files for the HaloTag-sensor complex, and AutoDock4 was used to optimise ligand binding.

The docking studies reveal that the bound sensors sit well outside of the enzyme binding pocket, and away from the protein surface (Figure 3.12). This indicates that the dielectric constant of water, 78, is an appropriate estimate for modelling the fluorescent properties of **FCR1-Tag** and **NpFR1-Tag**.

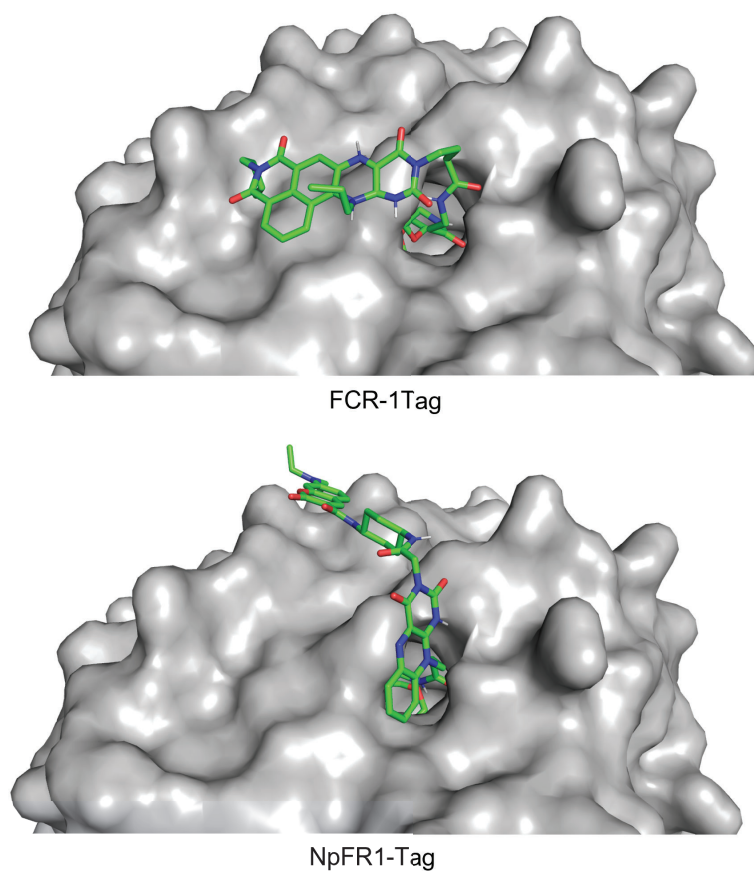


Figure 3.12: **FCR1-Tag** and **NpFR1-Tag** docked within the Halotag enzyme as determined using Autodock4.

3.6 FRET efficiency studies on FCR1

Förster resonance energy transfer

Förster resonance energy transfer (FRET) describes energy transfer between two fluorophores in a distance-dependent manner. This mechanism occurs when a donor fluorophore in an excited state, transfers energy to an acceptor fluorophore in the ground state. This induces excitation of the acceptor fluorophore, and non-radiative relaxation of the donor. As such, only emission of the acceptor is observed (Figure 3.13).

It is important to note that this is not a form of radiative transfer and thus, FRET is not the result of emission from the donor being absorbed by the acceptor. Rather, the process

of energy transfer occurs *via* through-space dipole-dipole intermolecular coupling. As such, the process of FRET can be thought of as an oscillating dipole that exchanges energy with another oscillating dipole of similar resonance frequency.^{292,293}

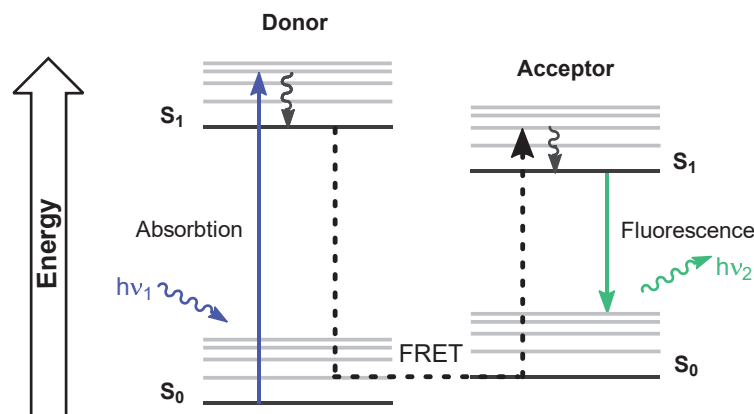


Figure 3.13: Jablonski diagram showing the process of FRET occurring as a non-radiative transfer of energy from the excited donor to the acceptor, resulting in emission from the acceptor alone, and return of the donor to the ground state.

FRET efficiency is determined by several factors including the distance between the donor and acceptor fluorophores, relative orientation of donor and acceptor transition dipole moments, and the extent of spectral overlap in the donor emission and acceptor absorption spectra (Figure 3.14).²⁹⁴

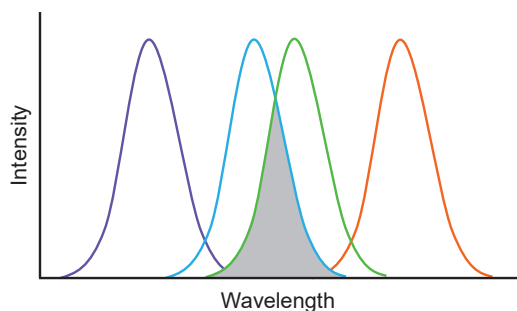


Figure 3.14: Simulated example of spectral overlap between two fluorophores. Donor absorbance and emission, and acceptor absorbance and emission are shown in purple, blue, green and orange respectively.

FCR1 was reported by Kaur *et al.* as a FRET based redox sensor.²⁹⁵ The probe is designed to contain a coumarin-based non-redox-responsive sensor in combination with

the redox-responsive flavin moiety. The two fluorophores are tethered by a cyclohexane linker allowing FRET to occur at the fixed distance (Figure 3.15). When exposed to a reducing environment, the flavin moiety undergoes a $2e^-$ reduction, causing it to adopt a bent conformation, and lose its fluorescence capabilities. In this state, the flavin moiety is no longer able to act as a FRET acceptor and fluorescence is emitted from the coumarin at a different wavelength. Upon reoxidation, the flavin returns to its fluorescent state and FRET is restored. This oxidation and reduction process can occur over multiple cycles, allowing **FCR1** to monitor transient changes in the redox environment. **FCR1** is a non-specific redox sensor and is capable of measuring the overall redox state of its environment.

The change in the ratio of fluorescence intensities between two emission wavelengths upon sensing changes is known as ratiometricity. Ratiometric sensors offer several advantages over intensity-based sensors. They have an internal reference for self-calibration because they use the ratio of two fluorescent signals. This reduces the impact of experimental variability such as sensor concentration. As such, ratiometric sensors offer a more reliable and accurate tool for measuring biological processes.

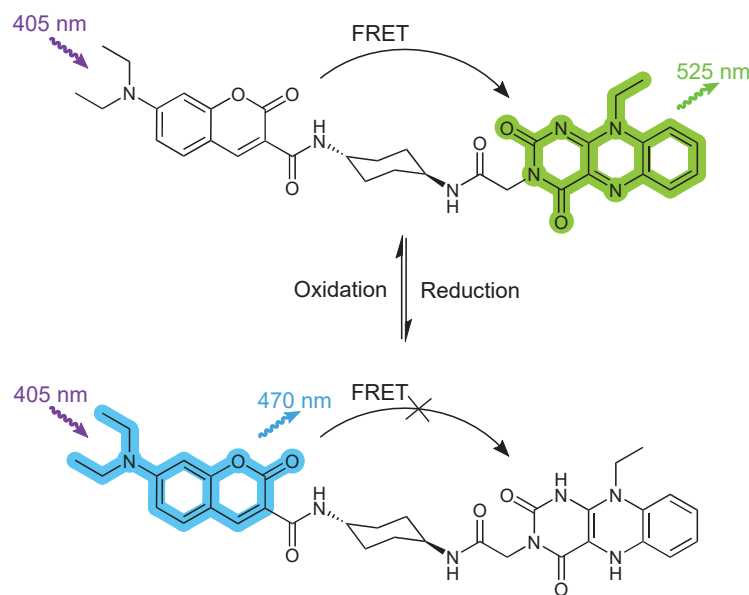


Figure 3.15: Chemical structure of **FCR1** in the oxidised and reduced forms highlighting the effect of FRET between the donor, coumarin, and the acceptor, flavin, and the effect on its fluorescence output.

The development of FRET-based probes is not straightforward, given the number of variables needed to ensure efficient energy transfer. In this section, these variables will be assessed for the **FCR1**, using a combination of experimental and quantum chemical methods. This will aid in understanding and potentially improving the structure of this FRET probe.

3.6.1 FRET efficiency calculations

FRET efficiency (E) is the probability of an energy transfer event occurring per donor excitation event and is defined by equation 3.1.

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} \quad (3.1)$$

where R is the distance between the centres of the donor and acceptor, and R_0 is the distance at which $E = 0.5$ (also known as the Förster radius). R_0 is defined by equation 3.2.

$$R_0 = 0.211[J(\lambda)\kappa^2n^{-4}\Phi_D]^{\frac{1}{6}} \quad (3.2)$$

where Φ_D is the quantum yield of the donor fluorophore and n is the refractive index of the medium. $J(\lambda)$ is the spectral overlap integral over all λ as given in equation 3.3.

$$J(\lambda) = \int_0^\infty I_D(\lambda)\epsilon_A(\lambda)\lambda^4d\lambda \quad (3.3)$$

Where I_D is the normalised emission spectrum of the donor, ϵ_A is the molar extinction spectrum of the acceptor, and λ is the wavelength. κ^2 is the orientation factor given in equation 3.4.

$$\kappa^2 = (\cos\theta_T - 3\cos\theta_D\cos\theta_A)^2 \quad (3.4)$$

where θ_A and θ_D are the angles of the acceptor's absorbance and donor's emission transition dipole moments vs distance axis, respectively and θ_T is the angle between these transition dipole moments.

3.6.2 Calculating the spectral overlap integral, J

To begin, the experimental spectral overlap of the donor, coumarin, and acceptor, flavin, was calculated. To do this, the absorbance and emission spectra of the independent fluorophores were measured in water. 7-Diethylamino coumarin and *N*-ethyl flavin were used for the measurement. The spectra show that there is indeed a good overlap between the emission spectrum of coumarin and the absorbance spectrum of flavin (Figure 3.16).

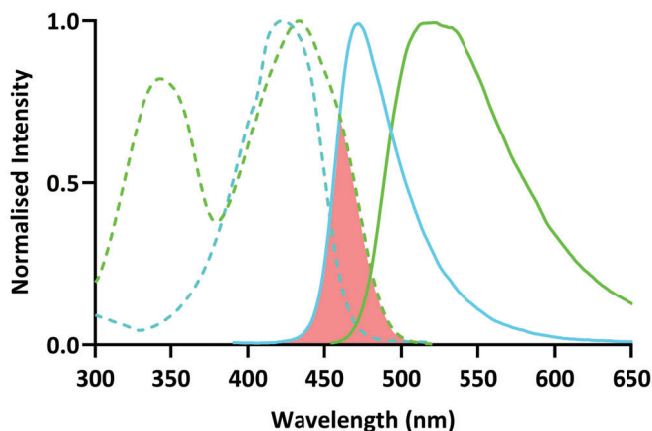


Figure 3.16: Normalised absorbance (dashed) and emission (solid) spectra of *N*-ethyl flavin (green) and 7-diethylamino coumarin (blue), respectively. Spectra were taken in MilliQ water. Red highlight indicates region of overlap.

In order to calculate the spectral overlap, J , the absorption spectrum of the acceptor, flavin, was converted to be expressed in units of $\text{M}^{-1}\text{cm}^{-1}$ vs λ . To do this, the extinction coefficient at the maximally absorbing wavelength was calculated using the Beer-Lambert law. The *N*-ethyl flavin absorbance spectrum was measured at a concentration of 0.100 mM. At this concentration, the maximum absorbance value of 0.868 au was found to be at 433 nm (Figure S1). Therefore, with a path length of 1.00 cm the extinction coefficient was calculated to be $8670 \text{ M}^{-1}\text{cm}^{-1}$. This value was used to convert the absorption spectrum to a plot of molar absorptivity. Next, the emission spectrum of the donor, coumarin, was normalised to the highest peak value. From here, the integral was calculated over wavelengths ranging from 380 to 600 nm. While the integral is defined from 0 - ∞ , the spectral overlap below 380 and above 600 nm is negligible and was therefore estimated to be zero. This gave

the final spectral overlap integral value of $1.02 \times 10^{14} \text{ nm}^4\text{M}^{-1}\text{cm}^{-1}$.

3.6.3 Calculating the inter-chromophore distance, R

Herein, the required values were determined computationally. Firstly, the minimised geometry of FCR1 was calculated. In order to speed up the calculations, ethyl groups on the coumarin and flavin were simplified to methyl groups (Figure 3.17). Because these groups do not contribute to the conjugation, they do not have a significant effect on fluorescence behaviour. No other simplifications to the structure were made.

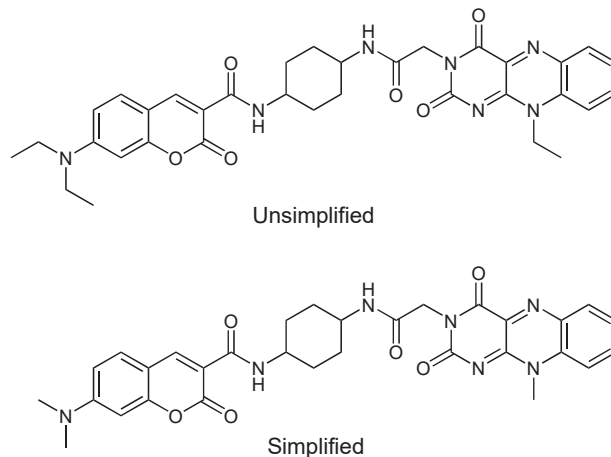


Figure 3.17: Structure simplification for **FCR1** calculations.

In order to find a minimum geometry, multiple starting geometries must be considered to sample positions along the potential energy surface. A Python script that samples by altering multiple dihedral angles was created, and used to generate 124 starting geometries. The feasibility of each structure was assessed and 16 were determined to be chemically invalid. The remaining structures were optimised using B3LYP/6-31G* level of theory that is very commonly used for the optimisation of small organic molecules.^{296–298} At first, the optimisations were performed in the gas phase due to the large number of files, in order to reduce computational time. Once calculated, the lowest 10 conformations from the gas phase were re-optimised in water using the conductor-like polarisable continuum model (CPCM) to emulate protein binding conditions.²⁹⁹ The lowest energy conformer was subjected to a frequency analysis to confirm that the structure was at a minimum on the potential energy surface, but was found to have 3 imaginary frequencies with amplitudes of

51.92, 45.13, and 25.49, indicating that the geometry is not at a minimum. Therefore, the lowest energy conformer geometry was further refined using a finer quadrature grid (17,674 vs 3,816 number of grid points). The frequency was then computed again, and 0 imaginary frequencies were found. Thus, the lowest energy geometry of **FCR1** in water was found.

The distance between the coumarin and flavin was determined by averaging the coordinates of each chromophore to find the mid-points. From this, distance between the donor and acceptor, R , was found to be 16.62 Å (Figure 3.18).

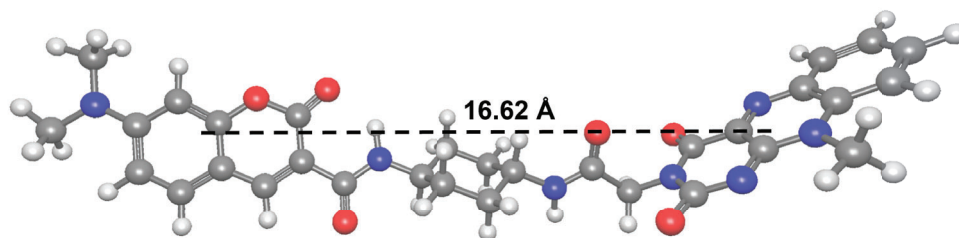


Figure 3.18: Inter-chromophore distance from optimised geometry of **FCR1**.

3.6.4 Calculating the orientation factor, κ^2

The orientation factor, κ^2 is dependent on the angles between the donor's excitation, and acceptor's ground transition dipole moments (Figure 3.19). This value can range from 0 when the transition dipole moment vectors are perpendicular, to 4 when they are parallel.

If the two fluorophores were unbound, and free to undergo isotropic reorientation in a time-frame less than that of the donor's excited state lifetime, the value can be approximated as $\kappa^2 = 2/3$.^{1,300} However, in the case of **FCR1**, the two fluorophores are bound by a rigid linker, and hence do not have freedom of reorientation. As such, κ^2 cannot be assumed to be 2/3. Therefore, κ^2 of the fluorophores in **FCR1** were determined computationally.

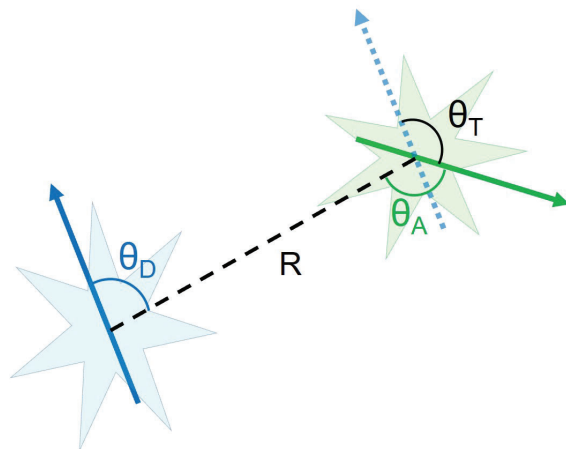


Figure 3.19: Angles between the donor excited transition dipole moment (blue arrow), the acceptor transition dipole moment (green arrow) with respect to the axis of the inter-fluorophore distance (R)

Firstly, the transition dipole moments of the individual chromophores were calculated. To do this, each chromophore was separated from the optimised **FCR1** structure for independent calculations. The point of separation was determined to be one carbon from the last atom involved in the chromophores' conjugation to ensure the entire structure involved in fluorescence was included in the calculation.

For these calculations, M11/6-31+G* level of theory was employed as it has been benchmarked against a number of conjugated systems, including fluorophores like coumarin, and was shown to calculate transition dipole moments of various excited states with the lowest angle of deviation.³⁰¹

As stated previously, the calculation of the orientation factor angles is based on the transition dipole moment of the ground state for the acceptor, and excited state for the donor. Thus, for the donor coumarin, TD-DFT was employed to find the most significant excited state. These calculations showed that the first excited state had the most significant intensity. (Table 3.1).

Table 3.1: Vertical excitation energies of various singlet excited states of coumarin.

S_n	Energy (eV)	Strength
1	3.82	1.04
2	4.11	0.00
3	4.75	0.04
4	5.15	0.03
5	5.22	0.00

Therefore, geometry optimisation of this first singlet excited state was carried out, after which TD-DFT was performed to calculate the transition dipole moment of this excited state. The transition dipole moment of flavin at the ground state was then calculated (Table 3.2). Ideally, these transition dipole moments should have large vectors in the same direction, which is favourable for the dipole-dipole interaction. The largest transition dipole moment vector is in the X-direction for both coumarin and flavin. The transition dipole moments are shown visually (Figure 3.20).

Table 3.2: Transition dipole moment vectors of flavin and coumarin calculated with M11/6-31+G* level of theory

	X	Y	Z
Coumarin (Emission)	-3.513	-0.680	0.679
Flavin (Absorbance)	-1.946	1.098	-0.874

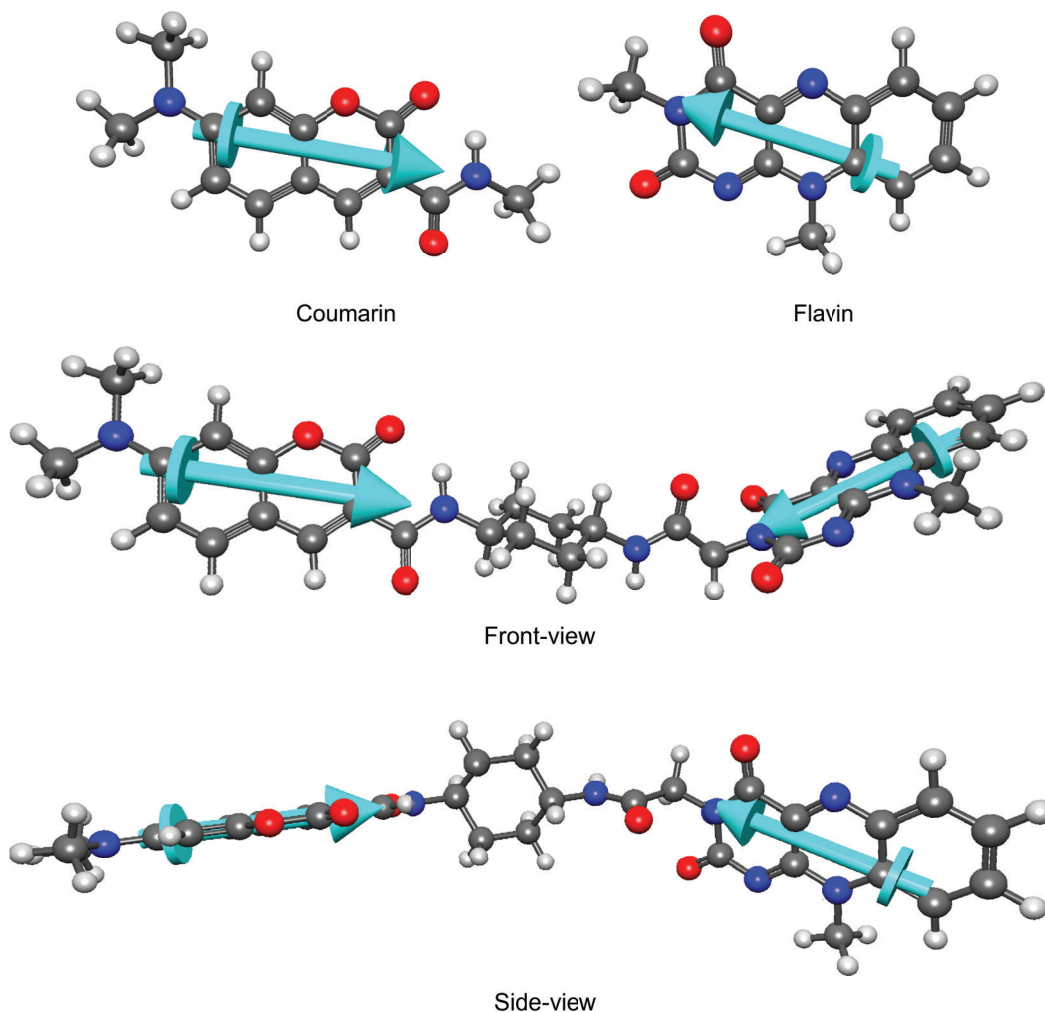


Figure 3.20: Transition dipole moments (blue arrow) of the isolated coumarin- and flavin-based chromophores independently (separated at 1 carbon beyond the last atom involved in the chromophores' conjugation) and when bound by the cyclohexane linker as calculated using M11/6-31+G* level of theory.

To calculate κ^2 , the angles between the transition dipoles and the distance axis were determined. Firstly, the vector of the distance axis was calculated by using the three-dimensional coordinates of the chromophore centre points, which was determined in section 3.6.3. Next, the magnitude of each vector was calculated - that is, the transition dipole moment vectors and the distance axis vectors. From here, the angle was determined by employing a rearranged dot product of two vectors (Table 3.3). As such, κ^2 was calculated

using equation 3.4 to give a value of 3.136, indicating excellent overlap.

Table 3.3: Calculated angles between the transition dipole moments of the specified angles.

	Angle ($^{\circ}$)	Angle (rad)
θ_D	16.91	0.295
θ_A	32.43	0.566
θ_T	49.33	0.861

3.6.5 Calculating the Förster radius, R_0

The Förster radius, R_0 , is the donor-acceptor distance at which half the energy of the donor is transmitted by energy transfer, and the other half decays by usual radiative and/or non-radiative pathways. More simply, R_0 is the distance at which $E = 0.5$. At distances below R_0 , the FRET efficiency rapidly moves toward 100%, while the converse is true for distances above R_0 (Figure 3.21).

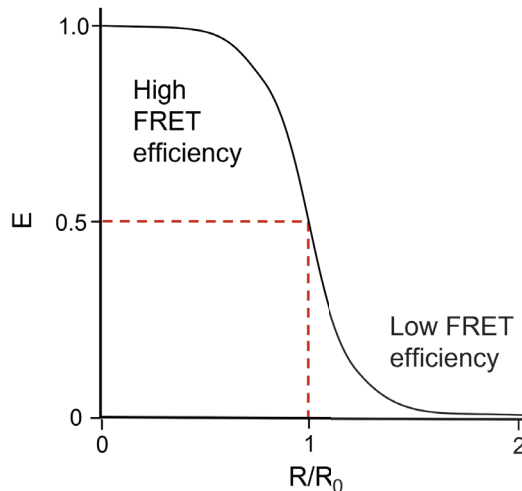


Figure 3.21: Distance dependence of FRET efficiency, E .

The Förster radius of FCR1 was calculated according to equation 3.2, using a Φ_D value of 0.219.³⁰² The refractive index value, n , was 1.33 which is the refractive index of water. From these values and the values calculated in the previous sections, the R_0 of FCR1 was calculated to be 35.44 Å.

3.6.6 Calculating FRET efficiency, E

With the values R and R_0 in hand, the FRET efficiency, E (equation 3.1), of **FCR1** in water was calculated to be 98.95%. This indicates that when bound to the HaloTag enzyme, **FCR1** should exhibit excellent transfer of donor energy to the acceptor, making it an ideal sensor for this application.

3.6.7 Designing future FRET systems

At first, it may be tempting to assume that FRET probes should be synthesised to have the shortest donor-acceptor distances possible in order to maximise FRET efficiency (Figure 3.19). However, at distances less than 10 Å other energy transfer mechanisms, such as Dexter energy transfer (DET), become competing factors. DET is considered to be a fluorescence quenching mechanism. Like FRET, DET is a distance dependant energy transfer mechanism, but unlike FRET, the rate of dexter energy transfer decays exponentially as the distance increases. Also unlike FRET, DET is an electron exchange interaction that requires overlap of the donor and acceptors molecular orbitals. As such, DET becomes near negligible at distances >10 Å. The inter-chromophore distance in **FCR1** is around 17 Å, which sits well outside of the significant DET distance range. Previous iterations of **FCR1** design, **FCR1B** were suggested based on a triazole linker (Figure 3.22), which would result in a decreased interchromophore distance (Amandeep Kaur, PhD thesis). However, due to synthetic difficulty with the click reaction, the cyclohexane linker design was pursued resulting in the final design of **FCR1**.

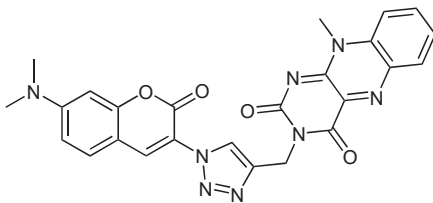


Figure 3.22: Structure of potential **FCR1B** using a triazole linker.

To test the feasibility of triazole linker for this FRET system, geometry optimisations were ensued in the same manner as performed for **FCR1** (Figure 3.23). The inter-chromophore distance was calculated as 10.39 Å. This may have resulted in a greater FRET

efficiency value, but this value is too close to the DET range, resulting in reduced FRET efficiency and inaccurate sensing.

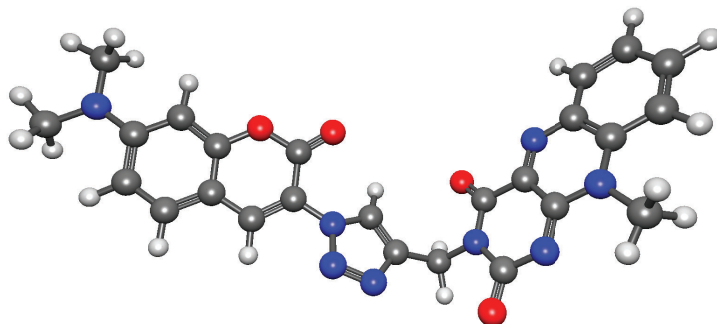


Figure 3.23: Optimised structure of **FCR1** with triazole linker.

Given the incidental nature of the **FCR1** design, this highlights the importance of molecular design and testing through computational chemistry. Even minor changes in molecular design can have a significant impact on the overall function of the molecule, and computational tools can help predict, and optimise these properties.

3.7 Benchmarking studies on a novel naphthalimide-flavin derivative

In computational chemistry, benchmarking refers to the process of evaluating the accuracy, and reliability of a computational method for a particular application. The benchmarking process can involve comparing the results of a given method to experimental data, to other computational methods, or to both. In computational chemistry, various properties of molecular systems can be calculated, such as energy levels, bond lengths and angles, vibrational frequencies, excitation energies, and more. The results of method benchmarking studies are often used to guide the development of new computational chemistry methods, or to optimise existing methods for specific applications. In this section, benchmarking of various methods for the evaluation of the excitation energies of **NpFR1** will be performed.

3.7.1 NpFR1

In 2014, Yeow *et al.* designed and developed a redox sensor based on the reversible redox response of the flavin moiety. The fluorescent scaffold was based on the fusion of a 1,8-naphthalimide and a flavin to form the novel derivative, **NpFR1** (Figure 3.25). This sensor exhibits a 125-fold increase in fluorescence signal in an oxidising environment, and near-zero fluorescence in a reducing environment. The redox mechanism of this sensor is based on the flavin moiety, making it reversible over many cycles. **NpFR1** is a redox sensor that is not specific to a particular ROS, and can measure the overall redox state of its environment.

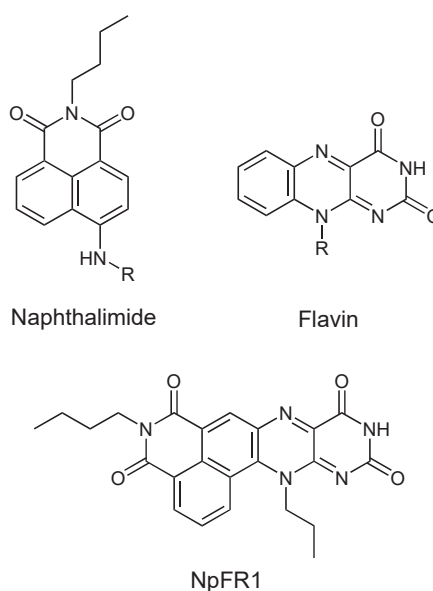


Figure 3.24: Chemical structures of the naphthalimide and flavin cores, and the novel derivative **NpFR1**.

The absorption spectrum of **NpFR1** contains four excitation transitions including a shoulder peak. To obtain more accurate values, a Gaussian fit was applied to each peak (Figure S2). By fitting a Gaussian curve to the peak, the true maximum can be calculated with greater accuracy. The values are reported in Table 3.4.

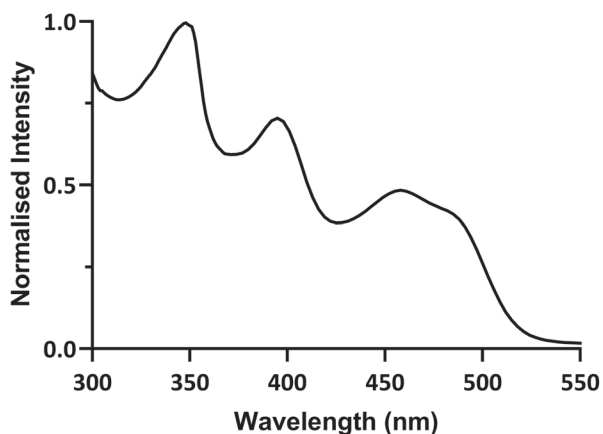


Figure 3.25: Normalised absorbance spectrum of **NpFR1**.

Table 3.4: Absorbance maxima of NpFR1 as determined by Gaussian fitting.

Transition no.	Wavelegnth (nm)	Energy (eV)
1	489	2.54
2	473	2.62
3	380	3.26
4	345	3.59

3.7.2 Choosing benchmarking methods

The two most important considerations when choosing an appropriate level of theory are accuracy and computational cost. Accuracy determines how well the calculated results match with what is seen experimentally. Because these methods are based on a range of different approximations and assumptions, the accuracy and applicability of a given method can vary depending on the type of calculation being performed. The cost of a calculation is determined by many factors including molecule size, accuracy of the employed method, size of the basis set, and the resources available to the computer performing these calculations, such as memory and processing power. The more accurate the method employed, the more computationally expensive, in that they require more resources and time to compute. In general, it is important to choose a method that is suitable for the specific calculation goals, while striking an appropriate balance between accuracy and cost.

In the case on **NpFR1**, which is a hybrid between two well known fluorophores with differing fluorescence mechanisms, selecting a computation method will involve sampling methods that have been established for both independent fluorophores. Previously, benchmarking studies have been performed on the individual naphthalimide and flavin scaffolds. These findings will inform the methods that will be used to benchmark **NpFR1** as the electronics of this scaffold are most likely to resemble one, or both of the independent scaffolds.

In previous computational studies of the naphthalimide scaffold, time dependent density functional theory (TD-DFT) was used to calculate the vertical excitation energies and fluorescence of the system. TD-PBE0/6-31+G* level of theory was employed for the calculations and the polarisable continuum model (PCM) model was used to model solvent effects. This yielded a mean absolute deviation of 0.13 eV for vertical excitation energies and 0.16 eV for fluorescence, when using solvents such as hexane, DCM, ethanol, and acetonitrile. It was discovered that the deviation was significantly larger for gas-phase calculations compared to solvated ones, particularly for fluorescence, highlighting the importance of performing calculations with solvent effects.³⁰³

In a separate study, TD-DFT was used to analyse the naphthalimide scaffold using the PBE0 method, and the accuracy of various Pople basis sets was evaluated. The smallest basis sets (6-31G, 6-31G(d), and 6-31+G) were found to give more accurate results for the first transition, but the correct excited-state nature was only obtained using the larger 6-31+G(d) basis set, which resulted in energies 0.033 to 0.15 eV away from experimental values. On the other hand, the 6-311++G(2d,2p) basis set gave results that were only 0.011 to 0.028 eV different from those obtained with the 6-31+G(d) basis set, but the calculation took five times longer. Additionally, configuration interaction singles (CIS) and Hartree Fock (HF) in conjunction with the 6-311++G** basis set, were assessed and showed that CIS lead to valid geometries, although the transition energies were much larger than the experimental values.³⁰⁴

The method used to obtain emission spectra for three 1,8-naphthalimide derivatives involved two steps: 1) optimising the geometry of the first excited state using CIS with the 6-31G** basis set, and 2) calculating vertical excitation energies using TD-B3LYP with 6-31G** and 6-311++G** basis sets. The results from the CIS/6-31G** level of theory did not agree with experimental results. On the other hand, the TD-DFT with B3LYP using

the 6-31G** and 6-311++G** basis sets showed good agreement with the experimental values.³⁰⁵

In another study, TD-B3LYP method using the 6-31G* basis set was used to calculate the excitation energy of a 1,8-naphthalimide. The calculations gave an excitation of 368 nm compared to the experimental values of 432 nm, giving a difference of 0.50 eV. However, it is noted that the calculations were performed in the gas state while the experimental spectra were taken in water, so the effect of solvation was not considered, which may explain the discrepancy.³⁰⁶

In the case of flavin-based molecules, there are many studies that have assessed the efficacy of various methods to calculate the excitation and emission energies. For example, Neiss *et al.* compared the results of CIS and TD-DFT methods with those of a multireference configuration interaction DFT method (DFT/MRCI) to calculate excitation and emission energies of various flavin derivatives. In the case of lumiflavin, the results showed that the CIS method consistently predicted larger excitation energies than the DFT-based methods, which were in better agreement with experimental values. The energies calculated with TD-B3LYP using either the 6-31G* or SV(P) basis sets gave very similar results. They also showed that TD-B3LYP and DFT/MRCI using B3LYP gave results with deviations rarely exceeding 0.2 eV.³⁰⁷

List *et al.* compared the performance of time-dependent density functional theory (TD-DFT) with the CAM-B3LYP functional and the cc-pVDZ+ basis set to the approximate coupled cluster singles and doubles (CC2) approach to model the excitation and emission profile of flavin mononucleotide. They found that TD-CAM-B3LYP/cc-pVDZ slightly overestimates excitation energies when compared to CC2 results, with deviations ranging from 0.1 to 0.3 eV when a polarisable embedding scheme was employed to model an aqueous environment.³⁰⁸

There have also been some computational studies performed on the various redox states of flavin. For example, Kabir *et al.* assessed the vertical excitation and emission profile of flavin derivatives in 5 different redox states. They employed TD-DFT with B3LYP/cc-pVTZ level of theory and the polarisable continuum model (PCM) to take into account solvation effects. The results showed that the computed spectra were consistent with the experimental spectra for the oxidised form of flavin and the $1e^{-1}$ reduced radical forms, with

an average energy difference of only 0.085 eV. However, the agreement was not as strong for the neutral and anionic forms, with an average energy deviation of 0.16 eV.³⁰⁹

Another study assessed utility of various methods including B3LYP, PBE0, M06-2X, M06, CAM-B3LYP, and LC-PBE. They found that B3LYP yielded the smallest mean deviation in excitation energies, while M06-2X provided reasonable band shapes.³¹⁰

Based on the considerations mentioned above, several methods were selected for the excited state calculations of **NpFR1**. For a rough estimation, the CIS and its perturbation counterpart CIS(D) were chosen. The maximum overlap method (MOM) was selected for the calculation of single transitions because of its accuracy with internal charge transfer mechanism, which are inherent to naphthalimide scaffolds. Finally, TD-DFT was chosen for its versatility. The specific functionals employed were selected from the predominating literature benchmarking, including B3LYP, CAM-B3LYP, PBE0, and M06-2X.

3.7.3 Geometry optimisation of the ground state

To begin, the ground-state geometry of oxidised **NpFR1** was calculated in its oxidised form using B3LYP/6-31G* level of theory. This level of theory is commonly used in the geometry optimisations of small organic molecules. Several initial geometries were explored to survey the potential energy surface (PES). The lowest energy conformer was subjected to a frequency analysis, which revealed no imaginary frequencies, signifying a minimum on the PES. The result showed that **NpFR1** has a planar structure and a extensive conjugated electron system (Figure 3.26).

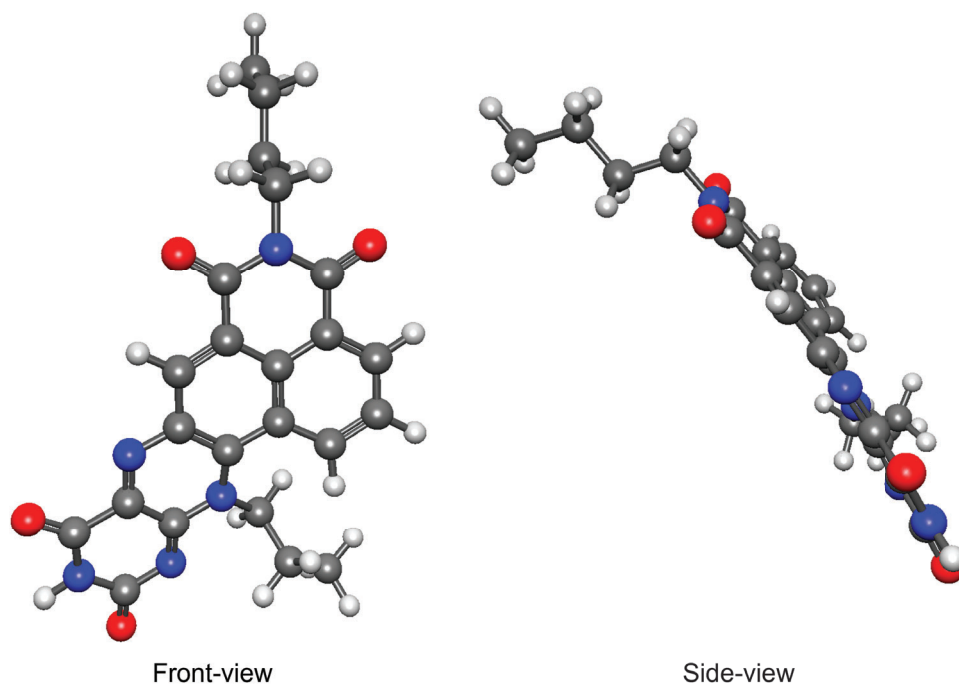


Figure 3.26: Optimised geometry of **NpFR1** in its oxidised form calculated using B3LYP/6-31G* level of theory showing the front and side views.

The highest occupied molecular orbital (HOMO) shows delocalised π orbitals across the conjugated scaffold. Lone pairs are also seen on the oxygens at the imide positions of both the naphthalimide and flavin portion of the scaffold. Similarly, the lowest unoccupied molecular orbital (LUMO) shows delocalised π orbitals across the scaffold and lone pairs on the imide oxygens (Figure 3.27).

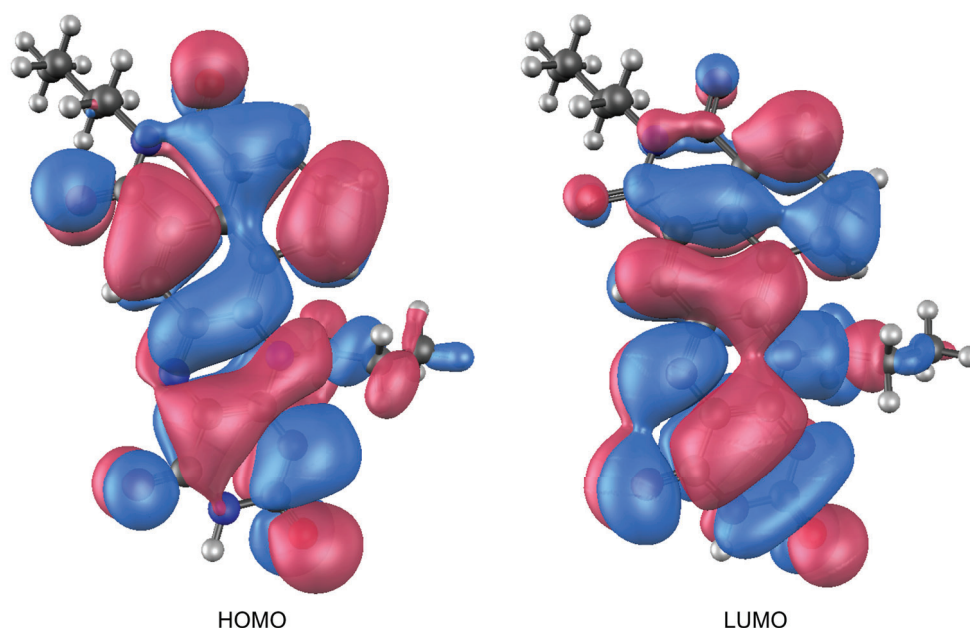


Figure 3.27: HOMO and LUMO orbitals of **NpFR1** calculated at B3LYP/6-31G* level of theory presented at $\pm 0.02 \text{ \AA}$ isovalues.

3.7.4 Assessment of configuration interaction singles (CIS)

To start the investigation of excited states, the configuration interaction singles (CIS) method was employed. The CIS method was selected as it is a simple, and "cheap" way of computing excited states. This method is also known to provide acceptable results in cases involving charge transfer mechanisms, which are characteristic of the naphthalimide scaffold. While the CIS method is a crude approximation, it serves as a useful starting point to provide quick and rough estimates of the excited states and their energies, which can then be used to guide more sophisticated calculations.

The CIS method was used in this study to evaluate the performance of two Pople basis sets, 6-31G* and 6-31+G*. Both basis sets are popular choices for molecular calculations. 6-31G* is a relatively small basis set that is commonly used in geometry calculations due to its speed and accuracy. The * indicates the addition of polarisation on heavy atoms. In some cases, this basis set is capable of accurately describing electronic transitions. On the other hand, the 6-31+G* basis set, includes an added component, +, that indicates diffuse functions. These diffuse functions are important for accurately describing electronic transi-

tions because they account for the delocalised nature of electrons in excited states. While this basis set is generally more accurate, calculations are more computationally expensive.

Additionally, the CIS method was utilised to evaluate the impact of solvation on electronic transitions. As such, both gas phase and aqueous transitions were calculated. The CPCM model was selected as the solvation model, and the excitation energies were calculated in water. CPCM is a commonly used solvation model that considers the solvent as a continuous dielectric medium surrounding the molecule. The results are summarised in (Table 3.5).

Table 3.5: Vertical excitation parameters of **NpFR1** calculated using CIS with 6-31G* or 6-31+G* basis sets in either the gas state, or solvated in water.

Basis Set	Energy (eV)	λ (nm)	Strength	Transition
6-31G* (Gas)	4.16	298	0.382	HOMO - LUMO
	4.59	270	0.438	HOMO - LUMO+1
	5.06	245	0.971	HOMO-1 - LUMO
6-31+G* (Gas)	4.10	303	0.381	HOMO - LUMO
	4.53	274	0.463	HOMO - LUMO+1
	4.98	249	0.956	HOMO-1 - LUMO
6-31G* (Water)	4.07	305	0.682	HOMO - LUMO
	4.52	275	0.514	HOMO - LUMO+1
	4.98	249	0.905	HOMO-1 - LUMO
6-31+G* (Water)	4.01	309	0.707	HOMO - LUMO
	4.44	279	0.530	HOMO - LUMO+1
	4.89	254	0.832	HOMO-1 - LUMO

From these calculations, it can be seen that the CIS method consistently overestimates the transition energies compared to experimental values. For all calculations, the orbitals involved in the most significant transitions are HOMO - LUMO, HOMO - LUMO+1, and HOMO-1 - LUMO.

Comparing the transition energies of the two basis sets in the gas phase, it was found that the 6-31+G* basis set generally yields lower, and thus more accurate transition energies when compared to the experimental values. However, the strengths of the transitions are relatively similar between both basis sets. When modelled in an aqueous environment, both basis sets calculate lower energies for the electronic transitions compared to the gas phase results. The strengths of the transitions also differ from those in the gas phase. The HOMO - LUMO transition in the aqueous calculations gave a higher oscillator strength, indicating higher transition probability when compared to the gas phase. This is in better agreement with the intensities of the peaks of the experimental spectrum.

As a result, it was determined that the inclusion of diffuse functions is important in improving the overestimation of excitation energies. The impact of solvent effects on oscillator strengths was also found to be significant. As a result, subsequent calculations were performed using a basis set of at least 6-31+G* in an aqueous environment.

3.7.5 Assessment of configuration interaction with single and double CIS(D)

Next, calculations were performed using the configuration interaction singles and doubles (CIS(D)) method. CIS(D) is a perturbative doubles correction to the uncorrelated CIS method. Correlated methods are generally more accurate due to their consideration of electron correlations. However, modelling these effects may not always have a significant effect on the calculated results of certain systems. If they are significant, it is important to select methods that include correlation effects to ensure this effect is calculated accurately. Therefore, CIS(D) was used to assess the effect of correlation on the excited-state calculations of **NpFR1**. The calculations were carried out using the 6-31+G* basis set in an aqueous environment.

The calculations using CIS(D) yielded almost no improvement over CIS (Table 3.6). The lack of difference between these methods suggests that correlated effects are not important for the study of **NpFR1**.

Table 3.6: Vertical excitation parameters of **NpFR1** calculated using CIS(D) with 6-31+G* basis sets in either the gas state, or solvated in water.

Basis Set	Energy (eV)	λ (nm)	Strength	Transition
6-31+G* (Gas)	4.10	303	0.382	HOMO - LUMO
	4.53	274	0.463	HOMO - LUMO+1
	4.98	249	0.956	HOMO-1 - LUMO
6-31+G* (water)	4.01	309	0.707	HOMO - LUMO
	4.44	279	0.530	HOMO - LUMO+1
	4.89	254	0.833	HOMO-1 - LUMO

3.7.6 Assessment of time-dependent density functional theory

Time-dependent density functional theory (TD-DFT) is a time-dependent density functional theory method that can be used to calculate transition energies. TD-DFT is a good method to use for electronic transitions that can be described by single excitations, such as those involving the promotion of a single electron from a occupied molecular orbital to an unoccupied orbital. However, TD-DFT can have limitations for more complex systems, such as those that involve multiple excitations or electron transfer processes.

The methods chosen for TD-DFT studies were as follows: B3LYP, a hybrid functional with Becke 1988 exchange functional and Lee, Yang and Parr correlation functional; PBE0, hybrid generalised gradient functional combines the Perdew-Burke-Ernzerhof functional with a proportion of the exact Hartree-Fock exchange; CAM-B3LYP, a hybrid functional that combines the properties B3LYP and the long-range correction, Coulomb-attenuating method (CAM); and M06-2X, that combines Becke 2006 exchange functional and the M06 correlation functional with a long-range correction term. All calculations assessing these four methods used the 6-31+G* basis set, and in aqueous solvent using the CPCM method (Table 3.7).

Table 3.7: Vertical excitation parameters of NpFR1 calculated using TD-DFT with the specified method and the 6-31+G* basis sets, solvated in water.

Method	Energy (eV)	λ (nm)	Strength	Transition
B3LYP	2.87	432	0.338	HOMO - LUMO
	3.29	377	0.227	HOMO - LUMO+1
	3.66	339	0.799	HOMO-1 - LUMO
PBE0	2.98	416	0.376	HOMO - LUMO
	3.41	364	0.273	HOMO - LUMO+1
	3.80	327	0.823	HOMO-1 - LUMO
CAM-B3LYP	3.26	380	0.552	HOMO - LUMO
	3.77	329	0.417	HOMO - LUMO+1
	4.14	299	0.772	HOMO-1 - LUMO
M06-2X	3.29	377	0.536	HOMO - LUMO
	3.77	329	0.386	HOMO - LUMO+1
	4.20	295	0.758	HOMO-1 - LUMO

The results show that the orbitals involved in these transitions are consistent between all the functionals, and with the previously determined CIS and CIS(D) methods.

The long-range corrected functionals M06-2X and CAM-B3LYP have a similar approach to predicting the transition intensities. Both of these functionals show a higher strength for the first and second transitions, compared to what is given by the B3LYP and PBE0 functionals. The difference in prediction may be due to the fact that long-range corrected functionals account for the effect of long-range interactions on the electron transition, which is not considered in B3LYP and PBE0.

In the comparison of the transition energies, it can be seen that the M06-2X functional had the largest overestimation of energies when compared to the experimental results. The CAM-B3LYP functional also overestimates the transition energies, although to a lesser extent than M06-2X. On the other hand, B3LYP and PBE0 performed relatively similarly in terms of transition strength. However, B3LYP had the least overestimated transition

energies compared to the experimental results. The differences in energy between the B3LYP transitions and the experimental values is relatively small, with a difference of 0.25 eV for the first transition, 0.026 eV for the second transition, and 0.063 eV for the third transition. This indicates that of the four functionals assessed, the B3LYP method is most suited to studying the vertical transition energies of **NpFR1**.

Next, to determine the effect of using a triple zeta basis set on the accuracy of the calculations, two basis sets were selected for comparison: 6-311++G** and aug-cc-PVTZ. In a triple zeta basis set, the valence orbitals are treated with three sets of Gaussian functions instead of the two sets in a double zeta basis set, like 6-31+G*. Both basis sets include polarisation and diffuse functions, which allow for the treatment of more diffuse electronic states, which are important for describing excited states. The ++ in this case indicates diffuse functions on both heavy atoms and hydrogen included. Additionally, the ** indicates polarisation on both heavy atoms and hydrogens. These parameters were selected to get a closer comparison to the "aug" function on the cc-PVTZ basis set which indicates diffuse function on all atoms, in addition to the inherent polarisation on all atoms of this basis set. The selected basis sets differ in terms of the number of basis functions, with aug-cc-PVTZ having a greater number per atom when compared to 6-311++G*. However, it is important to note that using a larger basis set can also increase the computational cost of the calculations. Both calculations were performed using the B3LYP method in aqueous solvent using the CPCM method (Table 3.8).

Table 3.8: TD-DFT calculations using the B3LYP method and the specified triple zeta basis sets.

Method	Time	Energy (eV)	λ (nm)	Strength	Transition
6-311++G**	0d 6h	2.87	432	0.334	HOMO - LUMO
		3.26	378	0.217	HOMO - LUMO+1
		3.64	339	0.809	HOMO-1 - LUMO
aug-cc-PVTZ	8d 2h	2.85	436	0.327	HOMO - LUMO
		3.26	380	0.219	HOMO - LUMO+1
		3.64	341	0.808	HOMO-1 - LUMO

The results indicate that the two basis sets produce comparable excitation energies for all three excited states. Additionally, the similar oscillator strengths suggest that the likelihood of electronic transition is consistent between the basis sets. Overall, the outcomes suggest that the using a larger triple zeta basis set does not significantly impact the accuracy of the predicted excitation energies. It is important to note, that the TD-DFT calculation using the aug-cc-PVTZ basis set took significantly longer to complete (>8 days) in comparison to the calculation utilising the 6-311++G** basis set (6 hours). Considering the significantly greater computational time required for the aug-cc-PVTZ basis set, and the minor improvement in accuracy, it was not considered to be worth the additional time, and resources. Also of note, despite the larger size and the inclusion of additional functions such as diffuse, and polarisation functions on the hydrogens, the 6-311++G** basis set did not lead to a significant improvement over the smaller double zeta 6-31+G* basis set.

Based on the results of these calculations, B3LYP/6-31+G* level of theory was deemed the most appropriate for vertical excited state descriptions for **NpFR1**. Among the various methods compared, B3LYP was found to give the least overestimation of vertical excitation energies, while the 6-31+G* basis set demonstrated an excellent balance between accuracy and computational cost when compared to other basis sets.

3.7.7 Assessment of the maximum overlap method

The maximum overlap method (MOM) is an alternative method to CIS and TD-DFT for calculating excited states. In MOM calculations, the orbitals involved in the transition are manually selected, and the specified occupancy is maintained throughout the calculation. MOM is particularly effective for calculating excited states that involve electron transfer, such as internal charge transfer. This makes MOM well-suited for calculating excited states of molecules that exhibit internal charge transfer such as naphthalimides - assuming the fluorescence mechanism of NpFR1 resembles that of a naphthalimide.

In order to sample multiple transitions of interest, a python script was created to generate files specifying orbital overlaps sequentially between orbital HOMO-3 to HOMO+3. These orbitals were identified as significant transitions from the previous CIS and TD-DFT calculations. For the calculations, B3LYP/6-31+G* level of theory was employed based on

it having the highest accuracy in TD-DFT calculations. Aqueous solvent was used by utilising the CPCM model. Of all the calculations performed, the ones with the most significant strength have been reported (Table 3.9).

Table 3.9: MOM with significant strengths in order of energy using B3LYP 6-31+G* in CPCM

Orbitals	Energy (eV)	λ (nm)	Strength
HOMO-1 - LUMO+3	4.53	274	0.244
HOMO-1 - LUMO+2	4.26	291	0.556
HOMO - LUMO+2	4.26	291	0.570
HOMO-2 - LUMO+1	3.25	382	0.371
HOMO - LUMO+3	2.32	535	0.419

The significant transitions found in the CIS and TD-DFT, namely the HOMO - LUMO, HOMO - LUMO+1, and HOMO-1 - LUMO transitions all collapsed back to the initial state when using the MOM method (Table S1). There are a few reasons why this might occur. One potential explanation is a small energy gap between the orbitals, but this does not seem to be the case here as the energy gaps between the relevant orbitals are approximately 3-4 eV, which is not considered a small gap in energy. Another possible reason for the collapse is an inaccurate descriptions of the orbitals. This can happen when the orbitals are not well-described by the basis set used in the calculation, or when the approximations made by the method are not suitable for the electronic structure of the molecule. This can be addressed by redefining orbitals using a larger basis set, but these calculations can become computationally expensive. Additionally, most previous studies have identified 6-31+G* as a sufficient basis set for excited state descriptions. Lastly, the collapse can also occur due to the use of a small quadrature grid. In this case, the calculations were performed using standard grid SG-0 which uses 3,816 points.³¹¹ This could be remedied by utilising a larger grid, such as SG-3 which uses 17,674 points, allowing for diffuse densities to be more accurately integrated.³¹² As such, MOM calculations for the significant transitions was re-calculated using the finer grid setting. However, this did

not resolve the problem, as the transitions still collapsed back to the ground state, as indicated by an overlap value of 1.(Table 3.10). Therefore, a variation of MOM, initial maximum overlap method (IMOM) was employed. This method has been shown to provide more reliable convergence for more difficult calculations that may involve charge-transfer mechanisms. However, this too resulted in collapsing states. As such, MOM was considered inappropriate for the calculation of excitation energies for the **NpFR1** scaffold.

Table 3.10: Overlap of significant transitions calculated using MOM and the SG-3 grid. IMOM calculations yielded the same results.

Transition	Orbital Overlap
HOMO - LUMO	1.00
HOMO - LUMO+1	1.00
HOMO-1 - LUMO	1.00

From these results, TD-B3LYP/6-31+G* was concluded to be the most accurate descriptor of **NpFR1** excitation transitions from all the methods tested.

3.8 Conclusions

Previous work highlighted the utility of a protein targeted redox sensor in revealing a previous undiscovered mechanism for copper homeostasis at CTR1. As such, the development of new and improved redox sensors for use with the HaloTag system was desired. Two fluorescent sensors, **FCR1** and **NpFR1** were selected as sensors of interest on which to perform computational studies to assess and develop methods to assess the suitability of these sensors for measuring the redox state with the HaloTag system.

Molecular docking studies using AutoDock4 revealed that the HaloTag-bound sensors sit well outside of the binding pocket, indicating that an aqueous solvated environment would be appropriate for the assessment of the sensors.

For **FCR1**, the FRET efficiency was calculated by using a combination of experimental and computational results. Experimental spectra were used to calculate the spectral overlap integral, while the rest of the calculations were obtained from computational results.

Geometry optimisations revealed the inter-chromophore distance between the donor and acceptor. The transition dipole moments of flavin in the ground state, and coumarin in the first excited state were determined using TD-M11/6-31+G*. The orientation factor was calculated using these values, which revealed an excellent overlap between the transition dipole moments. From these calculations, the FRET efficiency of **FCR1** was calculated to be 98.95%, indicating excellent energy transfer between the coumarin donor and flavin acceptor. This revealed that **FCR1** should exhibit excellent fluorescent properties when utilised as a label for the HaloTag enzyme, allowing ratiometric redox measurements to be performed at a protein of interest with great accuracy.

The computational vertical excitation energies of **NpFR1**, a novel redox active scaffold, were compared to experimental values to assess the accuracy of the computational method used. Due to the novelty of this scaffold, there is not yet a consensus on the most accurate method for calculating its excitation properties. CIS provided crude results revealing the most significant orbitals involved in the transitions but highly overestimated the transition energies. CIS(D) did not provide any improvement over CIS. TD-DFT proved to show the most consistently accurate results and indicated that the B3LYP functional in combination with the 6-31+G* basis set provided the best balance of accuracy and computational cost for describing excited energies of the novel scaffold, **NpFR1**

Chapter 4

Fluorescent method to determine peptide-receptor binding *in vitro*

4.1 Overview

One of the main goals in medicinal research is to maximise the safety and efficacy of drug therapy. In recent years, there has been an increasing interest in the use of peptides as therapeutics due to their high selectivity and low toxicity. However, identifying new peptide receptor targets is a challenging task, and traditional methods such as ligand binding assays can be time-consuming and resource-intensive. This chapter will outline and investigate a proposed ligand-receptor binding assay that utilises fluorescence to enable sensitive, and rapid detection of peptide-receptor interactions.

4.1.1 Cell receptors in health and disease

Cell receptors are responsible for recognising and responding to extracellular communication signals. They can be found on the cell surface (membrane-bound) or within the cytoplasm (soluble). Signalling molecules including cytokines, hormones, neurotransmitters, peptides and other ligands, bind to these receptors and induce a conformational change which triggers a signalling cascade in the cell. Due to the large variety of signalling molecules, cells must express a large diversity of receptors which respond to a specific ligand.³¹³

In certain diseases, cell receptor expression or structure can vary from the normal state. This includes changes in the number of certain receptors expressed or a mutation of receptors

which inhibits normal function (Figure 4.1).³¹⁴ These unique receptor characteristics in diseases can be exploited to develop targeted therapeutic or diagnostic agents.

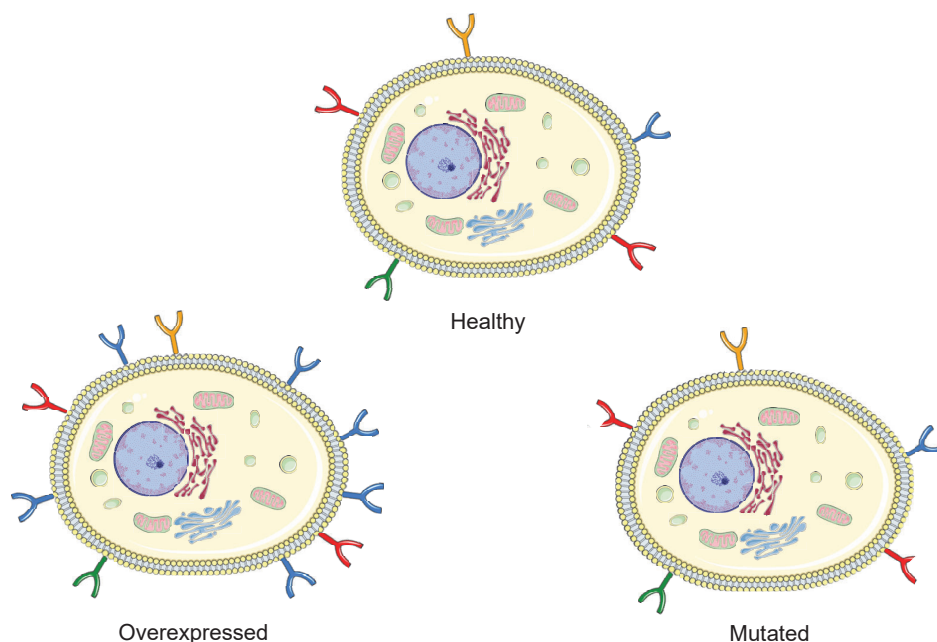


Figure 4.1: Varied receptor expression in healthy vs diseased cells

4.1.2 Peptides as targeting agents

During the 1980s, significant research was carried out to investigate the use of peptides as targeted clinical agents. For example, one of the first peptide therapeutics, Leuporelin, was launched onto the market in 1984. Peptide-based targeting agents are significantly smaller in size and mass than antibodies (Figure 4.2). They have been used both as therapeutic agents, as well as delivery agents.

Peptides offer several benefits over antibodies as clinical agents. Their smaller size offers a significant advantage in overcoming tumour penetration problems faced by the larger antibodies. The greater degree of synthetic flexibility of peptides allows for easy conjugation to a variety of chemical moieties such as small molecule drugs, transition metals, nanoparticles, diagnostic agents, and fluorophores.^{315–319} Peptides also have a lower potential for generating an immunogenic response, providing they do not contain an immunogenic sequence.³²⁰ Additionally, compared to antibodies, peptides have a higher activity per gram. Combined with considerable developments in the process of manufacturing, peptides therefore have a

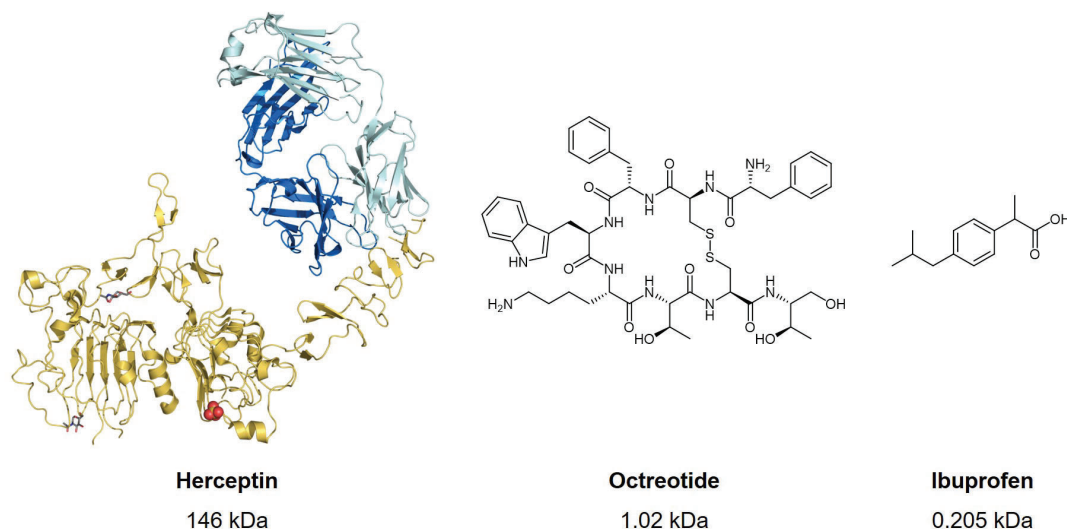


Figure 4.2: Size comparison of: Herceptin, an antibody-based drug; Octreotide, a peptide-based drug; and Ibuprofen, a small molecule drug (not shown to scale).

greater cost efficiency than antibodies.³²¹

While peptides are advantageous in comparison to antibodies, they too have some limitations. Many natural peptides are prone to enzymatic degradation in serum which reduces their half life and thus, their bioavailability. To address this, peptides can be synthetically modified in order to reduce their proteolytic susceptibility. For example, proteolysis is known to occur at both the N and C termini, and modifications such as N-acetylation and C-amidation have been shown to reduce a peptide's proteolytic susceptibility.³²² The use of unnatural D-amino acids has also been reported to reduce a peptide's affinity for proteolytic enzymes.³²³ Resistance to proteolysis can also be achieved through peptide cyclisation.³²⁴ Peptides can be cyclised head to tail, head to side chain, or side chain to side chain. While these modifications can enhance stability, it is important to ensure that these changes do not affect the peptide's interaction with the target receptor.

4.1.3 Novel fluorescent methods

Identification and determination of protein-receptor interactions is the most important step in the development of peptide-based targets. Having a reliable and efficient method to assess these interactions can greatly accelerate the process of identifying novel peptides or modified peptides that have potential for clinical use. This can help to quickly identify lead

compounds for further optimisation and development. Some fluorescent peptide-receptor detection methods have been reported.

In 2010, Ahmed *et al.* developed a peptide-whole cell screening assay by synthesising a library of 70 peptides on a cellulose membrane. The peptides were synthesised based on p160, a cancer targeting peptide. MDA-MB0435 and MCF-7 human cancer cell lines were allowed to adhere to the peptides. Cellular binding was determined using a fluorescent dye, CyQUANT, which detects cells. Three peptides were found to have an affinity of up to 3-fold higher than p160 for the tested cell lines.³²⁵ This array based binding assay is useful for screening a library of peptides. However, live-cell binding studies, such as microscopy, are not possible on the cellulose membrane.

In 2012, Frei *et al.* described this ligand-based receptor capture (LRP) technology to identify ligand-receptor interactions in live cells and tissues. They developed a chemical reagent with 3 moieties termed TRIPCEPS, that contained a group to bind glycosylated receptors, an amino-binding group, and a biotin tag to purify and identify the peptide receptor.³²⁶ While this method is useful for identifying important receptors, it lacks a reporter group and therefore cannot be used to visualise the ligand-receptor interaction *in vitro*. This system is also limited to only glycoprotein receptors.

In 2013, Zhang *et al.* developed a bioluminescence based peptide-receptor binding assay that was used to study the Insulin-like peptide 3 (INSL3) and relaxin family peptide receptor 2 (RXFP2) interaction. A luciferase bioluminescent reporter, NanoLuc, was modified to express a N-terminal cysteine residue.³²⁷ INSL3 was mutated to express an activated disulfide bond to which the modified NanoLuc reporter was conjugated. The INSL3-NanoLuc complex was shown to retain its high binding affinity to the RXFP2 receptor.³²⁸ The use of bioluminescence offers some advantages over radiolabelling such as higher sensitivity, improved safety and a longer shelf life. However, the Nanoluc reporter requires recombinant expression in *Escherichia coli* cells and further purification. Thus, this technique is labour intensive and inaccessible to the mainstream laboratory.

4.2 Aims

One of the main challenges in medicine is balancing drug efficacy with toxicity and poor bioavailability. Targeted drug therapy has the potential to address these issues by delivering

the therapeutic agent directly to the site of disease, minimising off-target effects. However, advances in the field of targeted medicine rely heavily on the discovery of novel, diseases-specific ligands. As such, the limiting factor in the successful development of these medicines is the advancement of ligand-receptor detection technologies. As such, the aims of this project were to design a fluorophore that can be incorporated into a peptide of interest with a functionalisation that allows for easy detection/isolation, and prepare a fluorophore-peptide conjugate.

4.3 Design and synthesis

4.3.1 Envisioned method of receptor discovery

A dual functionalised fluorescent sensor was designed with a solid-phase peptide synthesis-compatible group that would allow it to be incorporated into a peptide of interest. The fluorescent sensor would allow for the quick identification of peptide binding to a receptor. Additionally, a second group with a pull-down partner on the fluorophore would allow for easy isolation of the peptide-receptor complex for further investigation and characterisation of the interaction (Figure 4.3).

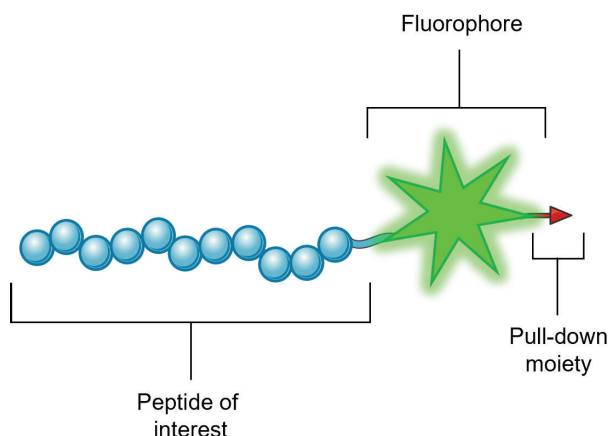


Figure 4.3: Design of the dual-functional fluorescent probe with the ability to be incorporated into the peptide of interest (blue chain) as well as containing a pull-down partner (red arrow).

For the process envisioned, the peptide-bound sensor would be incubated with the cell

line of interest. Binding can then be observed and confirmed through fluorescence analysis of the cells by using a suitable method such as flow cytometry, or microscopy (Figure 4.4).

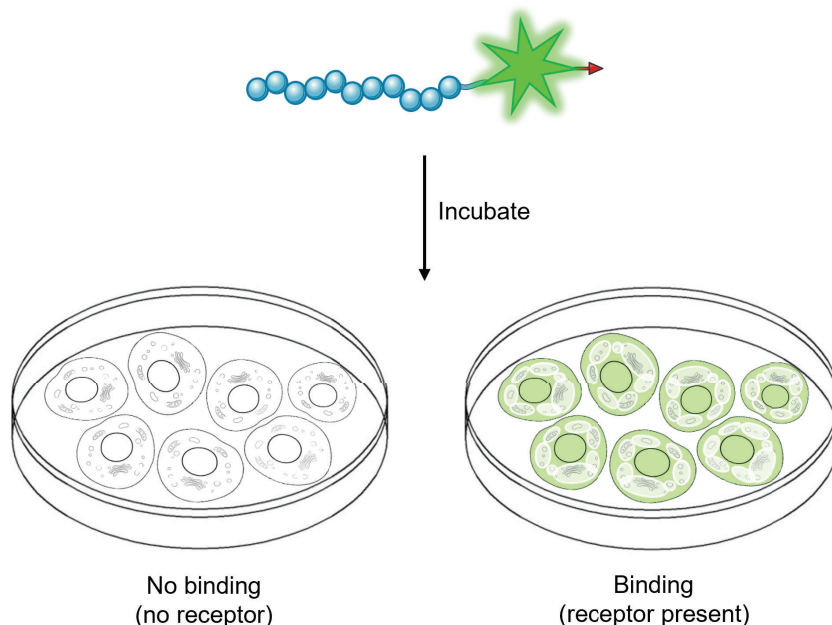


Figure 4.4: After incubating with a cell line of interest, binding of the peptide of interest to a cell line expressing a relevant receptor can be identified by the presence of fluorescence.

After binding between the peptide and receptor in the cell line is confirmed, a complementary pull-down partner can be attached to solid supports such as magnetic beads. This pull-down partner selectively binds to the fluorescently-labelled peptide-receptor complex. After binding to the magnetic beads, unbound contents are washed away. Depending on the nature of the peptide-receptor interaction, this will leave behind either the peptide alone (reversible interaction), or the peptide-receptor complex (irreversible interaction). The isolated peptide or complex can then be sent for further analysis, such as mass spectrometry, to determine the nature of the peptide-receptor interaction (Figure 4.5).

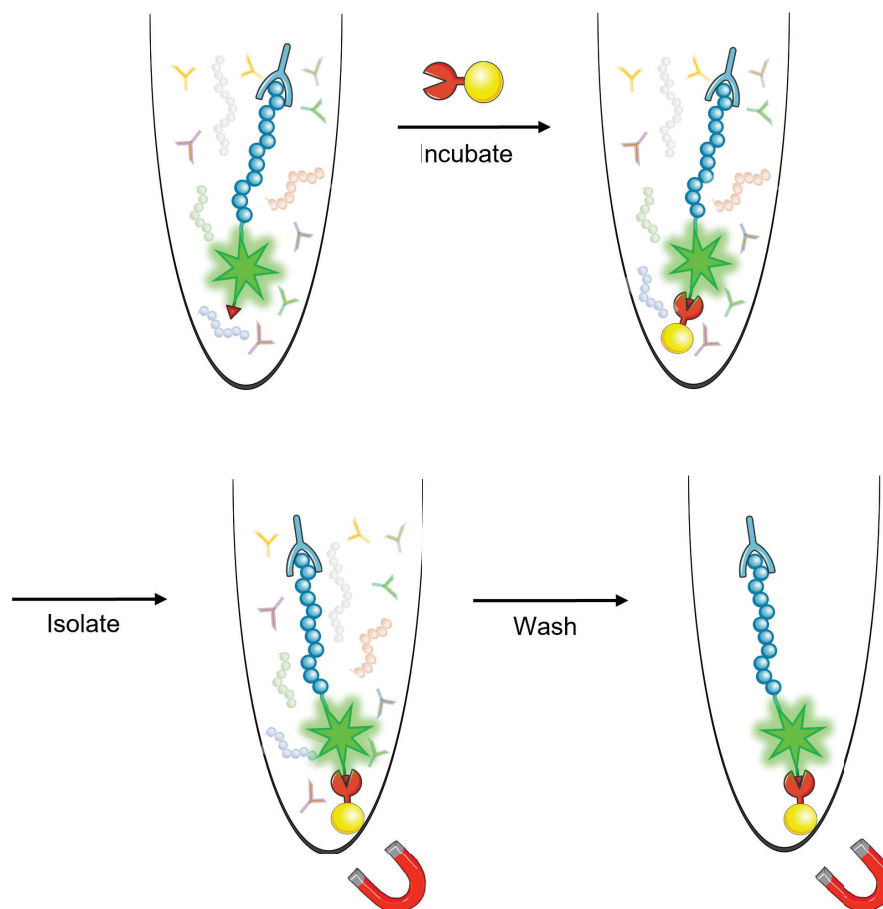


Figure 4.5: After identification of binding by fluorescence, the cell contents can be incubated with pull-down-partner-decorated magnetic solid beads. The bound peptide can be isolated using external magnets, after which unbound contents can be washed away to yield the isolated peptide alone (not shown) or the peptide-receptor complex.

4.3.2 Probe design

The probe was designed such that it contains: a fluorescent core used for easy identification; a carboxylic acid group as an attachment point that is compatible with solid phase peptide synthesis (SPPS); and a biotin group as a pull-down partner to streptavidin decorated magnetic beads.

A 4-amino-1,8-naphthalimide based core was chosen as the fluorescent scaffold. Naphthalimides are known for their high quantum yield, excellent photo-stability, and high Stokes' shifts, making them an excellent choice for biological studies.³²⁹ Additionally, naphthal-

imides feature two positions available for functionalisation: the imide position, that was functionalised with the carboxylic acid; and the 4 position, that was functionalised with the biotin group as shown in Figure 4.6.

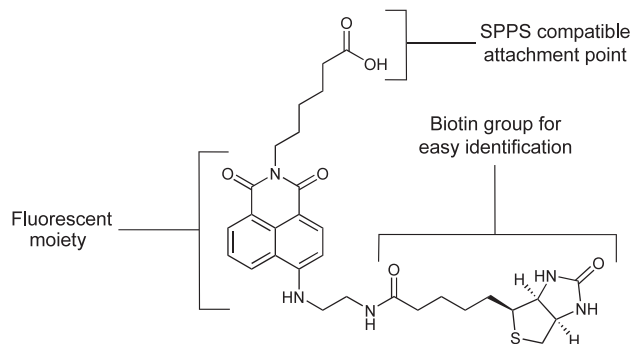
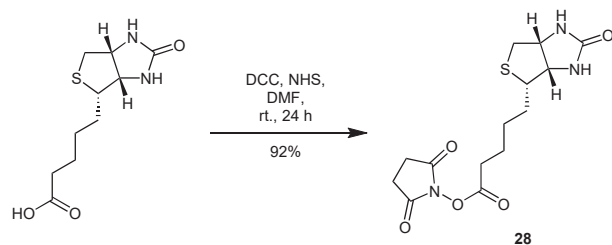


Figure 4.6: Chemical structure of the designed probe.

The carboxylic acid moiety enables the conjugation of the molecule to either the N-terminus or a lysine residue of a peptide. The choice attachment site should be determined by the specific amino acids that are essential for receptor activity, and must be considered carefully during the design phase of the peptide. This is crucial in ensuring that the peptide retains its desired biological activity and properties after modification, and also in avoiding any undesirable effects or interference with the receptor binding.

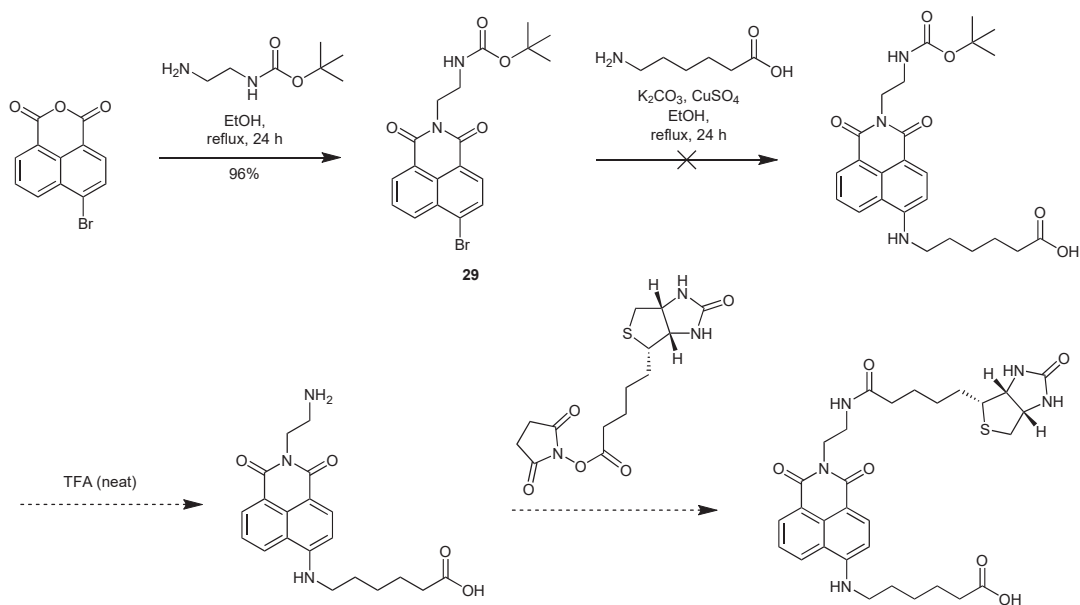
4.3.3 Synthesis

Synthesis of this probe began with the activation of the carboxylic acid of D-biotin *via* esterification with *N*-hydroxy succinimide to give **28**. This was achieved with the use of dicyclohexylcarbodiimide (DCC) as a coupling reagent as shown in Scheme 4.1. This activated ester will later be used to couple the biotin onto a free amine installed on the naphthalimide scaffold.



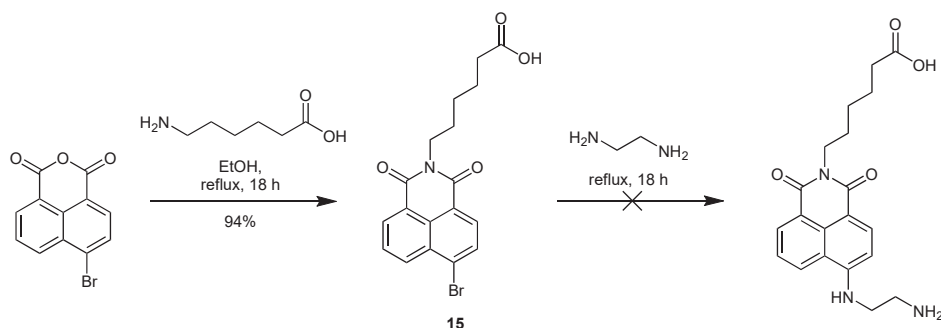
Scheme 4.1: Succinimidyl activation of biotin.

Initial attempts at synthesising the probe are outlined in Scheme 4.2. The amine handle was arbitrarily installed at the imide position of the naphthalimide, as the position of the handle is not expected to affect the sensors performance. The synthesis began with 4-bromo-1,8-naphthalic anhydride as the starting material. The anhydride position was substituted with mono-protected *N*-Boc-ethylenediamine to obtain **29**. Subsequently, substitution of the bromo position was carried out using 6-aminohexanoic acid via a copper(II)-catalyzed S_NAr reaction.³³⁰ However, this resulted in a mixture of products, possibly due to the deprotection of the Boc-protected amine under high temperatures, and basic conditions, leading to unwanted side reactions. As the reaction mixture was challenging to purify, an alternative synthetic route was explored.



Scheme 4.2: Attempted synthesis of the probe with the biotin at the imide position.

Another attempt at substitution of the aromatic bromine is shown in Scheme 4.3. In this case, 6-aminohexanoic acid was switched to the imide position of the naphthalimide core since this product could be synthesised in larger amounts at a cheaper cost. The synthesis began with substitution of the anhydride position of 4-bromo-1,8-naphthalic anhydride with 6-aminohexanoic acid to yield **15**. Nucleophilic aromatic substitution (S_NAR) of the bromine at the 4-position was attempted in neat ethylenediamine overnight. It was speculated that the vast excess of ethylenediamine would allow substitution to occur under relatively mild conditions to avoid side reactions. However, this was not the case, and it resulted in a mixture of products that could not be isolated or determined.



Scheme 4.3: Attempted S_NAR of the 4-position with ethylenediamine.

Attention was turned to finding suitable reaction conditions to substitute the aromatic bromine successfully. One equivalent of *N*-*boc*-ethylenediamine was reacted with **15** under the various conditions outlined in Table 4.1. Condition 1 was adapted from a known literature procedure where the two reagents were refluxed in ethanol in the presence of base and a copper sulfate as a catalyst.³³¹ These conditions resulted in a mixture of products. This may have been caused by inadvertent deprotection of the amine due to the high temperatures and basic conditions, resulting in side reactions. As such, this condition was reattempted in the absence of base, as given in condition 2. However, this reaction was also proved unsuccessful, resulting in a mixture of products. Condition 3 was adapted from a known literature procedure where *N*-*boc*-ethylenediamine was added dropwise to a solution of **15** in acetonitrile.²⁸⁸ No reaction occurred under these conditions. This reaction was attempted again using a different polar aprotic solvent, tetrahydrofuran, as given in condition 4. When left stirring overnight at room temperature, no reaction occurred. The

temperature was then increased to reflux, as shown in condition 5, but no reaction took place despite this. As such, an Ullmann-type reaction was attempted, as given in condition 6, where copper iodide was added as a catalyst and left to stir for a further 18 h. However, this also did not result in any reaction. A small amount of base, triethylamine, was then added to assist in deprotonating the free amine of *N*-boc-ethylenediamine, making it more likely for nucleophilic attack to occur. Despite this, there was still no reaction that occurred under the conditions shown in condition 7. The final reaction condition, 8, was adapted from a literature procedure where the reactants were stirred in dimethylsulfoxide at 70 °C overnight.³³² This reaction lead to the successful substitution of the aromatic bromine with *N*-boc-ethylenediamine.

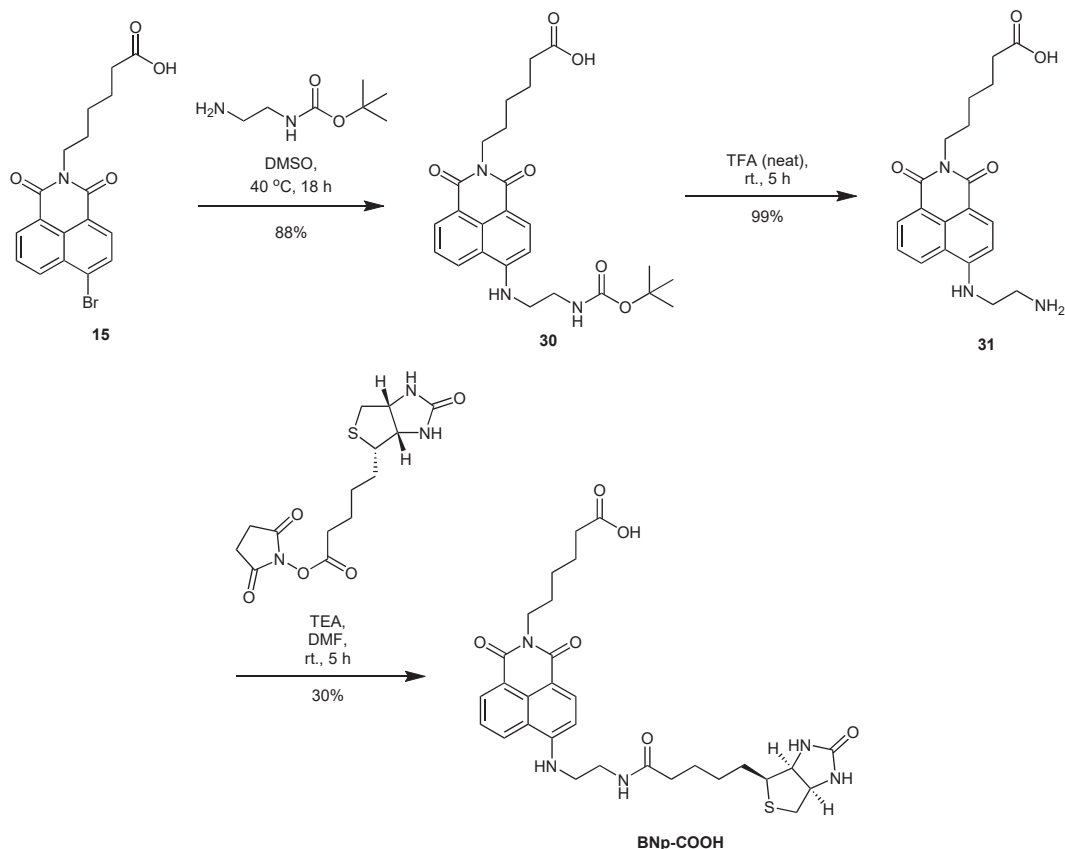
Table 4.1: Reaction conditions trialled for the synthesis of **30**.

	Additives	Solvent	Temperature	Time	Outcome
1	CuSO ₄ , KCO ₃	EtOH	reflux	48 h	Mixed products
2	CuSO ₄	EtOH	reflux	48 h	Mixed products
3	-	MeCN	rt.	18 h	No reaction
4	-	THF	rt.	18 h	No reaction
5	-	THF	reflux	18 h	No reaction
6	CuI	THF	reflux	18 h	No reaction
7	CuI, TEA	THF	reflux	18 h	No reaction
8	-	DMSO	70 °C	18 h	88% yield

The synthesis of the final probe was continued as shown in Scheme 4.4. Substitution of the bromine was performed in DMSO at 70 °C, as outlined above, to yield **30**. The Boc group was deprotected under acidic conditions to yield **31**. The free amine was subsequently

reacted with the activated ester, **28**, to give the final probe, **BNp-COOH**.

The pre-activated acid was chosen over standard solution-phase amide coupling conditions in order to avoid using coupling reagents. This was done to prevent activation of the unprotected acid moiety on **31**, and subsequent formation of side products such as polymerisation of the naphthalimide. Additionally, this route eliminates the need for extra protection and deprotection of the free acid, reducing the number of steps in this synthesis.



Scheme 4.4: Synthesis of **BNp-COOH**.

Fluorescence studies of **BNp-COOH** were performed to investigate the photophysical properties of the probe. Excitation and emission spectra were taken in HEPES buffer (pH 7.4) to emulate the buffered environment of the cellular conditions that this probe would be used in. The spectra show excitation and emission maxima at 455 and 547 nm, respectively (Figure 4.11). This indicates a suitable fluorescent profile for analysis with microscopy or flow cytometry as both the standard 405 nm standard laser lies within the excitation range, with minimal overlap from the emission. This will minimise the risk of excitation bleed-

through (fluorescence signal caused by the excitation wavelength), while maximising the emission collection window.

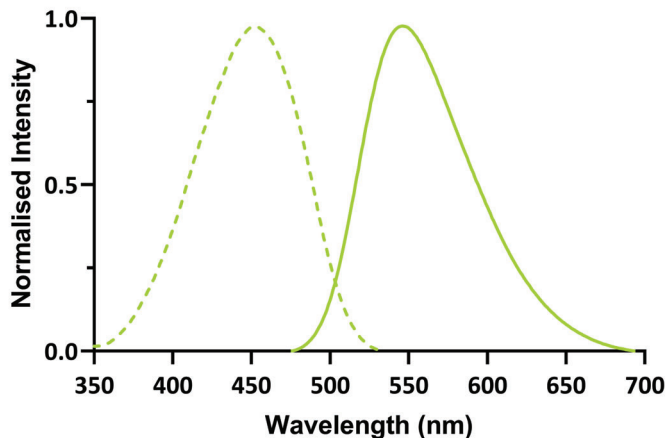


Figure 4.7: Normalised excitation (dashed) and emission (solid) spectra of **BNp-COOH** in 20 uM HEPES buffer ($\lambda_{ex} = 455$ nm, $\lambda_{em} = 547$ nm).

4.4 Proof of concept peptide

The first step in validating the proposed method is to test it with a known peptide and its receptor. In this case, the peptide of interest is octreotide, which is known to bind to the somatostatin receptor subtype 2 (SSTR2). By using a known peptide-receptor pair the feasibility of this method can be evaluated.

4.4.1 Targeting peptide

Somatostatin is a natural peptide known for its multiple targets in central nervous system, pituitary, endocrine and exocrine pancreas, as well as the gastrointestinal tract which are summarised in the review.³³³ This peptide has been used in targeted therapy based on findings that many neuroendocrine tumours expressed a high density of somatostatin receptors when compared to non-tumour cells.³³⁴ One example is a certain type of breast cancer which is characterised by the over-expression of somatostatin receptors, which are not found in healthy breast tissue.³³⁵ It has been found that somatostatin produces an inhibitory effect on somatostatin positive cancers, but its use as a therapeutic agent is limited by its short half-life in plasma. Therefore, stable peptide analogues are required for successful

application in clinical treatment.

Octreotide is a synthetic analogue of somatostatin which is more active against somatostatin receptors and has a longer half life (Figure 4.8).³³⁶ This synthetic analogue has been shown to produce anti-proliferative effects in somatostatin receptor positive breast cancers.³³⁷ Currently, octreotide is used as a therapeutic agent for these types of breast cancers.

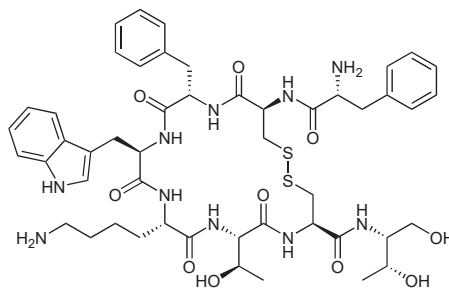


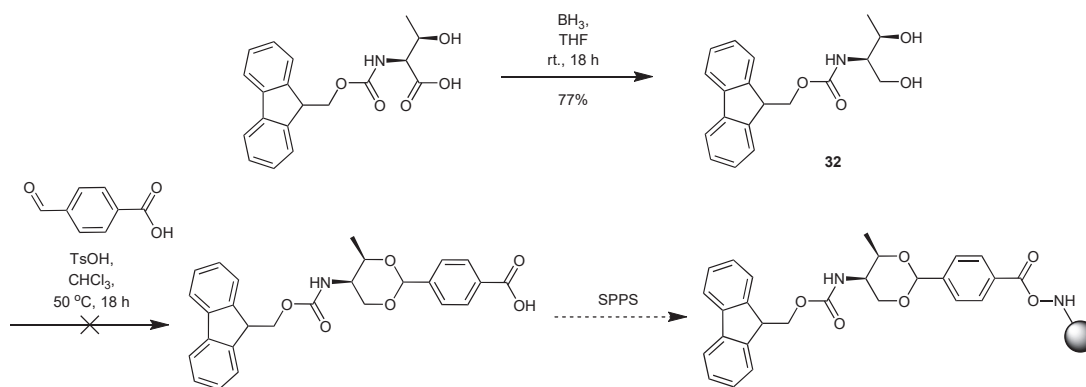
Figure 4.8: Chemical structure of octreotide.

4.4.2 Synthesis of fluorescent analogues of Octreotide

In general, peptide synthesis is performed on a solid resin, where the peptide is anchored through the carboxylic acid of the first amino acid.³³⁸ In the case of octreotide, the C-terminus exists as a hydroxyl rather than a carboxylic acid. As such, the first amino acid must be modified in such a way that will yield the desired peptide alcohol. Initial attempts at synthesising octreotide as a peptide alcohol were performed according to a known literature procedure. This procedure used 4-carboxybenzaldehyde to anchor Fmoc-Thr-ol to rink amide resin by forming an acetal between the two hydroxy groups, while tethering a carboxylic acid for use as a link between threonine and the resin. The acetal is shown to be stable under peptide synthesis conditions but is unstable under acidic conditions. This is advantageous since under the acidic conditions required to cleave the peptide from the resin, the acetal would also be cleaved, leaving behind the diol derivative, and hence a hydroxyl group at the c-terminus. This method was described by Hsieh *et al.* as a novel, facile procedure for the rapid synthesis of octreotide in high yields when compared to previously reported methods such direct amino alcohol conjugation.³³⁹

The attempted synthesis following this method shown in Scheme 4.5. To begin, the car-

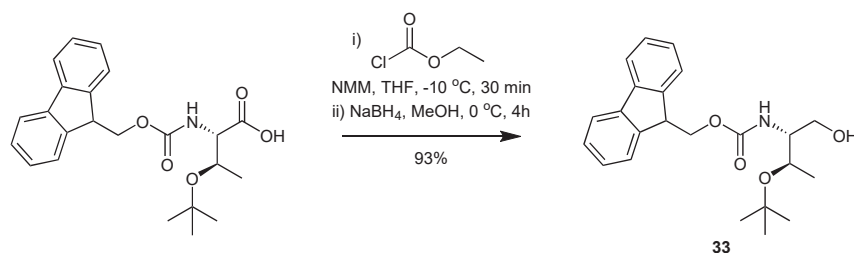
boxylic acid of Fmoc-Thr-OH was reduced to the primary alcohol to yield **32**. The hydroxyl side chain and the alcohol were subjected to acetal formation with 4-carboxybenzaldehyde using 4-toluenesulfonic acid as a catalyst. This reaction yielded the desired product. Unfortunately, the acetal group proved to be very unstable during purification and the product could not be isolated despite multiple attempts. This may have possibly been due to the acidic nature of the silica used in column chromatography. As such, purification was then attempted in a basic buffer (0.1% TEA) in order to mitigate degradation of the acetal. However, due to the presence of a free carboxylic acid, the product did not progress beyond the baseline of the silica, thus purification was not possible. Alumina was trialed as an alternative stationary phase for chromatography due to its neutral pH which may reduce the likelihood of acid-induced acetal degradation during purification. Unfortunately, product separation with alumina could not be achieved and purification could not be accomplished. As such, a different method of generating a terminal alcohol on resin was considered.



Scheme 4.5: Attempted synthesis of acetal bound threonine for conjugation to rink amide resin.

2-Chlorotrityl chloride resin is a common resin used for its ability to anchor several functional groups such as carboxylic acids, thiols, amines and alcohols. In this method, the amino alcohol, Fmoc-Thr(tBu)-ol, was anchored directly to the resin through the hydroxyl group which would yield the desired terminal alcohol upon peptide cleavage from resin. Thus, the synthesis began by reducing the carboxylic acid of Fmoc-Thr(tBu)-OH without cleaving the tert-butyl group protecting the hydroxyl side chain. This was achieved using ethyl chloroformate in the presence of base to form an anhydride at the carboxylic acid that

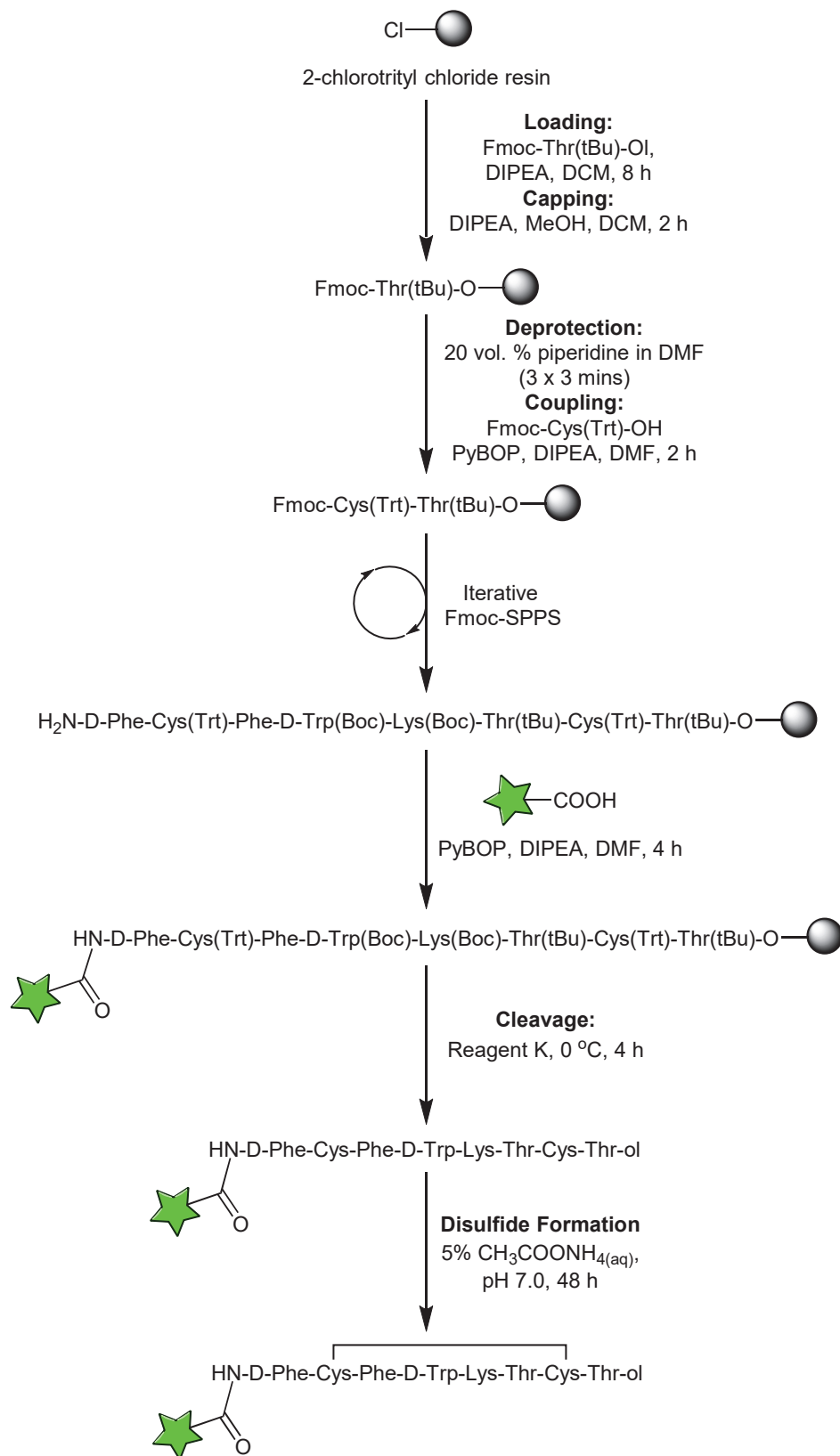
was immediately treated with methanol in the presence of sodium borohydride to yield the desired alcohol as **33**, as shown in Scheme 4.6



Scheme 4.6: Reduction of Fmoc-Thr-OH to Fmoc-Thr-ol.

Solid phase peptide synthesis was performed according to Scheme 4.7. **33** was anchored to 2-chlorotrityl chloride resin in the presence of *N,N*-diisopropylethylamine (DIPEA). The resin was then capped with methanol to prevent unreacted sites from reacting with subsequent amino acids. Peptide elongation was performed according to standard protocols, following an iterative process of Fmoc deprotection followed by amino acid coupling using standard coupling reagents PyBOP and DIPEA. Octreotide was synthesised with the following sequence: H₂N-D-Phe-Cys-Phe-D-Trp-Lys-Thr-Cys-Thr-ol. The carboxylic acid bearing fluorescent probe, **BNp-COOH**, was then coupled to the N-terminus using PyBOP as a coupling reagent in the presence of DIPEA. The fluorescent peptide was cleaved and side chains were deprotected in a single step.

Initial attempts at cleaving the peptide were performed with a standard cleaving cocktail consisting of TFA/TIPS/H₂O (17:1:1, v/v/v). However, this did not yield the desired peptide upon cleavage (Figure S3). This may have been due to the presence of cysteine and tryptophan residues that are known to be sensitive to side chain oxidation, resulting in multiple impurities upon treatment with the cleavage cocktail. An alternative cleavage cocktail, known as reagent K, was trialled instead, as it is designed for use with sensitive residues such as cysteine, methionine, tryptophan and thionine. Reagent K is comprised of TFA/H₂O/phenol/thioanisole/1,2-ethanediol (33:2:2:2:1 v/v/v/v/v). The peptide was stirred in this cocktail at $0\text{ }^\circ\text{C}$ for 4 h. This yielded the desired peptide, as verified by LCMS and analytical HPLC (Figure S4).



Scheme 4.7: SPPS of fluorescently labelled octreotide.

To determine whether the fluorescence output of **BNp-COOH** would be affected by peptide conjugation, the fluorescent properties of crude **BNp-COOH**-tagged octreotide in HEPES buffer were assessed. The spectra show an excitation and emission maximum at 452 and 545 nm, respectively (Figure 4.9). This indicates negligible changes to the fluorescent properties of **BNp-COOH** upon peptide conjugation, and confirmed that the 405 nm standard laser is still appropriate for testing of the peptide conjugated fluorophore.

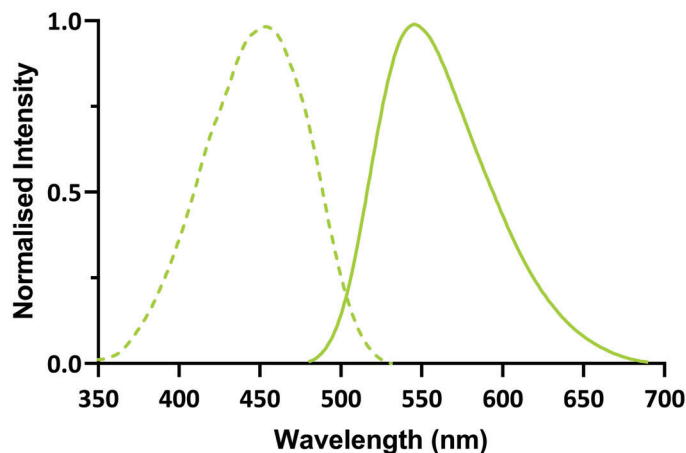
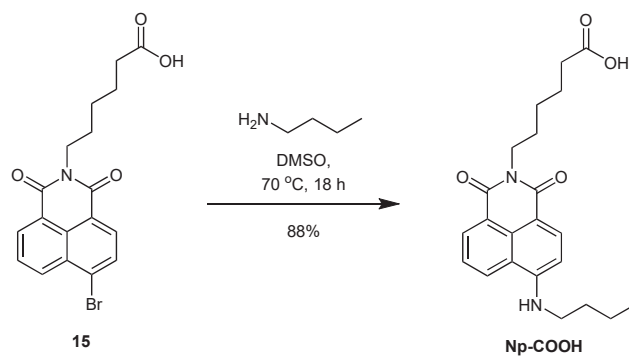


Figure 4.9: Normalised excitation (dashed) and emission (solid) spectra of **BNp-COOH**-tagged octreotide in 20 uM HEPES buffer ($\lambda_{ex} = 452$ nm, $\lambda_{em} = 545$ nm).

4.5 Synthesis of a control

A control fluorescent probe was also synthesised in which the biotin group of **BNp-COOH** was replaced with an alkyl group. This control probe was synthesised to investigate whether peptide binding would be affected by the presence of the biotin group as it is known to bind to surface receptors present on many cell lines of interest.

The synthesis began with **15** whose synthetic details are outlined in section 2.9, Scheme 2.10. S_NAR of the bromine with butyl amine was performed in dimethylsulfoxide for 18 h to yield the final control probe, **Np-COOH** as shown in Scheme 4.8.



Scheme 4.8: Synthesis of control probe, **Np-COOH**.

The fluorescence spectra of **Np-COOH** were measured in HEPES buffer. This revealed excitation and emission maxima of 451 and 548 nm respectively, which is within a similar range to **BNp-COOH**.

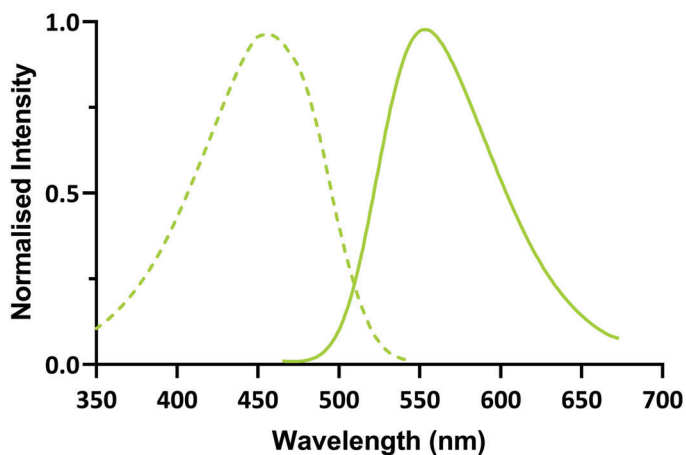


Figure 4.10: Normalised excitation (dashed) and emission (solid) spectra of **Np-COOH** in 20 μM HEPES buffer ($\lambda_{ex} = 451\text{ nm}$, $\lambda_{em} = 548\text{ nm}$).

After this, the probe was conjugated to octreotide in the same manner described in scheme Scheme 4.7. As previously described, the fluorescence properties of crude **Np-COOH** conjugated to octreotide were assessed to determine any effect of the peptide on the fluorescence output. The excitation and emission spectrum maxima of 450 and 540 nm respectively. This shows negligible changes to the fluorescence spectra of **Np-COOH** upon conjugation to the peptide. Additionally, the excitation and emission spectra are within

the same range as those of **BNp-COOH**, making **Np-COOH** a suitable control. It is important to use the same laser and collection window parameters for both the control and test sample to ensure valid comparisons. This is because any differences observed in the fluorescence signal can then be attributed to the presence or absence of the peptide-protein complex, rather than to differences in the experimental setup.

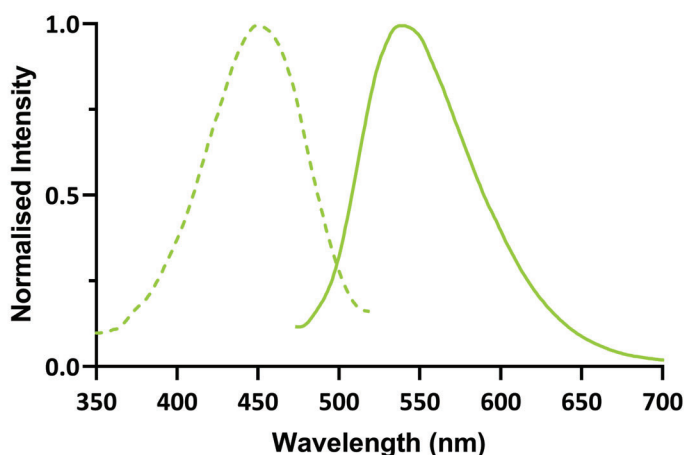


Figure 4.11: Normalised excitation (dashed) and emission (solid) spectra of **NpCOOH** labelled octreotide in HEPES buffer (20 mM).

4.5.1 HPLC challenges

To begin the purification process, initial analytical HPLC tests were performed on **Np-COOH** labelled octreotide in the non-cyclised form in order to find the optimal method. The linear peptide was dissolved in 50% acetonitrile in water and analytical HPLC was performed. The trace exhibited a peak at 38 minutes with a shoulder at a similar retention time when assessed at 20 - 100% acetonitrile with water in 0.1% TFA (Figure 4.12A). Using this same sample, various methods were set up to run overnight in order to optimise the separation of the peaks. However, upon inspection the following day, a new peak was observed at 2 minutes which slowly increased in intensity over time (Figure 4.12B). After 24 h assessment of the same sample again showed that this became the main peak in the HPLC trace (Figure 4.12C). This was thought to be the result of low solubility which is a possibility considering this fluorophore is planar, and contains multiple alkyl chains. As such, a fresh sample of **Np-COOH**-labelled octreotide was prepared in 100% DMSO to

ensure solubility. Assessment of various HPLC methods was attempted with this solution, however, the same issue was presented and after 24 h the HPLC trace gradually transitioned into a large peak at 2 min which occur on both C18 and C8 silica (Figure S5).

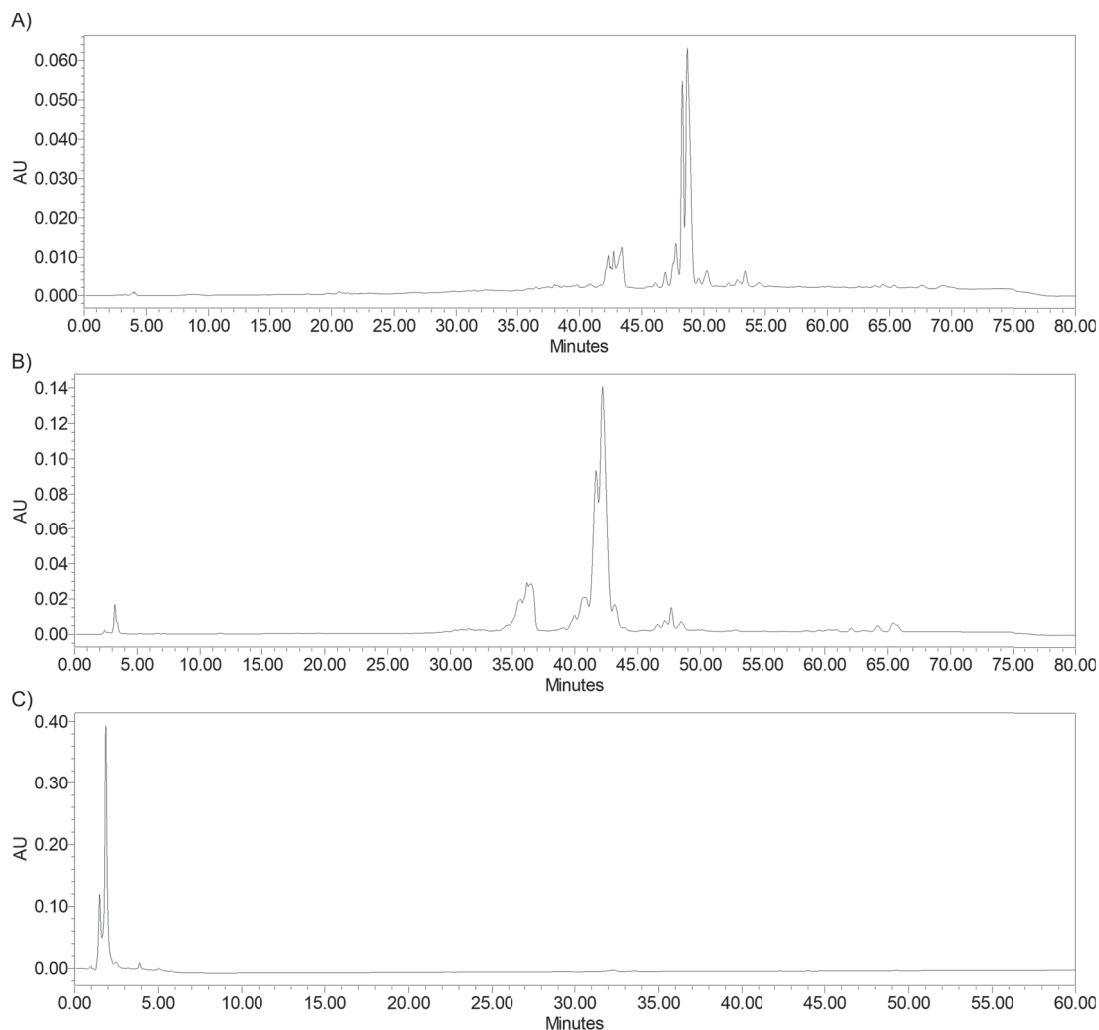


Figure 4.12: HPLC trace of **Np-COOH**-conjugated octreotide assessing the same sample after A) 0 h in 20 - 60% B) 2 h in 30 - 60%, and C) 24 h at 20 - 100% acetonitrile in water with 0.1% TFA over 60 min.

A possible explanation for the observed increase in peak intensity at 2 min could be attributed to the high concentration of non-cyclised peptide which may have undergone intermolecular disulfide bridge formation, resulting in the formation of a polymerised product that elutes at 2 min. As such, the crude peptide was cyclised prior to HPLC in order to eliminate this possibility. Disulfide bridge formation was performed at a low concentration

(< 5 mg/mL) in a solution of ammonium bicarbonate (50 mM) at pH 8 for 4 days. The low concentration of peptide in solution is necessary to reduce the likelihood of intermolecular cross-linkage. After this, the analytical HPLC was performed. The product was dissolved in DMSO and the solution was subjected to analytical HPLC. As before, multiple methods were trialled in an attempt to separate the shoulder peak from the product. However, the problem persisted and the peak at 2 min continued to increase over time, until it predominated the HPLC trace (Figure 4.13). The reason for the degradation could not be concluded from these tests.

Eventually, by utilising a fresh sample every few hours, a suitable method was found. By using an isocratic gradient of 38% acetonitrile in water, the shoulder peak could be resolved from the product. It was theorised that preparing a fresh sample and using it immediately in purification would lead to isolation of the peptide without the risk of degradation. Therefore, the crude sample was dissolved in DMSO and immediately injected into the preparative HPLC. Unfortunately, the HPLC system malfunctioned with the sample in it, and the product could not be recovered. Due to time constraints, the product was not resynthesised. Instead, molecular dynamic simulations were employed to assess the feasibility of the proposed method.

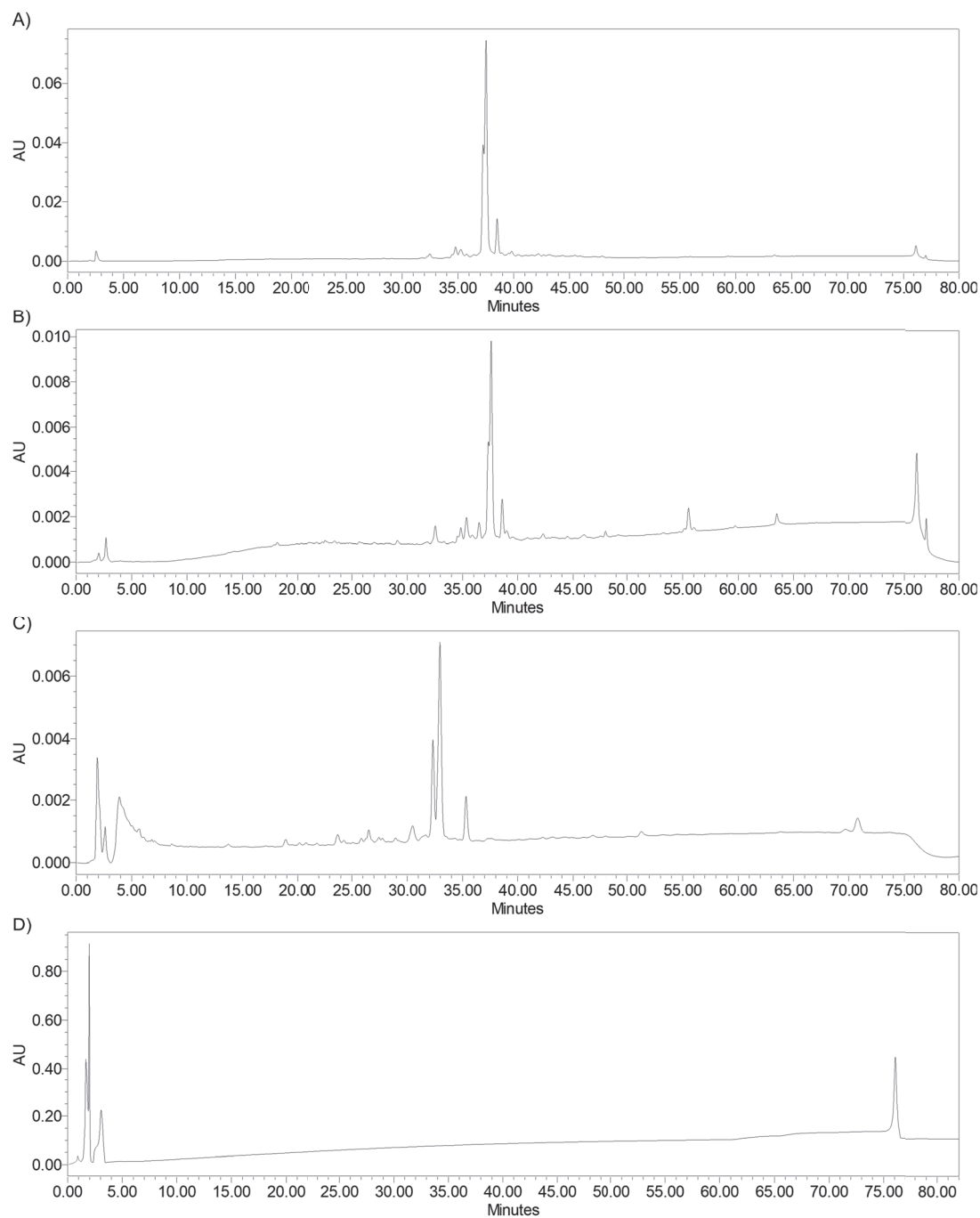


Figure 4.13: HPLC trace of cyclised **Np-COOH**-conjugated octreotide assessing the same sample after A) 0 h in 0 - 100% B) 6 h in 0 - 100%, C) 18 h in 30 - 70%, and D) 24 h in 0 - 100% acetonitrile in water with 0.1% TFA over 60 min.

4.6 Planned biological studies

Flow cytometry training was undertaken in preparation for studies of the binding assay. SSTR2- positive and negative cell lines, PC-3 and Huvec-TERET2 respectively, were identified using the Human Protein Atlas (<https://www.proteinatlas.org/>). The cells were obtained and readied for cell culture. These cells were to be assessed by flow cytometry using **BNp-COOH**-octreotide to determine the specificity of the binding between receptor-positive and receptor-negative cell lines. Further, collaborations with Professor Joel Mackay were set up for the assessment of the biotin-streptavidin pull-down assay, utilising this probe.

4.7 Molecular dynamics

Molecular dynamics (MD) is a computational technique used to predict, and analyse the interaction between a biomolecule and its target. In these simulations, the ligand and receptor are modelled as a set of interacting atoms, and their movements are tracked over time. During the simulation, the ligand and receptor can move, rotate, and interact with each other and with the solvent molecules surrounding them. From this, MD can be used to calculate the thermodynamic and kinetic properties of the system, such as the binding affinity.

Being a somatostatin analogue, octreotide is known to bind to the somatostatin receptor (SSTR) family with strongest binding observed for subtype 2, SSTR2, followed by weaker binding to SSTR5. MD simulations were employed to assess whether the conjugation of the designed probe, **BNp-COOH**, would affect the binding of octreotide within the SSTR2 binding pocket. The following work was performed with the assistance of Dr. Bryson Hawkins.

4.7.1 MD binding simulations

Binding simulations were first performed on untagged octreotide in order to assess the original binding conformation. Simulations were performed on octreotide docked within the SSTR2 binding pocket that was determined from X-ray crystallographic data obtained from the protein data bank (PDB). The simulations showed that the complex of octreotide within the SSTR2 receptor reaches equilibrium at 20 ns, and remains stable within the

binding pocket thereafter (Figure 4.14).

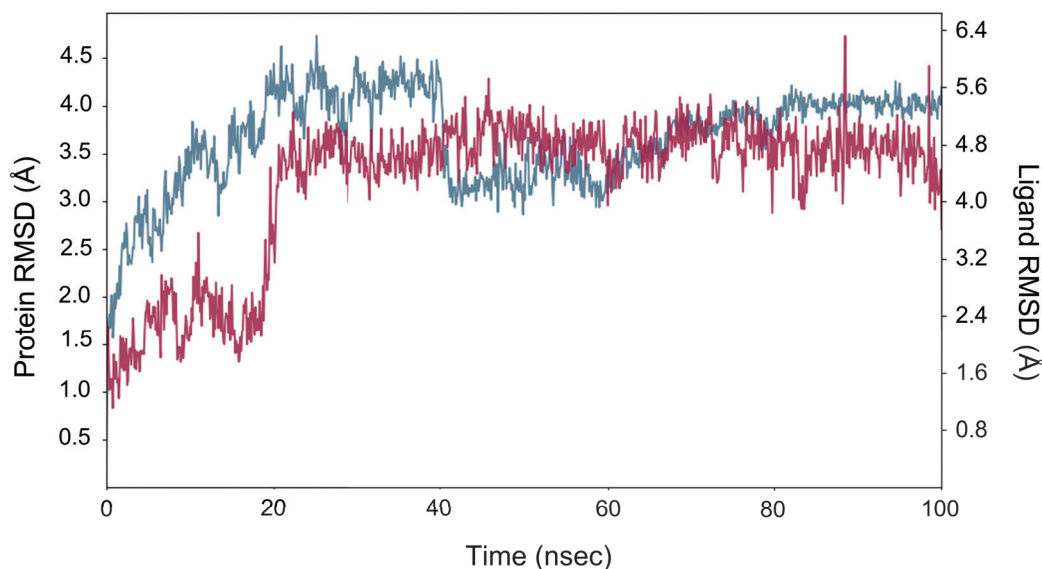
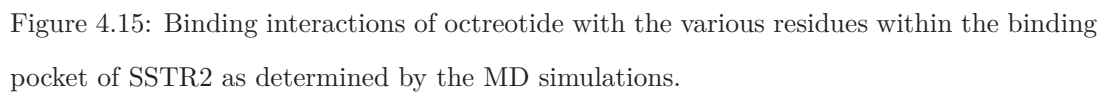


Figure 4.14: Changes in the RMSD of octreotide and the SSTR2 protein over time during MD binding simulations.

The residues responsible for the stabilisation of octreotide within the binding pocket were: ASN276, which contributed two key H-bonds to lock octreotide in its position for 99% of the simulation time; GLN126, which interacts with the NH_3^+ group of the Lys in octreotide; and ASP122, which also interacts with the NH_3^+ of the Lys residue. ARG184 served as a gatekeeper, its cation- π interaction with the D-Phe end prevented the whole peptide from dissociating from the binding pocket. Apart from these direct interactions, several dynamic water bridges also provided energetically favourable contact between the receptor and ligand, including MET119, ASP122, THR194, ILE195, SER279 and THR301 (Figure 4.15). This is comparable to results found in literature.^{340,341} Molecular mechanics with generalised Born and surface area solvation (MMGBSA) calculations based on the trajectory at 50 ns gave a binding free energy of -96.13 kJ, indicating good binding.



Next, binding simulations were performed on **BNp-COOH** conjugated octreotide within the SSTR2 binding pocket. The simulations showed that the complex of **BNp-COOH** conjugated octreotide within the SSTR2 receptor fluctuates until about 40 ns, when an equilibrium is reached (Figure 4.16).

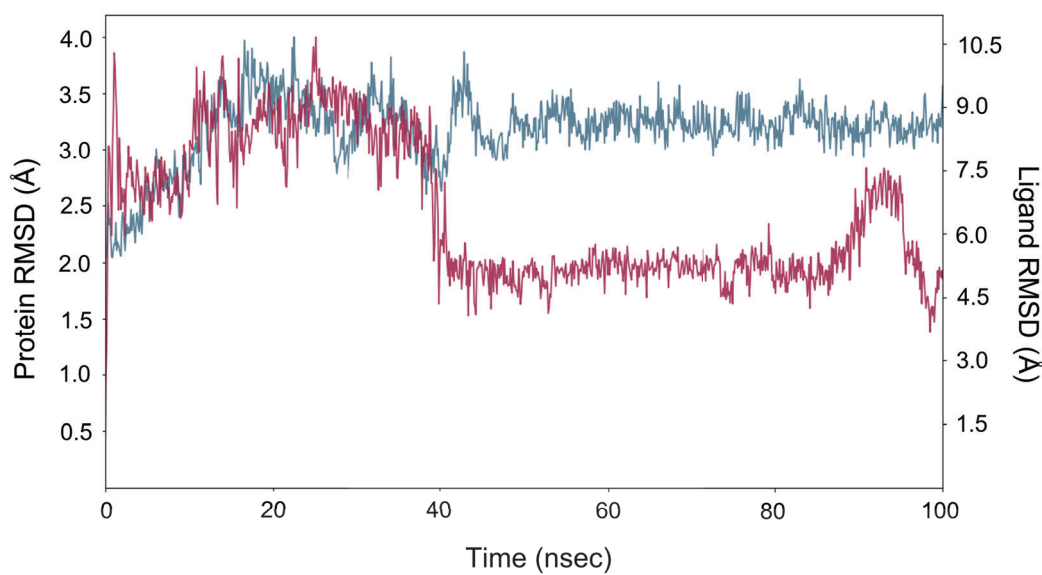


Figure 4.16: Changes in the RMSD of **BNp-COOH** conjugated octreotide and the SSTR2 protein over time during MD binding simulations.

Further, the simulations showed that the key interactions of octreotide remained consistent. However, by 50 ns the probe itself folds over and interacts with the receptor (Figure 4.17). This occurs through multiple pi-pi stacking with the aromatic rings of the naphthalimide core. ILE195 also appears to exhibit hydrogen bonding with the biotin moiety (Figure 4.18). MMGBSA calculations reveal a binding free energy of -187.46 kJ indicating that these additional binding interactions are energetically favourable.

These preliminary MD studies on the binding interaction between **BNp-COOH** conjugated octreotide and the SSTR2 receptor indicate that the native interactions of the peptide within the binding pocket remain consistent regardless of fluorescent probe conjugation. The weak interaction of the hydrogen bonding between the biotin and SSTR2 receptor residue would theoretically be negated by the high binding affinity of streptavidin to the biotin moiety, but this will need to be assessed and confirmed experimentally.

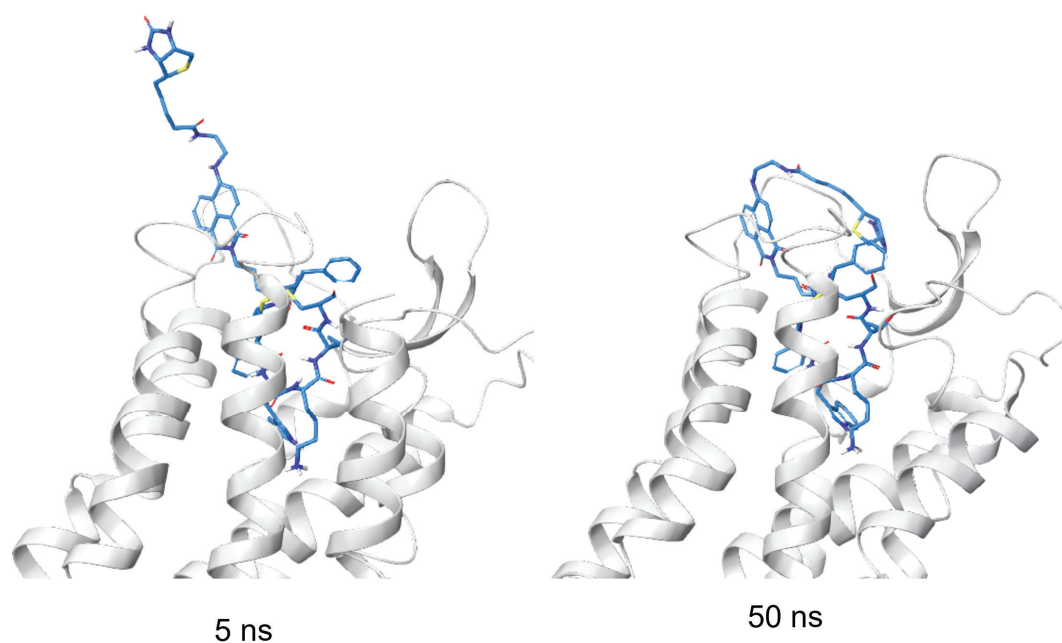


Figure 4.17: Binding conformation of **BNp-COOH** conjugated octreotide within the SSTR2 binding pocket at 5 and 50 ns.

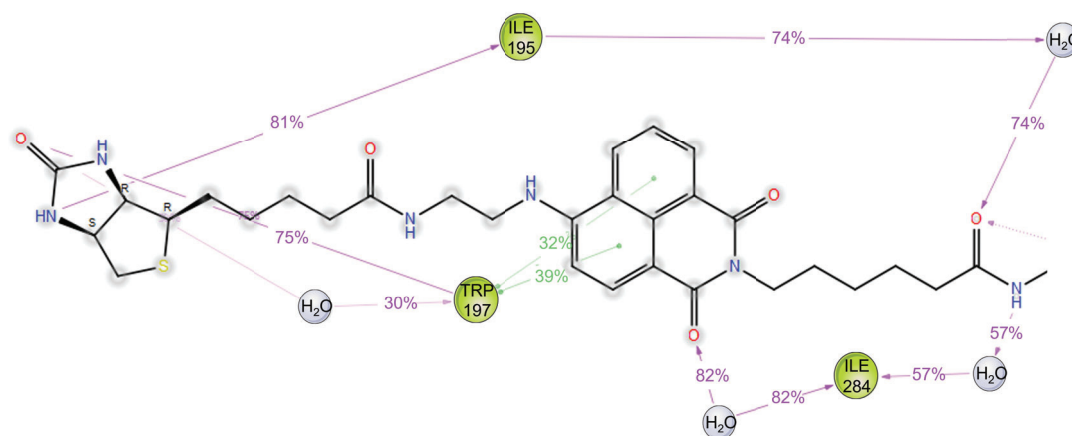


Figure 4.18: Binding interactions of **BNp-COOH** when bound to octreotide (not shown) with the various residues of the SSTR2 protein as determined by the MD simulations.

4.8 Conclusion

This work outlines the steps taken toward the development of a novel method to detect peptide-receptor binding interactions. The fluorescent probe was designed with two functions in mind: the ability to be incorporated into peptides synthesised by SPPS, and the incorporation of a pull-down moiety to allow for easy isolation of the peptide-receptor complex after binding. As such, **BNp-COOH** was designed to contain a carboxylic acid group for peptide synthesis and a biotin as a pull-down partner for the well known Streptavidin binding partner. The probe was shown to be successfully incorporated into a peptide of interest, octreotide, as confirmed by LCMS and analytical HPLC. However, this labelled peptide could not be isolated due to instrument failure.

Fluorescence studies revealed that **BNp-COOH** possesses excellent fluorescence properties that were well suited for use with standard microscope or flow cytometry lasers. Furthermore, it was observed that conjugating the probe to a peptide of interest did not compromise its fluorescence properties.

Additionally, to evaluate the specificity of **BNp-COOH**, a control probe with an alkyl chain replacing the biotin moiety, **Np-COOH**, was synthesised. The control probe exhibited a fluorescence profile similar to that of **BNp-COOH**, demonstrating its utility as a control. Furthermore, the compatibility of this control probe with SPPS was also confirmed.

The feasibility of the proposed method was initially evaluated through MD studies. These preliminary investigations indicated that the interaction between octreotide and its receptor is likely unaffected by the presence of the **BNp-COOH** probe. These results are encouraging, and further experimental validation of this approach is warranted.

Chapter 5

Conclusions

Fluorescence spectroscopy has proven to be a powerful technique in the field of biology, providing a robust imaging tool for investigating and understanding cellular function at the molecular level. Small molecule fluorophores, in particular, have become a prominent tool in this field as they offer several advantages over traditional biochemical methods. These fluorophores are relatively easy to synthesise and are characterised by their small size, and are generally biocompatible. By utilising the advantages of fluorescence imaging, valuable insights of peptide and protein interactions have been made, leading to a greater understanding of biological processes. In this work, progress was made toward the development of novel fluorescent methods that would allow for the study of peptide and protein interactions *in vitro*.

5.1 Development of a fluorescent sensor for multimodal mass spectrometry and fluorescence imaging

Mass spectrometry imaging is an effective method for analysing specific compounds in a biological sample. However, this technique suffers limitations such as the inability to detect larger molecules, as well as being a non-quantitative. By combining this technique with fluorescence imaging, novel interactions have been assessed, highlighting the utility of this excellent combination of techniques. In this work, **CyMT** was designed as a small molecule with properties that allowed it to be utilised in a multimodal mass spectrometry and fluorescence imaging protocol. The probe was capable of multimodal imaging in isolated protein

samples. A second probe, **NpMT**, was also synthesised but due to time constraints, could not be isolated and tested.

With the purification and testing of **NpMT**, further investigations of potential applications should be made. This would include exploring the use of multiplexed experiments involving both **CyMT**, and **NpMT** to assess the interplay between two biomolecules of interest. This innovative approach combines the unique properties of both probes including their distinct m/z values, and unique emission profiles which would enable multiplexed-multimodal imaging to be performed using a single sample. To the best of my knowledge, this type of imaging protocol has yet to be explored.

In this work, the bioorthogonal reaction between a succinimidyl ester and primary amine was exploited for specific biomolecule labelling. However, to enhance the versatility and applicability of this probe to a wider range of biomolecules, other biomolecular attachments should be considered. These can include but are not limited to maleimide orthogonality toward free cysteine residues, and even the incorporation of protein tagging substrates for systems like HaloTag. Additionally, the use of different bioorthogonal groups on each probe can allow for simultaneous and specific tagging of different biomolecules without the risk of cross-reactivity, and eliminating the need for sequential tagging steps.

Expanding on this, the successful demonstration of a small molecule fluorophore in a multimodal setup provides the basis for exploring the use of different types of fluorophores. For instance, the use of a fluorophore with sensing capabilities could provide an opportunity to unveil even more valuable information about the properties and behaviour of the biomolecule of interest. For example, this could include a reversible redox sensor that would be used to study the redox environment of the biomolecular interaction.

By demonstrating the capabilities of a small molecule sensor in a multimodal setup, new opportunities have been opened up for the development of novel imaging protocols for the exploration of previously undiscovered biomolecular interactions.

5.2 Investigation of novel redox sensors for protein tagging

Reactive oxygen species are known to play a significant role in both health and disease, yet many of the underlying mechanisms remain poorly understood. The visualisation of redox environments within sub-cellular organelles has been made possible by the use of targeted

fluorescent sensors. However, to gain a more comprehensive understanding of the role of ROS in biology, it is essential to measure the redox state of any sub-cellular location within a cell.

Previous work on this development has demonstrated the utility of a fluorescent redox sensor for the HaloTag system, **FlavinTag**, which uncovered a reductive event at a copper import protein. With the success of this approach, efforts were focused on identifying additional suitable sensors. To evaluate the effectiveness of these sensors, computational studies were conducted as a preliminary step.

The computational studies revealed that the ratiometric redox sensor, **FCR1** would exhibit excellent FRET efficiency if used in the Halotag imaging setup. Interestingly, previous iterations of **FCR1** were predicted to have poor FRET efficiency, underscoring the importance of computational testing in the design of new molecular probes. The incidental nature of the final design of **FCR1** highlights the importance of conducting computational chemistry simulations prior to experimental work. In light of these promising results, it will be important to assess the experimental utility of **FCR1**-Tag to further validate the findings of the computational studies. An acid-containing scaffold of **FCR1** has been designed and synthesised by Dr Liam Adair and as such, these experiments should soon ensue.

For the turn on sensor **NpFR1**, benchmarking performed on excitation calculations revealed that B3LYP/6-31+G* gave the most accurate computational results when compared to the experimental values. However, in order to gain a more comprehensive understanding of the scaffold’s properties, future work should involve benchmarking methods that can provide accurate fluorescence emission profiles. In having the accurate tools to describe both the excitation and emission profile of the novel scaffold, future studies on novel designs can take place computationally. For example, sulfonation of the 6-position on a naphthalimide has been shown to significantly red-shift the emission profile.¹⁹ While the electronic-character likeness of **NpFR1** to the naphthalimide core could not be fully elucidated from these experiments, a computational study on this modification can provide a fast way of assessing how this modification may affect the fluorescence profile. Red shifting of the fluorescence emission is particularly useful for whole tissue imaging studies as it allows for deeper light penetration. Protein-targeted redox imaging of whole tissue can provide more valuable information of the protein redox environment, and will allow for a more comprehensive

understanding of the roles of ROS in whole systems.

As such, this chapter has shown how computational chemistry can provide valuable insights on the properties of fluorescent sensors. This allows decisions to be made about a particular molecular design, to help save time and resources by avoiding time-consuming experiments that may not yield the desired results.

5.3 Studies towards a method for peptide-receptor binding analysis

Peptides have emerged as promising therapeutics due to their high specificity and low toxicity, making them attractive for targeted therapy and diagnostics. However, discovering new peptide receptor targets is a challenging task, and traditional techniques can be resource-intensive and time-consuming. This work outlined the design and synthesis of a biontitylated fluorescent sensor that would allow for the development of a novel peptide-receptor binding assay.

The design of this sensor was such that the presence of the fluorophore could allow for instant identification of peptide-receptor binding. This was envisioned to be achieved with the use of fluorescence microscopy or flow cytometry. Further, the biotin moiety was incorporated to allow for simple isolation of the peptide-receptor complex that would allow for further assessment of the interaction. The probe was proven to successfully be incorporated into a peptide of interest, octreotide. However, due to instrumental difficulties, purification and testing could not be performed. Preliminary molecular docking provided promising results that warrant further development of this method.

Therefore, at the first possible opportunity, **BNp-COOH** conjugated octreotide should be resynthesised and isolated. Collaborations with Prof. Joel Mackay have been set up to assess and validate a method for the streptavidin-binding pull-down of peptide bound **BNp-COOH**. SSTR2- positive and negative cell lines, PC-3 and Huvec-TERET2 respectively, were identified and obtained. These cell lines should be used to determine whether the peptide can selectively bind to SSTR2-positive cells as intended.

The probe was designed to serve as an attachment point to free amines such as the N-terminus as was performed on octreotide. Additionally, it was envisioned that the probe

could be designed as a biotinylated fluorescent amino acid, which would allow for greater flexibility in attaching the probe at any desired location during peptide synthesis, thereby avoiding known active regions of interest. The suggested structure of this biotinylated fluorescent amino acid, **Fmoc-BNp-OH**, is shown in Figure 5.1

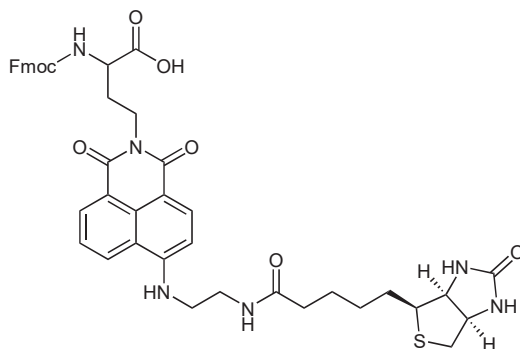


Figure 5.1: Suggested chemical structure of biotinylated fluorescent amino acid Fmoc-BNp-OH.

5.4 Outlook

This work has illustrated the development of novel fluorescent methods to detect and discover peptide and protein interactions *in vitro*. The combination of different imaging modalities with fluorescence imaging allows for a more comprehensive understanding of complex biological systems. Additionally, the use of computational chemistry allows for the assessment of fluorescent sensor designs prior to experimental work. This will optimise the process of development of novel sensors for specified applications.

With further development and refinement of these techniques, a more detailed understanding of these interactions, as well as the discovery of previously unknown interactions can be uncovered. In doing so, the discovery of new biological pathways and the identification of novel targets for drug development can be made. This could have a profound impact on the field of medicine, leading to the development of more effective treatments for the betterment of human health.

Chapter 6

Experimental methods

6.1 Synthetic methods

6.1.1 Instrumentation

^1H NMR spectra were obtained at a frequency of 300, 400 or 500 MHz using the Bruker Advance 300, Bruker Ascend 400 or Bruker Ascend 500 spectrometers respectively. ^{13}C NMR spectra were obtained at a frequency of 75, 101 or 126 MHz using the Bruker Advance 300, Bruker Ascend 400 or Bruker Ascend 500 spectrometers respectively. Samples were run in commercially obtained deuterated solvents. All NMR spectra are reported in parts per million (ppm) as downfield chemical shifts (δ) with reference to residual solvent peaks. Peak multiplicity abbreviations used are as follows: s = singlet; d = doublet; t = triplet; q = quartet; quin = quintuplet; sext = sextuplet; sept = septuplet; m = multiplet; br = broad. When needed, heteronuclear single quantum correlation spectroscopy (HSQC), correlation spectroscopy (COSY), and heteronuclear multiple-bond correlation spectroscopy (HMBC) experiments were used to help assign spectra.

Low-Resolution Mass Spectrometry (LRMS) was acquired on a Bruker Amazon SL mass spectrometer operating on Electrospray Ionisation (ESI) or Atmospheric Pressure Chemical Ionisation (APCI). High Resolution Mass Spectrometry (HRMS) was acquired on a Bruker Apex-Ultra spectrometer operating on ESI or APCI using an Apollo II ESI/MALDI dual source.

High Performance Liquid Chromatography (HPLC) was performed using a Waters prep HPLC with a 2545 quaternary pump, 2707 Autosampler, 2998 PDA detector (with semi

prep cell), and WFCIII fraction collector. Analytical HPLC was performed using a Waters Alliance e2690 with a 2998 PDA detector using a Sunfire C18 column (21 Å 150 mm).

6.1.2 Materials

For peptide synthesis, RCI Labscan anhydrous *N,N*-dimethylformamide (100ppm) was used. Dry solvents were obtained from the Pure-solv. All other solvents used were laboratory grade and used without further purification. Reagents were purchased from Sigma-Aldrich, Alfa Aesar, Merck, Fluka, Ajax, and Combi-blocks, and were used as purchased.

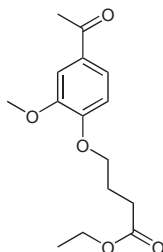
Analytical Thin Layer Chromatography (TLC) was performed using Merck Kieselgel 60 0.25 mm F254 silica plates. Flash column chromatography was performed using silica gel 60 (Merck, 0.040-0.063 mm) either manually, or using a Biotage Isolera One.

All reactions were performed under a stream of nitrogen.

6.2 Synthesis

6.2.1 Photo-cleavable linker

Ethyl 4-(4-acetyl-2-methoxyphenoxy)butanoate (**1**)



Acetovanillone (9.97 g, 60.0 mmol) and ethyl-4-bromobutyrate (10.8 mL, 75.5 mmol) were dissolved in dry *N,N*-dimethylformamide (100 mL). Potassium carbonate (12.5 g, 90.4 mmol) was added and the resulting suspension was stirred under nitrogen for 18 h. The reaction mixture was poured onto ice-cold water (600 mL), resulting the formation of a precipitate which was filtered and washed with ice-cold water (3 x 10 mL). The product was dried *in vacuo* to yield **1** (14.9 g, 53.4 mmol, 89%) as a white solid.

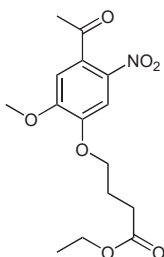
¹H NMR (500 MHz, CDCl₃): δ 7.56 - 7.52 (m, 2H) 6.89 (d, *J* = 8.1, 1H), 4.18 - 4.11 (m, 4H), 3.91 (s, 3H), 2.56 - 2.52 (m, 5H), 2.19 (quin, *J* = 6.7, 2H), 1.26 (t, *J* = 7.1, 3H).

^{13}C NMR (126 MHz, CDCl_3): δ 196.8, 173.1, 152.7, 149.4, 130.6, 123.3, 111.4, 110.6, 67.9, 60.5, 56.1, 30.6, 26.2, 24.4, 14.3.

HRMS (APCI $^+$) m/z : 281.1383 $[\text{M}+\text{H}]^+$ (calculated 281.1389 for $\text{C}_{15}\text{H}_{21}\text{O}_5^+$).

Characterisation data is in accordance with literature values.³⁴²

Ethyl 4-(4-acetyl-2-methoxy-5-nitrophenoxy)butanoate (2)



1 (4.11 g, 14.6 mmol) was added in small portions over 30 mins to nitric acid (150 mL, 70%) while stirring at 0 °C. The reaction mixture was then stirred for 2 h. The reaction was slowly diluted with water (600 mL) resulting in the formation of a precipitate which was filtered. The product was dried *in vacuo* to yield **2** (2.62 g, 8.06 mmol, 55%) as pale yellow crystals.

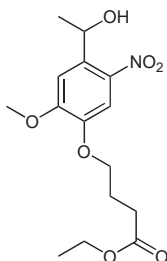
^1H NMR (500 MHz, CDCl_3): δ 7.61 (s, 1H), 6.75 (s, 1H), 4.18 - 4.15 (m 4H), 3.96 (s, 3H), 2.55 (t, $J = 7.2$, 2H), 2.50 (s, 3H), 2.20 (quin, $J = 6.7$, 2H), 1.27 (t, $J = 7.1$, 3H).

^{13}C NMR (126 MHz, CDCl_3): δ 200.0, 172.8, 154.3, 148.9, 138.4, 132.8, 108.8, 108.1, 68.5, 60.6, 56.6, 30.5, 30.4, 24.2, 14.2.

HRMS (APCI $^+$) m/z : 326.1234 $[\text{M}+\text{H}]^+$ (calculated 326.1240 for $\text{C}_{15}\text{H}_{20}\text{NO}_7^+$).

Characterisation data is in accordance with literature values.³⁴³

Ethyl 4-(4-(1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate (3)



2 (0.089 g, 0.273 mmol) was added to ethanol (50 mL) and cooled to 0 °C. Sodium borohydride (0.045 g, 1.2 mmol) was added in small portions over 10 min to the stirring solution.

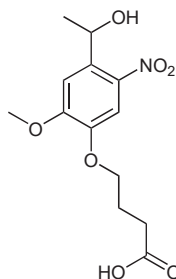
The reaction was warmed to room temperature and left to stir under nitrogen for 18 h. The reaction was quenched with water (100 mL) for 30 min. The ethanol was evaporated from the mixture, resulting in the formation of a precipitate which was filtered. The solid was washed with water (3 x 20 mL) then dried *in vacuo* to yield **3** (0.038 g, 0.11 mmol, 41%) as an off-white solid.

¹H NMR (500 MHz, CD₃OD): δ 7.60 (s, 1H), 7.42 (s, 1H), 5.48 (q, J = 7.5, 1H), 4.17 (q, J = 7.5, 4H), 4.12 (t, J = 5.0, 3H), 3.99 (s, 3H), 2.56 (t, J = 7.5, 2H), 2.14 (quin, J = 5.0, 2H), 1.50 (d, J = 5.0, 3H), 1.28 (t, J = 10.0, 3H).

¹³C NMR (126 MHz, CD₃OD): δ 175.0, 155.5, 148.3, 140.8, 139.0, 110.2, 110.0, 69.6, 66.2, 61.6, 56.8, 31.6, 25.6, 25.1, 14.5.

LRMS (ESI⁺) m/z : 350.08 [M+Na]⁺ (calculated 350.12 for [C₁₅H₂₁NNaO₇]⁺).

4-(4-(1-Hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoic acid (**4**)



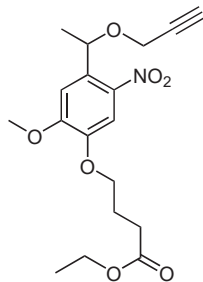
3 (0.250 g, 0.764 mmol) was dissolved in tetrahydrofuran (5 mL). 1 M Lithium hydroxide (5 mL) was added, and the two phase reaction was left to stir at room temperature for 3 h. The tetrahydrofuran was removed under a stream of nitrogen and the remaining solution was acidified to pH 3 by the addition of 1 M hydrochloric acid, resulting in the formation of a precipitate which was filtered. The solid was recrystallised in hot ethanol to yield **4** (0.083 g, 0.28 mmol, 36%) as a pale yellow powder.

¹H NMR (500 MHz, CD₃OD): δ 7.60 (s, 1H), 7.41 (s, 1H), 5.47 (q, J = 6.2, 1H), 4.11 (t, J = 6.1, 2H), 3.98 (s, 3H), 2.54 (t, J = 7.0, 2H), 2.16 – 2.05 (m, 2H), 1.48 (d, J = 6.2, 3H).

¹³C NMR (126 MHz, CD₃OD): δ 155.6, 148.3, 140.8, 139.0, 110.3, 110.0, 107.5, 69.6, 66.2, 56.7, 31.4, 25.7, 25.1.

LRMS (ESI[−]) m/z : 298.00 [M−H][−] (calculated 298.02 for [C₁₃H₁₆NO₇][−]).

Ethyl 4-(2-methoxy-5-nitro-4-(1-(prop-2-yn-1-yloxy)ethyl)phenoxy)butanoate (5)



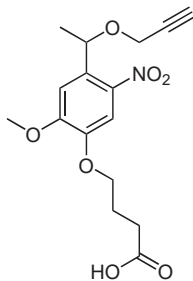
3 (0.101 g, 0.308 mmol) was dissolved in dry *N,N*-dimethylformamide (10 mL) and cooled to 0 °C. Sodium hydride (0.036 g, 1.5 mmol) was added and left to stir for 5 min. Propargyl bromide (0.10 mL, 1.1 mmol) and sodium iodide (0.016 g, 0.11 mmol) were added to the reaction mixture. The reaction was allowed to come to room temperature and was then stirred for 18 h. Water (50 mL) was added to quench the reaction. The pH was adjusted to 3 with 1 M hydrochloric acid, and the mixture was extracted with 1:4 isopropanol:chloroform (5 x 40 mL). The organic extracts were combined and washed with 10% sodium thiosulfate (2 x 50 mL), and brine (2 x 50 mL), then dried over magnesium sulfate. The solvent was removed *in vacuo* and the crude product was purified *via* column chromatography (SiO₂, 30% ethyl acetate in hexanes) to afford **5** (0.081 g, 0.240 mmol, 78%) as a yellow oil.

¹H NMR (400 MHz, CD₃OD): δ 7.63 (s, 1H), 7.28 (s, 1H), 5.89 (q, J = 6.3, 1H), 4.21 - 4.11 (m, 6H), 4.10 (s, 3H), 2.84 (t, J = 2.4, 1H), 2.58 (t, J = 7.3, 2H), 2.16 (quin, J = 7.1, 2H), 1.43 (d, J = 6.3, 3H), 1.29 (t, J = 7.1, 3H).

¹³C NMR (101 MHz, CD₃OD): δ 174.9, 155.8, 148.8, 142.2, 135.3, 110.3, 110.1, 80.4, 75.9, 73.5, 69.6, 61.6, 57.0, 56.8, 31.6, 25.6, 23.6, 14.5.

HRMS (ESI⁺) m/z : 388.1371 [M+Na]⁺ (calculated 388.1372 for [C₁₈H₂₃NNaO₇]⁺).

4-(2-Methoxy-5-nitro-4-(1-(prop-2-yn-1-yloxy)ethyl)phenoxy)butanoic acid (6)



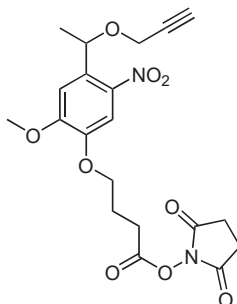
5 (0.226 g, 0.670 mmol) was dissolved in tetrahydrofuran (1.0 mL). 1 M lithium hydroxide (1.0 mL) was added and the two-phase reaction mixture was left to stir at room temperature for 4 h. The tetrahydrofuran was dried under a stream of nitrogen and the remaining aqueous layer was acidified to pH 3 using 2 M hydrochloric acid. The mixture was extracted into 1:4 isopropanol to chloroform (5 x 25 mL). Organic extracts were combined and washed with brine (2 x 25 mL) then dried over sodium sulfate. The solvent was removed under a stream of nitrogen to yield **6** (0.193 g, 0.539 mmol, 79%) as a dark yellow oil.

¹H NMR (500 MHz, CD₃OD): δ 7.63 (s, 1H), 7.26 (s, 1H), 5.37 (q, J = 6.3, 1H), 4.21 – 4.11 (m, 7H), 2.82 (t, J = 2.4, 1H), 2.54 (t, J = 7.3, 2H), 2.13 (m, J = 6.4, 2H), 1.52 (d, J = 6.3, 3H).

¹³C NMR (126 MHz, CD₃OD): δ 177.1, 155.6, 148.7, 142.0, 135.0, 110.1, 109.8, 80.2, 75.8, 73.3, 69.5, 56.8, 56.6, 31.4, 25.6, 23.4.

HRMS (ESI[−]) m/z : 336.1093 [M−H][−] (calculated 336.1083 for [C₁₆H₁₈NO₇][−]).

2,5-Dioxopyrrolidin-1-yl 4-(2-methoxy-5-nitro-4-(1-(prop-2-yn-1-yloxy)ethyl)phenoxy)butanoate (7)



6 (0.193 g, 0.529 mmol) was dissolved in dry *N,N*-dimethylformamide (10 mL). To the solution was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (0.532 g, 3.44 mmol), *N*-hydroxysuccinimide (0.262 g, 2.27 mmol), and diisopropylethylamine (0.70 mL, 4.0 mmol) after which the reaction mixture was left to stir overnight. The *N,N*-dimethylformamide was dried under a stream of nitrogen and the crude mixture was purified by column chromatography (SiO₂, 20-50% ethyl acetate in hexanes) to afford **7** (0.024 g, 0.055 mmol, 11%) as a pale yellow oil.

¹H NMR (500 MHz, CDCl₃): δ 7.59 (s, 1H), 7.20 (s, 1H), 5.37 (q, J = 6.2, 1H), 4.17 (t, J = 6.0, 2H), 4.13 – 3.92 (m, 5H), 2.90 – 2.84 (m, 6H), 2.40 (t, J = 2.4, 1H), 2.29 (quin,

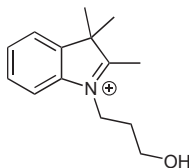
$J = 6.8$, 2H), 1.53 (d, $J = 6.3$, 3H).

^{13}C NMR (126 MHz, CDCl_3): δ 169.2, 168.3, 154.6, 147.1, 140.6, 135.0, 109.5, 108.8, 79.4, 74.8, 73.0, 67.6, 56.6, 56.5, 27.7, 25.7, 24.3, 23.6.

HRMS (ESI^+) m/z : 457.1214 $[\text{M}+\text{Na}]^+$ (calculated 457.1223 for $[\text{C}_{20}\text{H}_{22}\text{N}_2\text{NaO}_9]^+$).

6.2.2 CyMT

1-(3-Hydroxypropyl)-2,3,3-trimethyl-3H-indol-1-ium (8)

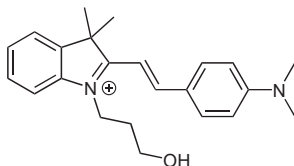


Trimethylindolenine (1.60 mL, 9.97 mmol) and 2-bromoethanol (1.50 mL, 18.8 mmol) were dissolved in acetonitrile (5 mL) and left to reflux for 24 h. The mixture was allowed to cool to room temperature after which cold diethyl ether was added slowly until a precipitate formed. The solid was filtered by vacuum filtration and dried under high vacuum to yield **8** (0.740 g, 3.39 mmol, 34%) as a purple solid.

^1H NMR (300 MHz, CD_3OD): δ 7.80 (m, 1H), 7.60 – 7.45 (m, 3H), 4.85 (t, $J = 6.8$, 2H), 3.78 (t, $J = 5.3$, 2H), 2.31 (quin, $J = 5.6$, 2H), 1.62 (s, 6H), 1.53 (s, 3H).

LRMS (ESI^+) m/z : 218.11 $[\text{M}]^+$ (calculated 218.15 for $[\text{C}_{14}\text{H}_{20}\text{NO}]^+$).

(E)-2-(4-(Dimethylamino)styryl)-1-(3-hydroxypropyl)-3,3-dimethyl-3H-indol-1-ium (9)

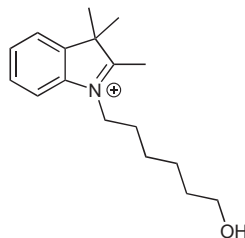


8 (0.103 g, 0.470 mmol) was dissolved in anhydrous ethanol to which 4-(dimethylamino) benzaldehyde (0.20 mL, 0.55 mmol) and methanesulfonic acid (0.02 mL, 0.21 mmol) were added. The reaction mixture was left to reflux overnight. After this, the ethanol was removed by rotary evaporation. The crude mixture was purified by column chromatography (SiO_2 , 0 -10% methanol in dichloromethane) to yield **9** (0.088 g, 0.25 mmol, 53%) as a purple solid.

¹H NMR (300 MHz, CDCl₃): δ 8.01 (d, J = 15.4, 2H), 7.66 (d, J = 15.4, 1H), 7.54 – 7.35 (m, 5H), 6.81 (d, J = 8.9, 2H), 4.79 (t, J = 7.5, 2H), 3.88 (t, J = 5.3, 2H), 3.19 (s, 6H), 2.26 – 2.17 (m, 2H), 1.76 (s, 6H).

LRMS (ESI⁺) m/z : 349.19 [M]⁺ (calculated 349.23 for [C₂₃H₂₉N₂O]⁺).

1-(6-Hydroxyhexyl)-2,3,3-trimethyl-3H-indol-1-ium (10)

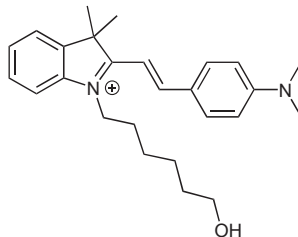


Trimethylindolenine (0.32 mL, 1.9 mmol) and 6-bromo-1-hexanol (0.26 mL, 2.2 mmol) were dissolved in acetonitrile (5 mL) and the mixture was left to reflux for 24 h. The mixture was cooled to room temperature after which ice cold diethyl ether was added until a precipitate was formed. The solid was filtered by vacuum filtration to yield **10** (0.095 g, 0.44 mmol, 22%) as a purple solid.

¹H NMR (300 MHz, CDCl₃): δ 7.85 (d, J = 3.5, 2H), 7.57 – 7.37 (m, 2H), 4.66 (t, J = 7.2, 2H), 3.51 (t, J = 5.4, 2H), 3.34 (t, J = 6.1, 2H), 1.94 - 1.76 (m, 2H), 1.58 (s, 10H), 1.49 (s, 3H).

LRMS (ESI⁺) m/z : 260.17 [M]⁺ (calculated 260.20 for [C₁₇H₂₆NO]⁺).

(E)-2-(4-(Dimethylamino)styryl)-1-(6-hydroxyhexyl)-3,3-dimethyl-3H-indol-1-ium (11)



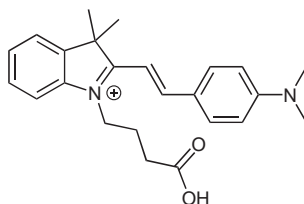
10 (0.020 g, 0.77 mmol) and 4-(dimethylamino)benzaldehyde (0.013 g, 0.085 mmol) were dissolved in absolute ethanol (30 mL). Methanesulfonic acid (0.20 mL, 0.21 mmol) was added and the mixture was heated to reflux overnight. The ethanol was then dried

in **vacuo** and the remaining solid was purified by column chromatography (SiO₂, 0 -10% methanol in dichloromethane) to yield **11** (0.14 g, 0.40 mmol, 52%) as a purple solid.

¹H NMR (300 MHz, CDCl₃): δ 8.02 (d, *J* = 15.2, 2H), 7.64 (d, *J* = 15.5, 1H), 7.54 - 7.38 (m, 5H), 6.85 (d, *J* = 8.6, 2H), 4.76 (t, *J* = 7.4, 2H), 3.65 (t, *J* = 5.2, 2H), 3.19 (s, 6H), 1.98 - 1.90 (m, 4H), 1.86 - 1.64 (m, 10H).

LRMS (ESI⁺) *m/z*: 392.24 [M]⁺ (calculated 392.27 for [C₁₇H₂₆NO]⁺).

(E)-1-(3-Carboxypropyl)-2-(4-(dimethylamino)styryl)-3,3-dimethyl-3H-indol-1-ium (12)



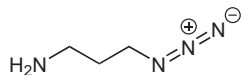
(E)-2-(4-(Dimethylamino)styryl)-1-(4-ethoxy-4-oxobutyl)-3,3-dimethyl-3H-indol-1-ium (0.603 g, 1.49 mmol), provided by Dr. Kylie Yang, was dissolved in 3 M hydrochloric acid (30 mL). The solution was heated to 80 °C and stirred under nitrogen for 18 h. Residual solvent was dried under nitrogen to yield **12** (0.556 g, 1.48 mmol) as a purple solid.

¹H NMR (400 MHz, DMSO-d₆): δ 8.52 (d, *J* = 15.6, 1H), 8.27 (d, *J* = 8.8, 2H), 7.94 (d, *J* = 3.3, 1H), 7.92 (d, *J* = 3.9, 1H), 7.70 (t, *J* = 7.7, 1H), 7.63 (t, *J* = 7.4, 1H), 7.48 (d, *J* = 15.6, 1H), 7.04 (d, *J* = 9.0, 2H), 4.64 (t, *J* = 7.4, 1H), 2.20 - 2.11 (m, 2H), 1.92 (s, 5H).

¹³C NMR (101 MHz, DMSO-d₆): δ 179.4, 173.9, 154.7, 154.6, 142.7, 141.1, 134.3, 130.7, 128.8, 127.4, 122.8, 122.4, 113.4, 112.3, 104.5, 50.9, 44.4, 30.2, 26.5, 22.8.

LRMS (ESI⁺) *m/z*: 377.21 [M]⁺ (calculated 377.22 for [C₂₄H₂₉N₂O₂]⁺).

3-Azidopropan-1-amine (13)



3-Chloropropyl-1-amine hydrochloride (0.499 g, 3.84 mmol) was dissolved in water (20 mL). Sodium azide (0.502 g, 7.72 mmol) was added to the reaction which was then heated to 80 °C for 18 h. The reaction mixture was cooled to room temperature before being poured

into 1 M potassium hydroxide (100 mL). The product was extracted with diethyl ether (3 x 30 mL) and subsequently washed with water (1 x 20 mL) then brine (1 x 20 mL). The product was concentrated in an ice bath under nitrogen to yield **13** (0.0836 g, 0.835 mmol) as a colourless oil.

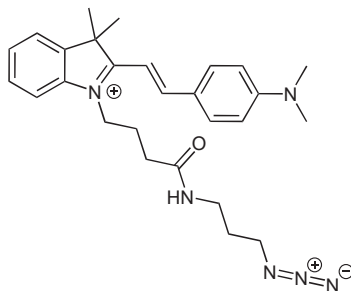
¹H NMR (300 MHz, CD₃OD): δ 3.40 (t, J = 6.6, 2H), 2.75 (t, J = 6.9, 2H), 1.76 (quin, J = 6.9, 2H).

¹³C NMR (75 MHz, CD₃OD): δ 50.0, 39.3, 32.4.

LRMS (ESI⁺) m/z : 101.13 [M+H]⁺ (calculated 101.07 for [C₃H₉N₄]⁺).

Characterisation data is in accordance with literature values.³⁴⁴

(E)-1-(4-((3-Azidopropyl)amino)-4-oxobutyl)-2-(4-(dimethylamino)styryl)-3,3-dimethyl-3H-indol-1-ium (14)



12 (0.098 g, 0.26 mmol) was dissolved in dry *N,N*-dimethylformamide (25 mL). Hydroxybenzotriazole (0.198 g, 1.29 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (0.212 g, 1.10 mmol) followed by **13** (0.084 g, 0.84 mmol) were added and the reaction mixture was stirred under nitrogen for 18 h. The product was poured into water (100 mL) and extracted with isopropanol in chloroform (1:5, 3 x 35 mL). The combined organic extracts were dried with magnesium sulfate then solvent was removed by rotary evaporation. The crude product was purified by column chromatography (SiO₂, 0 - 10% methanol in dichloromethane). The solvent was evaporated by rotary evaporation to afford **14** (0.020 g, 0.26 mmol, 17%) as a purple solid.

¹H NMR (500 MHz, CDCl₃): δ 8.00 (d, J = 15.3, 1H), 7.59 – 7.45 (m, 3H), 7.45 – 7.34 (m, 2H), 6.81 (d, J = 8.5, 2H), 4.64 – 4.46 (m, 2H), 3.48 – 3.27 (m, 4H), 3.19 (s, 6H), 2.81 (t, J = 6.7, 2H), 2.27 – 2.09 (m, 4H), 1.87 (quin, J = 6.8, 2H), 1.74 (s, 6H).

¹³C NMR (126 MHz, CDCl₃): δ 178.7, 172.9, 155.4, 155.2, 142.0, 141.1, 129.6, 127.9,

HRMS (ESI⁺) *m/z*: 459.2867 [M]⁺ (calculated 459.2872 for [C₂₇H₃₅N₆O]⁺).

CN(C)c1ccc(/C=C/[N+]2c3ccccc3C(C)(C)N2CCCCC(=O)NCCCN4C=NC=N4COC(C)c5ccc(OC)c(OCCOC(=O)ON6C(=O)CCC6=O)c5)cc1

¹H NMR (500 MHz, CD₃OD): δ 8.34 (d, *J* = 15.5, 1H, *CH*), 8.16 – 8.05 (m, 4H, 3×ArH, NH), 7.77 (d, *J* = 5.5, 2H, 2×ArH), 7.60 (s, 1H, ArH), 7.50 (dt, *J* = 7.5, 2H, 2×ArH), 7.37 (d, *J* = 15.6, 1H, *CH*), 7.26 (s, 1H, ArH), 6.86 (d, *J* = 8.9, 2H, 2×ArH), 5.14 (q, *J* = 6.0, 1H, *CH*), 4.51 – 4.40 (m, 2H, CH₂CH₂CH₂), 4.40 – 4.24 (m, 4H, CH₂, CH₂CH₂CH₂), 4.15 (t, *J* = 6.1, 2H, CH₂CH₂CH₂), 3.92 (s, 3H, OCH₃), 3.16 (s, 6H, 2×CH₃), 3.06 (q, *J* = 5.9, 2H, CH₂CH₂CH₂), 2.95 – 2.78 (m, 6H, 3×CH₂), 2.36 (t, *J* = 6.2, 2H, CH₂CH₂CH₂), 2.09 (quin, *J* = 6.8, 2H, CH₂CH₂CH₂), 2.02 (quin, *J* = 6.61, 2H, CH₂CH₂CH₂), 1.91 (quin, *J* = 6.8, 2H, CH₂CH₂CH₂), 1.75 (s, 6H, 2×CH₃), 1.41 (d, *J* = 6.2, 3H, CH₃).

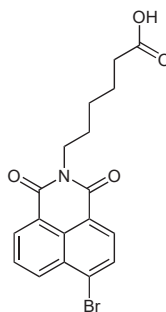
159

155.9, 153.0, 152.1, 150.6, 149.5, 143.9, 138.2, 136.9, 133.4, 132.2, 131.8, 127.0, 122.9, 121.7, 118.4, 114.2, 81.7, 76.7, 70.9, 65.6, 60.3, 56.6, 54.2, 45.3, 40.5, 39.3, 36.4, 35.9, 34.9, 34.7, 33.4, 32.6.

HRMS (ESI⁺) m/z : 893.4190 [M]⁺ (calculated 893.4192 for [C₄₇H₅₇N₈O₁₀]⁺).

6.2.3 NpMT

6-(6-Bromo-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoic acid (**15**)



4-Bromo-1,8-naphthalic anhydride (2.08 g, 7.50 mmol) and 6-aminohexanoic acid (1.43 g, 10.9 mmol) were suspended in absolute ethanol (50 mL). Triethylamine (1.40 mL, 10.0 mmol) was added and the reaction mixture was refluxed under nitrogen for 18 h. The solvent was removed and the crude solid was purified *via* column chromatography (SiO₂, 5% ethyl acetate in hexane) to afford **15** (2.76 g, 7.07 mmol, 94%) as a pale brown solid.

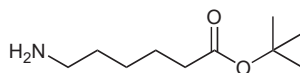
¹H NMR (300 MHz, CDCl₃): δ 8.66 (d, J = 7.5, 1H), 8.58 (d, J = 8.4, 1H), 8.42 (d, J = 7.8, 1H), 8.05 (d, J = 7.8, 1H), 7.85 (t, J = 7.8, 1H), 4.18 (t, J = 6.9, 2H), 2.38 (t, J = 7.2, 2H), 1.85-1.64 (m, 4H) 1.54 - 1.42 (m, 2H).

¹³C NMR (75 MHz, CDCl₃): δ 178.9, 163.6, 132.0, 130.9, 130.4, 130.0, 128.8, 123.0, 122.2, 40.7, 40.1, 28.1, 26.6, 24.2 (missing Ar signals attributed to symmetry).

LRMS (ESI⁻) m/z : 389.88 [M-H]⁻ (calculated 388.02 for [C₁₈H₁₅BrNO₄]⁻).

Characterisation data is in accordance with literature values.³⁴⁵

tert-Butyl 6-aminohexanoate (**16**)



6-Aminohexanoic acid (2.03 g, 12.5 mmol) was dissolved in thionyl chloride (3.50 mL, 48.0 mmol) and stirred at room temperature for 1.5 h. Excess thionyl chloride was then evapo-

rated under a stream of nitrogen after which a solution of sodium hydrogen carbonate (2.82 g, 33.5 mmol) in *tert*-butyl alcohol (2.80 g, 33.3 mmol) was added and the reaction mixture was stirred for 18 h. The mixture was concentrated and the remaining residue was diluted with 1 M sodium hydroxide (50 mL) and extracted with ethyl acetate (3 x 30 mL). The organic extracts were combined and dried over magnesium sulfate and solvent was removed by rotary evaporation to yield **16** (2.00 g, 4.49 mmol, 29%) as a colourless oil.

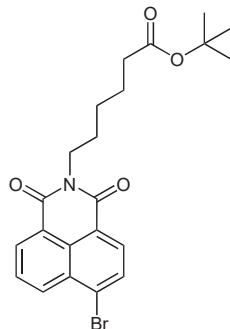
¹H NMR (300 MHz, CDCl₃): δ 2.70 (t, J = 7.0, 2H), 2.22 (t, J = 7.4, 2H), 1.63 – 1.45 (m, 4H), 1.44 - 1.28 (m, 11H).

¹³C NMR (75 MHz, CDCl₃): δ 172.8, 79.4, 41.7, 35.1, 32.9, 27.7, 26.1, 24.5.

LRMS (ESI⁺) m/z : 180.10 [M+H]⁺ (calculated 180.16 for [C₁₀H₂₂NO₂]⁺).

Characterisation data is in accordance with literature values.³⁴⁶

***tert*-Butyl 6-(6-bromo-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoate (17)**



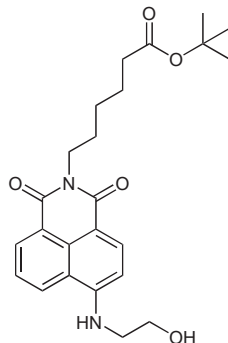
4-Bromo-1,8-naphthalic anhydride (0.455 g, 1.64 mmol) and **16** (0.350 g, 1.87 mmol) were added to absolute ethanol (50 mL) and refluxed under nitrogen for 18 h. The solvent was then dried and the crude product was purified by column chromatography (SiO₂, 20% ethyl acetate in hexane) to yield **17** (0.653 g, 1.64 mmol, 89%) as a pale brown solid.

¹H NMR (400 MHz, CDCl₃): δ 8.54 (d, J = 7.3, 1H), 8.43 (d, J = 8.5, 1H), 8.29 (d, J = 7.8, 1H), 7.93 (d, J = 7.8, 1H), 7.75 (t, J = 7.9, 1H), 4.19 – 4.02 (m, 2H), 2.20 (t, J = 7.4, 2H), 1.72 - 1.63 (m, 4H), 1.49 – 1.32 (m, 11H).

¹³C NMR (101 MHz, CDCl₃): δ 173.0, 163.6, 163.5, 133.2, 132.0, 131.2, 131.1, 130.7, 130.2, 129.0, 128.1, 123.2, 122.3, 80.0, 40.4, 35.5, 28.1, 27.8, 26.6, 24.8.

LRMS (ESI⁺) m/z : 468.07 [M+Na]⁺ (calculated 468.08 for [C₂₂H₂₄BrNNaO₄]⁺).

***tert*-Butyl 6-(6-((2-hydroxyethyl)amino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoate (18)**



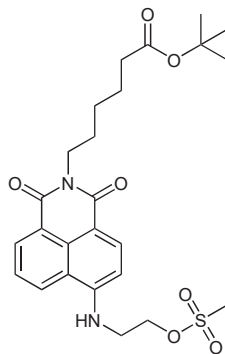
17 (0.600 g, 1.34 mmol) and ethanolamine (0.31 mL, 5.1 mmol) were dissolved in dimethylsulfoxide (3 mL) and heated to 70 °C under nitrogen for 18 h. The reaction mixture was allowed to cool before being diluted with 1 M hydrochloric acid (50 mL). The product was then extracted into diethyl ether (3 x 30 mL) then washed with brine (2 x 20 mL). The organic extract was dried over magnesium sulphate and dried on the rotary evaporator to yield **18** (0.527 g, 1.24 mmol, 92%) as a yellow solid.

¹H NMR (300 MHz, CDCl₃): δ 8.34 (t, J = 10.9, 1H), 8.26 (d, J = 8.4, 1H), 8.03 (d, J = 8.2, 1H), 7.56 – 7.38 (m, 1H), 6.58 (d, J = 8.5, 1H), 4.15 – 3.96 (m, 4H), 3.53 (s, 2H), 2.28 – 2.12 (m, 2H), 1.71 - 1.64 (m, 4H), 1.43 (d, J = 9.1, 11H).

¹³C NMR (75 MHz, CDCl₃): δ 173.5, 164.6, 164.2, 149.7, 134.3, 121.1, 126.2, 124.7, 122.8, 120.3, 110.2, 104.5, 80.3, 60.5, 45.6, 40.1, 35.6, 29.7, 28.7, 28.2, 28.0, 27.8, 26.7, 25.0.

LRMS (ESI⁺) m/z : 449.16 [M+Na]⁺ (calculated 449.21 for [C₂₄H₃₀N₂NaO₅]⁺).

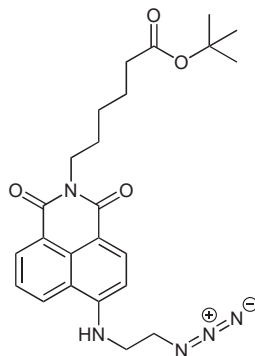
***tert*-Butyl 6-(6-((2-((methylsulfonyl)oxy)ethyl)amino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoate (19)**



18 (0.446 g, 1.05 mmol) was dissolved in chloroform (30 mL) and cooled to 0 °C. *N,N*-Diisopropylethylamine (1.65 mL, 9.47 mmol) was added dropwise to the solution, followed by methanesulfonyl chloride (0.48 mL, 6.2 mmol) which was added dropwisely. The solution was stirred for 30 mins after which the reaction was left to reach room temperature and stir for a further 1.5 h. The reaction mixture was washed with 1 M hydrochloric acid (1 x 30 mL) followed by 5% sodium carbonate (1 x 30 mL) then brine (1 x 30 mL). The organic phase was dried over magnesium sulfate then dried by rotary evaporation to yield **19** as a yellow solid which was used immediately in the next step.

LRMS (ESI⁺) *m/z*: 527.21 [M+Na]⁺ (calculated 527.18 for [C₂₅H₃₂N₂NaO₇S]⁺).

***tert*-Butyl 6-((2-azidoethyl)amino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoate (20)**



19 (0.660 g, 1.31 mmol) and sodium azide (0.280 g, 4.31 mmol) were dissolved in *N,N*-dimethylformamide (25 mL) and heated to 80 °C under nitrogen for 6 h. The reaction mixture was then allowed to cool to room temperature before being poured into water (100 mL) and extracted with ethyl acetate (3 x 30 mL). The combined organic extracts were subsequently washed with water (1 x 30 mL) followed by brine (1 x 30 mL) and dried over sodium sulfate. The solvent was removed by rotary evaporation to yield **20** (0.568 g, 1.05 mmol, 80%) as a yellow solid.

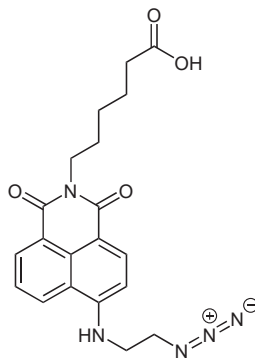
¹H NMR (500 MHz, CDCl₃): δ 8.55 (d, *J* = 7.3, 1H), 8.42 (d, *J* = 8.3, 1H), 8.18 (d, *J* = 8.4, 1H), 7.62 (t, *J* = 7.9, 1H), 6.68 (d, *J* = 8.4, 1H), 4.63 (t, *J* = 4.9, 2H), 4.13 (t, *J* = 7.5, 2H), 3.77 (t, *J* = 4.9, 2H), 3.10 (s, 3H), 2.21 (t, *J* = 7.5, 2H), 1.75 - 1.62 (m, 4H), 1.49 - 1.41 (m, 12H).

¹³C NMR (126 MHz, CDCl₃): δ 176.7, 164.2, 163.6, 133.9, 133.3, 132.1, 131.3, 131.1,

130.7, 122.7, 122.3, 48.4, 40.2, 39.6, 33.4, 29.7, 27.7, 26.5, 24.2, 22.1, 19.2.

LRMS (ESI⁺) *m/z*: 474.25 [M+Na]⁺ (calculated 474.21 for [C₂₄H₂₉N₅NaO₄]⁺).

6-(6-((2-Azidoethyl)amino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoic acid (21)



20 (0.254 g, 0.563 mmol) was dissolved in dichloromethane (2 mL). Trifluoroacetic acid (1.50 mL, 1.96 mmol) was added and the reaction mixture was stirred at room temperature for 1 h. Residual solvent was evaporated under a stream of nitrogen. The crude solid was washed with diethyl ether (3 x 100 mL) and dried *in vacuo* to yield **21** (0.198 g, 0.546 mmol, 98%) as a yellow solid.

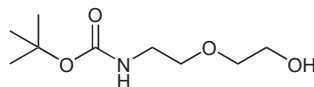
¹H NMR (500 MHz, DMSO-d₆): δ 8.68 (d, *J* = 8.5, 1H, Ar*H*), 8.45 (d, *J* = 6.7, 1H, Ar*H*), 8.28 (d, *J* = 8.5, 1H, Ar*H*), 7.90 (t, *J* = 4.5, 1H, NH), 7.71 (t, *J* = 7.5, 1H, Ar*H*), 6.91 (d, *J* = 8.6, 1H, Ar*H*), 4.00 (t, *J* = 7.3, 2H, NCH₂), 3.66 - 3.63 (m, 4H, 2×CH₂), 2.21 (t, *J* = 7.3, 2H, NCH₂), 1.62 - 1.51 (m, 4H, 2×CH₂), 1.36 - 1.29 (m, 2H, CH₂).

¹³C NMR (126 MHz, DMSO-d₆): δ 174.4, 163.7, 162.9, 150.2, 134.1, 130.7, 129.4, 128.5, 127.3, 124.5, 122.0, 120.3, 108.3, 104.1, 28.9, 42.3, 33.5, 27.4, 26.0, 24.2.

HRMS (ESI⁺) *m/z*: 418.1486 [M+Na]⁺ (calculated 418.1481 for [C₂₀H₂₁N₅NaO₄]⁺).

6.2.4 HaloTag linker

***tert*-Butyl (2-(2-hydroxyethoxy)ethyl)carbamate (22)**



2-(2-Aminoethoxyethanol) (1.52 mL, 15.0 mmol) was added to absolute ethanol (50 mL) and stirred at 0 °C for 30 min. Di-*tert*-butyl dicarbonate (3.30 g, 15.11 mmol) was added

and the reaction mixture was allowed to warm to room temperature before stirring for 2.5 h. Dichloromethane (50 mL) was added and the mixture was washed with brine (2 x 75 mL). The organic layer was dried over magnesium sulfate, and was concentrated *in vacuo* to yield **22** (2.61 g, 12.7 mmol, 85%) as a pale yellow oil.

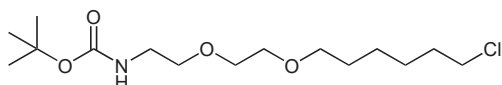
¹H NMR (300 MHz, DMSO-*d*₆): δ 5.22 (s, br, 1H), 3.69 - 3.66 (m, 2H), 3.53 - 3.48 (m, 4H), 3.26 (q, J = 5.1, 2H), 3.02 (s, br, 1H), 1.39 (s, 9H).

¹³C NMR (75 MHz, DMSO-*d*₆): δ 156.2, 79.3, 72.3, 70.4, 61.5, 40.3, 28.4.

LRMS (ESI⁺) m/z : 228.06 [M+Na]⁺ (calculated 228.12 for [C₉H₁₉NNaO₄]⁺).

Characterisation data is in accordance with literature values.³⁴⁷

***tert*-Butyl (2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamate (**23**)**



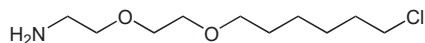
22 (1.042 g, 5.091 mmol) was dissolved in *N,N*-dimethylformamide (10 mL) and the solution was stirred at 0 °C. Sodium hydride (60% in mineral oil, 0.304 g, 12.5 mmol) was added the mixture was left to stir for 45 min. 1-Chloro-6-iodohexane (2.01 mL, 8.11 mmol) was added and the reaction mixture was warmed to room temperature before being stirred for 4 h. The reaction was quenched with saturated ammonium chloride solution (10 mL) then extracted with ethyl acetate (3 x 50 mL). The organic phase washed with brine (2 x 70 mL) then dried over magnesium sulfate. Solvent was removed *in vacuo* to yield **23** (0.172 g, 0.534 mmol, 25%) as a yellow oil.

¹H NMR (300 MHz, DMSO-*d*₆): δ 4.99 (s, br, 1H), 3.58 - 3.41 (m, 10H), 3.28 (t, J = 5.0, 2H), 1.74 (quin, J = 7.0, 2H), 1.58 (quin, J = 7.0, 2H), 1.41 - 1.21 (m, 13H).

¹³C NMR (75 MHz, DMSO-*d*₆): δ 155.9, 79.1, 71.2, 70.3, 70.2, 70.0, 45.0, 40.3, 32.5, 29.4, 28.4, 26.7, 25.4.

LRMS (ESI⁺) m/z : 346.16 [M+Na]⁺ (calculated 346.18 for [C₁₅H₃₀ClNNaO₄]⁺).

(2-(2-((Chlorohexyl)oxy)ethoxy)ethan-1-amine (24**))**



23 (0.350 g, 1.07 mmol) was dissolved in dichloromethane (5 mL). 2,2,2-Trifluoroacetic acid (0.80 mL, 9.9 mmol) was added and the reaction was stirred at room temperature for 5 h

before being quenched with potassium carbonate (0.523 g, 3.78 mmol) in methanol (6 mL) for 30 mins. The solvent was removed by rotary evaporation to yield **24** (1.47 g, 6.58 mol, 45%) as a pale orange oil.

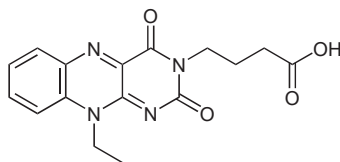
¹H NMR (300 MHz, CD₃OD): δ 3.78 - 3.68 (m, 6H), 3.62 (t, *J* = 6.7, 2H), 3.57 (t, *J* = 6.6, 2H), 3.21 - 3.18 (m, 2H), 1.85 (quin, *J* = 6.8, 2H), 1.67, (quin, *J* = 6.9, 2H), 1.59 - 1.41 (m, 4H).

¹³C NMR (75 MHz, DMSO-*d*₆): δ 71.4, 70.3, 69.9, 66.7, 45.1, 40.0, 32.6, 29.3, 26.7, 25.3.

LRMS (ESI⁺) *m/z*: 224.13 [M+H]⁺ (calculated 224.14 for [C₁₀H₂₃ClNO₂]⁺).

6.2.5 FlavinTag

4-(10-Ethyl-2,4-dioxo-4,10-dihydrobenzo[g]pteridin-3(2H)-yl)butanoic acid (**25**)



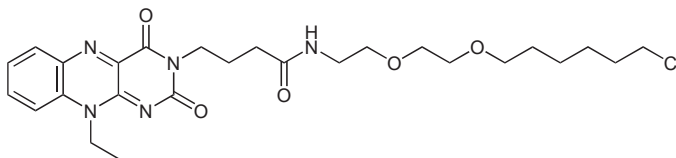
Methyl 4-(10-ethyl-2,4-dioxo-4,10-dihydrobenzo[g]pteridin-3(2H)-yl)butanoate (supplied by Dr. Kylie Yang, 0.222 g, 0.698 mmol) was suspended in conc. hydrochloric acid (40 mL) and refluxed overnight. The mixture was then diluted with water (50 mL) and extracted with dichloromethane (5 x 40 mL). The organic extracts were combined and washed with brine (2 x 30 mL) then dried over magnesium sulfate. The solvent was removed by rotary evaporation to yield **25** (0.212 g, 0.665 mmol, 94%) as a yellow solid.

¹H NMR (300 MHz, DMSO-*d*₆): δ 12.01 (s, br, 1H), 8.1 - 8.16 (m, 1H), 8.04 - 7.94 (m, 1H), 7.70 - 7.64 (m, 2H), 4.67 (q, *J* = 6.9, 2H), 3.94 (t, *J* = 6.8, 2H), 2.28 (t, *J* = 7.5, 2H), 1.83 (quin, *J* = 7.1, 2H), 1.34 (t, *J* = 7.1, 3H).

¹³C NMR (75 MHz, DMSO-*d*₆): δ 174.0, 159.5, 156.0, 148.8, 149.7, 135.2, 135.0, 132.1, 131.9, 126.0, 116.2, 40.2, 31.2, 30.7, 22.9, 11.9.

LRMS (ESI⁻) *m/z* 351.00 [M+Na]⁺ (calculated 351.11 for [C₁₆H₁₆N₄NaO₄]⁺).

N-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethyl)-4-(10-ethyl-2,4-dioxo-4,10-dihydrobenzo[g]pteridin-3(2H)-yl)butanamide (FlavinTag)



25 (0.072 g, 0.25 mmol) was dissolved into *N,N*-dimethylformamide (10 mL). 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (0.133 g, 0.701 mmol) followed by *N*-hydroxysuccinimide (0.112 g, 0.969 mmol) were added to the solution which was left to stir for 24 h. After this, the reaction mixture was diluted with water (50 mL) and extracted into chloroform (5 x 25 mL). The organic extracts were combined and washed with brine (2 x 20 mL) then dried over magnesium sulfate. The solvent was removed *in vacuo* to yield a bright yellow oil which was then dissolved in *N,N*-dimethylformamide. **24** was added to the solution and was left to stir for 24 h. After this, the mixture was diluted with 2 M sodium hydroxide (20 mL) and extracted with dichloromethane (5 x 50 mL). The organic extracts were combined and washed with brine (2 x 25 mL) then dried over magnesium sulfate. The solvent was removed by rotary evaporation to yield an oily yellow product which was purified by column chromatography (SiO₂, 0 - 5% methanol in dichloromethane). The solvent was removed by rotary evaporation to yield **FlavinTag** (0.033 g, 0.060 mmol, 32.0%) as a bright yellow solid.

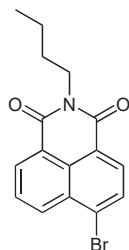
¹H NMR (400 MHz, CDCl₃): δ 8.34 (d, J = 8.7 Hz, 1H, ArH), 7.93 - 7.88 (t, J = 7.8 Hz, 1H, ArH), 7.68 - 7.61 (m, 2H, ArH), 4.79 (q, J = 7.0 Hz, 2H, NCH₂), 4.18 (q, J = 6.8 Hz, 2H, NCH₂), 3.58 - 3.42 (m, 14H, 8xCH₂), 3.32 (t, J = 8.3 Hz, 2 H, CH₂), 2.52 - 2.42 (m, 2H, CH₂), 2.13 - 2.06 (m, 2H, CH₂), 1.81 - 1.71 (m, 4H, 2xCH₂), 1.51 (t, J = 8.3 Hz, 3H, CH₃).

¹³C NMR (101 MHz, CDCl₃): δ 173.4, 136.0, 135.7, 133.6, 126.4, 115.0, 71.3, 70.2, 70.1, 69.9, 45.1, 41.8, 40.3, 40.2, 39.2, 37.3, 34.0, 32.3, 31.1, 29.7, 29.5, 26.7, 24.2, 23.5.

HRMS (ESI⁺) m/z : 556.2297 [M+Na]⁺ (calculated 556.2303 for [C₂₆H₃₆ClN₅O₅]⁺).

6.2.6 NaphTag

6-Bromo-2-butyl-1H-benzo[de]isoquinoline-1,3(2H)-dione (**26**)



1-Bromo-8-naphthalic anhydride (4.119 g, 14.87 mmol) was suspended in absolute ethanol (40 mL) after which *N*-butylamine (2.20 mL, 22.3 mmol) was added. The reaction mixture was heated to reflux for 24. After this, the reaction mixture was left to cool before the solvent was removed by rotary evaporation. The crude mixture was purified by column chromatography (SiO₂, 7.5% ethyl acetate in hexanes) to yield **26** as a pale yellow solid (1.099 g, 14.87 mmol, 22%)

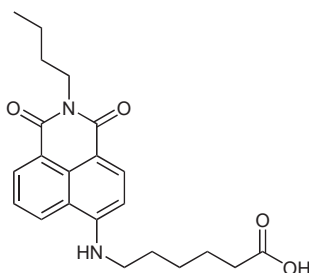
¹H NMR (300 MHz, CDCl₃): δ 8.65 (d, J = 7.2, 1H), 8.57 (d, J = 8.4, 1H), 8.41 (d, J = 7.8, 1H), 8.03 (d, J = 7.8, 1H), 7.85 (t, J = 7.8, 1H), 4.18 (t, J = 7.4, 2H), 1.72 (quin, J = 7.4, 2H), 1.45 (sext, J = 7.2, 2H), 0.98 (t, J = 7.2, 3H).

¹³C NMR (75 MHz, CDCl₃): δ 163.8, 133.4, 132.2, 131.4, 131.3, 130.8, 130.3, 129.2, 128.2, 123.4, 122.5, 40.5, 30.3, 20.5, 14.0.

LRMS (APCI⁺) m/z : 333.98 [M+H]⁺ (calculated 333.21 for [C₁₆H₁₅BrNO₂]⁺).

Characterisation data is in accordance with literature values.³⁴⁸

6-((2-Butyl-1,3-dioxo-2,3-dihydro-1H-benzo[de]isoquinolin-6-yl)amino)hexanoic acid (**27**)



A mixture of **26** (0.105 g, 0.316 mmol) and 6-aminohexanoic acid (0.450 g, 3.43 mmol) were dissolved in dimethyl sulfoxide and was heated to 70°C and left to react overnight. The

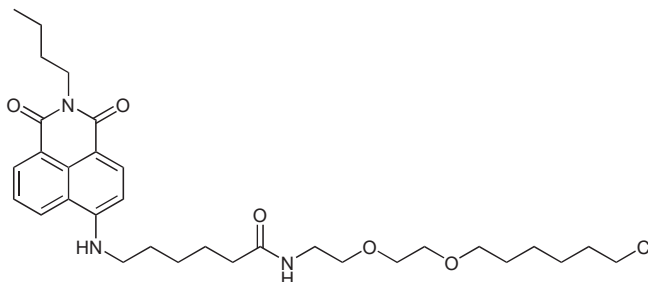
reaction mixture was concentrated under a stream of nitrogen. The residue was diluted with 0.1 M hydrochloric acid (20 mL) and extracted with chloroform (5 x 40 mL). The combined organic extracts were washed with water (2 x 20 mL) followed by brine (2 x 20 mL) then dried over magnesium sulfate. The solvent was removed by rotary evaporation to yield **27** (0.093 g, 0.316 mmol, 77%) as a bright yellow solid.

¹H NMR (300 MHz, CDCl₃): δ 8.57 (d, J = 7.0, 1H), 8.46 (d, J = 8.5, 1H), 8.12 (d, J = 8.5, 1H), 7.61 (t, J = 8.0, 1H), 6.72 (d, J = 8.5, 1H), 4.16 (t, J = 7.0, 2H), 3.43 (t, J = 6.5, 2H), δ 2.43 - 2.41 (m, 2H), 1.85 - 1.67 (m, 6H), 1.52 - 1.39 (m, 4H), 0.98 (t, J = 7.0, 3H).

¹³C NMR (75 MHz, CDCl₃): δ 165.0, 164.5, 149.6, 143.6, 131.3, 130.0, 126.1, 125.0, 123.5, 120.5, 110.6, 104.6, 43.6, 40.3, 30.6, 28.6, 26.7, 24.5, 22.9, 20.8, 13.9.

LRMS (ESI⁻) m/z : 381.29 [M-H]⁻ (calculated 381.18 for [C₂₂H₂₅N₂O₄]⁻).

6-((2-Butyl-1,3-dioxo-2,3-dihydro-1H-benzo[de]isoquinolin-6-yl)amino)-N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)hexanamide (NaphTag)



27 (0.042 g, 0.091 mmol) and **24** (0.100 g, 0.449 mmol) were dissolved in *N,N*-dimethyl formamide and stirred at 40°C for 48 h. The reaction mixture was concentrated under a stream of nitrogen after which it was diluted with sodium hydroxide solution (1M, 50 mL). The aqueous layer was extracted with dichloromethane (5 x 40 mL). The combined organic layers were washed with brine (2 x 50 mL) then dried over magnesium sulfate. The solvent was removed, and the remaining crude mixture was purified by column chromatography (SiO₂, 0 - 3% methanol in dichloromethane) to yield **NaphTag** (0.013 g, 0.23 mmol, 26%) as a bright yellow solid.

¹H NMR (500 MHz, CDCl₃): δ 8.57 (s, 1H, ArH), 8.45 (s, 1H, ArH), 8.24 (d, J = 8.2, 1H, ArH), 7.61 (t, J = 8.3, 1H, ArH), 6.72 (d, J = 8.4, 1H, ArH), 4.16 (t, J = 7.6, 2H,

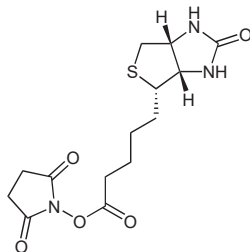
CH_2O), 3.60 – 3.44 (m, 16H, $8\times CH_2$), 2.26 (t, $J = 7.1$, 2H, NCH_2), 1.83 - 1.68 (m, 8H, $4\times CH_2$), 1.50 - 1.33 (m, 10H, $5\times CH_2$), 0.97 (t, $J = 7.4$, 3H, CH_3).

^{13}C NMR (126 MHz, $CDCl_3$): δ 173.0, 164.9, 164.3, 149.7, 134.6, 131.2, 126.4, 124.7, 123.3, 120.5, 110.4, 104.4, 71.4, 70.4, 70.2, 69.9, 45.2, 43.4, 40.1, 39.3, 36.2, 32.6, 30.5, 30.3, 29.8, 29.6, 28.2, 26.8, 26.6, 25.6, 24.8, 20.6, 20.5, 14.0, 14.0, 1.2.

HRMS (ESI^+) m/z : 610.3018 $[M+Na]^+$ (calculated 610.3018 for $[C_{32}H_{46}ClN_3NaO_5]^+$).

6.2.7 BNp-COOH

2,5-Dioxopyrrolidin-1-yl 5-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl) pentanoate (**28**)



D-Biotin (1.26 g, 5.14 mmol) was dissolved in *N,N*-dimethylformamide (50 mL). *N*-Hydroxy succinimide (0.587 g, 5.01 mmol) and 1,3-dicyclohexyl carbodiimide (1.08 g, 5.25 mmol) were added to the mixture and the resulting suspension was stirred under nitrogen for 24 h. The suspension was filtered and washed with diethyl ether. The product was dried *in vacuo* to yield **28** (1.57 g, 4.61 mmol, 92%) as a white solid.

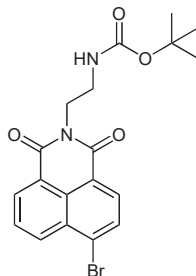
1H NMR (500 MHz, $DMSO-d_6$): δ 6.37 (d, $J = 8.6$, 2H), 4.32 - 4.27 (m, 1H), 4.15 - 4.12 (m, 1H), 3.13 - 3.06 (m, 2H), 2.85 - 2.83 (m, 5H), 2.66 (t, $J = 7.5$, 2H), 1.68 - 1.47 (m, 6H).

^{13}C NMR (126 MHz, CD_3OD): δ 172.0, 170.3, 166.2, 63.3, 61.7, 56.9, 56.8, 41.0, 34.8, 32.0, 31.4, 29.3, 26.5, 25.6.

LRMS (ESI^+) m/z : 364.12 $[M+Na]^+$ (calculated 364.09 for $[C_{14}H_{19}N_3NaO_5S]^+$).

Characterisation data is in accordance with literature values.³⁴⁹

***tert*-Butyl (2-(6-bromo-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)ethyl)carbamate (**29**)**

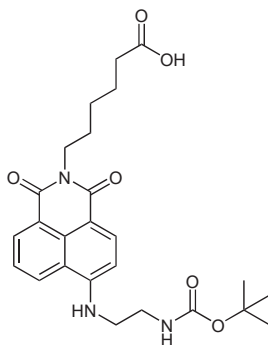


4-Bromo-1,8-naphthalic anhydride (3.11 g, 11.2 mmol) and *N*-boc-ethylenediamine (2.20 mL, 12.9 mmol) were suspended in absolute ethanol (70 mL). The mixture was heated to reflux for 24 h. The ethanol was removed. The brown residue was washed with diethyl ether (3 x 30 mL) to yield **29** (4.53 g, 10.8 mmol, 96%) as a pale yellow solid.

¹H NMR (400 MHz, CDCl₃): δ 8.63 (d, *J* = 7.3, 0.7, 1H), 8.54 (d, *J* = 8.5, 1H), 8.39 (d, *J* = 7.9, 1H), 8.01 (d, *J* = 7.9, 1H), 7.82 (dd, *J* = 8.3, 7.5, 1H), 4.34 (t, *J* = 5.7, 2H), 3.51 (t, *J* = 10.7, 2H), 1.27 (s, 9H).

HRMS (ESI⁺) *m/z*: 441.0420 [M+Na]⁺ (calculated 441.0426 for [C₁₉H₁₉BrN₂NaO₄]⁺). Characterisation data is in accordance with literature values.³⁵⁰

6-(6-((2-((*tert*-Butoxycarbonyl)amino)ethyl)amino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoic acid (30**)**



15 (0.606 g, 1.16 mmol) and *N*-Boc-ethylenediamine (1.00 mL, 16.8 mmol) were added to dimethylsulfoxide (4 mL) and heated to 70 °C under nitrogen for 18 h. The reaction mixture was cooled to room temperature before being poured into 0.1 M hydrochloric acid (150 mL). The aqueous mixture was extracted with dichloromethane (5 x 25 mL). Organic

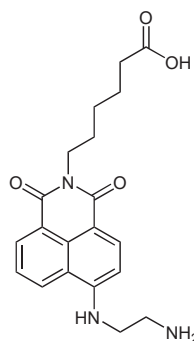
extracts were combined and dried over magnesium sulphate then concentrated to yield an oily residue that was washed with petroleum ether then dried *in vacuo*, yielding **30** (0.610 g, 1.02 mmol, 88%) as a yellow solid.

¹H NMR (500 MHz, CD₃OD): δ 8.51 (d, *J* = 7.2, 1H), 8.46 (d, *J* = 8.4, 1H), 8.34 (d, *J* = 8.6, 1H), 7.65 (t, *J* = 7.6, 1H), 6.84 (d, *J* = 8.5, 1H), 4.14 (t, *J* = 7.5, 2H), 3.54 (t, *J* = 5.7, 2H), 3.46 (t, *J* = 5.8, 2H), 2.33 (t, *J* = 5.7, 2H), 1.72 (sept, *J* = 5.7, 2H), 1.50 - 1.44 (m, 12H), 1.33 (s, 2H).

¹³C NMR (126 MHz, CD₃OD): δ 177.8, 166.3, 165.8, 159.9, 152.7, 136.0, 132.8, 131.3, 129.3, 125.6, 123.5, 121.9, 109.6, 105.0, 80.5, 45.2, 40.9, 40.8, 40.1, 35.0, 28.9, 28.7, 25.9.

LRMS (ESI⁻) *m/z*: 468.25 [M-H]⁻ (calculated 468.22 for [C₂₅H₃₀N₃O₆]⁻).

6-(6-((2-Aminoethyl)amino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoic acid (31)



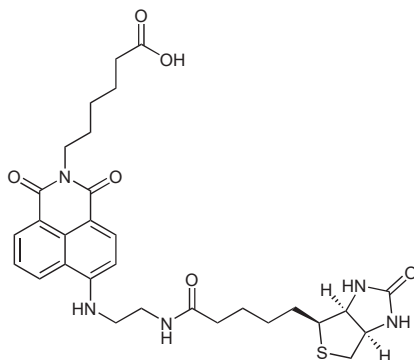
30 (0.113 g, 0.241 mmol) was added to trifluoroacetic acid (2.00 mL, 26.1 mmol) and stirred under nitrogen for 5 h. The trifluoroacetic acid was evaporated and the resulting residue was washed with diethyl ether. The product was dried under a stream of nitrogen to yield **31** (0.0881 g, 0.238 mmol, 99%) as a yellow solid.

¹H NMR (400 MHz, CD₃OD): δ 8.54 (t, *J* = 6.8, 2H), 8.42 (d, *J* = 8.4, 1H), 7.71 (t, *J* = 7.6, 1H), 4.15 - 4.11 (m, 3H), 3.79 (t, *J* = 6.0, 2H), 2.35 - 2.30 (m, 3H), 1.76 - 1.67 (m, 4H), 1.65 - 1.42 (m, 2H).

¹³C NMR (101 MHz, CD₃OD): δ 177.5, 165.8, 165.4, 151.5, 135.4, 132.1, 130.6, 129.2, 125.7, 123.2, 121.9, 110.8, 105.2, 31.6, 40.8, 39.2, 34.8, 28.8, 27.6, 25.7.

LRMS (ESI⁺) *m/z*: 370.18 [M+H]⁺ (calculated 370.42 for [C₂₀H₂₄N₃O₄]⁺).

6-(1,3-Dioxo-6-((2-(5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamido)ethyl)amino)-1*H*-benzo[*de*]isoquinolin-2(3*H*)-yl)hexanoic acid (BNp-COOH)



29 (0.110 g, 0.298 mmol) and **28** (0.120 g, 0.352 mmol) were dissolved in *N,N*-dimethylformamide (20 mL). Triethylamine (0.05 mL, 0.328 mmol) was added and the reaction mixture was stirred under nitrogen for 5 h. The reaction mixture was then poured into 1 M sodium hydroxide (150 mL) after which a precipitate was formed and filtered off. The remaining filtrate was acidified to pH 1 with hydrochloric acid after which a precipitate was formed. This precipitate was filtered and dried *in vacuo* yielding **BNp-COOH** (0.053 g, 0.090 mmol, 30%) as a yellow solid.

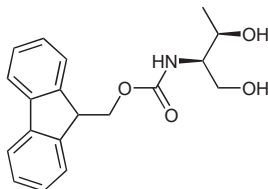
¹H NMR (500 MHz, DMSO-*d*₆): δ 8.61 (d, $J = 8.3$, 1H, Ar*H*), 8.44 (d, $J = 7.2$, 1H, Ar*H*), 8.27 (d, $J = 8.4$, 1H, Ar*H*), 8.13 (t, $J = 5.2$, 1H, NH), 7.83 (t, $J = 4.7$, 1H, NH), 7.70 (t, $J = 7.8$, 1H, Ar*H*), 6.84 (d, $J = 8.6$, 1H, Ar*H*), 6.35 (d, $J = 22.4$, 2H, 2×NH), 4.26 - 4.23 (m, 1H, CH), 4.04 - 3.98 (m, 5H, CH, 2×CH₂), 3.44 (CH₂NHCO)*, 2.94 - 2.90 (m, 1H, CH), 2.76 - 2.72 (m, 1H, CH₂), 2.54 (d, $J = 12.4$, 1H, CH₂'), 2.21 (t, $J = 7.3$, 2H, CH₂CO), 2.10 (t, $J = 7.2$, 2H, CH₂CO), 1.63 - 1.49 (m, 8H, 4×CH₂), 1.43 - 1.21 (m, 6H, 2×CH₂). *resolved by 2D NMR

¹³C NMR (126 MHz, DMSO-*d*₆): δ 175.1, 173.9, 168.6, 168.3, 164.3, 163.5, 163.3, 134.9, 131.4, 125.1, 122.0, 120.4, 108.2, 104.3, 74.1, 61.5, 59.7, 55.9, 35.7, 34.3, 34.0, 32.1, 28.5, 28.4, 28.1, 27.4, 26.5, 25.7, 25.4, 24.7.

LRMS (ESI[−]) m/z : 594.24 [M−H][−] (calculated 594.72 for [C₃₀H₃₆N₅O₆S][−]).

6.2.8 Modified amino acids

(9H-Fluoren-9-yl)methyl ((2R,3R)-1,3-dihydroxybutan-2-yl)carbamate (**32**)

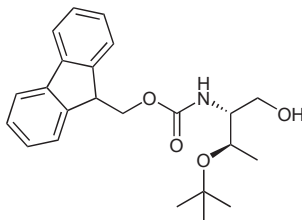


Fmoc-Thr-OH (1.00 g, 2.93 mmol) was dissolved in anhydrous tetrahydrofuran (50 mL). 1 M Borane tetrahydrofuran complex solution (50 mL) was added dropwise over 10 min after which the mixture was stirred at room temperature overnight. The reaction was quenched by adding 1 M hydrochloric acid and stirring for 30 min. This mixture was diluted into (30 mL) and extracted with a 1:1 mixture of tetrahydrofuran and diethyl ether (3 x 100 mL). The organic extracts were combined and washed with 1 M potassium hydroxide (2 x 50 mL) followed by brine (2 x 50 mL) then dried over magnesium sulfate. Solvent was removed by rotary evaporation and purified by column chromatography (SiO₂, 0 -10% methanol in dichloromethane with 0.1% acetic acid) to yield **32** (0.738 g, 2.26 mmol, 77%) as an off-white solid.

¹H NMR (300 MHz, CD₃OD): δ 7.81 (d, J = 7.4, 2H), 7.69 (d, J = 7.3, 2H), 7.37 (dt, J = 14.1, 7.3, 4H), 4.53 – 4.31 (m, 2H), 4.24 (t, J = 6.6, 1H), 4.00 (d, J = 4.1, 1H), 3.78 – 3.50 (m, 3H), 1.17 (d, J = 6.4, 3H).

LRMS (ESI⁺) m/z : 350.18 [M+Na]⁺ (calculated 350.14 for [C₁₉H₂₁NNaO₄]⁺).

(9H-fluoren-9-yl)methyl ((2R,3R)-3-(*tert*-butoxy)-1-hydroxybutan-2-yl)carbamate ((**33**))



A solution of Fmoc-Thr(*t*Bu)-OH (1.02 g, 2.57 mmol) in tetrahydrofuran (40 mL) was cooled to -10 °C. *N*-Methylmorpholine (0.280 mL, 2.55 mmol) and ethyl chloroformate

(0.240 mL, 2.51 mmol) were added successively. The reaction mixture was stirred for 30 min at -10 °C after which sodium borohydride (0.307 g, 8.12 mmol) was added. Methanol (30 mL) was slowly added in portions over 30 min and the reaction was left to stir at room temperature. After this, 1 M hydrochloric acid (50 mL) was added and stirred until the reaction mixture became transparent. The organic solvents were evaporated and the remaining aqueous solution was extracted with dichloromethane (3 x 30 mL). The organic layers were combined and dried over magnesium sulfate before the solvent was removed by rotary evaporation. The crude oil was purified by column chromatography (SiO₂, 20 - 50% ethyl acetate in hexanes) to yield **33** (0.941 g, 2.44 mmol, 98%) as a white solid.

¹H NMR (500 MHz, CDCl₃): δ 7.77 (d, J = 7.5, 2H), 7.61 (d, J = 7.4, 2H), 7.40 (t, J = 7.4, 2H), 7.32 (t, J = 7.4, 2H), 4.48 – 4.37 (m, 2H), 4.24 (t, J = 6.8, 1H), 3.96 (d, J = 3.4, 1H), 3.76 – 3.58 (m, 4H), 1.26 – 1.10 (m, 12H).

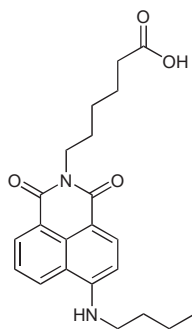
¹³C NMR (126 MHz, CDCl₃): δ 157.2, 144.1, 141.5, 127.8, 127.2, 125.2, 120.1, 74.5, 67.6, 67.0, 64.1, 57.3, 47.5, 28.8, 20.2.

LRMS (ESI⁻) m/z : 382.18 [M-H]⁻ (calculated 382.20 for [C₂₃H₂₈NO₄]⁻).

Characterisation data is in accordance with literature values.³⁵¹

6.2.9 Np-COOH

6-(6-(butylamino)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)hexanoic acid (Np-COOH)



15 (0.750 g, 1.93 mmol) and *N*-butylamine (0.95 mL 9.6 mmol) were dissolved in dimethylsulfoxide (5 mL) and stirred at 70 °C overnight. After this, the reaction mixture was concentrated under a stream of nitrogen. The concentrate was diluted into 2 M hydrochloric acid (20 mL) which induced a yellow precipitate. The mixture was extracted into

dichloromethane (5 x 25 mL). The organic extracts were combined and washed with water (3 x 25 mL) and brine (3 x 25 mL) then dried over magnesium sulfate. The solvent was evaporated *in vacuo* to yield a yellow oil. Petroleum ether was added slowly to induce precipitation of a bright yellow solid which was vacuum filtered. The solid was then purified by column chromatography (SiO₂, 5 - 7% methanol in dichloromethane) to yield **Np-COOH** (0.658 g, 1.70 mmol, 88%) as a bright orange solid.

¹H NMR (300 MHz, CDCl₃): δ 8.58 (d, *J* = 7.3 1H, Ar*H*), 8.44 (d, *J* = 8.4, 1H, Ar*H*), 8.07 (d, *J* = 8.4, 1H, Ar*H*), 7.62 (d, *J* = 8.3, 1H, Ar*H*), 6.72 (d, *J* = 8.5, 1H, Ar*H*), 4.16 (t, *J* = 7.5, 2H, NCH₂), 3.41 (t, *J* = 7.0, 2H, NCH₂), 2.37 (t, *J* = 7.5, 2H, CH₂COOH), 1.80 – 1.69 (m, 6H, 3×CH₂), 1.55 – 1.48 (m, 4H, 2×CH₂), 1.03 (t, *J* = 7.3, 3H, CH₃).

¹³C NMR (75 MHz, CDCl₃): δ 177.0, 164.7, 164.2, 149.4, 134.5, 131.1, 129.8, 125.7, 124.7, 123.2, 120.2, 110.3, 104.4, 43.4, 39.9, 33.5, 31.1, 27.8, 26.6, 24.4, 20.4, 13.9.

LRMS (ESI[−]) *m/z*: 381.18 [M−H][−] (calculated 381.18 for [C₂₂H₂₅N₂O₄][−]).

6.2.10 General peptide synthesis procedure

Loading an amino acid onto resin

Rink amide (100-200 mesh) resin (300 mg) was swelled for 30 min in dichloromethane. The Fmoc group was cleaved from the resin by rocking in a solution of piperidine in *N,N*-dimethyl formamide (20% v/v, 4 mL) for 1 h. The solution was discarded and the resin was washed with *N,N*-dimethylformamide (5 x 5 mL), dichloromethane (5 x 5 mL), then *N,N*-dimethylformamide (5 x 5 mL). The first amino acid (4 equiv.) was activated using DIPEA (8 equiv.) and PyBOP (8 equiv.) in *N,N*-dimethylformamide and left to activate for 2 min. The activated amino acid solution was added to the deprotected resin and shaken for 1.5 h. The solution was discarded and the resin was washed with *N,N*-dimethylformamide (5 x 5 mL), dichloromethane (5 x 5 mL), then *N,N*-dimethylformamide (5 x 5 mL). The loaded resin was then capped to inactivate unreacted amines.

A capping solution was made by combining acetic anhydride and pyridine (3:2). The solution was added to the resin and shaken for 30 min. The solution was discarded and the resin was washed with *N,N*-dimethylformamide (5 x 5 mL) dichloromethane (5 x 5 mL), then *N,N*-dimethylformamide (5 x 5 mL).

Loading an amino alcohol onto resin

2-Chlorotrityl Chloride resin (300 mg) was swelled in dichloromethane for 30 min. The amino alcohol (2 equiv.) was added to a solution of DIPEA (4 equiv.) in *N,N*-dimethylformamide. This solution was added to the resin and left to shake for 18 h. The solution was discarded and the resin was washed with *N,N*-dimethylformamide (5 x 5 mL) dichloromethane (5 x 5 mL), then *N,N*-dimethylformamide (5 x 5 mL). The loaded resin was then capped.

A capping solution of dichloromethane, methanol and DIPEA (17:2:1, 5 mL) was added to the resin and shaken for 1 hour. The solution was discarded and the resin was washed with *N,N*-dimethylformamide (5 x 5 mL) dichloromethane (5 x 5 mL), then *N,N*-dimethylformamide (5 x 5 mL).

Checking resin loading

A small portion of dried loaded resin was weighed exactly. Piperidine in *N,N*-dimethylformamide (20% v/v, 3 mL) was added and left to sit for at least 5 min after which the absorbance was measured at 290 nm. Loading was calculated using an extinction coefficient of $1.65 \text{ cm}^{-1}(\text{mmol/g})^{-1}$.

Manual peptide elongation

To deprotect the Fmoc from the resin bound amino acid, a solution of piperidine in *N,N*-dimethylformamide (20% v/v, 3 mL) was shaken for 3 mins (3 x 5 mL). The solution was discarded and the resin was washed with *N,N*-dimethylformamide (5 x 5 mL) dichloromethane (5 x 5 mL), then *N,N*-dimethylformamide (5 x 5 mL).

The amino acid (2 equiv.) was activated using DIPEA (4 equiv.) and PyBOP (4 equiv.) in *N,N*-dimethylformamide (5 mL) and left to shake for 5 min. The activated amino acid solution was added to the resin and shaken for 1.5 h. The solution was discarded and the resin was washed with *N,N*-dimethylformamide (5 x 5 mL) dichloromethane (5 x 5 mL), then *N,N*-dimethylformamide (5 x 5 mL).

The deprotection and amino acid coupling steps were repeated for each amino acid until the entire peptide was synthesised.

Coupling the fluorophore onto the N-terminus

A solution of piperidine in *N,N*-dimethylformamide (20% v/v) was used to protect the Fmoc of the resin bound amino acid. The acid functionalised fluorescent sensor (1.2 equiv.) was dissolved in *N,N*-dimethylformamide (5 mL) and activated using DIPEA (4 equiv.) and PyBOP (4 equiv.) which was left to stand for 2 min. The activated fluorophore was added to the resin and shaken overnight. The solution was discarded and the resin was washed with *N,N*-dimethylformamide (5 x 5 mL) dichloromethane (5 x 5 mL), then *N,N*-dimethylformamide (5 x 5 mL).

Cleaving and deprotecting the peptide off the resin

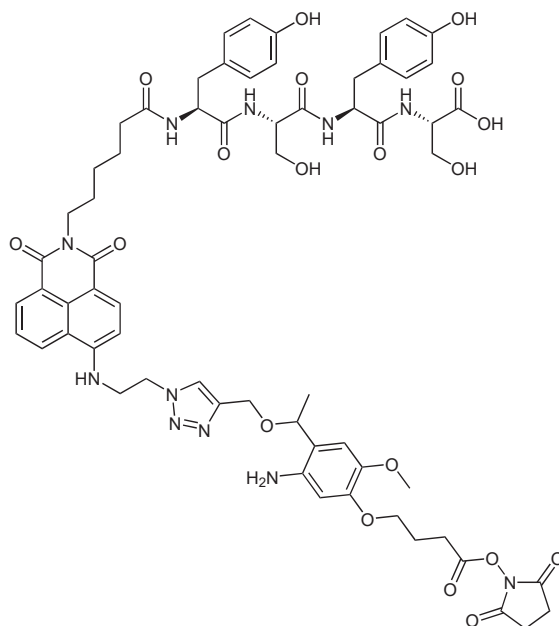
The resin bound **Np-COOH** and **BNp-COOH** conjugated octreotide was cleaved using a solution containing trifluoroacetic acid, water, phenol, thioanisole, and ethanedithiol (82.5:5:5:5:2.5 v/v/v/v/v, 5 mL). Resin bound **NpMT** was cleaved using a solution containing trifluoroacetic acid, water, and triisopropylsilane (95:2.5:2.5, v/v/v, 5 mL). The respective solution was added to the resin and shaken for 1 h. Excess solution was dried under a stream of nitrogen and the cleaved peptide was analysed by HPLC.

Peptide cyclisation

Peptide was dissolved at a concentration of <5 mg/mL in a solution of ammonium bicarbonate (50 mM) at pH 8 and shaken for 4 days. Cyclisation was confirmed using LCMS. The cyclised peptide was isolated by lyophilisation.

6.2.11 Peptides

NpMT

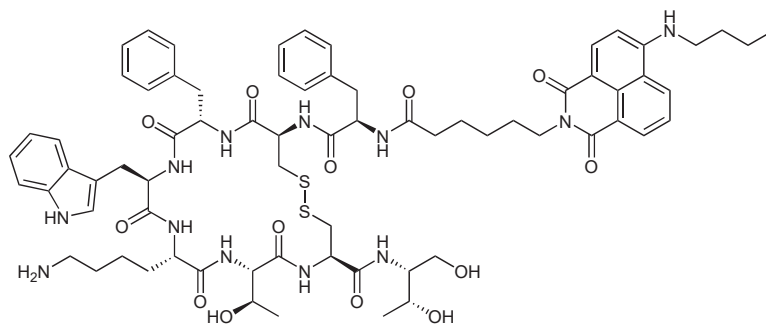


LRMS (ESI⁻) m/z : 1293.05 [M-6H]⁶⁻ (calculated 1293.46 for [C₆₄H₆₇N₁₁O₁₉]⁶⁻).

LRMS (ESI⁺) m/z : 1350.18 [M+2Na+4H]⁶⁺ (calculated 1350.36 for [C₆₄H₇₇N₁₁Na₂O₁₉]⁶⁺).

Analytical HPLC (C18 column, 0 - 100% acetonitrile in water over 30 minutes (0.1% TFA buffer)) retention time: 7.7 min.

Np-COOH octreotide (cyclised)



LRMS (ESI⁻) m/z : 1380.91 [M-2H]²⁻ (calculated 1380.60 for [C₇₁H₈₈N₁₂O₁₃S₂]²⁻).

LCMS (C18 column, 0 - 100% acetonitrile in water over 60 min) m/z : 1385 [M+2H]²⁺ (calculated 1385 for [C₇₁H₉₂N₁₂O₁₃S₂]²⁺). retention time: 42.5 min.

For the absorbance and emission spectra taken using the smeared product, **CyMT** or ethyl ester protected hemicyanine were dissolved in chloroform (1 mg/mL). 10 μ L of this solution was smeared onto a glass slide and was left to dry.

For spectra taken in a matrix dispersion, polymethylmethacrylate (PMMA) was dissolved in chloroform (20 mg/mL) and was stirred for 2h. **CyMT** was dissolved in the PMMA solution at a concentration of 10% (w/v). Dilutions were performed to give final concentrations of 5, 2.5, 1.25 and, 0.625% (w/v). 10 μ L of the solutions was spin coated onto a glass slide using the dynamic dispense technique at 1500 rpm.

The coated slides were mounted onto a 3D printed cradle and placed at 45° to the laser and absorbance paths. UV-Vis absorption spectra were collected on Aligent Technologies Cary 60 UV-Vis Spectrophotometer.

6.4 Electrochemical studies

Cyclic voltammetry

Cyclic voltammograms were obtained on the BAS 100B/W Electrochemical Analyser. Glassy carbon was used as the working electrode and a silver/silver chloride wire was used as the reference electrode. Probes were prepared at a concentration of 0.6 mM in acetonitrile with 0.1 M tetrabutylammonium chloride. The sample was degassed for 20 mins before taking measurements.

6.5 Insulin preparation

6.5.1 CyMT tagged insulin fibril formation

To a solution of human recombinant insulin (0.970 mg, 0.172 nmol) dissolved in *N,N*-dimethylformamide (100 μ L) was added *N,N*-Diisopropylethylamine (10 μ L, 0.13 μ mol). This solution was left to shake for 10 min at 170 rpm. A solution of **CyMT** (0.620 mg, 0.517 nmol) in dimethylsulfoxide (10 μ L) was added drop-wise to the insulin mixture which was then left to stir for 24 h. The mixture was extracted into phosphate buffered solution using a Zeba 7K MWCO (2 mL, Thermo Fisher) spin column. The concentration of the protein fractions were assessed using absorbance spectroscopy at a wavelength of 280 nm

using an extinction coefficient of $1.0675 \text{ cm}^{-1}(\text{mg/mL})^{-1}$. The tagged insulin was dissolved into glycine buffer (20 mM) at a concentration of 1 mg/mL and were fibrilised using a thermomixer at 60 °C and 700 rpm for 18 h.

6.5.2 CyMT tagged insulin fibril slide preparation

ITO slides were treated with poly-L-lysine to assist with adhesion. To do this, the ITO slides were cleaned as described in section 6.3.2. The slide was placed in a humidity chamber before poly-L-lysine solution (200 uL, 0.01% (w/v)) was placed across the entire slide, and incubated for 30 min. The solution was removed and excess was rinsed off with MilliQ water. The residual water was blown off with a stream of nitrogen.

CyMT-tagged insulin fibrils were synthesised as described in section 6.5.1. 20 uL of the fibril solution was dropped onto the poly-L-lysine-treated ITO-coated glass slides. The droplets were left to air dry in the dark.

6.6 Confocal microscopy

6.6.1 CyMT tagged insulin microscopy

CyMT tagged insulin fibrils were prepared as described in section 6.5.2. No imaging media was used for these studies. Fluorescence images were taken on the Olympus Fluoview FV3000 Confocal Microscope using the 561 nm laser. A lambda scan was taken to find the most appropriate collection window, which was determined to be 575 - 675 nm. Images were taken using a 40x objective. Images were processed using FIJI software.

6.6.2 HaloTag microscopy

Halo-CTR1 DLD1 cells were plated into MatTek glass bottom plates at 30% confluence per well and allowed to adhere for 24 h in a humidified incubator at 37 °C with 5% CO₂. BCS (50 µM) or CuSO₄ (30 µM) was added to the cells to incubate for 24 h. After this, the cells were incubated with the desired probe for 20 mins before being rinsed with phosphate-buffered saline (3 x 1 mL). After which FluoroBrite DMEM supplemented with 2% (v/v) foetal bovine serum was used for as an imaging medium.

Cells were imaged in a temperature controlled incubator at 37 °C with 5% CO₂. Images

were obtained on an Olympus FluoView FV1000 confocal microscope using a 60x water objective lens. The probes were excited using the 405 nm laser and a collection window of 500 - 600 nm.

6.7 MALDI studies

6.7.1 Low resolution MALDI-MS

A saturated solution of α -cyano-4-hydroxycinnamic acid (4CCA) in ethanol (0.25 μ L) was dropped onto a polished steel plate and left to dry. The sample was dissolved into TA30, then mixed 1:1 with a saturated solution of 4CCA in TA30. This mixture (0.25 μ L) was dropped on top of the dried droplet and was left to dry. MALDI-MS was taken on the Bruker Autoflex Speed MALDI-TOF.

6.7.2 MALDI-MSI sample preparation

For the images of the probe alone, **CyMT** was dissolved in dichloromethane (0.5 mg/mL). 50 μ L was dropped onto an indium tin oxide (ITO) coated glass slide, and the droplet was smeared to form a thin layer across the slide which was left to dry.

The prepared glass slides were placed in a MTP slide adaptor and a visual scan was taken at 4800 dpi using the Epson perfection v39.

6.7.3 Matrix sublimation

Matrix sublimation was performed by depositing 4CCA matrix (350 mg) in a glass petri dish that was then placed into a vacuum chamber. The glass slide was tape-mounted to the bottom of the cooling insert face-down. The cooling insert was filled with ice. The chamber was evacuated for 4 min, after which the chamber was heated to between 140 - 170 °C under vacuum for 10 min. The chamber was removed from the heat source and allowed to equilibrate to atmospheric pressure before the sample was removed.

6.7.4 MALDI-MSI

Prepared slides were placed on a Bruker solarix 2xR 7T Fourier Transform Ion Cyclotron Resonance Mass Spectrometer equipped with a smart beam II Nd:YAG MALDI laser. Plate

offset was set to 100.0 V. Deflector plate was set to 220.0 V. Laser power 36%. 200 laser shots at a frequency of 1000 Hz with small laser focus. Accumulation during detection was turned on. On-line calibration was turned on for the following m/z related to 4CCA responses: 379.092463, 399.03778, 568.135056, 845.105426. The API high voltage was turned off. The dry gas temperature was set to 30 °C. Funnel 1 voltage was set to 150 V Skimmer 1 was set to 15 V. Funnel RF Amplitude was set to 150 Vpp. Octopole frequency was 5 MHz at an amplitude of 350 Vpp. Collision cell RF frequency was set to 2 MHz with an amplitude of 1000 Vpp. Time of flight delay was set to 0.8 ms at a transfer optics frequency of 6 MHz and an amplitude of 350 Vpp. The transient size was 512 K mass range 150 - 3000. Average scan set to 1. Transients were transformed in magnitude mode with 1 zero fill and a data reduction factor of 95% and full sine apodisation function was applied. Sweep excitation power was set to 19%.

Spectral image was prepared using flexImaging software with 1 region selected that included the 4CCA matrix sublimed over the slide and a second slide without sublimation. 100 μm raster pattern was applied. Laser diameter was 50 μm to prevent oversampling.

6.8 Computational chemistry

All structure files were generated and visualized using IQmol 3.0.1. Quantum calculations were carried out using Q-chem 5.4 or Q-chem 6.0 software packages.^{352,353} The self-consistent field (SCF) method was performed using Pulay’s direct inversion in the iterative subspace (DIIS) algorithm, with a convergence criterion of 10^{-8} for geometry optimizations and 10^{-12} for excited-state calculations. The SG1 standard quadrature grid was used for all calculations unless otherwise stated, except for M11 functional calculations which were performed using SG3.

6.8.1 Geometry optimisations

For **NpFR1** Geometry optimisations were performed using density functional theory (DFT) by utilising the Becke’s three-parameter hybrid functional (B3YLP) method²⁹⁶ and the 6-31G* basis set with the inclusion of solvent by using the conductor-like polarisable continuum model (CPCM).^{299,354,355} A dielectric constant of 78.39 was used as a standard for water.

Several conformers were considered and generated for the geometry optimisations in order to sample multiple structures along the potential energy surface. Frequency analysis was used to determine the nature of the found stationary point. A conformer was concluded to be at a minimum point on the potential energy surface if the frequency analysis returned 0 imaginary frequencies.

To ensure that multiple starting geometries were sampled, a Python script was created for use with PyMol due to the high number of degrees of freedom in the **FCR1** structure. The dihedrals were specified and were generated at either 0°, 90°, 180°, or 270°, while the dihedrals specified for the rotation of the fluorophore were selected to be either 90° or 180°. This generated 128 starting geometry files. Of these, 20 files were discarded due to overlap of the chromophores. The remaining files were optimised using B3LYP/6-31G* level of theory in the gas phase. From here, the 10 lowest energy conformers were then re-optimised using a solvated model to emulate protein binding conditions by using CPCM. The lowest energy conformer was subjected to a frequency analysis to confirm that the structure was at a minimum on the potential energy surface. The lowest energy conformer was found to have 3 imaginary frequencies with amplitudes of 51.92, 45.13, and 25.49, indicating the geometry is not at a minimum. Therefore, the lowest energy conformer geometry was further refined using a finer quadrature grid (17,674 vs 3,816 number of grid points). The frequency was then computed again and 0 imaginary frequencies were found.

6.8.2 Spectral overlap integral

The absorbance spectrum of *N*-ethyl flavin was converted to be expressed in units of $\text{M}^{-1}\text{cm}^{-1}$ vs λ . The conversion of the absorbance units was performed by calculating the extinction coefficient which was used to scale the absorbance values. The emission spectrum of coumarin was normalised. The scaled spectra were imported into a|e - UV/VIS-IR Spectral Software and the spectral overlap integral was calculated from 380 - 600 nm using the 'Calculate spectral overlap integral' function.

6.8.3 Transition dipole moments

The optimised geometry of **FCR1** was used to extract the structures of the coumarin and flavin portions individually. To ensure the most accurate representation of the relevant

structures, the point of separation was determined to be one carbon atom away from the last molecule involved in the structure conjugation.

For the flavin portion, TD-M11/6-31+G* level of theory was used to calculate the first eight singlet excited states from the ground state. The first excited state was found to have the most significant transition strength. The transition dipole moment of this was extracted from the Q-chem output file.

For the coumarin portion, TD-M11/6-31+G* level of theory was used to calculate the first eight singlet excited states. The first excited state exhibited the most significant transition strength, thus a geometry optimisation was performed on this state. After optimisation, further TD-M11/6-31+G* was used to calculate the first five singlet excited states. The first excited state was found to have the most significant transition strength. The transition dipole moment of this state was extracted from the Q-chem output file.

6.8.4 Orientation factor calculation

The midpoint of each fluorophore in **FCR1** was determined by extracting their coordinates from the optimised structure. The extent of each independent fluorophore was defined as one carbon atom away from the last atom involved in the conjugation, to ensure accurate consideration of the relevant structure. The coordinates of each chromophore were averaged to obtain the midpoint 3D coordinates, which was converted into a vector. The vector of the transition dipole moments were determined as described in section 6.8.3. The angle between the two vectors was calculated using a rearrangement of the dot product formula given in equation 6.1.

$$\theta = \arccos \left(\frac{a \cdot b}{|a| \times |b|} \right) \quad (6.1)$$

where θ is the angle between two vectors, a is vector 1, and b is vector 2.

6.8.5 Excited state calculations

Excited state calculations were first performed using the CIS³⁵⁶ and CIS(D)³⁵⁷ methods using either 6-31G* or 6-31+G* basis sets. Then TD-DFT was performed using using PBE0³⁵⁸, CAM-B3LYP³⁵⁹, B3LYP²⁹⁶, and M06-2X³⁶⁰ functional. For these calculations,

6-31G*, 6-31+G*, 6-311++G**,^{361–364} and aug-cc-PVTZ³⁶⁵ basis sets were considered. MOM³⁶⁶ was performed at B3LYP/6-31+G* level of theory using the SG1 grid or the SG3 grid. IMOM³⁶⁷ was calculated using B3LYP/6-31+G* level of theory. Solvated calculations were performed using CPCM with a dielectric constant of 78.39.

6.9 Computational Biology

6.9.1 Autodock 4

Autodock4 calculations were performed in the manner described by Bianco *et al.*²⁹¹ The protein structure (PDBID 6U32) was prepared by removing the complexed ligand. The protein structure was prepared by removing implicit waters and adding all hydrogen atoms. The ligand structures of **NpFR1-Tag** and **FCR1-Tag** were generated in iQmol. An alignment file the ligands within the HaloTag binding pocket was generated. The flexible protein residue was specified to be APS106. Calculation parameter files were generated using Autogrid4 for both the receptor and ligand files. These files were used to optimise the ligand binding using Autodock4. Output files were visualised using PyMOL.

6.9.2 Molecular docking

The crystal structure of SSTR2 was downloaded from the protein database (PDBID: 7XNA). This was selected based on its overall score and because it was obtained using X-ray crystallography, which typically results in more accurate descriptions of the crystal structure. To prepare the structure for docking, the native ligand was extracted and removed from the protein binding pocket. Using Maestro software, the protein preparation wizard was used to correct for errors in the crystal structure including misalignments, overlaps, and missing side chains.³⁶⁸ Implicit waters in the structure were removed, then all hydrogens were added. The ligand structure of octreotide and **BNp-COOH** conjugated octreotide were drawn and imported into Maestro. The geometry of each ligand was optimised using molecular mechanics in Maestro by selecting the Macromodel minimisation program. The lowest energy conformations were used for molecular dynamic simulations.

Molecular dynamic simulations of the ligands within the SSTR2 binding pocket were performed using Desmond with the OPLSv4 forcefield.³⁶⁹ The simulation was performed

at pH 7 in a T3I3P solvent and 25°C for 100 ns. These interactions were visualised using mm-GBSA in Meastro, scanning across the simulation frames to elucidate bonding type (eg hydrogen bonding or pi-pi interactions etc).³⁷⁰ Free binding energy was calculated in the same manner, and are expressed as percentages.

References

- [1] Lakowicz, J. R. In *Principles of Fluorescence Spectroscopy*; Lakowicz, J. R., Ed.; Springer US: Boston, MA, 2006.
- [2] Sabnis, R. W. *Handbook of Fluorescent Dyes and Probes*; John Wiley & Sons, Inc: Hoboken, NJ, USA, 2015.
- [3] Yuan, L.; Lin, W.; Zheng, K.; He, L.; Huang, W. Far-red to near infrared analyte-responsive fluorescent probes based on organic fluorophore platforms for fluorescence imaging. *Chem. Soc. Rev.* **2013**, *42*, 622–661.
- [4] Lavis, L. D.; Raines, R. T. Bright Ideas for Chemical Biology. *ACS Chemical Biology* **2008**, *3*, 142–155.
- [5] Zheng, Q.; Lavis, L. D. Development of photostable fluorophores for molecular imaging. *Current Opinion in Chemical Biology* **2017**, *39*, 32–38.
- [6] Yeo, J. H.; New, E. J. *Molecular Fluorescent Sensors for Cellular Studies*; Wiley, 2022; pp 37–62.
- [7] Baeyer, A. Ueber eine neue Klasse von Farbstoffen. *Berichte der deutschen chemischen Gesellschaft* **1871**, *4*, 555–558.
- [8] Cavallo, C.; De Laurentis, C.; Vetrano, I. G.; Falco, J.; Broggi, M.; Schiariti, M.; Feroli, P.; Acerbi, F. The utilization of fluorescein in brain tumor surgery: a systematic review. *Journal of Neurosurgical Sciences* **2018**, *62*, 690–703.
- [9] Le Guern, F.; Mussard, V.; Gaucher, A.; Rottman, M.; Prim, D. Fluorescein Derivatives as Fluorescent Probes for pH Monitoring along Recent Biological Applications. *International Journal of Molecular Sciences* **2020**, *21*, 9217.

- [10] S, K.; Sam, B.; George, L.; N, S. Y.; Varghese, A. Fluorescein Based Fluorescence Sensors for the Selective Sensing of Various Analytes. *Journal of Fluorescence* **2021**, *31*, 1251–1276.
- [11] McKinnon, K. M. Flow Cytometry: An Overview. *Current Protocols in Immunology* **2018**, *120*, 5.1.1–5.1.11.
- [12] Butler, J. E. Enzyme-Linked Immunosorbent Assay. *Journal of Immunoassay* **2000**, *21*, 165–209.
- [13] Mahmoudian, J.; Hadavi, R.; Jeddi-Tehrani, M.; Mahmoudi, A. R.; Bayat, A. A.; Shaban, E.; Vafakhah, M.; Darzi, M.; Tarahomi, M.; Ghods, R. Comparison of the Photobleaching and Photostability Traits of Alexa Fluor 568- and Fluorescein Isothiocyanate- conjugated Antibody. *Cell journal* **2011**, *13*, 169–72.
- [14] Beija, M.; Afonso, C. A. M.; Martinho, J. M. G. Synthesis and applications of Rhodamine derivatives as fluorescent probes. *Chemical Society Reviews* **2009**, *38*, 2410.
- [15] Zhang, Q.; Wong, K. M.-C. Photophysical, ion-sensing and biological properties of rhodamine-containing transition metal complexes. *Coordination Chemistry Reviews* **2020**, *416*, 213336.
- [16] Chen, F.; Liu, W.; Li, H.; Deng, T.; Xing, B.; Liu, F. Rhodamine Fluorophores for STED Super-Resolution Biological Imaging. *Analysis & Sensing* **2022**, *2*, e202100066.
- [17] Papillon, Y.; Buffiere, L.; Butet, A. Rhodamine B as a collective marker for studying movements of small mammals. *Acta Theriologica* **2002**, *47*, 491–497.
- [18] McKay, I. C.; Forman, D.; White, R. G. A comparison of fluorescein isothiocyanate and lissamine rhodamine (RB 200) as labels for antibody in the fluorescent antibody technique. *Immunology* **1981**, *43*, 591–602.
- [19] Leslie, K. G.; Jacquemin, D.; New, E. J.; Jolliffe, K. A. Expanding the Breadth of 4-Amino-1,8-naphthalimide Photophysical Properties through Substitution of the Naphthalimide Core. *Chemistry - A European Journal* **2018**, *24*, 5569–5573.
- [20] Kumari, R.; Sunil, D.; Ningthoujam, R. S. Naphthalimides in fluorescent imaging of tumor hypoxia – An up-to-date review. *Bioorganic Chemistry* **2019**, *88*, 102979.

- [21] Ranjana, M.; Sunil, D. Naphthalimide derivatives as fluorescent probes for imaging endogenous gasotransmitters. *Chemico-Biological Interactions* **2022**, *363*, 110022.
- [22] Dong, H.-Q.; Wei, T.-B.; Ma, X.-Q.; Yang, Q.-Y.; Zhang, Y.-F.; Sun, Y.-J.; Shi, B.-B.; Yao, H.; Zhang, Y.-M.; Lin, Q. 1,8-Naphthalimide-based fluorescent chemosensors: recent advances and perspectives. *Journal of Materials Chemistry C* **2020**, *8*, 13501–13529.
- [23] Katerinopoulos, H. The Coumarin Moiety as Chromophore of Fluorescent Ion Indicators in Biological Systems. *Current Pharmaceutical Design* **2004**, *10*, 3835–3852.
- [24] Li, W.-h.; Zheng, G. Photoactivatable fluorophores and techniques for biological imaging applications. *Photochemical & Photobiological Sciences* **2012**, *11*, 460–471.
- [25] Raunio, H.; Pentikäinen, O.; Juvonen, R. O. Coumarin-Based Profluorescent and Fluorescent Substrates for Determining Xenobiotic-Metabolizing Enzyme Activities In Vitro. *International Journal of Molecular Sciences* **2020**, *21*, 4708.
- [26] Krzeszewski, M.; Vakuliuk, O.; Gryko, D. T. Color-Tunable Fluorescent Dyes Based on Benzo[c]coumarin. *European Journal of Organic Chemistry* **2013**, *2013*, 5631–5644.
- [27] Gorka, A. P.; Nani, R. R.; Schnermann, M. J. Harnessing Cyanine Reactivity for Optical Imaging and Drug Delivery. *Accounts of Chemical Research* **2018**, *51*, 3226–3235.
- [28] Laramie, M. D.; Levitz, A.; Henary, M. Cyanine and Squaric Acid Metal Sensors. *Sensors and Actuators B: Chemical* **2017**, *243*, 1191–1204.
- [29] Fei, G.; Ma, S.; Wang, C.; Chen, T.; Li, Y.; Liu, Y.; Tang, B.; James, T. D.; Chen, G. Imaging strategies using cyanine probes and materials for biomedical visualization of live animals. *Coordination Chemistry Reviews* **2021**, *447*, 214134.
- [30] Kand, D.; Liu, P.; Navarro, M. X.; Fischer, L. J.; Rousso-Noori, L.; Friedmann-Morvinski, D.; Winter, A. H.; Miller, E. W.; Weinstain, R. Water-Soluble BODIPY Photocages with Tunable Cellular Localization. *Journal of the American Chemical Society* **2020**, *142*, 4970–4974.

- [31] Chazotte, B. Labeling Lysosomes in Live Cells with LysoTracker. *Cold Spring Harbor Protocols* **2011**, 2011, pdb.prot5571.
- [32] Thorn, K. A quick guide to light microscopy in cell biology. *Molecular Biology of the Cell* **2016**, 27, 219–222.
- [33] Smith, C. L. Basic Confocal Microscopy. *Current Protocols in Neuroscience* **2011**, 56.
- [34] van Munster, E. B.; Gadella, T. W. J. *Microscopy Techniques*; 2005; pp 143–175.
- [35] Schermelleh, L.; Ferrand, A.; Huser, T.; Eggeling, C.; Sauer, M.; Biehlmaier, O.; Drummen, G. P. C. Super-resolution microscopy demystified. *Nature Cell Biology* **2019**, 21, 72–84.
- [36] Adan, A.; Alizada, G.; Kiraz, Y.; Baran, Y.; Nalbant, A. Flow cytometry: basic principles and applications. *Critical Reviews in Biotechnology* **2017**, 37, 163–176.
- [37] Cooper, G. M. *The Cell: A Molecular Approach*. 2nd edition; Sinauer Associates 2000, 2000.
- [38] Casey Cunningham, C. Actin structural proteins in cell motility. *Cancer and Metastasis Review* **1992**, 11, 69–77.
- [39] Rose, A. M.; Valdes, R. Understanding the sodium pump and its relevance to disease. *Clinical Chemistry* **1994**, 40, 1674–1685.
- [40] Thomas, C.; Lumb, A. B. Physiology of haemoglobin. *Continuing Education in Anaesthesia Critical Care & Pain* **2012**, 12, 251–256.
- [41] Bahar, A.; Ren, D. Antimicrobial Peptides. *Pharmaceuticals* **2013**, 6, 1543–1575.
- [42] Oyston, P. C. F.; Fox, M. A.; Richards, S. J.; Clark, G. C. Novel peptide therapeutics for treatment of infections. *Journal of Medical Microbiology* **2009**, 58, 977–987.
- [43] Zhang, L.; Parente, J.; Harris, S. M.; Woods, D. E.; Hancock, R. E. W.; Falla, T. J. Antimicrobial Peptide Therapeutics for Cystic Fibrosis. *Antimicrobial Agents and Chemotherapy* **2005**, 49, 2921–2927.
- [44] Gaebler, C. *et al.* Evolution of antibody immunity to SARS-CoV-2. *Nature* **2021**, 591, 639–644.

- [45] Lazazzera, B. A.; Grossman, A. D. The ins and outs of peptide signaling. *Trends in Microbiology* **1998**, *6*, 288–294.
- [46] Pollak, M. Insulin and insulin-like growth factor signalling in neoplasia. *Nature Reviews Cancer* **2008**, *8*, 915–928.
- [47] Lu, C.; Lam, H. N.; Menon, R. K. New Members of the Insulin Family: Regulators of Metabolism, Growth and Now ... Reproduction. *Pediatric Research* **2005**, *57*, 70R–73R.
- [48] Vizi, E.; Lendvai, B. Modulatory role of presynaptic nicotinic receptors in synaptic and non-synaptic chemical communication in the central nervous system. *Brain Research Reviews* **1999**, *30*, 219–235.
- [49] Flannery, S.; Bowie, A. G. The interleukin-1 receptor-associated kinases: Critical regulators of innate immune signalling. *Biochemical Pharmacology* **2010**, *80*, 1981–1991.
- [50] Boyman, O.; Sprent, J. The role of interleukin-2 during homeostasis and activation of the immune system. *Nature Reviews Immunology* **2012**, *12*, 180–190.
- [51] Gaestel, M.; Kotlyarov, A.; Kracht, M. Targeting innate immunity protein kinase signalling in inflammation. *Nature Reviews Drug Discovery* **2009**, *8*, 480–499.
- [52] Miller, E. J.; Lappin, S. L. *StatPearls*; 2022.
- [53] Tikellis, C.; Bernardi, S.; Burns, W. C. Angiotensin-converting enzyme 2 is a key modulator of the renin–angiotensin system in cardiovascular and renal disease. *Current Opinion in Nephrology and Hypertension* **2011**, *20*, 62–68.
- [54] Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270–273.
- [55] Zoufaly, A. *et al.* Human recombinant soluble ACE2 in severe COVID-19. *The Lancet Respiratory Medicine* **2020**, *8*, 1154–1158.
- [56] Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veasley, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281–292.e6.

- [57] Slamon, D.; Clark, G.; Wong, S.; Levin, W.; Ullrich, A.; McGuire, W. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **1987**, *235*, 177–182.
- [58] Slamon, D.; Godolphin, W.; Jones, L.; Holt, J.; Wong, S.; Keith, D.; Levin, W.; Stuart, S.; Udove, J.; Ullrich, A.; Et, A. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* **1989**, *244*, 707–712.
- [59] Hudis, C. A. Trastuzumab — Mechanism of Action and Use in Clinical Practice. *New England Journal of Medicine* **2007**, *357*, 39–51.
- [60] Seidman, A.; Hudis, C.; Pierri, M. K.; Shak, S.; Paton, V.; Ashby, M.; Murphy, M.; Stewart, S. J.; Keefe, D. Cardiac dysfunction in the trastuzumab clinical trials experience. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **2002**, *20*, 1215–21.
- [61] Guglin, M.; Hartlage, G.; Reynolds, C.; Chen, R.; Patel, V. Trastuzumab-Induced Cardiomyopathy: Not as Benign as it Looks? A Retrospective Study. *Journal of Cardiac Failure* **2009**, *15*, 651–657.
- [62] Quasthoff, S.; Hartung, H. P. Chemotherapy-induced peripheral neuropathy. *Journal of Neurology* **2002**, *249*, 9–17.
- [63] Verstappen, C. C. P.; Heimans, J. J.; Hoekman, K.; Postma, T. J. Neurotoxic Complications of Chemotherapy in Patients with Cancer. *Drugs* **2003**, *63*, 1549–1563.
- [64] Hood, A. F. Cutaneous Side Effects of Cancer Chemotherapy. *Medical Clinics of North America* **1986**, *70*, 187–209.
- [65] Monsuez, J.-J.; Charniot, J.-C.; Vignat, N.; Artigou, J.-Y. Cardiac side-effects of cancer chemotherapy. *International Journal of Cardiology* **2010**, *144*, 3–15.
- [66] King, P. D.; Perry, M. C. Hepatotoxicity of Chemotherapy. *The Oncologist* **2001**, *6*, 162–176.
- [67] Wang, X.; Guo, Z. Targeting and delivery of platinum-based anticancer drugs. *Chem. Soc. Rev.* **2013**, *42*, 202–224.

- [68] Patra, J. K.; Das, G.; Fraceto, L. F.; Campos, E. V. R.; Rodriguez-Torres, M. d. P.; Acosta-Torres, L. S.; Diaz-Torres, L. A.; Grillo, R.; Swamy, M. K.; Sharma, S.; Habtemariam, S.; Shin, H.-S. Nano based drug delivery systems: recent developments and future prospects. *Journal of Nanobiotechnology* **2018**, *16*, 71.
- [69] Pienta, K. Preclinical mechanisms of action of docetaxel and docetaxel combinations in prostate cancer. *Seminars in Oncology* **2001**, *28*, 3–7.
- [70] Chan, S. *et al.* Prospective Randomized Trial of Docetaxel Versus Doxorubicin in Patients With Metastatic Breast Cancer. *Journal of Clinical Oncology* **1999**, *17*, 2341–2341.
- [71] Bostwick, D. G.; Pacelli, A.; Blute, M.; Roche, P.; Murphy, G. P. Prostate specific membrane antigen expression in prostatic intraepithelial neoplasia and adenocarcinoma: a study of 184 cases. *Cancer* **1998**, *82*, 2256–61.
- [72] Maresca, K. P.; Hillier, S. M.; Femia, F. J.; Keith, D.; Barone, C.; Joyal, J. L.; Zimmerman, C. N.; Kozikowski, A. P.; Barrett, J. A.; Eckelman, W. C.; Babich, J. W. A Series of Halogenated Heterodimeric Inhibitors of Prostate Specific Membrane Antigen (PSMA) as Radiolabeled Probes for Targeting Prostate Cancer. *Journal of Medicinal Chemistry* **2009**, *52*, 347–357.
- [73] Autio, K. A. *et al.* Safety and Efficacy of BIND-014, a Docetaxel Nanoparticle Targeting Prostate-Specific Membrane Antigen for Patients With Metastatic Castration-Resistant Prostate Cancer. *JAMA Oncology* **2018**, *4*, 1344.
- [74] Bernard Stewart, P. K. World Cancer Report. **2003**,
- [75] Shaib, Y. H.; Davila, J. A.; El-Serag, H. B. The epidemiology of pancreatic cancer in the United States: changes below the surface. *Alimentary Pharmacology and Therapeutics* **2006**, *24*, 87–94.
- [76] Sohn, T.; Yeo, C.; Cameron, J.; Koniaris, L.; Kaushal, S.; Abrams, R.; Sauter, P.; Coleman, J.; Hruban, R.; Lillemoe, K. Resected adenocarcinoma of the pancreas?616 patients: results, outcomes, and prognostic indicators. *Journal of Gastrointestinal Surgery* **2000**, *4*, 567–579.

- [77] Stathis, A.; Moore, M. J. Advanced pancreatic carcinoma: current treatment and future challenges. *Nature Reviews Clinical Oncology* **2010**, *7*, 163–172.
- [78] Oberstein, P. E.; Olive, K. P. Pancreatic cancer: why is it so hard to treat? *Therapeutic Advances in Gastroenterology* **2013**, *6*, 321–337.
- [79] Chen, C.; Wu, C. Q.; Chen, T. W.; Tang, M. Y.; Zhang, X. M. Molecular Imaging with MRI: Potential Application in Pancreatic Cancer. *BioMed Research International* **2015**, *2015*, 1–10.
- [80] Andrianifahanana, M.; Moniaux, N.; Schmied, B. M.; Ringel, J.; Friess, H.; Hollingsworth, M. A.; Büchler, M. W.; Aubert, J. P.; Batra, S. K. Mucin (MUC) gene expression in human pancreatic adenocarcinoma and chronic pancreatitis: a potential role of MUC4 as a tumor marker of diagnostic significance. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2001**, *7*, 4033–40.
- [81] Wu, S.-C.; Chen, Y.-J.; Lin, Y.-J.; Wu, T.-H.; Wang, Y.-M. Development of a Mucin4-Targeting SPIO Contrast Agent for Effective Detection of Pancreatic Tumor Cells in Vitro and in Vivo. *Journal of Medicinal Chemistry* **2013**, *56*, 9100–9109.
- [82] Lee, D. Y.; Li, K. C. P. Molecular Theranostics: A Primer for the Imaging Professional. *American Journal of Roentgenology* **2011**, *197*, 318–324.
- [83] European Medicines Agency Authorization details for Lutathera® in Europe. 2017; <https://www.ema.europa.eu/en/medicines/human/EPAR/lutathera#authorisation-details-section>.
- [84] Food and Drug Administration FDA Letter of Approval for LUTATHERA®. 2018; https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2018/2087000rig1s0001tr.pdf.
- [85] Reubi, J. C. Somatostatin and Other Peptide Receptors as Tools for Tumor Diagnosis and Treatment. *Neuroendocrinology* **2004**, *80*, 51–56.
- [86] Strosberg, J. *et al.* Health-Related Quality of Life in Patients With Progressive Midgut

- Neuroendocrine Tumors Treated With 177 Lu-Dotatate in the Phase III NETTER-1 Trial. *Journal of Clinical Oncology* **2018**, *36*, 2578–2584.
- [87] Engvall, E.; Perlmann, P. Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry* **1971**, *8*, 871–874.
- [88] KURIEN, B.; SCOFIELD, R. Western blotting. *Methods* **2006**, *38*, 283–293.
- [89] Hogrefe, W.; Su, X.; Song, J.; Ashley, R.; Kong, L. Detection of Herpes Simplex Virus Type 2-Specific Immunoglobulin G Antibodies in African Sera by Using Recombinant gG2, Western Blotting, and gG2 Inhibition. *Journal of Clinical Microbiology* **2002**, *40*, 3635–3640.
- [90] Magaki, S.; Hojat, S. A.; Wei, B.; So, A.; Yong, W. H. An Introduction to the Performance of Immunohistochemistry. *Methods in molecular biology (Clifton, N.J.)* **2019**, *1897*, 289–298.
- [91] Fiore, C.; Bailey, D.; Conlon, N.; Wu, X.; Martin, N.; Fiorentino, M.; Finn, S.; Fall, K.; Andersson, S.-O.; Andren, O.; Loda, M.; Flavin, R. Utility of multispectral imaging in automated quantitative scoring of immunohistochemistry. *Journal of Clinical Pathology* **2012**, *65*, 496–502.
- [92] Koh, J.; Kwak, Y.; Kim, J.; Kim, W. H. High-Throughput Multiplex Immunohistochemical Imaging of the Tumor and Its Microenvironment. *Cancer research and treatment* **2020**, *52*, 98–108.
- [93] Smith, G. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **1985**, *228*, 1315–1317.
- [94] Cwirla, S. E.; Peters, E. A.; Barrett, R. W.; Dower, W. J. Peptides on phage: a vast library of peptides for identifying ligands. *Proceedings of the National Academy of Sciences* **1990**, *87*, 6378–6382.
- [95] Smith, G. P.; Petrenko, V. A. Phage Display. *Chemical Reviews* **1997**, *97*, 391–410.
- [96] Kehoe, J. W.; Kay, B. K. Filamentous Phage Display in the New Millennium. *Chemical Reviews* **2005**, *105*, 4056–4072.

- [97] Brown, K. Peptidic Tumor Targeting Agents: The Road from Phage Display Peptide Selections to Clinical Applications. *Current Pharmaceutical Design* **2010**, *16*, 1040–1054.
- [98] Minter, R. Phage Display – A Practical Approach. *Briefings in Functional Genomics and Proteomics* **2005**, *3*, 391–392.
- [99] Hoess, R. H. Protein Design and Phage Display. *Chemical Reviews* **2001**, *101*, 3205–3218.
- [100] Hanes, J.; Pluckthun, A. In vitro selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences* **1997**, *94*, 4937–4942.
- [101] Wilson, D. S.; Keefe, A. D.; Szostak, J. W. The use of mRNA display to select high-affinity protein-binding peptides. *Proceedings of the National Academy of Sciences* **2001**, *98*, 3750–3755.
- [102] Lamla, T.; Erdmann, V. A. Searching Sequence Space for High-affinity Binding Peptides using Ribosome Display. *Journal of Molecular Biology* **2003**, *329*, 381–388.
- [103] Cho, G.; Keefe, A. D.; Liu, R.; Wilson, D. S.; Szostak, J. W. Constructing high complexity synthetic libraries of long ORFs using In Vitro selection. *Journal of Molecular Biology* **2000**, *297*, 309–319.
- [104] Wolff, J.; Covelli, I. Factors in the Iodination of Histidine in Proteins. *European Journal of Biochemistry* **1969**, *9*, 371–377.
- [105] Tsomides, T.; Eisen, H. Stoichiometric Labeling of Peptides by Iodination on Tyrosyl or Histidyl Residues. *Analytical Biochemistry* **1993**, *210*, 129–135.
- [106] Schumacher, T. N.; Tsomides, T. J. In Vitro Radiolabeling of Peptides and Proteins. *Current Protocols in Protein Science* **1995**, *00*.
- [107] Chait, B. T. Mass Spectrometry: Bottom-Up or Top-Down? *Science* **2006**, *314*, 65–66.

- [108] Ossipova, E.; Fenyő, D.; Eriksson, J. Optimizing search conditions for the mass fingerprint-based identification of proteins. *PROTEOMICS* **2006**, *6*, 2079–2085.
- [109] Wardyn, J. D.; Jeyasekharan, A. D. *eLS*; John Wiley & Sons, Ltd: Chichester, UK, 2018; pp 1–9.
- [110] Miller, D. M.; Shakes, D. C. *Immunofluorescence Microscopy*; 1995; Chapter 16, pp 365–394.
- [111] Shimomura, O.; Johnson, F. H.; Saiga, Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *Journal of Cellular and Comparative Physiology* **1962**, *59*, 223–239.
- [112] Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **1992**, *111*, 229–233.
- [113] Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W.; Prasher, D. Green fluorescent protein as a marker for gene expression. *Science* **1994**, *263*, 802–805.
- [114] Inouye, S.; Tsuji, F. I. Aequorea green fluorescent protein. *FEBS Letters* **1994**, *341*, 277–280.
- [115] Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiological Reviews* **2010**, *90*, 1103–1163.
- [116] Tsien, R. Y. The Green Fluorescent Protein. *Annual Review of Biochemistry* **1998**, *67*, 509–544.
- [117] Cranfill, P. J.; Sell, B. R.; Baird, M. A.; Allen, J. R.; Lavagnino, Z.; de Gruiter, H. M.; Kremers, G.-J.; Davidson, M. W.; Ustione, A.; Piston, D. W. Quantitative assessment of fluorescent proteins. *Nature Methods* **2016**, *13*, 557–562.
- [118] Montecinos-Franjola, F.; Bauer, B. L.; Mears, J. A.; Ramachandran, R. GFP fluorescence tagging alters dynamin-related protein 1 oligomerization dynamics and creates disassembly-refractory puncta to mediate mitochondrial fission. *Scientific Reports* **2020**, *10*, 14777.

- [119] Miller, L. W.; Cornish, V. W. Selective chemical labeling of proteins in living cells. *Current Opinion in Chemical Biology* **2005**, *9*, 56–61.
- [120] Griffin, B. A. Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* **1998**, *281*, 269–272.
- [121] Webb, J. L. *Enzyme and metabolic inhibitors*.; Academic Press,; New York,, 1963; pp 595–819.
- [122] Whittaker, V. P. An experimental investigation of the “ring hypothesis” of arsenical toxicity. *Biochemical Journal* **1947**, *41*, 56–62.
- [123] Andresen, M.; Schmitz-Salue, R.; Jakobs, S. Short Tetracysteine Tags to β -Tubulin Demonstrate the Significance of Small Labels for Live Cell Imaging. *Molecular Biology of the Cell* **2004**, *15*, 5616–5622.
- [124] Stroffekova, K.; Proenza, C.; Beam, K. The protein-labeling reagent FLASH-EDT 2 binds not only to CCXXCC motifs but also non-specifically to endogenous cysteine-rich proteins. *Pflügers Archiv European Journal of Physiology* **2001**, *442*, 859–866.
- [125] Martin, B. R.; Giepmans, B. N. G.; Adams, S. R.; Tsien, R. Y. Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nature Biotechnology* **2005**, *23*, 1308–1314.
- [126] Scheck, R. A.; Schepartz, A. Surveying Protein Structure and Function Using Bis-Arsenical Small Molecules. *Accounts of Chemical Research* **2011**, *44*, 654–665.
- [127] Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. New Biarsenical Ligands and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. *Journal of the American Chemical Society* **2002**, *124*, 6063–6076.
- [128] Adams, S. R.; Tsien, R. Y. Preparation of the membrane-permeant biarsenicals FAsH-EDT2 and ReAsH-EDT2 for fluorescent labeling of tetracysteine-tagged proteins. *Nature Protocols* **2008**, *3*, 1527–1534.
- [129] Hoffmann, C.; Gaietta, G.; Bünemann, M.; Adams, S. R.; Oberdorff-Maass, S.; Behr, B.; Vilardaga, J.-P.; Tsien, R. Y.; Ellisman, M. H.; Lohse, M. J. A FAsH-

- based FRET approach to determine G protein-coupled receptor activation in living cells. *Nature Methods* **2005**, *2*, 171–176.
- [130] Gelman, H.; Wirth, A. J.; Gruebele, M. ReAsH as a Quantitative Probe of In-Cell Protein Dynamics. *Biochemistry* **2016**, *55*, 1968–1976.
- [131] Roberti, M. J.; Bertoncini, C. W.; Klement, R.; Jares-Erijman, E. A.; Jovin, T. M. Fluorescence imaging of amyloid formation in living cells by a functional, tetracysteine-tagged α -synuclein. *Nature Methods* **2007**, *4*, 345–351.
- [132] Bhunia, A. K.; Miller, S. C. Labeling Tetracysteine-Tagged Proteins with a SplAsH of Color: A Modular Approach to Bis-Arsenical Fluorophores. *ChemBioChem* **2007**, *8*, 1642–1645.
- [133] Tour, O.; Adams, S. R.; Kerr, R. A.; Meijer, R. M.; Sejnowski, T. J.; Tsien, R. W.; Tsien, R. Y. Calcium Green FAsH as a genetically targeted small-molecule calcium indicator. *Nature Chemical Biology* **2007**, *3*, 423–431.
- [134] Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature Biotechnology* **2003**, *21*, 86–89.
- [135] Juillerat, A.; Gronemeyer, T.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K. Directed Evolution of O6-Alkylguanine-DNA Alkyltransferase for Efficient Labeling of Fusion Proteins with Small Molecules In Vivo. *Chemistry & Biology* **2003**, *10*, 313–317.
- [136] Juillerat, A.; Heinis, C.; Sielaff, I.; Barnikow, J.; Jaccard, H.; Kunz, B.; Tersikh, A.; Johnsson, K. Engineering Substrate Specificity of O6-Alkylguanine-DNA Alkyltransferase for Specific Protein Labeling in Living Cells. *ChemBioChem* **2005**, *6*, 1263–1269.
- [137] Gronemeyer, T.; Chidley, C.; Juillerat, A.; Heinis, C.; Johnsson, K. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling. *Protein Engineering Design and Selection* **2006**, *19*, 309–316.

- [138] Mollwitz, B.; Brunk, E.; Schmitt, S.; Pojer, F.; Bannwarth, M.; Schiltz, M.; Rothlisberger, U.; Johnsson, K. Directed Evolution of the Suicide Protein O⁶-Alkylguanine-DNA Alkyltransferase for Increased Reactivity Results in an Alkylated Protein with Exceptional Stability. *Biochemistry* **2012**, *51*, 986–994.
- [139] Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K. An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chemistry & Biology* **2008**, *15*, 128–136.
- [140] Los, G. V. *et al.* HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chemical Biology* **2008**, *3*, 373–382.
- [141] Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. Ligand-directed tosyl chemistry for protein labeling in vivo. *Nature Chemical Biology* **2009**, *5*, 341–343.
- [142] Steinberg, I.; Huland, D. M.; Vermesh, O.; Frostig, H. E.; Tummers, W. S.; Gambhir, S. S. Photoacoustic clinical imaging. *Photoacoustics* **2019**, *14*, 77–98.
- [143] Wang, L. V.; Hu, S. Photoacoustic Tomography: In Vivo Imaging from Organelles to Organs. *Science* **2012**, *335*, 1458–1462.
- [144] Barolet, D.; Boucher, A. Radiant near infrared light emitting Diode exposure as skin preparation to enhance photodynamic therapy inflammatory type acne treatment outcome. *Lasers in Surgery and Medicine* **2010**, *42*, 171–178.
- [145] Knox, H. J.; Chan, J. Acoustogenic Probes: A New Frontier in Photoacoustic Imaging. *Accounts of Chemical Research* **2018**, *51*, 2897–2905.
- [146] Wang, H.; Yang, Y.; Huang, F.; He, Z.; Li, P.; Zhang, W.; Zhang, W.; Tang, B. In Situ Fluorescent and Photoacoustic Imaging of Golgi pH to Elucidate the Function of Transmembrane Protein 165. *Analytical Chemistry* **2020**, *92*, 3103–3110.
- [147] Zhang, Z.; Nair, S.; McMurry, T. Gadolinium Meets Medicinal Chemistry: MRI Contrast Agent Development. *Current Medicinal Chemistry* **2005**, *12*, 751–778.
- [148] Pan, D.; Schmieder, A. H.; Wickline, S. A.; Lanza, G. M. Manganese-based MRI contrast agents: past, present, and future. *Tetrahedron* **2011**, *67*, 8431–8444.

- [149] Chen, H.-Y.; Wilson, C. B.; Tycko, R. Enhanced spatial resolution in magnetic resonance imaging by dynamic nuclear polarization at 5 K. *Proceedings of the National Academy of Sciences* **2022**, *119*.
- [150] Mizukami, S.; Takikawa, R.; Sugihara, F.; Shirakawa, M.; Kikuchi, K. Dual-Function Probe to Detect Protease Activity for Fluorescence Measurement and ^{19}F MRI. *Angewandte Chemie International Edition* **2009**, *48*, 3641–3643.
- [151] Chen, K.; Li, Z.-B.; Wang, H.; Cai, W.; Chen, X. Dual-modality optical and positron emission tomography imaging of vascular endothelial growth factor receptor on tumor vasculature using quantum dots. *European Journal of Nuclear Medicine and Molecular Imaging* **2008**, *35*, 2235–2244.
- [152] Rijpkema, M.; Oyen, W. J.; Bos, D.; Franssen, G. M.; Goldenberg, D. M.; Boerman, O. C. SPECT- and Fluorescence Image-Guided Surgery Using a Dual-Labeled Carcinoembryonic Antigen-Targeting Antibody. *Journal of Nuclear Medicine* **2014**, *55*, 1519–1524.
- [153] Nyquist, R. *Interpreting Infrared, Raman, and Nuclear Magnetic Resonance Spectra*; Elsevier, 2001; pp 1–24.
- [154] Hackett, M. J.; DeSouza, M.; Caine, S.; Bewer, B.; Nichol, H.; Paterson, P. G.; Colbourne, F. A New Method To Image Heme-Fe, Total Fe, and Aggregated Protein Levels after Intracerebral Hemorrhage. *ACS Chemical Neuroscience* **2015**, *6*, 761–770.
- [155] Matthäus, C.; Bird, B.; Miljković, M.; Chernenko, T.; Romeo, M.; Diem, M. Chapter 10: Infrared and Raman microscopy in cell biology. *Methods in cell biology* **2008**, *89*, 275–308.
- [156] van Wilderen, L. J. G. W.; Brunst, H.; Gustmann, H.; Wachtveitl, J.; Broos, J.; Bredenbeck, J. Cyano-tryptophans as dual infrared and fluorescence spectroscopic labels to assess structural dynamics in proteins. *Physical Chemistry Chemical Physics* **2018**, *20*, 19906–19915.
- [157] Kopittke, P. M.; Punshon, T.; Paterson, D. J.; Tappero, R. V.; Wang, P.; Blamey, F. P. C.; van der Ent, A.; Lombi, E. Synchrotron-Based X-Ray Fluorescence Microscopy

- as a Technique for Imaging of Elements in Plants. *Plant Physiology* **2018**, *178*, 507–523.
- [158] Newville, M. Fundamentals of XAFS. *Reviews in Mineralogy and Geochemistry* **2014**, *78*, 33–74.
- [159] Lin, J.; Trinh, N.; New, E. *Molecular Fluorescent Sensors for Cellular Studies*; Wiley, 2022; pp 241–284.
- [160] Bronzel, J. L.; Milagre, C. D.; Milagre, H. M. Analysis of low molecular weight compounds using MALDI- and LDI-TOF-MS: Direct detection of active pharmaceutical ingredients in different formulations. *Journal of Mass Spectrometry* **2017**, *52*, 752–758.
- [161] Stoeckli, M.; Chaurand, P.; Hallahan, D. E.; Caprioli, R. M. Imaging mass spectrometry: A new technology for the analysis of protein expression in mammalian tissues. *Nature Medicine* **2001**, *7*, 493–496.
- [162] Prideaux, B.; Stoeckli, M. Mass spectrometry imaging for drug distribution studies. *Journal of Proteomics* **2012**, *75*, 4999–5013.
- [163] Spengler, B. Mass Spectrometry Imaging of Biomolecular Information. *Analytical Chemistry* **2015**, *87*, 64–82.
- [164] Eberlin, L. S.; Liu, X.; Ferreira, C. R.; Santagata, S.; Agar, N. Y.; Cooks, R. G. Desorption Electrospray Ionization then MALDI Mass Spectrometry Imaging of Lipid and Protein Distributions in Single Tissue Sections. *Analytical Chemistry* **2011**, *83*, 8366–8371.
- [165] Tipton, J. D.; Tran, J. C.; Catherman, A. D.; Ahlf, D. R.; Durbin, K. R.; Keller, N. L. Analysis of Intact Protein Isoforms by Mass Spectrometry. *Journal of Biological Chemistry* **2011**, *286*, 25451–25458.
- [166] Lahiri, S.; Sun, N.; Buck, A.; Imhof, A.; Walch, A. MALDI imaging mass spectrometry as a novel tool for detecting histone modifications in clinical tissue samples. *Expert Review of Proteomics* **2016**, *13*, 275–284.
- [167] Brunetti, A. E.; Marani, M. M.; Soldi, R. A.; Mendonça, J. N.; Faivovich, J.; Cabrera, G. M.; Lopes, N. P. Cleavage of Peptides from Amphibian Skin Revealed by

- Combining Analysis of Gland Secretion and in Situ MALDI Imaging Mass Spectrometry. *ACS Omega* **2018**, *3*, 5426–5434.
- [168] Patterson, N. H.; Thomas, A.; Chaurand, P. Monitoring time-dependent degradation of phospholipids in sectioned tissues by MALDI imaging mass spectrometry. *Journal of Mass Spectrometry* **2014**, *49*, 622–627.
- [169] Amstalden van Hove, E. R.; Smith, D. F.; Heeren, R. M. A concise review of mass spectrometry imaging. *Journal of Chromatography A* **2010**, *1217*, 3946–3954.
- [170] Olejnik, J.; Krzymańska-Olejnik, E.; Rothschild, K. J. [8] *Photocleavable affinity tags for isolation and detection of biomolecules*; 1998; pp 135–154.
- [171] Olejnik, J. Photocleavable peptide-DNA conjugates: synthesis and applications to DNA analysis using MALDI-MS. *Nucleic Acids Research* **1999**, *27*, 4626–4631.
- [172] Lemaire, R.; Stauber, J.; Wisztorski, M.; Van Camp, C.; Desmons, A.; Deschamps, M.; Proess, G.; Rudlof, I.; Woods, A. S.; Day, R.; Salzter, M.; Fournier, I. Tag-Mass: Specific Molecular Imaging of Transcriptome and Proteome by Mass Spectrometry Based on Photocleavable Tag. *Journal of Proteome Research* **2007**, *6*, 2057–2067.
- [173] Ma, W.; Xu, S.; Nie, H.; Hu, B.; Bai, Y.; Liu, H. Bifunctional cleavable probes for in situ multiplexed glycan detection and imaging using mass spectrometry. *Chemical Science* **2019**, *10*, 2320–2325.
- [174] Dai, C.; Cazares, L. H.; Wang, L.; Chu, Y.; Wang, S. L.; Troyer, D. A.; Semmes, O. J.; Drake, R. R.; Wang, B. Using boronolactin in MALDI-MS imaging for the histological analysis of cancer tissue expressing the sialyl Lewis X antigen. *Chemical Communications* **2011**, *47*, 10338.
- [175] Lorey, M.; Adler, B.; Yan, H.; Soliymani, R.; Ekström, S.; Yli-Kauhaluoma, J.; Laurell, T.; Baumann, M. Mass-Tag Enhanced Immuno-Laser Desorption/Ionization Mass Spectrometry for Sensitive Detection of Intact Protein Antigens. *Analytical Chemistry* **2015**, *87*, 5255–5262.
- [176] Yagnik, G.; Liu, Z.; Rothschild, K. J.; Lim, M. J. Highly Multiplexed Immunohis-

- tochemical MALDI-MS Imaging of Biomarkers in Tissues. *Journal of the American Society for Mass Spectrometry* **2021**, *32*, 977–988.
- [177] Yang, J.; Chaurand, P.; Norris, J. L.; Porter, N. A.; Caprioli, R. M. Activity-Based Probes Linked with Laser-Cleavable Mass Tags for Signal Amplification in Imaging Mass Spectrometry: Analysis of Serine Hydrolase Enzymes in Mammalian Tissue. *Analytical Chemistry* **2012**, *84*, 3689–3695.
- [178] Delaroche, D.; Aussedat, B.; Aubry, S.; Chassaing, G.; Burlina, F.; Clodic, G.; Bolbach, G.; Lavielle, S.; Sagan, S. Tracking a New Cell-Penetrating (W/R) Nonapeptide, through an Enzyme-Stable Mass Spectrometry Reporter Tag. *Analytical Chemistry* **2007**, *79*, 1932–1938.
- [179] Wang, Y.; Zhang, K.; Huang, X.; Qiao, L.; Liu, B. Mass Spectrometry Imaging of Mass Tag Immunoassay Enables the Quantitative Profiling of Biomarkers from Dozens of Exosomes. *Analytical Chemistry* **2021**, *93*, 709–714.
- [180] Han, J.; Sun, J.; Song, S.; Beljaars, L.; Groothuis, G. M. M.; Permentier, H.; Bischoff, R.; Halmos, G. B.; Verhoeven, C. J.; Amstalden van Hove, E. R.; Horvatovich, P.; Casini, A. Targeted imaging of integrins in cancer tissues using photocleavable Ru(II) polypyridine complexes as mass-tags. *Chemical Communications* **2020**, *56*, 5941–5944.
- [181] Taylor, A. J.; Dexter, A.; Bunch, J. Exploring Ion Suppression in Mass Spectrometry Imaging of a Heterogeneous Tissue. *Analytical Chemistry* **2018**, *90*, 5637–5645.
- [182] Thomas, A.; Chaurand, P. Advances in tissue section preparation for MALDI imaging MS. *Bioanalysis* **2014**, *6*, 967–982.
- [183] Gessel, M. M.; Norris, J. L.; Caprioli, R. M. MALDI imaging mass spectrometry: Spatial molecular analysis to enable a new age of discovery. *Journal of Proteomics* **2014**, *107*, 71–82.
- [184] Kompauer, M.; Heiles, S.; Spengler, B. Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4- μ m lateral resolution. *Nature Methods* **2017**, *14*, 90–96.

- [185] Rabe, J.-H.; A. Sammour, D.; Schulz, S.; Munteanu, B.; Ott, M.; Ochs, K.; Hohenberger, P.; Marx, A.; Platten, M.; Opitz, C. A.; Ory, D. S.; Hopf, C. Fourier Transform Infrared Microscopy Enables Guidance of Automated Mass Spectrometry Imaging to Predefined Tissue Morphologies. *Scientific Reports* **2018**, *8*, 313.
- [186] Tian, X.; Xie, B.; Zou, Z.; Jiao, Y.; Lin, L.-E.; Chen, C.-L.; Hsu, C.-C.; Peng, J.; Yang, Z. Multimodal Imaging of Amyloid Plaques: Fusion of the Single-Probe Mass Spectrometry Image and Fluorescence Microscopy Image. *Analytical Chemistry* **2019**, *91*, 12882–12889.
- [187] Patterson, N. H.; Tuck, M.; Lewis, A.; Kaushansky, A.; Norris, J. L.; Van de Plas, R.; Caprioli, R. M. Next Generation Histology-Directed Imaging Mass Spectrometry Driven by Autofluorescence Microscopy. *Analytical Chemistry* **2018**, *90*, 12404–12413.
- [188] Rujchanarong, D.; Lefler, J.; Saunders, J. E.; Pippin, S.; Spruill, L.; Bethard, J. R.; Ball, L. E.; Mehta, A. S.; Drake, R. R.; Ostrowski, M. C.; Angel, P. M. Defining the Tumor Microenvironment by Integration of Immunohistochemistry and Extracellular Matrix Targeted Imaging Mass Spectrometry. *Cancers* **2021**, *13*, 4419.
- [189] Prade, V. M.; Kunzke, T.; Feuchtinger, A.; Rohm, M.; Lubner, B.; Lordick, F.; Buck, A.; Walch, A. De novo discovery of metabolic heterogeneity with immunophenotype-guided imaging mass spectrometry. *Molecular Metabolism* **2020**, *36*, 100953.
- [190] Dufresne, M.; Guneyasu, D.; Patterson, N. H.; Marcinkiewicz, M. M.; Regina, A.; Demeule, M.; Chaurand, P. Multimodal detection of GM2 and GM3 lipid species in the brain of mucopolysaccharidosis type II mouse by serial imaging mass spectrometry and immunohistochemistry. *Analytical and Bioanalytical Chemistry* **2017**, *409*, 1425–1433.
- [191] Van de Plas, R.; Yang, J.; Spraggins, J.; Caprioli, R. M. Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping. *Nature Methods* **2015**, *12*, 366–372.
- [192] Neumann, E. K.; Comi, T. J.; Spegazzini, N.; Mitchell, J. W.; Rubakhin, S. S.; Gillette, M. U.; Bhargava, R.; Sweedler, J. V. Multimodal Chemical Analysis of the

- Brain by High Mass Resolution Mass Spectrometry and Infrared Spectroscopic Imaging. *Analytical Chemistry* **2018**, *90*, 11572–11580.
- [193] Chughtai, S.; Chughtai, K.; Cillero-Pastor, B.; Kiss, A.; Agrawal, P.; MacAleese, L.; Heeren, R. M. A multimodal mass spectrometry imaging approach for the study of musculoskeletal tissues. *International Journal of Mass Spectrometry* **2012**, *325-327*, 150–160.
- [194] Ellis, S. R.; Cappell, J.; Potočník, N. O.; Balluff, B.; Hamaide, J.; Van der Linden, A.; Heeren, R. M. A. More from less: high-throughput dual polarity lipid imaging of biological tissues. *The Analyst* **2016**, *141*, 3832–3841.
- [195] Kaya, I.; Zetterberg, H.; Blennow, K.; Hanrieder, J. Shedding Light on the Molecular Pathology of Amyloid Plaques in Transgenic Alzheimer’s Disease Mice Using Multimodal MALDI Imaging Mass Spectrometry. *ACS Chemical Neuroscience* **2018**, *9*, 1802–1817.
- [196] Kaya, I.; Sämfors, S.; Levin, M.; Borén, J.; Fletcher, J. S. Multimodal MALDI Imaging Mass Spectrometry Reveals Spatially Correlated Lipid and Protein Changes in Mouse Heart with Acute Myocardial Infarction. *Journal of the American Society for Mass Spectrometry* **2020**, *31*, 2133–2142.
- [197] Fincher, J. A.; Korte, A. R.; Yadavilli, S.; Morris, N. J.; Vertes, A. Multimodal imaging of biological tissues using combined MALDI and NAPA-LDI mass spectrometry for enhanced molecular coverage. *The Analyst* **2020**, *145*, 6910–6918.
- [198] Ahlf, D. R.; Masyuko, R. N.; Hummon, A. B.; Bohn, P. W. Correlated mass spectrometry imaging and confocal Raman microscopy for studies of three-dimensional cell culture sections. *The Analyst* **2014**, *139*, 4578.
- [199] Svirskova, A.; Turyanskaya, A.; Perneczky, L.; Streli, C.; Marchetti-Deschmann, M. Multimodal imaging of undecalcified tissue sections by MALDI MS and μ XRF. *The Analyst* **2018**, *143*, 2587–2595.
- [200] Chughtai, K.; Jiang, L.; Post, H.; Winnard, P. T.; Greenwood, T. R.; Raman, V.; Bhujwalla, Z. M.; Heeren, R. M. A.; Glunde, K. Mass Spectrometric Imaging of Red

- Fluorescent Protein in Breast Tumor Xenografts. *Journal of the American Society for Mass Spectrometry* **2013**, *24*, 711–717.
- [201] Huber, K.; Kunzke, T.; Buck, A.; Langer, R.; Lubner, B.; Feuchtinger, A.; Walch, A. Multimodal analysis of formalin-fixed and paraffin-embedded tissue by MALDI imaging and fluorescence in situ hybridization for combined genetic and metabolic analysis. *Laboratory Investigation* **2019**, *99*, 1535–1546.
- [202] Choi, S. K. *Photonanotechnology for Therapeutics and Imaging*; Elsevier, 2020; pp 243–275.
- [203] Il'ichev, Y. V.; Schwörer, M. A.; Wirz, J. Photochemical Reaction Mechanisms of 2-Nitrobenzyl Compounds: Methyl Ethers and Caged ATP. *Journal of the American Chemical Society* **2004**, *126*, 4581–4595.
- [204] Gaplovsky, M.; Il'ichev, Y. V.; Kamdzhilov, Y.; Kombarova, S. V.; Mac, M.; Schwörer, M. A.; Wirz, J. Photochemical reaction mechanisms of 2-nitrobenzyl compounds: 2-Nitrobenzyl alcohols form 2-nitroso hydrates by dual proton transfer. *Photochemical & Photobiological Sciences* **2005**, *4*, 33–42.
- [205] Jackowska, A.; Gryko, D. Vitamin B 12 Derivatives Suitably Tailored for the Synthesis of Photolabile Conjugates. *Organic Letters* **2021**, *23*, 4940–4944.
- [206] Kaneko, S.; Nakayama, H.; Yoshino, Y.; Fushimi, D.; Yamaguchi, K.; Horiike, Y.; Nakanishi, J. Photocontrol of cell adhesion on amino-bearing surfaces by reversible conjugation of poly(ethylene glycol) via a photocleavable linker. *Physical Chemistry Chemical Physics* **2011**, *13*, 4051.
- [207] Qi, J.; Hu, X.; Dong, X.; Lu, Y.; Lu, H.; Zhao, W.; Wu, W. Towards more accurate bioimaging of drug nanocarriers: turning aggregation-caused quenching into a useful tool. *Advanced Drug Delivery Reviews* **2019**, *143*, 206–225.
- [208] Zhang, Y.; Xiang, J.; Tang, Y.; Xu, G.; Yan, W. Aggregation behaviour of two thiacyanine dyes in aqueous solution. *Dyes and Pigments* **2008**, *76*, 88–93.
- [209] Sawicka, M. J. The self-aggregation of some 7H-indolo[1,2-a]quinolinium dyes in aqueous solution. *Journal of Molecular Structure* **2015**, *1098*, 26–33.

- [210] Hestand, N. J.; Spano, F. C. Expanded Theory of H- and J-Molecular Aggregates: The Effects of Vibronic Coupling and Intermolecular Charge Transfer. *Chemical reviews* **2018**, *118*, 7069–7163.
- [211] Cao, W.; Sletten, E. M. Fluorescent Cyanine Dye J-Aggregates in the Fluorous Phase. *Journal of the American Chemical Society* **2018**, *140*, 2727–2730.
- [212] Wu, F.; Lu, Y.; Mu, X.; Chen, Z.; Liu, S.; Zhou, X.; Liu, S.; Li, Z. Intriguing H-Aggregates of Heptamethine Cyanine for Imaging-Guided Photothermal Cancer Therapy. *ACS Applied Materials & Interfaces* **2020**, *12*, 32388–32396.
- [213] Nassier, L. F.; Shinen, M. H. Study of the optical properties of poly (methyl methacrylate) (PMMA) by using spin coating method. *Materials Today: Proceedings* **2022**, *60*, 1660–1664.
- [214] Momotake, A.; Lindegger, N.; Niggli, E.; Barsotti, R. J.; Ellis-Davies, G. C. R. The nitrodibenzofuran chromophore: a new caging group for ultra-efficient photolysis in living cells. *Nature Methods* **2006**, *3*, 35–40.
- [215] Neveu, P.; Aujard, I.; Benbrahim, C.; Le Saux, T.; Allemand, J.-F.; Vriz, S.; Bensimon, D.; Jullien, L. A Caged Retinoic Acid for One- and Two-Photon Excitation in Zebrafish Embryos. *Angewandte Chemie International Edition* **2008**, *47*, 3744–3746.
- [216] Wong, P. T.; Roberts, E. W.; Tang, S.; Mukherjee, J.; Cannon, J.; Nip, A. J.; Corbin, K.; Krummel, M. F.; Choi, S. K. Control of an Unusual Photo-Claisen Rearrangement in Coumarin Caged Tamoxifen through an Extended Spacer. *ACS Chemical Biology* **2017**, *12*, 1001–1010.
- [217] Iannuzzi, C.; Borriello, M.; Portaccio, M.; Irace, G.; Sirangelo, I. Insights into Insulin Fibril Assembly at Physiological and Acidic pH and Related Amyloid Intrinsic Fluorescence. *International Journal of Molecular Sciences* **2017**, *18*, 2551.
- [218] Graziotto, M. E.; Adair, L. D.; Kaur, A.; Vérité, P.; Ball, S. R.; Sunde, M.; Jacquemin, D.; New, E. J. Versatile naphthalimide tetrazines for fluorogenic bioorthogonal labelling. *RSC Chemical Biology* **2021**, *2*, 1491–1498.

- [219] Auten, R. L.; Davis, J. M. Oxygen Toxicity and Reactive Oxygen Species: The Devil Is in the Details. *Pediatric Research* **2009**, *66*, 121–127.
- [220] Bhatt, S.; Nagappa, A. N.; Patil, C. R. Role of oxidative stress in depression. *Drug Discovery Today* **2020**, *25*, 1270–1276.
- [221] Popa-Wagner, A.; Mitran, S.; Sivanesan, S.; Chang, E.; Buga, A.-M. ROS and Brain Diseases: The Good, the Bad, and the Ugly. *Oxidative Medicine and Cellular Longevity* **2013**, *2013*, 1–14.
- [222] Alfadda, A. A.; Sallam, R. M. Reactive Oxygen Species in Health and Disease. *Journal of Biomedicine and Biotechnology* **2012**, *2012*, 1–14.
- [223] Yang, S.; Lian, G. ROS and diseases: role in metabolism and energy supply. *Molecular and Cellular Biochemistry* **2020**, *467*, 1–12.
- [224] D’Autréaux, B.; Toledano, M. B. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nature Reviews Molecular Cell Biology* **2007**, *8*, 813–824.
- [225] Ray, P. D.; Huang, B.-W.; Tsuji, Y. Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular Signalling* **2012**, *24*, 981–990.
- [226] Sun, S. G.; Ding, H.; Yuan, G.; Zhou, L. An efficient TP-FRET-based lysosome-targetable fluorescent probe for imaging peroxynitrite with two well-resolved emission channels in living cells, tissues and zebrafish. *Analytica Chimica Acta* **2020**, *1100*, 200–207.
- [227] Moldovan, L.; Moldovan, N. I. Oxygen free radicals and redox biology of organelles. *Histochemistry and Cell Biology* **2004**, *122*, 395–412.
- [228] Zhang, Y.; Chen, S.-Y.; Hsu, T.; Santella, R. M. Immunohistochemical detection of malondialdehyde-DNA adducts in human oral mucosa cells. *Carcinogenesis* **2002**, *23*, 207–211.
- [229] Lin, J.; Yang, K.; New, E. J. Strategies for organelle targeting of fluorescent probes. *Organic & Biomolecular Chemistry* **2021**, *19*, 9339–9357.

- [230] Chen, H.; Yu, Z.; Ren, S.; Qiu, Y. Fluorescent Probes Design Strategies for Imaging Mitochondria and Lysosomes. *Frontiers in pharmacology* **2022**, *13*, 915609.
- [231] Wang, B.; Yuan, F.; Wang, S.; Duan, R.; Ren, W. X.; Hou, J.-T. Detection of atherosclerosis-associated HOCl using a mitochondria-targeted fluorescent probe. *Sensors and Actuators B: Chemical* **2021**, *348*, 130695.
- [232] Teng, H.; Tian, J.; Sun, D.; Xiu, M.; Zhang, Y.; Qiang, X.; Tang, H.; Guo, Y. A mitochondria-specific fluorescent probe based on triazolopyridine formation for visualizing endogenous hypochlorous acid in living cells and zebrafish. *Sensors and Actuators B: Chemical* **2020**, *319*, 128288.
- [233] Liu, J. F.; Zhai, Z. Y.; Niu, H. W.; Zhang, Y. R.; Song, X. Z.; Zhang, P. K.; Ye, Y. Endoplasmic reticulum-targetable fluorescent probe for visualizing HClO in EC1 cells. *Tetrahedron Letters* **2020**, *61*, 152301.
- [234] Feng, Y.; Li, S.; Li, D.; Wang, Q.; Ning, P.; Chen, M.; Tian, X.; Wang, X. Rational design of a diaminomaleonitrile-based mitochondria – targeted two-photon fluorescent probe for hypochlorite in vivo: Solvent-independent and high selectivity over Cu²⁺. *Sensors and Actuators B: Chemical* **2018**, *254*, 282–290.
- [235] Xu, H.; Zhong, J.; Zhuang, W.; Jiang, J.; Ma, B.; He, H.; Li, G.; Liao, Y.; Wang, Y. A bifunctional mitochondrial targeting AIE-active fluorescent probe with high sensitivity to hydrogen peroxide and viscosity for fatty liver diagnosis. *New Journal of Chemistry* **2021**, *45*, 12138–12144.
- [236] Mu, S.; Jiang, L.; Gao, H.; Zhang, J. L.; Sun, H. P.; Shi, X. Z.; Liu, X. Y.; Zhang, H. X. A novel fluorescent probe with large Stokes shift for accurate detection of HOCl in mitochondria and its imaging application. *Analytica Chimica Acta* **2022**, *1191*.
- [237] Huang, Y.; Zhang, Y.; Huo, F.; Liu, Y.; Yin, C. Mitochondrial-targeted near-infrared “dual mode” fluorescent dyes with large Stokes shift for detection of hypochlorous acid and its bioimaging in cell and mice. *Dyes and Pigments* **2020**, *179*, 108387.
- [238] Shu, W.; Wu, Y.; Shen, T.; Cui, J.; Kang, H.; Jing, J.; Zhang, X. A mitochondria-targeted far red fluorescent probe for ratiometric imaging of endogenous peroxynitrite. *Dyes and Pigments* **2019**, *170*, 107609.

- [239] Song, X.; Wang, Y.; Ru, J.; Yang, Y.; Feng, Y.; Cao, C.; Wang, K.; Zhang, G.; Liu, W. A mitochondrial-targeted red fluorescent probe for detecting endogenous H₂S in cells with high selectivity and development of a visual paper-based sensing platform. *Sensors and Actuators B: Chemical* **2020**, *312*, 127982.
- [240] Zhou, L.; Lu, D.; Wang, Q.; Liu, S.; Lin, Q.; Sun, H. Molecular engineering of a mitochondrial-targeting two-photon in and near-infrared out fluorescent probe for gaseous signal molecules H₂S in deep tissue bioimaging. *Biosensors and Bioelectronics* **2017**, *91*, 699–705.
- [241] Deng, T.; Wang, X.; Wu, S.; Hu, S.; Liu, W.; Chen, T.; Yu, Z.; Xu, Q.; Liu, F. A new FRET probe for ratiometric fluorescence detecting mitochondria-localized drug activation and imaging endogenous hydroxyl radicals in zebrafish. *Chemical Communications* **2020**, *56*, 4432–4435.
- [242] Qin, F.; Zhang, Y.; Zhu, J.; Li, Y.; Cao, W.; Ye, Y. A mitochondrial-targeted fluorescent probe to sense pH and HOCl in living cells. *Sensors and Actuators B: Chemical* **2019**, *291*, 207–215.
- [243] Gu, T.; Mo, S.; Mu, Y.; Huang, X.; Hu, L. Detection of endogenous hydrogen peroxide in living cells with para-nitrophenyl oxoacetyl rhodamine as turn-on mitochondria-targeted fluorescent probe. *Sensors and Actuators B: Chemical* **2020**, *309*, 127731.
- [244] Jiang, C.; Li, Y.; Yan, L.; Ye, A.; He, Q.; Yao, C. A ratiometric fluorescence mitochondrial-targeted probe for imaging HOCl in vitro and in vivo. *Dyes and Pigments* **2022**, *198*, 109975.
- [245] Bao, X.; Ai, K.; Cao, X.; Chen, D.; Zhou, B.; Huo, C. A dual-site and dual-turn-on fluorescence probe for imaging mitochondrial HClO and SO₂. *Dyes and Pigments* **2022**, *197*, 109928.
- [246] Niu, H.; Tang, J.; Zhu, X.; Li, Z.; Zhang, Y.; Ye, Y.; Zhao, Y. A three-channel fluorescent probe to image mitochondrial stress. *Chemical Communications* **2020**, *56*, 7710–7713.

- [247] Wen, Y.; Liu, K.; Yang, H.; Liu, Y.; Chen, L.; Liu, Z.; Huang, C.; Yi, T. Mitochondria-Directed Fluorescent Probe for the Detection of Hydrogen Peroxide near Mitochondrial DNA. *Analytical Chemistry* **2015**, *87*, 10579–10584.
- [248] Tang, F.; Gao, C.; Liu, J. Y.; Lu, Z. L.; He, L.; Ding, A. X. Lysosome-targeting BODIPY-derived Hantzsch ester for nitric oxide detection and imaging in live cells. *Sensors and Actuators B: Chemical* **2021**, *339*, 129880.
- [249] Mao, G. J.; Wang, Y. Y.; Dong, W. P.; Meng, H. M.; Wang, Q. Q.; Luo, X. F.; Li, Y.; Zhang, G. A lysosome-targetable two-photon excited near-infrared fluorescent probe for visualizing hypochlorous acid-involved arthritis and its treatment. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2021**, *249*, 119326.
- [250] Ma, S.; Ma, Y.; Liu, Q.; Lin, W. A two-photon fluorescent probe with lysosome targetability for imaging endogenous superoxide anion in living cells, zebrafish and pneumonia tissue. *Sensors and Actuators B: Chemical* **2021**, *332*, 129523.
- [251] Liu, J.; Niu, P.; Rong, Y.; Chen, W.; Liu, X.; Wei, L.; Song, X. A phenothiazine coumarin based ratiometric fluorescent probe for real-time detection of lysosomal hypochlorite in living cell and zebra fish. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2021**, *261*, 120024.
- [252] Gu, B.; Wu, C.; Zhang, C.; He, S.; Tang, S.; Li, H.; Shen, Y. A morpholino hydrazone-based lysosome-targeting fluorescent probe with fast response and high sensitivity for imaging peroxynitrite in living cells. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2021**, *262*, 120100.
- [253] Sun, S.-G.; Ding, H.; Yuan, G.; Zhou, L. An efficient TP-FRET-based lysosome-targetable fluorescent probe for imaging peroxynitrite with two well-resolved emission channels in living cells, tissues and zebrafish. *Analytica Chimica Acta* **2020**, *1100*, 200–207.
- [254] Wang, K.; Liu, Y.; Liu, C.; Zhu, H.; Li, X.; Zhang, F.; Gao, N.; Pang, X.; Sheng, W.; Zhu, B. A simple pyridine-based highly specific fluorescent probe for tracing hypochlorous acid in lysosomes of living cells. *New Journal of Chemistry* **2021**, *45*, 14548–14553.

- [255] Xu, C.; Qian, Y. A selenamorpholine-based redox-responsive fluorescent probe for targeting lysosome and visualizing exogenous/endogenous hydrogen peroxide in living cells and zebrafish. *Journal of Materials Chemistry B* **2019**, *7*, 2714–2721.
- [256] Zhao, L.; Huang, Z.; Ma, D.; Yan, Y.; Zhang, X.; Xiao, Y. A nucleus targetable fluorescent probe for ratiometric imaging of endogenous NO in living cells and zebrafishes. *The Analyst* **2021**, *146*, 4130–4134.
- [257] Ikeda, M.; Nakagawa, H.; Suzuki, T.; Miyata, N. Novel bisbenzimidazole-nitroxides for nuclear redox imaging in living cells. *Bioorganic & Medicinal Chemistry Letters* **2012**, *22*, 1949–1952.
- [258] Srikun, D.; Albers, A. E.; Nam, C. I.; Iavarone, A. T.; Chang, C. J. Organelle-Targetable Fluorescent Probes for Imaging Hydrogen Peroxide in Living Cells via SNAP-Tag Protein Labeling. *Journal of the American Chemical Society* **2010**, *132*, 4455–4465.
- [259] Wang, C.; Song, X.; Han, Z.; Li, X.; Xu, Y.; Xiao, Y. Monitoring Nitric Oxide in Subcellular Compartments by Hybrid Probe Based on Rhodamine Spirolactam and SNAP-tag. *ACS Chemical Biology* **2016**, *11*, 2033–2040.
- [260] Wen, Y.; Liu, K.; Yang, H.; Li, Y.; Lan, H.; Liu, Y.; Zhang, X.; Yi, T. A Highly Sensitive Ratiometric Fluorescent Probe for the Detection of Cytoplasmic and Nuclear Hydrogen Peroxide. *Analytical Chemistry* **2014**, *86*, 9970–9976.
- [261] Dickinson, B. C.; Tang, Y.; Chang, Z.; Chang, C. J. A Nuclear-Localized Fluorescent Hydrogen Peroxide Probe for Monitoring Sirtuin-Mediated Oxidative Stress Responses In Vivo. *Chemistry & Biology* **2011**, *18*, 943–948.
- [262] Ding, S.; Hong, Y. The fluorescence toolbox for visualizing autophagy. *Chemical Society Reviews* **2020**, *49*, 8354–8389.
- [263] Zhou, Y.; Li, P.; Fan, N.; Wang, X.; Liu, X.; Wu, L.; Zhang, W.; Zhang, W.; Ma, C.; Tang, B. In situ visualization of peroxisomal peroxynitrite in the livers of mice with acute liver injury induced by carbon tetrachloride using a new two-photon fluorescent probe. *Chemical Communications* **2019**, *55*, 6767–6770.

- [264] Li, J.; Peng, S.; Li, Z.; Zhao, F.; Han, X.; Liu, J.; Cao, W.; Ye, Y. Visualization of peroxynitrite in cyclophosphamide-induced oxidative stress by an activatable probe. *Talanta* **2022**, *238*, 123007.
- [265] Li, Z.; Yang, Y.; Yin, P.; Yang, Z. M.; Zhang, B.; Zhang, S. J.; Han, B. Y.; Lv, J. W.; Dong, F. H.; Ma, H. C. A New Lipid-Droplets–Targeted Fluorescence Probe with Dual–Reactive Sites for Specific Detection of ClO– in Living Cells. *Chemistryselect* **2022**, *7*.
- [266] Wu, W. L.; Ma, H. L.; Xi, L. L.; Huang, M. F.; Wang, K. M.; Miao, J. Y.; Zhao, B. X. A novel lipid droplets-targeting ratiometric fluorescence probe for hypochlorous acid in living cells. *Talanta* **2019**, *194*, 308–313.
- [267] Ye, M. T.; Hu, W.; He, M.; Li, C. C.; Zhai, S. Y.; Liu, Z. H.; Wang, Y. Y.; Zhang, H. J.; Li, C. Y. Deep imaging for visualizing nitric oxide in lipid droplets: discovering the relationship between nitric oxide and resistance to cancer chemotherapy drugs. *Chemical Communications* **2020**, *56*, 6233–6236.
- [268] Zhang, J.; An, Q.; Li, W.; Chai, L.; Hu, W.; Wang, Y.; Su, S.; He, Y.; Li, C.; Sun, D. Rationally designed lipid droplets-selective two-photon nitric oxide probe for high-fidelity neuroinflammation evaluation. *Sensors and Actuators B: Chemical* **2021**, *345*, 130329.
- [269] Zhang, X. F.; Wang, B. L.; Xiao, Y.; Wang, C.; He, L. Targetable, two-photon fluorescent probes for local nitric oxide capture in the plasma membranes of live cells and brain tissues. *Analyst* **2018**, *143*, 4180–4188.
- [270] Chvanov, M.; Huang, W.; Jin, T.; Wen, L.; Armstrong, J.; Elliot, V.; Alston, B.; Burdyga, A.; Criddle, D. N.; Sutton, R.; Tepikin, A. V. Novel lipophilic probe for detecting near-membrane reactive oxygen species responses and its application for studies of pancreatic acinar cells: Effects of pyocyanin and L-ornithine. *Antioxidants and Redox Signaling* **2015**, *22*, 451–464.
- [271] Greene, L. E.; Lincoln, R.; Cosa, G. Spatio-temporal monitoring of lipid peroxyl radicals in live cell studies combining fluorogenic antioxidants and fluorescence microscopy methods. *Free Radical Biology and Medicine* **2018**, *128*, 124–136.

- [272] Abo, M.; Minakami, R.; Miyano, K.; Kamiya, M.; Nagano, T.; Urano, Y.; Sumimoto, H. Visualization of Phagosomal Hydrogen Peroxide Production by a Novel Fluorescent Probe That Is Localized via SNAP-tag Labeling. *Analytical Chemistry* **2014**, *86*, 5983–5990.
- [273] Underhill, D. M. Macrophage recognition of zymosan particles. *Journal of endotoxin research* **2003**, *9*, 176–180.
- [274] Albrett, A. M.; Ashby, L. V.; Dickerhof, N.; Kettle, A. J.; Winterbourn, C. C. Heterogeneity of hypochlorous acid production in individual neutrophil phagosomes revealed by a rhodamine-based probe. *Journal of Biological Chemistry* **2018**, *293*, 15715–15724.
- [275] Kenmoku, S.; Urano, Y.; Kojima, H.; Nagano, T. Development of a highly specific rhodamine-based fluorescence probe for hypochlorous acid and its application to real-time imaging of phagocytosis. *Journal of the American Chemical Society* **2007**, *129*, 7313–7318.
- [276] Bassøe, C. F.; Li, N.; Ragheb, K.; Lawler, G.; Sturgis, J.; Robinson, J. P. Investigations of phagosomes, mitochondria, and acidic granules in human neutrophils using fluorescent probes. *Cytometry Part B: Clinical Cytometry* **2003**, *51B*, 21–29.
- [277] Lindberg, E.; Winssinger, N. High Spatial Resolution Imaging of Endogenous Hydrogen Peroxide in Living Cells by Solid-State Fluorescence. *ChemBioChem* **2016**, *17*, 1612–1615.
- [278] Thekkan, S.; Jani, M. S.; Cui, C.; Dan, K.; Zhou, G.; Becker, L.; Krishnan, Y. A DNA-based fluorescent reporter maps HOCl production in the maturing phagosome. *Nature Chemical Biology* **2019**, *15*, 1165–1172.
- [279] Li, J.; Tang, J.; Yang, X.; Xie, P.; Liu, J.; Zhang, D.; Ye, Y. A novel aggregation-induced emission fluorescent probe to visualize peroxynitrite levels during Golgi stress. *Sensors and Actuators B: Chemical* **2022**, *358*, 131513.
- [280] Feng, S.; Zheng, Z.; Gong, S.; Feng, G. A unique probe enables labeling cell membrane and Golgi apparatus and tracking peroxynitrite in Golgi oxidative stress and drug-induced liver injury. *Sensors and Actuators B: Chemical* **2022**, *361*, 131751.

- [281] He, Z.; Liu, D.; Liu, Y.; Li, X.; Shi, W.; Ma, H. Golgi-Targeted Fluorescent Probe for Imaging NO in Alzheimer’s Disease. *Analytical Chemistry* **2022**, *94*, 10256–10262.
- [282] Wang, H.; He, Z.; Yang, Y.; Zhang, J.; Zhang, W.; Zhang, W.; Li, P.; Tang, B. Ratiometric fluorescence imaging of Golgi H₂O₂ reveals a correlation between Golgi oxidative stress and hypertension. *Chemical Science* **2019**, *10*, 10876–10880.
- [283] Yan, Z.; Tang, Z.; Wang, X.; Zheng, Z.; Tian, Z.; Geng, X.; Li, Y.; Jiang, H. A novel Golgi-targetable fluorescent probe for imaging peroxynitrite in Golgi stress and sepsis-induced acute lung injury. *Sensors and Actuators B: Chemical* **2022**, *369*, 132352.
- [284] Rong, X.; Liu, C.; Li, M.; Zhu, H.; Zhang, Y.; Su, M.; Wang, X.; Li, X.; Wang, K.; Yu, M.; Sheng, W.; Zhu, B. An Integrated Fluorescent Probe for Ratiometric Detection of Glutathione in the Golgi Apparatus and Activated Organelle-Targeted Therapy. *Analytical Chemistry* **2021**, *93*, 16105–16112.
- [285] Zhang, W.; Zhang, J.; Li, P.; Liu, J.; Su, D.; Tang, B. Two-photon fluorescence imaging reveals a Golgi apparatus superoxide anion-mediated hepatic ischaemia-reperfusion signalling pathway. *Chemical Science* **2019**, *10*, 879–883.
- [286] Sun, Y.; Wang, R.; Wang, J.; Wei, H.; Chen, Q.; Wang, Y.; Dong, B. Construction of a ratiometric two-photon ER-targeting fluorescent probe for the imaging of peroxynitrite in living systems. *Sensors and Actuators B: Chemical* **2022**, *370*, 132439.
- [287] Vassallo, N.; Herms, J. Cellular prion protein function in copper homeostasis and redox signalling at the synapse. *Journal of Neurochemistry* **2003**, *86*, 538–544.
- [288] Yeow, J.; Kaur, A.; Anscomb, M. D.; New, E. J. A novel flavin derivative reveals the impact of glucose on oxidative stress in adipocytes. *Chem. Commun.* **2014**, *50*, 8181–8184.
- [289] Sinnecker, S.; Neese, F. QM/MM calculations with DFT for taking into account protein effects on the EPR and optical spectra of metalloproteins. Plastocyanin as a case study. *Journal of Computational Chemistry* **2006**, *27*, 1463–1475.
- [290] Li, L.; Li, C.; Zhang, Z.; Alexov, E. On the Dielectric “Constant” of Proteins: Smooth

- Dielectric Function for Macromolecular Modeling and Its Implementation in DelPhi. *Journal of Chemical Theory and Computation* **2013**, *9*, 2126–2136.
- [291] Bianco, G.; Forli, S.; Goodsell, D. S.; Olson, A. J. Covalent docking using autodock: Two-point attractor and flexible side chain methods. *Protein Science* **2016**, *25*, 295–301.
- [292] Clegg, R. M. Fluorescence resonance energy transfer. *Current Opinion in Biotechnology* **1995**, *6*, 103–110.
- [293] Andrews, D. L.; Bradshaw, D. S.; Dinshaw, R.; Scholes, G. D. *Photonics*; Wiley, 2015; pp 101–127.
- [294] King, C.; Barbiellini, B.; Moser, D.; Renugopalakrishnan, V. Exactly soluble model of resonant energy transfer between molecules. *Physical Review B* **2012**, *85*, 125106.
- [295] Kaur, A.; Haghighatbin, M. A.; Hogan, C. F.; New, E. J. A FRET-based ratiometric redox probe for detecting oxidative stress by confocal microscopy, FLIM and flow cytometry. *Chemical Communications* **2015**, *51*, 10510–10513.
- [296] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.
- [297] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *The Journal of Physical Chemistry* **1994**, *98*, 11623–11627.
- [298] Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *The Journal of Chemical Physics* **1972**, *56*, 2257–2261.
- [299] Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *The Journal of Physical Chemistry A* **1998**, *102*, 1995–2001.
- [300] Dale, R. E.; Eisinger, J.; Blumberg, W. E. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophysical journal* **1979**, *26*, 161–93.

- [301] Robinson, D. Comparison of the Transition Dipole Moments Calculated by TDDFT with High Level Wave Function Theory. *Journal of Chemical Theory and Computation* **2018**, *14*, 5303–5309.
- [302] Li, Y.; Tang, Q.; Xie, Y.; He, D.; Yang, K.; Zheng, L. Synthesis of mitochondria-targeted coumarin-3-carboxamide fluorescent derivatives: Inhibiting mitochondrial TrxR2 and cell proliferation on breast cancer cells. *Bioorganic & Medicinal Chemistry Letters* **2021**, *33*, 127750.
- [303] Jacquemin, D.; Perpète, E. A.; Scalmani, G.; Ciofini, I.; Peltier, C.; Adamo, C. Absorption and emission spectra of 1,8-naphthalimide fluorophores: A PCM-TD-DFT investigation. *Chemical Physics* **2010**, *372*, 61–66.
- [304] Jacquemin, D.; Perpète, E. A.; Scalmani, G.; Frisch, M. J.; Ciofini, I.; Adamo, C. Fluorescence of 1,8-naphthalimide: A PCM-TD-DFT investigation. *Chemical Physics Letters* **2007**, *448*, 3–6.
- [305] de la Peña O’Shea, V. A.; Pardo, A.; Poyato, J. M. L. DFT study of electronic spectra and excited-state properties of some 1,8-naphthalimide derivatives. *International Journal of Quantum Chemistry* **2003**, *91*, 446–450.
- [306] Dhar, S.; Singha Roy, S.; Rana, D. K.; Bhattacharya, S.; Bhattacharya, S.; Bhattacharya, S. C. Tunable Solvatochromic Response of Newly Synthesized Antioxidative Naphthalimide Derivatives: Intramolecular Charge Transfer Associated with Hydrogen Bonding Effect. *The Journal of Physical Chemistry A* **2011**, *115*, 2216–2224.
- [307] Neiss, C.; Saalfrank, P.; Parac, M.; Grimme, S. Quantum Chemical Calculation of Excited States of Flavin-Related Molecules. *The Journal of Physical Chemistry A* **2003**, *107*, 140–147.
- [308] List, N. H.; Pimenta, F. M.; Holmegaard, L.; Jensen, R. L.; Etzerodt, M.; Schwabe, T.; Kongsted, J.; Ogilby, P. R.; Christiansen, O. Effect of chromophore encapsulation on linear and nonlinear optical properties: the case of “miniSOG”, a protein-encased flavin. *Physical Chemistry Chemical Physics* **2014**, *16*, 9950.
- [309] Kabir, M. P.; Orozco-Gonzalez, Y.; Gozem, S. Electronic spectra of flavin in different

- redox and protonation states: a computational perspective on the effect of the electrostatic environment. *Physical Chemistry Chemical Physics* **2019**, *21*, 16526–16537.
- [310] Charaf-Eddin, A.; Planchat, A.; Mennucci, B.; Adamo, C.; Jacquemin, D. Choosing a Functional for Computing Absorption and Fluorescence Band Shapes with TD-DFT. *Journal of Chemical Theory and Computation* **2013**, *9*, 2749–2760.
- [311] Chien, S.-H.; Gill, P. M. W. SG-0: A small standard grid for DFT quadrature on large systems. *Journal of Computational Chemistry* **2006**, *27*, 730–739.
- [312] Dasgupta, S.; Herbert, J. M. Standard grids for high-precision integration of modern density functionals: SG-2 and SG-3. *Journal of Computational Chemistry* **2017**, *38*, 869–882.
- [313] Uings, I. J.; Farrow, S. N. *Cell receptors and cell signalling*; 2000; Vol. 53; pp 295–299.
- [314] Amino, N. Receptors in disease: an overview. *Clinical Biochemistry* **1990**, *23*, 31–36.
- [315] Majumdar, S.; Siahaan, T. J. Peptide-mediated targeted drug delivery. *Medicinal Research Reviews* **2012**, *32*, 637–658.
- [316] van Zutphen, S.; Robillard, M. S.; van der Marel, G. A.; Overkleeft, H. S.; den Dulk, H.; Brouwer, J.; Reedijk, J. Extending solid-phase methods in inorganic synthesis: the first dinuclear platinum complex synthesised via the solid phase. Electronic supplementary information (ESI) available: Definitions for abbreviations used, and synthesis of 1a, 1b, 4 and 7. See <http://>. *Chemical Communications* **2003**, 634–635.
- [317] Accardo, A.; Morelli, G. Review peptide-targeted liposomes for selective drug delivery: Advantages and problematic issues. *Biopolymers* **2015**, *104*, 462–479.
- [318] Okarvi, S. Peptide-based radiopharmaceuticals: Future tools for diagnostic imaging of cancers and other diseases. *Medicinal Research Reviews* **2004**, *24*, 357–397.
- [319] Tung, C.-H. Fluorescent peptide probes for in vivo diagnostic imaging. *Biopolymers* **2004**, *76*, 391–403.
- [320] Yano, A.; Onozuka, A.; Matin, K.; Imai, S.; Hanada, N.; Nisizawa, T. RGD motif enhances immunogenicity and adjuvanicity of peptide antigens following intranasal immunization. *Vaccine* **2003**, *22*, 237–243.

- [321] Bray, B. L. Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nature Reviews Drug Discovery* **2003**, *2*, 587–593.
- [322] Adessi, C.; Soto, C. Converting a Peptide into a Drug: Strategies to Improve Stability and Bioavailability. *Current Medicinal Chemistry* **2002**, *9*, 963–978.
- [323] Tugyi, R.; Uray, K.; Ivan, D.; Feller, E.; Perkins, A.; Hudecz, F. Partial D-amino acid substitution: Improved enzymatic stability and preserved Ab recognition of a MUC2 epitope peptide. *Proceedings of the National Academy of Sciences* **2005**, *102*, 413–418.
- [324] Kluskens, L. D.; Nelemans, S. A.; Rink, R.; de Vries, L.; Meter-Arkema, A.; Wang, Y.; Walther, T.; Kuipers, A.; Moll, G. N.; Haas, M. Angiotensin-(1-7) with Thioether Bridge: An Angiotensin-Converting Enzyme-Resistant, Potent Angiotensin-(1-7) Analog. *Journal of Pharmacology and Experimental Therapeutics* **2009**, *328*, 849–854.
- [325] Ahmed, S.; Mathews, A. S.; Byeon, N.; Lavasanifar, A.; Kaur, K. Peptide Arrays for Screening Cancer Specific Peptides. *Analytical Chemistry* **2010**, *82*, 7533–7541.
- [326] Frei, A. P.; Jeon, O.-Y.; Kilcher, S.; Moest, H.; Henning, L. M.; Jost, C.; Plückthun, A.; Mercer, J.; Aebersold, R.; Carreira, E. M.; Wollscheid, B. Direct identification of ligand-receptor interactions on living cells and tissues. *Nature Biotechnology* **2012**, *30*, 997–1001.
- [327] Hall, M. P. *et al.* Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chemical Biology* **2012**, *7*, 1848–1857.
- [328] Zhang, W.; Li, P.; Yang, F.; Hu, X.; Sun, C.; Zhang, W.; Chen, D.; Tang, B. Dynamic and reversible fluorescence imaging of superoxide anion fluctuations in live cells and in vivo. *Journal of the American Chemical Society* **2013**, *135*, 14956–14959.
- [329] Duke, R. M.; Veale, E. B.; Pfeffer, F. M.; Kruger, P. E.; Gunnlaugsson, T. Colorimetric and fluorescent anion sensors: an overview of recent developments in the use of 1,8-naphthalimide-based chemosensors. *Chemical Society Reviews* **2010**, *39*, 3936.
- [330] Tian, Z.; Liu, Y.; Tian, B.; Zhang, J. The synthesis and fluorescence properties of

- novel 1,8-naphthalimide derivatives. *Research on Chemical Intermediates* **2015**, *41*, 1157–1169.
- [331] Alexiou, M.; Tyman, J. Reactions of Isomeric Dinitronaphthalic-1,8-Anhydrides with Alkylamines. *Journal of Chemical Research* **2001**, *2001*, 59–62.
- [332] Langdon-Jones, E. E.; Symonds, N. O.; Yates, S. E.; Hayes, A. J.; Lloyd, D.; Williams, R.; Coles, S. J.; Horton, P. N.; Pope, S. J. Fluorescent Rhenium-Naphthalimide Conjugates as Cellular Imaging Agents. *Inorganic Chemistry* **2014**, *53*, 3788–3797.
- [333] Reichlin, S. Somatostatin. *New England Journal of Medicine* **1983**, *309*, 1495–1501.
- [334] Reubi, J. C. Neuropeptide receptors in health and disease: the molecular basis for in vivo imaging. *Journal of Nuclear Medicine* **1995**, *36*, 1825–1835.
- [335] Reubi, J. Somatostatin Receptors in Human Endocrine Tumors. *Cancer Research* **1987**, *47*, 551–558.
- [336] Bauer, W.; Briner, U.; Doepfner, W.; Haller, R.; Huguenin, R.; Marbach, P.; Petcher, T. J.; Pless, J. SMS 201–995: A very potent and selective octapeptide analogue of somatostatin with prolonged action. *Life Sciences* **1982**, *31*, 1133–1140.
- [337] Setyono-Han, B.; Henkelman, M. S.; Foekens, J. A.; Klijn, J. G. Direct Inhibitory Effects of Somatostatin (Analogues) on the Growth of Human Breast Cancer Cells. *Cancer Research* **1987**, *47*, 1566–1570.
- [338] Fields, G. B.; Noble, R. L. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *International journal of peptide and protein research* **1990**, *35*, 161–214.
- [339] Hsieh, H.-P.; Wu, Y.-T.; Chen, S.-T.; Wang, K.-T. Direct solid-phase synthesis of octreotide conjugates. *Bioorganic & Medicinal Chemistry* **1999**, *7*, 1797–1803.
- [340] Zhao, W. *et al.* Structural insights into ligand recognition and selectivity of somatostatin receptors. *Cell Research* **2022**, *32*, 761–772.

- [341] Robertson, M. J.; Meyerowitz, J. G.; Panova, O.; Borrelli, K.; Skiniotis, G. Plasticity in ligand recognition at somatostatin receptors. *Nature Structural & Molecular Biology* **2022**, *29*, 210–217.
- [342] Cao, S. S.; Kaufman, R. J. Endoplasmic Reticulum Stress and Oxidative Stress in Cell Fate Decision and Human Disease. *Antioxidants & Redox Signaling* **2014**, *21*, 396–413.
- [343] Panda, S.; Bhol, C. S.; Bhutia, S. K.; Mohapatra, S. PEG–PEI-modified gated N-doped mesoporous carbon nanospheres for pH/NIR light-triggered drug release and cancer phototherapy. *Journal of Materials Chemistry B* **2021**, *9*, 3666–3676.
- [344] Fang, Z.; Lu, Z.; Han, S.; Zhou, Y.; Yang, W.; Zhang, X.; Zhou, X. The Transcriptome-Wide Mapping of 2-Methylthio- N⁶-isopentenyladenosine at Single-Base Resolution. *Journal of the American Chemical Society* **2023**, *145*, 5467–5473.
- [345] Ma, Z.; Zhang, P.; Yu, X.; Lan, H.; Li, Y.; Xie, D.; Li, J.; Yi, T. Sugar based nanotube assembly for the construction of sonication triggered hydrogel: an application of the entrapment of tetracycline hydrochloride. *Journal of Materials Chemistry B* **2015**, *3*, 7366–7371.
- [346] Mohamed, Z. H.; Rhein, C.; Schmid, B.; Tripal, P.; Kornhuber, J.; Arenz, C. Synthesis and characterization of a new two photon excitable acid sphingomyelinase FRET probe. *Bioorganic & Medicinal Chemistry* **2021**, *44*, 116303.
- [347] Liu, J.; Zhou, S.; Ren, J.; Wu, C.; Zhao, Y. A lysosome-locating and acidic pH-activatable fluorescent probe for visualizing endogenous H₂O₂ in lysosomes. *Analyst* **2017**, *142*, 4522–4528.
- [348] Reddy, T. S.; Maragani, R.; Dhokale, B.; Mobin, S. M.; Misra, R. Heteroatom-connected ferrocenyl substituted naphthalimides. *RSC Advances* **2016**, *6*, 7746–7754.
- [349] Wan, C.; Feng, Y.; Hou, Z.; Lian, C.; Zhang, L.; An, Y.; Sun, J.; Yang, D.; Jiang, C.; Yin, F.; Wang, R.; Li, Z. Electrophilic Sulfonium-Promoted Peptide and Protein Amidation in Aqueous Media. *Organic Letters* **2022**, *24*, 581–586.

- [350] Zhang, X.; Fu, Y.; Qian, G.; Zhang, R.; Xu, Z. P. An artificial protein-probe hybrid as a responsive probe for ratiometric detection and imaging of hydrogen peroxide in cells. *Journal of Materials Chemistry B* **2020**, *8*, 5420–5424.
- [351] Schottelius, M.; Wester, H.-J.; Reubi, J. C.; Senekowitsch-Schmidtke, R.; Schwaiger, M. Improvement of Pharmacokinetics of Radioiodinated Tyr 3 -Octreotide by Conjugation with Carbohydrates. *Bioconjugate Chemistry* **2002**, *13*, 1021–1030.
- [352] Shao, Y. *et al.* Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Molecular Physics* **2015**, *113*, 184–215.
- [353] Epifanovsky, E. *et al.* Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *The Journal of Chemical Physics* **2021**, *155*, 084801.
- [354] Truong, T. N.; Stefanovich, E. V. A new method for incorporating solvent effect into the classical, ab initio molecular orbital and density functional theory frameworks for arbitrary shape cavity. *Chemical Physics Letters* **1995**, *240*, 253–260.
- [355] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *Journal of Computational Chemistry* **2003**, *24*, 669–681.
- [356] Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *The Journal of Physical Chemistry* **1992**, *96*, 135–149.
- [357] Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. A doubles correction to electronic excited states from configuration interaction in the space of single substitutions. *Chemical Physics Letters* **1994**, *219*, 21–29.
- [358] Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of Chemical Physics* **1999**, *110*, 6158–6170.
- [359] Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional us-

- ing the Coulomb-attenuating method (CAM-B3LYP). *Chemical Physics Letters* **2004**, *393*, 51–57.
- [360] Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theoretical Chemistry Accounts* **2008**, *120*, 215–241.
- [361] McLean, A. D.; Chandler, G. S. Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, $Z=11-18$. *The Journal of Chemical Physics* **1980**, *72*, 5639–5648.
- [362] Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *The Journal of Chemical Physics* **1980**, *72*, 650–654.
- [363] Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li-F. *Journal of Computational Chemistry* **1983**, *4*, 294–301.
- [364] Frisch, M. J.; Pople, J. A.; Binkley, J. S. Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *The Journal of Chemical Physics* **1984**, *80*, 3265–3269.
- [365] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *The Journal of Chemical Physics* **1992**, *96*, 6796–6806.
- [366] Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method (MOM). *The Journal of Physical Chemistry A* **2008**, *112*, 13164–13171.
- [367] Barca, G. M. J.; Gilbert, A. T. B.; Gill, P. M. W. Simple Models for Difficult Electronic Excitations. *Journal of Chemical Theory and Computation* **2018**, *14*, 1501–1509.

- [368] Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design* **2013**, *27*, 221–234.
- [369] Lu, C.; Wu, C.; Ghoreishi, D.; Chen, W.; Wang, L.; Damm, W.; Ross, G. A.; Dahlgren, M. K.; Russell, E.; Von Bargen, C. D.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *Journal of Chemical Theory and Computation* **2021**, *17*, 4291–4300.
- [370] Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery* **2015**, *10*, 449–461.

Appendix

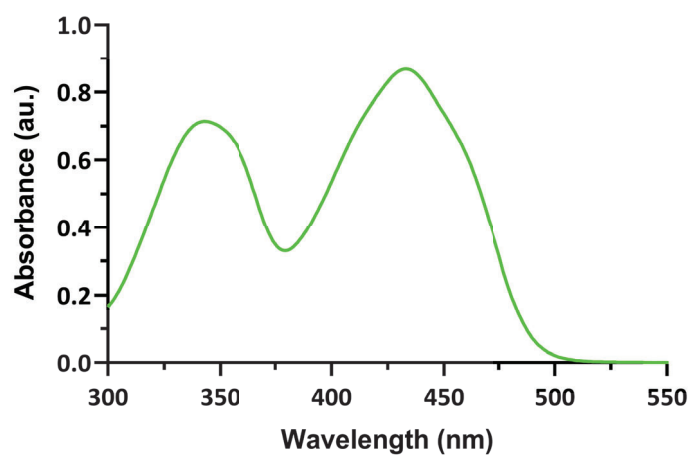


Figure S1: Absorbance spectrum of *N*-ethyl flavin (0.100 mM) in water.

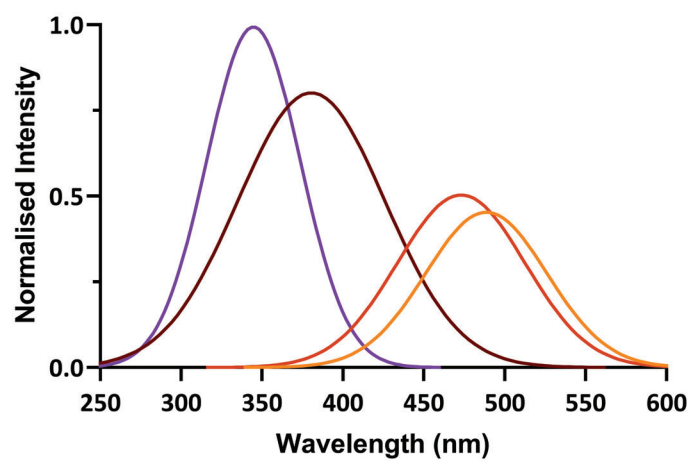


Figure S2: Fitted Gaussians for each **NpFR1** absorbance peak.

Table S1: Overlap of significant transitions calculated using MOM and the SG-0 grid.

Transition	Orbital Overlap
HOMO - LUMO	1.00
HOMO - LUMO+1	1.00
HOMO-1 - LUMO	1.00

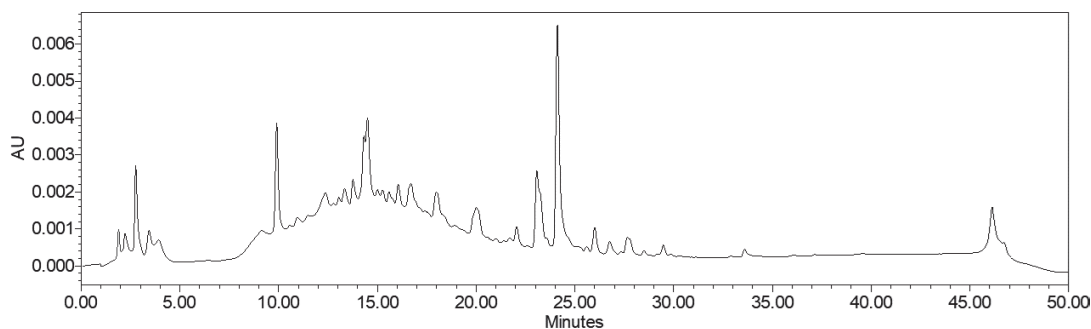


Figure S3: HPLC trace of **BNp-COOH**-labelled octreotide in 0 - 100% acetonitrile in water with 0.1% TFA in water over 60 mins after using standard cleavage solution.

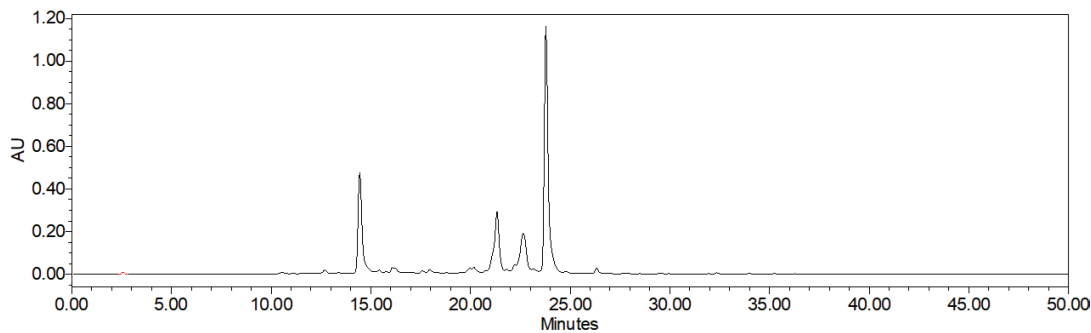


Figure S4: HPLC trace of **BNp-COOH**-labelled octreotide in 0 - 100% acetonitrile in water with 0.1% TFA in water over 30 mins after using reagent K.

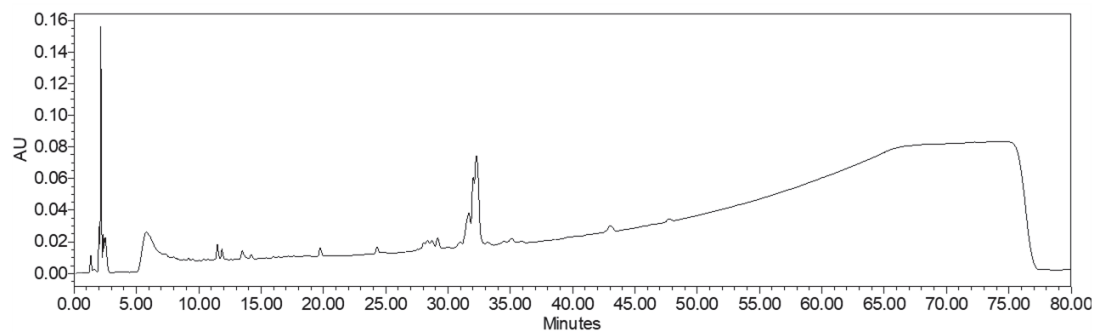


Figure S5: HPLC trace of **Np-COOH**-labelled octreotide dissolved in DMSO after sitting for 12 h. Run in 15 - 100% acetonitrile in water with 0.1% TFA in water over 60 mins using C8.

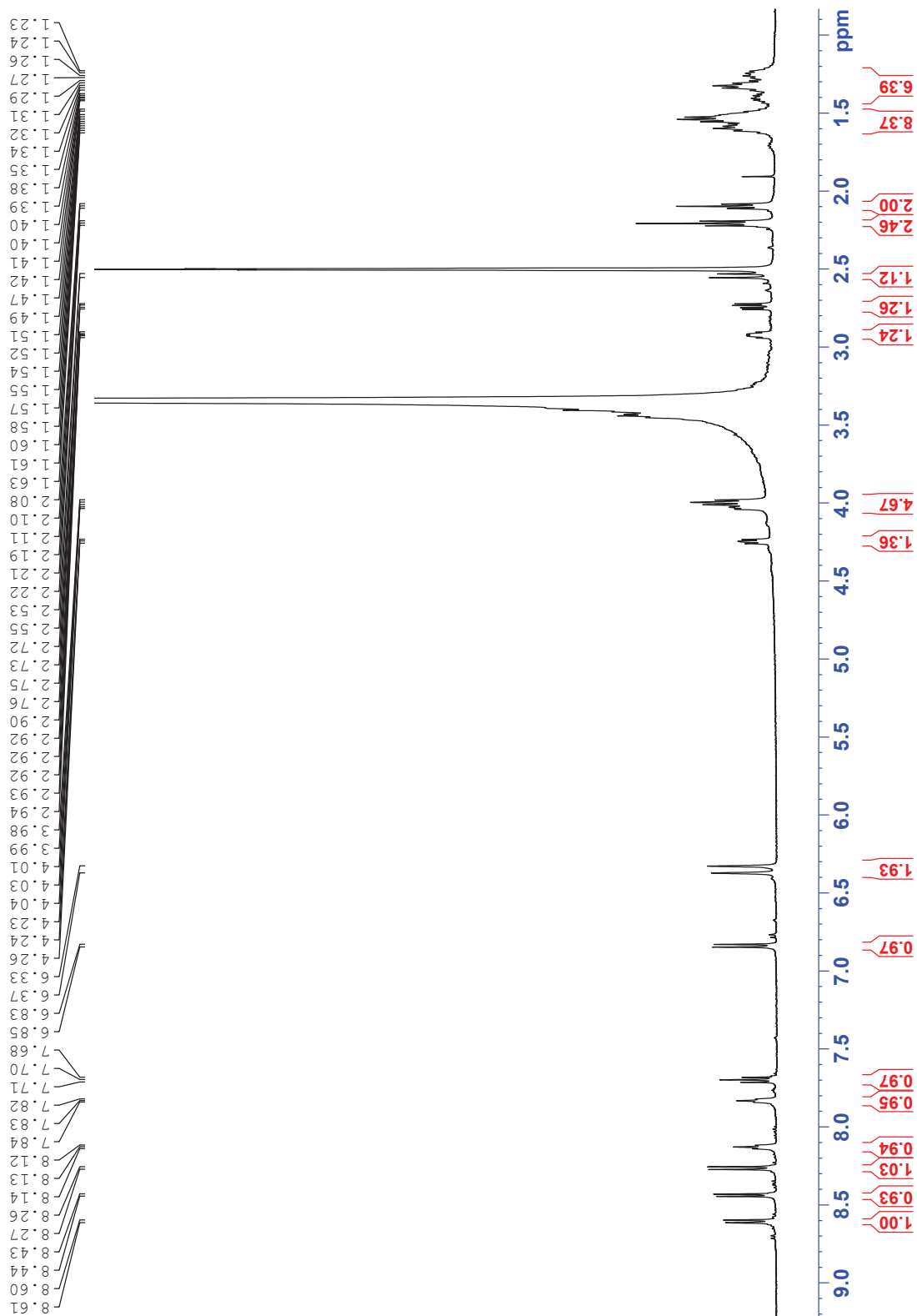


Figure S6: ^1H NMR spectrum of **CyMT**.

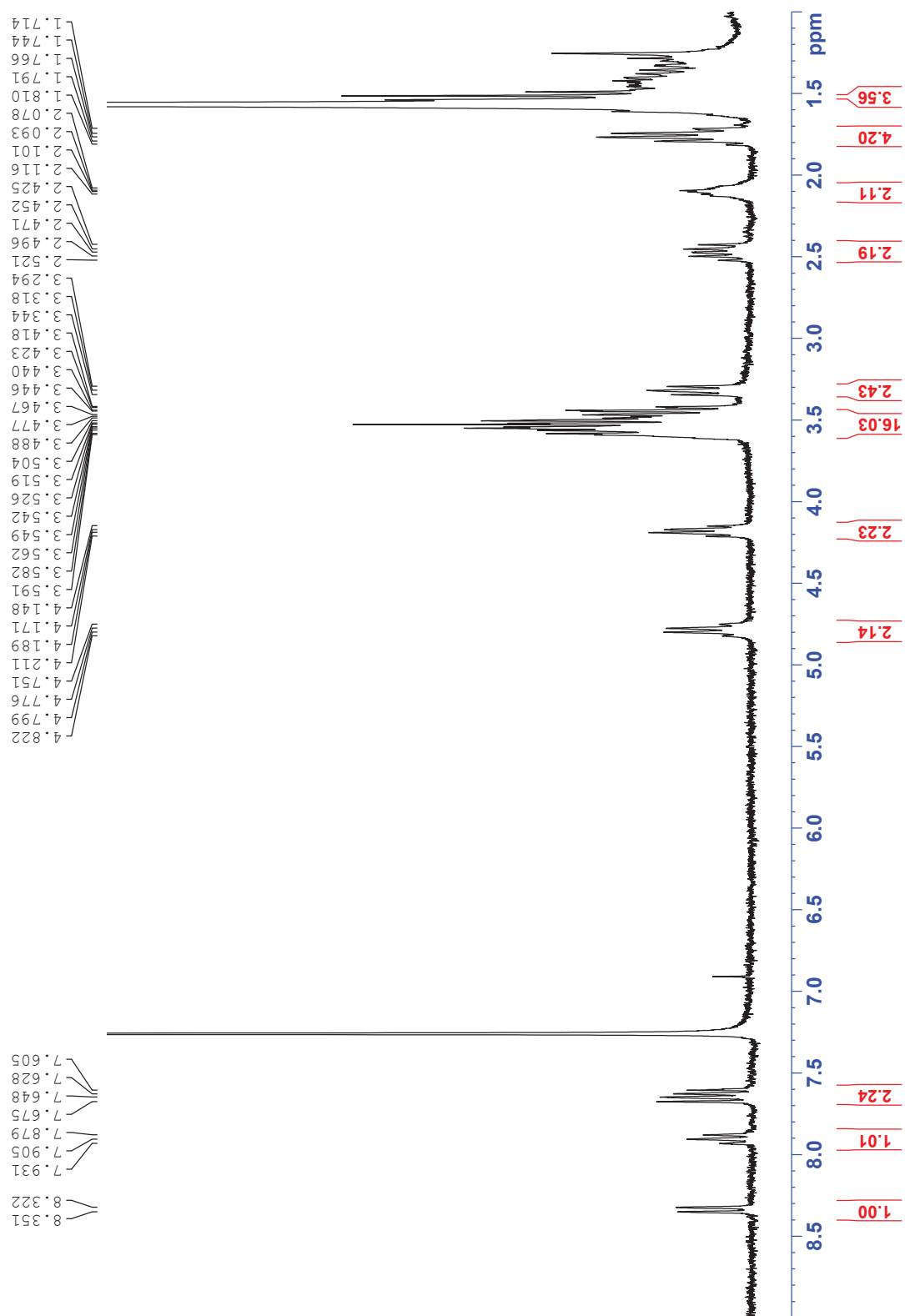
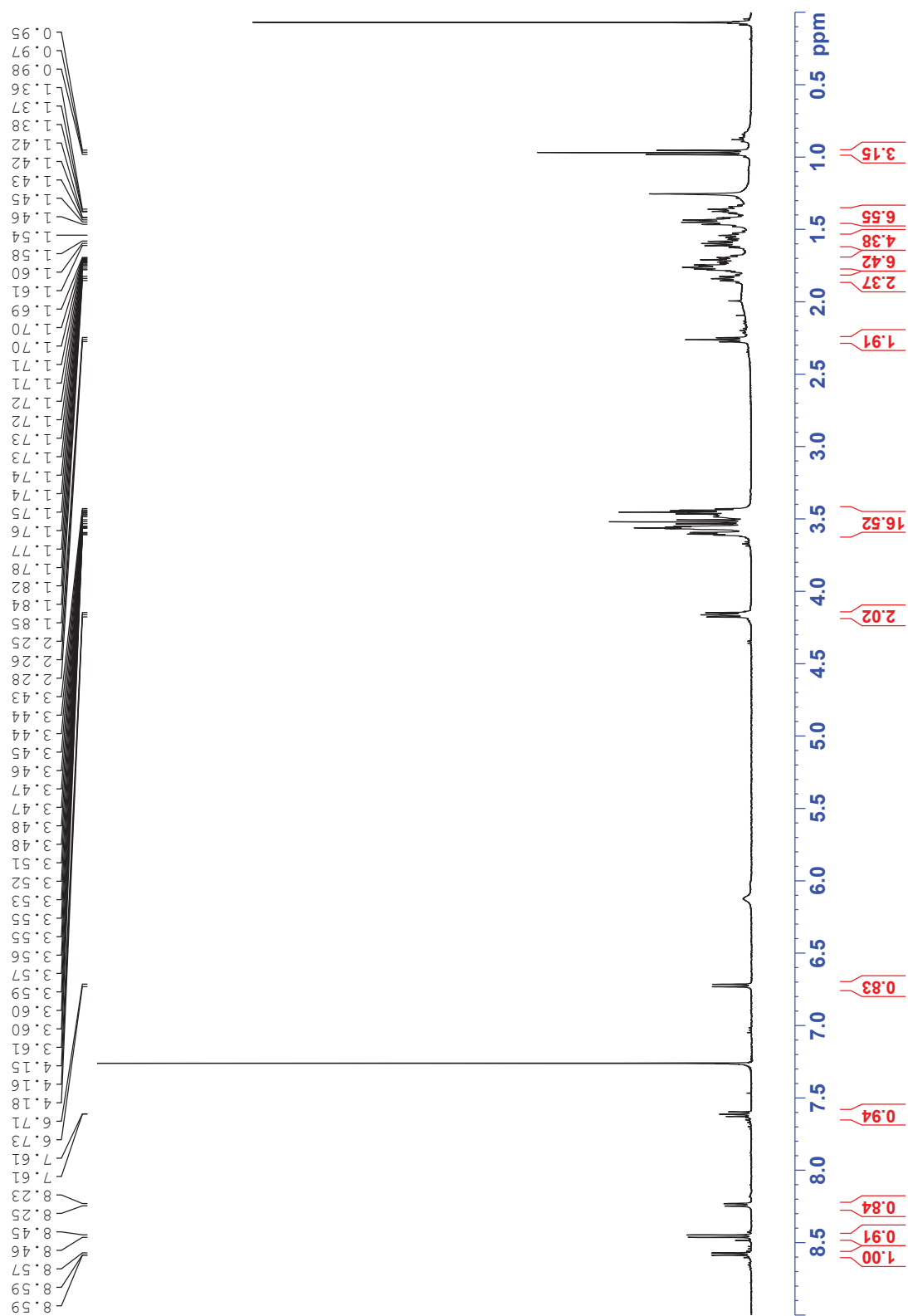


Figure S8: ^1H NMR spectrum of **FlavinTag**.



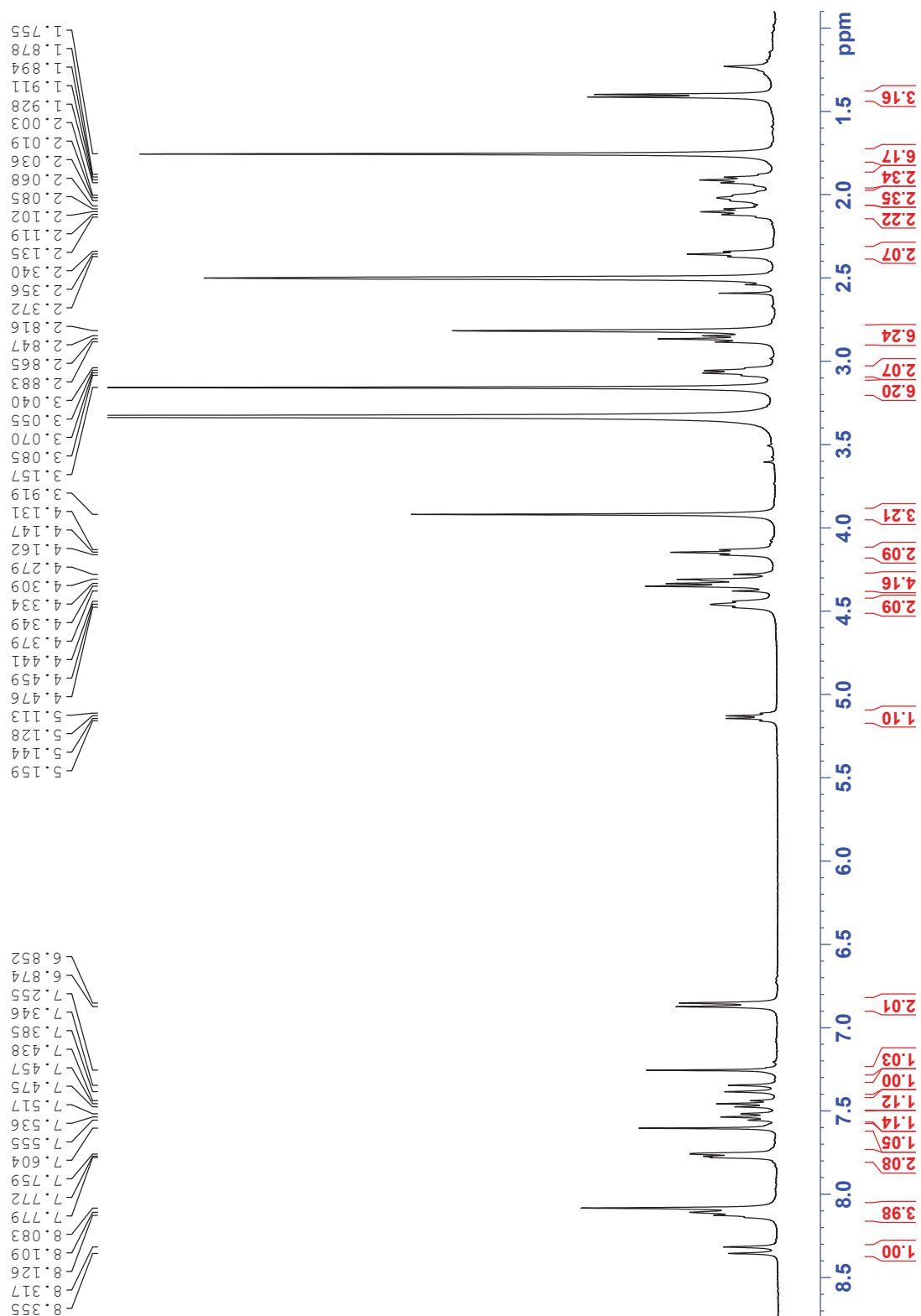


Figure S10: ^1H NMR spectrum of **BNp-COOH**.

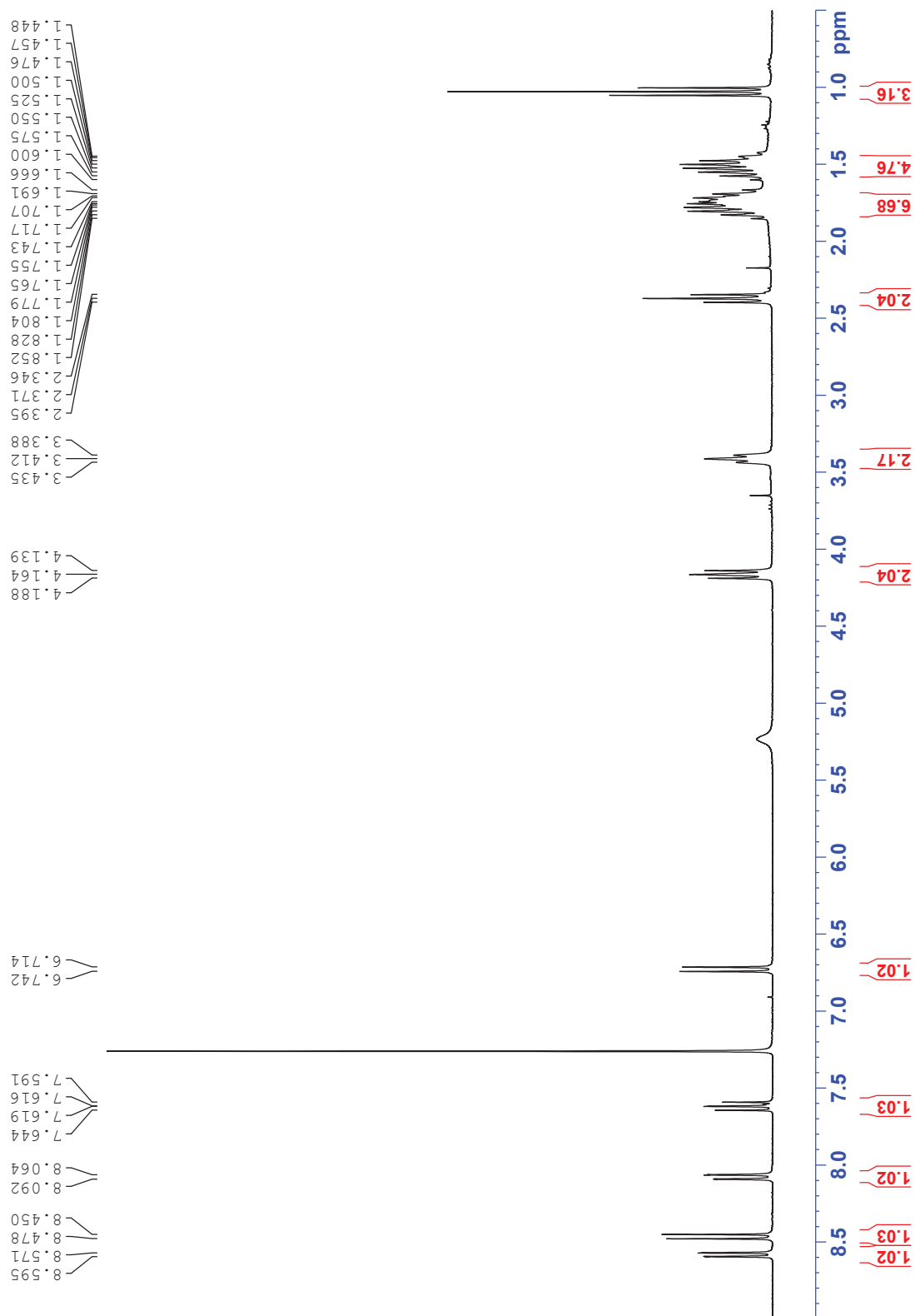


Figure S11: ^1H NMR spectrum of **Np-COOH**.

Examiners comment	Correction	Page/ref
There was a lack of NMR and MS data provided as Figures to support the syntheses. These will need to be added as evidence of the conclusions.	Added	Appendix
Abbreviation corrections: Chloroform-d4 should be chloroform-d Copper catalysed azide alkyne click should be copper catalysed azide-alkyne Huisgen cycloaddition HEPES should be 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid LCMS should be liquid chromatography mass spectrometry UV-Vis is just ultraviolet-visible	These have been amended	xi - xv
The section on multimodal imaging was a bit light on information and it was not always clear what specific benefits dual-modality brings in each case. Specific examples of what dual modality images look like and how they visually complement each other should be included to demonstrate these benefits.	Rewrote literature examples to highlight more clearly the benefit of the dual modal studies.	28
	Added: "Such multimodal probes allow for the detection of complimentary macroscopic, and microscopic information."	29
	"This dual-modal probe not only enabled fluorescent localisation of VEGFR-2 receptors in whole mice, but the addition of PET capability also allowed for accurate quantification of targeting efficacy."	30
	"While useful at providing some level of localisation, SPECT imaging alone, does not allow for precise tumour delineation during surgery. By incorporating the fluorophore, it was found that greater delineation of the tumour could be achieved during surgery, which allowed for excellent results from the radical excision of lymph-node tumour tissue in mouse models."	30
'spacial' should be 'spatial'	Amended	8, 29, 36, 38
amino acids are 'connected' by amide bonds rather than 'tethered'	Amended	9
Hereceptin should be Herceptin.	Amended	12
should be single-photon emission 'computed' tomography (SPECT)	Amended	14
radiolabelling doesn't always involve peptides.	Corrected to "Radiolabelling involves incorporating a radionuclide onto a compound of interest, and is commonly used to label peptides and proteins." for clarification	18
an explanation of 'high specific activity' and its importance is needed.	Added "defined as the number of radioactive emissions per unit mass".	18
Add in the tissue penetration limits of different wavelengths of light.	Added "This is in contrast to fluorescence, which is limited by optical scattering. However, penetration depth of fluorescence is largely dependent on wavelength where longer wavelengths allow for deeper penetration." With reference	28
Provide a definition for temporal resolution.	Added "the duration of time between successive data measurements"	28
It would be useful to provide the structures of probes such as CPH (page 27) and Gd-DOTA-DEVD-AFC (page 28).	Added probe descriptors including "CPH, a cyanine-based probe" and "gadolinium complex, Gd-DOTA-DEVD-AFC"	29
It is worth noting that MRI is versatile, is commonly used in diagnostic medicine and research and provides images of anatomy and physiological processes.	Added "MRI is a versatile technique that is frequently used in research and diagnostic medicine to provide information about anatomy and physiological processes. However, MRI is a relatively insensitive technique, and often requires the use of contrast agents."	29
Paramagnetic ions' instead of just 'paramagnetic' and Gd3+ should be Gd ³⁺ .	Amended	29
Some background on ¹⁹ F spin and use in MRI is required and it should be explained why both modalities (fluorescence imaging and MRI) was required for detecting caspase-3 enzyme activity.	Added "Other contrast agents include heavy isotopes of atoms with a non-zero net spin, such as ¹⁹ F."	29
SPECT radiotracers are not necessarily always 'taken up by the cells in the body' and it certainly is not a requirement for emitting gamma rays.	Omitted "are taken up by cells in the body"	30
Clarify what is meant by 'correlation and quantified	Typo. Corrected to "correlated and quantified"	30

information'.		
I think it is important to add information regarding the large differences in sensitivity, tissue penetration and resolution between fluorescence and radioimaging.	Added "Fluorescence imaging offers superior spatial resolution (~ 200 nm) in comparison to radioimaging (~ 5 mm). However, a drawback of fluorescence imaging is its limited tissue penetration depth, which varies based on the wavelength of the fluorophore employed. On the other hand, radioimaging has the advantage of deeper tissue penetration, enabling imaging of structures at greater depths."	30
Check consistency of spelling for 'tumours' and 'tumors'.	All corrected to 'tumour(s)'	30 and throughout
. 'ual-modality should be 'dual-modality.	Amended	30
PET 'recognition' would be better described as 'detection' or 'imaging.	Changed 'recognition' to 'detection'	30
The example of the dual PET and fluorescence imaging probe does not describe the different complementary information that was gathered from the two techniques so this should be included.	This comment was addressed in a previous correction	30
Does fluorescence provide complementary temporal as well as spatial information to the vibrational technique?	Added "and temporal"	31
The thesis would benefit from visual representations of published multi-modal images to appreciate the benefits of each approach.	Literature examples of the presentation of multi-modal images have already been given in the review segment of the introduction.	Section 2.1
it was never made clear to me what potential proteins were to be analysed and why	Added "allowing it to be attached to any protein of interest containing a primary amine."	40
It was also not clear how an NHS ester could be used to label a protein unless the protein was already isolated.	This was touched on in section 2.8 (planned tissue studies as well as the conclusion.	Section 2.8
The applications for these probes, including the proteins to be studied and the conditions for attaching the probes, should be clarified.	Added "This would allow the analysis of any protein of interest."	39
A low-res MS was mentioned but it should be included as a Figure.	No page was mentioned on this comment, so I'm unsure what this is in relation to. Figures of the low-res studies based on CyMT have already been provided.	Not specified
mass to charge ratio or mass-to-charge ratio (<i>m/z</i>)?	Changed to 'mass-to-charge'.	33
'detected' instead of 'detector'.	Amended	33
'record' instead of 'recorded'.	Amended	34
'immunohistochemistry' instead of 'immunohistochemsitry'	Amended	37
'demonstrated' not 'demonstration'	Amended	37
should it read 'mass-tags and MSI' rather than 'mass-tags and multimodality'?	This sentence is in reference to using mass-tags for multimodality, and thus was not changed.	39
Are there any requirements for the fluorescent moiety? Such as excitation and emission wavelengths?	Added "This fluorophore must not exhibit emission within the range of the photo-cleavable linker, so as to avoid self-induced cleavage."	40
It is better to state that NHS-activated esters react with primary amines rather than are 'specific to primary amines'	Changed to "react to primary amines, specifically".	40
Figure 2.3 Indicate that the ONB linker is in blue and that the NHS is in red.	Added "(blue)" and "(red)" specification in the caption.	Figure 2.3
what is the significance of the 'presence of water' that enables cyclisation to occur?	The original reference to this section was added and "in the presence of water" was removed to avoid confusion.	42
Fig 2.4 -The charge on the nitro group N in the top right corner should be +.	Amended	Fig 2.4
Inherent positive charge allows for the tag to be 'readily detected' by which technique?	Changed to "readily detected by MS".	42
What wavelengths are in the orange region? How deep is 'deeper tissue imaging'? And what tissue are you intending to analyse?	Added citation to literature reference describing penetration depths of difference wavelengths of light.	42
Figure 2.5, you could label this molecule as CyMT	Added CyMT to figure caption	43
Was sodium iodide needed for the reaction to work? Why?	This was already addressed by reference to the 'Finkelstein-type' reaction. Therefore, no changes were made.	43
how was the formation of the cyclised product confirmed?	These characterisations were preliminary and will need to be repeated before publication.	46
The C:N ratio should be ≤ 3, or between 1 and 3 are not synthesised.....	Amended	48
Include the absorbance spectra in solution and compare with the solid state.	This has been noted as a comparison to be performed in future work.	49

The sentence 'If medium-free....' is unclear. Does 'washing and drying' refer to removing the medium and will this occur after fluorescence imaging?	Fixed to say "Medium-free imaging eliminates the need for removal of imaging medium post fluorescence imaging, reducing the risk of sample disruption or contamination that can impact the accuracy of subsequent analysis."	50
Literature references about the different ways aggregation causes quenching would strengthen the argument.	Reference added	50
'It is well known....' Provide references.	References added	51
Does the solution phase absorbance provide any indication for J- or H-aggregation?	Added "This cannot be elucidated from the solution-phase absorbance either, as the spectroscopic properties and behaviour of molecules in the solid state can differ significantly from those in solution."	51
How much excess of CyMT was used?	This was already given in the methods	61
You could use 'eluted' rather than 'run through'	Changed to "eluted from"	62
What is the magnification of the image? Provide some discussion on why the image in Figure 2.20 looks so different to this one?	The magnification could not be determined as it is taken from a photo-image	Figure 2.21
What different information could be garnered from the multimodal imaging in this example? Or is it enough in this preliminary study that there was something observed using both techniques?	This was a preliminary assessment of whether it would be possible to observe signal from both techniques sequentially using the same sample.	Whole of chapter 2
What are the limitations of this probe and the images gathered thus far? How would you address these?	These were assessed within the conclusion chapter.	Section 5.1
it is disappointing that instrument failure prevented further studies, but it would still be useful to describe the intended study in some detail to display the planning and knowledge gathering that was presumably done. What protein was being studied? How was it going to be modified selectively in a tissue sample?	Added "...conjugation of CyMT to an isolated protein using immunohistochemical techniques, allowing it to be assessed within a tissue sample for fluorescence imaging..."	64
The choice of amino acid sequence seems rather random. There are many combinations of amino acids that would provide water solubility and ionisation potential so why was this one chosen in particular? Provide a reference if you can.	This was a random selection.	65
'Scheme Scheme 2.9' is a typo.	Amended	66
'Wang' has a capital W.	Amended	68
Peptide synthesis sounds better than 'elongation'.	Amended	68
Please include the low-res MS spectrum of NpMT	This was only preliminary characterisation and will need to be repeated for publication.	69
Preparative HPLC was mentioned as completed but 'purification and isolation...was not performed'. Can you clarify this statement?	Typo. This was changed to 'analytical HPLC' rather than 'preparative HPLC'.	68
There is a lack of figures in this section. I recommend adding a figure of the cell pointing out the areas/organelles and some of the sensors discussed in this section.	Added structures for some of the mentioned probes.	Chapter 3 intro
Several of the chemical abbreviations are mis-represented such as 'H2O2' should 'H ₂ O ₂ '.	Amended	71
'Alzheimer's' should be 'Alzheimer's disease'.	Amended	71
Retention rather than retainment.	Amended	73
Hypochlorous rather than hypochlorus.	Amended	74
The names in the Figure don't match the caption.	Amended	Figure 3.3.
Does AutoDock software have a reference? Who makes/publishes it?	This is referenced in the next sentence but I have moved the reference to the earlier sentence.	85
'molar extinction spectrum' should be 'molar extinction coefficient'?	Molar extinction spectrum is correct. No change was made.	89
'taken' should be 'produced or analysed'	Changed to 'measured'	90
denote the absorbance (dashed) and emission (solid line).	"absorbance (dashed) and emission (solid)" added to figure caption.	Figure 3.8
What does it mean to 'scale the absorption spectrum'?	Changed to "This value was used to convert the absorption spectrum to a plot of molar absorptivity"	90
Is it 17.6 or 16.62?	Changed in text to '16.62'	Figure 3.10.
Split the images into (a) and (b) and indicate in the caption which molecule is bound by the cyclohexane group.	Added to caption "separated at 1 carbon beyond the last atom involved in the chromophores' conjugation"	Figure 3.12
References for positive sstr2 expression on PC3 cells and for the use of Huvec-TERET2 cells need to be provided. I was under the impression that sstr2 expression was minimal for PC3 cells (Cancers, 2022, 14(10): 2513, doi: 10.3390/cancers14102513).	Added "identified using the Human Protein Atlas (https://www.proteinatlas.org/)"	138
The first sentence is not clear.	Added "was" for clarity.	Section 4.3.1
Provide a literature reference for the copper(II)-catalysed SnAr reaction.	Changed to 'Ullman-type reaction' – see comment from examiner 2.	124

'KCO3' should 'K ₂ CO ₃ '. Bioin should be biotin in the caption.	Amended	Scheme 4.2
Provide a literature reference for the statement 'copper catalyst may induce formation of product'?	In relation to comment for page 119.	124
Do you have an explanation for why DMSO at reflux was successful? a mechanism? a reference? some discussion? solubility? temperature? DMSO and THF are both polar aprotic solvents.	Reference to the paper that used this was already given. Mechanism is not fully understood.	125
Provide a definition for 'bleed-through'.	Added "(fluorescence signal caused by the excitation wavelength)"	126
There is no LCMS data in Figure S4. There needs to be some confirmation that this peptide conjugate has been synthesised. Please include the data as a figure in this chapter.	This was given by LCMS data in the methods. Also, Figure S4 is the analytical HPLC trace.	Experimental and appendix.
How was the peak at 38 minutes confirmed as the product?	Added "as confirmed by LCMS".	143
Isochratic should be isocratic.	Amended	136
Check for several spelling typos in the first paragraph. For example, 'Streptavidin' should be 'Streptavidin'.	Amended	Section 4.8.
I think the statement 'The probe was shown to be successfully incorporated into a peptide of interest, octreotide' might not be entirely accurate based on the evidence provided. See comments on chapter 6.	Amended in the comment presented further down. (pg 175)	Section 4.8.
'in vacuo' should be 'in vacuo'	Amended	Methods section
'precipitate which was filtered' should be 'precipitate which was collected by filtration.'	Amended	155
'dried by rotary evaporation' should be 'solvent was removed by rotary evaporation'.	Amended	158
'DMSO-d ₆ ' should be 'DMSO-d ₆ '	Amended	methods
'MeOH' should be 'MeOH-d ₄ '	All changed to CD ₃ OD for consistency	methods
'aminohexanoic acid' should '6-aminohexanoic acid'	Amended	168
More detail is required to explain how the loading was calculated. What was the extinction coefficient? Provide a reference if possible.	Added "Loading was calculated using an extinction coefficient of 1.65 cm ⁻¹ (mmol/g) ⁻¹ ."	177
I have concerns about the data presented here for the peptides. The LRMS characterisation of NpMT doesn't add up. The ESI- m/z value of 1293.05 can't be a 6- charge and the calculated value is 1297 not 1293 as stated. Same for ESI+, the value of 1350.18 can't be a 6+ charge.	These were all typos. Corrected formula from 'C ₆₄ H ₇₁ N ₁₁ O ₁₉ ' to 'C ₆₄ H ₆₇ N ₁₁ O ₁₉ ' bringing the mass to 1293 with a 6 ⁻ charge, as stated. Therefore, the 6+ charge is also correct in this light.	179-180
The LCMS data for Np-COOH octreotide (cyclised) has a different molecular formula (2 extra N atoms) compared to the given LRMS formula. Also, the calculated mass is 1410 not 1385.	Corrected formula from 'C ₇₁ H ₉₀ N ₁₄ O ₁₃ S ₂ ' to 'C ₇₁ H ₉₂ N ₁₂ O ₁₃ S ₂ ' making 1385 correct.	
The LCMS data for BNP-COOH octreotide can't be a 5+ ion with a value of 1600.15.	Corrected from 'C ₇₁ H ₉₂ N ₁₂ O ₁₃ S ₂ ' to 'C ₇₉ H ₁₀₆ N ₁₅ O ₁₅ S ₃ ' and corrected from [M+5H] ⁵⁺ to [M+2H] ²⁺ , making 1600.15 correct.	
This data needs to be checked and provided in the relevant chapters because the data as it stands is not convincing.		
'uL' should be 'µL' and 'um' should 'µm'	Amended	181
References 8, 11, 16, 32, 105, 225, 228 are missing page numbers.	Amended	References
Reference 308 has typos	Amended	218
Examiner2		
1.3 ideal characteristics for a fluorophore – toxicity should be mentioned, or (slightly different concept) unwanted biological activity.	Added "Additionally, the fluorophore should also have low toxicity."	1.3
it is not clear the distinction between therapeutics and targeting agents as outlined in the Thesis, the latter is essentially dual-function? It is not clear and does not seem appropriate to have in different sections unless the distinction is clearer, e.g. the HER2 example seems the same principle as the ACE2 delivery example. This is a central theme to the Thesis so needs to be clear in the readers mind.	Changed location of title "targeting agents" to the section after HER2, which is now placed under the "therapeutics" title.	1.6.1 and 1.6.3
Page 112-115, are peptides discussed as direct 'drugs' i.e. the example in fig 5.2, or does "targeting agents" mean the peptide is the delivery vector and an additional "drug" is attached to the peptide? It seems to move between these concepts and is not explained well.	Added "They have been used both as therapeutic agents, as well as delivery agents." for clarity.	116
The aims on page 114, is "targeted drug therapy" simply a selective drug that is a peptide? Or adrug attached to a 'targeted' peptide ?	Also changed within the paragraphs from 'therapeutic' to 'clinical'	115

Page 123, I am still confused as the what "targeted medicine" is i.e. "breakthrough in the field of targeted medicine", it seems this "targeted medicine" is simply targeting a receptor overexpressed in a disease or such? This is not new a new concept (?) discovered in the 1980's?	Sentence was not relevant in the context of that paragraph and was therefore deleted.	116
Figures such as 1.11 and others – label what the cartoon is of. What is the blue ball? What is the bright green callout/star? This level of detail/legend is missing from a lot of the figures.	Added 'blue chain' and 'red arrow' in caption	119
Mid Page 20, "The advantage of using small fluorescent molecules is that they are small enough to permeate the cell membrane " - this is not correct as a general statement, many small molecule fluorophores do not permeate the cell membrane due to charges or other factors.	Changed to 'enter the cell membrane without the need for permeabilisation'	21
1.8.1 labelling proteins and 'self-labelling' – ligand-directed labelling should be mentioned and some references quoted. For example, https://pubs.acs.org/doi/10.1021/ja401956b ,	Added section on ligand-directed labelling: Ligand-directed labelling is another method for conjugating fluorescent sensors to a POI. One widely utilised approach involves the use of tosylates, wherein a ligand specifically designed for the POI is synthesized with a tosyl ester attached to a fluorescent label. The labelled ligand selectively binds to the active binding site of the target protein. Subsequently, an S _N 2-type reaction takes place between the tosylate group and a nearby nucleophilic amino acid, resulting in the transfer of the fluorescent probe to the target protein through acyl transfer. \cite{Tsuchiji2009} Compared to other labelling methods, ligand-directed chemistry is considered less invasive since it doesn't require modification of the POI. This techniques also enables labelling in close proximity to the active site of the protein.	28
1.9 Thesis Aims, I was expecting the aims to be a bit more specific, even narrowed down to a protein family, for example, aim 2, "Create and investigate new sensors designed to measure the redox state of <i>particular proteins</i> ." – what proteins? This sounds like a very vague aim without more details. To allow for what? Create new sensors better than existing ones? That can be used to do different things? Aim 3 is also very broad – is this a probe applicable to every peptide-receptor binding that occurs in a cell? Or a class of receptors? If it is broad (i.e. any peptide or receptor) this advantage should be specified/championed in the aim. In summary, at the end of the intro/aims, it is not obvious what the Thesis is going to be about, specifically. abstract also never mentions the name of a protein or receptor.	Added "isolated insulin proteins" in the abstract. The general aims of the thesis were not changed as individual chapter aims are more in depth.	Abstract
Starting page 42 and throughout the Thesis – it should be obvious what compounds have already been reported in the literature and what are novel compounds. It does not matter in terms of Thesis novelty if the first 2/3 of every scheme are compounds already reported in the literature, since the novelty lies mostly in the application. But this should be made clear within the Thesis, it currently is not clear either in the main chapter or in the experimental section. For literature compounds, does the NMR data match literature?	Added literature comparisons for compounds that were previously reported within the experimental section.	Experimental
Page 42, the nitration reaction – how was the regiochemistry of the product determined? 2D NMR should be discussed, or a literature reference provided and that the compound data matches. What type of reaction is this? Why was nitric acid alone enough for nitration, as opposed to the more common combination of nitric and sulfuric acid, or acetic anhydride and nitric acid?	I did not take a 2D NMR. Regiochemistry was not regarded as important to the function of the compound and was not assessed.	44
Page 45, the "cyclised product" at the end of scheme 2.3 – what data led to this structure being proposed? 1H NMR spectra? MS? In order to claim (or propose) this compound has formed some data needs to be presented. Data for this proposed compound is not presented in the experimental section either. What is the mechanism of how this proposed compound formed? "It is likely that activation of the hydroxyl group with the mesylate rendered it more likely to partake in	Amended in previous comment	46

<i>nucleophilic attack of the iminium.</i> ” But how is a mesylated alcohol more nucleophilic than an alcohol, when the Ms group is electron withdrawing and OMs is often used as a leaving group (which is actually what was planned in the next step)? Maybe depending on the reagent addition order, as soon as the TEA goes in the reaction occurs, nothing to do with MsCl (or OMs). It is a simple ether cyclisation. In fact, if the mesylate could be made/was made, it would prob severely limit the cyclisation reaction (?) since a nucleophilic alcohol would not be present.		
Although a commentary of all reaction yields is too much for a Thesis, there should be some comments (true reaction conversion versus tricky purification) regarding standard reactions that are usually high yielding that gave low yields. E.g. Scheme 2.6 product 14, why did a ‘standard’ amide coupling give a 17% yield? Why were yields in scheme 2.1 so low, i.e. one would not expect a LiOH (aq.) ester cleavage to yield 36%, the ester cleavage in scheme 2.2 is much higher yielding, why?	2.6 – Added “The lower than expected yield of this reaction were attributed to the use of an old batch of EDC.” 2.1/2.2 – the reaction in scheme 2.1 was performed on a larger scale and was not optimised, yielding lower than the one performed on a smaller scale in scheme 2.2	47/48
I might have missed something, but Fig 2.19, why was any ‘tagged’ insulin detected, i.e. why was the photocleavage not quantitative? And page 54, ‘showed the intact and cleaved’ molecules – why would 100% cleavage not be expected, why is it expected to see both after the MALDI laser? Where is this discussed? This seems important.	This was mentioned in the first instance of matrix assisted low resolution mass spectrometry. “This is expected, given that the matrix absorbs a lot of the energy from the laser, resulting in milder conditions.”	56
Fig 3.8, what solvent was this done in? Is the normalised intensity the same in all solvents? The solvent should be stated in the text or fig caption.	Added “Spectra were taken in MilliQ water.”	90
Page 114, what’s a “solid membrane” ?	Changed to ‘cellulose membrane’	118
Page 116, “This pull-down partner selectively binds to the fluorescently-labelled peptide-receptor complex. After binding to the magnetic beads, unbound contents are washed away, leaving only the target peptide-receptor complex.” How is this possible unless the peptide-receptor is a covalent and/or irreversible interaction? Otherwise the ‘unbound contents are washed away’ will also wash the non-covalently bound ligand from the receptor to a large extent. This warrants better explanation.	Added “Depending on the nature of the peptide-receptor interaction, this will leave behind either the peptide alone (reversible interaction), or the peptide-receptor complex (irreversible interaction).” Fixed caption to include “The bound peptide can be isolated using external magnets, after which unbound contents can be washed to yield the isolated peptide alone (not shown) or the peptide-receptor complex.”	120-121
Page 119 to 121, nucleophilic aromatic substitution – it is well known that SNAr needs a strong EWG (like R-NO ₂) ortho or para to the leaving group, and if this is not the case, then a very strong base like molten NaOH or NaNH ₂ is required. So was it surprising that it didn’t work (ie scheme 4.3)?	Added citation to a paper using the conditions in scheme 4.3	124
When Cu was tried, this then is not a normal SNAr mechanistically but a ‘Ullmann’ type reaction with oxidative addition and reductive elimination etc. It would be useful to at least touch on this.	Changed to “As such, an Ullmann-type reaction was attempted, as given in condition 6, where copper iodide was added as a catalyst...”	124
The reaction that did work, entry 8 in table 4.1. – what is the mechanism of this, is it a ‘normal’ SNAr? Why did these conditions work? This needs some explanation.	This mechanism is not well understood but a reference was given for the first use of these conditions to achieve S _N AR of a bromo-naphthlaimide.	125
Table 4.1 Entry 8, “successful reaction” - the yield should be given in the table/somewhere in 4.1	Changed to ‘88% yield’	125
Table 4.1 the ‘mixed products’ - this should at least state if the desired product was part of the mixture, or not.	Unsure- too messy to determine.	125
Page 124 or 127, why was the N-terminus chosen for fluorophore/probe attachment, and not, for example, the Lys side-chain, was this based on peptide SAR / sites required for activity/binding? This is not discussed and should be. For cyclic peptides like this, the N-terminus is not always the obvious/non-interfering position for probe attachment.	Added to probe design segment instead “The carboxylic acid moiety enables the conjugation of the molecule to either the N-terminus or a lysine residue of a peptide. The choice attachment site should be determined by the specific amino acids that are essential for receptor activity, and must be considered carefully during the design phase of the peptide. This is crucial in ensuring that the peptide retains its desired biological activity and properties after modification, and also in avoiding any undesirable effects or interference with the receptor binding.”	122
Page 124 ‘design’ didn’t actually outline the design of WHERE on the peptide the probe would be put. The concept that the small peptide+probe has different binding to the peptide could also be introduced. This is an important	Addressed in previous comment.	122

concept to discuss when 'tagging' things. It is mentioned for the first time of page 134, but ideally should come into the 'design' discussion too.		
Section 4.5.1 HPLC challenges, this all comes as a shock (that these products were never made in pure form?) since earlier it is declared products have been synthesised and even fluorescence spectroscopy has been carried out on the products, so at this stage, the reader assumes these products were pure. So what was measured in the spectra in 4.11 and 4.9? It should be made clear earlier on that this is of a crude mixture? What else is present? Also in the synthesis section it should not read as if the products were made and purified, which is how it reads.	Added "crude" for clarity.	132/133
The sentence "The probe was shown to be successfully (spelt wrong) incorporated into a peptide of interest, octreotide." on page 139 is inappropriate since it could not be purified. Also the control probe was not really 'synthesised' either i.e. isolated in pure form, this should be fleshed out in the conclusions.	Changed to "The probe was shown to be successfully incorporated into a peptide of interest, octreotide, as confirmed by LCMS and analytical HPLC. However, this labelled peptide could not be isolated due to instrument failure." The control probe alone was isolated and therefore was 'synthesised' – no changes made.	143
Section 4.7.1. is missing a lot of comparison and discussion of the literature, doing a quick internet search there are many reports of others doing MD with octreotide and SSTR2, and analogue ligands. This should be discussed and compared to the results in the Thesis. The whole literature area is completely missing here.	Added some literature comparisons of other MD simulations.	139
Page 139, cannot conclude unequivocally that "These preliminary investigations indicated that the interaction between octreotide and its receptor is unaffected by the presence of the BNP-COOH probe." There should be a COULD or IS LIKELY in here. Without biological data this cannot be declared, the modelling just indicates it is likely to be tolerated.	Added "likely".	143
Experimental – it is usual for PhD candidates to assign 1H NMR spectra (at least to functional groups if not to exact hydrogens) to demonstrate their understanding of NMR assignments etc. e.g. compound 1, at least writing ArH not just 'H' for the aromatic protons is pretty standard in a PhD thesis. It is a lot of work to write up the assignments all the compounds, so I suggest the candidate fully assigns a handful of the final products/larger compounds as a means to demonstrate their understanding of 1H NMR spectra peak assignment.	Assignments were performed on the following final compounds: - CyMT - 21 - FlavinTag - NaphTag - BNP-COOH - Np-COOH	Experimental
Some of the NMR data also seems a little odd, page 155, a J value of 15.5 Hz – what H is this? This compound should be assigned.	These have been assigned.	159
Page 156, the J values and multiplicities of the ArH's do not make sense? I would have expected to see dd's for a lot of the H's, esp. since a lot of the compounds were run at 500 MHz. Putting this structure into SciFinder tells me there are 42 references that contain/make this compound.	Added literature reference. Many naphthalimide Ar-H's only split as doublets.	experimental
Does the NMR data in the Thesis match literature?	NMR spectra for 3,4,8,11,24,32 were given in literature using different solvents and thus were not included in text.	Experimental
Page 156, 13C for compound 15, are some of the quaternary Ar carbons not showing up? This should be commented on/stated.	Added "(missing Ar signals attributed to symmetry)."	160
Other analogues seem to have more C's listed that could be Ar or carbonyl. There are several 1H NMR signals listed as a 'd' but with two J values, this should be a 'dd' I assume.	Two J values were typos. Amended.	Experimental
Putting some copies of 1H NMR spectra in an Appendix to show a symmetrical 't' in that Ar region would also alleviate doubts that it is a true triplet.	Added.	Appendix
The peptide synthesis in the experimental chapter does not have any indication of scale, it is all very generic. What mmol scale was this done on? Coupling to the fluorophore page 174, information not given about the key sensor, in equiv. or mmol.	Added some scale and equivalents for indication.	Experimental
In 1.3.3 when discussing fluorescent properties of the various scaffolds "...and DNA interchelating anticancer	Removed the DNA interchelating example.	6

agents..” – is this a biological activity in its own right regardless of fluorescence? This is an important distinction to make. Fluorescent sensor versus biologically active ligand, or dual.		
1.3.4. coumarins are based on the naturally occurring non-fluorescent parent compound – which is what compound? Specify this, this is vague.	Moved figure caption reference to first sentence to specify that the parent compound is the one shown in the figure.	6
It would be good to show one example of a commonly used fluorophore in each class along with the ‘core’ structure, esp for figure 1.8.	References to papers that utilise these probes in biology have already been given in text. No changes were made.	5-7
How were the most commonly used fluorophores from 1.3.1 onwards selected, and why did BODIPY not make this list?	Added a segment on BODIPY.	5-7
1.7 conventional methods to study peptide and protein, isn't MS ie MS/MS a common way to study peptides? From what follows in the Thesis the answer is yes, but it should not be missing/not mentioned in this section. The Thesis Aims 1.9, aim one involves multimodal MS imaging, yet the topic of MS is lacking from the introduction..... it should be stated this is covered in the intro to Chapter 2.	Added section on mass spectrometry: Mass spectrometry is a powerful technique that is frequently used to analyse peptides and proteins. Protein analysis using mass spectrometry can be performed using bottom-up or top-down methods. In the bottom-up approach, proteins are enzymatically cleaved into peptides before analysis, providing information about peptide masses and fragmentation patterns. This approach is popular due to its simplicity and ease of peptide analysis, but results in loss of information such as protein modifications or splice variants. On the other hand, the top-down approach involves the direct analysis of intact proteins, allowing for a more comprehensive characterisation of the protein. However, this approach faces challenges such as low sensitivity and difficulties in analysing larger proteins. In both scenarios, the spectrum generates peptide or protein fragments that need to be matched to databases for identification, which requires prior knowledge of peptides and proteins for accurate analysis. Moreover, the technique is destructive, leading to the loss of information regarding the native function and location of the peptide or protein.	19
Toxicity comments 21-23 about the arsenic containing probes – where does the arsenic eventually end up?	These comments are about arsenic toxicity in the context of the mechanism (binding to thiols) which is how the arsenic probes function. I have also mentioned “EDT ligands remain complexed to any unreacted probe, ensuring negligible arsenic toxicity when using the FIAsh system.” Therefore, no changes were made.	21
Fig 1.12, missing the structure of ReAsH-EDT2 that is listed in the figure description.	Removed ReAsH-EDT ₂ from caption. This compound is later presented in Figure 1.15	22
It is not good practice to represent things in diff ways esp. so close, fig 1.12 and 1.13 have the FIAsh-EDT2 written in CO ₂ H then COOH and the phenol and carbonyl shown on different sides of the page.	Changed Figure 1.12 to COOH	22
Fig 1.15– it is ReAsH-EDT2, so what is “ReAsH” listed in the figure description?	Changed to ReAsH-EDT ₂	25
Also ‘+’ typo - should this be Ca ²⁺ and not Ca ² + meaning and sensor?	Amended	25
Page 38, criteria 1, needs rewording / English fixed.	This sounds ok to me. No changes made.	39
Criteria 1 – a reporter cleaves selectively, criteria 2 the compound is released, is the reporter also the compound? Fix language to be consistent, this is confusing to follow. Then point 4 “reported molecule” is used (?)	Criteria 2 – changed compound to reporter for consistency	39
Page 39, “designed by Stevens et al. (unpublished work)” who is Stevens? If this an inhouse group in Sydney or another research group somewhere? How is this known if it is unpublished? A little more info is required here for this to make sense. At least state Stevens full name and affiliation.	Changed to “Katherine Stevens from the Puakala lab at the University of Adelaide”	40
Page 39 “specifically an N-hydroxysuccinimide that is specific to primary amines.” It is NHS esters not just “N-hydroxysuccinimide”, which does not react at all with amines.	Changed N-hydroxysuccinimide to succinimidyl ester	40
Page 45 Scheme 2.3, and throughout, it is usual to number all structures in a scheme, to allow for easy and specific	Noted. Amendments will be made to any future publications.	44

referral in the text. I am not suggesting this is changed throughout as this is a big job, but there are key points, like when discussing scheme 2.3, that a number is really required. "reaction results in formation of the cyclised product" without compound code labelling in Scheme 2.3 this is vague, there are cyclic products all through the scheme.		
Is there a reason that alkylation with e.g. 6-azido-1-bromohexane was not considered? Is it because of the reflux/azide safety? It would add to the discussion to say why this was not considered an alternative approach.	This route didn't occur to me at the time.	47
Page 46, is linker flexibility a concern for this application? As switching to an amide has shape/flexibility differences, which is not touched on.	No. Criteria has mentioned what is important already.	48
Scheme 2.7, compound 7 should be coded as such.	Amended	49
Figure 2.6, to collect this data what sort of time did was the sample in water for? Since in the 'click' reaction section earlier it says water was avoided because of NHS-ester stability. If this would not influence the fig 2.6 measurements then state this (NHS ester stability) was not of concern.	Added "Any cleavage of the NHS ester in water should not affect fluorescent readings as it is not associated with the fluorescent core."	49
One page 67 an NHS ester is submitted to 4 h resin cleavage with water – is it stable to these conditions?	Added "The short exposure time to water during this reaction is assumed to have minimal effect on the NHS ester. However, this should be assessed more thoroughly in future syntheses."	68
Page 64 "Naphthalimide cores also contain two positions available for functionalisation." But there are literature examples with 3 positions substituted?	Added "at least two" and a reference to Kate Leslie's paper about naphthalimide substitution.	64
Scheme 2.10, reaction to make compound 16, how did this work with an unprotected amine? I would have thought many side products would form and it is surprising any product did. This should be discussed in the text.	Added "The lower yields of this reaction may have been attributed to the unprotected amine forming side-products."	66
Pages 70-75 read like a review without figures, although well written in terms of English, it is very hard to follow and really get much out of the examples without any figures/chemdraw.	Amended as per previous examiners comment.	Section 3.1
Fig 3.3 and in text, is NpFR1 and NpFR1-Tag the same chemical structure? It is confusing with both formats used, if one includes an extra "Tag" or something	This was already outlined in text that the structures of NpFR1 and FCR1 with the HaloTag ligand would be named NpFR1-Tag and FCR1-Tag. No changes were made.	80/81
. On page 83 near the end FCR1 is introduced again like it appears for the first time with its full name then abbreviation – is this something else different to FCR1 above?	Full name and abbreviation moved to first mention on page 80.	83
Is there any literature on quantum chemical methods for FCR1?	No. This is a novel compound.	Section 3.6
Fig 3.14, structure of a potential FCR1 derivative ? The compound coding (or lack of) is confusing.	Labelled as FCR1B	97
Fig 3.16 title, why is NpFR1 novel, in the sense it was reported in 2014 by others (?)	Novelty in terms of the scaffold. Removed "novel" from first sentence and kept novel when referring to the scaffold: In 2014, Yeow <i>et al.</i> designed and developed a redox sensor based on the reversible redox response of the flavin moiety. The fluorescent scaffold was based on the fusion of a 1,8-naphthalimide and a flavin to form the novel derivative, NpFR1	99
Page 110, Chapter 3 conclusions, it is not clear what the future work/recommendations are, so FCR1 already contains almost the optimal distance between FRET partners so distance optimisation is not recommended/necessary? I.e. there are not any derivatives of FCR1 proposed to be better?	Conclusions make it clear that FCR1 is optimal for use in future biological studies with the HaloTag system, which addresses the aim. Future work is mentioned in final conclusion about use in biological studies with HaloTag. No changes were made.	113-114
Abstract, "Fluorescent properties of FCR1.... "	Amended	i
Minor comment/style comment, compounds don't have feelings and are not people, so the language, for example talking about fluorophores allowing 'them' to respond to 'their' environment, could be worded to be more scientific throughout. Use of 'they' to describe compounds.	This has been noted for future publications.	Throughout
Page 23, it is not G-coupled protein receptors, but it is G protein coupled receptor	Amended	24
Page 36, "as demonstration by"	Amended	37
Page 44, carboxylic (with two l's)	Amended	45
Page 70, "various methods can used to target..."	Changed to "can be used"	72
Page 95, dexhibits	Changed to "exhibits"	99
Minor English comment, sentences like "No reaction	Noted.	Throughout

occurred under these conditions" should be "A reaction did not occur.." because 'nothing' can't actually 'occur'. It is like saying no peak WAS observed, you can't 'see' no peak.		
---	--	--

Can't attach Supervisor Confirmation of Corrections.PNG, please review file online.