# Development of High-Affinity Peptide-Drug Conjugates for Investigating the TRIM24 PHD and Bromodomain




**Michael Arran Platt**

Worcester College, University of Oxford

Thesis submitted for the degree

*Doctor of Philosophy in Chemical Biology*

Trinity Term 2024

Under the supervision of

Prof. Stuart Conway

**Department of Chemistry**

# DECLARATION OF AUTHORSHIP

**Name: Michael Arran Platt** **Candidate number: 1009207**

**College: Worcester College** **Supervisor: Prof. Stuart Conway**

**Title of thesis: Development of High-Affinity Peptide-Drug Conjugates for Investigating the TRIM24 PHD and Bromodomain**

*Please tick to confirm the following:*

| | |
|---|---|
| I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*The University Student Handbook* Section 7.7) | [✓] |
| I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism | [✓] |
| The thesis I am submitting is entirely my own work except where otherwise indicated. | [✓] |
| It has not been submitted, either partially or in full, either for this Honour School or qualification or for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution. | [✓] |
| I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs. | [✓] |
| I have clearly indicated the presence of all paraphrased material with appropriate references. | [✓] |
| I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor. | [✓] |
| I have not copied from the work of any other candidate. | [✓] |
| I have not used the services of any agency providing specimen, model or ghost-written work, or unauthorised artificial intelligence in the preparation of this thesis. See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere. Sections 3 and 4 of Proctors' Regulations 1 of 2003 dealing with unauthorised use of AI, also apply. | [✓] |
| I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted. | [✓] |
| I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism. | [✓] |

Candidate's signature:  Date: 10/10/24

# Table of Contents


**Abstract** v

**Acknowledgements** vii

**Abbreviations** ix

**Chapter 1 Introduction** 1

**1.1 Epigenetics and the Histone Code** 1

1.1.1 Epigenetics 1

1.1.2 Chromatin Structure 1

1.1.3 The Histone Code 2

**1.2 Epigenetic Reader Domains** 3

1.2.1 Bromodomains (BRDs) 3

1.2.2 Plant Homeodomain (PHD) Fingers 5

**1.3 Tandem Reader Domains and Combinatorial Readout** 7

**1.4 Tripartite Motif-containing Proteins (TRIMs)** 9

**1.5 TRIM24** 10

1.5.1 Structure 10

1.5.2 Functions 13

1.5.3 TRIM24 and Cancer 13

1.5.4 Ligands and Degraders 14

**1.6 Peptide-Drug Conjugates (PDCs)** 15

**1.7 Binding Assays** 17

1.7.1 AlphaScreen® 17

1.7.2 Fluorescence Anisotropy 18

1.7.3 Surface Plasmon Resonance (SPR) 20

1.7.4 BROMO*scan* 21

**1.7 Hypothesis and Project Aims** 22

**Chapter 2 Peptidomimetic TRIM24 PHD Ligands** 24

**2.1 Introduction and Aims** 24

**2.2 H3K9Me$_3$-mimicking Peptides** 25

| | | |
|---|---|---|
| 2.2.1 | Cysteine-selective Alkylation Reactions | 25 |
| 2.2.2 | Validation of the Cysteine Alkylation Strategy | 25 |
| 2.2.3 | $KMe_3$ Mimic Design | 29 |
| 2.2.4 | Synthesis of H3K9$Me_3$-mimicking Peptides | 30 |
| 2.2.5 | Evaluation of H3K9$Me_3$-mimicking Peptides | 36 |
| **2.3** | **Acyl Hydrazone Dynamic Ligation Screen** | **39** |
| 2.3.1 | Ligand-directed Dynamic Fragment Ligation | 39 |
| 2.3.2 | Anchor Peptide Design | 40 |
| 2.3.3 | Tracer Design | 41 |
| 2.3.4 | Anchor Peptide Competition Assay | 44 |
| 2.3.5 | Dynamic Ligation Screen | 45 |
| **2.4** | **Summary** | **48** |
| **Chapter 3** | **Synthesis of TRIM24 Peptide-Drug Conjugates** | **49** |
| **3.1** | **Introduction and Aims** | **49** |
| **3.2** | **TRIM24 BRD Ligand Synthesis** | **50** |
| **3.3** | **Unnatural Amino Acid Synthesis** | **51** |
| 3.3.1 | Pyrrolizidine | 51 |
| 3.3.2 | Isoindoline | 55 |
| **3.4** | **Histone-mimicking Peptide Synthesis** | **60** |
| **3.5** | **Assembly of Peptide-Drug Conjugates** | **61** |
| **3.6** | **Control Compound Synthesis** | **62** |
| **3.7** | **Summary** | **64** |
| **Chapter 4** | **Biological Evaluation of Peptide-Drug Conjugates** | **65** |
| **4.1** | **Introduction and Aims** | **65** |
| **4.2** | **Biophysical Characterisation** | **66** |
| 4.2.1 | Investigation of TRIM24 PHD and BRD Inhibition | 66 |
| 4.2.2 | Determination of Binding Affinities and Kinetics | 69 |
| 4.2.3 | Elucidation of Binding Modes | 75 |
| 4.2.4 | Evaluation of Selectivity | 83 |
| **4.3** | **Evaluation of Phenotypic Effects** | **84** |
| **4.4** | **Summary** | **91** |

**Chapter 5 Conclusions and Future Work** **93**

**5.1 Conclusions** **93**

**5.2 Future Work** **95**

**Chapter 6 Experimental** **97**

**6.1 General Chemistry Experimental** **97**

**6.2 General Chemistry Procedures** **101**

**6.3 Synthetic Procedures and Compound Characterisation** **103**

**6.4 General Biology Experimental** **169**

**6.5 Histone-mimicking Peptide Alkylations** **173**

6.5.1 Alkylation Buffer 173

6.5.2 Test Reaction 174

6.5.3 Peptide Alkylation Reactions 174

**6.6 Protein Production** **178**

6.6.1 Constructs, Plasmids, and Cell Lines 178

6.6.2 Bacterial Transformation 178

6.6.3 Protein Expression 179

6.6.4 $His_6$-TRIM24 Purification 179

6.6.5 $His_6$-SUMO-TRIM24 Purification and Tag Cleavage 181

**6.7 Protein Crystallography and X-ray Diffraction** **183**

6.7.1 TRIM24 Crystallisation in the *apo* State 183

6.7.2 TRIM24 Co-crystallisation 184

6.7.3 TRIM24 Crystal Soaking 184

6.7.4 Diffraction Data Collection and Analysis 185

**6.8 AlphaScreen®** **185**

6.8.1 General AlphaScreen® Experimental 185

6.8.2 Signal-response Assay 186

6.8.3 Dose-response Assay 186

6.8.4 Ternary Complex Detection Assay 187

6.8.5 TruHits Assay 187

**6.9 Fluorescence Anisotropy** **188**

6.9.1 General Fluorescence Anisotropy Experimental 188

6.9.2 Stability Tests 190

| | | |
|---|---|---|
| 6.9.3 | Tracer Titration Assay | 190 |
| 6.9.4 | Direct Titration Assay | 190 |
| 6.9.5 | Competition Assay | 190 |
| 6.9.6 | Acyl Hydrazone Dynamic Ligation Screen | 191 |
| **6.10** | **BROMO*scan*** | **192** |
| **6.11** | **Surface Plasmon Resonance (SPR)** | **193** |
| 6.11.1 | General SPR Experimental | 193 |
| 6.11.2 | Binding Saturation Assay (Equilibrium Analysis) | 194 |
| 6.11.3 | Multi-cycle Kinetics Assay (Kinetic Analysis) | 194 |
| **6.12** | **Cell Permeability Assay** | **195** |
| **6.13** | ***in vitro* Cellular Assays** | **196** |
| 6.13.1 | Cell Lines | 197 |
| 6.13.2 | Cell Culture | 197 |
| 6.13.3 | MTT Assay | 197 |
| 6.13.4 | Cell Proliferation Assay | 198 |
| 6.13.5 | Clonogenic Assay | 198 |

**References** **199**

**Appendices** **212**

| | | |
|---|---|---|
| Appendix 1 | Protein Purification and Characterisation | 212 |
| Appendix 2 | Protein Crystallography | 222 |
| Appendix 3 | Histone-mimicking Peptide Characterisation | 227 |
| Appendix 4 | AlphaScreen® | 246 |
| Appendix 5 | Fluorescence Anisotropy | 257 |
| Appendix 6 | BROMO*scan* | 279 |
| Appendix 7 | Surface Plasmon Resonance (SPR) | 283 |
| Appendix 8 | Cell Permeability Assay | 286 |
| Appendix 9 | NMR Spectra of Novel Compounds | 292 |
| Appendix 10 | LC-MS Chromatograms for Compounds 58–60 and 67 | 339 |
| Appendix 11 | HPLC Chromatograms for Compounds 30, 58–63, and 67 | 343 |
| Appendix 12 | Image Permissions | 351 |

# Abstract

Tripartite motif-containing protein 24 (TRIM24) is an epigenetic transcriptional regulator involved in reading histone 3 (H3) methylated ($KMe_x$) and acetylated (KAc) lysine post-translational modifications (PTMs) through an adjacent plant homeodomain-bromodomain (PHD-BRD) cassette, respectively. The PHD recognises trimethylated K9 ($K9Me_3$) through a non-canonical cation-π interaction with a solvent-exposed tryptophan (W828); while the BRD reads K18/23Ac *via* a hydrogen bonding interaction with an asparagine (N980).[1,2] TRIM24 has been implicated in breast cancer development, and genetic knockdown of the protein was shown to have an antiproliferative effect in breast cancer cell lines, making it an attractive therapeutic target.[2] Currently, only two high-affinity TRIM24 BRD ligands exist in the literature, and no small-molecule TRIM24 PHD inhibitors have been reported.[3,4] The two BRD ligands are also non-selective over BRPF1 (a homologous BRD-containing protein), and importantly, have no significant cytotoxic or antiproliferative effects on breast cancer cell lines.[3,4]

While TRIM24 BRD inhibition alone has proven ineffective against breast cancer cells, simultaneous PHD and BRD inhibition remains unexplored. The aims of this project were to synthesise and evaluate a series of bivalent TRIM24 chemical tools in order to investigate the effects of inhibiting *both* the PHD and BRD on breast cancer cells.

Given the close proximity of the PHD and BRD, and lack of TRIM24 PHD ligands, it was decided that a TRIM24 BRD ligand would be linked to a H3-mimicking peptide containing a $KMe_3$ isostere, to generate bivalent peptide-drug conjugates (PDCs) (Figure 0.1a). $KMe_3$ mimics, which could interact with W828 in the TRIM24 PHD, were screened by incorporation into a K9C mutant H3 peptide using cysteine-selective alkylation reactions (Figure 0.1b).[5–7] An AlphaScreen® competition assay identified isoindoline as one of the

highest affinity and most synthetically tractable $KMe_3$ mimics, which was subsequently taken forward for incorporation into an unnatural amino acid in the PDCs (Figure 0.1c).

The PDCs showed low nanomolar AlphaScreen® $IC_{50}$ values for displacement of a TRIM24 PHD- and BRD-binding peptide (H3K9Me$_3$K18Ac), and exhibited avidity for TRIM24, which arises from simultaneous in *cis* binding to the PHD and BRD. Surface plasmon resonance (SPR) assays revealed the PDCs have picomolar affinity for TRIM24, resulting from a slow dissociation rate, consistent with a bivalent binding mode. Furthermore, by BROMO*scan* the PDCs demonstrated 13-fold selectivity for TRIM24 (PHD-BRD) over BRPF1 (BRD). PDC5 demonstrated some cell permeability, despite its peptidomimetic structure, and so was evaluated in proliferation assays. Although, no significant antiproliferative effects were observed in the $ER^+$ MCF-7 breast cancer cell line, the triple negative CAL-51 cell line appeared sensitive to simultaneous TRIM24 PHD and BRD inhibition. Future work with these high-affinity bivalent PDCs will enable investigation of TRIM24 function in more detail.

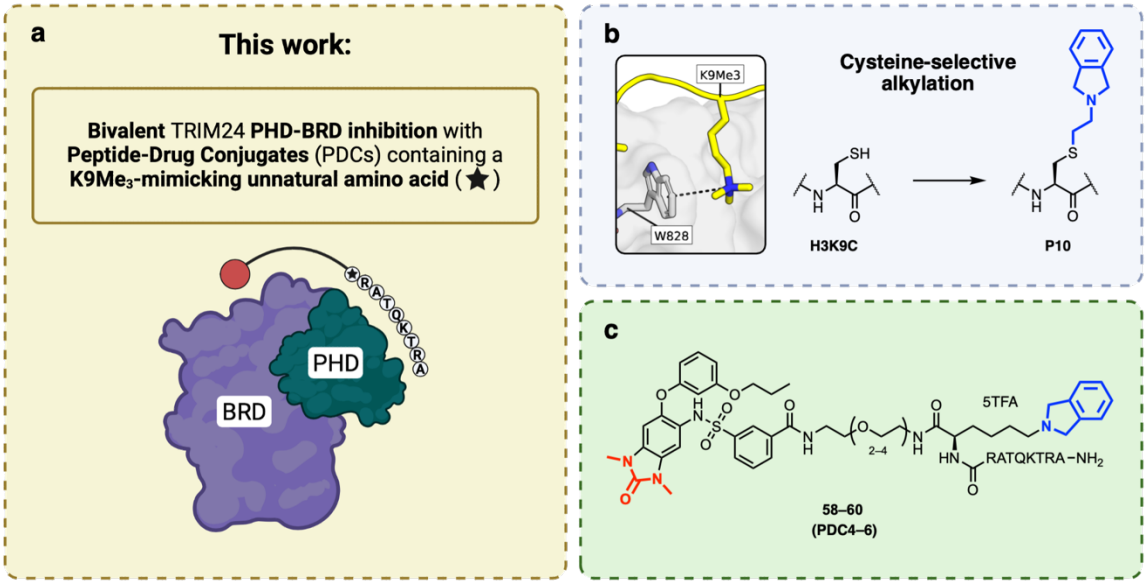



**Figure 0.1:** a) Schematic representation of the project; b) Cysteine-selective alkylation of H3K9C to identify $KMe_3$ mimics that could interact with W828; c) Structure of the PDCs showing the $KMe_3$ and KAc mimics in blue and red, respectively. Figure created with BioRender.com

# Acknowledgements

I would like to begin by thanking my supervisor, **Professor Stuart Conway**, for his encouragement and guidance throughout my DPhil. Your mentorship has been instrumental in shaping me into the chemical biologist I am today.

To all members of the **Conway group**, both past and present, thank you for creating such a welcoming and friendly atmosphere in the lab every day. It's been a pleasure to work with you all over the last few years. Thank you, **Bernie** and **Keefe** (Team TRIM), for sharing in the ups and downs (mostly downs) of working with TRIM proteins and AlphaScreen. Thank you, **Harriet** and **Oli**, for being a constant source of gossip and laughter in the lab. Our conference trip to Rome was one of the highlights of my DPhil. Thanks to **Alistair**, **Glenn** and **Marta** for all the pub quiz trips. Even though a win was hard to come by, I thoroughly enjoyed all the evenings spent together at the Cape of Good Hope. Thank you, **Antoine**, for all the fascinating scientific discussions throughout the years. Your enthusiasm and expertise in chemical biology is truly inspiring. Thank you, **Ceri**, for being such a lovely member of the lab. I'll miss the lunchtime crosswords and trips to Pieminister and Gloucester Green. Thanks to **Jenny**, **Darius**, and **Katrina** for being such friendly lab mates and for all the valuable tips and tricks you shared with me along the way. Thanks to **Patrick** for keeping everyone's spirits up with your 'pun-ishing' dad jokes. I'd also like to thank the **Part II students** over the years, for being fantastic additions to the lab. Especially **Jonny**, for being such an excellent student to supervise. A big thank you also to **Carole** from the Russell group for doing so much proof reading of my thesis. Our chats in the office undoubtedly kept me sane during the write-up!

I would like to thank all my collaborators that I've worked with over the last four years. A big thank you to **Professor Ester Hammond** and **Louise Martin** (Oncology Dept.,

University of Oxford) for carrying out the cellular assays. Thanks to **Professor Andrew Wilson** and **Yue Li** (University of Birmingham) for collaborating with the Dynamic Ligation Screen; especially Yue for synthesising the anchor peptide. Many thanks to **Eleanor Williams** and **Korvus Wang** (Centre for Medicines Discovery, University of Oxford) for providing access to the protein production facility and for troubleshooting tips on protein expression and purification. I'm very grateful to **Lizbé Koekemoer** and **Kate Kot** (Centre for Medicines Discovery, University of Oxford) for helping me to solve my first protein crystal structures. A special mention goes to Kate for noticing that a crystal appeared after 2 months, just as we were about to give up. Thank you to **Mikhail Kutuzov** (Sir William Dunn School of Pathology, University of Oxford) for training on the Biacore machine and for your expert advice about SPR. Thanks to **Vaishnavi Ravikumar** and **Marjorie Fournier** (Biochemistry Dept., University of Oxford) for carrying out the nanoLC-MS/MS experiments. I would also like to thank the **Schofield group** (Chemistry Dept., University of Oxford) for providing access to the PHERAstar plate reader, Liberty Blue peptide synthesiser, and protein production/purification equipment.

I would like to thank the **UKRI-BBSRC** and **Oxford Interdisciplinary Bioscience DTP** for studentship support (BB/T009794/1). Thanks also to **Diamond Light Source** for access to beamline I03; and for diffraction data collection and processing.

Finally, I would like to say a special thank you to all my **friends** and **family** who have been there for me throughout my studies. In particular, thank you **Liwen** for being so supportive and for encouraging me to try so many new things (I really didn't expect to be jumping out of planes with you by the end of my DPhil!). Thanks also to my **mum** and **sister**, for always being there to listen when I needed to vent my frustrations. I could not have got through this without you. I'm very lucky to have such loving and caring people around me.

# Abbreviations

| | |
|---|---|
| σ | Standard deviation |
| Ahx | 6-Aminohexanoic acid |
| AR | Androgen Receptor |
| BET | Bromodomain and extra-terminal domain |
| Boc | *tert*-Butyloxycarbonyl |
| BPTF | Bromodomain and PHD finger-containing transcription factor |
| BRD | Bromodomain |
| BRD2(2) | Second bromodomain of bromodomain-containing protein 2 |
| BRD4 | Bromodomain-containing protein 4 |
| BRPF1 | Bromodomain and plant homeodomain finger-containing protein 1 |
| BSA | Bovine serum albumin |
| BTEAC | Benzyltriethylammonium chloride |
| Cbz | Benzyloxycarbonyl |
| CDI | 1,1'-Carbonyldiimidazole |
| CHAPS | 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate |
| CREBBP | CREB-binding protein |
| CV | Column volume |
| DABCO | 1,4-Diazabicyclo[2.2.2]octane |
| DIBAL-H | Diisobutylaluminium hydride |
| DIC | *N*,*N*-Diisopropylcarbodiimide |
| DIPEA | *N*,*N*-Diisopropylethylamine |
| DLS | Diamond light source |
| DMAP | 4-Dimethylaminopyridine |
| DMEM | Dulbecco's modified Eagle's medium |
| DMF | *N*,*N*-Dimethylformamide |

| | |
|---|---|
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| DPF3b | Double PHD fingers 3b |
| DTT | Dithiothreitol |
| *E. coli* | *Escherichia coli* |
| EDC | 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide |
| EDTA | Ethylenediaminetetraacetic acid |
| EI | Electron ionisation |
| ER | Oestrogen receptor |
| ESI | Electrospray ionisation |
| FAC | Final assay concentration |
| FC | Flow cell |
| FGI | Functional group interconversion |
| FITC | Fluorescein isothiocyanate |
| Fmoc | Fluorenylmethyloxycarbonyl |
| GFP | Green fluorescent protein |
| GKAP | Guanylate kinase-associated protein |
| GST | Glutathione S-transferase |
| H2/H3/H4 | Histone 2/3/4 |
| HAT | Histone acetyl transferase |
| HATU | Hexafluorophosphate azabenzotriazole tetramethyl uronium |
| HBS | HEPES-buffered saline |
| HDAC | Histone deacetylase |
| HEPES | 2-[4-(2-Hydroxyethyl)-1-piperazin-1-yl]ethanesulfonic acid |
| HFIP | Hexafluoro-2-propanol |
| $His_6$ | Hexa histidine-tag |
| HIV-1 | Human Immunodeficiency Virus 1 |

| | |
|---|---|
| HP1 | Heterochromatin protein 1 |
| HPBD | Heterochromatin protein family binding domain |
| HPLC | High-performance liquid chromatography |
| HRMS | High resolution mass spectrometry |
| $IC_{50}$ | Half maximal inhibitory concentration |
| IMAC | Immobilised metal affinity chromatography |
| IPA | Isopropanol |
| IPTG | Isopropyl-1-thio-β-D-galactopyranoside |
| ITC | Isothermal titration calorimetry |
| JARID1A | Jumonji/ARID domain-containing protein 1A |
| KAc | Acetylated lysine |
| KD | Catalytic kinase domain |
| $K_d$ | Equilibrium dissociation constant |
| KDM5A | Lysine-specific demethylase 5A |
| $KMe_x$ | Methylated lysine |
| LC-MS | Liquid chromatography-mass spectrometry |
| LDA | Lithium diisoproylamide |
| LRMS | Low resolution mass spectrometry |
| m.p. | Melting point |
| MALDI | Matrix-assisted laser desorption/ionisation |
| MES | 2-(*N*-Morpholino)ethanesulfonic acid |
| MTT | 3-(4,5-Dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide |
| MW | Molecular weight |
| MWCO | Molecular weight cut-off |
| NHS | *N*-Hydroxysuccinimide |
| NMR | Nuclear Magnetic Resonance |
| NTA | Nitrilotriacetic acid |

| | |
|---|---|
| *p*-TsOH | *p*-Toluenesulfonic acid |
| p53 | Tumour protein P53 |
| PAINs | Pan assay interference compounds |
| PBD | Polo box domain |
| Pbf | 2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl |
| PBS | Phosphate-buffered saline |
| PCAF | P300/CREBBP-associated factor |
| Pd/C | Palladium on carbon |
| PDB | Protein data bank |
| PDC | Peptide-drug conjugate |
| PDZ | Post-synaptic density-95/Discs large/Zonula occludens-1 domain |
| PEG | Polyethylene glycol |
| PEI | Polyethylenimine |
| PHD | Plant homeodomain |
| PHF1 | PHD finger protein 1 |
| pKaH | pKa of the conjugate acid |
| PLK1 | Polo-like kinase 1 |
| PPI | Protein-protein interaction |
| PROTAC | Proteolysis targeting chimera |
| PTM | Post-translational modification |
| Pygo2 | Pygopus homolog 2 |
| PZP | PHD-zinc knuckle-PHD domain |
| qPCR | Quantitative polymerase chain reaction |
| *r* | Anisotropy |
| RING | Really interesting new gene |
| RMS | Root mean square |
| rt | Room temperature |

| | |
|---|---|
| RU | Response units |
| SDS-PAGE | Sodium dodecyl sulfate-polyacrylamide gel electrophoresis |
| SEC | Size exclusion chromatography |
| SENP1 | Sentrin/SUMO1-specific peptidase 1 |
| SHANK1 | SH3 domain and ankyrin repeat-containing protein 1 |
| SOC | Super optimal broth with catabolite repression |
| SPPS | Solid phase peptide synthesis |
| SPR | Surface Plasmon Resonance |
| SUMO | Small ubiquitin-like modifier |
| TAF1 | TATA-box binding protein associated factor 1 |
| TBAB | Tetrabutylammonium bromide |
| TBAI | Tetrabutylammonium iodide |
| $^{t}$Bu | *tert*-Butyl |
| TEMPO | 2,2,6,6-Tetramethyl-piperidine 1-oxyl |
| TEV | Tobacco etch virus |
| TFA | Trifluoroacetic acid |
| THF | Tetrahydrofuran |
| TIF1 | Transcriptional intermediary factor 1 |
| TIPS | Triisopropyl silane |
| TLC | Thin Layer Chromatography |
| TMS | Trimethylsilyl |
| TOF | Time of flight |
| TRIM | Tripartite motif-containing protein |
| Trt | Trityl |
| Ts | Tosyl |
| TSS | TIF1 signature sequence |
| UHRF1 | Ubiquitin-like, containing PHD and RING finger domains, protein 1 |

| | |
|---|---|
| UPLC | Ultra-performance liquid chromatography |
| UV | Ultra-violet |
| WT | Wild-type |
| YT | Yeast extract tryptone |

# 1 Introduction

## 1.1 Epigenetics and the Histone Code

### 1.1.1 Epigenetics

The field of *Epigenetics* (literally meaning 'above' genetics) involves the study of "heritable changes in gene expression or phenotype that are stably maintained through cell divisions, and in some cases, across generations, without alterations to the underlying DNA sequence" (Arrowsmith *et al.* 2012).[8] Epigenetic changes are ubiquitous in nature from the metamorphosis of the caterpillar into a butterfly, to the differentiation of stem cells, and even the phenotypic divergence of aging identical twins.[9–11] The epigenome adds another layer of information beyond the DNA base sequence, providing instructions on how the genetic code should be read. Furthermore, epigenetic mechanisms enable lifestyle choices and environmental factors to influence gene expression, thus, providing a link between nature and nurture.[12]

### 1.1.2 Chromatin Structure

In the nuclei of eukaryotic cells, genomic DNA is packaged into protein-DNA complexes, called nucleosomes (Figure 1.1).[13] Each nucleosome unit consists of DNA supercoiled around an octamer of histone proteins, containing two molecules of each histone (2× H2A, 2× H2B, 2× H3, and 2× H4).[14] Nucleosomes are linked together through small segments of linker DNA to form polynucleosome fibres, called chromatin, which is further condensed through chromatin looping to form chromosomes (Figure 1.1).[13] The transcription of genes is heavily influenced by the higher-order structure of chromatin, with epigenetic mechanisms governing the switch between a condensed, transcriptionally-silent state (Heterochromatin) and an open, transcriptionally-active state (Euchromatin).[15]

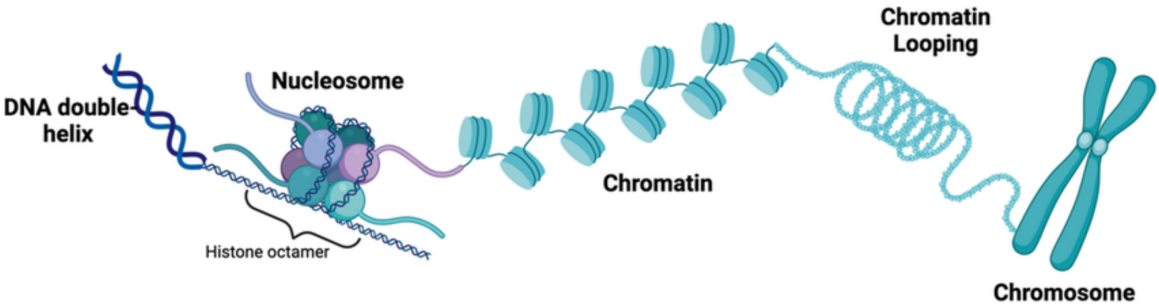



**Figure 1.1:** Illustration of the levels of DNA organisation within eukaryotic cells. DNA is wrapped around histone proteins, which are organised into an octamer unit called a nucleosome. The nucleosomes compact together to form chromatin, which condenses further to form chromosomes. Figure reprinted with permission from Clarke *et al.* 2022.[16] Copyright 2022 The Authors. Published by John Wiley & Sons Ltd on behalf of the Federation of European Biochemical Societies.

### 1.1.3 The Histone Code

One mechanism of epigenetic control on gene expression is through reversible post-translational modifications (PTMs) on histone tails, including acetylation, methylation, and phosphorylation (Figure 1.2).[17] For many years, it was believed that histone PTMs directly affected chromatin structure, and DNA accessibility, by modulating the electrostatic interactions between nucleosomes.[18] However, the revelation by Dhalluin *et al.* of proteins binding to acetylated lysines (KAc), suggested that something more complex was occurring.[19,20] This led to the development of a new theory – the 'histone code' hypothesis.

It is now widely accepted that the combinations of PTMs on histone tails form a 'histone code', which acts as a signal governing the downstream expression or repression of genes in response to environmental factors.[20] The code is dynamically regulated by 'writer' and 'eraser' proteins, which add and remove PTMs, respectively (Figure 1.2).[21] 'Reader' proteins (Figure 1.2) can bind to the PTMs and recruit chromatin-remodelling/transcription factors to alter the structure of chromatin and ultimately influence gene transcription.[21] Aberrant expression and mutation of epigenetic proteins have been associated with the upregulation of oncogenes, making them potential therapeutic targets.[22]

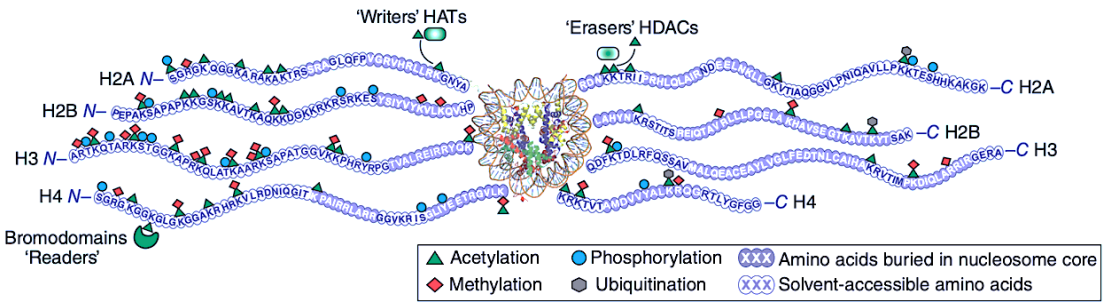



**Figure 1.2:** Schematic representation of a nucleosome showing the known PTMs on histone tails, and the writer, reader, and eraser proteins involved in adding (HATs), recognising (Bromodomains), and removing (HDACs) lysine acetylation marks, respectively. Figure adapted and reprinted with permission from Jennings *et al.* 2014.[23] Copyright 2014 Future Science Group.

## 1.2 Epigenetic Reader Domains

### 1.2.1 Bromodomains (BRDs)

Bromodomains (BRDs) are the primary epigenetic readers of acetylated lysine (KAc) residues on histone tails.[24] Since the first BRD structure was determined for PCAF in 1999, 61 unique BRDs have been identified across 46 different proteins in the human proteome.[19,25] Despite low overall sequence homology, all BRDs possess a conserved 'BRD fold', consisting of four α-helices and two loop regions, that form a deep hydrophobic cavity for the binding of KAc (Figure 1.3).[25] Within the binding site, KAc interacts with a highly conserved asparagine residue and a hydrogen bond network of tightly bound water molecules at the base of the pocket (Figure 1.3).[26]

Due to their implication in cancer, efforts have been made to develop BRD inhibitors for cancer therapy. The first reported small molecule BRD inhibitor was (+)-JQ1 (Figure 1.4), which targeted the bromodomain and extra-terminal domain (BET) protein family.[27] Since then, this ligand has been used to develop proteolysis targeting chimeras (PROTACs), which selectively degrade BRD4 (Figure 1.4).[28] Although early efforts primarily targeted the BET BRDs, a growing number of inhibitors are now being reported for proteins outside the BET family including BRPF1, PCAF, and CREBBP (Figure 1.4).[29,30]

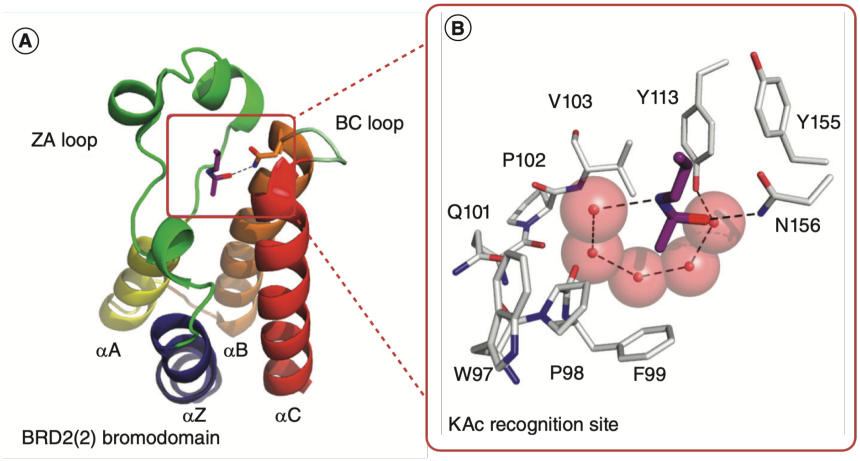



**Figure 1.3:** A) X-ray crystal structure of the canonical BRD2(2) bromodomain (PDB ID: 2DVQ[31]) showing the 'bromodomain fold' formed of four α-helices [αA (yellow), αB (orange), αC (red), αZ (blue)] and two loop regions [ZA loop (green), BC loop (light green)]. B) KAc (carbon = purple) interacts with a highly conserved asparagine residue in a hydrophobic cavity, and several tightly bound water molecules (red spheres) at the base of the pocket. Figure adapted and reprinted with permission from Jennings *et al.* 2014.[23] Copyright 2014 Future Science Group.

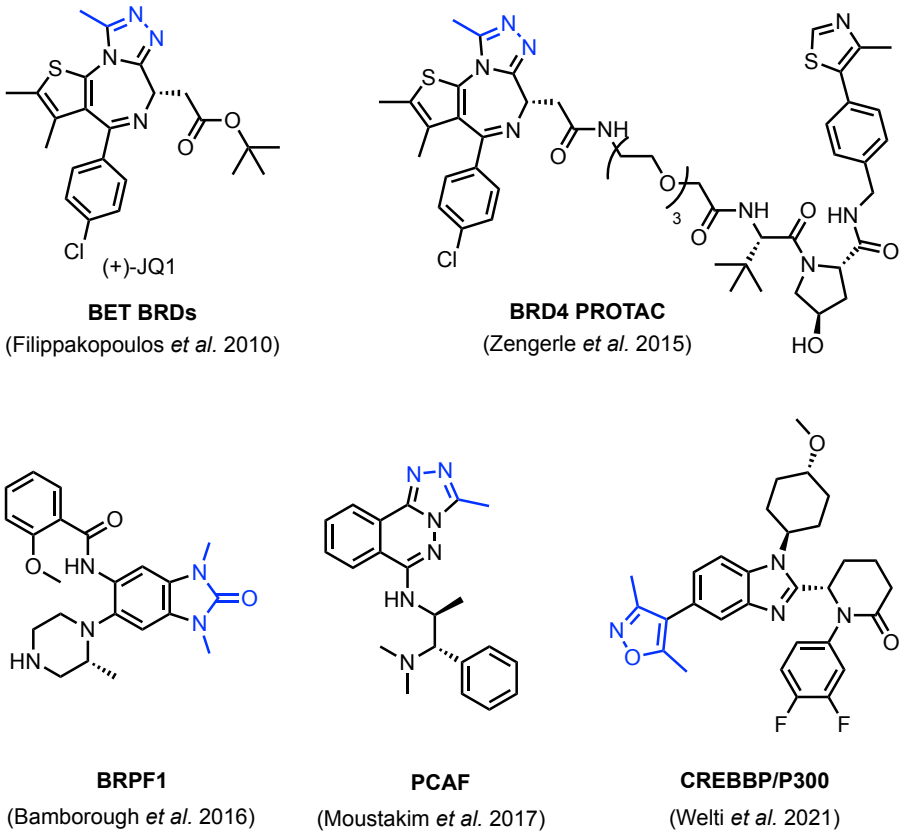



**Figure 1.4:** Structures of selected BRD inhibitors/degraders and their targets. The presumed KAc-mimicking moieties are shown in blue.[27,28,32–34]

### 1.2.2 Plant Homeodomain (PHD) Fingers

Plant homeodomain (PHD) fingers are characterised by a conserved $Cys_4HisCys_3$ zinc-binding motif which stabilises two anti-parallel β-sheets (Figure 1.5).[35] Although it has long been known that PHDs perform structural functions, their role as epigenetic methyl readers was first uncovered in 2006, with the discovery that the BPTF PHD binds to histone 3 lysine 4 trimethylated ($H3K4Me_3$).[36] Since then, numerous PHDs have been identified, which recognise the various methylation states of lysine and, to a lesser extent, arginine on the N-terminal tail of histone H3.[35,37]

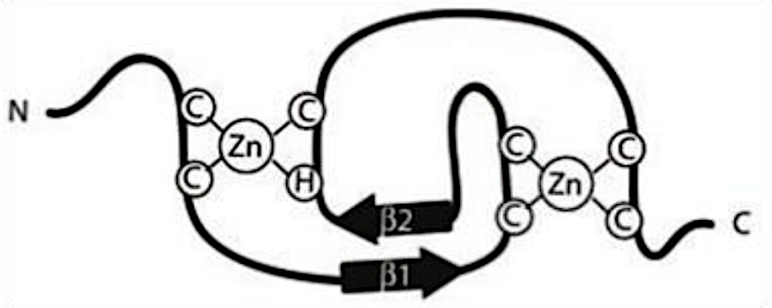



**Figure 1.5:** Cartoon representation of the characteristic structure of PHD fingers. PHDs are characterised by two zinc ions bound to a $Cys_4HisCys_3$ motif, which stabilises two antiparallel β-sheets. Figure reprinted with permission from Musselman *et al.* 2009.[35] Copyright 2009 The American Society for Pharmacology and Experimental Therapeutics.

PHD fingers usually use a surface-groove binding mode for interacting with histone tails and recognise trimethylated lysines ($KMe_3$) through cation-π and hydrophobic interactions with an aromatic cage.[38,39] This is exemplified by the BPTF PHD, which recognises $KMe_3$ using an aromatic cage formed of three tyrosines and a tryptophan residue (Figure 1.6).[36] The replacement of aromatic residues with hydrogen bond acceptors, coupled with a reduction in size of the aromatic cage, allows PHDs to select for lower lysine methylation states ($KMe_2$ and KMe).[40] Some PHDs, with a binding preference for unmethylated lysine ($KMe_0$) such as BHC80 (Figure 1.6), lack an aromatic cage altogether, instead relying on an acidic pocket of hydrogen bond acceptor residues for interaction with the N-H donors.[37]

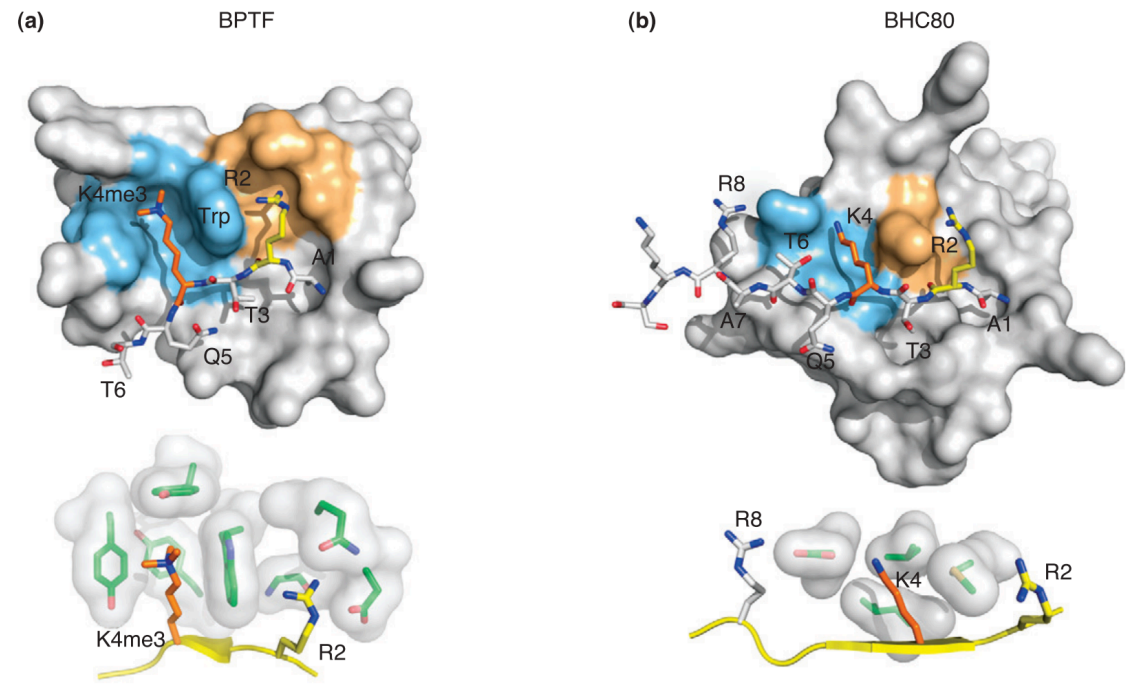



**Figure 1.6:** Protein surface and stick representations showing a) the recognition of H3K4Me$_3$ by the BPTF PHD and b) the recognition of unmethylated H3K4 by the BHC80 PHD. Figure reprinted with permission from Sanchez *et al.* 2011.[37] Copyright 2011 Elsevier Ltd.

In contrast to BRDs, PHDs have received less attention in drug discovery, presumably due to their low predicted drugability.[41] Nonetheless, a number of small molecule ligands have been developed targeting the UHRF1, Pygo2, and JARID1A PHDs (Figure 1.7).[42–44] Despite these successes, most PHDs possess a broad polar surface groove for recognition of *N*-terminal histone tails, which poses a significant challenge for small molecule inhibition.[41] Therefore, a number of studies have used peptidomimetic approaches instead involving KMe$_3$ isosteres (Figure 1.7).[45,46] Intriguingly, Zhang *et al.* used proximity-reactive cyclopeptide inhibitors to target the KDM5A-PHD3 (Figure 1.7).[46] This covalent strategy took advantage of two atypical solvent-exposed lysine residues, not normally found in PHDs with a binding preference for H3K4Me$_3$.[47] Little is known about the biological consequences of methyl lysine readout, therefore, developing chemical probes and inhibitors for PHD fingers could prove useful in discovering more about this versatile class of epigenetic reader.

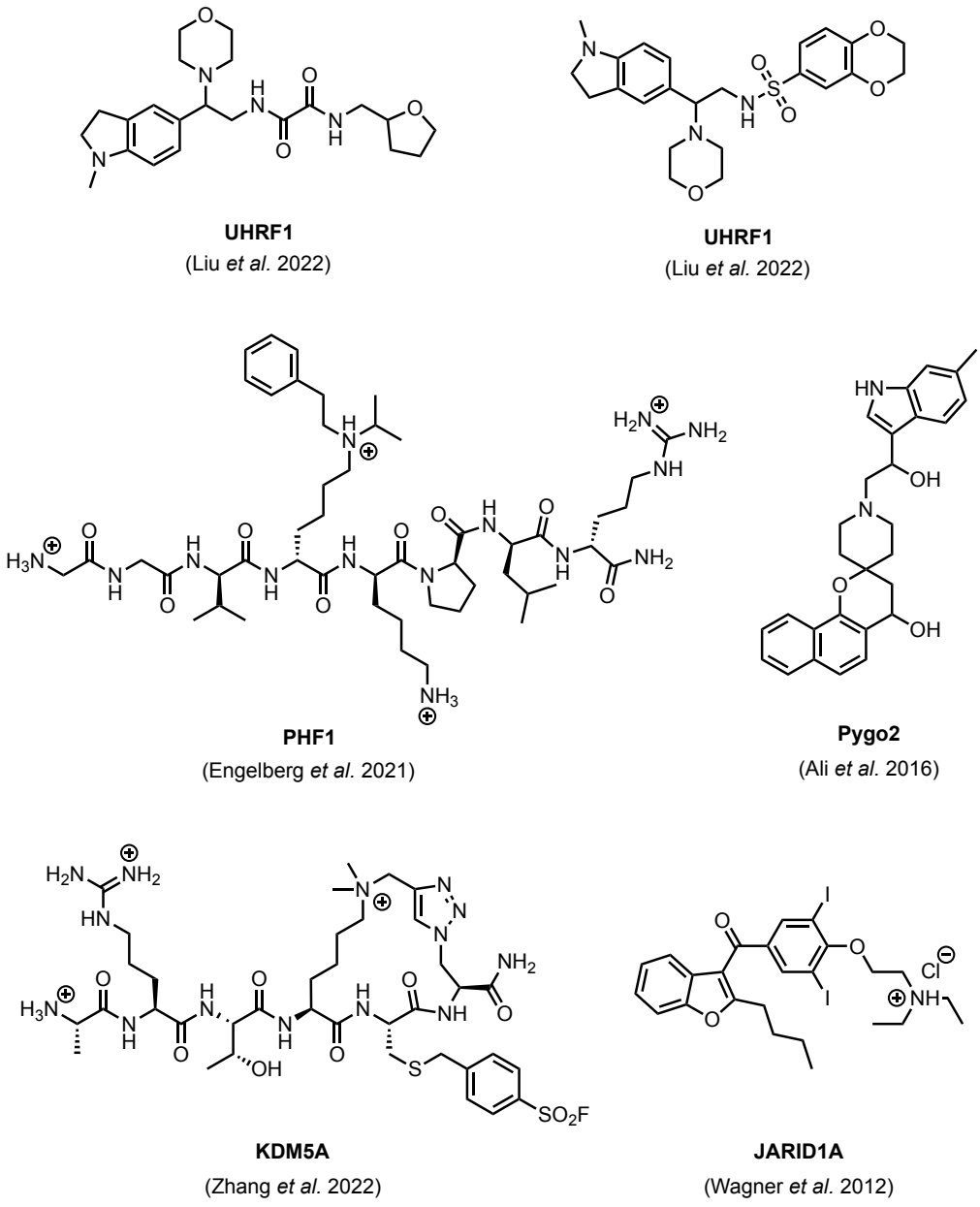



**Figure 1.7:** Structures of PHD inhibitors and their targets.[42–46]

## 1.3 Tandem Reader Domains and Combinatorial Readout

The majority of PHD fingers are found in proteins containing multiple epigenetic reader domains, which are often situated close to each other.[48] These tandem reader domains enable multivalent engagement of PTMs and combinatorial readout of the histone code through a number of different intra- and inter-nucleosomal mechanisms (Figure 1.8).[49] Notable examples include the *in cis* binding of H4(KAc)$_2$ to the tandem BRDs of TAF1 (Figure 1.9), and the *in trans* engagement of H3K4Me$_3$ and H4KAc by the BPTF

PHD-BRD (Figure 1.9).[36,50] The interplay between histone modifications does not necessarily always involve positive feedback, and can lead to different outcomes depending on the biological context. For instance, the protein HP1 normally binds to $H3K9Me_3$, facilitating heterochromatin formation; however, during mitosis, phosphorylation of H3S10 prevents the binding of histones to HP1, allowing mitotic progression.[49,51]

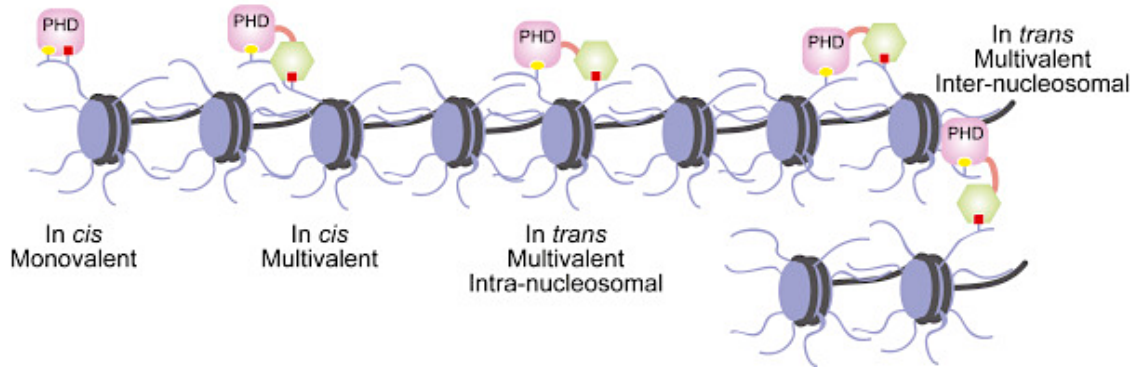



**Figure 1.8:** Schematic of the different mechanisms by which epigenetic reader domains can recognise histone PTMs. Figure reprinted with permission from Morrison *et al.* 2016.[39] Copyright 2016 Elsevier Inc.

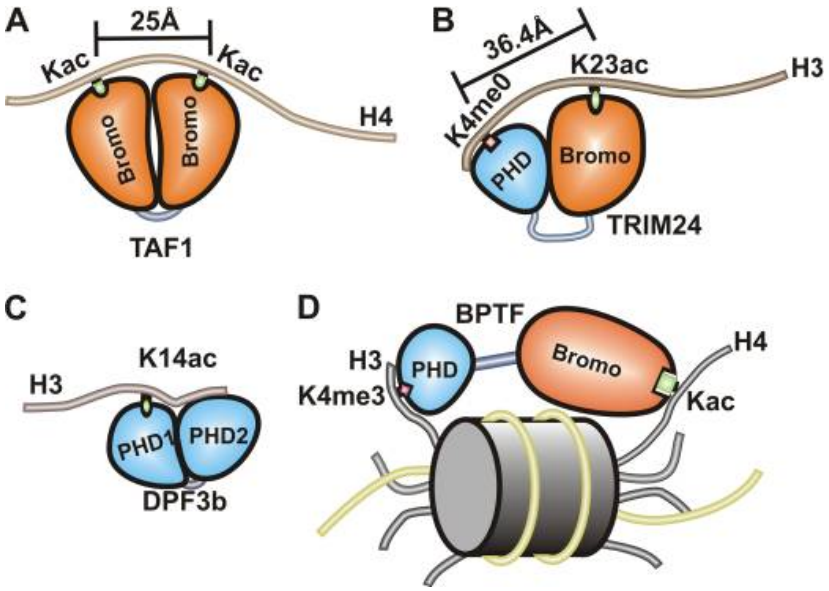



**Figure 1.9:** Examples of proteins with paired reader domains engaging in combinatorial epigenetic readout. A) The TAF1 BRD-BRD reads $H4(KAc)_2$; B) The TRIM24 PHD-BRD reads H3K4 and H3K23Ac; C) The DPF3b PHD-PHD reads *N*-terminal H3 and H3K14Ac; D) The BPTF PHD-BRD reads $H3K4Me_3$ and H4KAc. Figure reprinted with permission from Wang *et al.* 2011.[52] Copyright 2011 The American Society for Biochemistry and Molecular Biology.

In recent years several bivalent BRD inhibitors have been identified (Figure 1.10), but none have been reported for tandem reader domains containing PHDs.[53–56] It is imperative that further chemical tools are developed to probe the relationships between different reader domains, and evaluate the functional consequences of combinatorial epigenetic readout; which could help improve our understanding of the histone code.

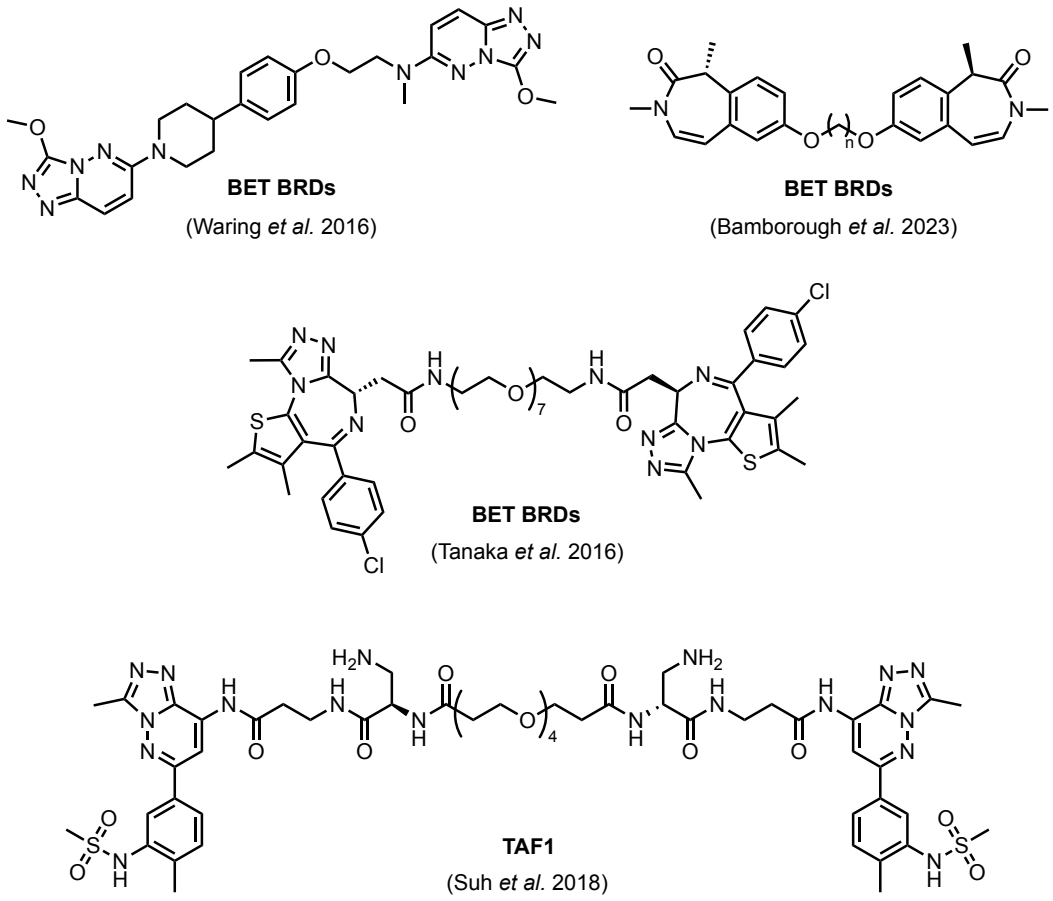



**Figure 1.10:** Structures of bivalent tandem BRD ligands and their targets.[53–56]

## 1.4 Tripartite Motif-containing Proteins (TRIMs)

Tripartite motif-containing proteins (TRIMs) comprise a superfamily of approximately 80 members in humans, that is characterised by a highly conserved 'RBCC' motif consisting of a really interesting new gene (RING) domain, one or two B-box domains, and a coiled-coil (Figure 1.11).[57] TRIM proteins are further classified into 12 subfamilies based on their variable *C*-terminal domain structure.[57] The *N*-terminal RING domain is common to almost

all TRIM proteins and confers E3 ubiquitin ligase activity.[58] While the coiled-coil domain is known to be involved in homo- and hetero-oligomerisation of the TRIM family members, the function of the B-box domains is poorly understood, with studies suggesting their involvement in higher-order complex formation.[59,60] TRIM proteins have numerous functions including intracellular signalling, transcriptional regulation, and cell proliferation; and are particularly involved in the innate immune response to viral infections such as HIV-1 (Figure 1.11).[59,61] Given their involvement in transcriptional regulation and cell proliferation, it is unsurprising that TRIM proteins have been implicated in several diseases such as multiple sclerosis, cardiovascular disease, and cancer (Figure 1.11).[62,63]

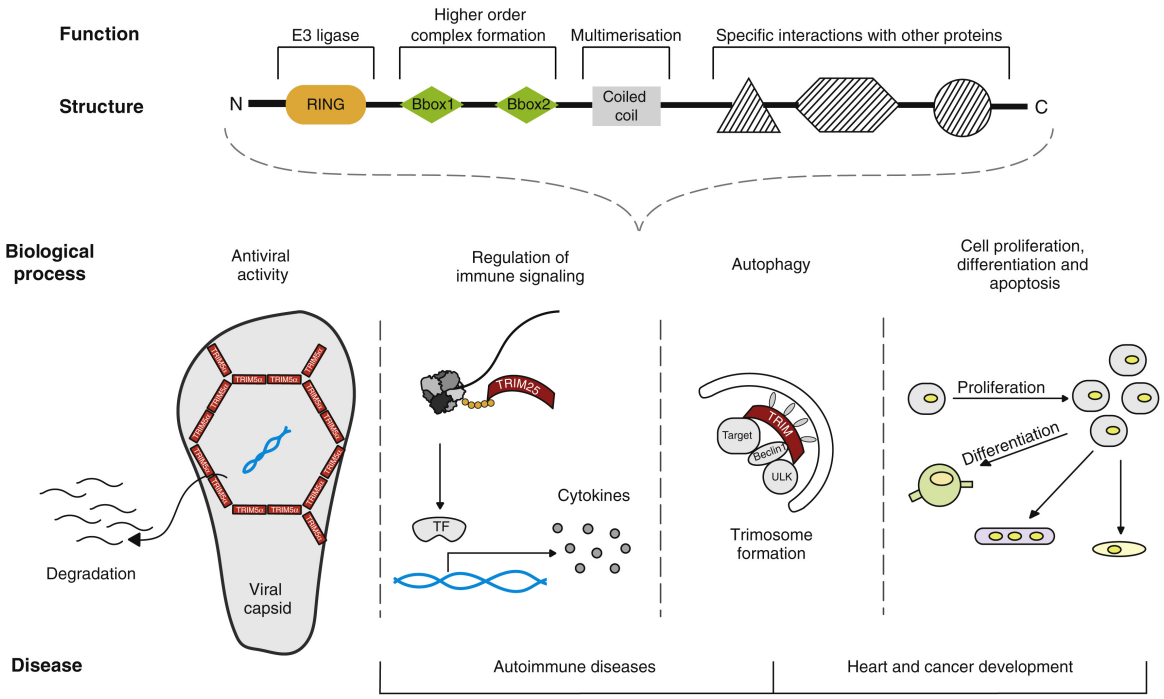



**Figure 1.11:** TRIM protein domain structure, function, and associated disease. Figure reprinted with permission from Vunjak *et al.* 2019.[59] Copyright 2019 Elsevier Ltd.

## 1.5 TRIM24

### 1.5.1 Structure

TRIM24 belongs to the C-VI subfamily, along with TRIM28 and TRIM33, which all contain a *C*-terminal tandem PHD-BRD in addition to the 'RBCC' hallmark of TRIM

proteins (Figure 1.12).[64] TRIM66 also contains a PHD and BRD but lacks the RING domain of the other proteins (Figure 1.12).[65] TRIM24 (TIF1α), TRIM28 (TIF1β), TRIM33 (TIF1γ), and TRIM66 (TIF1δ) can alternatively be classified into the transcriptional intermediary factor 1 (TIF1) protein family (Figure 1.12), which is defined by a 25 amino acid TIF1 signature sequence (TSS).[66]

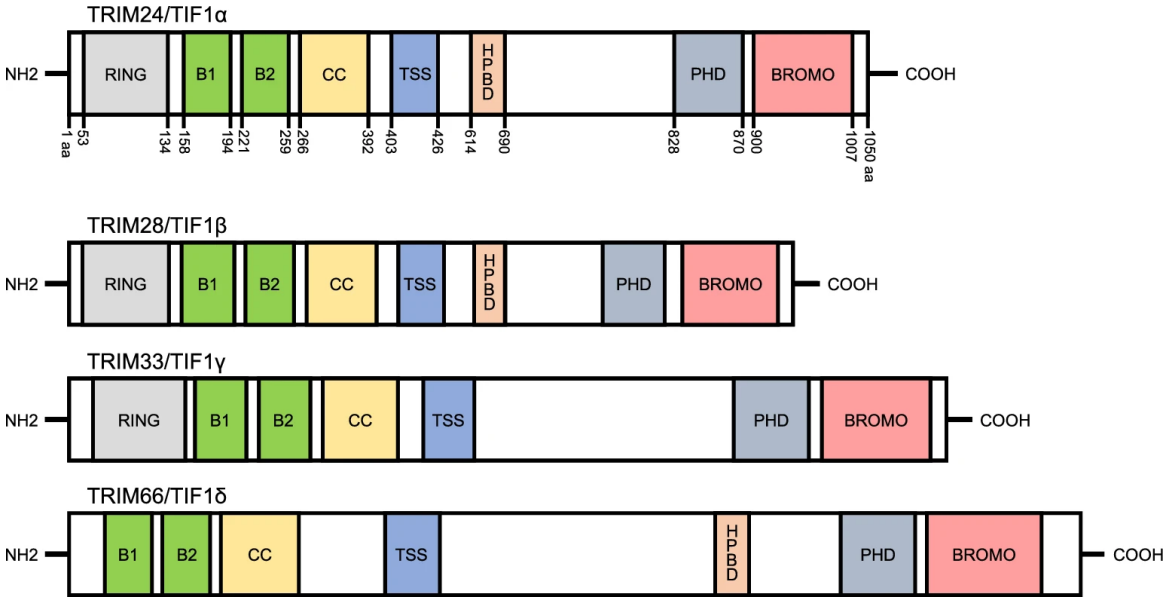



**Figure 1.12:** Schematic representation of the domains for TRIM proteins in the TIF1 subfamily (RING, RING domain; B1 and B2, B-box domains; CC, coiled-coil domain; TSS, TIF1 signature sequence; HPBD, heterochromatin protein family binding domain; PHD, plant homeodomain finger; BROMO, bromodomain). Figure adapted and reprinted with permission from Yao *et al.* 2024.[66] Copyright BioMed Central Ltd.

The tandem PHD-BRD in TRIM24 is a chromatin-binding module that reads epigenetic marks on histone H3 tails. The TRIM24 BRD possesses the classical structure consisting of four alpha helices linked by 2 loop regions (Figure 1.13), while the PHD contains the usual $Cys_4HisCys_3$ zinc-binding motif and two anti-parallel β-sheets (Figure 1.13).[2] The PHD recognises H3R2K4K9$Me_3$ using a broad polar surface groove (Figure 1.14a); whereas the BRD binds to H3K18/K23Ac through a conserved asparagine (N980) (Figure 1.14b).[1,2] As with other K4$Me_0$-binding PHDs, K4 methylation inhibits histone binding.[2]

However, unlike most PHDs that bind $KMe_3$, TRIM24 lacks an aromatic cage and instead relies on a cation-π interaction with a solvent-exposed tryptophan residue (W828) (Figure 1.14a).[1] A study by Tsai *et al.* also showed that the close proximity of the PHD and BRD (25.5 Å) allows combinatorial readout of unmodified K4 and K23Ac on histone H3.[2]

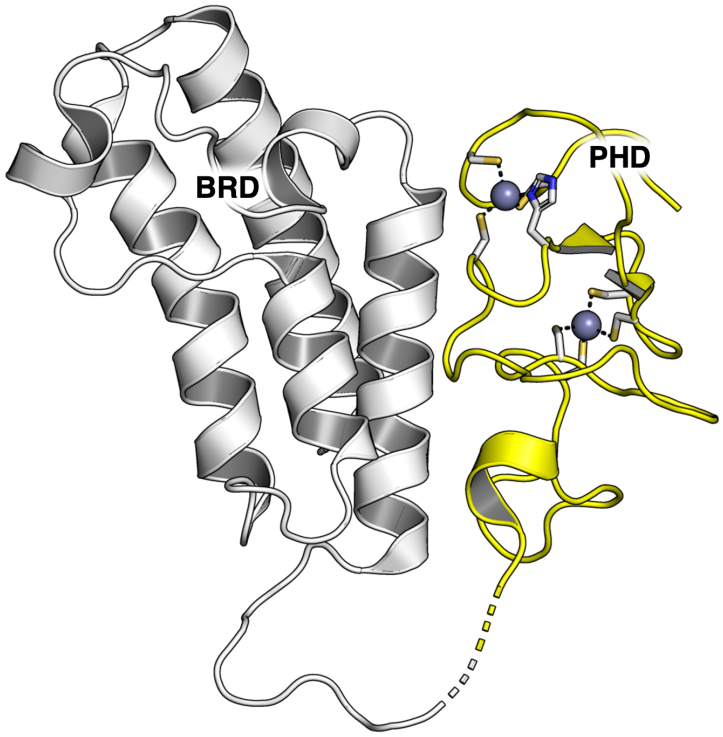



**Figure 1.13:** New X-ray crystal structure of the TRIM24 BRD (white) and PHD (yellow) - PDB ID: 9GD5. The $Cys_4HisCys_3$ motif is shown in the PHD. Grey spheres represent $Zn^{2+}$ ions.

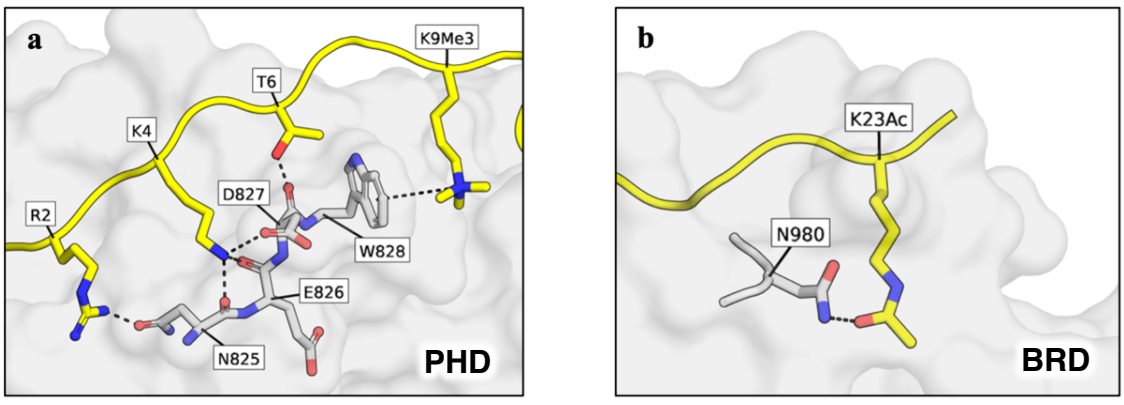



**Figure 1.14:** New *apo* crystal structure of TRIM24 PHD-BRD (1.7 Å, PDB ID: 9GD5; surface and carbon = grey) overlaid with the peptides (carbon = yellow) from the crystal structures of a) TRIM33α (which has an identical PHD domain) in complex with $H3_{(1–22)}K9Me_3K14AcK18Ac$ (PDB ID: 3U5O[67]) or b) TRIM24 in complex with $H3_{(13–32)}K23Ac$ (PDB ID: 3O34[2]).

### 1.5.2 Functions

TRIM24 is known to act as a transcriptional coregulator of numerous transcription factors including the p53 tumour suppressor.[68] For instance, the TRIM24 RING domain can function as an E3 ligase, downregulating p53 by targeting it for degradation.[68,69] Complementary to this, the TRIM24 PHD can also restrict p53-mediated chromatin remodelling and subsequent gene transcription, by localising to p53-associated chromatin sites near unmethylated H3K4.[70] TRIM24 has also been identified as a ligand-dependent coregulator of nuclear receptors, interacting with the retinoic acid receptor, androgen receptor (AR), and oestrogen receptor (ER) through an 'LxxLL' motif.[71–74] In a separate study, a unique role was discovered for the tandem PHD-BRD in promoting the SUMOylation of TRIM24, which influences cell adhesion.[75] Like other members of the TIF1 family, TRIM24 plays a role in the DNA damage repair pathway, specifically in the response to DNA double-strand breaks.[64,76] Several studies have also suggested its involvement in the immune response to viral infections such as HIV-1.[77–79]

### 1.5.3 TRIM24 and Cancer

Overexpression of TRIM24 has been implicated in the development and progression of numerous cancers including acute myeloid leukemia,[80,81] glioma,[82] colorectal cancer,[83] prostate cancer,[84,85] ovarian cancer,[86] cervical cancer,[87] and breast cancer.[2,88] In breast cancer, upregulation of TRIM24 correlates with poor prognosis (Figure 1.15b), and knockdown of the protein was shown to have an antiproliferative effect on $ER^+$ MCF-7 breast cancer cells (Figure 1.15a).[2] It is hypothesised that TRIM24 promotes the proliferation of $ER^+$ breast cancer by recruiting ER transcription factors to chromatin, resulting in the upregulation of oestrogen-dependent oncogenes.[2] This is mediated through combinatorial readout of H3K4K23Ac by the PHD-BRD cassette.[2] TRIM24 was also shown to drive proliferation in $ER^-$ breast cancer, although the mechanism is unknown.[2]

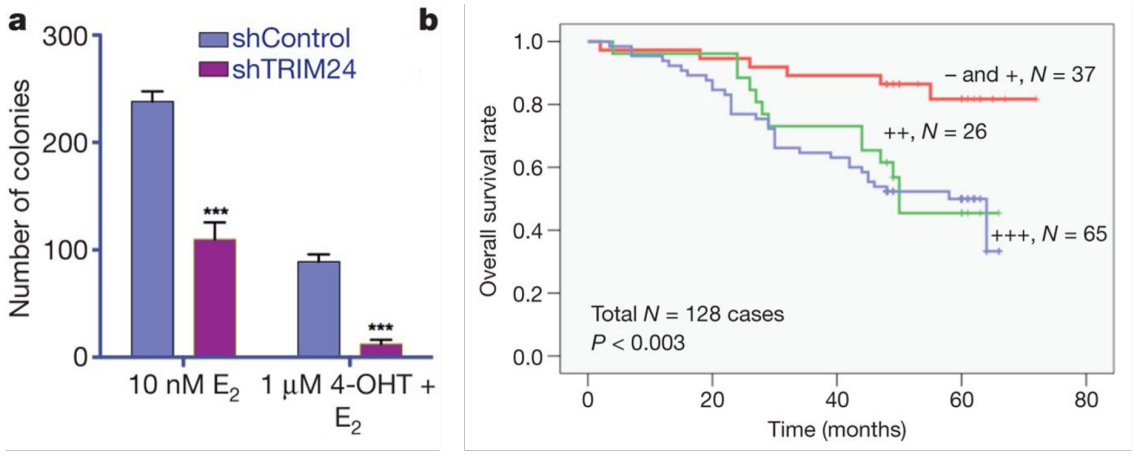
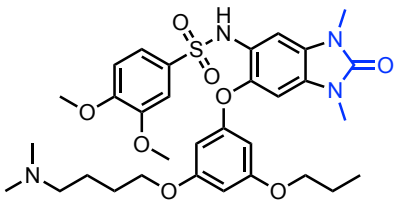


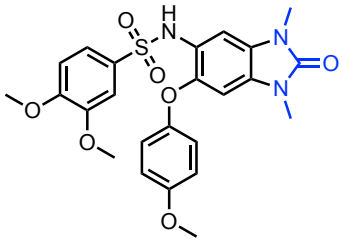

**Figure 1.15:** a) shControl and shTRIM24 MCF7 cells with $E_2$ (Estradiol, ER agonist) or $E_2$ + 4-OH-tamoxifen (4-OHT, ER antagonist), as indicated. Each bar represents the averaged results for three independent colony formation assays in triplicate plates; error bars, s.d. *$P < 0.0001$. b) The overall survival rate of 128 patients with non-metastatic disease, classified by TRIM24 expression (determined by immunostaining: strong, +++; moderate, ++; weak or slightly above background, +; none, −) plotted by the Kaplan–Meier method. Figure adapted and reprinted with permission from Tsai *et al.* 2010.[2] Copyright 2010, Springer Nature Limited.

### 1.5.4 Ligands and Degraders

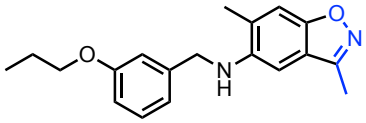
Two high-affinity TRIM24 BRD ligands, containing a benzimidazolone scaffold, have been reported in the literature (Figure 1.16).[3,4] IACS9571, developed by Palmer *et al.*, is the current highest affinity TRIM24 BRD ligand with a reported ITC $K_d$ of 31 nM.[4] Despite their high affinity, both ligands were non-selective over a closely-related BRD in BRPF1 and displayed negligible effects on breast cancer cell proliferation.[3,4,89] In recent years, a number of alternative scaffolds and KAc mimics have been explored (Figure 1.16), however, these compounds demonstrated weak affinity for the TRIM24 BRD and remained non-selective over BRPF1.[90,91] A PROTAC (dTRIM24) has also been developed by Gechijian *et al.* that degrades TRIM24 (Figure 1.16).[92] Although this PROTAC was based on the dual TRIM24-BRPF1 BRD ligands, dTRIM24 demonstrated remarkably selective degradation of TRIM24.[92] Furthermore, while IACS9571 demonstrated no phenotypic effects in acute myeloid leukemia cell lines, dTRIM24 caused antiproliferative effects.



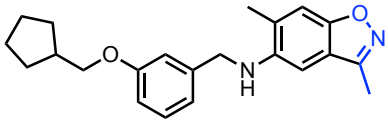
TRIM24 $K_d$ 31 nM
BRPF1 $K_d$ 14 nM
(Palmer *et al.* 2015)



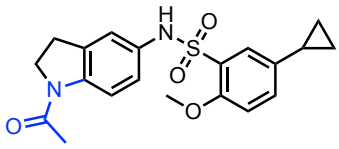
TRIM24 $K_d$ 0.22 μM
BRPF1 $K_d$ 0.14 μM
(Bennett *et al.* 2015)



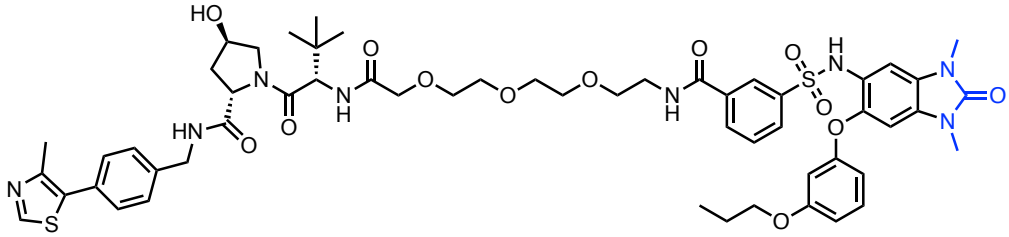
TRIM24 $IC_{50}$ 1.88 μM
BRPF1 $IC_{50}$ –
(Hu *et al.* 2020)



TRIM24 $IC_{50}$ 2.53 μM
BRPF1 $IC_{50}$ –
(Hu *et al.* 2020)



TRIM24 $IC_{50}$ 0.98 μM
BRPF1 $IC_{50}$ 1.16 μM
(Xiang *et al.* 2022)



TRIM24 $K_d$ 42 nM
BRPF1 $K_d$ 52 nM
(Gechijian *et al.* 2018)

**Figure 1.16:** Structures of the reported TRIM24 BRD ligands and the dTRIM24 PROTAC.[3,4,90–92] The presumed KAc-mimicking moieties are shown in blue. A dash (–) indicates that the affinity was not reported.

## 1.6 Peptide-Drug Conjugates (PDCs)

Peptide-drug conjugates (PDCs) are an emerging modality in drug discovery, linking a small molecule drug or ligand to a peptide through a cleavable or non-cleavable linker.

PDCs offer several advantages over their unlinked counterparts by combining the benefits of both peptides and small molecules.[93] For instance they can exhibit enhanced cell permeability compared to free peptides and reduced off-target promiscuity – a common issue with small molecule drugs that often leads to side effects.[93]

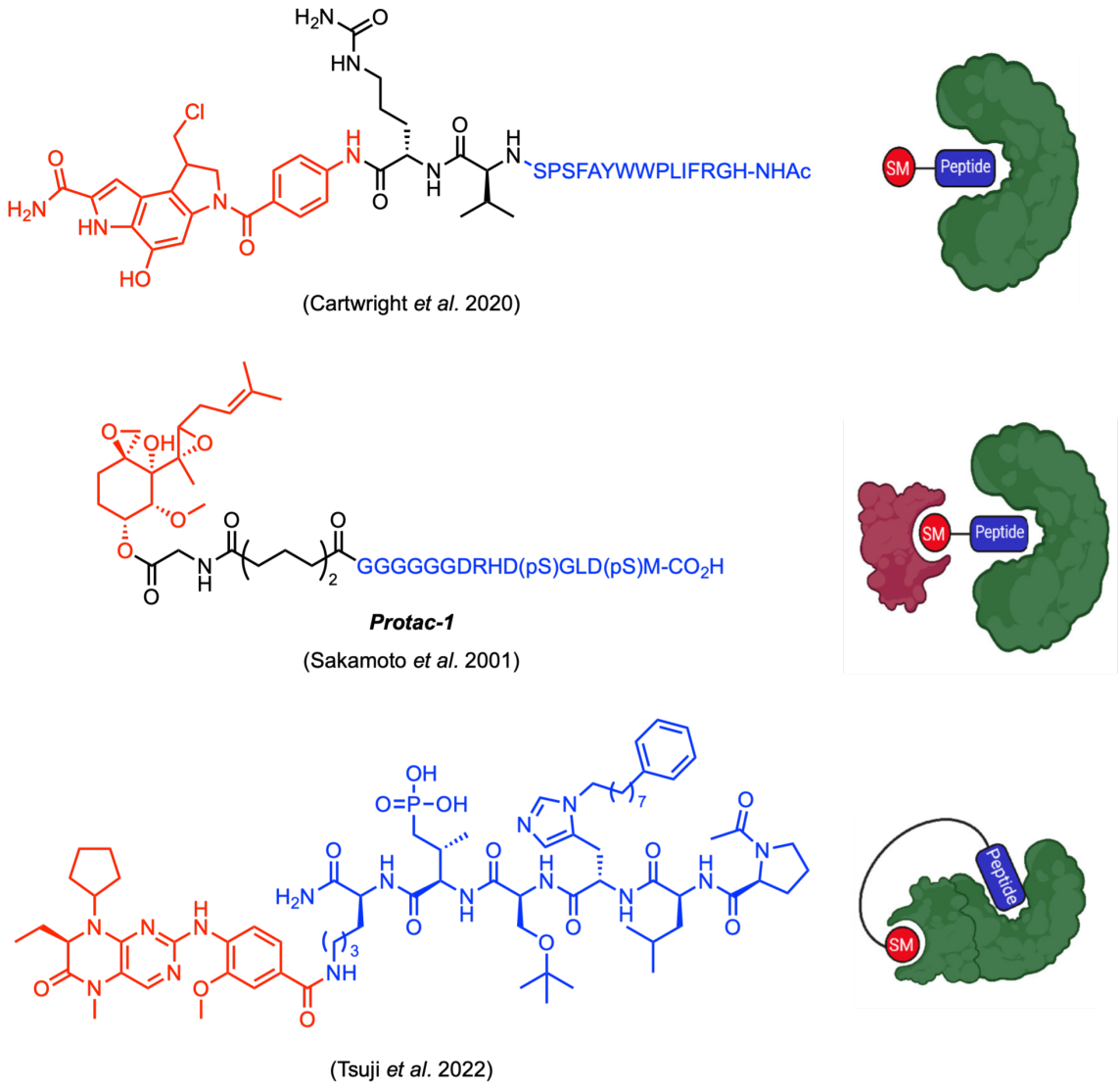



**Figure 1.17:** Structures of a selection of PDCs from the literature (pS = phosphoserine).[94–96] The small molecule and peptide sections of the molecules are coloured. Figure reprinted and adapted with permission from Dean *et al.* 2024.[93] Copyright 2024, American Chemical Society.

PDCs have found numerous applications in drug discovery for example in the targeted delivery of cytotoxic payloads to tumours, as demonstrated by Cartwright *et al.* (Figure 1.17).[94] Furthermore, PDCs have been used in the development of PROTACs, a notable

example being Protac-1 (Figure 1.17) – the first reported PROTAC from the seminal paper by Sakamoto *et al.*.[95] Various studies have leveraged the bifunctionality of PDCs to target multiple binding sites on a single protein, resulting in substantial increases in both affinity and selectivity. For example, Tsuji *et al.* disclosed a PDC (Figure 1.17) that simultaneously binds the catalytic kinase (KD) and polo-box (PBD) domains of polo-like kinase 1 (PLK1) with 2-orders of magnitude higher affinity than the parent monovalent PBD-binding peptide.[96] Since peptides can bind to protein surfaces, PDCs can also inhibit so-called 'undruggable' protein-protein interactions (PPIs), with the small molecule serving as an anchor in binding sites with higher druggability.[97]

## 1.7 Binding Assays

### 1.7.1 AlphaScreen®

AlphaScreen® (Amplified Luminescent Proximity Homogeneous Assay Screen) is a high-throughput bead-based assay used for evaluating the strength of binding interactions between a substrate (e.g. peptide or small molecule) and a biological target (e.g. protein). The assay was originally developed from the Luminescent Oxygen Channelling Immunoassay (LOCI) methodology reported by Ullman *et al.* in 1994.[98]

In an AlphaScreen® assay each interaction partner is conjugated to either the donor or acceptor beads, respectively. Excitation of the donor beads at 680 nm causes photosensitisers on the bead surface to excite ambient oxygen to the singlet state, producing around 60,000 singlet oxygen species per second (Figure 1.18).[99] When the substrate interacts with the target, the donor and acceptor beads are brought into close proximity, enabling singlet oxygen to diffuse a short distance (200 nm max.) through solution to the acceptor beads (Figure 1.18).[100] Reaction of singlet oxygen with chemicals on the acceptor bead surface causes a chemiluminescent signal to be produced. The intensity of emission

is dependent on the number of target-substrate complexes formed and hence the strength of the binding interaction. The assay can be adapted into a competition format by adding an inhibitor to disrupt the interaction between the substrate and its target. Singlet oxygen is rapidly quenched in solution; therefore, a disruption of the interaction will lead to a dose-dependent reduction in the emission signal, allowing $IC_{50}$ values to be determined (Figure 1.18).

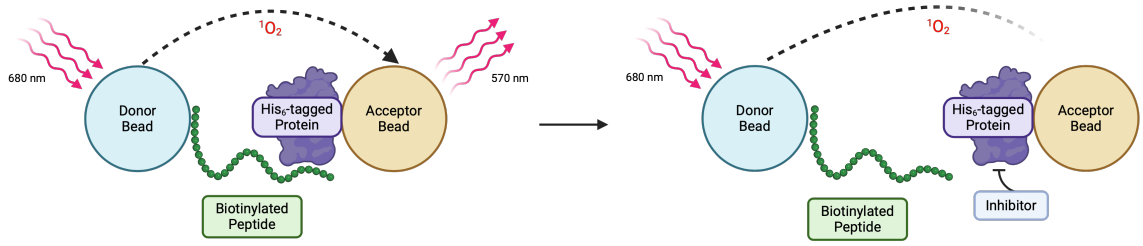



**Figure 1.18:** Schematic representation of an AlphaScreen® competition assay. Figure created with BioRender.com.

One of the key advantages of this technology is its low background signal, resulting due to the short lifetime of singlet oxygen (4 ms).[100] However, the assay can be susceptible to effects from pan assay interference compounds (PAINs) through mechanisms such as singlet oxygen quenching and aggregation.[101] To overcome this an AlphaScreen® TruHits assay is usually run in parallel to exclude any false positives.[101]

### 1.7.2 Fluorescence Anisotropy

Fluorescence anisotropy is another common technique used to investigate the binding interactions of proteins, peptides, and small molecules. It relies on the principle that when a fluorescently-labelled molecule is excited with plane-polarised light, the emission is polarised to a degree that is inversely proportional to the rate of molecular rotation.[102] A free fluorescent tracer in solution will have a fast rate of rotation, resulting in largely depolarised emission and a low anisotropy (Figure 1.19).[103] However when the tracer is

bound to a large molecule such as a protein, its emission will be polarised to a greater degree as a consequence of the slower rotational speed of the high molecular weight protein (Figure 1.19).[103]

The anisotropy (r) can be calculated from the measured parallel ($F_{\parallel}$) and perpendicular intensities ($F_{\perp}$); and for protein-tracer mixtures, it is proportional to the fraction of tracer bound to the protein.[104] Therefore, a protein titration can be used to determine the tracer binding affinity ($K_d$) (Figure 1.19). In the competition assay format, a non-fluorescent molecule is added to displace the tracer from the protein, and reduce the anisotropy in a dose-dependent manner, enabling an $IC_{50}$ value to be determined.

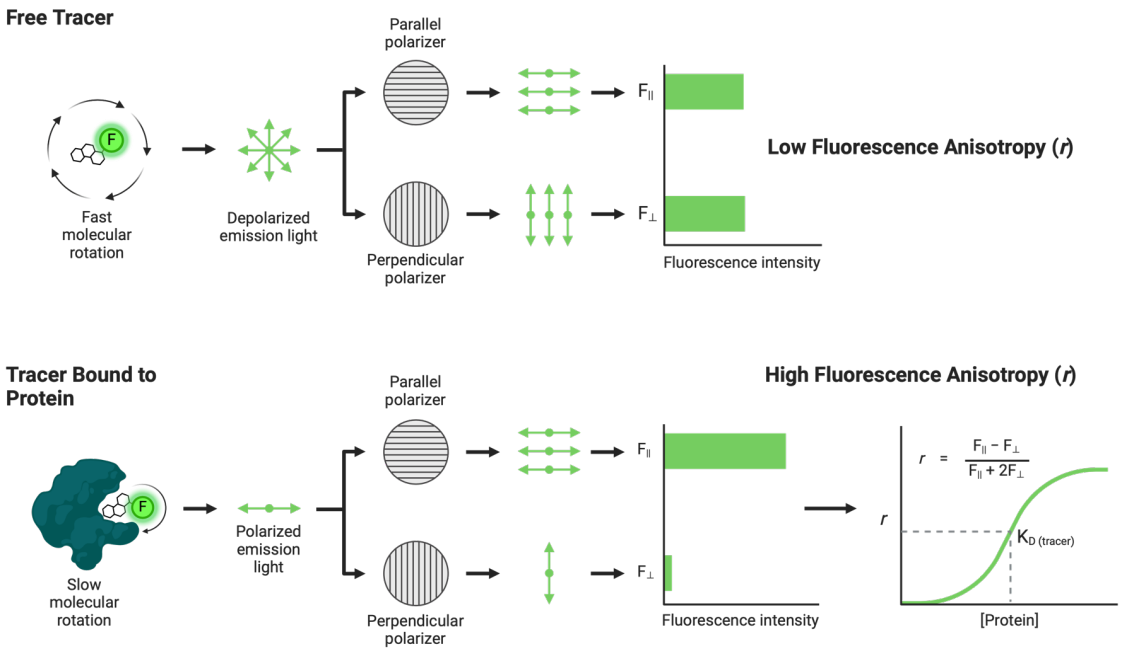



**Figure 1.19:** Schematic representation of the Fluorescence Anisotropy assay. Small molecule or peptide fluorescent tracers undergo fast molecular rotation in solution leading to a depolarisation of the emission light and low fluorescence anisotropy (*r*). Whereas when the tracer is bound to a large molecule such as a protein, which undergoes slow molecular rotation, the emission light is polarised leading to an increased *r*. As the concentration of protein is increased *r* will increase in a dose-dependent manner allowing a $K_d$ to be determined ($F_{\parallel}$ / $F_{\perp}$ = fluorescence intensities in the parallel and perpendicular planes). Figure created with BioRender.com.

### 1.7.3 Surface Plasmon Resonance (SPR)

Surface plasmon resonance (SPR) is a highly sensitive, label-free biophysical technique that measures changes in the refractive index near a biosensor surface due to biomolecular interactions.[105] This technique relies on the phenomenon where photons striking a metal, at a critical angle of incidence, transfer part of their energy to electrons in the metal surface, causing them to oscillate and create a plasmon wave.[105] The critical angle of incidence at which resonance occurs is determined by the refractive index near the metal surface.[105] In a typical SPR experiment an analyte solution is continuously passed over biomolecules immobilised on a gold surface (Figure 1.20). As the analyte binds, the refractive index changes, leading to a shift in the critical resonance angle (Figure 1.20).[106] This shift is detected as a change in the reflected light intensity and is recorded in response units (RU) in the form of a sensogram (Figure 1.20).[106] Analyte molecules with higher molecular weights can generate larger changes in the refractive index leading to greater resonance angle shifts.[106] A reference surface with no immobilised biomolecules is included to enable removal of the contributions from the buffer, capture method, and non-specific binding of the analyte to the sensor surface.[107]

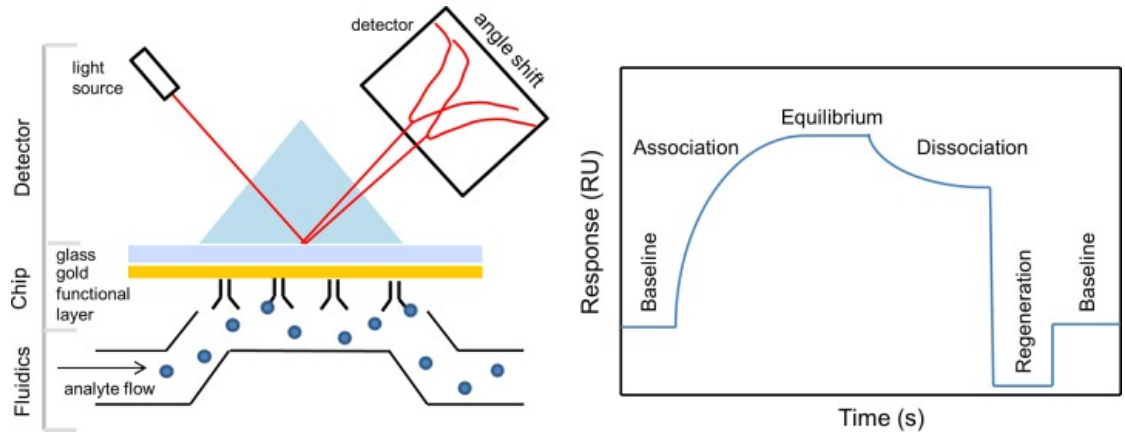



**Figure 1.20:** Schematic representation of the SPR set-up and a sensogram showing the key stages during an experiment. Figure reprinted with permission from Damborský *et al.* 2016.[108] Copyright 2016 The Author(s). Published by Portland Press Ltd on behalf of the Biochemical Society.

Affinities ($K_d$) can be determined by plotting binding saturation curves of the maximum response at equilibrium (upper plateau) for different analyte concentrations.[109] Real-time monitoring enables the determination of kinetic parameters from the shape of the sensograms.[109] It is important to highlight that $k_{on}$ cannot be directly measured, due to the analyte freely dissociating during the association phase. The association curve is, therefore, a function of both $k_{on}$ and $k_{off}$, with an observed rate constant known as $k_{obs}$.[109] On the other hand, during the dissociation phase the flow of the analyte is halted, and so this part of the sensogram is purely a function of $k_{off}$. $k_{on}$ must instead be derived using the measured values for $k_{obs}$ and $k_{off}$. A $K_d$ value can then alternatively be obtained using the law of mass action ($K_d = k_{off}/k_{on}$).[110]

### 1.7.4 BROMO*scan*

BROMO*scan* is a competition-based assay that is used to assess the affinities and selectivities of small molecule inhibitors across a panel of BRD targets.[111] The assay was developed from the pre-existing KINOME*scan*™ (Eurofins DiscoverX corporation) methodology, which involves the competition between a 'free' novel ligand and a known inhibitor immobilised on the surface of magnetic beads (Figure 1.21).[112] For BROMOscan, this competition occurs for the binding site of a DNA-tagged target BRD (Figure 1.21).[112] If the novel small-molecule binds to the BRD KAc-binding site, it will displace the known ligand from the protein, thereby removing the protein from the magnetic bead.[112] The beads are then isolated and washed to remove any unbound protein. The amount of DNA-tagged protein bound to the known inhibitor on the bead can then be quantified using qPCR.[112] A titration of the novel ligand will, therefore, lead to a dose-dependent decrease in the qPCR signal, as an increasing number of protein molecules are displaced from the bead surface. This enables a binding affinity to be determined for the novel ligand.

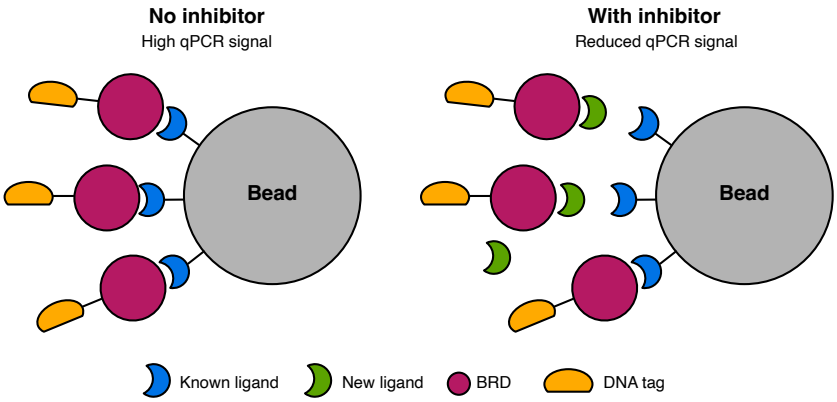



**Figure 1.21:** Schematic representation of the BROMO*scan* assay. Figure reprinted with permission from the creator, Dr Bernadette Lee.

BROMOscan has several advantages including high sensitivity, label-free detection and the generation of quantitative binding data. It's worth highlighting that the technology also has a number of disadvantages such as its reliance on having a known ligand for the target BRD and its incompatibility with allosteric binders.[111,112]

## 1.8 Hypothesis and Project Aims

TRIM24 overexpression is believed to drive breast cancer cell proliferation through a number of different mechanisms. For instance, TRIM24 can bind to chromatin through the combinatorial action of its PHD and BRD, facilitating the recruitment of transcription factors that upregulate oncogenes associated with cell proliferation.[2] Furthermore, TRIM24 may promote cancer proliferation through other mechanisms, such as enhancing cell adhesion *via* BRD-dependent self-SUMOylation, or PHD-dependent regulation of the p53 tumour suppressor.[70,75] As a result of its implication in cancer progression, the TRIM24 PHD-BRD cassette presents an attractive target for therapeutic intervention.

Two high-affinity TRIM24 BRD inhibitors have been developed, unfortunately, neither showed any effect on the proliferation of breast cancer cells.[3,4] Therefore, BRD inhibition alone appears insufficient to induce an antiproliferative phenotype on breast cancer cells.

Since TRIM24 can also bind to chromatin through its PHD domain, simultaneous inhibition of *both* the PHD and BRD might be necessary to achieve a noticeable antiproliferative effect on breast cancer cells.

Given the proximity of the PHD and BRD in TRIM24, a bivalent strategy was proposed to inhibit both domains simultaneously by conjugating a previously reported BRD ligand with an H3-mimicking peptide (Figure 1.22). The resulting peptide-drug conjugates (PDCs) could then be used to investigate the effects of simultaneous TRIM24 PHD and BRD inhibition on breast cancer cells.

In summary, the main aims of this project were to:

- Identify histone-mimicking peptides with enhanced affinity for the TRIM24 PHD (Chapter 1)
- Synthesise bivalent PDCs targeting the TRIM24 PHD and BRD (Chapter 2)
- Evaluate the PDCs in biophysical assays to characterise their binding to the TRIM24 PHD and BRD (Chapter 3)
- Evaluate the PDCs in cellular assays to identify any phenotypic effects on breast cancer cells. (Chapter 3)

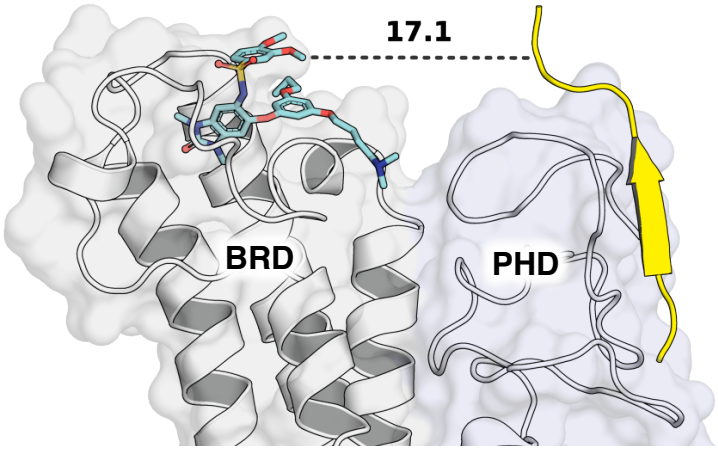



**Figure 1.22:** X-ray crystal structure of the TRIM24 PHD-BRD in complex with IACS-9571 (PDB ID: 4YC9[4]) overlaid with the peptide (carbon = yellow) from the crystal structure of TRIM24 in complex with $H3_{(1\text{-}10)}$ (PDB ID: 3O37[2]). The distance shown is given in Å.

# 2 Peptidomimetic TRIM24 PHD Ligands

## 2.1 Introduction and Aims

While two high-affinity TRIM24 BRD ligands have been reported in the literature, no small-molecule TRIM24 PHD ligands existed prior to this work.[3,4,89] Therefore, it was necessary to identify effective PHD ligands before embarking on the development of compounds that simultaneously inhibit the TRIM24 PHD and BRD. The TRIM24 PHD recognises H3R2K4K9Me$_3$ using a protein surface groove involving residues N825, E826, D827, and W828 (Figure 2.1). This work focussed on inhibiting a non-canonical cation-π interaction between H3K9Me$_3$ and a solvent-exposed tryptophan residue (W828) (Figure 2.1) in the PHD.[1,2] The TRIM24 PHD lacks a defined binding pocket and so fragments would likely have prohibitively low affinities for detection in a fragment screen. Instead, to ensure a detectable affinity, a peptidomimetic approach was used involving histone-mimicking peptides. Detailed in this chapter are two different peptidomimetic strategies that were used to identify moieties that could interact with W828 and enhance affinity of a H3-mimicking peptide for the TRIM24 PHD.

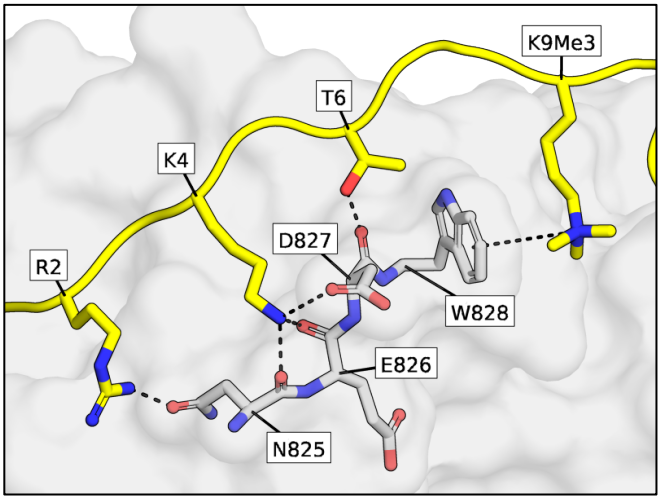



**Figure 2.1:** New *apo* crystal structure of TRIM24 (1.7 Å, PDB ID: 9GD5; surface and carbon = grey) overlaid with the peptide from the crystal structure of TRIM33α (which has an identical PHD domain) in complex with H3$_{(1–22)}$K9Me$_3$K14AcK18Ac (PDB ID: 3U5O; carbon = yellow). The structure shows the interactions between a histone-mimicking peptide and the TRIM24 PHD.

## 2.2 H3K9Me$_3$-mimicking Peptides

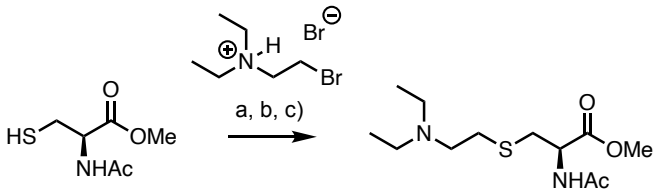
### 2.2.1 Cysteine-selective Alkylation Reactions

The first strategy for identifying groups that bind to W828 in the TRIM24 PHD involved the use of histone-mimicking peptides containing KMe$_3$ mimics. Inspiration was taken from three studies which used cysteine-selective alkylation as a method for the incorporation of KAc (Sekirnik *et al.*) and KMe$_x$ (Simon *et al.*, Pieters *et al.*) mimics into histones.[5–7,113,114] It was decided that a similar approach would be used to incorporate KMe$_3$ bioisosteres into H3 peptides containing a cysteine at position 9 (H3K9C). A cysteine-selective alkylation approach was favoured over other site-specific modification strategies, such as dehydroalanine alkylation, as it should not lead to epimerisation at the α-carbon, which could reduce affinity for the PHD.[115]

### 2.2.2 Validation of the Cysteine Alkylation Strategy

Before pursuing the cysteine-alkylation strategy further, it was necessary to validate its suitability for identifying groups that could interact with W828 in the TRIM24 PHD. An alkylation buffer, recommended by Simon *et al.*, was used which contained 1 M HEPES (pH 7.8), 10 mM methionine, and 4 M guanidine·HCl.[5] The alkylation buffer is crucial to the success of cysteine-selective alkylation as it controls the reaction pH to enhance the nucleophilicity of cysteines and ensure lysines are protonated to minimise unwanted over-alkylation. During the course of the alkylation reaction the pH decreases, as the alkylating agent is consumed and acid is produced. Therefore, high concentrations of buffer are required to prevent acidification of the reaction mixture, which would hamper the reaction through protonation of cysteines.[5] A test reaction was carried out between *N*-acetyl-L-cysteine methyl ester and (2-bromoethyl)trimethylammonium bromide to ensure that a reaction pH of 7.8 and the alkylation buffer were sufficient for effective alkylation on cysteine (Scheme 2.1). The reactions demonstrated excellent conversion to the desired

product by LC-MS and $^1$H NMR (Figures 2.2 and 2.3), indicating that this buffer could be used for cysteine alkylation reactions to incorporate $KMe_3$ mimics into histone peptides.




**Scheme 2.1:** Alkylation of *N*-acetyl-L-cysteine methyl ester in alkylation buffer (pH 7.8). *Reagents and conditions*: a) DTT, alkylation buffer, 37 °C, 1 h; then b) DTT, $[BrCH_2CH_2NH(Et)_2]^+Br^-$, 50 °C, in the dark, 5 h; then c) $HOCH_2CH_2SH$, rt, 30 min (alkylation buffer = 1 M HEPES (pH 7.8), 10 mM methionine, 4 M guanidine·HCl).[5]

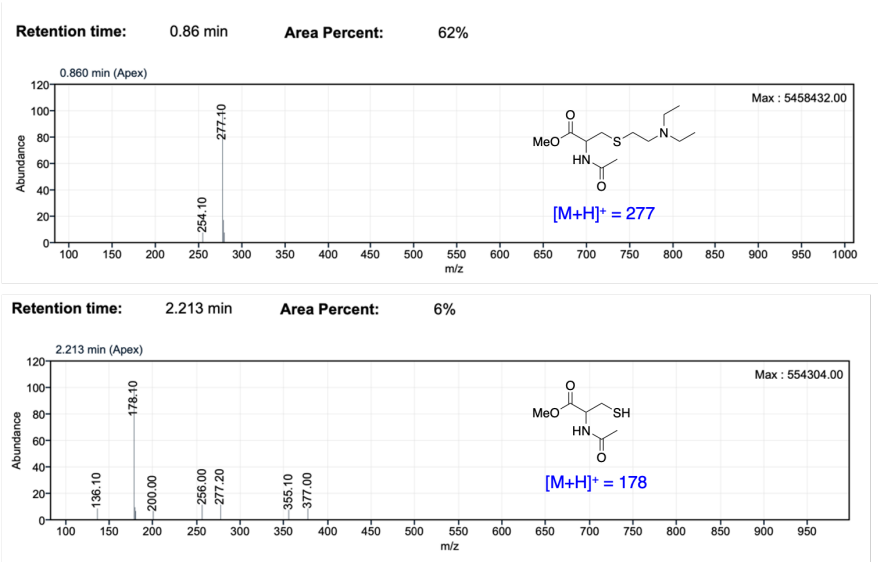



**Figure 2.2:** LC-MS spectra and peak area percentages for the test alkylation reaction after 5 h.

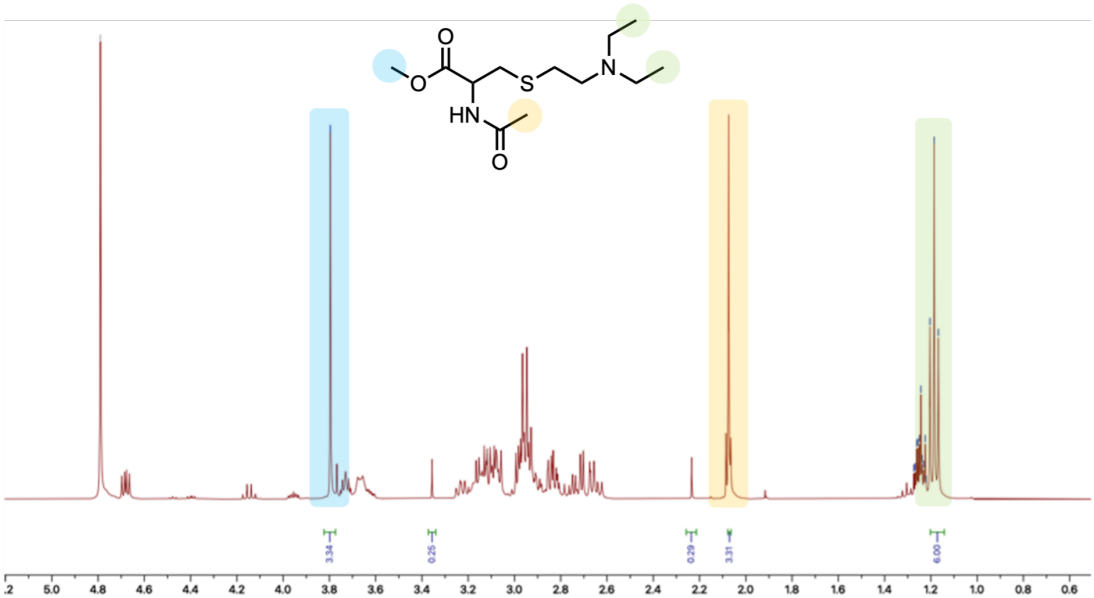
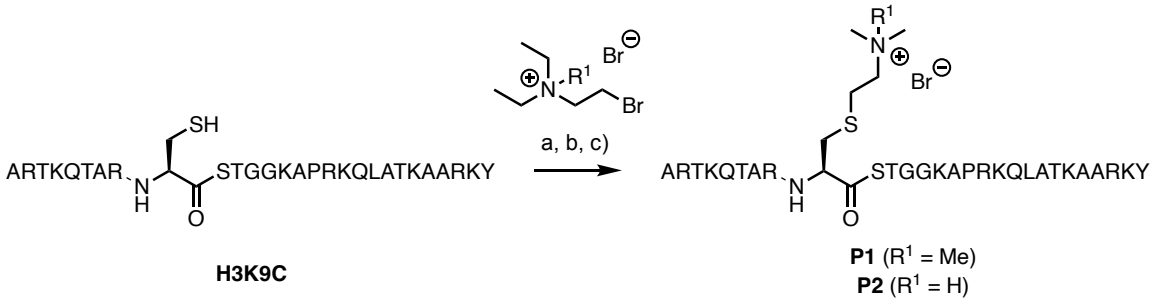


**Figure 2.3:** Crude $^1$H NMR ($D_2O$, 400 MHz) spectrum of the test alkylation reaction products.

A disadvantage of cysteine alkylation is that several properties of the lysine side-chain are altered by the introduction of a sulfur atom; for instance the side-chain geometry is changed due to different bond angles ($\angle_{CSC}$ = 99° *vs* $\angle_{CCC}$ = 109°).[114,116] Replacement of trimethyllysine ($KMe_3$) with its sulfur-containing analogue ($K_CMe_3$) can have detrimental effects on the binding of histone peptides to their targets. For example it was reported that $H3K_C4Me_3$ has a 13-fold weaker binding affinity to the BPTF PHD finger, than its natural $KMe_3$ analogue.[117] It was also important to investigate the effects of switching from a quaternary ammonium to a tertiary amine moiety, which would enable a larger library of $KMe_3$ mimics to be screened. The tertiary amine would be protonated under physiological pH and so could still engage with W828 through a cation-π interaction.




**Scheme 2.2:** Cysteine-selective alkylation of a cysteine mutant H3 peptide (H3K9C). *Reagents and conditions*: a) DTT, alkylation buffer, 37 °C, 1 h; then b) DTT, $[BrCH_2CH_2N(Me)_2R^1]^+Br^-$, rt/50 °C, in the dark, 5 h; then c) $HOCH_2CH_2SH$, rt, 30 min [alkylation buffer = 1 M HEPES (pH 7.8), 10 mM methionine, 4 M guanidine·HCl].

To investigate these effects a $H3_{(1–27)}K9C$ mutant peptide was alkylated according to a previously reported procedure by Simon *et al.*, to yield $H3_{(1–27)}K_C9Me_3$ (**P1**) and $H3_{(1-27)}K_C9Me_2$ (**P2**) (Scheme 2.2).[5] Pleasingly, efficient conversion to the monoalkylated product was confirmed by LC-MS (Appendix 3, Figure A3.1) for peptide **P1**. However, significant overalkylation was observed at 50 °C for peptide **P2**, therefore, the reaction temperature and alkylating agent equivalents were reduced (Appendix 3, Figure A3.1). The sites of alkylation were then evaluated using NanoLC-MS/MS (Appendix 3, Figure A3.2);

with alkylation solely occurring at cysteine-9 for peptide **P1**. The MS/MS data suggested additional sites of alkylation for peptide **P2**, including lysine-23, and the *N*-terminus. The peptides were then purified using analytical reverse phase HPLC (Appendix 3, Figure A3.3), and their binding affinities were evaluated using an adapted version of a TRIM24 AlphaScreen® competition assay, which had been previously developed in the Conway group (Figure 2.4).[1] In this assay, compounds were evaluated for their ability to displace biotinylated $H3_{(1\text{-}27)}K9Me_3$ from the PHD finger of recombinantly-expressed TRIM24 (PHD-BRD), with lower $IC_{50}$ values indicating a stronger binding affinity for the TRIM24 PHD. $IC_{50}$ values for $H3_{(1–27)}K9Me_3$ and $H3_{(1–27)}K_C9Me_3$ (**P1**) were 4.2 ± 0.6 μM and 5.0 ± 0.5 μM, respectively (Figure 2.5), which are in good agreement with the reported ITC $K_d$ of the $H3_{(1–27)}K9Me_3$ peptide (3.9 ± 0.4 μM).[1] As expected, a comparison of the binding affinities for $H3_{(1–27)}K_C9Me_3$ (**P1**, 5.0 ± 0.5 μM) and $H3_{(1–27)}K_C9Me_2$ (**P2**, 7.8 ± 0.7 μM) revealed a slight reduction in TRIM24 affinity upon switching from a quaternary to a tertiary amine (Figure 2.5). Nonetheless, these data suggest that incorporation of sulfur, and switching from quaternary to tertiary amines, does not substantially affect binding of histone peptides to the TRIM24 PHD finger.

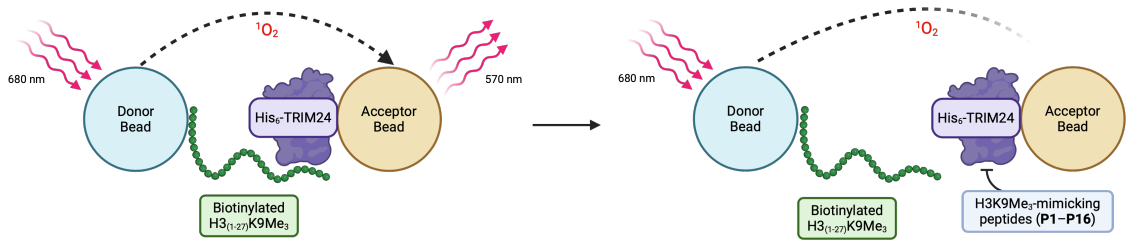
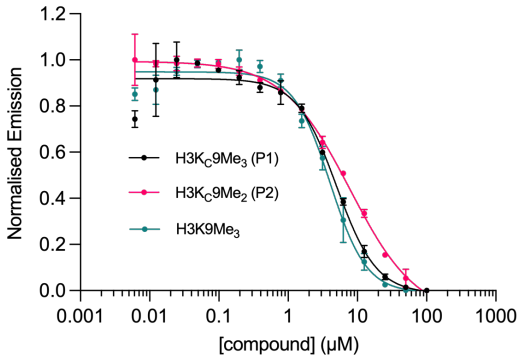



**Figure 2.4:** Graphical representation of the AlphaScreen® competition assay.[1] Briefly, histone-mimicking peptides are added to inhibit the interaction between biotinylated $H3_{(1\text{-}27)}K9Me_3$, attached to the donor beads, and $His_6$-TRIM24 (PHD-BRD), attached to the acceptor beads. Inhibition of this interaction prevents the beads from coming into close proximity and causes a reduction in the AlphaScreen® emission signal that is proportional to the inhibitor binding affinity.




| Compound | AlphaScreen® $IC_{50}$ (μM) |
|---|---|
| $H3_{(1–27)}K9Me_3$ | 4.2 ± 0.6 |
| $H3_{(1–27)}K_C9Me_3$ (**P1**) | 5.0 ± 0.5 |
| $H3_{(1–27)}K_C9Me_2$ (**P2**) | 7.8 ± 0.7 |

**Figure 2.5:** AlphaScreen® competition assay dose-response curves and $IC_{50}$ values for inhibition of the interaction between biotinylated $H3_{(1–27)}K9Me_3$ and $His_6$-TRIM24 (PHD-BRD) by unbiotinylated $H3_{(1–27)}K9Me_3$, $H3_{(1–27)}K_C9Me_3$ (**P1**), and $H3_{(1–27)}K_C9Me_2$ (**P2**). $IC_{50}$ values are quoted as the mean of triplicate data ± the standard error of the mean. Error bars indicate the standard deviation.

### 2.2.3 $KMe_3$ Mimic Design

Having validated the cysteine-alkylation strategy, a number of putative $KMe_3$ mimics (Figure 2.6) were chosen for incorporation into histone-mimicking peptides. Previous studies have used quaternary amines as isosteres of $KMe_3$ such as pyridinium,[6,118] although the majority of work has focused on tertiary amines due to their increased cell permeability, for example diethyl, pyrrolidinyl, and piperidinyl derivatives (Figure 2.6).[119–124] A number of alternative moieties were proposed for their potential as $KMe_3$ mimics, for example isoindoline and tetrahydroisoquinoline, which could have enhanced affinity for the tryptophan through additional π-π interactions (Figure 2.6). The pyridine (pKaH = 5.9[125]) and pyrazole (pKaH = 2.7[125]) moieties are unlikely to be protonated under a physiological pH of 7.4 and were chosen to investigate if π-π interactions can compensate for the loss of the cation-π interaction. The morpholine and piperazine groups have the potential to form additional hydrophilic contacts with other residues in the PHD, as well as maintaining the cation-π interaction with W828 (Figure 2.6). It has been reported that sugars, such as β-D-galactose (Figure 2.6), form CH-π interactions with tryptophan residues due to the polarised nature of the axial C-H bonds, and could serve as alternative $KMe_3$ mimics.[126]

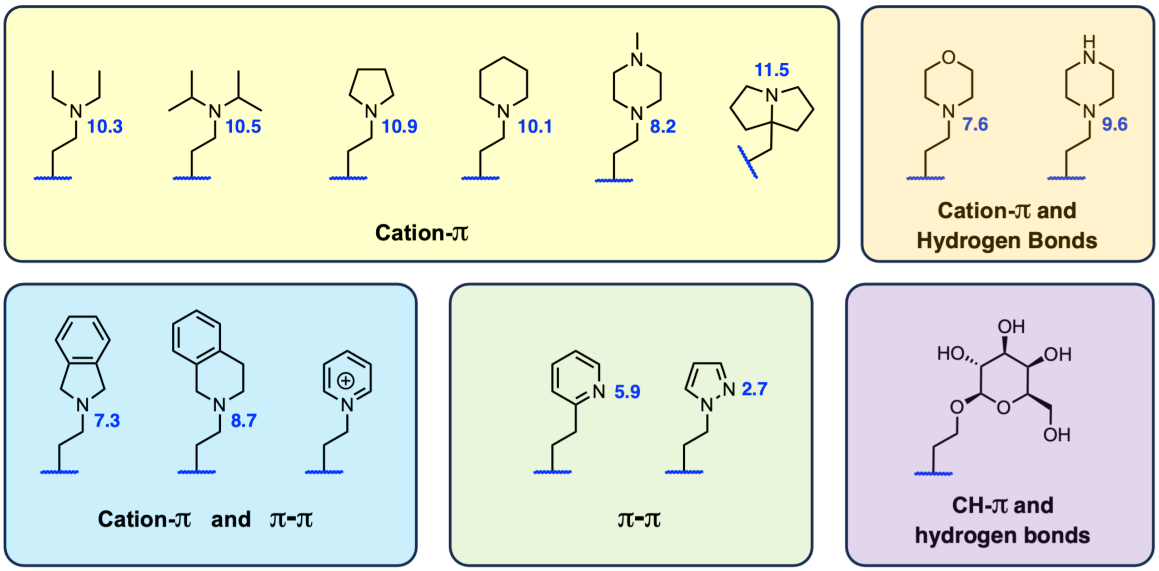
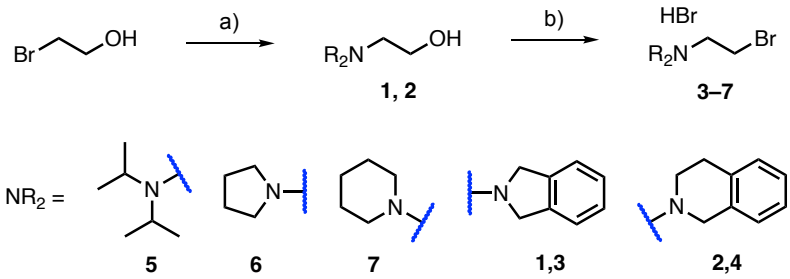


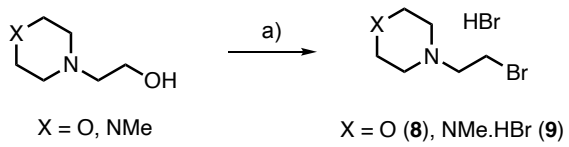

**Figure 2.6:** Structures of the $KMe_3$ mimics used in this study and the possible interactions with W828 and other residues in the TRIM24 PHD finger. Predicted pKaH values (MolGpKa) are stated for the most basic nitrogen atom in each structure.[125]

### 2.2.4 Synthesis of H3K9Me3-mimicking Peptides

Several $KMe_3$ mimics were readily available as the corresponding amino alcohols, whilst others could be converted to amino alcohols in a single step from 2-bromoethanol in 44–74% yield (Scheme 2.3a). Alkylating agents could then be synthesised by heating the amino alcohols under reflux in conc. aqueous HBr (Scheme 2.3b) to obtain alkyl bromide salts **3**–**7** in 34–78% yield. Alternatively, morpholine and methyl piperazine alkylating agents **8** and **9** were synthesised from their respective amino alcohols, in 44–74% yield, using an Appel-type reaction with $PPh_3Br_2$ (Scheme 2.4).[127]

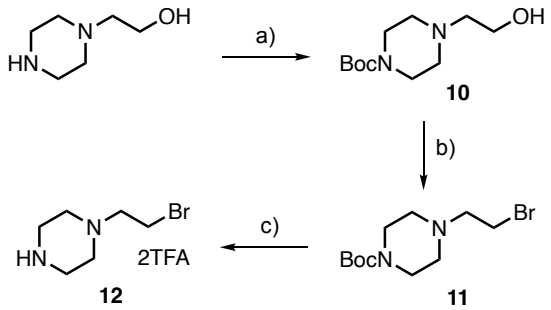



**Scheme 2.3:** Synthesis of alkylating agents **3**–**7**. *Reagents and conditions*: a) $R_2NH$, 2-bromoethanol, $NEt_3$, toluene, rt–120 °C, 6–20 h, (**1**) 44%, (**2**) 74%; b) conc. aqueous HBr, 0–135 °C, 20 h, (**3**) 34%, (**4**) 52%, (**5**) 39%, (**6**) 61%, (**7**) 78%.

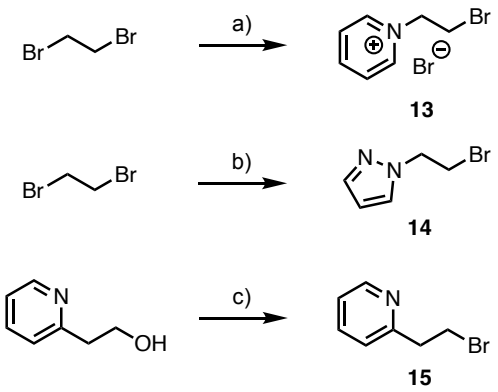



**Scheme 2.4:** Synthesis of alkylating agents **8** and **9**. *Reagents and conditions*: a) $PPh_3Br_2$, $CH_2Cl_2$, 0 °C–rt, 6–20 h, (**8**) 70%, (**9**) 44%.

The two previous approaches (Schemes 2.3 and 2.4) were also attempted for the synthesis of piperazine alkylating agent **12**, however, this resulted in the formation of DABCO. This was presumably due to favoured intramolecular 6-*exo-tet* cyclisation of the secondary amine onto the alkyl bromide that had formed during the reaction. Instead, a protecting group strategy was employed which involved Boc protection of the secondary amine (Scheme 2.5). Boc protected amine **10** was obtained in 99% yield by reaction with $Boc_2O$ (Scheme 2.5) and was subsequently brominated using Appel conditions with $CBr_4$ and $PPh_3$ to obtain compound **11** in 90% yield (Scheme 2.5). Boc-protected amine **11** was then deprotected using TFA, which protonates the nitrogen and prevents cyclisation, to yield alkylating agent **12** in 84% yield (Scheme 2.5). Pyridinium salt **13** and pyrazole **14** were synthesised from 1,2-dibromoethane (Scheme 2.6) according to procedures by Pieters *et al.* and López *et al*., respectively.[6,128] TBAB was used, in the synthesis of **14** as a phase transfer catalyst due to the immiscibility of 1,2-dibromoethane and water. Pyridine **15** was obtained in 45% yield using another Appel reaction, which has previously been reported by Mambwe *et al*.[129]



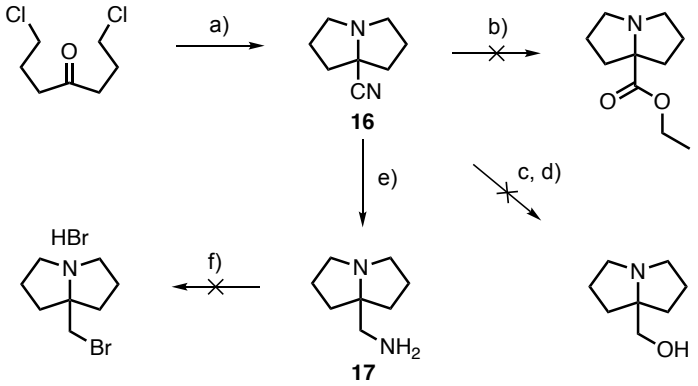

**Scheme 2.5:** Synthesis of alkylating agent **12**. *Reagents and conditions*: a) $Boc_2O$, THF, 0 °C, 2 h, 99%; b) $CBr_4$, $PPh_3$, $CH_2Cl_2$, 0 °C–rt, 20 h, 90%; c) TFA, $CH_2Cl_2$, 0 °C–rt, 2 h, 84%.



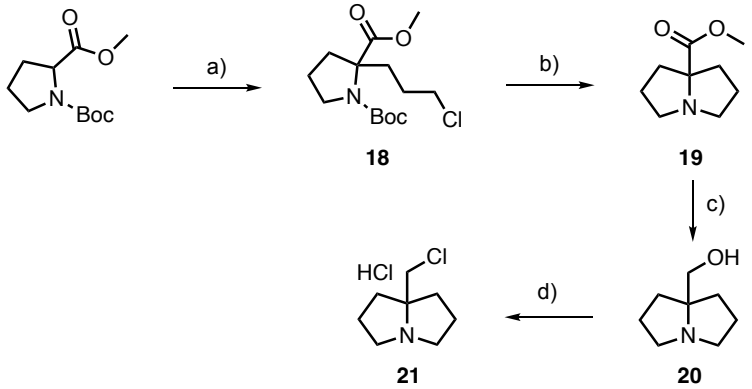

**Scheme 2.6:** Synthesis of alkylating agents **13**–**15**. *Reagents and conditions*: a) Pyridine, acetone, rt, 60 °C, 16 h, 74%; b) Pyrazole, TBAB (10 mol%), aqueous NaOH, rt, 24 h, 62%; (c) $CBr_4$, $PPh_3$, THF, 0 °C–rt, 20 h, 45%.

Synthesis of a pyrrolizidine alkylating agent proved more challenging and required a number of different routes to be investigated. Initially, 1,7-dichloroheptan-4-one was converted to nitrile **16** in 74% yield (Scheme 2.7), as reported by Oka *et al.*.[130] In this reaction, an imine is formed, which subsequently undergoes double 5-*exo*-*tet* cyclisation. The resulting bicyclic iminium intermediate is trapped by a cyanide anion, which is released from the breakdown of acetone cyanohydrin. The aim was then to convert nitrile **16** to an alcohol, which could be brominated with conc. aqueous HBr to yield the desired alkylating agent. However, attempts at converting nitrile **16** to the ester, or reducing it to the alcohol *via* the aldehyde, both failed (Scheme 2.7). A different route was then attempted involving $LiAlH_4$ reduction of nitrile **16** to amine **17** (Scheme 2.7), which was achieved in 80% yield. The intention was then to convert amine **17** to the desired alkylating agent *via* a diazonium intermediate, by reaction with $NaNO_2$ and HBr (Scheme 2.7). However, this resulted in a complex mixture of products, from which the desired product could not be isolated. Ultimately, a successful route was established starting from *N*-Boc-proline methyl ester using reactions disclosed in two patents.[131,132] The first step involved enolate formation of the methyl ester using LDA, followed by chemoselective alkylation to afford alkyl chloride **18** in 68% yield (Scheme 2.8). In the second step, deprotection of the Boc group by HCl, produced *in situ* from the reaction of TMSCl and MeOH, enabled 5-*exo*-*tet*

cyclisation to form the azabicyclic scaffold (Scheme 2.8) of compound **19** in 79% yield. Pleasingly, reduction of methyl ester **19** to alcohol **20** with $LiAlH_4$ in 80% yield, followed by chlorination with $SOCl_2$, afforded alkylating agent **21** in 91% yield.



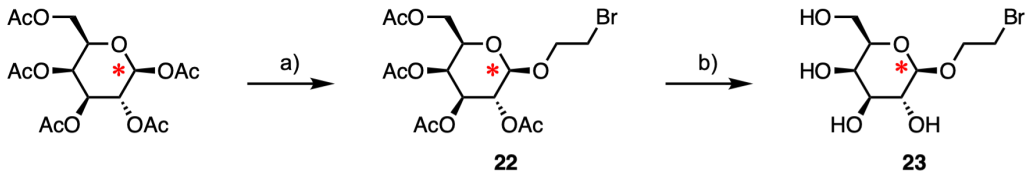

**Scheme 2.7:** Initial attempts to synthesise an alkylating agent containing a pyrrolizidine moiety. *Reagents and conditions*: a) acetone cyanohydrin, $NH_3$ in MeOH (7 M), 0 °C–rt, 24 h, 77%; b) HCl in EtOH (1.25 M), $H_2O$, 100 °C, 16 h, 0%; c) DIBAL-H, $CH_2Cl_2$, −78 °C, 30 min; then aqueous work-up; d) $NaBH_4$, MeOH, 0 °C–rt, 20 h, 0%; e) $LiAlH_4$, $Et_2O$, 0 °C–rt, 1 h, 80%; f) $NaNO_2$, HBr in AcOH (33% *v*/*v*), 0 °C–rt, 24 h, 0%.




**Scheme 2.8:** Synthesis of alkylating agent **21**. *Reagents and conditions*: a) LDA, THF, −78 °C, 2 h, then $ClCH_2CH_2CH_2I$, −78 °C–rt, 2 h, 68%; b) TMSCl, MeOH, 0 °C–rt, 16 h, 79%; c) $LiAlH_4$, THF, −10 °C–rt, 1 h, 88%; d) $SOCl_2$, toluene, 0–80 °C, 3 h, 91%.

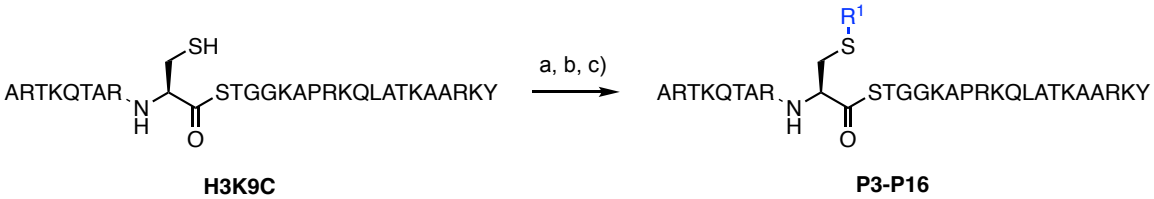
Alkylating agent **23** was synthesised from β-D-galactose pentaacetate using a modified version of a procedure reported by Hwu *et al.* (Scheme 2.9).[133] In the first step (Scheme 2.9) co-ordination of $BF_3{\cdot}OEt_2$ promotes elimination of acetate from the anomeric carbon (*) allowing nucleophilic attack by 2-bromoethanol to yield compound **22**. Anchimeric assistance from the neighbouring acetate group hinders nucleophilic attack from the bottom face of the tetrahydropyran ring, resulting in substitution with retention of configuration.

Acetyl deprotection with NaOMe afforded alkylating agent **23** in 57% yield. Compounds **22** and **23** were both assigned as the β-anomers due to the relatively large coupling constants of the anomeric proton (*), which is in accordance with the literature.[133]




**Scheme 2.9:** Synthesis of alkylating agent **23**. *Reagents and conditions*: a) $BF_3 \cdot OEt_2$, 2-bromoethanol, $CH_2Cl_2$, 0 °C, 3 h, 58%; b) NaOMe, MeOH, rt, 4 h, 57%. The anomeric carbon is indicated by *.

With the desired alkylating agents in hand, the H3K9Me$_3$-mimicking peptides could now be synthesised using the cysteine-selective alkylation reactions, described previously (Scheme 2.10). The majority of alkylating agents were soluble in the aqueous alkylation buffer. However, reactions involving alkylating agents **3**, **4**, **14**, and **15** were supplemented with 2% DMSO to improve solubility. Reaction conditions such as temperature and alkylkating agent equivalents were optimised to minimise overalkylation. LC-MS analysis (Appendix 3, Figure A3.1) revealed reactions occurred with only a small degree of overalkylation, as exemplified by the deconvoluted mass spectra of peptides **P6** and **P10** (Figures 2.7a and 2.8a). After purification (Figures 2.7b and 2.8b; Appendix 3, Figure A3.3) compounds were characterised by MALDI-TOF MS (Figures 2.7c and 2.8c; Appendix 3, Figure A3.4).




**H3K9C** **P3-P16**

**Scheme 2.10:** Synthesis of H3K9Me$_3$-mimicking peptides **P3**–**P16**. *Reagents and conditions*: a) DTT, alkylation buffer, 37 °C, 1 h; then b) DTT, $R^1X$ (X = Cl, Br), rt/50 °C, in the dark, 5 h; then c) $HOCH_2CH_2SH$, rt, 30 min. $R^1X$ = alkylating agents : 13 synthesised (**3**–**9**, **12**–**15**, **21**, **23**) and 3 commercially available [Alkylation buffer = 1 M HEPES (pH 7.8), 10 mM methionine, 4 M guanidine·HCl].

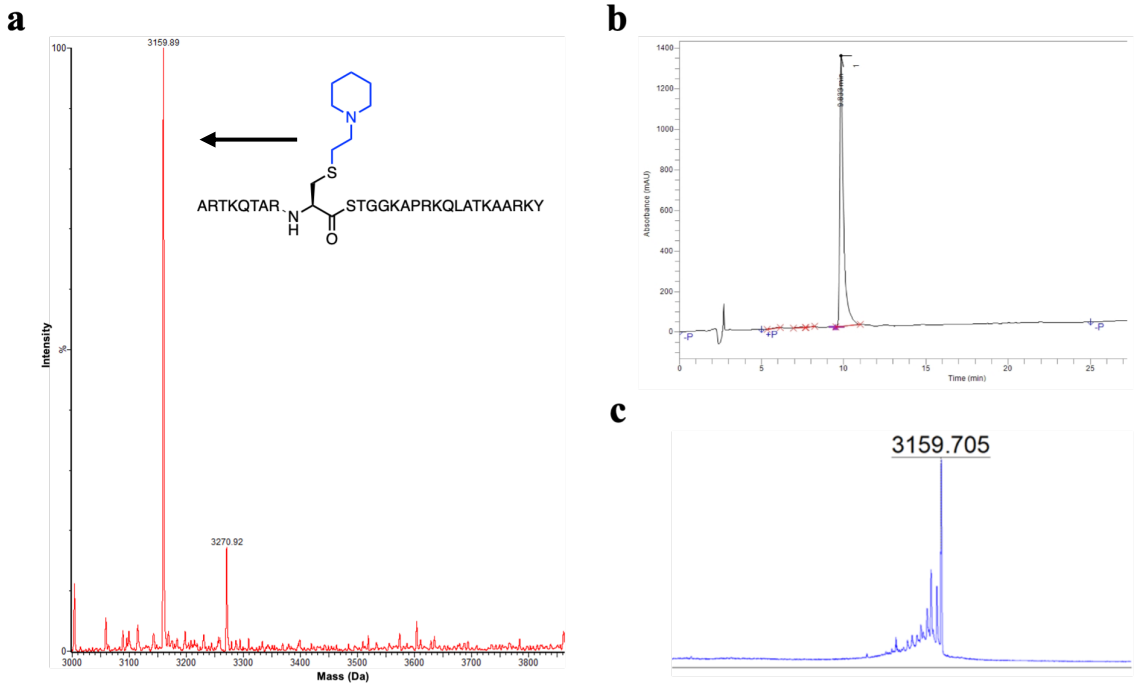



**Figure 2.7:** a) Deconvoluted ESI$^+$ mass spectrum of the reaction products for **P6** (structure shown). Theoretical MWs = 3159.76 (×1), 3270.95 (×2); Observed MWs = 3159.89 Da (×1), 3270.92 (×2); b) Analytical HPLC chromatogram [5–65% MeCN/$H_2O$ (+0.1% TFA), 25 min, 220 nm] for **P6** after purification; c) MALDI-TOF MS spectrum for **P6** after purification.

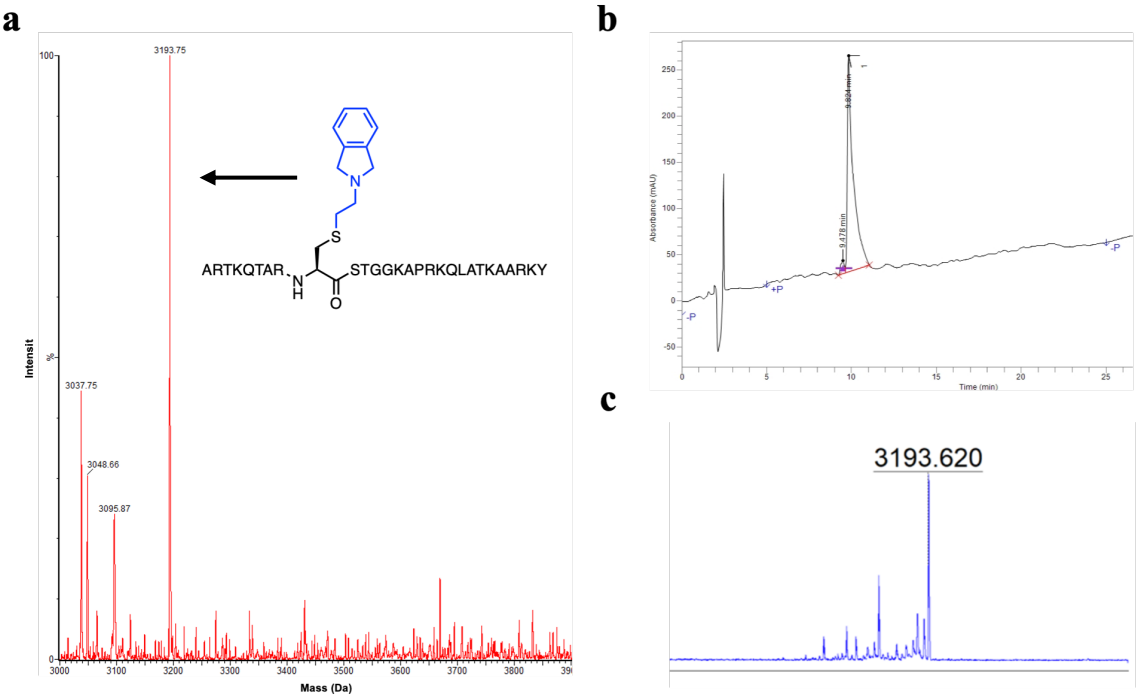
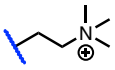
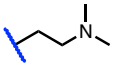
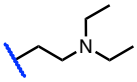
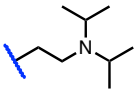
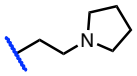
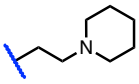
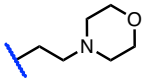
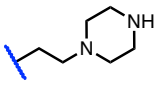
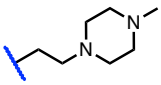
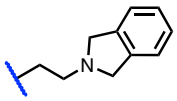
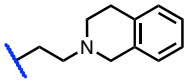
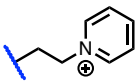
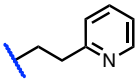
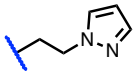
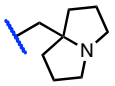
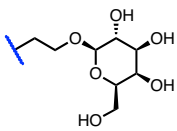



**Figure 2.8:** a) Deconvoluted ESI$^+$ mass spectrum of the reaction products for **P10** (structure shown). Theoretical MW = 3193.78 (×1), Observed MW = 3193.75 (×1); b) Analytical HPLC chromatogram [5–65% MeCN/$H_2O$ (+0.1% TFA), 25 min, 220 nm] for **P10** after purification; c) MALDI-TOF MS spectrum for **P10** after purification.

### 2.2.5 Evaluation of H3K9Me$_3$-mimicking Peptides

An AlphaScreen® competition assay, previously described in section 2.2.2, was used to evaluate the peptides' ability to displace biotinylated H3$_{(1-27)}$K9Me$_3$ from the TRIM24 PHD finger (Dose-response curves - Appendix 4, Figure A4.4). Analysis of the $IC_{50}$ values for unbiotinylated H3$_{(1-27)}$K9Me$_3$ (4.2 μM), H3$_{(1-27)}$K9Me$_0$ (10.0 μM), and H3$_{(1-27)}$K9C (12.7 μM) emphasised the importance of the cation-π interaction for binding to the PHD (Figure 2.9), which is in accordance with previous work by Sekirnik *et al.*.[1] A general trend was observed of increasing affinity with increasing basicity of the nitrogen atom (e.g. **P2**–**P4**, Table 2.1). This was exemplified by morpholine peptide **P7**, which exhibited a significantly weaker binding affinity than piperidine peptide **P6** (Table 2.1, Figure 2.9). This was likely due to the electron-withdrawing inductive effect from the oxygen, reducing the proportion of protonated molecules and, therefore, weakening the cation-π interaction. Peptides containing charge neutral mimics (**P13**, **P14,** and **P16**) demonstrated low affinities (Table 2.1, Figure 2.9) suggesting that π-π and CH-π interactions alone are insufficient to compensate for loss of the cation-π interaction with W828. Basicity and cation-π interaction strength are likely not the only factors influencing affinity. For example, the isoindoline and tetrahydroisoquinoline peptides (**P10** and **P11**) demonstrated increased affinity compared to peptides **P2**-**P6**, despite having lower basicities (Table 2.1, Figure 2.9). This was thought to be due to additional π-π interactions with W828. The pyrrolizidine in **P15** showed the most significant increase in affinity, with an approximate 4-fold boost relative to KMe$_3$ (Table 2.1, Figure 2.9). The pyrrolizidine moiety possesses a highly conformationally-constrained nitrogen atom and has the highest basicity, which likely contributes to the increased affinity. Therefore, it was concluded that pyrrolizidine (**P15**), isoindoline (**P10**), and tetrahydroisoquinoline (**P11**) are suitable alternatives to KMe$_3$ that could be incorporated into the PHD-binding section of the dual PHD-BRD ligands.

**Table 2.1:** AlphaScreen® $IC_{50}$ values for the inhibition of the interaction between biotinylated $H3_{(1\text{-}27)}K9Me_3$ and $His_6$-TRIM24 by **P1**–**P16**. $IC_{50}$ values are quoted as the mean of triplicate data ± standard error of the mean. Hot colours indicate stronger binders and cold colours indicate weaker binders. Predicted pKaH values (MolGpKa) are shown for the most basic N atom.[125]

| Compound | Peptide Modification | AlphaScreen® $IC_{50}$ (μM) | Predicted pKaH[125] |
|---|---|---|---|
| **P1** |  | 4.99 ± 0.52 | – |
| **P2** |  | 7.75 ± 0.72 | 10.0 |
| **P3** |  | 5.79 ± 0.83 | 10.3 |
| **P4** |  | 4.55 ± 0.21 | 10.5 |
| **P5** |  | 4.71 ± 0.66 | 10.9 |
| **P6** |  | 3.56 ± 0.30 | 10.1 |
| **P7** |  | 23.8 ± 8.9 | 7.6 |
| **P8** |  | 6.38 ± 0.92 | 9.6 |
| **P9** |  | 5.56 ± 0.70 | 8.2 |
| **P10** |  | 2.97 ± 0.21 | 7.3 |
| **P11** |  | 3.14 ± 0.33 | 8.7 |
| **P12** |  | 12.2 ± 2.4 | – |
| **P13** |  | 63.5 ± 34.9 | 5.9 |
| **P14** |  | 24.1 ± 4.0 | 2.7 |
| **P15** |  | 1.13 ± 0.08 | 11.5 |
| **P16** |  | 27.6 ± 8.4 | – |

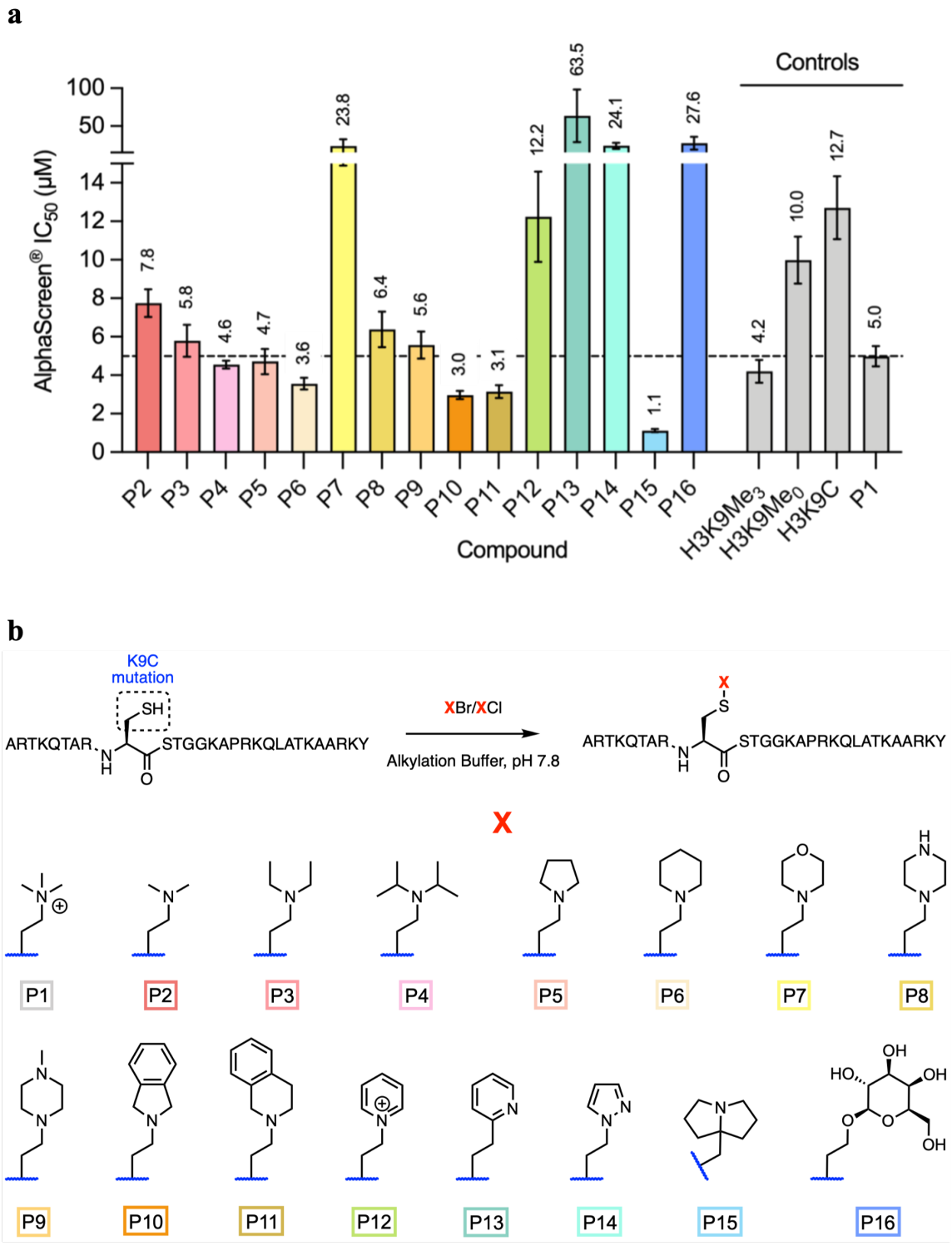



**Figure 2.9:** a) Bar graph showing the AlphaScreen® $IC_{50}$ values for inhibition of the interaction between biotinylated $H3_{(1–27)}K9Me_3$ and $His_6$-TRIM24. The dashed line represents the mean $IC_{50}$ value of peptide **P1** - the sulfur containing analogue of $H3_{(1-27)}K9Me_3$. Mean $IC_{50}$ values (in μM) from triplicate data are shown above the bars. Error bars indicate the standard deviation. Dose-response curves can be found in Appendix 4, Figure A4.4; b) Structures of the $H3K9Me_3$-mimicking peptides.

## 2.3 Acyl Hydrazone Dynamic Ligation Screen

### 2.3.1 Ligand-directed Dynamic Fragment Ligation

In the previous cysteine-alkylation strategy, $KMe_3$ mimics were predominantly designed with tertiary amine groups. However, it would be advantageous to move away from amine moieties towards scaffolds with improved physicochemical properties, such as increased cell permeability. Therefore, an alternative strategy was sought to investigate chemical space more broadly and identify additional functionalities that could interact with the TRIM24 PHD finger.

Ligand-directed dynamic fragment ligation using acyl hydrazone exchange, appeared a promising method for targeting 'undruggable' protein-protein interactions (PPIs), such as the H3/TRIM24 PHD interaction.[134,135] In this method, a truncated anchor peptide containing a *C*- or *N*-terminal hydrazone functionality, which is derived from the original interaction partner, is incubated with an aldehyde fragment in the presence of a protein template (Figure 2.10a). The reversible hydrazone exchange equilibrium will be shifted in favour of any high affinity peptide-fragment hybrids which bind preferentially to the protein (Figure 2.10b) over the lower affinity truncated anchor peptide, in accordance with Le Chatelier's principle. *In situ* screening using a fluorescence anisotropy competition assay (Figure 2.10b) subsequently allows identification of higher affinity hit compounds, which will displace the tracer, resulting in a lower anisotropy relative to the initial truncated peptide. Inspiration was taken from work by Hegedüs and coworkers, who used this approach to identify peptide-fragment hybrids that could inhibit the GKAP/SHANK1-PDZ protein-protein interaction (PPI).[135] Encouraged by the fact that this method was used to target a challenging β-strand/β-sheet mediated PPI, which shares similarity to the H3/TRIM24 PHD interaction, this methodology was chosen to screen for fragments that could interact with W828 in the TRIM24 PHD finger.

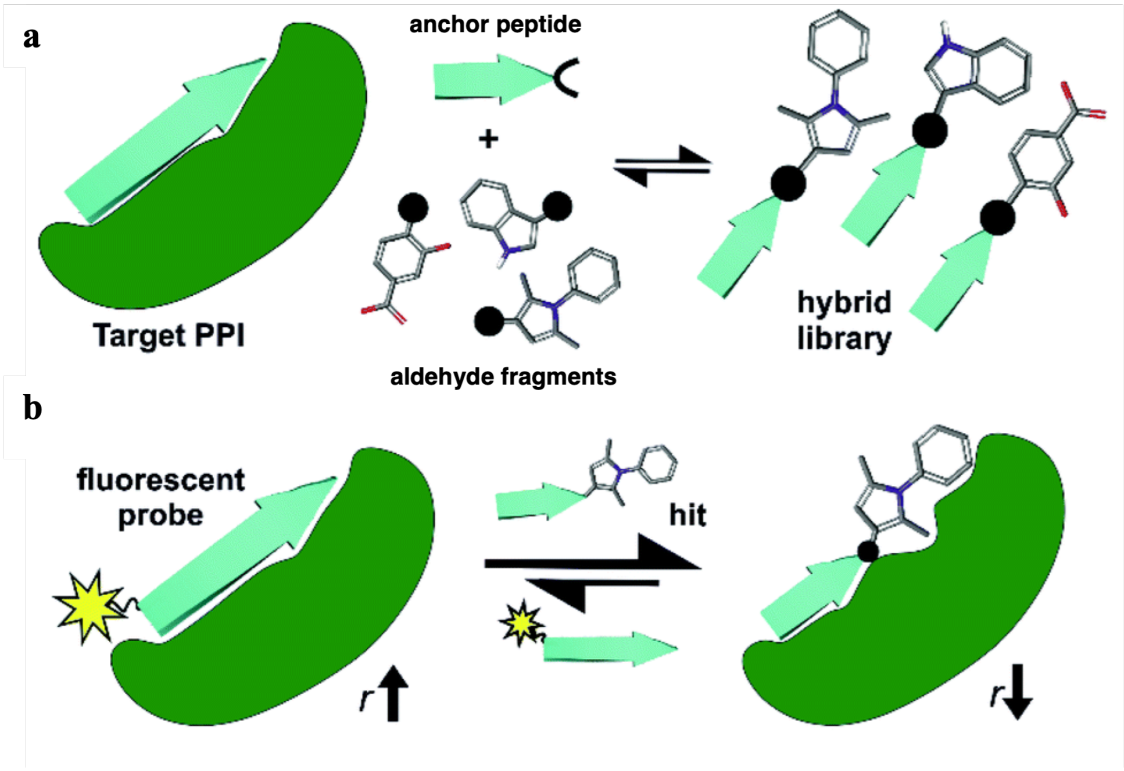



**Figure 2.10:** Principle of ligand-directed dynamic fragment ligation using acyl hydrazone exchange. a) Schematic showing the assembly of peptide-fragment hydrids from a hydrazone-functionalised truncated anchor peptide, derived from the endogenous interaction partner of the target protein; b) Schematic of the fluorescence anisotropy (*r*) competition assay. The peptide-fragment hybrid, formed *in situ*, competes with a fluorescently labelled tracer peptide for binding to the target protein. Figure reproduced and modified from Hegedüs *et al.* 2021.[135] Copyright 2021 The Royal Society of Chemistry.

### 2.3.2 Anchor Peptide Design

The anchor peptide for the dynamic ligation screen was derived from a $H3_{(1–9)}K9Me_3$ peptide (Figure 2.11). Truncation of the peptide at position R8 and replacement of $K9Me_3$ with a hydrazone (Figure 2.11) would enable the identification of fragments that could interact with W828 using the ligand-directed dynamic fragment ligation method. *N*-Terminal acetyl and *C*-terminal benzylimine groups (Figure 2.11) were incorporated to aid HPLC purification of the anchor peptide by increasing hydrophobicity.

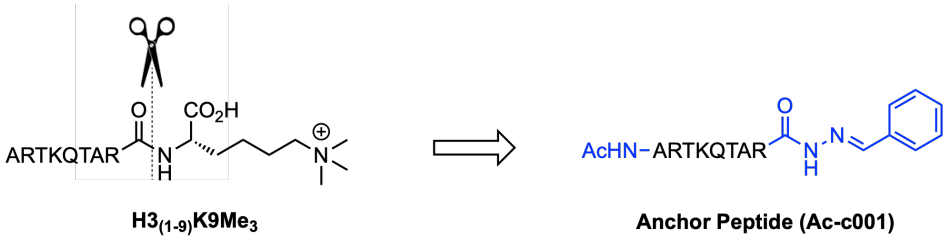



**Figure 2.11:** Design of the hydrazone-functionalised anchor peptide (Ac-c001) by truncation of $H3_{(1-9)}K9Me_3$, which binds to the TRIM24 PHD. The *N*-terminal acetyl and *C*-terminal acyl hydrazone groups are highlighted in blue.

After designing the anchor peptide (Ac-c001), it was necessary to verify its compatibility with the assay conditions. Ac-c001 contains a lysine residue at position 4, which could react with aldehydes to form a Schiff base during the screen and affect the results. To evaluate this a $H3_{(1-8)}$ peptide (exact mass = 930.54) was incubated in the assay buffer (pH 6.5) with excess benzaldehyde for 24 h. Pleasingly, after 24 h incubation LC-MS analysis revealed excellent stability with no peaks observed corresponding to addition of benzaldehyde (+ 88.03) or degradation of the hydrazone (Figure 2.12; Appendix 5, Figure A5.1). Similarly, TRIM24 was also stable under the assay conditions (Figure 2.13; Appendix 5, Figure A5.2) suggesting that both anchor peptide and TRIM24 protein are compatible with the dynamic ligation screen.

### 2.3.3 Tracer Design

The design of the tracer is another important aspect of the assay. This is because a higher affinity tracer gives a wider assay window, allowing a greater range of inhibitor affinities to be measured.[136] The affinity of the tracer is dependent on a number of factors including the fluorescent label, the linker, the attachment point, and the ligand. Fluoresceine isothiocyanate (FITC) was selected as the fluorescent label for the tracer as it is widely used in fluorescence assays and also has a long emission wavelength (535 nm), which reduces problems arising from protein background fluorescence (350 nm).[136]

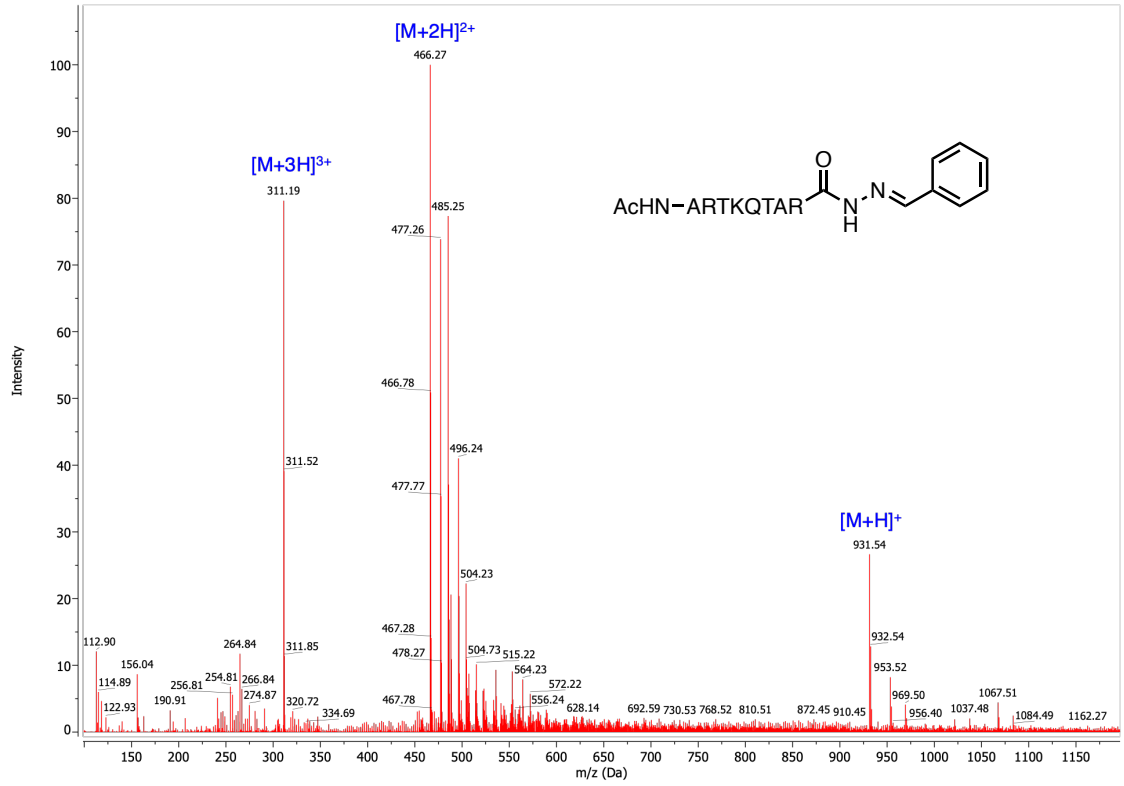



**Figure 2.12:** ESI$^+$ mass spectrum of the anchor peptide (structure shown) after 24 h incubation with benzaldehyde. Anchor peptide exact mass = 930.54 Da.

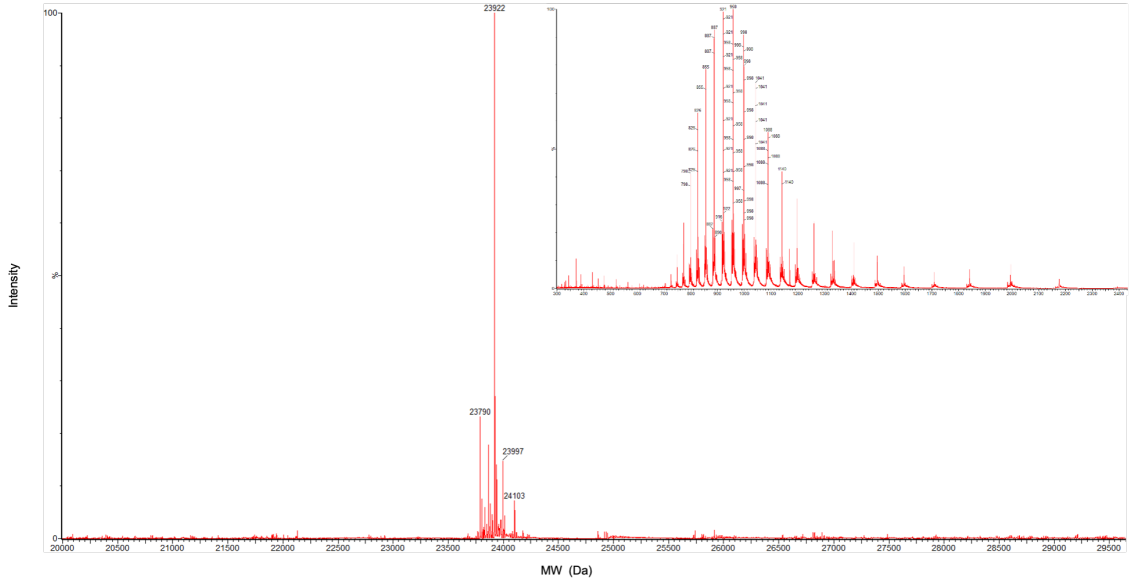



**Figure 2.13:** Deconvoluted and raw ESI$^+$ mass spectra of $His_6$-TRIM24 (PHD-BRD) after 24 h incubation with benzaldehyde. $His_6$-TRIM24 (PHD-BRD) MW = 23922 Da.

Initially the FITC fluorophore was attached to the *N*-terminus of a $H3_{(1\text{-}9)}K9Me_3$ peptide (tracer 1, Figure 2.14a) and included an Ahx (6-aminohexanoic acid) linker as a spacer to prevent the Edman degradation reaction during SPPS.[137] If the spacer were not included, the thiourea group could cyclise onto the terminal carbonyl of residue A1, leading to elimination of the FITC fluorophore and N-terminal residue under acidic conditions. However, unfortunately tracer 1 had low affinity for the TRIM24 PHD ($K_d$ > 100 μM, Figure 2.14b) and would require too much protein for it to be applicable in the dynamic ligation screen.

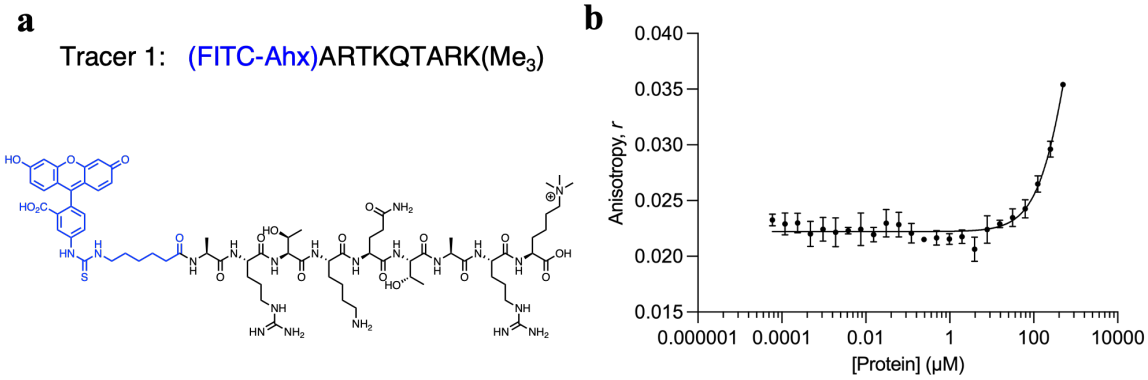



**Figure 2.14:** a) Structure and amino acid sequence for FITC tracer 1. The FITC fluorophore and Ahx linker are highlighted in blue; b) Fluorescence anisotropy curve for the direct titration of TRIM24 against FITC tracer 1 (100 nM). Anisotropy values were obtained in triplicate and are background subtracted. Error bars indicate the standard deviation.

It was hypothesised that the low affinity of tracer 1 could be due to the attachment position (*N*-terminus) and also the short length of the H3 amino acid sequence (9 residues). It has been reported that H3 does not tolerate modifications at position K4, and so attachment of FITC at this position was avoided.[1,2] Given the fact that a $H3_{(1\text{-}27)}K9Me_3K27$(biotin) peptide was used in the TRIM24 AlphaScreen® assays, it was hypothesised that residue K27 could be an alternative attachment point for the FITC fluorophore as biotin was tolerated at this position. Therefore, the tracer was redesigned with a longer H3 amino acid sequence and with the FITC fluorophore attached to the side-chain of K27 (Figure 2.15a).

Pleasingly, this tracer demonstrated considerably higher affinity in a fluorescence anisotropy direct titration assay ($K_d$ = 555 ± 16 nM, Figures 2.15b and 2.15c) and so was taken forward for use in the dynamic ligation screen. For the screen, the protein concentration was fixed around the $K_d$ of the tracer, [TRIM24] = 500 nM, to ensure sufficient signal, whilst also avoiding the effects of ligand depletion.

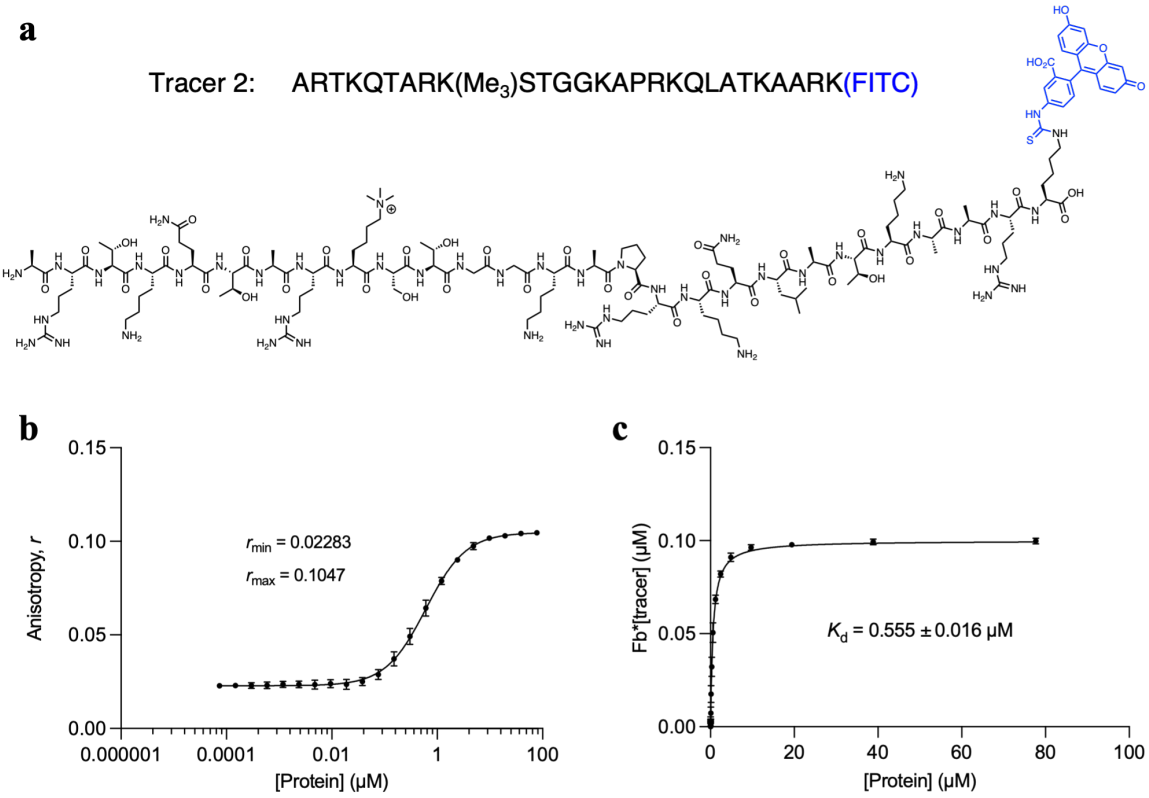



**Figure 2.15:** a) Structure and amino acid sequence for FITC tracer 2. The FITC fluorophore is highlighted in blue; b) Fluorescence anisotropy curve for the direct titration of TRIM24 against FITC tracer 2 (100 nM). Anisotropy values were obtained in triplicate and are background subtracted. Error bars indicate the standard deviation; c) Direct titration assay binding saturation curve for FITC tracer 2. $K_d$ values are quoted as the mean of triplicate data ± standard error of the mean. Error bars indicate the standard deviation ($F_b$ = fraction ligand bound).

### 2.3.4 Anchor Peptide Competition Assay

Having identified a suitable tracer, we next wanted to evaluate the concentrations at which the anchor peptide (Ac-c001), synthesised by Yue Li (University of Birmingham), could outcompete the tracer for binding to the TRIM24 PHD. It was decided that a fluorescence

anisotropy competition assay would be used in which the anchor peptide was titrated at constant protein and tracer concentration. However, the anchor peptide struggled to outcompete the tracer, and only did so above a concentration of approximately 62.5 µM, as indicated by the decrease in anisotropy (Figure 2.16). This concentration is relatively high, presumably due to the high affinity of the tracer peptide and low affinity of the anchor peptide. Nonetheless this anchor peptide could still be used for the dynamic ligation screen, as any higher affinity peptide-fragment hybrids would shift the curve to the left and reduce the anisotropy relative to the anchor peptide control, allowing hits to be identified (Figure 2.16). Therefore, it was decided that the anchor peptide concentration would be fixed at 62.5 µM for the dynamic ligation screen.

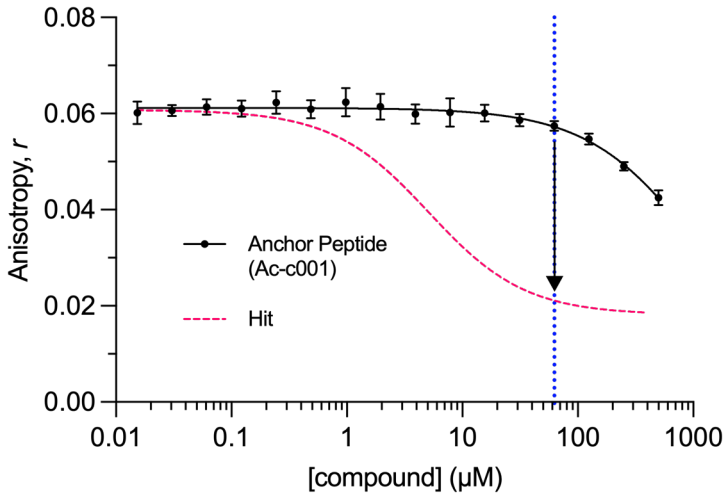
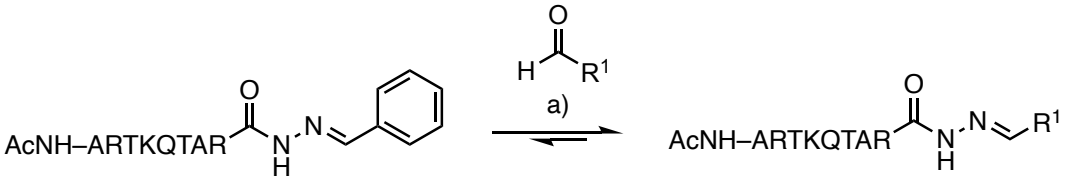



**Figure 2.16:** Competition assay dose-response curve for titration of the anchor peptide (Ac-c001). Error bars indicate the standard deviation of triplicate data. The pink dashed line represents a hypothetical curve for a higher affinity hit. The blue dotted line represents a concentration of 62.5 µM.

### 2.3.5 Dynamic Ligation Screen

With a suitable anchor peptide and tracer in hand, the dynamic ligation screen could now be carried out using 5 diverse libraries comprising a total of 165 aldehyde fragments (Appendix 5, Table A5.1) provided by the lab of Prof. Andrew Wilson. Each aldehyde was added to wells containing TRIM24 and the anchor peptide. The hydrazone exchange

reactions (Scheme 2.11), carried out in duplicate, were allowed to reach equilibrium for 24 h, before adding FITC tracer 2 to enable analysis by fluorescence anisotropy. *N*-Terminally acetylated $H3_{(1\text{-}9)}K9Me_3$ (Ac-$H3_{(1–9)}K9Me_3$) and a buffer control (protein + tracer only) were used as the positive and negative controls, respectively. Other controls included an anchor peptide control (Ac-c001 + protein + tracer) and a benzaldehyde control (Ac-c001 + benzaldehyde + protein + tracer) to validate that any benzaldehyde produced from the hydrazone exchange reactions does not interfere with the assay. After subtracting background fluorescence and normalising anisotropies to the negative buffer control, a hit threshold ($r_{rel,+} + 3\sigma_+$) was then defined in terms of the relative anisotropy of the positive control ($r_{rel,+}$) and the standard deviation of the positive control ($\sigma_+$).




**Scheme 2.11:** General representation of the acyl hydrazone exchange reaction in the dynamic ligation screen. *Reagents and conditions*: a) 500 nM TRIM24 (PHD-BRD), $R^1CHO$, 50 mM aq. $NH_4OAc$, 10 mM aqueous aniline, pH 6.5, rt, 24 h. ($R^1CHO$ = aldehyde fragment).

Interestingly, the screen identified a number of compounds with reduced anisotropy relative to the negative buffer control and anchor peptide (Figure 2.17; Appendix 5, Figure A5.6). This suggested that the products of the acyl hydrazone exchange reactions had increased affinity for TRIM24 and could displace more tracer. Only one fragment (**A033**) fell below the hit threshold that was set-up (Figure 2.17) and, unfortunately, due to time constraints on the project, could not be validated further in an orthogonal assay. Nonetheless it is not unreasonable to envision the aminothiazole moiety of **A033** interacting with W828 in the TRIM24 PHD through cation-π (when protonated) or π-π interactions. Therefore, in future this hit could prove useful in developing additional inhibitors for the TRIM24 PHD.

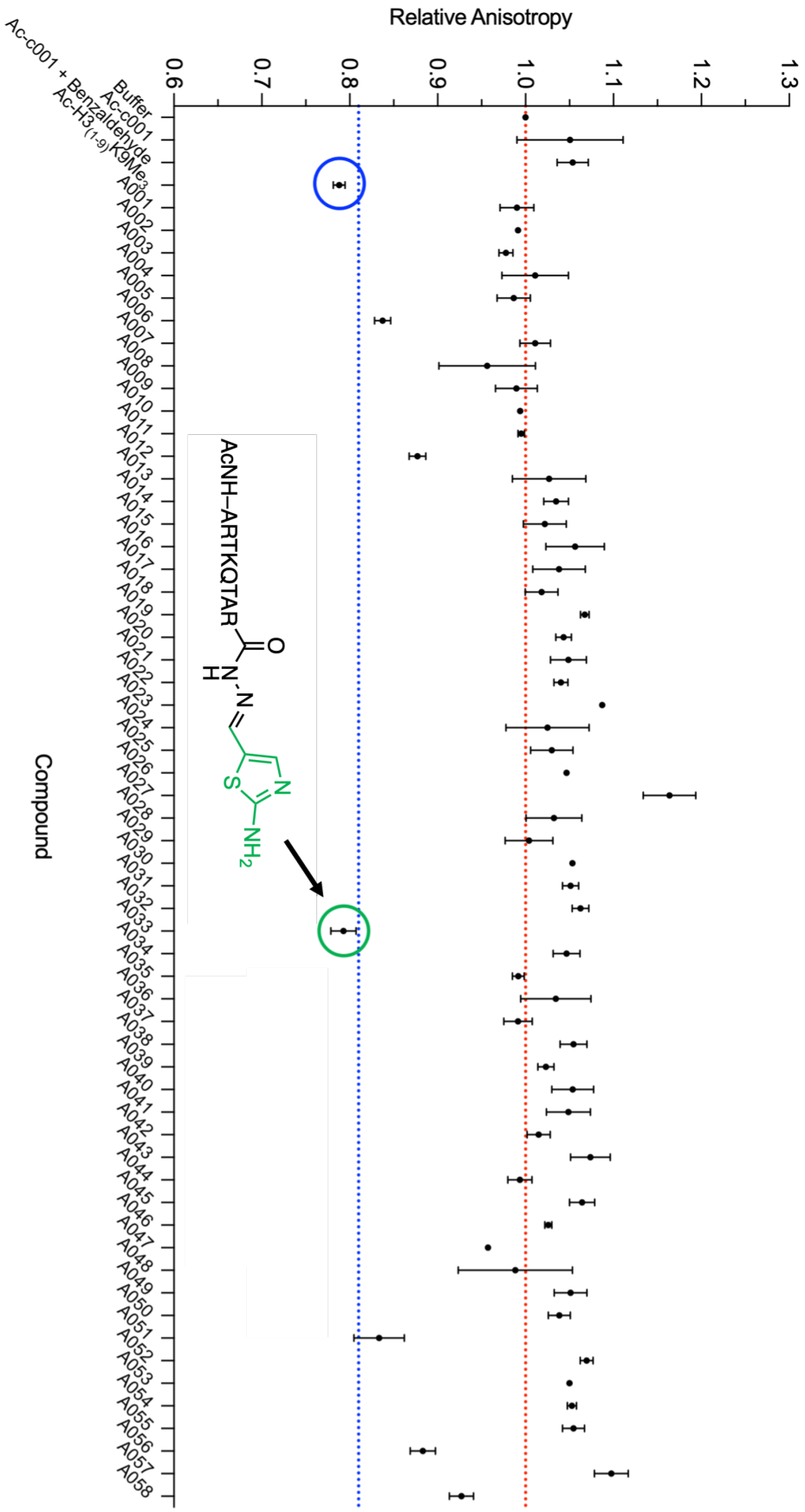



**Figure 2.17:** Results of the dynamic ligation screen for aldehyde fragments **A001-058**. The red dashed line represents the negative buffer control and the blue dashed line represents the hit threshold (relative anisotropy of positive control + [3 × standard deviation of the positive control]). Error bars indicate the standard deviation of duplicate data. The positive control is circled in blue. Hit fragment **A033** is circled in green. Raw fluorescence data can be found in Appendix 5, Figures A5.4 and A

## 2.4 Summary

In summary, a series of histone-mimicking peptides were synthesised using a combination of cysteine-selective alkylation and acyl-hydrazone ligand-directed dynamic fragment ligation. These peptides were screened for binding to the TRIM24 PHD using AlphaScreen® and fluorescence anisotropy assays. In the cysteine alkylation screen, novel $KMe_3$ mimics were identified, including isoindoline (**P10**) and pyrrolizidine (**P15**), which showed increased affinity relative to $KMe_3$. These mimics were chosen to be taken forward for incorporation into dual TRIM24 PHD and BRD inhibitors. An aminothiazole hit compound (**A033**) was also identified in the hydrazone dynamic ligation screen which could prove useful for future work in developing TRIM24 PHD inhibitors with improved physicochemical properties.

# 3 Synthesis of TRIM24 Peptide-Drug Conjugates

## 3.1 Introduction and Aims

To simultaneously inhibit the TRIM24 PHD and BRD, it was decided that a TRIM24 BRD ligand would be linked to a PHD-binding, histone-mimicking, peptide to generate a series of peptide-drug conjugates (PDCs). The TRIM24 BRD ligand from the dTRIM24 PROTAC was chosen for the BRD-binding section as it offered a solvent-exposed exit vector from the benzenesulfonamide ring for the attachment of linkers (Figure 3.1).[92,138] This vector also points towards W828 in the PHD domain (Figure 3.1). The histone-mimicking peptide was designed with an unnatural amino acid, containing one of the $KMe_3$ mimics identified in Chapter 2 (2.2.5, Figure 2.9), that could interact with W828. A $H3_{(1-8)}$-mimicking peptide was also included to engage with other key residues in the PHD, including N825, E826, and D827, to increase affinity. The syntheses of the PDCs and control compounds are described in this chapter.

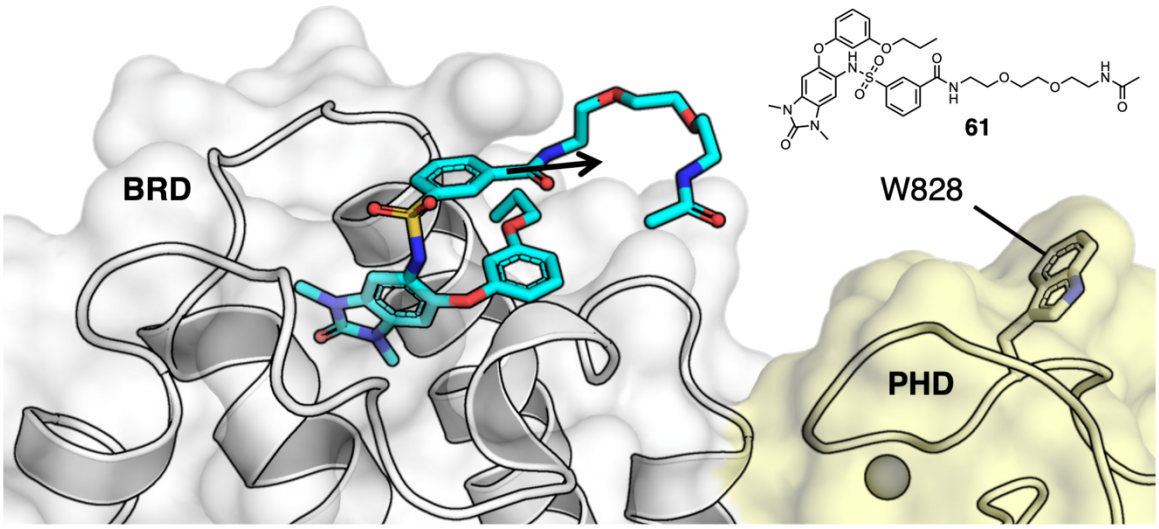
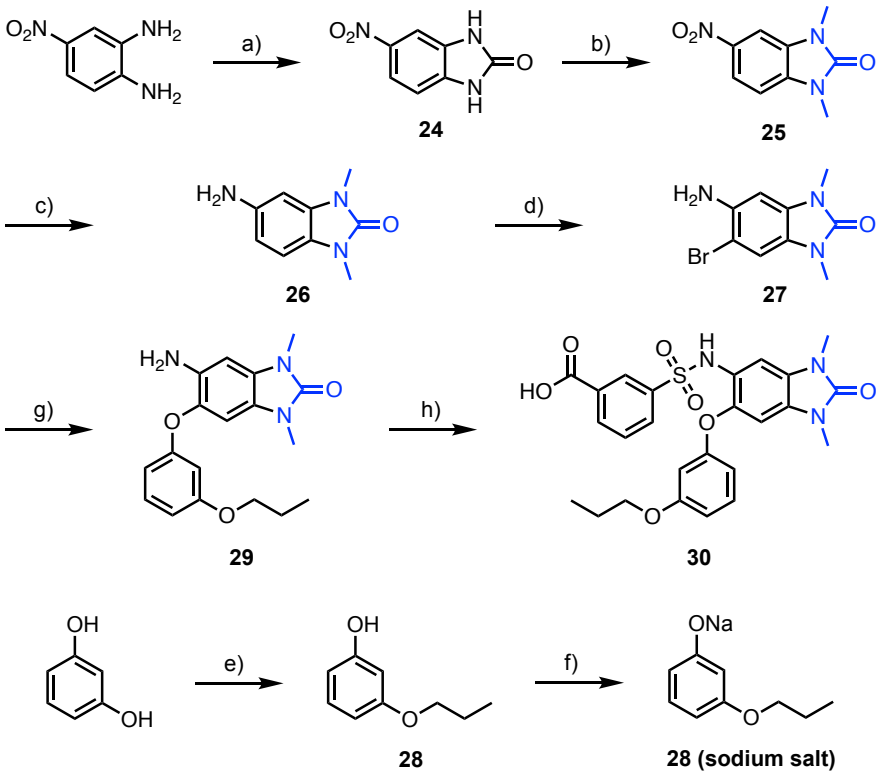



**Figure 3.1:** New crystal structure (1.5 Å, PDB ID: 9GDG) of TRIM24 (PHD-BRD) in complex with TRIM24 BRD ligand **61**, derived from the dTRIM24 PROTAC.[92,138] The structure shows an exit vector from the benzenesulfonamide ring (black arrow), that points towards W828 in the TRIM24 PHD. The TRIM24 PHD and BRD are shaded in yellow and white, respectively. Structure obtained by Michael Platt and Ekaterina Kot (Centre for Medicines Discovery, University of Oxford).

## 3.2 TRIM24 BRD Ligand Synthesis

TRIM24 BRD ligand **30** was synthesised using an adapted patent procedure in 21% yield over six steps (Scheme 3.1).[139] Firstly, 4-nitro-*o*-phenyldiamine was reacted with CDI to form the benzimidazolone core (compound **24**) of the BRD ligand in 99% yield (Scheme 3.1). The KAc mimic (Scheme 3.1, highlighted in blue) was then installed by methylation with MeI, to afford compound **25** in 97% yield (Scheme 3.1). In the next step, to improve the yield for reduction of the nitro group, a small excess of conc. aqueous HCl was added to protonate the aniline product and prevent poisoning of the hydrogenation catalyst. This reduction, followed by a basic wash, yielded compound **26** in 97% yield (Scheme 3.1).




**Scheme 3.1:** TRIM24 BRD ligand **30** synthesis. The KAc mimic is shown in blue. *Reagents and conditions*: a) CDI, DMF, rt, 20 h, 99%; b) $K_2CO_3$, MeI, DMF, 0 °C–rt, 16 h, 97%; c) Pd/C (10 mol%), $H_{2(g)}$, conc. aqueous HCl, EtOH, rt, 16 h, 97%; d) $Br_2$, $CHCl_3$, AcOH, −20 °C, 20 min, 73%; e) $CH_3CH_2CH_2Br$, $K_2CO_3$, DMF, rt–50 °C, 16 h, 69%; f) NaOH, MeOH, rt, 1 h, 100%; g) Compound **28** sodium salt, $Cs_2CO_3$, CuI (10 mol%), 8-hydroxyquinloline, diglyme, 130 °C, 72 h, 62%; h) 3-(Chlorosulfonyl)benzoic acid, pyridine, $CH_2Cl_2$, rt, 24 h, 71%.

In the next step bromination can occur at both the 4- and 6-positions due to *ortho*/*para* directing effects from the aniline and amide functional groups (Scheme 3.1). However, at low temperature the 6-bromo derivative (compound **27**) was formed in 73% yield (Scheme 3.1). It is presumed that this product was formed selectively due to greater steric hinderance at the 4-position. The sodium salt of compound **28**, formed from the monoalkylation of resorcinol (Scheme 3.1), was coupled to aryl bromide **27** to afford biaryl ether **29** in 62% yield (Scheme 3.1). Previous work in the Conway group demonstrated that the phenoxide salt of compound **28** (Scheme 3.1) was crucial for the success of the Ullmann reaction (Scheme 3.1).[140,141] Finally, compound **29** was coupled to 3-(chlorosulfonyl)benzoic acid, to obtain TRIM24 BRD ligand **30** in 71% yield (Scheme 3.1).

## 3.3 Unnatural Amino Acid Synthesis

### 3.3.1 Pyrrolizidine

Pyrrolizidine was identified as the highest affinity $KMe_3$ mimic in the previous chapter (2.2.5, Figure 2.9). Initial efforts were, therefore, focussed on incorporating this moiety into an unnatural amino acid. Although side chain sulfur oxidation was not observed for peptides in the previous chapter, it could still pose stability issues for the PDCs. Therefore, a cysteine alkylation strategy was avoided and alternative methods were used for the incorporation of $KMe_3$ mimics into unnatural amino acids.[5] A retrosynthesis was proposed, involving olefin metathesis as a key step, starting from an allyl glycine amino acid and an alkene-functionalised pyrrolizidine (Figure 3.2). A patent procedure was followed to synthesise the pyrrolizidine metathesis partner from dicyclopropylketone (Scheme 3.2).[142]

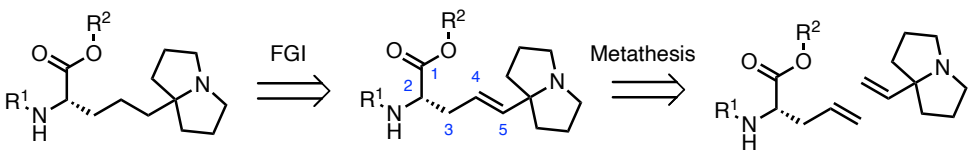
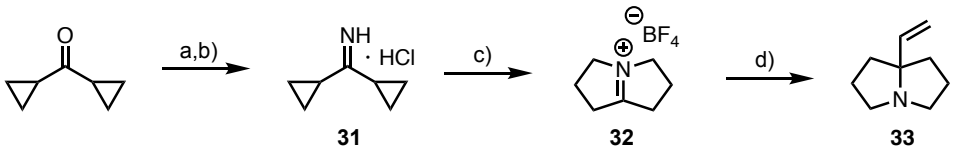


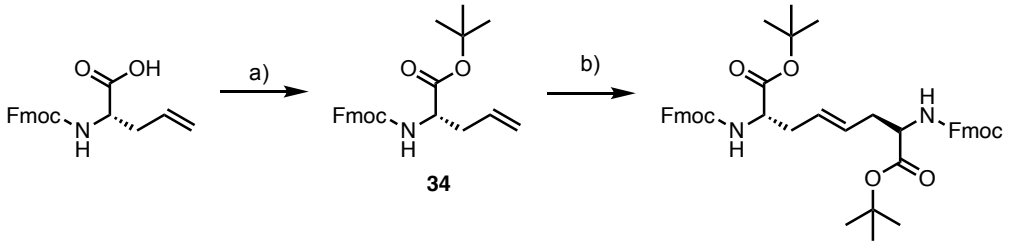

**Figure 3.2:** Retrosynthesis 1 for the synthesis of a pyrrolizidine-containing unnatural amino acid.

Initially the ketone group was activated using $TiCl_4$ and ammonia gas was passed through the reaction mixture to form iminium chloride **31**, in 48% yield over two steps (Scheme 3.2). Compound **31** could then undergo a thermally-promoted double Cloke-Wilson type rearrangement to form bicyclic iminium salt **32** in 39% yield (Scheme 3.2).[143,144] Pyrrolizidinium salt **32** was reacted with vinylmagnesium bromide, following a procedure analogous to that reported by Miyano *et al.*, to afford compound **33** in 63% yield (Scheme 3.2).[145]




**Scheme 3.2:** Synthesis of pyrrolizidine **33**. *Reagents and conditions*: a) $TiCl_4$, $NH_{3(g)}$, benzene, $CH_2Cl_2$, −10–0 °C, 1 h; b) HCl in dioxane (4 M), rt, 1 h, 48% over 2 steps; c) $NH_4BF_4$, DMSO, 100 °C, 6 h, 39%; d) Vinyl magnesium bromide, THF, 0–60 °C, 20 h, 63%.

The other alkene metathesis partner (compound **34**) was synthesised by $^t$Bu protection of Fmoc-L-allyl glycine, in 69% yield (Scheme 3.3). Metathesis of alkenes **33** and **34**, however, proved unsuccessful under a variety of conditions (Table 3.1), predominantly resulting in homodimer formation of amino acid **34** (Scheme 3.3 and Table 3.1). Amines are known to interfere during metathesis reactions and so there was concern that the tertiary amine of pyrrolizidine **33** could cause degradation of the Grubbs catalyst.[146] Therefore, *p*-TsOH was added to protect the amine as an organic-soluble quaternary ammonium salt.[147] However, these conditions also failed to yield the desired cross-metathesis product.




**Scheme 3.3:** Attempted synthesis of a pyrrolizidine-containing unnatural amino acid. *Reagents and conditions*: a) *tert*-Butyl 2,2,2-trichloroacetimidate, $BF_3{\cdot}OEt_2$, cyclohexane, $CH_2Cl_2$, rt, 6 h, 69%; b) compound **33**, compound **34**, see Table 3.1 for other conditions investigated for this reaction.

**Table 3.1:** Reagents and conditions screened for the metathesis reaction of compounds **33** and **34** (Scheme 3.3b). Reaction monitoring was conducted using LC-MS for a period of 24 h.

| Catalyst (5 mol%) | Solvent | Temperature (°C) | Additive | Major Product |
|---|---|---|---|---|
| Grubbs II | $CH_2Cl_2$ | 40 | – | No reaction |
| | Toluene | 100 | – | No reaction |
| | $CH_2Cl_2$ | 40 | *p*-TsOH | No reaction |
| | Toluene | 100 | *p*-TsOH | No reaction |
| Hoveyda-Grubbs II | $CH_2Cl_2$ | 40 | – | Homodimer |
| | Toluene | 100 | – | Homodimer |
| | $CH_2Cl_2$ | 40 | *p*-TsOH | Homodimer |
| | Toluene | 100 | *p*-TsOH | Homodimer |

It was hypothesised that a steric clash, between the pyrrolizidine ring of compound **33** and the Grubbs catalyst, could be hindering formation of the cross-metathesis product. Therefore, an alternative route was pursued, where the alkene was shifted further away from the pyrrolizidine ring, to alleviate steric hindrance during the metathesis reaction (Figure 3.3). To synthesise the alkene-containing amino acid (compound **36**), protected aspartic acid was reduced, chemoselectively, using $BH_3{\cdot}SMe_2$, to afford compound **35** in 57% yield (Scheme 3.4).[148] A Grieco elimination was then used to convert the resulting primary alcohol to an alkene, to obtain compound **36** in 43% yield (Scheme 3.4).[149]

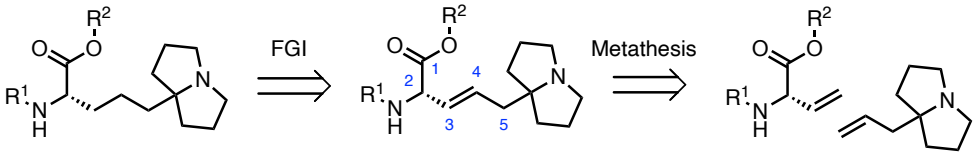
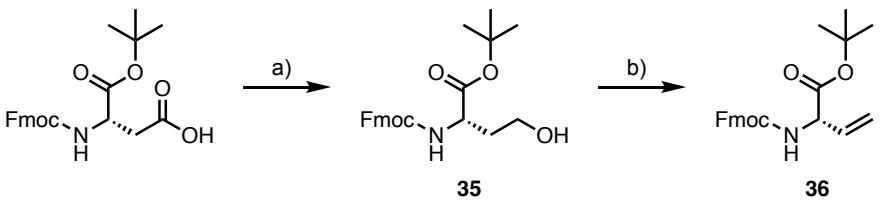
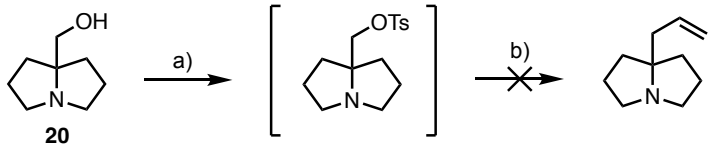



**Figure 3.3:** Retrosynthesis 2 for the synthesis of a pyrrolizidine-containing unnatural amino acid.



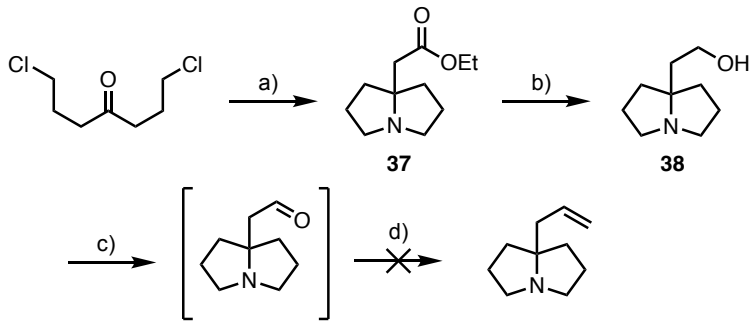

**Scheme 3.4:** Synthesis of alkene **36**. *Reagents and conditions*: a) $BH_3{\cdot}SMe_2$, THF, 0 °C–rt, 24 h, 57%; b) $PBu_3$, 2-nitrophenyl selenocyanate, THF, rt, 3 h; then aqueous $H_2O_2$, 0 °C–rt, 2 h, 43%.

Problems with sourcing an allymagneisum Grignard reagent meant it was necessary to develop alternative strategies to synthesise the pyrrolizidine metathesis partner. Initially a two-step procedure involving mesylation (Scheme 3.5) and a copper-catalysed Grignard cross-coupling reaction (Scheme 3.5) was attempted. However, this resulted in a complex mixture of products, from which the desired alkene could not be isolated.




**Scheme 3.5:** Attempted synthesis of an allyl-substituted pyrrolizidine using a Cu-catalysed Grignard coupling reaction. *Reagents and conditions*: a) TsCl, DMAP, $NEt_3$, $CH_2Cl_2$, 0 °C–rt, 1 h; b) Vinyl magnesium bromide, CuI (10 mol%) or $Li_2CuCl_4$ (5 mol%), THF, −78 °C–rt, 24 h.

A second strategy was proposed involving a final step Wittig reaction, to install the alkene (Scheme 3.6). Firstly, 1,7-dichloroheptan-4-one was converted to compound **37** in 82% yield, using a procedure first reported by Oka *et al.* (Scheme 3.6).[150] The imine formed undergoes double 5-*exo-tet* cyclisation to yield a bicylic iminium species, which is then trapped by the product of the decarboxylation of ethyl potassium malonate. Reduction of compound **37** to the aldehyde with DIBAL-H was attempted, however, this resulted in over reduction to the alcohol **38** (Scheme 3.6). Therefore, the reaction temperature was increased to promote full conversion to the alcohol **38** in 90% yield. Oxidation back to the aldehyde was then achieved using a Swern-oxidation with oxalyl chloride (Scheme 3.6).[151] Due to instability of the aldehyde, the product was carried forward without further purification. Finally, a Wittig reaction was performed, however, this failed to produce the desired alkene.[152] Unfortunately, due to time constraints on the project, the decision was taken to abandon pyrrolizidine as a $KMe_3$ mimic and pursue a more synthetically tractable alternative.




**Scheme 3.6:** Attempted synthesis of an allyl-substituted pyrrolizidine using a Wittig reaction. *Reagents and conditions*: a) Ethyl potassium malonate, $NH_3$ in MeOH (7 M), 0 °C–rt, 24 h, 82%; b) DIBAL-H, toluene, $CH_2Cl_2$, 0 °C–rt, 6 h, 90%; c) DMSO, $(COCl)_2$, $CH_2Cl_2$, −78 °C, 2 h; then $NEt_3$, −78 °C–rt, 16 h; d) $MePPh_3Br$, $^n$BuLi, THF, 0 °C–rt, 24 h.

### 3.3.2 Isoindoline

The AlphaScreen® assay, described in section 2.1.5, also identified isoindoline as an effective $KMe_3$ mimic. Although this mimic had reduced affinity compared to pyrrolizidine, it still maintained higher affinity than $KMe_3$, and appeared significantly easier to functionalise. An initial single step alkylation strategy was carried out to synthesise an isoindoline-containing unnatural amino acid from lysine (Figure 3.4). However, alkylation of lysine with 1,2-*bis*-(bromomethyl)benzene and TBAI, as a nucleophilic catalyst, resulted in a complex mixture of products, which proved difficult to purify (Scheme 3.7).[153]

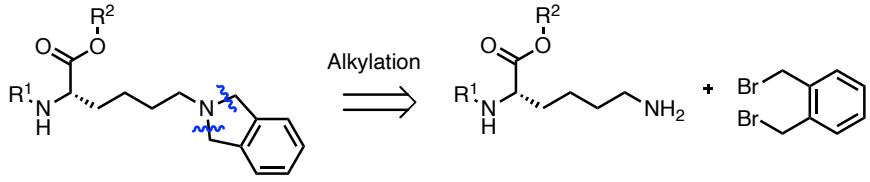
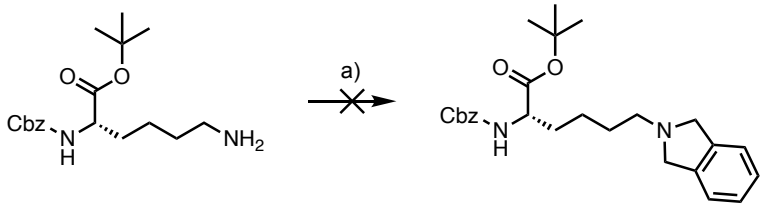



**Figure 3.4:** Retrosynthesis 1 for the synthesis of an isoindoline-containing unnatural amino acid.




**Scheme 3.7:** Attempted one-step synthesis of an isoindoline-containing unnatural amino acid from lysine. *Reagents and conditions*: a) 1,2-*bis*-(Bromomethyl)benzene, $K_2CO_3$, TBAI, DMF, rt, 20 h.

An alternative two-step retrosynthetic strategy was proposed, involving reductive amination (Figure 3.5). Firstly, the ε-amino group of the lysine side chain was converted to an alcohol in 38% yield, *via* a diazonium intermediate using sodium nitroprusside [$Na_2Fe(CN)_5NO$], as reported by Baldwin *et al.* (Scheme 3.8).[154] The low yield was due to the formation of a major side product, which was later identified as the alkene elimination product. Particular care was taken during the work-up for this reaction as acidification of the reaction mixture causes the release of HCN from the breakdown of excess sodium nitroprusside. Carboxylic acid **39** was then protected as the $^t$Bu ester using 2-bromo-2-methylpropane in 78% yield (Scheme 3.8).[155] However, subsequent oxidation of the hydroxyl group with substoichiometric TEMPO, and trichloroisocyanuric acid as a stoichiometric secondary oxidant (Scheme 3.8), formed a complex mixture of products.[156] This observation suggested that the aldehyde was either degrading or undergoing further reaction. Therefore, it was concluded that the aldehyde was too reactive to be used as an intermediate in this synthesis and that a different method should be pursued.

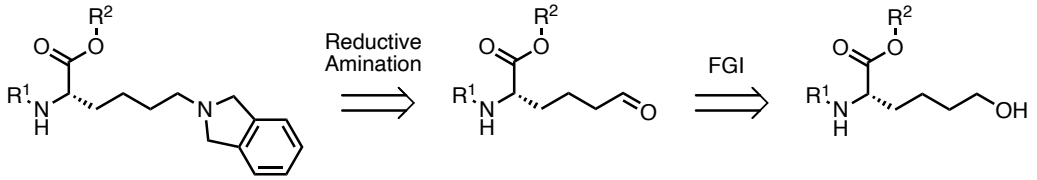
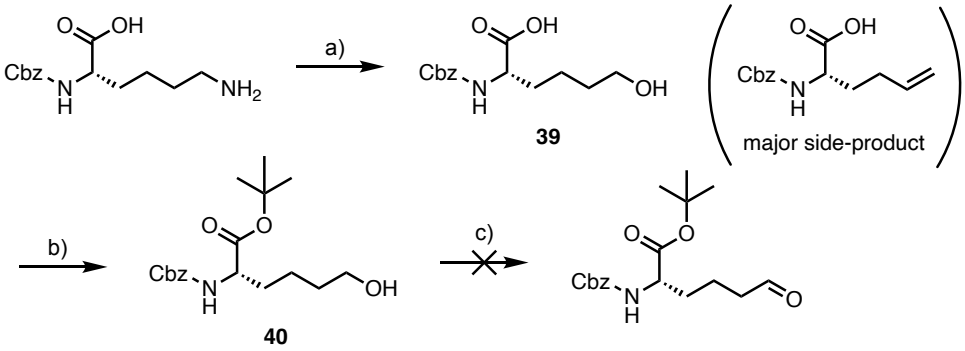



**Figure 3.5:** Retrosynthesis 2 for the synthesis of an isoindoline-containing unnatural amino acid.




**Scheme 3.8:** Synthesis of alcohol **40**. *Reagents and conditions*: a) $Na_2Fe(CN)_5NO$, aqueous NaOH (4 M), $H_2O$, pH 9–10, 60 °C, 6 h, 38%; b) 2-Bromo-2-methylpropane, $K_2CO_3$, BTEAC, MeCN, 55 °C, 24 h, 78%; c) TEMPO (5 mol%), trichloroisocyanuric acid, $CH_2Cl_2$, rt, 2 h.

A third retrosynthetic strategy was proposed involving a combination of olefin metathesis and alkylation (Figure 3.6). This method offered the benefit of avoiding the low yielding and potentially hazardous sodium nitroprusside diazotisation step from the previous strategy, whilst also utilising less reactive intermediates.

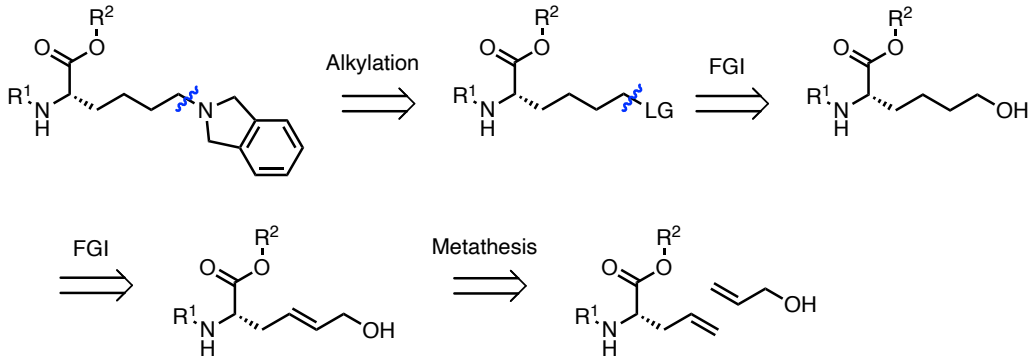
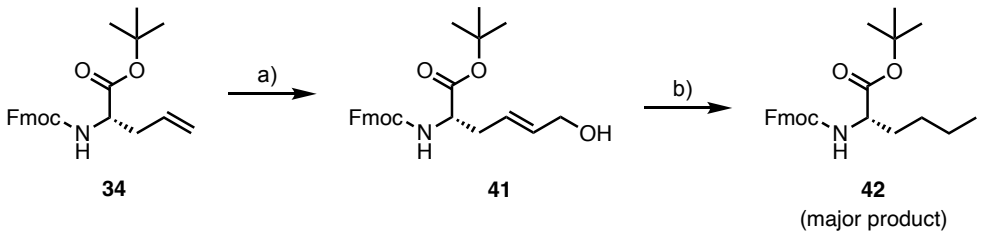
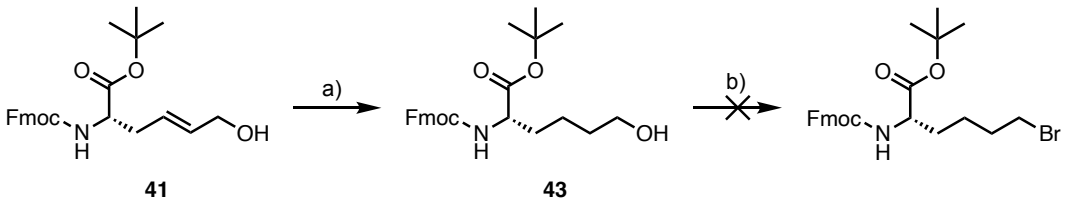



**Figure 3.6:** Retrosynthesis 3 for the synthesis of an isoindoline-containing unnatural amino acid (LG = leaving group).

Firstly, compound **34** was converted to allylic alcohol **41** in 70% yield using a metathesis reaction with Grubbs second-generation catalyst and *cis*-butene-1,4-diol (Scheme 3.9). This reaction is highly *E*-selective as indicated by the magnitude of the coupling constant ($J$ = 15.2 Hz) of the alkene protons. Hydrogenation of alkene **41** was then performed using a Pd/C catalyst, however, this resulted in dehydroxylation, as well as reduction of the double bond, to form compound **42** in 37% yield (Scheme 3.9). Fanning *et al.* reported a similar observation upon reduction of an analogous compound using Pd/C, however, they were able to overcome this by switching to a $PtO_2$ catalyst.[157] Pleasingly, by using $PtO_2$ as the hydrogenation catalyst, compound **43** was obtained in 75% yield (Scheme 3.10). Bromination of alcohol **43** was then attempted using standard Appel conditions with $CBr_4$ and $PPh_3$.[158] Although disappearance of the starting material was observed, the corresponding alkyl bromide could not be identified from the crude. It was hypothesised that cyclisation of the amide nitrogen onto the alkyl bromide may have occurred to form the six-membered ring.[157]

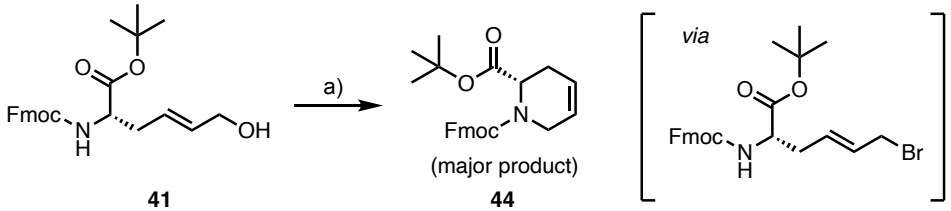



**Scheme 3.9:** Attempted alternative synthesis of an alcohol-containing unnatural amino acid. *Reagents and conditions*: a) *cis*-Butene-1,4-diol, Grubbs' II cat. (5 mol%), $CH_2Cl_2$, rt–40 °C, 6 h, 70%; b) Pd/C (10 mol%), $H_2$ (g), EtOH, rt, 16 h, 37%.




**Scheme 3.10:** Attempted synthesis of an alkyl bromide-containing unnatural amino acid. *Reagents and conditions*: a) $PtO_2$ (10 mol%), $H_2$ (g), EtOAc, rt, 20 h, 75%; b) $CBr_4$, $PPh_3$, $CH_2Cl_2$, 0 °C–rt, 24 h.

In an attempt to avoid cyclisation, the order of steps was reversed so that bromination was performed prior to the hydrogenation step, as it was anticipated that the *E*-double bond would prevent cyclisation to from the six-membered ring. Intriguingly, in spite of this, bromination of compound **41** afforded cyclised product **44** in 63% yield (Scheme 3.11).




**Scheme 3.11:** Attempted synthesis of an allyl bromide-containing unnatural amino acid. *Reagents and conditions*: a) $CBr_4$, $PPh_3$, $CH_2Cl_2$, rt, 0 °C–rt, 2 h, 63%.

For this reaction to take place an isomerisation of the double bond is required; therefore, a potential mechanism was proposed involving nucleophilic phosphine catalysis (Figure 3.7). Firstly, bromide is eliminated *via* an $S_N1$-like mechanism to form a resonance stabilised carbocation. The more stable secondary carbocation will likely be favoured due to increased hyperconjugative stabilisation. Phosphine could then function as a nucleophilic catalyst to provide further stabilisation of the carbocation. A 6-*endo-trig* cyclisation of the

amide nitrogen onto the double bond, likely *via* an $S_N2$' mechanism, could then occur to yield compound **44** containing the *Z*-alkene.

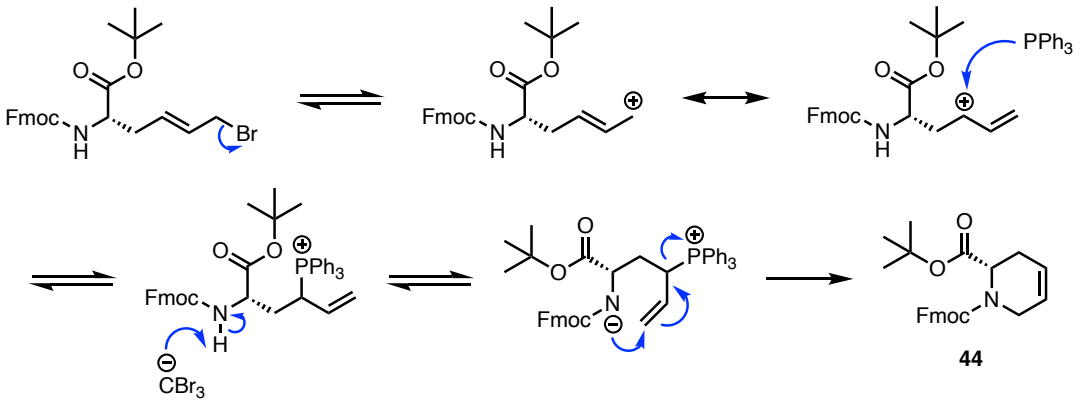
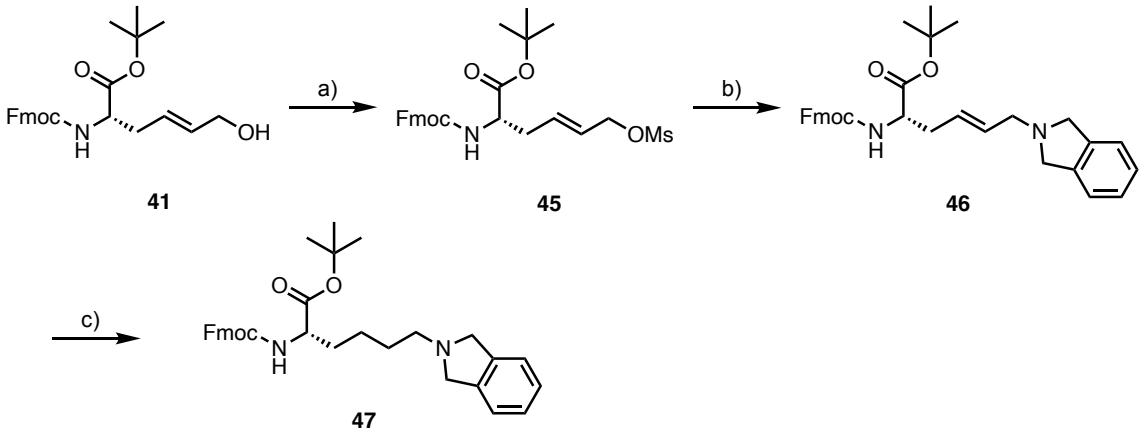



**Figure 3.7:** Proposed mechanism for the formation of compound **44**.

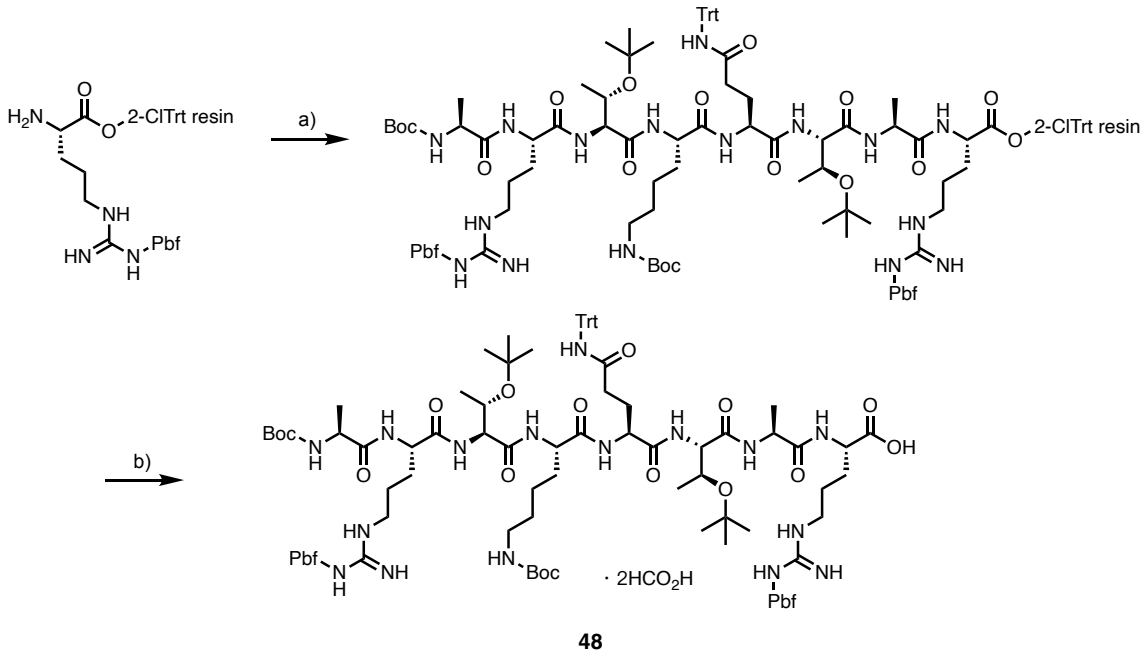
Based on the proposed mechanism (Figure 3.7), the presence of phosphine was hypothesised to drive the formation of the cyclised product, through stabilisation of the carbocation intermediate. To avoid the use of $PPh_3$, alcohol **41** was converted the mesylate (Scheme 3.12). This method pleasingly afforded compound **45** in 83% yield, with no cyclised product observed. Compound **45** was subsequently alkylated with isoindoline to obtain compound **46** in 71% yield (Scheme 3.12). Hydrogenation of the double bond, using a $PtO_2$ catalyst, afforded the desired isoindoline-containing amino acid **47** in 83% yield (Scheme 3.12).




**Scheme 3.12:** Synthesis of unnatural amino acid **47** from allyl alcohol **41**. *Reagents and conditions*: a) $MeSO_2Cl$, $NEt_3$, $CH_2Cl_2$, 0 °C–rt, 1 h, 83%; b) Isoindoline·HCl, $K_2CO_3$, KI, DMF, rt, 1 h, 71%; c) $PtO_2$ (10 mol%), $H_{2(g)}$, EtOAc, rt, 6 h, 83%.

## 3.4 Histone-mimicking Peptide Synthesis

Prior to assembling the TRIM24 PDCs, it was necessary to synthesise a peptide that could interact with key residues in the TRIM24 PHD. Solid phase peptide synthesis (SPPS) was performed to couple Fmoc-protected amino acids to 2-ClTrt-H-Arg(Pbf) resin from the *C*- to the *N*-terminus (Scheme 3.13). A coupling mixture of DIC and Oxyma Pure was used as it supresses racemisation and also has a lower explosion risk than benzotriazole-based coupling reagents.[159] Reactive amino acid sidechains were masked with acid-labile protecting groups (Boc, Pbf, Trt, $^t$Bu) to reduce side reactions during SPPS and the subsequent PDC assembly. 2-ClTrt resin was chosen as it allows a cleavage step in mild acid, without removing the acid-labile protecting groups.[160] Resin cleavage was achieved with HFIP to afford protected peptide **48** (Scheme 3.13), which is primed for solution phase amide coupling reactions to form the PDCs.[161]

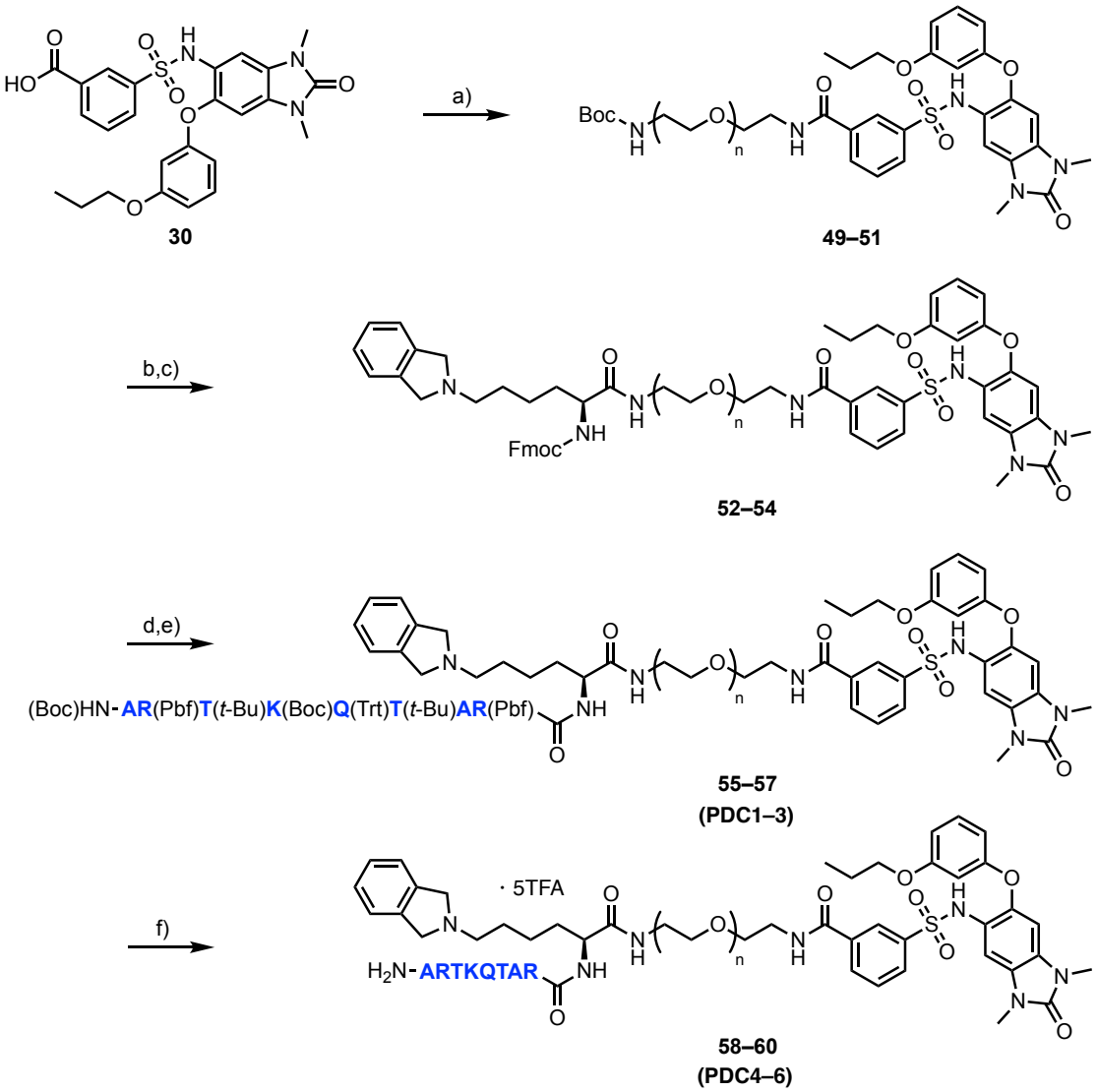



**48**

**Scheme 3.13:** Solid phase peptide synthesis (SPPS) of protected histone 3-mimicking peptide **48** using a Liberty Blue™ peptide synthesiser. *Reagents and conditions*: a) (Coupling) Fmoc-protected amino acid, DIC, Oxyma Pure, DIPEA, DMF, rt; (Fmoc deprotection) piperidine, DMF, rt; b) (Resin cleavage) HFIP, $CH_2Cl_2$, rt, 1 h.

## 3.5 Assembly of Peptide-Drug Conjugates

A series of TRIM24 PDCs with varying linker lengths were assembled from their constituent H3 peptides, isoindoline-containing amino acids, and TRIM24 BRD ligands using a six-step synthesis (Scheme 3.14).




**Scheme 3.14:** Assembly of TRIM24 PDCs 4–6 (compounds **58**–**60**). *Reagents and conditions*: a) HATU, DIPEA, DMF, rt, 24 h, 66% (n = 2), 80% (n = 3), 88% (n = 4); b) Compound **47**, HCl, dioxane, rt, 24 h; c) HATU, DIPEA, DMF, rt, 24 h, 64% (n = 2), 49% (n = 3), 77% (n = 4); d) Piperidine, DMF, rt, 1 h; e) Peptide **48**, HATU, DIPEA, $CH_2Cl_2$, rt, 24 h, 52% (n = 2), 57% (n = 3), 42% (n =4); f) TFA/TIPS/$H_2O$ (38:1:1), rt, 4 h, 32% (n = 2), 51% (n = 3), 50% (n = 4). Single letter amino acid codes are shown in **bold blue**.

Firstly, TRIM24 BRD ligand **30** was conjugated to flexible PEG linkers of varying lengths (n = 2–4), using HATU as the amide coupling reagent, to afford compounds **49**–**51** in 66–88% yield (Scheme 3.14). The Boc-protected amines of compounds **49**–**51** and $^{t}$Bu ester of amino acid **47** were then deprotected using 4 M HCl in dioxane (Scheme 3.14), before being subjected to HATU amide coupling conditions to form compounds **52**–**54** in 49–77% yield (Scheme 3.14). HCl in dioxane was favoured over TFA for $^{t}$Bu ester/Boc deprotection as residual TFA could react in subsequent amide coupling reactions if not fully removed. The *N*-terminal Fmoc groups of compounds **52**–**54** were removed using piperidine (Scheme 3.14), before coupling the resulting free amines and protected H3 peptide **48** to obtain protected PDCs 1–3 (compounds **55**–**57**) in 42–57% yield (Scheme 3.14). Initially, the Fmoc deprotected compounds were reacted directly with the H3 peptide without purification, however, this resulted in residual piperidine coupling onto the *C*-terminus of the H3 peptide. An HPLC purification step was, therefore, introduced prior to the amide coupling reactions to avoid this. Pleasingly, a final global deprotection step with TFA and TIPS, as a cation scavenger, afforded PDCs 4–6 (compounds **58**–**60**) as the pentatrifluoroacetic acid salts in 32–51% yield (Scheme 3.14).

## 3.6 Control Compound Synthesis

Several monovalent control compounds (Figure 3.8), designed to bind to either the PHD or BRD (not both domains), were also synthesised as they could prove useful for investigating the binding mode of PDCs 4–6 (compounds **58**–**60**). The BRD controls were synthesised from compounds **49**–**51**, *via* Boc-deprotection and acetylation with AcCl, to afford compounds **61**–**63** in 43–72% yield (Scheme 3.15). For the synthesis of the PHD control peptide, Cbz-protected allyl-L-glycine was first converted to methyl amide **64** in 94% yield using HATU as the amide coupling reagent (Scheme 3.16). Olefin metathesis with Grubbs second-generation catalyst was used to install an allyl alcohol functional group (Scheme

3.16). This compound proved difficult to purify using either normal or reverse phase chromatography, and so the crude product was mesylated and alkylated with isoindoline (Scheme 3.16) to yield compound **65** as the HCl salt in 36% yield over 3 steps. Hydrogenation of the double bond and removal of the Cbz protecting group was achieved in single step using a Pd/C catalyst (Scheme 3.16). The resulting free primary amine was then coupled to H3 peptide **48** using HATU (Scheme 3.16) to obtain compound **66** as the triformate salt in 28% yield over 2 steps. Global deprotection with TFA, as previously described for the PDCs, afforded PHD control peptide **67** as the pentatrifluoroacetate salt in 70% yield (Scheme 3.16).

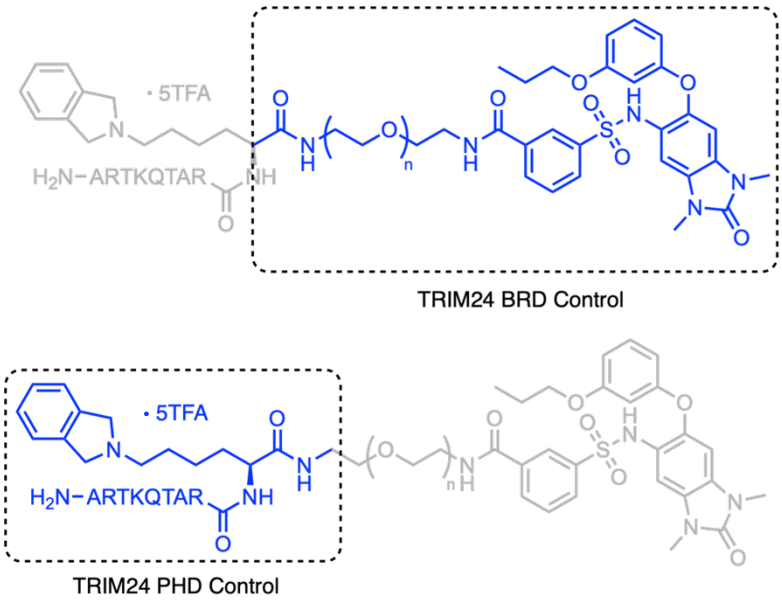
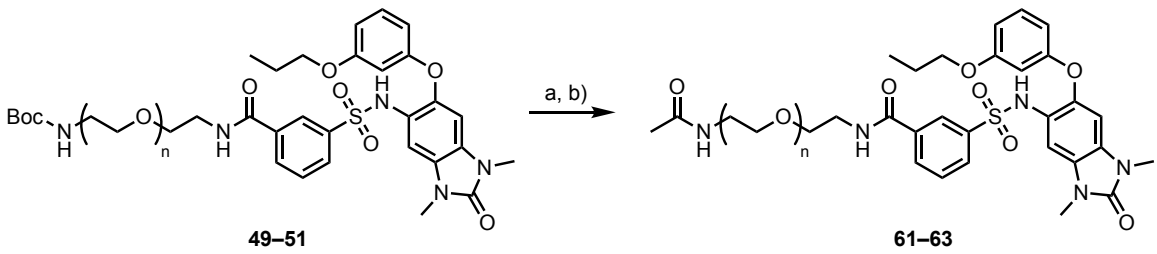
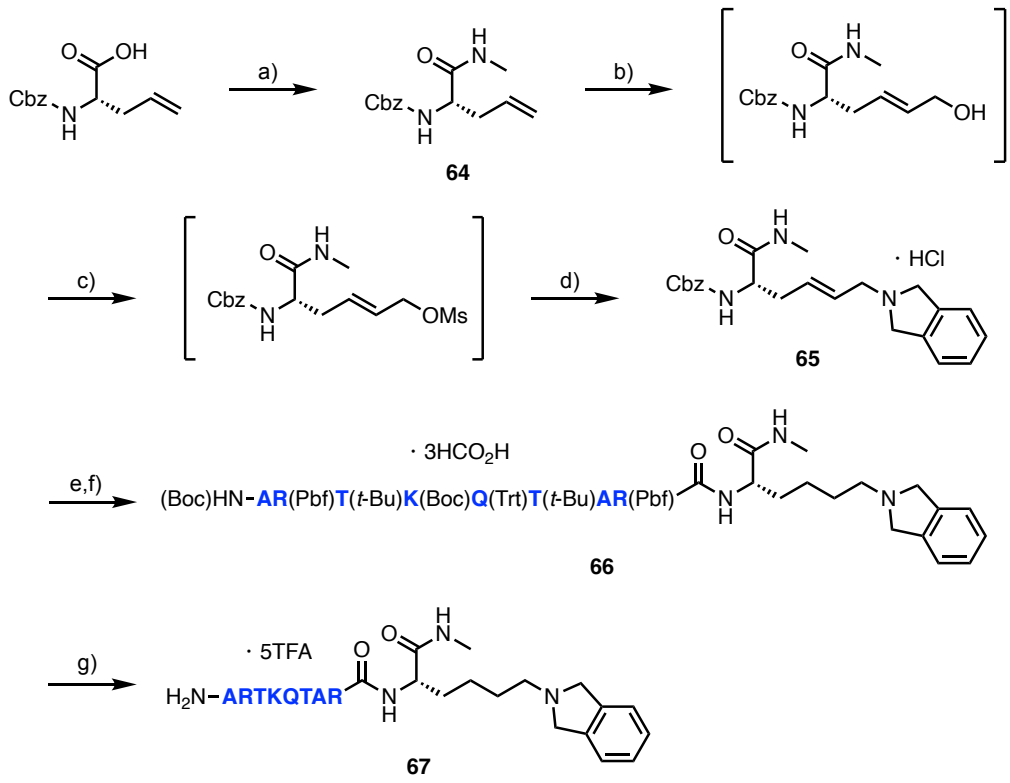



**Figure 3.8:** Design of the TRIM24 PHD and BRD control compounds.




**Scheme 3.15:** Synthesis of parent TRIM24 BRD control compounds **61**–**63**. *Reagents and conditions*: a) HCl in dioxane, rt, 2 h; b) AcCl, $NEt_3$, $CH_2Cl_2$, 0 °C–rt, 20 min, 72% (n = 2), 63% (n = 3), 43 % (n = 4).




**Scheme 3.16:** Synthesis of parent TRIM24 PHD control **67**. *Reagents and conditions*: a) $MeNH_2$, HATU, DIPEA, THF, DMF, rt, 2 h, 94%; b) *cis*-Butene-1,4-diol, Grubbs' II cat. (5 mol%), $CH_2Cl_2$, rt–40 °C, 6 h; c) $MeSO_2Cl$, $NEt_3$, $CH_2Cl_2$, 0 °C–rt, 1 h; d) Isoindoline, DMF, rt, 1 h, 36% (over 3 steps); e) 10% Pd/C (10 mol%), $H_{2(g)}$, MeOH, rt, 6 h; f) Peptide 48, HATU, DIPEA, $CH_2Cl_2$, rt, 6 h, 28%; g) TFA/TIPS/$H_2O$ (38:1:1), rt, 4 h, 70%. Single letter amino acid codes are shown in **bold blue**.

## 3.7 Summary

In summary, the first PDCs targeting the TRIM24 PHD and BRD have been designed and synthesised in 16 steps. This involved the synthesis of TRIM24 BRD ligand **30**, derived from the dTRIM24 PROTAC; a novel unnatural amino acid (compound **47**) containing the isoindoline $KMe_3$ mimic, identified in Chapter 2; and a protected H3 peptide (compound **48**), synthesised by SPPS on 2-ClTrt resin. Additionally, a number of control compounds, targeting either the BRD (compounds **61**–**63**) or PHD (compound **67**), were synthesised that were later used to interrogate the binding mode of the PDCs.

# 4 Biological Evaluation of Peptide-Drug Conjugates

## 4.1 Introduction and Aims

Epigenetic transcriptional regulation by the TRIM24 PHD and BRD has been implicated in breast cancer cell proliferation.[2] However, TRIM24 BRD inhibition has proven ineffective against cancer cells. For this project, it was hypothesised that simultaneous inhibition of the PHD and BRD might be required for antiproliferative effects to be observed. However, before testing this hypothesis using the newly synthesised PDCs, it was necessary to identify if they could simultaneously inhibit the TRIM24 PHD and BRD *in vitro*. Biophysical assays including AlphaScreen®, BROMO*scan*, and surface plasmon resonance (SPR) were used to characterise the binding of the PDCs to the TRIM24 PHD and BRD; evaluating affinities, selectivities, kinetics, and binding modes. The PDCs were then assessed in breast cancer cell lines to test the hypothesis and identify any phenotypic effects. The results of the biophysical and cellular assays for the PDCs and control compounds (Figure 4.1) are described in this chapter.

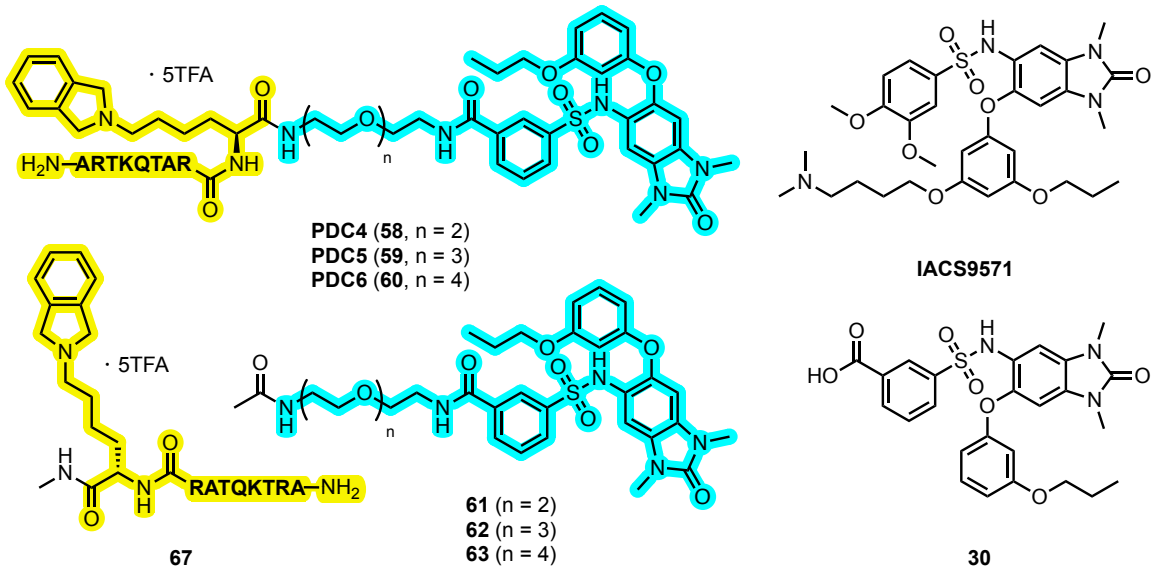



**Figure 4.1:** Structures of ligands evaluated in the biophysical and cellular assays. Single letter amino acid codes are shown on the structures in **bold**. IACS9571 is the highest affinity TRIM24 BRD ligand reported in the literature.[4]

## 4.2 Biophysical Characterisation

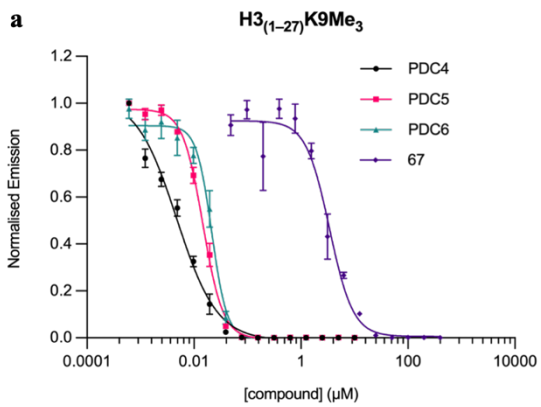
### 4.2.1 Investigation of TRIM24 PHD and BRD Inhibition

Firstly, AlphaScreen® competition assays, were carried out to investigate if the PDCs could displace biotinylated histone-mimicking peptides from the TRIM24 PHD and BRD, specifically $H3_{(1–27)}K9Me_3$ (Figure 4.2), $H3_{(1–27)}K18Ac$ (Figure 4.3), and $H3_{(1\text{-}27)}K9Me_3K18Ac$ (Figure 4.4). Unbiotinylated analogues of the competitor peptides were used as positive controls, and $H3_{(1–15)}K4Me_3$ was used as a technical negative control as it has previously been shown that K4 methylation abolishes binding to TRIM24.[1,2] Signal-response assays were conducted to verify the functional activity of the PHD and BRD in recombinantly expressed $His_6$-TRIM24 (Appendix 4, Figure A4.1) and a TruHits assay was performed to exclude assay interference (Appendix 4, Figure A4.3).

PDCs 4–6 (compounds **58**–**60**) demonstrated low nanomolar $IC_{50}$ values (4.7–21.5 nM) for displacement of the PHD-binding biotinylated $H3_{(1–27)}K9Me_3$ peptide (Figure 4.2), with a general trend of decreasing $IC_{50}$ with decreasing PEG linker length. PDCs 4–6 showed substantially lower $IC_{50}$ values than the positive control (5.03 μM) and parent PHD ligand **67** (3.41 μM). This was particularly emphasised by PDC4 which showed 1070- and 726-fold decreases in $IC_{50}$ compared to the positive control and compound **67**, respectively.

In the $H3_{(1–27)}K18Ac$ assay (Figure 4.3) the PDCs were compared to the current highest affinity TRIM24 BRD ligand in the literature, IACS9571 ($K_d$ = 31 nM);[4] and parent BRD ligands **61**–**63**. Pleasingly, the PDCs demonstrated lower $IC_{50}$ values relative to IACS9571 and compounds **61**–**63** (Figure 4.3b). This was exemplified by PDC4 which showed 4- and 22-fold reductions, respectively, compared to IACS9571 and the corresponding parent BRD ligand **61**. BRD ligand **30**, which lacks a PEG chain, exhibited a substantially higher $IC_{50}$ than BRD ligands **61**–**63**, suggesting the possibility of PEG linker interactions.

a

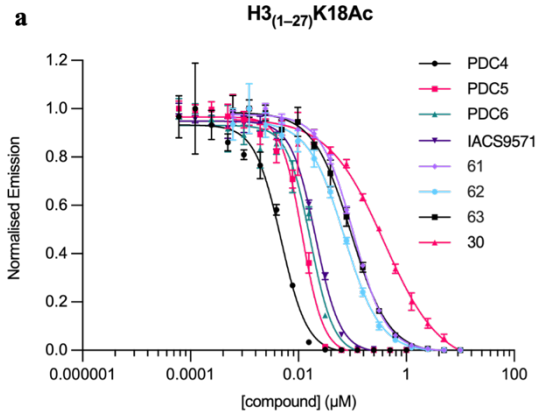



b

| Compound | AlphaScreen® $IC_{50}$ (μM) |
|---|---|
| **PDC4 (58)** | 0.0047 ± 0.0005 |
| **PDC5 (59)** | 0.0146 ± 0.0003 |
| **PDC6 (60)** | 0.0215 ± 0.0008 |
| **67** | 3.41 ± 0.29 |
| **$H3_{(1–27)}K9Me_3$ (+)** | 5.03 ± 0.40 |
| **$H3_{(1–15)}K4Me_3$ (−)** | no binding |

**Figure 4.2:** a) AlphaScreen® competition assay dose-response curves for inhibition of the interaction between biotinylated $H3_{(1–27)}K9Me_3$ and $His_6$-TRIM24 (PHD-BRD). Error bars indicate the standard deviation of triplicate data. Unnormalised dose-response curves can be found in Appendix 4, Figure A4.5; b) Table of $IC_{50}$ values from the AlphaScreen® competition assay. The data are quoted as the mean of triplicate data ± the standard error of the mean. Hot colours indicate stronger binders and cold colours indicate weaker binders [no binding = $IC_{50}$ > 400 μM; (+) = positive control; (−) = negative control].

a




b

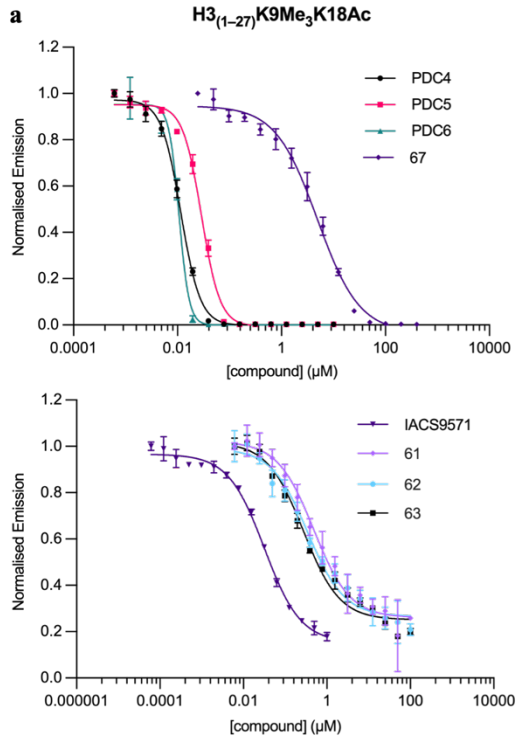
| Compound | AlphaScreen® $IC_{50}$ (μM) |
|---|---|
| **PDC4 (58)** | 0.0047 ± 0.0003 |
| **PDC5 (59)** | 0.0118 ± 0.0005 |
| **PDC6 (60)** | 0.0167 ± 0.0011 |
| **IACS9571** | 0.0190 ± 0.0017 |
| **30** | 0.388 ± 0.047 |
| **61** | 0.105 ± 0.005 |
| **62** | 0.0710 ± 0.0041 |
| **63** | 0.0942 ± 0.0046 |

c

| Compound | AlphaScreen® $IC_{50}$ (μM) |
|---|---|
| **$H3_{(1–27)}K18Ac$ (+)** | 0.266 ± 0.009 |
| **$H3_{(1–15)}K4Me_3$ (−)** | No binding |

**Figure 4.3:** a) AlphaScreen® competition assay dose-response curves for inhibition of the interaction between biotinylated $H3_{(1–27)}K18Ac$ and $His_6$-TRIM24 (PHD-BRD). Error bars indicate the standard deviation of triplicate data. Unnormalised dose-response curves can be found in Appendix 4, Figure A4.6; b, c) Tables of $IC_{50}$ values from the AlphaScreen® competition assay. The data are quoted as the mean of triplicate data ± the standard error of the mean. Hot colours indicate stronger binders and cold colours indicate weaker binders [no binding = $IC_{50}$ > 400 μM; (+) = positive control; (−) = negative control].

PDCs 4–6 were also shown to displace a dual modified $H3_{(1-27)}K9Me_3K18Ac$ peptide (Figure 4.4) with low nanomolar $IC_{50}$ values (10.7–28.8 nM). The data illustrated that the PDCs have higher affinity than the parent PHD (compound **67**) and BRD ligands (compounds **61**–**63**), which is in accordance with the previous assays (above). For example, PDC4 displayed 422- and 42-fold lower $IC_{50}$ values, compared to the corresponding parent PHD (compound **67**) and BRD ligands (compound **61**), respectively. It should be noted that the AlphaScreen® signal was observed to plateau significantly above baseline for IACS9571 and parent BRD ligands **61**–**63** (Figure 4.4a).




b

| Compound | AlphaScreen® $IC_{50}$ (μM) |
|---|---|
| **PDC4 (58)** | 0.0116 ± 0.0003 |
| **PDC5 (59)** | 0.0288 ± 0.0008 |
| **PDC6 (60)** | 0.0107 ± 0.0002 |
| **IACS9571** | 0.0318 ± 0.0020 |
| **61** | 0.484 ± 0.071 |
| **62** | 0.348 ± 0.049 |
| **63** | 0.274 ± 0.029 |
| **67** | 4.90 ± 0.35 |
| **$H3_{(1–27)}K9Me_3K18Ac$ (+)** | 0.242 ± 0.014 |
| **$H3_{(1–15)}K4Me_3$ (−)** | no binding |

**Figure 4.4:** a) AlphaScreen® competition assay dose-response curves for inhibition of the interaction between biotinylated $H3_{(1–27)}K9Me_3K18Ac$ and $His_6$-TRIM24 (PHD-BRD). Error bars indicate the standard deviation of triplicate data. Unnormalised dose-response curves can be found in Appendix 4, Figure A4.7; b) Table of $IC_{50}$ values from the AlphaScreen® competition assay. The data are quoted as the mean of triplicate data ± the standard error of the mean. Hot colours indicate stronger binders and cold colours indicate weaker binders [no binding = $IC_{50}$ > 400 μM; (+) = positive control; (−) = negative control].

The data from the AlphaScreen® assays (Figures 4.2, 4.3, and 4.4) show that the PDCs can inhibit *both* the TRIM24 PHD and BRD more effectively than the parent PHD and BRD ligands. The data also suggest that the PDCs have higher affinity for TRIM24 than IACS9571. The dual-modified $H3_{(1-27)}K9Me_3K18Ac$ peptide has been reported as one of the highest affinity histone-mimicking peptides for the TRIM24 PHD and BRD.[1] It was suggested that this could be due to avidity, arising from binding to the PHD and BRD on the same protein (*in cis* binding mode).[1] Therefore, only compounds that simultaneously inhibit the PHD and BRD, could fully displace biotinylated $H3_{(1-27)}K9Me_3K18Ac$. Whilst IACS9571 and BRD ligands **61**–**63** can block binding of K18Ac to the BRD, the biotinylated peptide could still bind to the PHD, albeit with lower affinity, through the $H3_{(1-9)}K9Me_3$ motif. However, the PDCs bind to both domains and so could completely inhibit binding of $H3_{(1–27)}K9Me_3K18Ac$. This could explain why the PDCs reduce the AlphaScreen® signal down to the baseline, whereas a residual signal remains for IACS9571 and parent BRD ligands **61**–**63** (Figure 4.4a). Intriguingly, parent PHD ligand **67** also brings the signal down to baseline, suggesting that it can weakly bind to the BRD.

### 4.2.2 Determination of Binding Affinities and Kinetics

Having identified that the PDCs can displace peptides mimicking the endogenous histone binding partner of the TRIM24 PHD and BRD, it was next decided that the binding affinities and kinetics would be investigated using surface plasmon resonance (SPR). Firstly, it was necessary to identify a method for capturing $His_6$-TRIM24 (PHD-BRD) onto the surface of the SPR sensor chip. Two reversible capture methods were evaluated, including immobilising an anti-His tag antibody or Ni-NTA chelate onto the surface of the chip. Reversible techniques have the advantage that the chip can be stripped of protein and reused. However, despite an initial increase in protein levels after injection, the level rapidly reduced over time as buffer was passed over surface, resulting in insufficient protein

capture (Figure 4.5). An irreversible covalent cross-linking approach was then investigated, which involved reaction of protein surface lysines with carboxylates on the sensor chip surface after activation with an EDC/NHS coupling mixture. This method captured sufficient protein on the chip surface (Figures 4.5 and 4.6) and so was used as the capture method for all subsequent experiments.

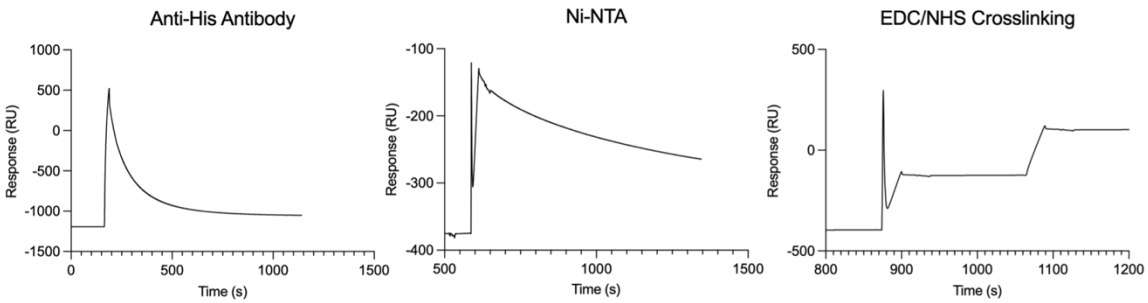



**Figure 4.5:** SPR Sensograms showing the capture of $His_6$-TRIM24 (PHD-BRD) onto the surface of a CM5 sensor chip using three different capture techniques: Anti-His antibody, Ni-NTA chelate, and EDC/NHS crosslinking.

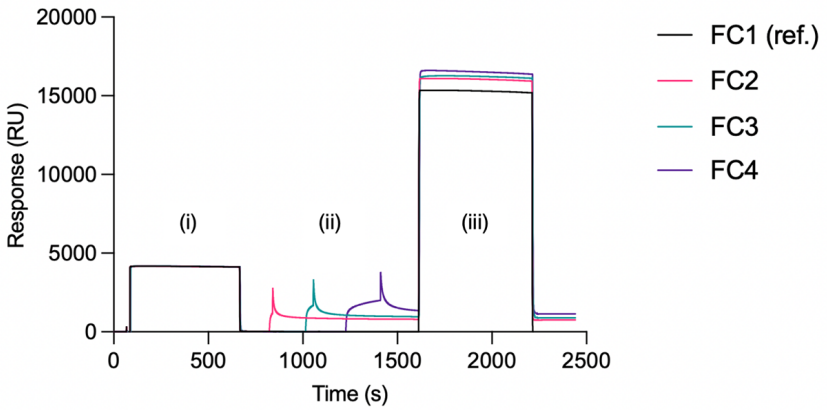



**Figure 4.6:** SPR Sensograms for the immobilisation of $His_6$-TRIM24 onto the surface of a sensor chip: (i) EDC/NHS activation, (ii) protein immobilisation, (iii) ethanolamine capping (FC = flow cell).

There was concern that the covalent capture method could block binding sites in the TRIM24 PHD and BRD due to inhomogeneity in the orientation of protein molecules captured on the surface. The binding affinity of a positive control peptide, $H3_{(1\text{-}27)}K9Me_3K18Ac$, was assessed to validate the assay. Intriguingly, a negative signal

was produced when using a CM5 sensor chip (Figure 4.7a) for the assay. A significant increase in signal was observed for the reference flow cell (Figure 4.7c) despite the fact that no protein was immobilised on this surface, which suggested non-specific binding of $H3_{(1-27)}K9Me_3K18Ac$. It was hypothesised that carboxymethyl dextran, a polar polysaccharide on the surface of the CM5 sensor chip, could be interacting with the peptide to produce a non-specific binding signal. Therefore, a C1 sensor chip was used instead as it does not contain dextran but still has carboxylates on the surface to capture the protein. Pleasingly, the C1 sensor chip demonstrated minimal non-specific binding (Figure 4.7c) and produced the expected positive response in the SPR sensogram (Figure 4.7b).

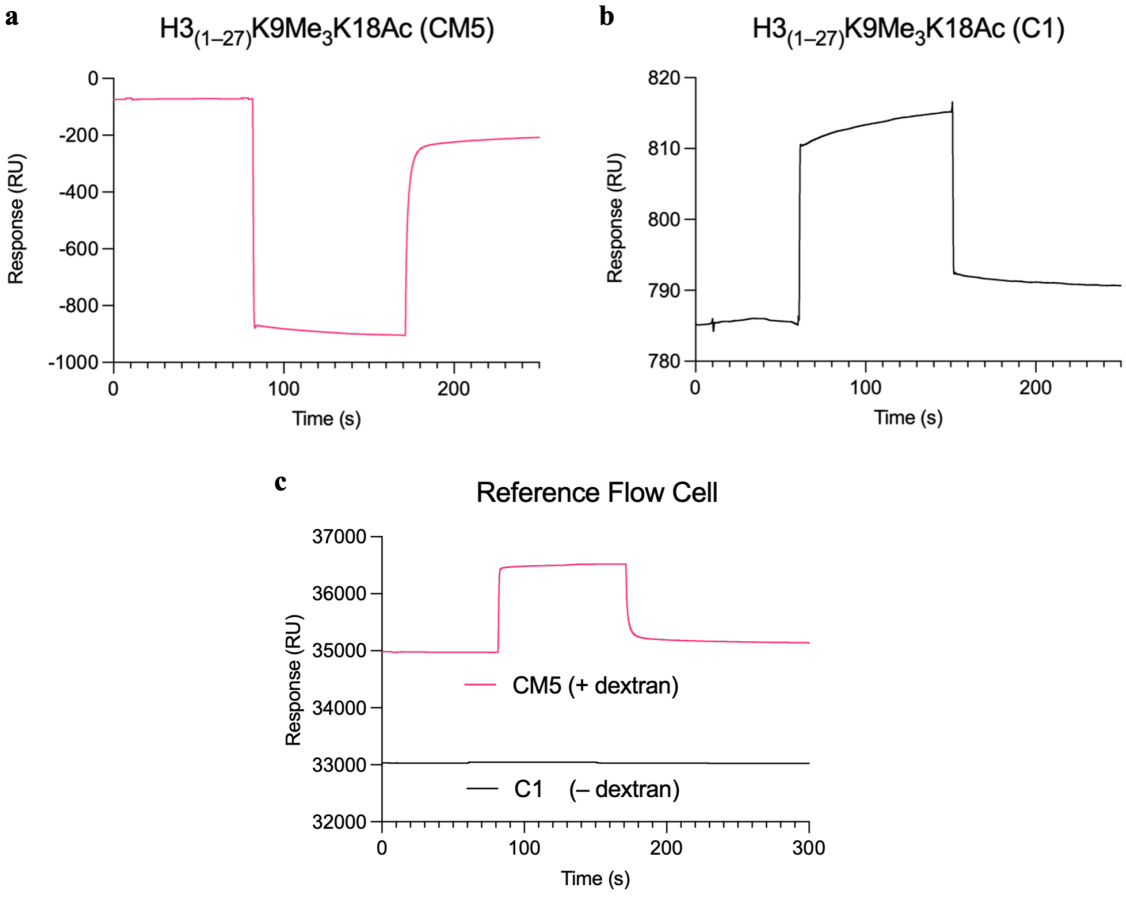



**Figure 4.7:** a,b) Reference-subtracted SPR sensograms for $H3_{(1–27)}K9Me_3K18Ac$ (20 µM) binding to $His_6$-TRIM24 (PHD-BRD) covalently immobilised on a a) CM5 or b) C1 sensor chip; c) Reference flow cell SPR sensograms [no $His_6$-TRIM24 (PHD-BRD)] showing non-specific binding of $H3_{(1–27)}K9Me_3K18Ac$ (20 µM) to the surface of a C1 (no dextran) or CM5 sensor chip (contains dextran).

After optimising the assay, the binding affinity of the positive control $H3_{(1-27)}K9Me_3K18Ac$ peptide could now be evaluated. The SPR sensogram for $H3_{(1-27)}K9Me_3K18Ac$ demonstrated fast $k_{on}$ and $k_{off}$ rates which reached the limit of detection (Figure 4.8a), as expected for a weak-binding (μM) peptide. To determine a value for the $K_d$, the mean equilibrium response (sensogram plateau) was plotted as a function of the peptide concentration and the data were fitted to a binding saturation curve (Figure 4.8b) using a one-site binding model. In reality, a two-site binding model would be more appropriate as the peptide binds to both the PHD and BRD, however, this would overcomplicate the data analysis. The $K_d$ value for the positive control was determined to be 1.14 ± 0.20 μM, which is in reasonable agreement with the previously reported ITC $K_d$ value (1.94 ± 0.11 μM).[1] This suggests that the covalent protein capture method does not significantly affect binding of compounds to the TRIM24 PHD and BRD.

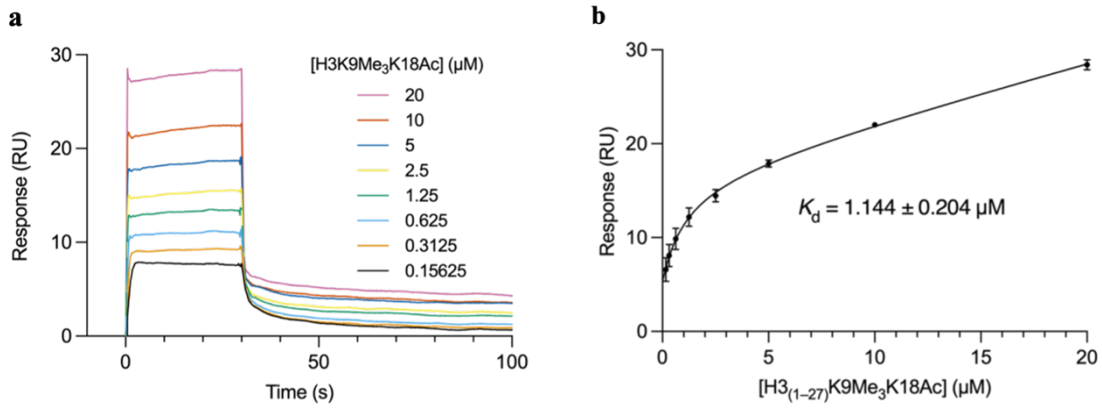
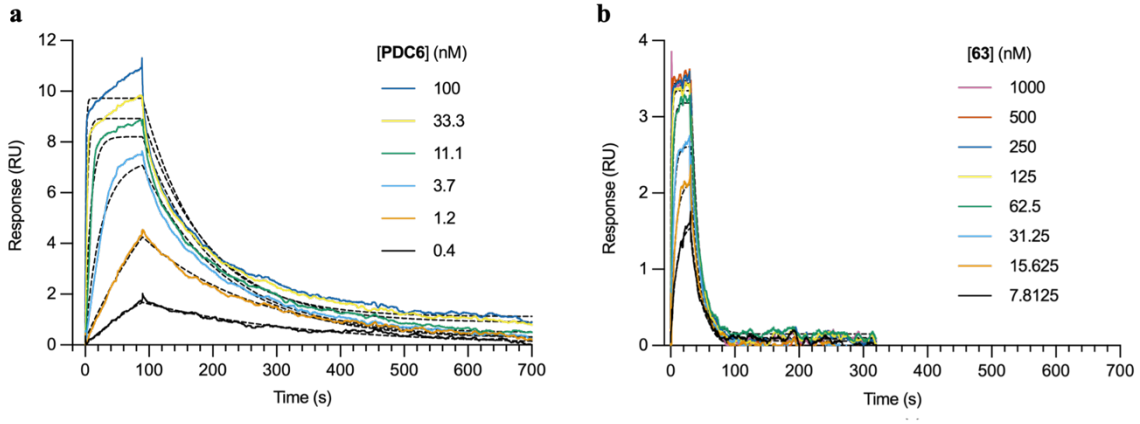



**Figure 4.8:** a) Representative reference-subtracted SPR sensograms for the binding of $H3_{(1-27)}K9Me_3K18Ac$ to $His_6$-TRIM24 (PHD-BRD); b) Binding saturation curve for the binding of $H3_{(1-27)}K9Me_3K18Ac$ to $His_6$-TRIM24 (PHD-BRD). $K_d$ values (equilibrium analysis) are quoted as the mean of triplicate data ± standard error of the mean. Error bars indicate the standard deviation.

Having validated the SPR assay using the positive control peptide, the binding affinities and kinetics of the PDCs and parent BRD ligands could now be compared. Due to time constraints on the project only PDC6 (compound **60**) and its corresponding parent BRD ligand (compound **63**) were investigated, although based on the results of the AlphaScreen®

assay, the other PDCs would likely produce similar results. Again, for these assays a one-site binding model was used as an approximation to simplify data analysis. The sensogram data were fitted to association and dissociation curves (Figures 4.9a and b) to determine values for $k_{on}$ and $k_{off}$, which were subsequently used to calculate a value for the $K_d$ (Figure 4.9c).




c

| Compound | $k_{on}$ ($M^{-1}$ $s^{-1}$) | $k_{off}$ ($s^{-1}$) | $K_d$ (nM) |
|---|---|---|---|
| **PDC6 (60)** | $1.3\times10^{7} \pm 1.4\times10^{6}$ | $8.2\times10^{-3} \pm 3.1\times10^{-4}$ | $0.6 \pm 0.2$ |
| **63** | $6.1\times10^{6} \pm 4.6\times10^{5}$ | $6.1\times10^{-2} \pm 3.2\times10^{-3}$ | $10.0 \pm 7.0$ |

**Figure 4.9:** a, b) Representative reference-subtracted SPR sensograms for the binding of PDC6 (compound **60**) and BRD ligand **63** to $His_6$-TRIM24 (PHD-BRD). Dashed lines represent fitted curves which were used to calculated $K_d$ values and rate constants. Individual reference-subtracted SPR sensograms for each replicate can be found in Appendix 7, Figures A7.1 and A7.2; c) SPR $K_d$ values (kinetic analysis) and rate constants for the binding of PDC6 and BRD ligand **63** to $His_6$-TRIM24 (PHD-BRD). $K_d$ values and rate constants were obtained from fitted curves for each concentration. The data were averaged across the concentrations before averaging across the replicates ($n$ = 3). Data are quoted as the mean of means for triplicate data ± standard error of the mean.

PDC6 demonstrated picomolar affinity (0.6 ± 0.2 μM) for the TRIM24 PHD and BRD; and a 17-fold increase in affinity compared to its parent BRD ligand (compound **63**). The affinity of compound **63** was also determined using the equilibrium analysis (Appendix 7, Figure A7.5) method used for the positive control, and was found to be in excellent

agreement with the $K_d$ value from kinetic analysis [$K_{d\ (kinetic)}$ = 10.0 ± 7.0 nM; $K_{d\ (equilibrium)}$ = 11.0 ± 2.3 nM]. It was evident from the sensograms (Figures 4.9a and b) that the high affinity of the PDCs was driven by a slow $k_{off}$. This was supported by the calucluated values from the fitted curves which revealed a 7-fold slower $k_{off}$ for PDC6 compared to parent BRD ligand **63** (Figure 4.9c). These data, combined with the AlphaScreen® data from the previous section (4.2.1), are consistent with simultaneous PHD and BRD binding *via* a bivalent binding mode.

PDC engagement with W828 in the PHD was also explored using W828A mutant TRIM24 (provided by Dr Angelina Sekirnik). Initially, protein activity was verified using the positive control, which still bound to W828A mutant TRIM24, albeit with 4-fold lower affinity than the wild type (Table 4.1), which is in accordance with previous work.[1] PDC6 demonstrated 2-fold lower affinity for the mutant than the wild-type (WT), indicating that it does engage with W828 in the WT protein (Table 4.1). These data also suggest that the cation-π interaction with the isoindoline $KMe_3$ mimic is important for the picomolar affinity of the PDCs.

**Table 4.1:** SPR $K_d$ values and rate constants for binding of $H3_{(1–27)}K9Me_3K18Ac$ and PDC6 (compound **60**) to WT/W828A mutant $His_6$-TRIM24 (PHD-BRD). PDC6 $K_d$ values and rate constants were obtained from fitted curves for each concentration (kinetic analysis); and the data were averaged across the concentrations before averaging across the replicates ($n$ = 3). PDC6 data are quoted as the mean of means for triplicate data ± standard error of the mean. $H3_{(1-27)}K9Me_3K18Ac$ $K_d$ values were obtained from binding saturation curves (equilibrium analysis) and are quoted as the mean of triplicate data ± standard error of the mean. W828 mutant sensograms and binding saturation curves can be found in Appendix 7; Figures A7.3, A7.4, A7.6.

| TRIM24 | Compound | $k_{on}$ ($M^{-1}$ $s^{-1}$) | $k_{off}$ ($s^{-1}$) | $K_d$ |
|---|---|---|---|---|
| **WT** | **$H3_{(1–27)}K9Me_3K18Ac$** | – | – | 1.1 ± 0.2 μM |
| | **PDC6 (60)** | $1.3\times10^{7} \pm 1.4\times10^{6}$ | $8.2\times10^{-3} \pm 3.1\times10^{-4}$ | 0.6 ± 0.2 nM |
| **W828A** | **$H3_{(1–27)}K9Me_3K18Ac$** | – | – | 4.4 ± 0.5 μM |
| | **PDC6 (60)** | $9.9\times10^{6} \pm 6.4\times10^{5}$ | $1.2\times10^{-2} \pm 2.6\times10^{-4}$ | 1.2 ± 0.4 nM |

### 4.2.3 Elucidation of Binding Modes

In principle, due to their bifunctional nature, the PDCs can bind to TRIM24 in a number of different ways (Figure 4.10). Therefore, to probe the binding mode further, co-crystallisation of the PDCs with TRIM24 (PHD-BRD) was investigated.

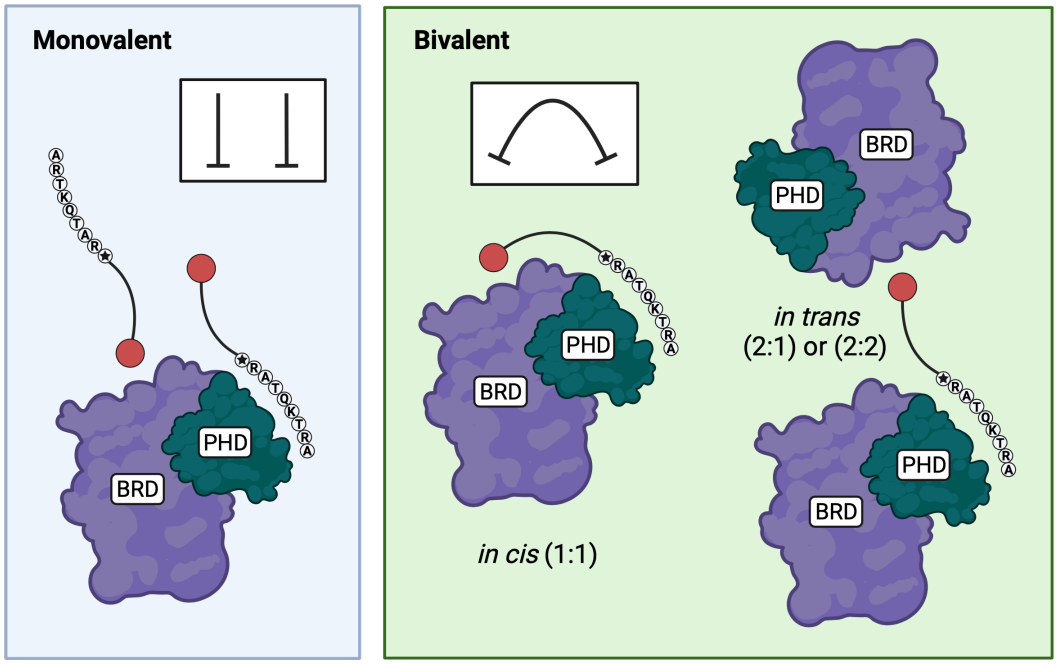



**Figure 4.10:** Schematic representation of the possible monovalent and bivalent binding modes of the PDCs to the TRIM24 PHD and BRD. The red circle represents the TRIM24 BRD ligand and white circles with single letter amino acid codes together represent the PHD-binding peptide. A 'star' represents the unnatural amino acid. Figure created with BioRender.com.

Firstly, an attempt was made to crystalise $His_6$-TRIM24 (PHD-BRD) in the *apo* state, however, this failed to yield any crystals in a variety of in-house crystallisation screens. It was hypothesised that the His-tag could be hindering crystalisation, therefore it was decided that an untagged construct would be used instead. Previous work identified that the PHD domain was unstable to His-tag cleavage with TEV protease.[162] Therefore, instead a $His_6$-SUMO-TRIM24 construct was expressed, which was previously shown to be stable to SUMO tag cleavage.[162] The $His_6$-SUMO tag enabled purification using immobilised metal affinity chromatography (IMAC) (Figure 4.11). SUMO tag cleavage was then achieved using $His_6$-SENP1 – a SUMO protease enzyme (Figure 4.11). A second IMAC

purification was then carried out to separate untagged TRIM24 from the SUMO protease enzyme and cleaved $His_6$-SUMO tag (Figure 4.11). Size exclusion chromatography (SEC) was then used as a polishing step to remove any aggregated protein and obtain the high purity untagged TRIM24 required for crystallisation (Figure 4.12; Appendix 1, Figures A1.10 and A1.11).

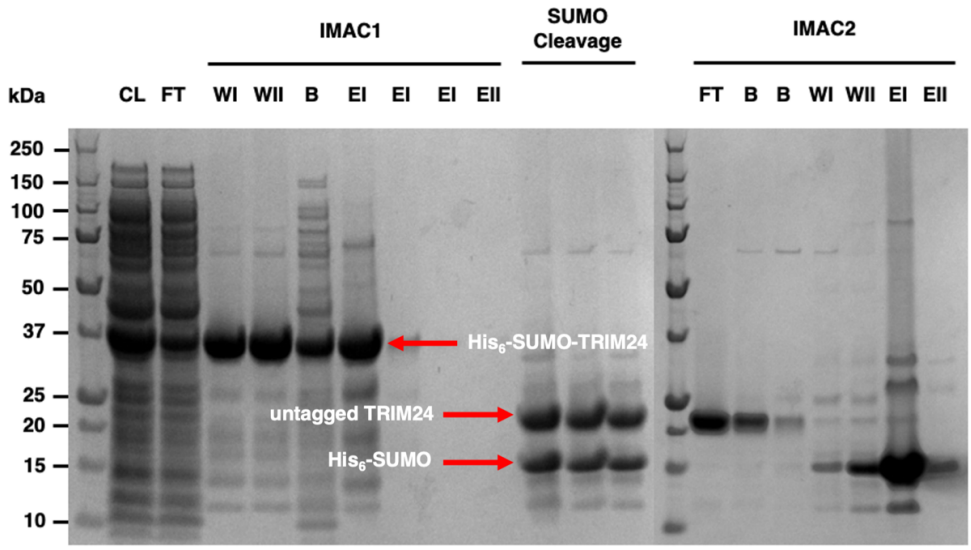



**Figure 4.11:** Images of SDS-PAGE gels for the purification of $His_6$-SUMO-TRIM24 and untagged TRIM24 [$His_6$-SUMO-TRIM24 (33.8 kDa); $His_6$-SUMO (12.3 kDa); untagged TRIM24 (21.5 kDa)] [CL, cell lysate; FT, flowthrough; W, wash buffer; B, binding buffer; E, elution buffer].

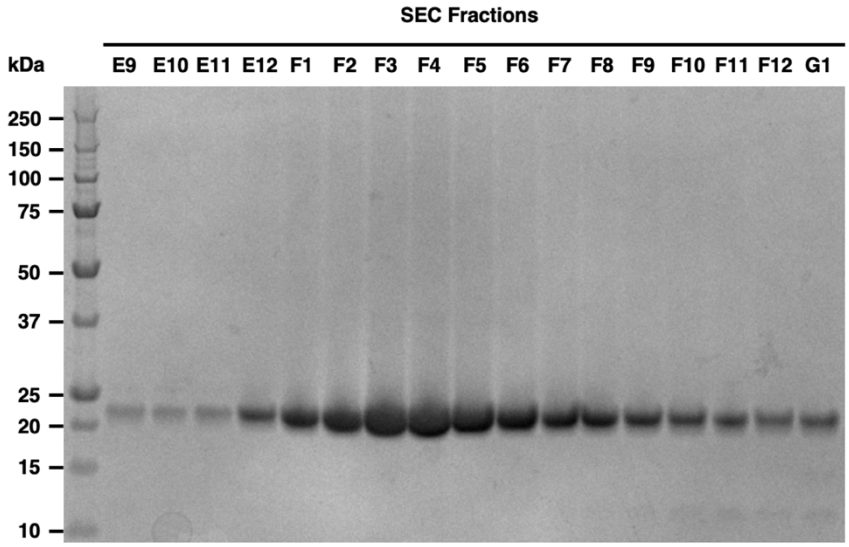



**Figure 4.12:** Image of the SDS-PAGE gel for the purification of untagged TRIM24 by size-exclusion chromatography (SEC).

Crystallisation of *apo* TRIM24 was then carried out at 20 °C using a fine screen based on previous crystallisation conditions.[2,4] Crystals appeared after two days and reached full size within a week (Figure 4.13). Single crystals were then mounted and sent to Diamond Light Source for X-ray diffraction data collection and processing.

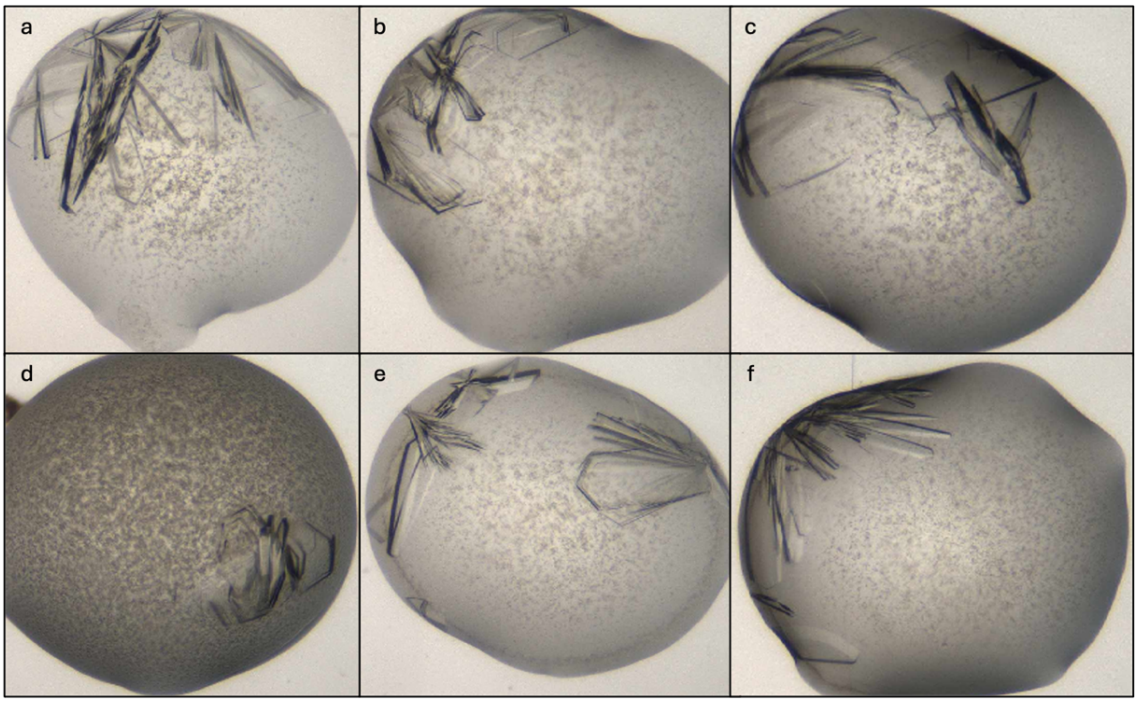



**Figure 4.13:** A selection of *apo* untagged TRIM24 (PHD-BRD) crystals obtained in the fine screen. *Conditions:* a) 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.4), PEG-400 5 % *v*/*v*, 2:1; b) 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.4), PEG-400 5 % *v*/*v*, 1:1; c) 2.0 M ammonium sulfate, 0.1 M HEPES (pH 7.7), PEG-400 2% *v*/*v*, glycerol 10% *v*/*v*, 1:1; d) 2.0 M ammonium sulfate, 0.1 M HEPES (pH 7.9), PEG-400 2% *v*/*v*, glycerol 8% *v*/*v*, 1:2; e) 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.7), PEG-400 2% *v*/*v*, 1:1; f) 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.1), PEG-400 5% *v*/*v*, 1:2. Ratios are given as protein : reservoir. Reservoir solutions are aqueous.

Pleasingly, the diffraction data revealed a new crystal form for *apo* TRIM24 in the C121 space group with each unit cell containing a dimer (Figure 4.14; Appendix 2, Table A2.1). Previously, *apo* TRIM24 had only been crystallised in the P1 space group as a monomer (PDB ID: 3O33).[2] The data confirmed that the protein purity was sufficient to obtain high resolution crystal structures of TRIM24 and so could be used for co-crystallisation of the PDCs.

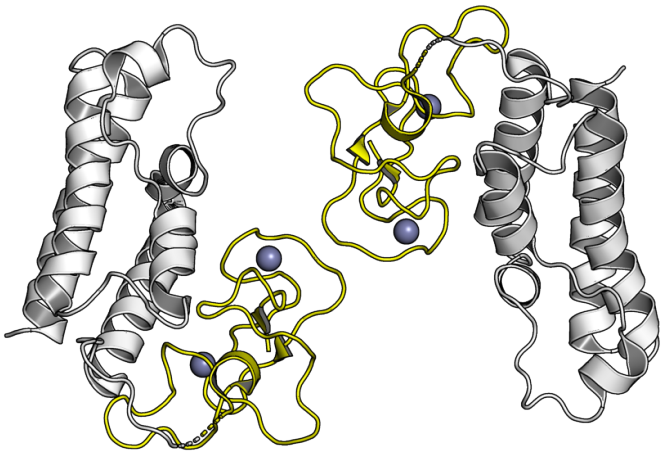


**Figure 4.14:** Unit cell for the *apo* crystal structure of TRIM24 in the C121 space group (1.7 Å, PDB ID: 9GD5). The PHD and BRD are coloured in yellow and white, respectively. Grey spheres represent $Zn^{2+}$ ions. Dashed lines represent unresolved residues in disordered regions.

Co-crystallisation of PDCs 4–6 (compounds **58**–**60**) with TRIM24 (PHD-BRD) was then investigated, however, no crystals were observed after 2 months under a variety of conditions using either the *apo* TRIM24 fine screen or in-house coarse screens. The *apo* TRIM24 crystal structure revealed that the PHD and BRD binding sites were solvent-exposed (Figure 4.14) and so it was reasoned that a crystal-soaking strategy could also be possible. Therefore, TRIM24 crystals, obtained from the fine screen, were soaked in reservoir solutions containing PDCs 4–6 for 48 h. Unfortunately, no TRIM24 crystals remained after 48 h, which suggested that the PDCs were hindering crystalisation.

Intriguingly whilst co-crystallisation of parent PHD peptide **67** proved unsuccessful, parent BRD ligand **61** was found to co-crystallise with TRIM24 to produce high quality crystals (Figure 4.15). The refined crystal structure (PDB ID: 9GDG, 1.5 Å) revealed the expected hydrogen bonding interaction between N980 in the BRD and the KAc mimic in compound **61** (Figure 4.16). The structure also confirmed the solvent-exposed linker attachment point on the benzenesulfonamide ring (Figure 4.16). It was therefore hypothesised that the PHD-binding peptide in the PDCs might be solubilising the protein and/or obstructing the specific protein-protein interactions, necessary for crystallisation to occur.

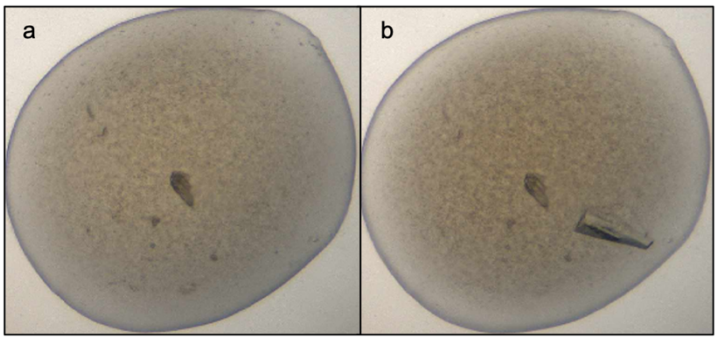



**Figure 4.15:** Images of crystalisation wells for the co-crystallisation of BRD ligand **61** with TRIM24 (PHD-BRD) after a) 28 days and b) 56 days. *Conditions*: 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.3), PEG-400 5 % *v/v*, 2:1 (protein : reservoir). Reservoir solutions are aqueous.

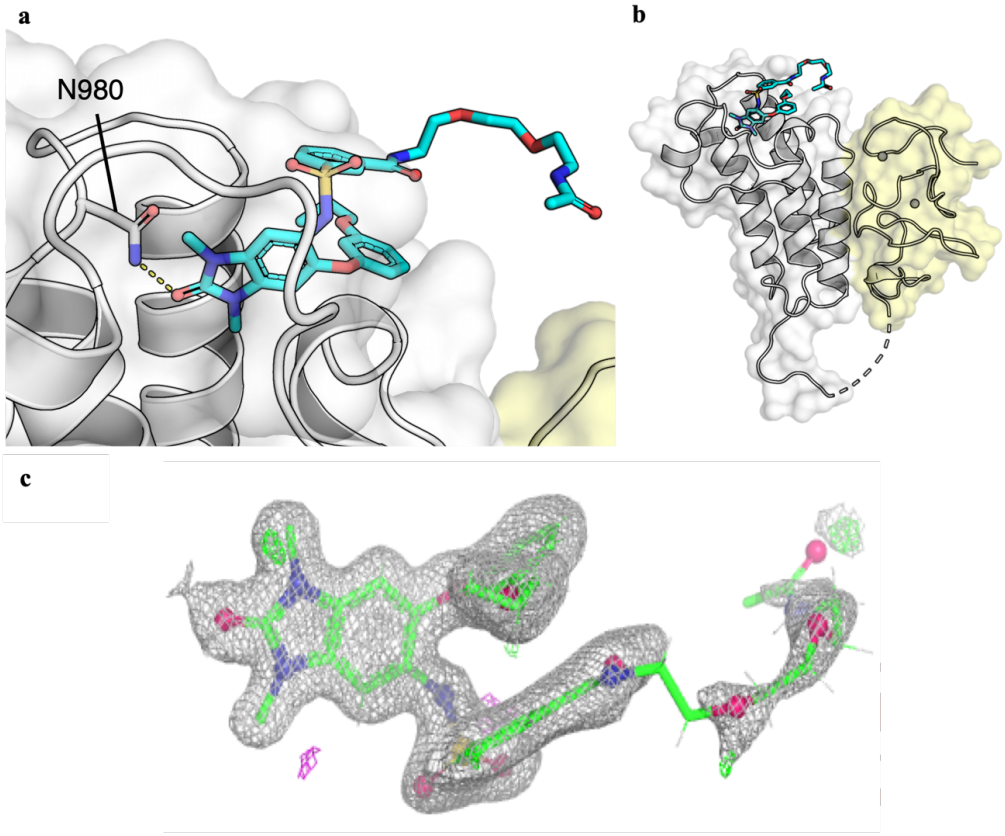
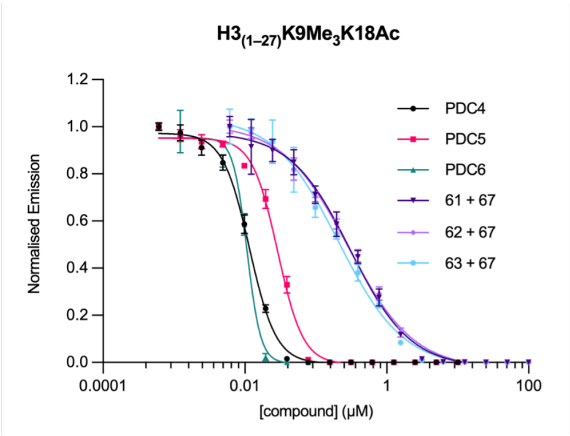



**Figure 4.16:** a) Co-crystal structure of TRIM24 in complex with BRD ligand **61** (1.5 Å, PDB ID: 9GDG). The structure shows a hydrogen bond between N980 and a carbonyl in the benzimidazolone core of BRD ligand **61**; b) Unit cell for the crystal structure of TRIM24 in complex with BRD ligand **61**. The PHD and BRD are coloured in yellow and white, respectively. Grey spheres represent $Zn^{2+}$ ions. Dashed lines represent unresolved residues in disordered regions; c) Experimental electron density around BRD ligand **61** from the PDB validation report. Limited electron density was observed for the PEG chain and so it was placed in an arbitrary position in the crystal structure.

Since crystallography failed to reveal the binding mode of the PDCs, AlphaScreen® was pursued as an alternative approach. In this assay the PDCs were compared to mixtures of the unlinked parent PHD and BRD ligands (e.g. PDC4 *vs* **61** + **67**; PDC5 *vs* **62** + **67** etc.) for their ability to displace biotinylated $H3_{(1–27)}K9Me_3K18Ac$. If the PDCs were binding bivalently to the PHD and BRD (*in cis* or *in trans*), a lower $IC_{50}$ would be expected compared to mixtures of the unlinked parent ligands due to avidity.

In the AlphaScreen® assay, PDCs 4–6 each demonstrated a significantly lower $IC_{50}$ value than their corresponding parent ligand mixtures (Figure 4.17). This was exemplified by PDC4, which displayed a 26-fold lower $IC_{50}$ than the mixture containing compounds **61** and **67**. These data provided further evidence that the PDCs were binding simultaneously to the PHD and BRD *via* a bivalent binding mode. The data also suggested that shorter PEG linkers enhance bivalent binding, as PDC4, which features the shortest PEG chain, showed the greatest reduction in $IC_{50}$ compared to its unlinked parent ligands.




| Compound | AlphaScreen® $IC_{50}$ (μM) |
|---|---|
| **PDC4 (58)** | 0.0116 ± 0.0003 |
| **PDC5 (59)** | 0.0288 ± 0.0008 |
| **PDC6 (60)** | 0.0107 ± 0.0002 |
| **61 + 67** | 0.305 ± 0.025 |
| **62 + 67** | 0.284 ± 0.018 |
| **63 + 67** | 0.210 ± 0.019 |

**Figure 4.17:** a) AlphaScreen® competition assay dose-response curves for inhibition of the interaction between biotinylated $H3_{(1–27)}K9Me_3K18Ac$ and $His_6$-TRIM24 (PHD-BRD). Error bars indicate the standard deviation of triplicate data. For compound mixtures (e.g. **61** + **67**), each data point corresponds to a mixture of both compounds at the specified concentration. Unnormalised dose-response curves can be found in Appendix 4, Figure A4.7; b) Table of $IC_{50}$ values from the AlphaScreen® competition assay. The data are quoted as the mean of triplicate data ± the standard error of the mean. Hot colours indicate stronger binders and cold colours indicate weaker binders.

Given the mounting evidence for bivalent PDC binding, we next sought to determine whether this was occurring through an *in cis* or *in trans* binding mode. To investigate the binding mode further an AlphaScreen® assay developed in the Conway group, was used to assess the ability of the PDCs to form a ternary complex between GST-tagged and His-tagged TRIM24 (Figure 4.18). If the PDCs were binding *in trans* to the PHD of one protein and the BRD of another, then a ternary complex would be formed, resulting in the production of an emission signal (Figure 4.18a). On the other hand, if the PDCs were binding *in cis* to the same protein then no ternary complex would form and no signal would be produced (Figure 4.18b).

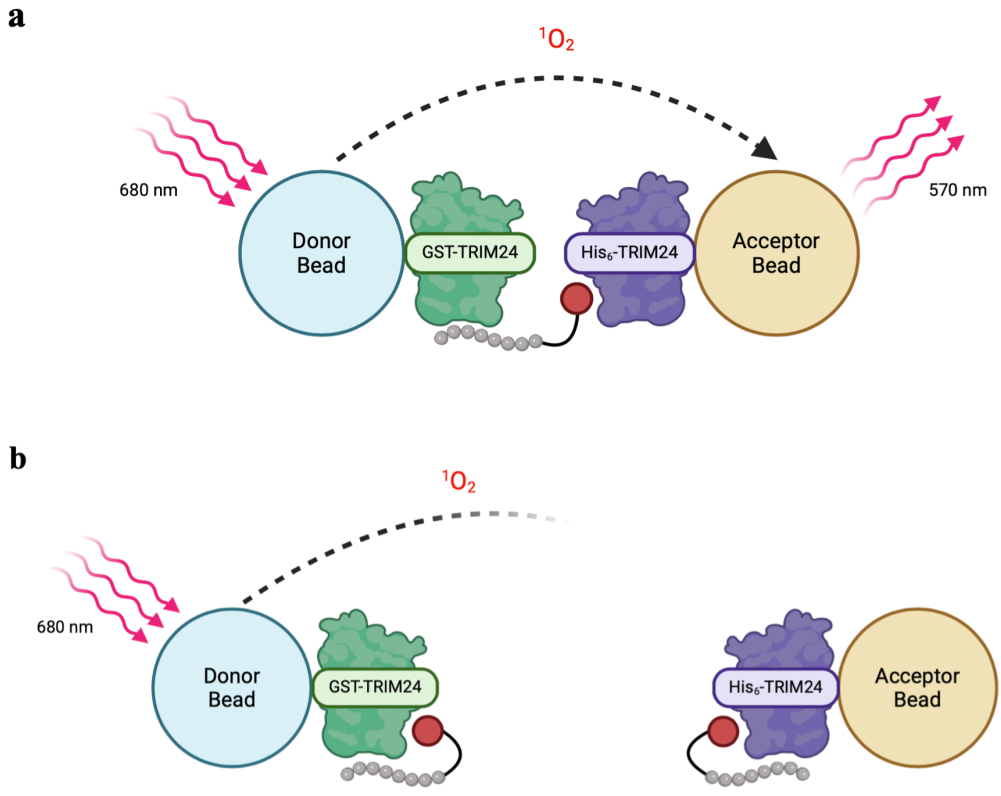



**Figure 4.18:** Schematic representations of the AlphaScreen® ternary complex detection assay for two binding scenarios - a) *in trans* and b) *in cis* PDC binding. In the absence of a ternary complex no emission signal is produced due to singlet oxygen quenching. GST-TRIM24 (provided by Dr Angelina Sekirnik) and His$_6$-TRIM24 were attached to glutathione donor and $Ni^{2+}$ chelate acceptor beads, respectively. The BRD ligand of the PDCs is represented by red spheres and grey spheres represent the amino acids in the PHD-binding peptide. Figure created with BioRender.com.

Initially, the functional activities of the PHD and BRD in GST-TRIM24 (provided by Dr Angelina Sekirnik) were assessed using signal-response assays with biotinylated $H3_{(1\text{-}27)}K9Me_3$ and $H3_{(1\text{-}27)}K18Ac$. This was carried out to verify that a lack of emission signal was not due to occlusion of the binding sites or degradation of the protein which would prevent binding of the PDCs. Dose-dependent signals were produced for each peptide (Appendix 4, Figure A4.2) indicating that both the PHD and BRD were active. In the absence of a *bona fide* positive control, $His_6$-GST was used as a technical positive control without any protein added to illustrate the expected signal produced when a ternary complex is formed (Figure 4.19). As the concentration of $His_6$-GST increases ternary complexes begin to form between the donor and acceptor beads, which causes an increase in the emission signal. Above a critical concentration the donor and acceptor beads become saturated, which hinders formation of ternary complexes and favours binary complex formation. This leads to a reduction in the emission signal known as the 'Hook effect'.[163] Contrastingly, PDCs 4–6 (compounds **58**–**60**) failed to produce an emission signal indicating that no ternary complex was formed. These data suggested that the PDCs were binding to the TRIM24 PHD and BRD *via* an *in cis* binding mode.

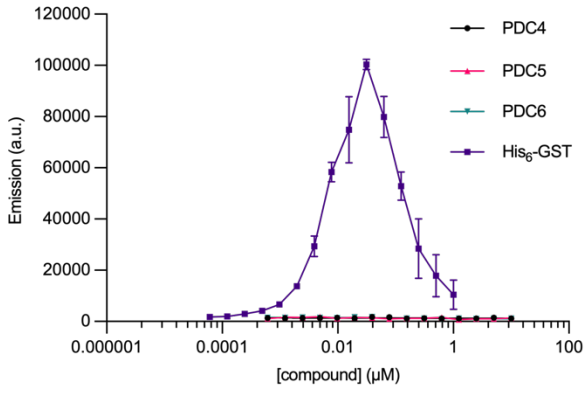



**Figure 4.19:** AlphaScreen® ternary complex detection assay dose-response curves for PDCs 4–6 (compounds **58**–**60**). $His_6$-GST has been included as a technical positive control to illustrate the 'Hook effect' that would be expected if the PDCs are able to bring the assay beads in close proximity through formation of a ternary complex. Error bars indicate the standard deviation of triplicate data. Individual dose-response curves can be found in Appendix 4, Figure A4.8.

### 4.2.4 Evaluation of Selectivity

Selectivity still remains a major challenge for TRIM24 BRD inhibition. For instance IACS9571, the highest affinity reported TRIM24 BRD ligand, in fact demonstrates 2-fold higher affinity for the BRD of BRPF1 than that of TRIM24.[4] Interestingly, the PHDs in the PZP domain of BRPF1 are further away from the BRD than in TRIM24 (Figure 4.20). This raises the possibility of gaining selectivity through avidity using the PDCs. BRPF1 still contains reader domains that recognise H3, however these are not involved in the recognition of K9Me$_3$. Therefore, although the BRD ligand alone is non-selective over BRPF1, the PDCs could have some selectivity for TRIM24 through simultaneous binding to the PHD and BRD, which is not possible in BRPF1.

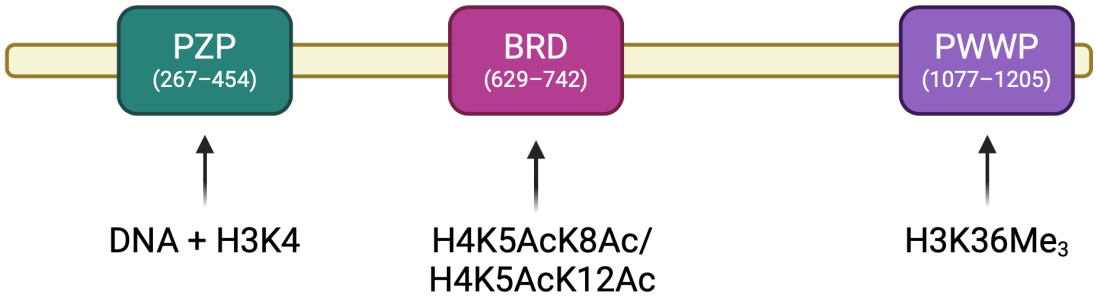



**Figure 4.20:** Domain composition of BRPF1. The PZP domain (residues 267–454) binds bivalently to DNA and unmodified H3K4, while the BRD (residues 629–742) and PWWP (residues 1077–1205) domains bind to diacetylated H4 and H3K36Me$_3$, respectively.[164–166] Figure created with BioRender.com.

BROMO*scan* (performed by Eurofins DiscoverX corporation) was used to identify any selectivity that could be gained over BRPF1, through bivalent binding to the TRIM24 PHD and BRD.[167] The PDCs again demonstrated picomolar $K_d$ values for TRIM24, and substantial increases in affinity compared to the parent PHD and BRD ligands (Table 4.2), which is in agreement with the SPR data (section 4.2.2). Crucially, whereas BRD ligand **63** showed no selectivity over BRPF1, PDCs 4–6 (compounds **58**–**60**) demonstrated 8.9–12.8-fold selectivity for TRIM24. These data suggest that it might be possible to gain

selectivity for TRIM24 through bivalent binding to the PHD and BRD. It is worth highlighting that although expression of full-length BRPF1 was beyond the scope of this work, in future it would be interesting to investigate binding of the PDCs to a full-length construct to see if selectivity is maintained.

**Table 4.2:** BROMO*scan* $K_d$ values (Eurofins DiscoverX corporation) for binding to TRIM24 (PHD-BRD) and BRPF1 (BRD). Fold selectivities were calculated by dividing $K_d$ (BRPF1) by $K_d$ (TRIM24). $K_d$ values are quoted as the mean of duplicate data ± 95% confidence interval. Hot colours indicate stronger binders and cold colours indicate weaker binders. Dose-response curves can be found in Appendix 6, Figures A6.1 and A6.2.

| Compound | TRIM24 (PHD-BRD) $K_d$ (nM) | BRPF1 (BRD) $K_d$ (nM) | Fold Selectivity for TRIM24 |
|---|---|---|---|
| **PDC4 (58)** | 0.27 ± 0.19 | 2.4 ± 0.9 | 8.9 |
| **PDC5 (59)** | 0.22 ± 0.05 | 2.2 ± 0.5 | 10.0 |
| **PDC6 (60)** | 0.18 ± 0.04 | 2.3 ± 0.4 | 12.8 |
| **63** | 2.0 ± 0.1 | 2.2 ± 0.5 | 1.1 |
| **67** | 3350 ± 686 | >10000 | > 3.0 |

## 4.3 Evaluation of Phenotypic Effects

Having confirmed that the PDCs simultaneously bind to the TRIM24 PHD and BRD with high affinity, they were next evaluated using *in vitro* cellular assays to identify any cytotoxic or antiproliferative effects on breast cancer cell lines. Given that TRIM24 is involved in switching on oestrogen-dependent genes linked to proliferation and breast cancer development, it was decided that the PDCs would be assessed for antiproliferative effects in an oestrogen receptor positive ($ER^+$) cell line.[2] The MCF-7 cell line was chosen for this purpose as it expresses the oestrogen receptor and has previously proven sensitive to TRIM24 knockdown.[2] Unless otherwise stated, cell culture, MTT, proliferation, and clonogenic assays were carried out by Louise Martin (University of Oxford).

Due to their peptidomimetic structure, it was necessary to evaluate the cell permeability of the PDCs, prior to carrying out the phenotypic assays. A previously developed cell permeability assay was used in which MCF-7 breast cancer cells (cell culture carried out Dr Antoine Wallabrègue) were incubated with the compound of interest (20 μM) for 6 h, before being washed with HBS to remove any extracellular compound.[168,169] Cells were then lysed by sonication and the lysates were analysed by LC-MS. All UV chromatograms were blank-subtracted using a DMSO control sample. A qualitative binary assessment of the cell permeability was deemed sufficient for these compounds as they were solely designed as tool compounds to study the biology of TRIM24, and not designed as drugs.

The IACS9571 positive control was detected in the cell lysate ($n$ = 2; Appendix 8, Figures A8.1–A8.3), as expected, indicating that the assay was functioning correctly. Evaluation of the PDCs demonstrated that PDC5 (compound **59**) was cell-permeable ($n$ = 2; Appendix 8, Figures A8.7–A8.9). However, PDCs 4 and 6 (compounds **58** and **60**) were not detected in the cell lysate ($n$ = 2). The parent BRD (compound **62**) and PHD (compound **67**) ligands were also assessed. While BRD ligand **62** showed cell permeability ($n$ = 2; Appendix 8, Figures A8.4–A8.6), peptide **67** was not observed in the cell lysate by LC-MS ($n$ = 2). Notably, ligand **62** was detected as the acetyl-deprotected compound (Appendix 8, Figures A8.5 and A8.6), suggesting that it might have been metabolised by endogenous deacetylating enzymes. Since PDC5, IACS9571, and BRD ligand **62** were found to be cell permeable, these compounds were taken forward to the cellular assays to evaluate the phenotypic effects of simultaneous PHD and BRD inhibition.

Firstly, it was necessary to evaluate compound cytotoxicity using an MTT assay. Due to issues with the procurement and culture of the MCF-7 cell line at the point of carrying out the MTT assays, an alternative $ER^+$ cell line was used (BT-474). A triple negative breast

cancer cell line (CAL-51) was also included for comparison. In line with previous data, TRIM24 BRD inhibition with IACS9571 had no effect on cell viability (Figure 4.21).[4] PDC5 also demonstrated no cytotoxic effects at 10 μM treatment in either cell line (Figure 4.21).

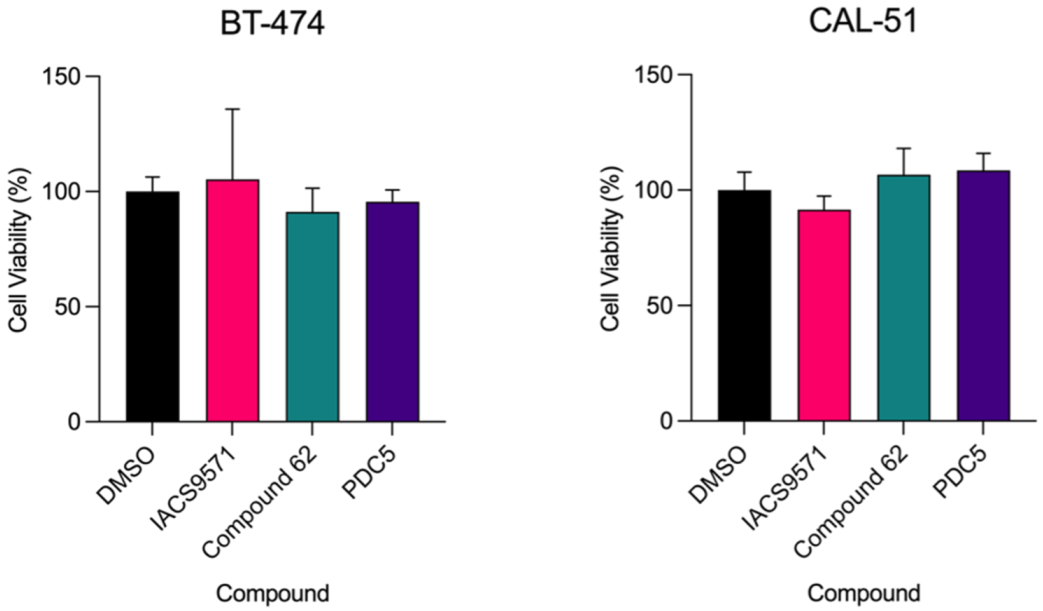



**Figure 4.21:** Bar graphs showing the results of the MTT assay for two breast cancer cell lines, BT-474 ($ER^+$) and CAL-51 (triple negative), treated with the indicated compound (10 μM). The data are represented as the mean of means for two biological repeats ($n = 2$), each consisting of three technical repeats. Error bars indicate the standard error of the mean.

Compounds were next assessed in proliferation assays to identify any antiproliferative effects on MCF-7 cells. With single compound treatment, PDC5 demonstrated no significant antiproliferative effects over 6 days (Figure 4.22). In accordance with the literature, IACS9571 also caused no change in proliferation (Figure 4.22).[4] It was hypothesised that PDC5 could be metabolised overtime in cells due to its peptidomimetic structure, therefore, a multiple dose experiment was also carried out to mitigate these effects. Nevertheless, PDC5 treatment caused no difference in proliferation compared to the vehicle control, IACS9571, or parent BRD ligand **62** (Figure 4.23).

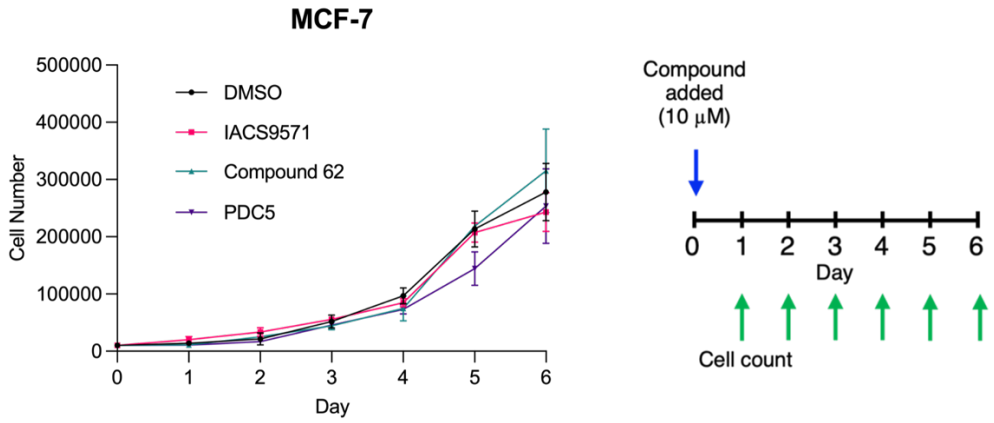



**Figure 4.22:** Graph showing the results of the proliferation assay for MCF-7 cells treated with the indicated compound (10 μM). Compound was added after seeding on day 0 only and counting was carried out each day for 6 days. Data are represented as the mean of triplicate data from one biological repeat. Error bars indicate the standard deviation.

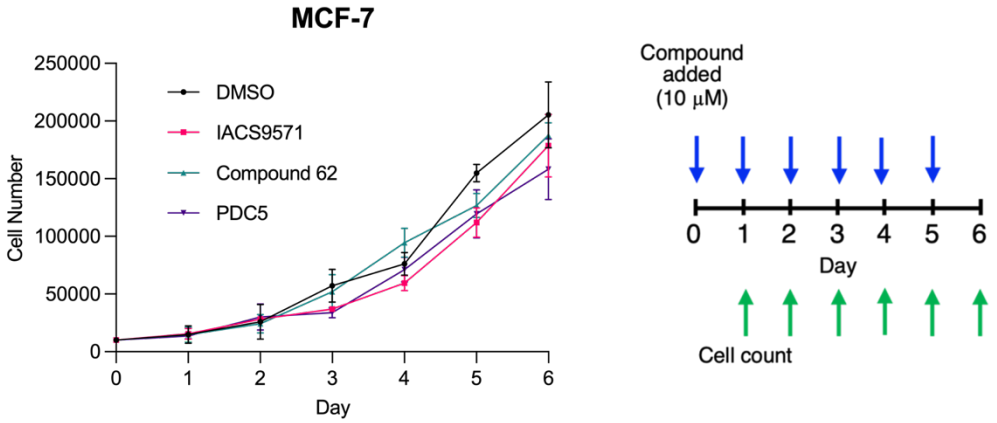



**Figure 4.23:** Graph showing the results of the proliferation assay for MCF-7 cells repeatedly treated with the indicated compound (10 μM) over 6 days. Compound was added each day for 6 days and counting was carried out prior to addition of compound on the following day. Data are represented as the mean of triplicate data from one biological repeat. Error bars indicate the standard deviation.

For completeness, clonogenic assays were carried out to evaluate the ability of cells to grow into colonies after compound treatment. For both the single and double dose experiments, no significant difference in the number of colonies was observed between cells treated with PDC5 and those treated with the DMSO control (Figure 4.24). This was emphasised by images of the stained colonies which clearly showed no antiproliferative effects from any of the tested compounds (Figure 4.24).

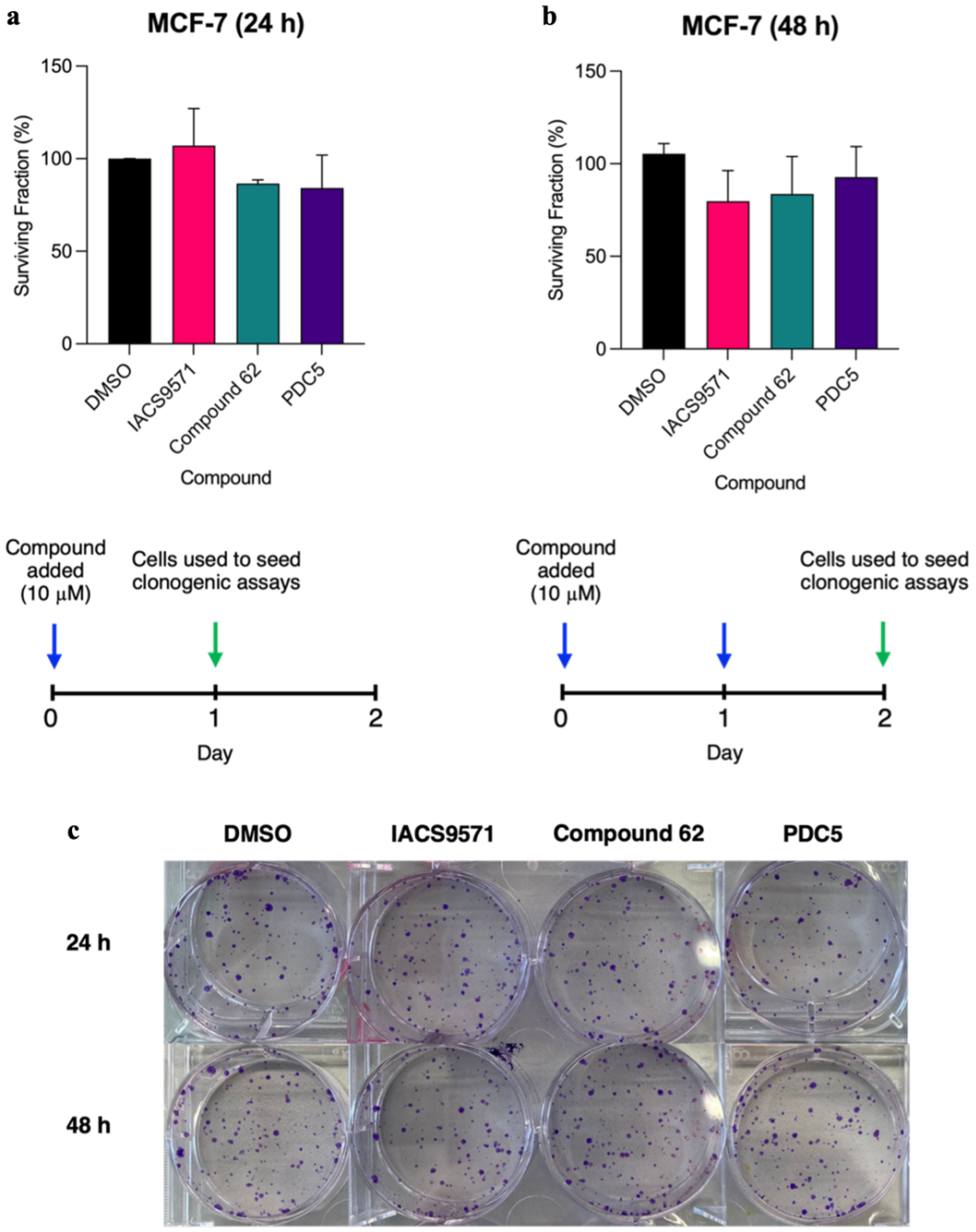



**Figure 4.24:** a) Bar graph showing the results of the single dose clonogenic assay for MCF-7 cells after treatment with the indicated compound (10 µM). Cells were treated with compound on day 0 only and after 24 h were used to seed plates for the clonogenic assay; b) Bar graph showing the results of the double dose clonogenic assay for MCF-7 cells after treatment with the indicated compound (10 µM). Cells were treated with compound on day 0 and day 1; and after a total of 48 h were used to seed plates for the clonogenic assay. The data are represented as the mean of means for two biological repeats ($n$ = 2), each consisting of three technical replicates. Error bars indicate the standard error of the mean; c) Representative images of the colonies from the MCF-7 clonogenic assays.

For comparison, compounds were also evaluated in the triple negative CAL-51 breast cancer cell line. A small statistically significant decrease in cell number was observed for PDC5 treatment compared to the vehicle (Figure 4.25). However, due to dwindling compound stocks only one biological repeat consisting of three technical replicates was carried out.

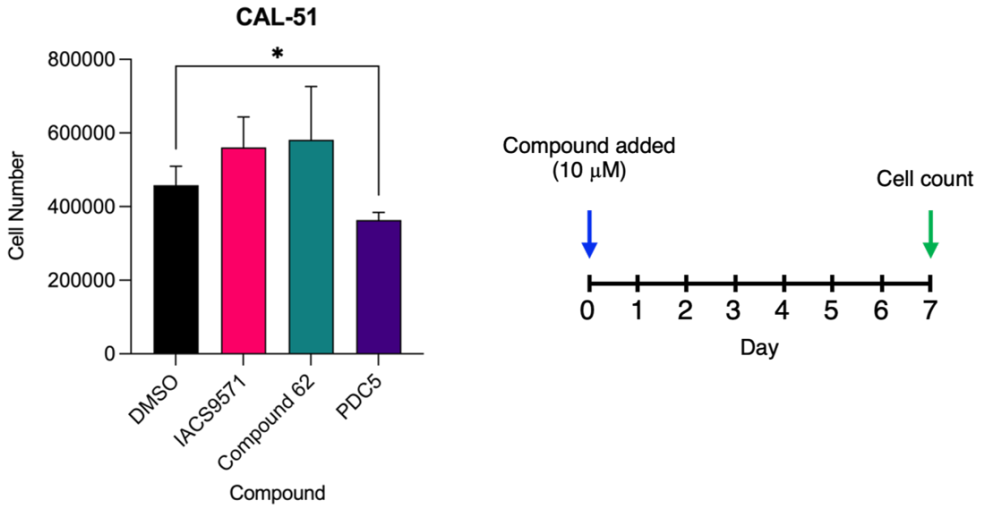



**Figure 4.25:** Bar graph showing the results of the proliferation assay for CAL-51 cells treated with the indicated compound (10 μM) for 7 days. Cells were treated with compound on day 0 only. The data are represented as the mean of triplicate data from one biological repeat, with error bars indicating the standard deviation. Statistical significance compared to the vehicle control was evaluated using an unpaired parametric t-test ($*$; $p \leq 0.05$).

A clonogenic assay was also carried out using the CAL-51 cell line. PDC5 again demonstrated a trend of reduced growth compared to the BRD ligands and DMSO control, however this difference was not statistically significant due to large errors in the data (Figure 4.26a). The images of the colonies tentatively suggest that PDC5 could be having some antiproliferative effect on CAL-51 cells (Figure 4.26b) *via* an ER-independent mechanism, however, further biological repeats would be needed to verify these data.

The lack of a significant antiproliferative phenotype could be due to metabolisation of PDC5 in cells. For instance PDC5 could be hydrolysed back to the respective monovalent inhibitors, which have significantly lower affinity for the TRIM24 PHD and BRD.

Alternatively, the MCF-7 cells used in these phenotypic assays may exhibit significantly lower H3K4K23Ac and TRIM24 expression levels compared to those used for the assays reported in the literature. Elevated TRIM24 and H3K4K23Ac expression is reported to drive $ER^+$ breast cancer cell proliferation by facilitating the recruitment of transcription factors to chromatin. Therefore, reduced expression levels may be limiting our ability to detect proliferation changes resulting from TRIM24 PHD-BRD inhibition.

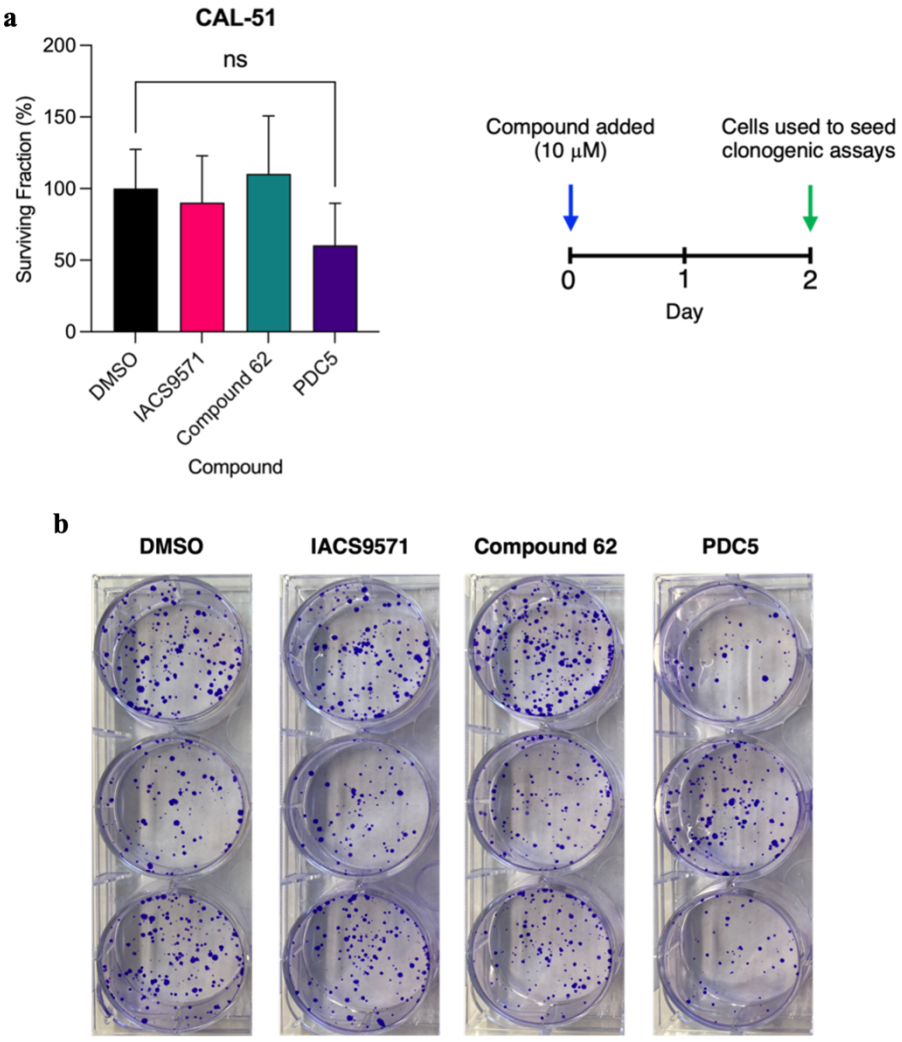



**Figure 4.26:** a) Bar graph showing the results of the clonogenic assay for CAL-51 cells after treatment with the indicated compound (10 μM) for 2 days. Cells were treated with compound on day 0 only and after 2 days were used to seed plates for the clonogenic assay. The data are represented as the mean of triplicate data from one biological repeat with error bars indicating the standard deviation. Statistical significance compared to the vehicle control was evaluated using an unpaired parametric t-test (ns; $p > 0.05$); b) Images of the colonies from the CAL-51 clonogenic assay. Each horizontal line of plates represents a technical replicate.

Nevertheless, given its exceptionally high affinity, sufficient PDC5 could still be present and inhibiting the binding of H3K4K23Ac to TRIM24. However, other domains in TRIM24, such as the RING domain, might also be responsible for driving MCF-7 cell proliferation and so PHD and BRD inhibition has no observable effect.

## 4.4 Summary

In summary, PDCs 4–6 (compounds **58**–**60**) were evaluated in biophysical assays to assess their affinities, selectivities, kinetics, and binding modes. In AlphaScreen® assays, the PDCs successfully inhibited the binding of histone-mimicking peptides to the PHD and BRD with low nanomolar $IC_{50}$ values. Notably, the PDCs could effectively inhibit the binding of $H3_{(1–27)}K9Me_3K18Ac$ (a dual PHD and BRD-binding peptide), whereas IACS9571 and parent BRD ligands **61**–**63** could not fully displace the peptide from the TRIM24 PHD and BRD. Although protein crystallography was unable to determine the binding mode of the PDCs, AlphaScreen® revealed avidity for TRIM24, which arises from an *in cis* binding mode. Furthermore, PDC6 displayed picomolar affinity for TRIM24 by SPR and BROMO*scan*, that stems from a slow $k_{off}$, which is consistent with a bivalent binding mode. BROMO*scan* also suggested that it might be possible to gain selectivity over the major off-target (BRFP1) through simultaneously binding to the TRIM24 PHD and BRD.

PDC5 showed cell permeability and so was taken forward for assessment in the cellular assays. Despite the previously established link between TRIM24 and breast cancer cell proliferation,[2] PDC5 showed no significant cytotoxic or antiproliferative effects on MCF-7 cells. Although PDC5 tentatively showed an antiproliferative effect on the triple negative CAL-51 cell line, further biological repeats would be needed to confirm conclusions from this data. The lack of a significant antiproliferative phenotype in MCF-7 cells could be due to a number of reasons, including poor physicochemical properties and metabolic stability.

Low expression of H3K4K23Ac and TRIM24 in the cell lines used may also contribute to the absence of an observed antiproliferative effect, although this would need to be validated using a combination of qPCR and Western Blotting. The data could also support the idea that other domains in TRIM24 might be responsible for regulating MCF-7 breast cancer cell proliferation, through an alternative ER-independent mechanism. Nevertheless, despite the lack of an induced phenotype, the work in this chapter demonstrates the proof of concept that PDCs can be used to simultaneously inhibit the TRIM24 PHD and BRD with high affinity. Future improvements to the physicochemical properties of these ligands could enable evaluation of the biology of TRIM24 in more detail.

# 5 Conclusions and Future Work

## 5.1 Conclusions

The aim of this study was to design and synthesise chemical tools to investigate the phenotypic effects of inhibiting *both* the TRIM24 PHD and BRD on breast cancer cells. This was achieved by developing a series of peptide-drug conjugates (PDCs), which could simultaneously bind to the TRIM24 PHD and BRD.

No small-molecule TRIM24 PHD ligands existed prior to this work, therefore, in Chapter 2 the objective was to develop histone-mimicking peptides with enhanced affinity for the TRIM24 PHD. Two peptidomimetic strategies were used to identify moieties that could interact with W828 and inhibit the cation-π interaction with $K9Me_3$ on histone H3. In the first method a focussed library of putative $KMe_3$ mimics were incorporated into histone-mimicking peptides, using cysteine-selective alkylation reactions, and screened for binding to the TRIM24 PHD using an AlphaScreen® competition assay. The data revealed that several mimics had increased affinity relative to the $KMe_3$ control peptide (**P1**, 4.99 ± 0.52 μM), including isoindoline (**P10**, 2.97 ± 0.21 μM) and pyrrolizidine (**P15**, 1.13 ± 0.08 μM). These mimics were taken forward into the synthesis of dual TRIM24 PHD and BRD-binding chemical tools. A second screen, involving ligand-directed dynamic fragment ligation, was also employed to probe a broader region of chemical space. This screen identified an aminothiazole moiety, which could prove useful in future for developing TRIM24 PHD inhibitors with improved physicochemical properties.

Chapter 3 described the design and synthesis of PDC inhibitors of the TRIM24 PHD and BRD. For the PHD-binding moiety, one of the highest affinity $KMe_3$ mimics (isoindoline), identified from Chapter 2, was incorporated into an unnatural amino acid using a 5-step

synthesis involving a key olefin metathesis step. The $KMe_3$-mimicking amino acid was attached, *via* PEG linkers, to a TRIM24 BRD ligand derived from the dTRIM24 PROTAC.[92] The PHD-binding motif was then extended to include a $H3_{(1-8)}$-mimicking peptide, which afforded a series of PDCs (PDCs 4–6, compounds **58**–**60**) targeting the TRIM24 PHD and BRD.

In Chapter 4, PDCs 4–6 (compounds **58**–**60**) were comprehensively assessed in biophysical assays to establish if they could simultaneously inhibit the TRIM24 PHD and BRD. In AlphaScreen® competition assays, PDCs 4–6 ($IC_{50}$ = 10.7–28.8 nM) inhibited the binding of a H3-mimicking peptide, bearing *both* $K9Me_3$ and K18Ac epigenetic marks, more effectively than the monovalent parent BRD (compounds **61**–**63**, $IC_{50}$ = 274–484 nM) and PHD (compound **67**, $IC_{50}$ = 4.90 μM) ligands. Furthermore, SPR, BROMO*scan*, and AlphaScreen® assays revealed that these PDCs have picomolar affinity [PDC6 $K_{d\ (SPR)}$ = 0.60 ± 0.20 nM; PDC6 $K_{d\ (BROMOscan)}$ = 0.18 ± 0.04 nM] for the TRIM24 PHD and BRD, that is driven by a combination of avidity and a slow $k_{off}$, resulting from an *in cis* bivalent binding mode. In subsequent cellular assays, the cell-permeant PDC5 (compound **59**) showed no antiproliferative effect on MCF-7 breast cancer cells, despite the fact that TRIM24 knockdown purportedly causes a reduction in MCF-7 cell proliferation. This suggested that other domains in TRIM24 (e.g. RING domain) could be driving proliferation in the MCF-7 cell line. Intriguingly, preliminary data for triple negative CAL-51 cells indicated that this cell line might be sensitive to treatment with PDC5, however, further biological repeats are needed to validate these data.

In conclusion, this work identified several $KMe_3$ bioisosteres, which enhanced the affinity of H3-mimicking peptides for the TRIM24 PHD. This enabled the development of the first bivalent PDC inhibitors of the TRIM24 PHD and BRD, with picomolar affinity, and some

selectivity over the analogous BRD of BRPF1. Subsequently, these chemical tools were used to evaluate the phenotypic effects of inhibiting *both* the PHD and BRD on breast cancer cells. Promisingly, in contrast to the TRIM24 BRD inhibitors, preliminary data suggested that PDC5 has an antiproliferative effect on the CAL-51 triple negative breast cancer cell line.

## 5.2 Future Work

The overarching goal of this work has been to validate TRIM24 as a therapeutic target by identifying if simultaneous PHD and BRD inhibition can cause an antiproliferative effect on breast cancer cells. Preliminary data has suggested a potential antiproliferative phenotype in the triple negative CAL-51 breast cancer cell line, however, the physicochemical properties of the PDCs such as metabolic stability and cell permeability might need to be improved for further investigation of TRIM24 as a therapeutic target. For instance, it is likely that the peptide and linker regions have poor metabolic stability meaning the PDCs could be cleaved back to the monovalent PHD and BRD ligands by endogenous peptidase enzymes. Metabolite identification (MetID) studies involving LC-MS analysis of cell lysates treated with the PDCs could help in identifying metabolic hotspots, which could then be optimised, for example by switching to non-proteinogenic D-amino acids, to improve overall compound stability.[170–172] On the other hand, cell permeability could be improved by methylating the peptide amide backbone and exchanging the PEG linker for an alkyl chain.[173,174]

The other proteins (TRIM33α/β, TRIM28 and TRIM66) in the TIF1 subfamily of TRIM proteins also contain a *C*-terminal adjacent PHD and BRD cassette.[1] Previous unpublished work in the Conway group has focussed on developing selective small molecule inhibitors of the TRIM33α/β BRDs.[141,175] Although these ligands were selective over the TRIM24

BRD, they exhibited low affinity for the TRIM33α/β BRDs, with sub-micromolar affinity proving elusive.[141,175] The PHD in TRIM33α/β is identical to that in TRIM24, therefore, a bivalent PDC approach targeting the TRIM33 PHD and BRD could be a potential gateway to higher affinity TRIM33 ligands.

The bivalent PDC strategy presented in this work could prove useful for developing inhibitors of other PHD domains, beyond those in TRIM proteins. Despite their high prevalence in the human proteome, PHDs represent a largely neglected class of epigenetic reader due to their poor ligandability.[41,48] However, many PHD fingers form part of a tandem reader cassette involving other reader domains with higher druggability such as BRDs, Tudor domains, and Chromodomains (Figure 5.1).[41,49,176] Moreover in 12 PHD-containing proteins a PHD is linked to a neighbouring BRD by fewer than 30 amino acids.[176] Although PHDs are challenging to target in isolation, a bivalent strategy, similar to that used in this project, could be employed using neighbouring domains as an anchor to enhance affinity for PHDs through avidity.

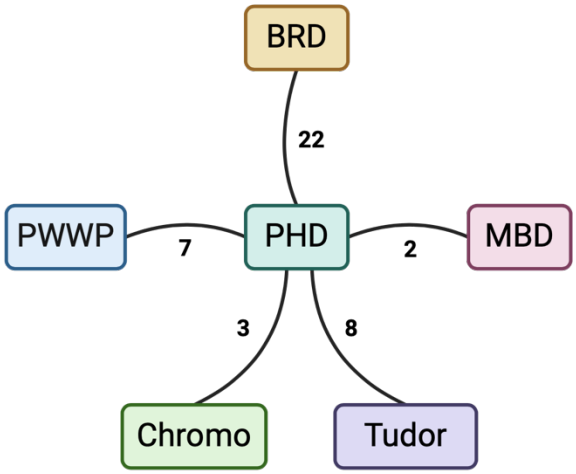



**Figure 5.1:** Schematic showing the occurrences of PHD domains that are linked to other reader domains within the human proteome. The number of times a PHD and another reader domain are found within the same protein is shown by the connecting line.[49] Figure created with BioRender.com

# 6 Experimental

## 6.1 General Chemistry Experimental

**Compound Names** were generated by ChemDraw 22.0.0 following IUPAC nomenclature.

**Reagents and Solvents** used were of commercially available reagent grade and were used without further purification, unless otherwise stated. All non-aqueous reactions were carried out under a nitrogen atmosphere in a flame-dried flask unless otherwise stated. Anhydrous solvents were obtained from an MBRAUN Solvent Purification System 5 and stored under an argon atmosphere over 3 Å molecular sieves. Degassed solvents were obtained by passing a steady stream of argon through the liquid, under sonication for 30 min. Where possible reactions were quenched/worked up after TLC analysis indicated complete consumption of the starting material. Concentration *in vacuo* was performed using a Buchi™ rotary evaporator with a water bath set at 40 °C, unless otherwise stated. Brine refers to saturated aqueous NaCl. Celite® refers to Celite® 545 filter aid, treated with sodium carbonate, flux-calcined (Sigma Aldrich). Lyophilisation was carried out on a Christ Alpha 3-4 LSCbasic Laboratory freeze dryer. Samples were dissolved in $H_2O$/MeCN and flash frozen in liquid nitrogen, before being lyophilised overnight.

**$^1$H NMR** spectra were measured on a Bruker AVIIIHD 400 (400 MHz), Bruker AVIIIHD 500 (500 MHz), or Bruker NEO 600 (600 MHz) spectrometer in the stated solvent. Chemical shifts (δ) are quoted in parts per million (ppm) relative to tetramethylsilane ($\delta_H$ = 0.00 ppm). Spectra were calibrated using residual protic solvent peaks and data provided by Fulmer *et al.*.[177] NMR data were processed using MestReNova 14.2.0 or Bruker Topspin 4.3.0 software. $^1$H NMR spectra are reported as follows: chemical shift (number of protons, multiplicity (s (singlet); br s (broad singlet); d (doublet); t (triplet); q (quartet); m (multiplet) or combinations thereof), coupling constants *J* (to the nearest 0.1

Hz, with identical coupling constants averaged), assignment). Spectra were assigned using COSY, HSQC and HMBC experiments as necessary.

**$^{13}$C NMR** spectra were measured on a Bruker AVIIIHD 500 (126 MHz) or Bruker NEO600 (151 MHz) spectrometer in the stated solvent and are broadband proton decoupled. Chemical shifts (δ) are quoted in parts per million (ppm) relative to tetramethylsilane ($\delta_C$ = 0.00 ppm). The spectra were calibrated using the solvent peak and data provided by Fulmer *et al.*.[177] NMR data were processed using MestReNova 14.2.0 or Bruker Topspin 4.3.0 software. The multiplicity of each signal is singlet unless otherwise stated. Spectra were assigned using HSQC and HMBC experiments as necessary.

**$^{19}$F NMR** spectra were recorded on a Bruker AVIIIHD 500 (470 MHz) or Bruker NEO600 (565 MHz) spectrometer in the stated solvent and are broadband proton decoupled. Chemical shifts (δ) are quoted in parts per million (ppm). NMR data were processed using MestReNova 14.2.0 or Bruker Topspin 4.3.0 software. The multiplicity of each signal is singlet unless otherwise stated.

**Mass Spectra** were acquired on either an Agilent 6120 (low resolution), Waters LCT Premier XE bench-top orthogonal acceleration TOF LC-MS (low resolution), Bruker microToF (high resolution), or Waters BioAccord LC-MS (high resolution) spectrometer using electrospray ionisation (ESI) or electron ionisation (EI) from solutions of MeOH, MeCN or $H_2O$. *m/z* values are reported in Daltons and followed by their percentage abundance in parentheses.

**Liquid Chromatography-Mass Spectrometry (LC-MS)** was carried out on an Agilent 1260 Infinity II system with a Quaternary LC Pump, DADWR, and InfinityLab LC/MSD. An Agilent InfinityLab Poroshell 120 SB-C18 [2.7 μm, 2.1 mm × 50 mm] column was used with a constant flow rate of 0.4 mL $min^{-1}$ and a gradient method of 10 min from 5–

95% MeCN/$H_2O$ +0.1% formic acid with a 5 min hold. Samples were prepared by dissolving in MeOH, MeCN or $H_2O$, and filtered. LC-MS data were processed using Agilent OpenLab ChemStation software.

**Melting Points** were determined using a Griffin capillary tube melting point apparatus and are uncorrected. The crystallisation solvent is given in parentheses.

**Infrared Spectra** were obtained from thin films or neat samples, using a diamond ATR module. The spectra were recorded on a Bruker Tensor 27 spectrometer. Absorption maxima ($\tilde{\nu}_{max}$) are reported in wavenumbers ($cm^{-1}$) and are reported as s (strong), m (medium), w (weak) or br (broad).

**Specific Optical Rotations** were measured using a Schmidt Haensch Unipol polarimeter, using a sodium lamp at 589 nm and a path length of 1.0 dm. The concentration (*c*) is expressed in g/100 mL (equivalent to g/0.1 $dm^3$). Specific rotations are quoted as an average of 10 measurements. They are denoted $[\alpha]_D^T$ and are given in implied units of $10^{-1}$ ° $cm^2$ $g^{-1}$ (where T = 25 °C).

**Analytical HPLC** was performed on an Agilent 1260 Infinity II system with a Quaternary LC Pump, DADWR and Agilent Infinity Fraction Collector. For determination of purity, an Agilent InfinityLab Poroshell 120 EC-C18 [4 μm, 4.6 mm × 150 mm] reverse phase column was used with a constant flow rate of 1.0 mL $min^{-1}$ and gradient method of 10 min from 5–95% MeCN/$H_2O$ +0.1% formic acid with a 5 min hold. Samples were prepared by dissolving in MeOH, MeCN or $H_2O$, and filtered. HPLC data were processed using Agilent OpenLab ChemStation software.

**Semi-Preparative Reverse Phase HPLC** was performed on an Agilent 1260 Infinity II Semi-Prep system with a binary LC pump, DADWR and Agilent Infinity Fraction

Collector. Samples were prepared by dissolving in MeOH, MeCN, $H_2O$, or DMF, and filtered. Injection volumes of 500–1500 μL were used depending upon the resolution of the chromatograms.

**Table E1:** Gradient Method 1.

| Time (min) | %$H_2O$ (+0.1% FA) | %MeCN (+0.1% FA) |
|---|---|---|
| 0.00 | 95 | 5 |
| 1.00 | 95 | 5 |
| 11.00 | 5 | 95 |
| 18.00 | 5 | 95 |

**Table E2:** Gradient Method 2.

| Time (min) | %$H_2O$ (+0.1% FA) | %MeCN (+0.1% FA) |
|---|---|---|
| 0.00 | 95 | 5 |
| 1.00 | 95 | 5 |
| 11.00 | 5 | 95 |
| 15.00 | 5 | 95 |

**Table E3:** Gradient Method 3.

| Time (min) | %$H_2O$ (+0.1% TFA) | %MeCN (+0.1% TFA) |
|---|---|---|
| 0.00 | 95 | 5 |
| 2.00 | 95 | 5 |
| 2.01 | 80 | 20 |
| 7.00 | 60 | 40 |
| 10.00 | 5 | 95 |

**Table E4:** Gradient Method 4.

| Time (min) | %$H_2O$ (+0.1% FA) | %MeCN (+0.1% FA) |
|---|---|---|
| 0.00 | 95 | 5 |
| 4.00 | 95 | 5 |
| 4.01 | 50 | 50 |
| 10.00 | 5 | 95 |
| 15.00 | 5 | 95 |

**Analytical Thin-Layer Chromatography (TLC)** was carried out on normal phase or reverse phase Merck silica ($SiO_2$) 60 $F_{254}$ aluminium-supported thin layer chromatography sheets. Visualisation was carried out by absorption of UV light ($\lambda_{max}$ = 254 nm) or thermal development after staining with an ethanolic ninhydrin solution, basic aqueous $KMnO_4$ or acidic aqueous 2,4-dinitrophenylhydrazine.

**Normal Phase Flash Column Chromatography** was carried out on Merck Geduran® silica gel 60 (40–63 μm) or on an automated Biotage® Selekt system with Biotage® Sfär silica columns.

**Solid Phase Extraction** was carried out using the specified columns (Biotage®). Columns were equilibrated with $CH_2Cl_2$ (1 CV) before samples were loaded. The column was then washed with $CH_2Cl_2$ (2 CV) and the purified product was eluted with the specified solvent.

## 6.2 General Chemistry Procedures

**General Bromination Procedure :** The specified amine (1.0 eq.) was added dropwise to aqueous HBr (48% *w/w*) at 0 °C. The reaction solution was stirred and heated under reflux at 135 °C for 20 h. Excess HBr was removed by distillation at atmospheric pressure under a stream of nitrogen. The residue was cooled to rt, dissolved in MeOH (10 mL), and concentrated *in vacuo*. The residue was washed with IPA (30 mL) to yield the brominated product.

**General Amide Coupling Procedure 1:** Anhydrous DIPEA (3.0 eq.) and HATU (1.0 eq.) were added to a solution of the specified carboxylic acid (1.0 eq.) and amine (1.0 eq.) in anhydrous DMF (5 mL) at rt. The reaction mixture was stirred at rt for 24 h. The reaction mixture was diluted with EtOAc (25 mL) and washed with aqueous LiCl (0.5 M, 25 mL), aqueous citric acid (10% *w/v*, 25 mL), and brine (25 mL). The organic component was dried ($Na_2SO_4$), filtered, and concentrated *in vacuo* to yield the crude amide.

**General Amide Coupling Procedure 2:** Anhydrous DIPEA (5.0 eq.) and HATU (1.0 eq.) were added to a stirring solution of the specified crude deprotected carboxylic acid (1.0 eq.) in anhydrous DMF (3 mL). A solution of the specified crude deprotected amine (1.0 eq.) in anhydrous DMF (2 mL) was added dropwise, and the reaction mixture was stirred at rt for 24 h. The reaction mixture was diluted with EtOAc (25 mL) and washed with aqueous LiCl (0.5 M, 25 mL), aqueous citric acid (10% *w/v*, 25 mL), and brine (25 mL). The organic component was dried ($Na_2SO_4$), filtered, and concentrated *in vacuo* to yield the crude amide

**General Amide Coupling Procedure 3:** HATU (1.0 eq.) was added to a stirring solution of the specified deprotected amine (1.0 eq.), peptide **48** (1.0 eq.), and anhydrous DIPEA (13.2 µL, 75.5 µmol, 5.0 eq.) in anhydrous $CH_2Cl_2$ (1 mL) at rt. The reaction solution was stirred at rt for 24 h. The reaction solution was concentrated *in vacuo*.

**General Boc Deprotection Procedure:** The specified Boc protected amine (1.0 eq.) was dissolved in a solution of HCl in dioxane (4 M, 1 mL) and stirred at rt for 2 h. The reaction mixture was concentrated *in vacuo* to yield the crude amine as the HCl salt.

**General Acetylation Procedure:** The specified crude deprotected amine was dissolved in $CH_2Cl_2$ (2 mL), and $NEt_3$ (2.0 eq.) was added. The reaction solution was cooled to 0 °C and acetyl chloride (1.0 eq.) was added dropwise. The reaction solution was warmed to rt and stirred for 20 min. The reaction solution was quenched with $H_2O$ (500 µL) and concentrated *in vacuo* to yield the crude acetylated product.

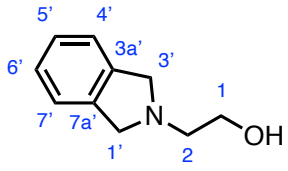
**General $^t$Bu Ester Deprotection Procedure:** The specified $^t$Bu carboxylate ester (1.0 eq.) was dissolved in a solution of HCl in dioxane (4 M, 1 mL) at rt. The reaction solution was stirred at rt for 24 h and concentrated *in vacuo.* The residue was dried overnight under high vacuum to yield the crude carboxylic acid.

**General Fmoc Deprotection Procedure:** The specified Fmoc-protected amine was dissolved in a solution of piperidine in anhydrous DMF (20% *v/v*, 1 mL) and stirred at rt for 1 h. The reaction solution was concentrated *in vacuo*. The crude amine was purified using semi-preparative reverse phase HPLC (gradient method 2, 20 mL $min^{-1}$, Agilent Prep C18 column [5 µm, 21.2 mm × 50 mm]) and lyophilised to yield the purified amine as the diformate salt.

**Global Deprotection Procedure:** The specified protected peptide (1 eq.) was dissolved in a solution of TFA/TIPS/$H_2O$ (38:1:1, 1 mL) and stirred at rt for 4 h. The reaction solution was concentrated *in vacuo* to yield the crude product.

## 6.3 Synthetic Procedures and Compound Characterisation

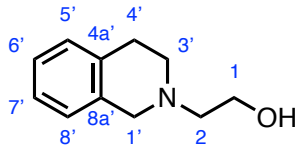
### (1) 2-(Isoindolin-2'-yl)ethan-1-ol




Isoindoline (952 µL, 8.39 mmol, 1.0 eq.) was added dropwise to a solution of 2-bromoethanol (595 µL, 8.39 mmol, 1.0 eq.), and $NEt_3$ (2.30 mL, 16.8 mmol, 2.0 eq.) in toluene (50 mL) at rt. The reaction solution was stirred and heated under reflux at 120 °C for 6 h. The reaction mixture was cooled to rt, filtered, and concentrated *in vacuo*. The residue was dissolved in EtOAc (30 mL) and washed with $H_2O$ (50 mL). The organic components were dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 2–20% MeOH/$CHCl_3$) to yield the title compound as a brown oil (600 mg, 44%): $R_f$ 0.29 (MeOH:$CHCl_3$, 1:9); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.24–7.20 (4H, m, H-4', 5', 6', 7'), 4.03 (4H, s, H-1', 3'), 3.73 (2H, t, *J* 5.4, H-1), 2.96 (2H, t, *J* 5.4, H-2),

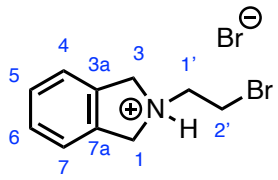
2.64 (1H, br s, O<u>H</u>); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 139.7 (C-3a', 7a'), 127.1 (C-4', 7' / C-5', 6'), 122.4 (C-4', 7' / C-5', 6'), 59.9 (C-1), 59.1 (C-1', 3'), 57.6 (C-2); LRMS *m/z* (ESI$^+$) 164 ([M+H]$^+$, 89%), 349 ([2M+Na]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[178]

**(2) 2-(3',4'-Dihydroisoquinolin-2'(1'*H*)-yl)ethan-1-ol**



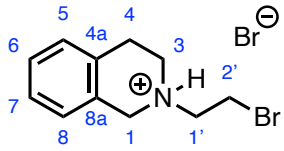

1,2,3,4-Tetrahydroisoquinoline (940 µL, 7.51 mmol, 1.0 eq.) was added dropwise to a solution of 2-bromoethanol (532 µL, 7.51 mmol, 1.0 eq.), and $NEt_3$ (2.10 mL, 15.0 mmol, 2.0 eq.) in toluene (50 mL) at rt. The reaction solution was stirred and heated under reflux at 120 °C for 20 h. The reaction mixture was cooled to rt, filtered, and concentrated *in vacuo*. The residue was dissolved in EtOAc (30 mL) and washed with $H_2O$ (50 mL). The organic components were dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 2–20% MeOH/$CHCl_3$) to yield the title compound as a yellow oil (978 mg, 74%): $R_f$ 0.48 (MeOH:$CHCl_3$, 1:9); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 3389 (O-H, br, w), 2922 (C-H, w), 741 (Ar C-H, s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.18–7.07 (3H, m, H-5', 6', 7'), 7.05–7.00 (1H, m, H-8'), 3.73–3.68 (4H, m, H-1, 1'), 2.92 (2H, t, *J* 5.9, H-4'), 2.82 (2H, t, *J* 5.9, H-3'), 2.74–2.70 (2H, m, H-2), 2.43 (1H, br s, O<u>H</u>); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 134.6 (C-4a' / C-8a'), 134.4 (C-4a' / C-8a'), 128.8 (C-5'), 126.7 (C-8'), 126.4 (C-7'), 125.9 (C-6'), 59.2 (C-2), 58.2 (C-1), 55.9 (C-1'), 50.8 (C-3'), 29.2 (C-4'); HRMS *m/z* (ESI$^+$) [Found: 178.1228, $C_{11}H_{16}NO$ requires [M+H]$^+$ 178.1226]; LRMS *m/z* (ESI$^+$) 178 ([M+H]$^+$, 39%), 222 ([M+2Na-H], 41%). The spectroscopic data are in good agreement with the literature values.[179]

**(3) 2-(2'-Bromoethyl)isoindoline hydrobromide**




Compound **1** (600 mg, 3.68 mmol, 1.0 eq.) was reacted with aqueous HBr (48% *w/w*, 20 mL) according to the general bromination procedure and then crystalised from hot MeOH to yield the title compound as a beige crystalline solid (389 mg, 34%): $R_f$ 0.85 (MeOH:$CH_2Cl_2$, 1:9); m.p. 182–185 °C (from MeOH); $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 3381 (N-H, br, w), 2464 (C-H, w), 767 (Ar C-H, s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 13.01 (1H, br s, NH), 7.43–7.37 (2H, m, H-5, 6), 7.35–7.29 (2H, m, H-4, 7), 5.13 (2H, dd, *J* 13.9, 6.0, H-1$^a$, 3$^a$), 4.37 (2H, dd, *J* 13.9, 5.6, H-1$^b$, 3$^b$), 3.96 (2H, t, *J* 7.1, H-2'), 3.79–3.71 (2H, m, H-1'); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 132.3 (C-3a, 7a), 129.8 (C-5, 6), 123.2 (C-4, 7), 59.0 (C-1, 3), 56.9 (C-1'), 22.7 (C-2'); HRMS *m/z* (ESI$^+$) [Found: 226.0228, $C_{10}H_{13}N^{79}Br$ requires $[M-HBr+H]^+$ 226.0226]; LRMS *m/z* (ESI$^+$) 226 ($[M(^{79}Br)-HBr+H]^+$, 57%), 228 ($[M(^{81}Br)-HBr+H]^+$, 59%).

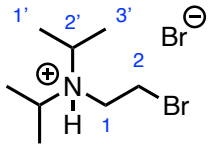
**(4) 2-(2'-Bromoethyl)-1,2,3,4-tetrahydroisoquinoline hydrobromide**




Compound **2** (978 mg, 5.52 mmol, 1.0 eq.) was reacted with aqueous HBr (48% *w/w*, 20 mL) according to the general bromination procedure and then crystallised from hot MeOH to yield the title compound as a colourless crystalline solid (929 mg, 52%): $R_f$ 0.85 (MeOH:$CH_2Cl_2$, 1:9); m.p. 229–232 °C (from MeOH); $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 3658 (N-H, w), 2980 (C-H, s), 753 (Ar C-H, w); $^1$H NMR (600 MHz, $D_6$-DMSO) $\delta_H$ 10.17 (1H, br s, NH), 7.34–7.23 (3H,

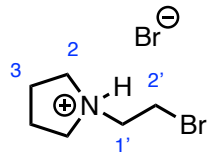
m, H-5, 6, 7), 7.22–7.16 (1H, m, H-8), 4.61 (1H, d, *J* 15.5, H-1$^{a}$), 4.47–4.35 (1H, m, H-1$^{b}$), 3.96–3.64 (5H, m, H-1', 2' H-3$^{a}$), 3.42$^{*}$ (1H, m, H-3$^{b}$), 3.25–3.00 (2H, m, H-4$^{a}$, 4$^{b}$); $^{13}$C NMR (151 MHz, D$_6$-DMSO) $\delta_C$ 131.1 (C-4a / C-8a), 128.6 (C-5), 128.0 (C-4a / C-8a), 127.9 (C-8), 126.7 (C-7), 126.5 (C-6), 55.5 (C-1'), 52.0 (C-1), 49.1 (C-3), 24.7 (C-4), 24.2 (C-2'); HRMS *m/z* (ESI$^{+}$) [Found: 240.0383, $C_{11}H_{15}N^{79}Br$ requires $[M-HBr+H]^{+}$ 240.0382], [Found: 242.0362, $C_{11}H_{15}N^{81}Br$ requires $[M-HBr+H]^{+}$ 242.0362]; LRMS *m/z* (ESI$^{+}$) 240 ($[M(^{79}Br)-HBr+H]^{+}$, 64%), 242 ($[M(^{81}Br)-HBr+H]^{+}$, 66%). *Peak obscured by NMR solvent peak and was assigned by HSQC.

### (5) *N*-(2-Bromoethyl)-*N*-isopropylpropan-2'-amine hydrobromide



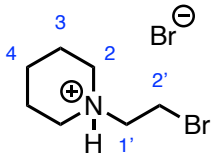

*N,N*-Diisopropylaminoethanol (1.00 mL, 5.69 mmol, 1.0 eq.) was reacted with aqueous HBr (48% *w/w*, 20 mL) according to the general bromination procedure to yield the title compound as a colourless crystalline solid (647 mg, 39%): $R_f$ 0.28 (MeOH:$CH_2Cl_2$, 1:9); m.p. 129–132 °C (from MeOH) [lit.[180] 141 °C (from acetone), lit.[181] 140–142 °C (solvent unspecified)]; $\tilde{\nu}_{max}$ (thin film)/cm$^{-1}$ 2925 (C-H, w), 2612 (w), 2506 (w); $^{1}$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 10.97 (1H, br s, N<u>H</u>), 4.02–3.92 (2H, m, H-2), 3.78–3.62 (2H, m, H-2'), 3.38–3.28 (2H, m, H-1), 1.61 (6H, d, *J* 6.1, H-1' / H-3'), 1.49 (6H, d, *J* 6.1, H-1' / H-3'); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 55.7 (C-2'), 49.0 (C-1), 24.6 (C-2), 18.8 (C-1' / C-3'), 17.2 (C-1' / C-3'); HRMS *m/z* (ESI$^{+}$) [Found: 208.0696, $C_8H_{19}N^{79}Br$ requires $[M-HBr+H]^{+}$ 208.0695]; LRMS *m/z* (ESI$^{+}$) 208 ($[M(^{79}Br)-HBr+H]^{+}$, 100%), 210 ($[M(^{81}Br)-HBr+H]^{+}$, 93%).

**(6) 1-(2'-Bromoethyl)pyrrolidine hydrobromide**




1-(2-Hydroxyethyl)pyrrolidine (1.00 mL, 8.55 mmol, 1.0 eq.) was reacted with aqueous HBr (48% *w/w*, 20 mL) according to the general bromination procedure to yield the title compound as a colourless crystalline solid (1.35 g, 61%): $R_f$ 0.42 (MeOH:$CH_2Cl_2$, 1:9); m.p. 162–165 °C (from MeOH); $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 2947 (C–H, w), 2846 (C-H, w), 2666 (w), 2583 (w), 2490 (w), 1453 (w); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 11.87 (1H, br s, NH), 3.95–3.81 (4H, m, H-2[a], 2'), 3.60–3.50 (2H, m, H-1'), 3.00–2.86 (2H, m, H-2[b]), 2.32–2.06 (4H, m, H-3[a], 3[b]); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 56.2 (C-1'), 54.3 (C-2), 23.4 (C-3), 23.0 (C-2'); HRMS *m/z* ($ESI^+$) [Found: 178.0227, $C_6H_{13}N^{79}Br$ requires $[M-HBr+H]^+$ 178.0226], [Found: 180.0205, $C_6H_{13}N^{81}Br$ requires $[M-HBr+H]^+$ 180.0205]; LRMS *m/z* ($ESI^+$) 178 ($[M(^{79}Br)-HBr+H]^+$, 96%), 180 ($[M(^{81}Br)-HBr+H]^+$, 100%).

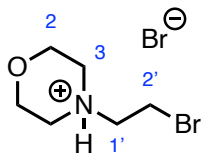
**(7) 1-(2'-Bromoethyl)piperidine hydrobromide**




1-(2-Hydroxyethyl)piperidine (1.00 mL, 7.53 mmol, 1.0 eq.) was reacted with aqueous HBr (48% *w/w*, 20 mL) according to the general bromination procedure to yield the title compound as a colourless crystalline solid (1.61 g, 78%): $R_f$ 0.55 (MeOH:$CH_2Cl_2$, 1:9); m.p. 226–229 °C (from MeOH) [lit.[182] 238 °C (solvent unspecified), lit.[183] 242–244 (from EtOH)]; $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 2922 (C-H, w); 2660 (w), 2575 (w), 2547 (w), 1459 (w); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 11.65 (1H, br s, NH), 3.94–3.88 (2H, t, *J* 7.3, H-2'), 3.68–3.60 (2H,

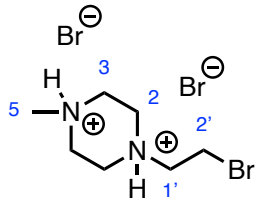
m, H-2[a]), 3.46–3.38 (2H, m, H-1'), 2.84–2.74 (2H, m, H-2[b]), 2.39–2.26 (2H, m, H-3[a]), 1.98–1.86 (3H, m, H-3[b], 4[a]), 1.52–1.42 (1H, m, H-4[b]); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 58.3 (C-1'), 53.9 (C-2), 22.6 (C-3), 22.0 (C-4), 21.9 (C-2'); HRMS *m/z* ($ESI^+$) [Found: 192.0382, $C_7H_{15}N^{79}Br$ requires $[M-HBr+H]^+$ 192.0382], [Found: 194.0361, $C_7H_{15}N^{81}Br$ requires $[M-HBr+H]^+$ 194.0362]; LRMS *m/z* ($ESI^+$) 192 ($[M(^{79}Br)-HBr+H]^+$, 30%), 194 ($[M(^{81}Br)-HBr+H]^+$, 51%).

**(8) 4-(2'-Bromoethyl)morpholine hydrobromide**



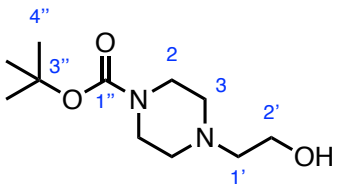

4-(2-Hydroxyethyl)morpholine (1.00 mL, 8.26 mmol, 1.0 eq.) was added dropwise to a solution of triphenylphosphine dibromide (4.18 g, 9.91 mmol, 1.2 eq.) in anhydrous $CH_2Cl_2$ (50 mL) at 0 °C. The reaction solution was warmed to rt and stirred for 6 h. The precipitate was collected by filtration and washed with $CH_2Cl_2$ (30 mL) to yield the title compound as an off-white solid (1.59 g, 70%): $R_f$ 0.59 (MeOH:$CH_2Cl_2$, 1:9); m.p. 201–204 °C (from $CH_2Cl_2$); $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 2971 (C-H, w), 1114 (s), 873 (s); $^1$H NMR (600 MHz, D$_6$-DMSO) $\delta_H$ 10.09 (1H, br s, NH), 4.04–3.91 (2H, m, H-2[a]), 3.82 (2H, t, *J* 7.1, H-2'), 3.78–3.56 (4H, m, H-1', H-2[b]), 3.25–3.06 (2H, m, H-3[a]), 3.16* (2H, m, H-3[b]); $^{13}$C NMR (151 MHz, D$_6$-DMSO) $\delta_C$ 63.1 (C-2), 56.2 (C-1'), 51.1 (C-3), 23.8 (C-2'); HRMS *m/z* ($ESI^+$) [Found: 194.0174, $C_6H_{13}NO^{79}Br$ requires $[M-HBr+H]^+$ 194.0175], [Found: 196.0152, $C_6H_{13}NO^{81}Br$ requires $[M-HBr+H]^+$ 196.0155]; LRMS *m/z* ($ESI^+$) 194 ($[M(^{79}Br)-HBr+H]^+$, 64%), 196 ($[M(^{81}Br)-HBr+H]^+$, 67%). The spectroscopic data are in good agreement with the literature values.[127] *Peak obscured by NMR solvent peak and was assigned by HSQC.

**(9) 1-(2'-Bromoethyl)-4-methylpiperazine dihydrobromide**




1-(2-Hydroxyethyl)-4-methylpiperazine (1.00 g, 6.93 mmol, 1.0 eq.) was added dropwise to a solution of triphenylphosphine dibromide (3.51 g, 8.32 mmol, 1.2 eq.) in $CH_2Cl_2$ (50 mL) at 0 °C. The reaction mixture was warmed to rt and stirred for 20 h. The reaction mixture was filtered, and the precipitate was washed with $CH_2Cl_2$ (30 mL) and MeOH (50 mL). The product was crystalised from hot MeOH to yield the title compound as a colourless crystalline solid (630 mg, 44%): $R_f$ 0.15 (MeOH:$CH_2Cl_2$, 1:9); m.p. > 250 °C (from MeOH); $\tilde{\nu}_{max}$ (thin film)/cm$^{-1}$ 2975 (C-H, w), 1454 (m), 1021 (m), 957 (s); $^1$H NMR (600 MHz, D$_6$-DMSO) $\delta_H$ 3.90–3.58 (12H, m, H-2, 3, 1', 2'), 3.26 (3H, s, H-5); $^{13}$C NMR (151 MHz, D$_6$-DMSO) $\delta_C$ 54.7 (C-1'), 52.1 (C-2), 51.4 (C-5), 43.1 (C-3), 24.6 (C-2'); HRMS *m/z* (ESI$^+$) [Found: 207.0490, $C_7H_{16}N_2{}^{79}Br$ requires [M−2HBr+H]$^+$ 207.0491], [Found: 209.0469, $C_7H_{16}N_2{}^{81}Br$ requires [M−2HBr+H]$^+$ 209.0471]; LRMS *m/z* (ESI$^+$) 207 ([M($^{79}$Br)−2HBr+H]$^+$, 45%), 209 ([M($^{81}$Br)−2HBr+H]$^+$, 46%).

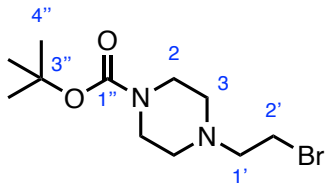
**(10) *tert*-Butyl 4-(2'-hydroxyethyl)piperazine-1-carboxylate**




Di-*tert*-butyl dicarbonate (3.56 g, 16.3 mmol, 1.0 eq.) was added portionwise to a solution of 1-(2-hydroxyethyl)piperazine (2.00 mL, 16.3 mmol, 1.0 eq.) in THF (30 mL) at 0 °C. The reaction solution was stirred at 0 °C for 2 h, after which time the solvent was

evaporated *in vacuo* and the residue was triturated with petroleum ether 40–60 °C. The mixture was filtered, and the filtrate was concentrated *in vacuo*. The residue was crystalised overnight at −20 °C to yield the title compound as a colourless crystalline solid (3.71 g, 99%): $R_f$ 0.35 (MeOH:$CH_2Cl_2$, 1:9); m.p. 31–34 °C (from petroleum ether 40–60 °C) [lit.[184] 45–52 °C (solvent unspecified)]; $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 3.62 (2H, t, *J* 5.4, H-2'), 3.43 (4H, t, *J* 4.8, H-2), 2.54 (2H, t, *J* 5.4, H-1'), 2.44 (4H, t, *J* 4.8, H-3), 1.45 (9H s, H-4''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 154.8 (C-1''), 79.8 (C-3''), 59.5 (C-1'), 57.9 (C-2'), 52.8 (C-3), 44.2* (C-2$^I$), 43.4* (C-2$^{II}$), 28.6 (C-4''); HRMS *m/z* (ESI$^+$) [Found: 231.1705, $C_{11}H_{23}N_2O_3$ requires $[M+H]^+$ 231.1703]; LRMS *m/z* (ESI$^+$) 231 ($[M+H]^+$, 57%), 253 ($[M+Na]^+$, 58%), 483 ($[2M+Na]^+$, 100%). The spectroscopic data are in good agreement with the literature values.[184,185] *Peaks coalesce to form a singlet at 363 K.

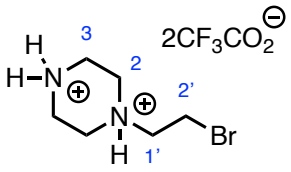
### (11) *tert*-Butyl 4-(2'-bromoethyl)piperazine-1-carboxylate




A solution of tetrabromomethane (1.58 g, 4.76 mmol, 1.1 eq.) in anhydrous $CH_2Cl_2$ (10 mL) was added dropwise to a solution of compound **10** (1.00 g, 4.34 mmol, 1.0 eq.) and triphenylphosphine (1.23 g, 4.69 mmol, 1.1 eq.) in anhydrous $CH_2Cl_2$ (10 mL) at 0 °C. The reaction solution was warmed to rt and stirred for 20 h. The solvent was evaporated *in vacuo* and the product was purified using flash silica column chromatography (gradient elution with 10–80% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a yellow crystalline solid (1.15 g, 90%): $R_f$ 0.42 (EtOAc:petroleum ether 40–60 °C, 2:3); m.p. 45–48 °C (from $CH_2Cl_2$); $\tilde{\nu}_{max}$ (thin film)/cm$^{-1}$ 2980 (C-H, w), 2814 (C-H, w), 1685 (C=O, s), 1415 (m), 1243 (m), 1165 (s), 1130 (s), 1000 (m), 772 (w), 657 (w); $^1$H NMR

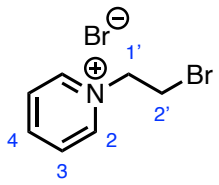
(600 MHz, $CDCl_3$) $\delta_H$ 3.47–3.38 (6H, m, H-2, 2'), 2.79 (2H, t, *J* 7.4, H-1'), 2.45 (4H, t, *J* 5.1, H-3), 1.45 (9H, s, H-4''); $^{13}C$ NMR (151 MHz, $CDCl_3$) $\delta_C$ 154.8 (C-1''), 79.9 (C-3''), 60.0 (C-1'), 52.9 (C-3), 44.1* (C-$2^I$), 43.2* (C-$2^{II}$), 28.9 (C-2'), 28.6 (C-4''); HRMS *m/z* ($ESI^+$) [Found: 293.0856, $C_{11}H_{22}N_2O_2{}^{79}Br$ requires $[M+H]^+$ 293.0859], [Found: 295.0836, $C_{11}H_{22}N_2O_2{}^{81}Br$ requires $[M+H]^+$ 295.0839]; LRMS *m/z* ($ESI^+$) 293 ($[M(^{79}Br)+H]^+$, 99%), 295 ($[M(^{81}Br)+H]^+$, 100%). The spectroscopic data are in good agreement with the literature values.[186] * Peaks coalesce to form a singlet at 363 K.

### (12) 1-(2'-Bromoethyl)piperazine ditrifluoroacetate



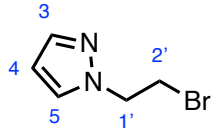

TFA (5 mL) was added dropwise to a solution of compound **11** (1.00 g, 3.41 mmol, 1.0 eq.) in $CH_2Cl_2$ at 0 °C. The reaction solution was warmed to rt and stirred for 2 h. The reaction solution was concentrated *in vacuo* and the residue was left under high vacuum overnight. The product was crystalised from $Et_2O$ and collected by filtration to yield the title compound as an off-white solid (1.21 g, 84%): $R_f$ 0.22 (MeOH:$CH_2Cl_2$, 1:9); m.p. 104–107 °C (from $Et_2O$); $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 3033 (w), 2471 (w), 1668 (s), 1119 (s), 796 (s), 721 (s); $^1H$ NMR (600 MHz, $D_6$-DMSO) $\delta_H$ 8.98–8.56 (3H, m, NH), 3.61 (2H, t, *J* 7.1, H-2'), 3.21–3.10 (4H, m, H-3), 2.97 (2H, t, *J* 7.1, H-1'), 2.90–2.77 (4H, m, H-2); $^{13}C$ NMR (151 MHz, $D_6$-DMSO) $\delta_C$ 57.8 (C-1'), 48.7 (C-2), 42.2 (C-3), 28.8 (C-2'); $^{19}F$ NMR (565 MHz, $D_6$-DMSO) $\delta_F$ −74.40 (s, TFA $CF_3$); HRMS *m/z* ($ESI^+$) [Found: 193.0336, $C_6H_{14}N_2{}^{79}Br$ requires $[M-2TFA+H]^+$ 193.0335], [Found: 195.0315, $C_6H_{14}N_2{}^{81}Br$ requires $[M-2TFA+H]^+$ 195.0314]; LRMS *m/z* ($ESI^+$) 193 ($[M(^{79}Br)-2TFA+H]^+$, 99%), 195 ($[M(^{81}Br)-2TFA+H]^+$, 100%).

**(13) 1-(2'-Bromoethyl)pyridinium bromide**




1,2-Dibromoethane (2.15 mL, 25.0 mmol, 1.0 eq.) was added dropwise to a solution of anhydrous pyridine (2.00 mL, 25.0 mmol, 1.0 eq.) in acetone (15 mL) at rt. The reaction solution was stirred and heated under reflux at 60 °C for 16 h. The resulting suspension was cooled to rt and the solvent was evaporated *in vacuo*. The precipitated crystals were washed with cold acetone (10 mL) and dried in a desiccator under high vacuum to yield the title compound as a colourless solid (4.89 g, 74%): $R_f$ 0.07 (MeOH:$CH_2Cl_2$, 1:9); m.p. 114–117 °C (from acetone) [lit.[187] 95–102 °C (from ethanol/acetone)]; $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 2981 (C-H, w), 1633 (m), 1486 (m), 773 (Ar C-H, s), 692 (s), 676 (s); $^1$H NMR (600 MHz, $CD_3OD$) $\delta_H$ 9.13–9.08 (2H, m, H-2), 8.74–8.67 (1H, m, H-4), 8.25–8.16 (2H, m, H-3), 5.12 (2H, t, *J* 5.8, H-1'), 4.07 (2H, t, *J* 5.8, H-2'); $^{13}$C NMR (151 MHz, $CD_3OD$) $\delta_C$ 147.9 (C-4), 146.4 (C-2), 129.5 (C-3), 63.3 (C-1'), 31.1 (C-2'); LRMS *m/z* ($ESI^+$) 186 ([M($^{79}$Br)−Br]$^+$, 34%), 188 ([M($^{81}$Br)−Br]$^+$, 53%). The spectroscopic data are in good agreement with the literature values.[6]

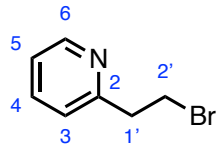
**(14) 1-(2'-Bromoethyl)-1*H*-pyrazole**




1,2-Dibromoethane (50.7 mL, 588 mmol, 20.0 eq.) was added dropwise to a solution of pyrazole (2.00 g, 29.4 mmol, 1.0 eq.) and tetrabutylammonium bromide (948 mg, 2.94 mmol, 0.1 eq.) in aqueous NaOH (40% *w/v*, 15 mL) at rt. The reaction mixture was

stirred at rt for 24 h. $H_2O$ (30 mL) was added, and the organic phase was separated and washed with $H_2O$ (50 mL). The organic phase was dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 5–40% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a pale yellow oil (3.17 g, 62%): $R_f$ 0.36 (EtOAc:petroleum ether 40–60 °C, 1:4); $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 7.55 (1H, d, *J* 1.9, H-3), 7.46 (1H, d, *J* 2.2, H-5), 6.26 (1H, dd, *J* 2.2, 1.9, H-4), 4.50 (2H, t, *J* 6.4, H-1'), 3.72 (2H, t, *J* 6.4, H-2'); LRMS *m/z* (ESI$^+$) 175 ([M($^{79}$Br)+H]$^+$, 98%), 177 ([M($^{81}$Br)+H]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[128,188]

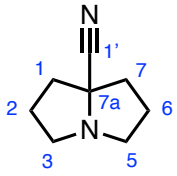
**(15) 2-(2'-Bromoethyl)pyridine**




Tetrabromomethane (3.55 g, 10.7 mmol, 1.2 eq.) and triphenylphosphine (2.81 g, 10.7 mmol, 1.2 eq.) were added to a solution of 2-(2-hydroxyethyl)pyridine (1.00 mL, 8.88 mmol, 1.0 eq.) in THF (20 mL) at 0 °C. The reaction mixture was warmed to rt and stirred for 20 h; before being diluted with $Et_2O$ (50 mL). The resulting suspension was filtered, and the filtrate was concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 7–60% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a yellow oil (740 mg, 45%): $R_f$ 0.35 (EtOAc:petroleum ether 40–60 °C, 3:7); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 3010 (C-H, w), 2970 (C-H, w), 1592 (m), 1474 (m), 1437 (m), 752 (Ar C-H, s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 8.59–8.53 (1H, m, H-5), 7.66–7.59 (1H, m, H-6), 7.22–7.13 (2H, m, H-3, 4), 3.77 (2H, t, *J* 7.1, H-2'), 3.33 (2H, t, *J* 7.1, H-1'); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 158.6

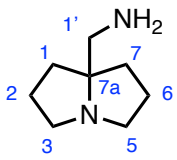
(C-2), 149.8 (C-5), 136.6 (C-6), 123.7 (C-3), 122.0 (C-4), 41.3 (C-1'), 31.8 (C-2'); HRMS *m/z* (ESI$^+$) [Found: 185.9913, $C_7H_9N^{79}Br$ requires $[M+H]^+$ 185.9913], [Found: 187.9891, $C_7H_9N^{81}Br$ requires $[M+H]^+$ 187.9892]; LRMS *m/z* (ESI$^+$) 186 ([M($^{79}$Br)+H]$^+$, 100%), 188 (M($^{81}$Br)+H]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[129]

### (16) Tetrahydro-1*H*-pyrrolizine-7a(5*H*)-carbonitrile



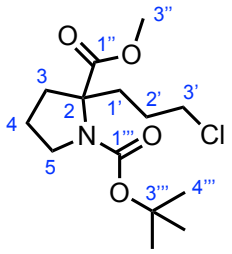

Acetone cyanohydrin (7.50 mL, 82.0 mmol, 5.0 eq.) was added dropwise to a solution of 1,7-dichloroheptan-4-one (3.00 g, 16.4 mmol, 1.0 eq.) in $NH_3$/MeOH (7 M, 50 mL) at 0 °C. The reaction solution was warmed to rt and stirred for 24 h. The mixture was concentrated *in vacuo*, and the residue was diluted with aqueous NaOH (0.1 M, 50 mL). The aqueous phase was extracted with $CH_2Cl_2$ (3 × 100 mL), and the organic components were combined, dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 10–50% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a yellow oil (1.72 g, 77%): $R_f$ 0.15 (EtOAc:petroleum ether 40–60 °C, 3:7); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2968 (C-H, w), 2871 (C-H, w), 2231 (C≡N, w); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 3.24–3.15 (2H, m, H-3$^a$, 5$^a$), 2.60–2.51 (2H, m, H-3$^b$, 5$^b$), 2.37–2.28 (2H, m, H-1$^a$, 7$^a$), 2.06–1.74 (6H, m, H-1$^b$, 2, 6, 7$^b$); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 124.6 (C-1'), 67.2 (C-7a), 55.5 (C-3, 5), 38.6 (C-1, 7), 26.5 (C-2, 6); HRMS *m/z* (ESI$^+$) [Found: 137.1071, $C_8H_{13}N_2$ requires $[M+H]^+$ 137.1073]. The spectroscopic data are in good agreement with the literature values.[130]

**(17) (Tetrahydro-1*H*-pyrrolizin-7a(5*H*)-yl)methanamine**




A solution of compound **16** (856 mg, 6.29 mmol, 1.0 eq.) in $Et_2O$ (10 mL) was added dropwise to a stirring suspension of $LiAlH_4$ (1.43 g, 37.7 mmol, 6.0 eq.) in $Et_2O$ (20 mL) at 0 °C. The reaction mixture was warmed to rt and stirred for 1 h. The reaction mixture was cooled to 0 °C and quenched sequentially with $H_2O$ (2 mL), aqueous NaOH (15% *w/w*, 2 mL), and $H_2O$ (6 mL). The mixture was warmed to rt and stirred for 15 min. Anhydrous $Na_2SO_4$ (10 g) was added, and the mixture was stirred at rt for 15 min. The mixture was filtered, and the filter cake was washed with $Et_2O$ (30 mL). The filtrate was concentrated *in vacuo* to yield the title compound as a pale yellow oil which was used in the next step without further purification (704 mg, 80%): $R_f$ 0.23 (MeOH:$CH_2Cl_2$, 1:9); $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 2.98–2.53 (4H, m, H-3, 5), 2.49 (2H, s, H-1'), 1.84–1.42 (8H, m, H-1, 2, 6, 7); HRMS *m/z* (ESI$^+$) [Found: 141.1384, $C_8H_{17}N_2$ requires $[M+H]^+$ 141.1386]; LRMS *m/z* (ESI$^+$) 141 ($[M+H]^+$, 15%). The spectroscopic data are in good agreement with the literature values.[189]

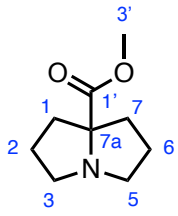
**(18) 1-(*tert*-Butyl) 2-methyl 2-(3'-chloropropyl)pyrrolidine-1,2-dicarboxylate**




A solution of 1-(*tert*-butyl) 2-methyl pyrrolidine-1,2-dicarboxylate (10.0 g, 43.6 mmol, 1.0 eq.) in anhydrous THF (30 mL) was added dropwise to a solution of LDA (1 M in THF,

65.4 mL, 65.4 mmol, 1.5 eq.) at −78 °C under an argon atmosphere. The reaction solution was stirred at −78 °C for 2 h. 1-Chloro-3-iodopropane (8.89 mL, 82.8 mmol, 1.9 eq.) was added dropwise; and the reaction mixture was warmed to rt, and stirred for 2 h. The reaction mixture was cooled to 0 °C and sat. aqueous $NH_4Cl$ (100 mL) was added dropwise. The aqueous phase was extracted with EtOAc (3 × 100 mL); and the organic components were combined, washed with brine (100 mL), dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 7–60% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a yellow oil (9.02 g, 68%): $R_f$ 0.32 (EtOAc:petroleum ether 40–60 °C, 1:4); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2976 (C-H, w), 1742 (C=O, s), 1698 (C=O, s), 1391 (s), 1164 (C-O, s); $^1$H NMR (600 MHz, $D_6$-DMSO) $\delta_H$ 3.75–3.44 (7H, m, H-3', 3'', 5$^a$, 5$^b$), 2.28–2.16 (1H, m, H-3$^a$), 2.12–1.70 (6H, m, H-1', 2', 3$^b$, 4$^a$), 1.61–1.46 (1H, m, H-4$^b$), 1.32 (9H, s, H-4'''); LRMS *m/z* (ESI$^+$) 328 ([M($^{35}$Cl)+Na]$^+$, 58%), 330 ([M($^{37}$Cl)+Na]$^+$, 33%), 633 ([2M($^{35}$Cl)+Na]$^+$, 100%), 635 ([2M($^{37}$Cl)+Na]$^+$, 92%). The spectroscopic data are in good agreement with the literature values.[131]

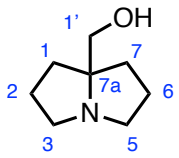
### (19) Methyl tetrahydro-1*H*-pyrrolizine-7a(5*H*)-carboxylate




TMSCl (18.8 mL, 148 mmol, 5.0 eq.) was added dropwise to a solution of compound **18** (9.02 g, 29.6 mmol, 1.0 eq.) in MeOH (100 mL) at 0 °C. The reaction solution was warmed to rt and stirred for 16 h. The reaction solution was basified to pH 8.0 with sat. aqueous $NaHCO_3$ and the reaction solution was concentrated *in vacuo*. The aqueous phase was

extracted with $CH_2Cl_2$ (3 × 100 mL); and the organic components were combined, dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 5–15% MeOH/$CH_2Cl_2$) to yield the title compound as a brown oil (3.95 g, 79%): $R_f$ 0.23 (MeOH:$CH_2Cl_2$, 1:11.5); $\tilde{\nu}_{max}$ (neat)/$cm^{-1}$ 2953 (C-H, w), 1744 (C=O, s), 1222 (m), 732 (s); $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 3.86 (3H, s, H-3'), 3.70–3.40 (4H, m, H-3, 5), 2.55–1.60 (8H, m, H-1, 2, 6, 7); LRMS *m/z* ($ESI^+$) 170 ([M+$H^+$], 100%). The spectroscopic data are in good agreement with the literature values.[132]

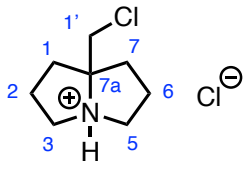
**(20) (Tetrahydro-1*H*-pyrrolizin-7a(5*H*)-yl)methanol**




$LiAlH_4$ (4.38 g, 116 mmol, 5.0 eq.) was added portionwise to a solution of compound **19** (3.90 g, 23.1 mmol, 1.0 eq.) in anhydrous THF (25 mL) at −10 °C. The reaction mixture was warmed to rt and stirred for 1 h. The reaction mixture was cooled to 0 °C and quenched sequentially with $H_2O$ (4.38 mL), aqueous NaOH (15% *w/w*, 4.38 mL), and $H_2O$ (13.1 mL). The mixture was warmed to rt and stirred for 15 min. Anhydrous $Na_2SO_4$ (5 g) was added, and the mixture was stirred at rt for 15 min. The mixture was filtered, and the filter cake was washed with THF (30 mL). The filtrate was concentrated *in vacuo* to yield the title compound as a yellow oil which was deemed pure enough for use in subsequent steps without further purification (2.86 g, 88%): $R_f$ 0.52 (MeOH:$CH_2Cl_2$, 1:9); $^1$H NMR (400 MHz, $CD_3OD$) $\delta_H$ 3.43 (2H, s, H-1'), 3.20–3.09 (2H, m, H-3$^a$, 5$^a$), 2.88–2.78 (2H, m, H-3$^b$, 5$^b$), 2.10–1.55 (8H, m, H-1, 2, 6, 7); HRMS *m/z* ($ESI^+$) [Found 142.1227, $C_8H_{16}NO$

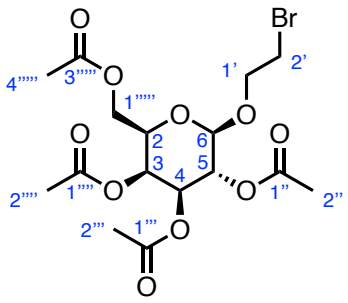
requires $[M+H]^+$ 142.1226]; LRMS *m/z* ($ESI^+$) 142 ($[M+H]^+$, 100%). The spectroscopic data are in good agreement with the literature values.[131]

### (21) 7a-(Chloromethyl)hexahydro-1*H*-pyrrolizine hydrochloride




Thionyl chloride (15.5 mL, 212 mmol, 3.0 eq.) was added dropwise to a solution of compound **20** (10.0 g, 70.8 mmol, 1.0 eq.) in toluene (200 mL) at 0 °C. The reaction mixture was warmed to 80 °C and stirred for 3 h. The reaction mixture was concentrated *in vacuo*. The residue was redissolved in $CH_2Cl_2$ and triturated with $Et_2O$ to yield the title compound as a beige solid (12.7 g, 91%): $R_f$ 0.30 ($MeOH:CH_2Cl_2$, 1:9); m.p. 199–201 °C (from $CH_2Cl_2/Et_2O$); $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 3393 (N-H, m, br), 2953 (C-H, w), 742 (C-Cl, s); $^1H$ NMR (600 MHz, $CDCl_3$) $\delta_H$ 12.47 (1H, br s, N*H*), 3.94 (2H, s, H-1'), 3.84–3.72 (2H, m, H-3$^a$, 5$^a$), 3.06–2.95 (2H, m, H-3$^b$, 5$^b$), 2.34–2.24 (2H, m, H-1$^a$, 7$^a$), 2.18–2.03 (4H, m, H-2, 6), 2.02–1.90 (2H, m, H-1$^b$, 7$^b$); $^{13}C$ NMR (151 MHz, $CDCl_3$) $\delta_C$ 80.8 (C-7a), 55.8 (C-3, 5), 47.6 (C-1'), 35.6 (C-1, 7), 24.1 (C-2, 6); HRMS *m/z* ($ESI^+$) [Found: 160.0889, $C_8H_{15}N^{35}Cl$ requires $[M+H]^+$ 160.0888]; LRMS *m/z* ($ESI^+$) 160 ($[M(^{35}Cl)+H]^+$, 88%).

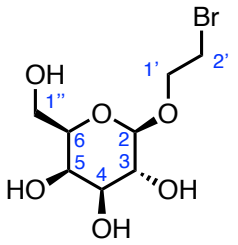
**(22) (2*R*,3*S*,4*S*,5*R*,6*R*)-2-(Acetoxymethyl)-6-(2'-bromoethoxy)tetrahydro-2*H*-pyran-3,4,5-triyl triacetate**




Boron trifluoride diethyl etherate (3.63 mL, 7.69 mmol, 1.2 eq.), followed by 2-bromoethanol (0.55 mL, 7.69 mmol, 1.2 eq.), were added dropwise to a solution of β-D-galactose pentaacetate (2.50 g, 6.40 mmol, 1.0 eq.) in anhydrous $CH_2Cl_2$ (20 mL) at 0 °C. The reaction solution was warmed to rt and stirred for 3 h. $K_2CO_3$ (1.33 g, 9.60 mmol, 1.5 eq.) was added and the reaction mixture was stirred at rt for 1 h. The reaction mixture was concentrated *in vacuo* and the residue was resuspended in brine (25 mL). The aqueous phase was extracted with $CHCl_3$ (3 × 50 mL), and the organic components were combined, dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 20–50% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a pale yellow oil which crystalised on standing to give a colourless amorphous solid (1.68 g, 58%): $R_f$ 0.21 (EtOAc:petroleum ether 40–60 °C, 3:7); $[\alpha]_D^{25}$ = +19.9 (*c* 1.0, $CHCl_3$); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 5.39 (1H, dd, *J* 5.7, 5.0, H-3), 5.22 (1H, dd, *J* 10.5, 5.7, H-4), 5.02 (1H, dd, *J* 10.5, 7.9, H-5), 4.53 (1H, d, *J* 7.9, H-6), 4.20–4.05 (3H, m, H-1'$^a$, 1'''''$^a$, 1'''''$^b$), 3.91 (1H, td, *J* 6.7, 5.0, H-2), 3.81 (1H, ddd, *J* 11.3, 7.2, 6.6, H-1'$^b$), 3.52–3.42 (2H, m, H-2'$^a$, 2'$^b$), 2.15 (3H, s, H-2'' / H-2''''), 2.08 (3H, s, H-2'''), 2.05 (3H, s, H-4'''''), 1.98 (3H, s, H-2'' / H-2''''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 170.5 (C-3'''''), 170.34 (C-1'' / C-1''''), 170.26 (C-1'' / C-1''''), 169.7 (C-1'''), 101.7 (C-6), 71.0 (C-5), 70.9 (C-2), 69.9 (C-1'), 68.7 (C-4), 67.1 (C-3), 61.4 (C-1'''''), 30.1 (C-2'), 21.0 ($CH_3$), 20.80 ($CH_3$), 20.78 ($CH_3$), 20.7 ($CH_3$); LRMS *m/z*

(ESI$^+$) 477 ([M($^{79}$Br)+Na]$^+$, 95%), 479 ([M($^{81}$Br)+Na]$^+$, 96%). The spectroscopic data are in good agreement with the literature values.[190,191]

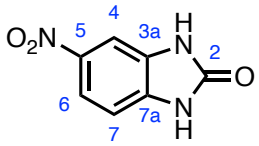
**(23) (2*R*,3*R*,4*S*,5*R*,6*R*)-2-(2'-Bromoethoxy)-6-(hydroxymethyl)tetrahydro-2*H*-pyran-3,4,5-triol**




A solution of NaOMe in MeOH (25% *w/w*, 39 µL, 0.18 mmol, 0.05 eq.) was added dropwise to a stirring solution of compound **22** (1.67 g, 3.67 mmol, 1.0 eq.) in MeOH (20 mL) at rt. The reaction solution was stirred at rt for 4 h, before being filtered through Dowex® 50WX8 ion exchange resin (H$^+$ form, 50–100 mesh, 2 × 5 cm plug). The resin was washed with MeOH (3 × 10 mL), and the combined filtrates were concentrated *in vacuo*. The product was purified using flash silica column chromatography (isocratic elution with 30% EtOH/$CH_2Cl_2$) to yield the title compound as a pale yellow oil which crystalised on standing to form an amorphous off-white solid (766 mg, 57%): $R_f$ 0.46 (EtOH:$CH_2Cl_2$, 3:7); $[\alpha]_D^{25} = -2.2$ (*c* 1.0, MeOH); $\tilde{\nu}_{max}$ (thin film)/cm$^{-1}$ 3366 (O-H, br, s), 2892 (C-H, w), 1071 (C-O, s); $^1$H NMR (600 MHz, $CD_3OD$) $\delta_H$ 4.32 (1H, d, *J* 7.6, H-2), 4.18–4.10 (1H, m, H-1'$^a$), 3.99–3.91 (1H, m, H-1'$^b$), 3.88–3.84 (1H, m, H-5), 3.82–3.71 (2H, m, H-1''$^a$, 1''$^b$), 3.68–3.52 (4H, m, H-2'$^a$, 2'$^b$, 3, 6), 3.50 (1H, dd, *J* 9.7, 3.4, H-4); $^{13}$C NMR (151 MHz, $CD_3OD$) $\delta_C$ 105.2 (C-2), 76.8 (C-6), 74.9 (C-4), 72.4 (C-3), 70.9 (C-1'), 70.3 (C-5), 62.5 (C-1''), 31.0 (C-2'); HRMS *m/z* (ESI$^+$) [Found: 308.9943, $C_8H_{15}O_6{}^{79}BrNa$ requires [M+Na]$^+$ 308.9944], [Found: 310.9922, $C_8H_{15}O_6{}^{81}BrNa$ requires [M+Na]$^+$

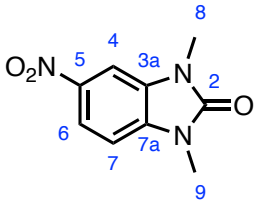
310.9924]; LRMS *m/z* (ESI$^+$) 309 ([M($^{79}$Br)+Na]$^+$, 78%), 311 ([M($^{81}$Br)+Na]$^+$, 78%). The spectroscopic data are in good agreement with the literature values.[133]

**(24) 5-Nitro-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one**




1,1'-Carbonyldiimidazole (10.6 g, 65.2 mmol, 2.0 eq.) was added portionwise to a solution of 4-nitro-*o*-phenyldiamine (5.00 g, 32.6 mmol, 1.0 eq.) in anhydrous DMF (50 mL) at rt. The reaction mixture was stirred at rt for 20 h, before being concentrated *in vacuo* to ~20 mL. The reaction mixture was diluted with $H_2O$ (200 mL) and filtered. The precipitate was washed with $H_2O$ (3 × 100 mL) and dried overnight under high vacuum to yield the title compound as a yellow solid (5.82 g, >99%): $R_f$ 0.59 (1% $NEt_3$; MeOH:$CH_2Cl_2$, 1:9); m.p. >300 °C (from $H_2O$/MeCN) [lit.[192] 308–309 °C (solvent unspecified)]; $^1$H NMR (400 MHz, $D_6$-DMSO) $\delta_H$ 11.41 (1H, br s, NH-1 / NH-3), 11.18 (1H, br s, NH-1 / NH-3), 7.94 (1H, dd, *J* 8.7, 2.3, H-6), 7.71 (1H, d, *J* 2.3, H-4), 7.10 (1H, d, *J* 8.7, H-7); LRMS *m/z* (ESI$^-$) 178 ([M−H]$^-$, 100%). The spectroscopic data are in good agreement with the literature values.[192]

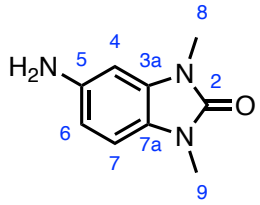
**(25) 1,3-Dimethyl-5-nitro-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one**




$K_2CO_3$ (6.17 g, 44.7 mmol, 4.0 eq.) and MeI (2.78 mL, 44.7 mmol, 4.0 eq.) were added to a solution of compound **24** (2.00 g, 11.2 mmol, 1.0 eq.) in anhydrous DMF (20 mL) at 0 °C.

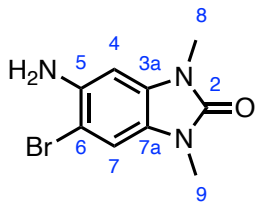
The reaction mixture was warmed to rt and stirred for 16 h. The reaction mixture was diluted with $H_2O$ (100 mL) and filtered. The precipitate was washed with $H_2O$ (3 × 50 mL) and dried overnight under high vacuum to yield the title compound as a yellow solid (2.25 g, 97%): $R_f$ 0.76 (1% $NEt_3$; MeOH:$CH_2Cl_2$, 1:19); m.p. 200–202 °C (from $H_2O$) [lit.[193] 202–204 °C, lit.[194] 220 °C]; $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 1708 (C=O, s), 1513 (N-O, s); $^1H$ NMR (400 MHz, $D_6$-DMSO) $\delta_H$ 8.10–8.05 (2H, m, H-4, 6), 7.40–7.32 (1H, m, H-7), 3.43 (3H, s, H-8), 3.41 (3H, s, H-9); LRMS *m/z* ($ESI^+$) 208 ([M+$H^+$], 100%). The spectroscopic data are in good agreement with the literature values.[193,194]

**(26) 5-Amino-1,3-dimethyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one**



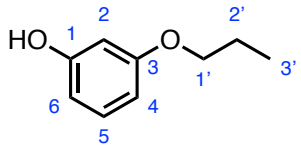

Conc. aqueous HCl (37% *w/w*, 200 µL) was added to a suspension of compound **25** (3.20 g, 15.5 mmol, 1.0 eq.) in EtOH (200 mL) under an argon atmosphere at rt. 10% Pd/C (1.65 g, 1.55 mmol, 0.1 eq.) was added and the flask was purged with hydrogen gas *via* 3 successive evacuate/fill cycles. The reaction mixture was stirred vigorously under a hydrogen atmosphere at rt for 16 h. The reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo* to yield the title compound as a pink solid (2.67 g, 81%): $R_f$ 0.35 (1% $NEt_3$; MeOH:$CH_2Cl_2$, 1:19); m.p. 140–142 °C (from $CH_2Cl_2$) [lit.[194] 160 °C, lit.[193] 135–136 °C (from benzene)]; $^1H$ NMR (400 MHz, $D_6$-DMSO) $\delta_H$ 6.88 (1H, d, *J* 8.2, H-7), 6.51 (1H, d, *J* 2.1, H-4), 6.46 (1H, dd, *J* 8.2, 2.1, H-6), 6.28 (2H, br s, $N\underline{H}_2$), 3.24 (3H, s, H-8), 3.23 (3H, s, H-9); LRMS *m/z* ($ESI^+$) 178 ([M+H]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[194]

**(27) 5-Amino-6-bromo-1,3-dimethyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one**




Bromine (290 µL, 5.65 mmol, 1.0 eq.) was added dropwise to a solution of compound **26** (1.00 g, 5.65 mmol, 1.0 eq.) in anhydrous $CHCl_3$ (20 mL) and AcOH (20 mL) at −20 °C. The reaction mixture was stirred at −20 °C for 20 min and quenched with aqueous $Na_2S_2O_3$ (10% *w/v*, 25 mL). The mixture was stirred for 30 min and diluted with $H_2O$ (25 mL). The aqueous phase was extracted with $CHCl_3$ (3 × 50 mL) and the organic components were combined, washed with sat. aqueous $NaHCO_3$ (50 mL), brine (50 mL), dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 20–50% EtOAc/$CH_2Cl_2$) to yield the title compound as an off-white solid (1.06 g, 73%): $R_f$ 0.21 (EtOAc:$CH_2Cl_2$, 1:4), m.p. 183–185 °C (dec., from EtOAc) [lit.[195] 209–210 °C (solvent unspecified)]; $^1$H NMR (400 MHz, $D_6$-DMSO) $\delta_H$ 7.18 (1H, s, H-7), 6.58 (1H, s, H-4), 4.96 (2H, br s, N$\underline{H}_2$), 3.22 (3H, s, H-8), 3.21 (3H, s, H-9); LRMS *m/z* (ESI$^+$) 256 ([M($^{79}$Br)+H]$^+$, 100%), 258 ([M($^{81}$Br)+H]$^+$, 99%). The spectroscopic data are in good agreement with the literature values.[4,140]

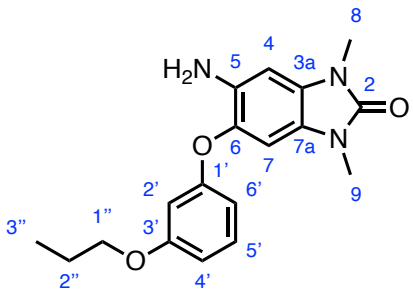
**(28) 3-Propoxyphenol**




$K_2CO_3$ (13.8 g, 100 mmol, 3.0 eq.) was added to a solution of resorcinol (11.0 g, 100 mmol, 3.0 eq.) in DMF (30 mL) at rt. 1-Bromopropane (3.00 mL, 33.3 mmol, 1.0 eq.) was added

dropwise over 5 min at rt. The reaction mixture was stirred vigorously at 50 °C for 16 h. The reaction mixture was quenched with aqueous HCl (1 M, 150 mL) and the aqueous phase was extracted with EtOAc (4 × 100 mL). The organic components were combined, washed with aqueous LiCl (0.5 M, 100 mL), brine (100 mL), dried ($MgSO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 7–60% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a pale-yellow oil (3.50 g, 69%): $R_f$ 0.61 (EtOAc:petroleum ether 40–60 °C, 3:7); $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 7.16–7.08 (1H, dd, *J* 8.4, 8.4, H-5), 6.49 (1H, ddd, *J* 8.4, 2.3, 1.0, H-4), 6.45–6.38 (2H, m, H-2, 6), 5.23 (1H, s, OH), 3.89 (2H, t, *J* 6.6, H-1'), 1.88–1.72 (2H, m, H-2'), 1.02 (3H, t, *J* 7.4, H-3'); LRMS *m/z* (ESI$^+$) 305 ([2M+H]$^+$, 38%). The spectroscopic data are in good agreement with the literature values.[139,140]

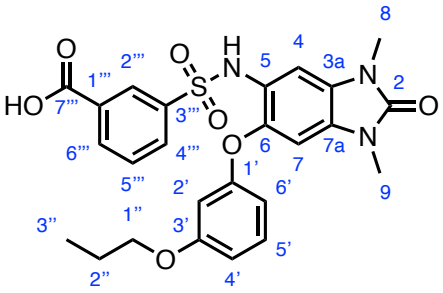
**(29) 5-Amino-1,3-dimethyl-6-(3'-propoxyphenoxy)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one**




Sodium hydroxide (920 mg, 23.0 mmol, 1.0 eq.) was added to a solution of compound **28** (3.50 g, 23.0 mmol, 1.0 eq.) in MeOH (10 mL). The reaction mixture was stirred at rt for 1 h. The mixture was concentrated *in vacuo*, and the residue was dissolved in toluene and concentrated *in vacuo* to yield sodium 3-propoxyphenolate as a white solid (4.00 g, 100%). The product was dried by lyophilisation and used directly in the next step.

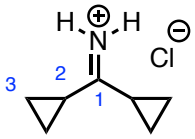
CuI and $Cs_2CO_3$ were dried in an Abderhalden's drying pistol under high vacuum at 160 °C for 24 h. A Schlenk tube was charged with compound **27** (300 mg, 1.17 mmol, 1.0 eq.), sodium 3-propoxyphenolate (306 mg, 1.76 mmol, 1.5 eq.), $Cs_2CO_3$ (1.14 g, 3.51 mmol, 3.0 eq.), CuI (22.3 mg, 0.12 mmol, 0.1 eq.), and 8-hydroxyquinoline (34.0 mg, 0.23 mmol, 0.2 eq.). The tube was sealed and purged with argon *via* 3 successive evacuate/fill cycles. Anhydrous diglyme (20 mL) was degassed by 3 successive freeze-pump-thaw cycles under argon before being added to the reaction mixture. The reaction mixture was stirred vigorously at 130 °C for 72 h. The reaction mixture was cooled to rt and filtered through a silica plug. The silica plug was washed with 10% MeOH/$CH_2Cl_2$ (100 mL) and the filtrate was concentrated *in vacuo* at 60 °C. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 0–6% MeOH/$CHCl_3$) to yield the title compound as a brown solid (235 mg, 62%): $R_f$ 0.26 (MeOH:$CHCl_3$, 3:97); m.p. 134–136 °C (from $CHCl_3$) [lit.[141] 137–139 °C (from $CH_2Cl_2$)]; $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 7.17 (1H, dd, *J* 8.6, 8.6, H-5'), 6.63 (1H, s, H-7), 6.61–6.56 (1H, m, H-6'), 6.54–6.43 (3H, m, H-4, 2', 4'), 3.88 (2H, t, *J* 6.6, H-1''), 3.37 (3H, s, H-8), 3.31 (3H, s, H-9), 1.84–1.70 (2H, m, H-2''), 1.01 (3H, t, *J* 7.5, H-3''); LRMS *m/z* (ESI$^+$) 328 ([M+H]$^+$, 65%), 655 ([2M+H]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[4,141]

**(30) 3'''-(*N*-(1,3-Dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)benzoic acid**



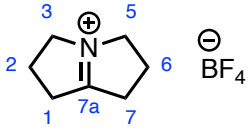
3-(Chlorosulfonyl)benzoic acid (324 mg, 1.47 mmol, 1.2 eq.) and anhydrous pyridine (198 µL, 2.44 mmol, 2.0 eq.) were added to a solution of compound **29** (400 mg, 1.22 mmol, 1.0 eq.) in anhydrous $CH_2Cl_2$ (20 mL) at rt. The rection solution was stirred at rt for 24 h and quenched by addition of aqueous HCl (1 M, 20 mL). The aqueous phase was extracted with 25% MeOH in $CH_2Cl_2$ (3 × 50 mL); and the organic components were combined, dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 0–10% MeOH/$CH_2Cl_2$) to yield the title compound as a brown solid (442 mg, 71%): $R_f$ 0.14 (MeOH:$CH_2Cl_2$, 3:47); m.p. 120–122 °C (dec., from $CH_2Cl_2$) [lit.[141] 116–118 °C (from $CH_2Cl_2$)]; $^1$H NMR (400 MHz, $CD_3OD$) $\delta_H$ 8.29 (1H, t, *J* 1.8, H-2'''), 8.13–8.05 (1H, m, H-6'''), 7.80–7.72 (1H, m, H-4'''), 7.42–7.32 (2H, m, H-4, 5'''), 7.02 (1H, t, *J* 8.2, H-5'), 6.59–6.50 (2H, m, H-4', 7), 6.11–6.04 (1H, m, H-6'), 6.01 (1H, t, *J* 2.4, H-2'), 3.80 (2H, t, *J* 6.5, H-1''), 3.43 (3H, s, H-8), 3.23 (3H, s, H-9), 1.82–1.69 (2H, m, H-2''), 1.02 (3H, t, *J* 7.4, H-3''); LRMS *m/z* (ESI$^+$) 512 ([M+H]$^+$, 56%), 1045 ([2M+Na]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[141]

### (31) Dicyclopropylmethaniminium chloride




Dicyclopropyl ketone (10.0 mL, 88.7 mmol, 1.0 eq.) was added dropwise to a stirring solution of titanium tetrachloride (10.7 mL, 97.6 mmol, 1.1 eq.) in anhydrous benzene (200 mL) and anhydrous $CH_2Cl_2$ (100 mL) at −10 °C. A steady stream of ammonia gas was bubbled through the reaction mixture at 0 °C for 1 h. The reaction mixture was filtered, and the filtrate was warmed to rt. HCl (4 M in dioxane, 25 mL) was added and the mixture was stirred at rt for 1 h. The mixture was concentrated *in vacuo* and the residue was resuspended in anhydrous $CHCl_3$ (100 mL). The suspension was filtered, and the filtrate was concentrated *in vacuo*. The residue was triturated with $Et_2O$ to yield the title compound as a colourless crystalline solid (6.23 g. 48%): $R_f$ 0.18 (MeOH:$CH_2Cl_2$, 1:9); m.p. 131–133 °C (from $CHCl_3$/$Et_2O$) [lit.[196] 135–137 °C (from $CHCl_3$/$Et_2O$)]; $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 1.98–1.90 (2H, m, H-2), 1.67–1.56 (4H, m, H-3$^a$), 1.45–1.34 (4H, m, H-3$^b$); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 200.0 (C-1), 16.0 (C-2), 13.9 (C-3); LRMS *m/z* (ESI$^+$) 110 ([M−HCl+H]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[196]

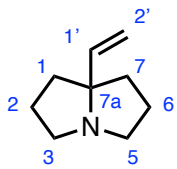
### (32) 1,2,3,5,6,7-Hexahydropyrrolizin-4-ium tetrafluoroborate




Ammonium tetrafluoroborate (7.19 g, 68.6 mmol, 2.0 eq.) was added to a solution of compound **31** (5.00 g, 34.3 mmol, 1.0 eq.) in DMSO (50 mL). The reaction solution was

stirred and heated at 100 °C for 6 h. The solvent was distilled out *in vacuo* and $^{t}$BuOH (100 mL) was added to the residue. The precipitated crystals were filtered and washed with $^{t}$BuOH (2 × 50 mL). The product was recrystalised from EtOH to yield the title compound as a brown solid (2.63 g, 39%): $R_f$ 0.05 (MeOH:$CHCl_3$, 1:9); m.p. 180–183 °C (from EtOH); $\tilde{\nu}_{max}$ (thin film)/cm$^{-1}$ 2980 (C-H, w), 1697 (C=N, m) 1050 (B-F, s); $^1$H NMR (500 MHz, $D_6$-DMSO) $\delta_H$ 3.92–3.80 (4H, m, H-3, 5), 3.00–2.90 (4H, m, H-1, 7), 2.53–2.43 (4H, m, H-2, 6); $^{13}$C NMR (126 MHz, $D_6$-DMSO) $\delta_C$ 197.9 (C-7a), 51.9 (C-3, 5), 30.9 (C-1, 7), 24.1 (C-2, 6); $^{19}$F NMR (470 MHz, $D_6$-DMSO) $\delta_F$ −148.38 (s, $BF_4$); HRMS *m/z* (ESI$^+$) [Found: 110.0969, $C_7H_{12}N^+$ requires $[M-BF_4]^+$ 110.0964]; LRMS *m/z* (ESI$^+$) 110 ($[M-BF_4]^+$, 100%), 219 ($[M-BF_4-H+Na]^+$, 95%). The spectroscopic data are in good agreement with the literature values.[142]

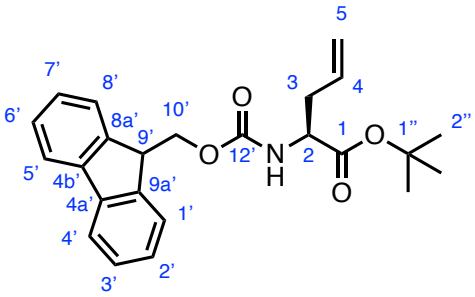
### (33) 7a-Vinylhexahydro-1*H*-pyrrolizine




Vinylmagnesium bromide solution (1 M in THF, 12.2 mL, 12.2 mmol, 1.2 eq.) was added dropwise to a suspension of compound **32** (2.00 g, 10.2 mmol, 1.0 eq.) in anhydrous THF at 0 °C. The reaction mixture was heated under reflux at 60 °C for 20 h. The reaction mixture was cooled to 0 °C and quenched with aqueous NaOH (20% *w/v*, 25 mL). The solvent was removed *in vacuo* and the aqueous phase was extracted with EtOAc (3 × 50 mL). The organic components were combined, washed with brine (50 mL), dried ($Na_2SO_4$), filtered, and concentrated *in vacuo* to yield the title compound as a yellow oil (874 mg, 63%), which was taken forward without further purification: $R_f$ 0.65 (1% TFA; $H_2O$:MeCN, 1:1); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2959 (C-H, w), 2866 (C-H, w), 1639 (C=C, w); $^1$H NMR

(600 MHz, $CDCl_3$) $\delta_H$ 5.95 (1H, dd, *J* 17.2, 10.5, H-1'), 5.15 (1H, dd, *J* 17.2, 1.5, H-2'$^a$), 4.95 (1H, dd, *J* 10.5, 1.5, H-2'$^b$), 3.21–3.10 (2H, m, H-3$^a$, 5$^a$), 2.66–2.58 (2H, m, H-3$^b$, 5$^b$), 1.94–1.60 (8H, m, H-1, 2, 6, 7); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 145.3 (C-1'), 110.8 (C-2'), 75.1 (C-7a), 55.7 (C-3, 5), 37.4 (C-1, 7), 25.4 (C-2, 6); HRMS *m/z* ($ESI^+$) [Found: 138.1278, $C_9H_{16}N$ requires $[M+H]^+$ 138.1277]; LRMS *m/z* ($ESI^+$) 138 ($[M+H]^+$, 100%).

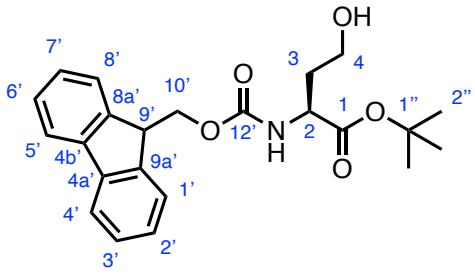
**(34) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)pent-4-enoate**




A solution of *tert*-butyl 2,2,2-trichloroacetimidate (10.6 mL, 59.3 mmol, 2.0 eq.) in cyclohexane (100 mL) was added dropwise over 30 min to a solution of Fmoc-L-allyl-glycine (10.0 g, 29.6 mmol, 1.0 eq.) in $CH_2Cl_2$ (100 mL) at rt. Boron trifluoride diethyl etherate (365 µL, 2.96 mmol, 0.1 eq.) was added and the reaction mixture was stirred at rt for 6 h. The reaction mixture was cooled to 0 °C and $NaHCO_3$ (8.20 g, 97.7 mmol, 3.3 eq.) was added. The reaction mixture was warmed to rt and stirred for 10 min. The reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 5–40% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a colourless oil (8.08 g, 69%): $R_f$ 0.48 (EtOAc:petroleum ether 40–60 °C, 1:4); $[\alpha]_D^{25}$ = +10.3 (*c* 1.0, $CHCl_3$); $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 7.80–7.74 (2H, m, H-4', 5'), 7.63–7.57 (2H, m, H-1', 8'), 7.44–7.37 (2H, m, H-3', 6'), 7.35–7.28 (2H, m, H-2', 7'), 5.80–5.64 (1H, m, H-4), 5.36 (1H, d, *J* 8.2, NH), 5.19–5.08 (2H, m, H-5), 4.51–4.31 (3H, m, H-2, 10'),

4.23 (1H, t, *J* 7.1, H-9'), 2.66–2.44 (2H, m, H-3), 1.48 (9H, s, H-2''); LRMS *m/z* ($ESI^+$) 416 ($[M+Na]^+$, 86%), 809 ($[2M+Na]^+$, 100%). The spectroscopic data are in good agreement with the literature values.[197]

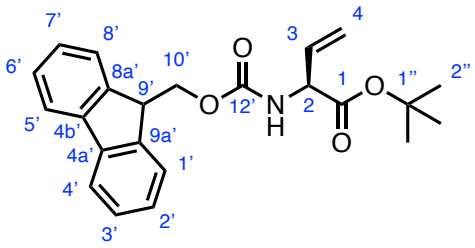
**(35) *tert*-Butyl (((9'*H*-fluoren-9'-yl)methoxy)carbonyl)-*L*-homoserinate**




Borane dimethyl sulfide complex (2 M in THF, 18.2 mL, 36.5 mmol, 5.0 eq.) was added to a solution of Fmoc-*L*-aspartic acid 4-*tert*-butyl ester (3.00 g, 7.29 mmol, 1.0 eq.) in anhydrous THF (30 mL) over 20 min at 0 °C. The reaction solution was warmed to rt and stirred for 24 h. The reaction solution was cooled to 0 °C and quenched with sat. aqueous $NH_4Cl$ (30 mL). The aqueous phase was extracted with EtOAc (3 × 50 mL) and the organic components were combined, washed with brine (50 mL), dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 12–100% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a colourless oil (1.64 g, 57%): $R_f$ 0.39 (EtOAc:petroleum ether 40–60 °C, 1:1); $[\alpha]_D^{25} = -2.5$ (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/$cm^{-1}$ 3368 (O-H, br, w), 2936 (C-H, w), 1723 (C=O, s), 1157 (C-O, s), 740 (s); $^1H$ NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.79–7.74 (2H, m, H-4', 5'), 7.62–7.57 (2H, m, H-1', 8'), 7.43–7.37 (2H, m, H-3', 6'), 7.35–7.29 (2H, m, H-2', 7'), 5.62 (1H, d, *J* 7.8, NH), 4.52–4.32 (3H, m, H-2, 10'), 4.22 (1H, t, *J* 6.9, H-9'), 3.75–3.54 (2H, m, H-4), 3.06–2.96 (1H, m, OH), 2.24–2.11 (1H, m, H-3$^a$), 1.67–1.59 (1H, m, H-3$^b$), 1.48 (9H, s, H-2''); $^{13}C$ NMR (151 MHz, $CDCl_3$)

$\delta_C$ 171.9 (C-1), 157.1 (C-12'), 143.9, 143.7 (C-8a' , 9a'), 141.47, 141.46 (C-4a', 4b'), 127.9 (C-3', 6'), 127.2 (C-2', 7'), 125.22, 125.15 (C-1', 8'), 120.17, 120.13 (C-4', 5'), 82.7 (C-1''), 67.3 (C-10'), 58.4 (C-4), 51.6 (C-2), 47.3 (C-9'), 36.3 (C-3), 28.1 (C-2''); HRMS *m/z* ($ESI^+$) [Found: 420.1776, $C_{23}H_{27}NO_5Na$ requires $[M+Na]^+$ 420.1781]; LRMS *m/z* ($ESI^+$) 420 ($[M+Na]^+$, 68%), 817 ([2M+Na], 100%). The spectroscopic data are in good agreement with the literature values.[198]

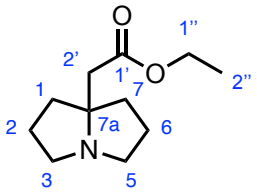
**(36) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)but-3-enoate**




Tributylphosphine (1.53 mL, 6.11 mmol, 1.5 eq.) was added dropwise to a solution of compound **35** (1.62 g, 4.08 mmol, 1.0 eq.) and 2-nitrophenyl selenocyanate (1.39 g, 6.11 mmol, 1.5 eq.) in anhydrous THF (15 mL) at rt. The reaction mixture was stirred at rt for 3 h; at which point TLC analysis indicated complete consumption of the starting material. The reaction mixture was cooled to 0 °C and aqueous $H_2O_2$ (30% *v/v*, 4.17 mL, 40.8 mmol, 10.0 eq.) was added. The reaction mixture was warmed to rt and stirred for 2 h. The reaction mixture was diluted with $H_2O$ (50 mL) and the aqueous phase was extracted with $Et_2O$ (3 × 50 mL). The organic components were combined, dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 2–20% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a yellow oil (674 mg, 43%): $R_f$ 0.29 (EtOAc:petroleum ether 40–60 °C, 1:9); $[\alpha]_D^{25} = -0.7$ (*c* 1.0, $CHCl_3$); $^1H$ NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.79–7.74 (2H, m, H-4', 5'), 7.65–7.57 (2H, m, H-1', 8'), 7.43–7.38 (2H, m,

H-3', 6'), 7.35–7.29 (2H, m, H-2', 7'), 5.91 (1H, ddd, *J* 16.7, 10.3, 5.9, H-3), 5.49 (1H, d, *J* 7.2, NH), 5.34 (1H, d, *J* 16.7, H-4$^{a}$), 5.25 (1H, d, *J* 10.3, H-4$^{b}$), 4.82 (1H, dd, *J* 7.2, 5.9, H-2), 4.47–4.36 (2H, m, H-10'), 4.24 (1H, t, *J* 7.2, H-9'), 1.48 (9H, s, H-2''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 169.7 (C-1), 155.7 (C-12'), 144.1, 144.0 (C-8a', 9a'), 141.5 (C-4a', 4b'), 133.2 (C-3), 127.9 (C-3', 6'), 127.2 (C-2', 7'), 125.27, 125.25 (C-1', 8'), 120.14, 120.12 (C-4', 5'), 117.2 (C-4), 82.9 (C-1''), 67.2 (C-10'), 56.7 (C-2), 47.4 (C-9), 28.1 (C-2''); LRMS *m/z* (ESI$^{+}$) 402 ([M+Na]$^{+}$, 47%), 781 ([2M+Na], 100%). The spectroscopic data are in good agreement with the literature values.[199]

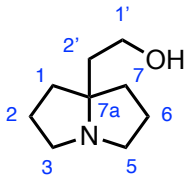
### (37) Ethyl 2'-(tetrahydro-1*H*-pyrrolizin-7a(5*H*)-yl)acetate




Ethyl potassium malonate (7.86 g, 46.2 mmol, 3.0 eq.) was added to a solution of 1,7-dichloroheptan-4-one (2.82 g, 15.4 mmol, 1.0 eq.) in $NH_3$/MeOH (7 M, 50 mL) at 0 °C. The reaction mixture was warmed to rt and stirred for 24 h. The reaction mixture was concentrated *in vacuo* and redissolved in EtOH (50 mL). The reaction mixture was stirred and heated at 70 °C for 3 h. The reaction mixture was cooled to 0 °C and filtered. The filtrate was concentrated *in vacuo* and diluted with aqueous $K_2CO_3$ (1 M, 25 mL). The aqueous phase was extracted with $CH_2Cl_2$ (3 × 50 mL); and the organic components were combined, dried ($Na_2SO_4$), filtered, and concentrated *in vacuo* to yield the crude title compound as a yellow oil (2.49 g, 82%) which was deemed pure enough for use in the next step without further purification: $R_f$ 0.25 (acetone:pentane, 3:2); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2957 (C-H, w), 1734 (C=O, s); $^{1}$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 4.11 (2H, q, *J* 7.2, H-1''), 3.08–2.98 (2H, m, H-3$^{a}$, 5$^{a}$), 2.61–2.51 (2H, m, H-3$^{b}$, 5$^{b}$), 2.43 (2H, s, H-2'), 2.05–1.55 (8H, m,

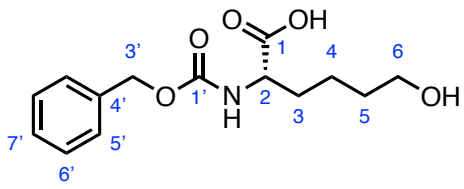
H-1, 2, 6, 7), 1.25 (3H, t, *J* 7.2, H-2’’); LRMS *m/z* (ESI$^+$) 198 ([M+H]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[200]

**(38) 2’-(Tetrahydro-1*H*-pyrrolizin-7a(5*H*)-yl)ethan-1’-ol**



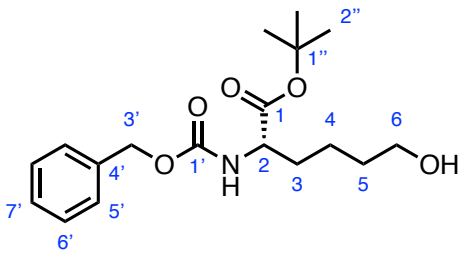

DIBAL-H (1.0 M in toluene, 26.3 mL, 26.3 mmol, 2.1 eq.) was added dropwise to a solution of compound **37** (2.47 g, 12.5 mmol, 1.0 eq.) in anhydrous $CH_2Cl_2$ (50 mL) at 0 °C. The reaction mixture was warmed to rt and stirred for 6 h. The reaction mixture was cooled to 0 °C and quenched sequentially with $H_2O$ (1.05 mL), aqueous NaOH (15% *w/w*, 1.05 mL), and $H_2O$ (2.63 mL). The mixture was warmed to rt and stirred for 15 min. Anhydrous $Na_2SO_4$ was added, and the mixture was stirred for a further 15 min at rt. The mixture was filtered, and the filtrate was concentrated *in vacuo*. The product was purified by solid phase extraction (ISOLUTE® SCX-2, elution with $NH_3$/MeOH (7 M, 1 CV) and $CH_2Cl_2$ (1 CV)) to yield the title compound as a yellow oil (1.75 g, 90%): $R_f$ 0.00 (MeOH:$CH_2Cl_2$, 1:4); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 3358 (O-H, br, w), 2955 (C-H, w), 1456 (w), 1070 (w); $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 3.83–3.77 (2H, m, H-1’), 3.10–2.95 (2H, m, H-3$^a$, 5$^a$), 2.65–2.50 (2H, m, H-3$^b$, 5$^b$), 1.90–1.55 (10H, m, H-1, 2, 2’, 6, 7); LRMS *m/z* (ESI$^+$) 156 ([M+H]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[145]

**(39) (*S*)-2-(((Benzyloxy)carbonyl)amino)-6-hydroxyhexanoic acid**



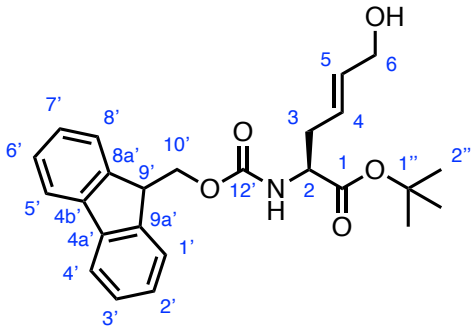

A solution of $N_\alpha$-(carbobenzyloxy)-*L*-lysine (7.00 g, 25.0 mmol, 1.0 eq.) in $H_2O$ (100 mL) was basified to pH 9–10 using aqueous NaOH (4 M) and heated to 60 °C. Sodium nitroprusside dihydrate (13.4 g, 44.9 mmol, 1.8 eq.) was added portionwise over 30 min, whilst maintaining pH 9–10 by addition of aqueous NaOH (4 M). The reaction mixture was stirred at 60 °C and pH 9–10 for 6 h. The reaction mixture was cooled to rt and filtered through Celite®. The filtrate was carefully acidified to pH 1 and the aqueous phase was extracted with EtOAc (5 × 100 mL). The organic components were combined, washed with brine (250 mL), dried ($MgSO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 2–20% MeOH/$CH_2Cl_2$) to yield the title compound as a colourless oil (1.15 g, 38%): $R_f$ 0.21 (MeOH:$CH_2Cl_2$, 1:9), $[\alpha]_D^{25}$= +13.8 (*c* 1.0, $CHCl_3$) [lit.[154] $[\alpha]_D^{20}$ = +15.0 (*c* 1.07, $CHCl_3$)]; $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 7.38–7.28 (5H, m, H-5', 6', 7'), 5.10 (2H, s, H-3'), 4.45–4.33 (1H, m, H-2), 3.64 (2H, t, *J* 5.9, H-6), 1.96–1.35 (6H, m, H-3, 4, 5); LRMS *m/z* ($ESI^-$) 280 ($[M-H]^-$, 100%). The spectroscopic data are in good agreement with the literature values.[154]

**(40) *tert*-Butyl (*S*)-2-(((benzyloxy)carbonyl)amino)-6-hydroxyhexanoate**




Benzyltriethylammonium chloride (932 mg, 4.09 mmol, 1.0 eq.) and $K_2CO_3$ (14.7 g, 106 mmol, 26.0 eq.) were added to a solution of compound **39** (1.15 g, 4.09 mmol, 1.0 eq.) in anhydrous MeCN (50 mL) at rt. Bromo-2-methylpropane (22.0 mL, 196 mmol, 48.0 eq.) was added and the reaction mixture was stirred at 55 °C for 24 h. The reaction mixture was cooled to rt, and cold $H_2O$ (10 mL) was added. The reaction mixture was concentrated *in vacuo* and the residue was redissolved in $H_2O$ (50 mL). The aqueous phase was extracted with EtOAc (3 × 50 mL) and the organic components were combined, washed with brine (100 mL), dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 0–10% MeOH/$CH_2Cl_2$) to yield the title compound as a yellow oil (1.07 g, 78%): $R_f$ 0.32 (MeOH:$CH_2Cl_2$, 1:19); $[\alpha]_D^{25}$ = +4.8 (*c* 1.0, $CHCl_3$) [lit.[201] $[\alpha]_D^{24}$ = +6.5 (*c* 1.0, $CHCl_3$)]; $^1$H NMR (400 MHz, $CDCl_3$) $\delta_H$ 7.40–7.28 (5H, m, H-5’, 6’, 7’), 5.10 (2H, s, H-3’), 4.32–4.19 (1H, m, H-2), 3.69–3.58 (2H, m, H-6), 1.90–1.35 (15H, m, H-2’’, 3, 4, 5); LRMS *m/z* (ESI$^+$) 360 ([M+Na]$^+$, 39%), 697 ([2M+Na]$^+$, 100%). The spectroscopic data are in good agreement with the literature values.[202]

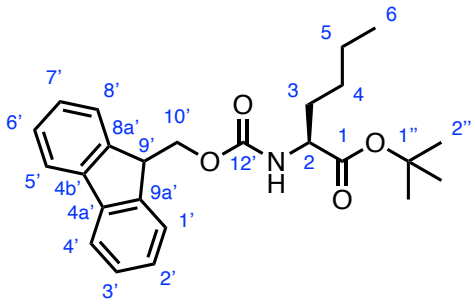
**(41) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-hydroxyhex-4-enoate**




*cis*-Butene-1,4-diol (2.81 mL, 34.3 mmol, 3.00 eq.) and Grubbs 2$^{nd}$ generation catalyst (484 mg, 0.57 mmol, 0.05 eq.) were added to a solution of compound **34** (4.50 g, 11.4 mmol, 1.00 eq.) in anhydrous, degassed $CH_2Cl_2$ (50 mL) under an argon atmosphere at rt. The reaction solution was stirred and heated under reflux at 40 °C for 6 h. The reaction solution was cooled to rt and filtered through Celite®; and the filtrate was concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 12–100% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a yellow oil (3.56 g, 70%): $R_f$ 0.36 (EtOAc:petroleum ether 40–60 °C, 1:1); $[\alpha]_D^{25}$ = +14.4 (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 3327 (O-H, br, s), 2978 (C-H, w), 1718 (C=O, s), 1248 (m), 1155 (C-O, s), 1047 (m), 741 (s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.77 (2H, d, *J* 7.6, H-4', 5'), 7.60 (2H, d, *J* 7.6, H-1', 8'), 7.40 (2H, t, *J* 7.6, H-3', 6'), 7.32 (2H, t, *J* 7.6, H-2', 7'), 5.74 (1H, dt, *J* 15.2, 6.0, H-5), 5.66–5.56 (1H, m, H-4), 5.38 (1H, d, *J* 7.9, NH), 4.44–4.29 (3H, m, H-2, 10'), 4.23 (1H, t, *J* 7.0, H-9'), 4.09 (2H, t, *J* 6.0, H-6), 2.66–2.42 (2H, m, H-3), 1.48 (9H, s, H-2''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 170.9 (C-1), 155.8 (C-12'), 144.1, 144.0 (C-8a', 9a'), 141.5 (C-4a', 4b'), 133.7 (C-5), 127.9 (C-3', 6'), 127.2 (C-2', 7'), 126.0 (C-4), 125.3, 125.2 (C-1' 8'), 120.14, 120.13 (C-4', 5'), 82.5 (C-1''), 67.1 (C-10'), 63.4 (C-6), 54.0 (C-2), 47.3 (C-9'), 35.7 (C-3), 28.2 (C-2''); HRMS *m/z* (ESI$^+$)

[Found: 446.1934, $C_{25}H_{29}NO_5Na$ requires $[M+Na]^+$ 446.1938]; LRMS *m/z* ($ESI^+$) 446 ($[M+Na]^+$, 85%), 869 ($[2M+Na]^+$, 100%); HPLC retention time: 10.4 min, 93.5% (220 nm), 93.8% (254 nm), 93.7% (280 nm).

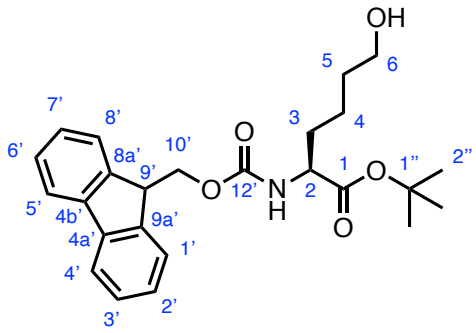
**(42) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)hexanoate**




10% Pd/C (916 mg, 0.86 mmol, 0.1 eq.) was added to a solution of compound **41** (3.65 g, 8.62 mmol, 1.0 eq.) in EtOH (100 mL) under an argon atmosphere at rt. The flask was purged with hydrogen gas *via* 3 successive evacuate/fill cycles and the reaction mixture was stirred vigorously under a hydrogen atmosphere at rt for 16 h. The reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 2–20% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a colourless oil (1.32 g, 37%): $R_f$ 0.32 (EtOAc:petroleum ether 40–60 °C, 1:9); $[\alpha]_D^{25}$ = + 4.3 (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/$cm^{-1}$ 2959 (C-H, w), 1724 (C=O, s), 1156 (C-O, s), 740 (s); $^1H$ NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.77 (2H, d, *J* 7.6, H-4', 5'), 7.64–7.58 (2H, m, H-1', 8'), 7.40 (2H, t, *J* 7.6, H-3', 6'), 7.32 (2H, t, *J* 7.6, H-2', 7'), 5.31 (1H, d, *J* 8.3, NH), 4.39 (2H, d, *J* 7.3, H-10'), 4.31–4.21 (2H, m, H-2, 9'), 1.90–1.62 (2H, m, H-3), 1.48 (9H, s, H-2''), 1.42–1.24 (4H, m, H-4, 5), 0.91 (3H, t, *J* 7.1, H-6); $^{13}C$ NMR (151 MHz, $CDCl_3$) $\delta_C$ 172.0 (C-1), 156.0 (C-12'), 144.1, 144.0 (C-8a', 9a'), 141.5 (C-4a', 4b'), 127.8 (C-3', 6'), 127.2 (C-2', 7'), 125.3 (C-1', 8'), 120.11, 120.09 (C-4', 5'), 82.1 (C-1''), 67.0 (C-10'), 54.5

(C-2), 47.4 (C-9'), 32.7 (C-3), 28.2 (C-2''), 27.3 (C-4), 22.5 (C-5), 14.0 (C-6); HRMS *m/z* (ESI$^+$) [Found: 432.2142, $C_{25}H_{31}NO_4Na$ requires $[M+Na]^+$ 432.2145]; LRMS *m/z* (ESI$^+$) 432 ($[M+Na]^+$, 18%), 841 ($[2M+Na]^+$, 100%); HPLC retention time: 13.5 min, 96.0% (220 nm), 97.9% (254 nm), 96.7% (280 nm).

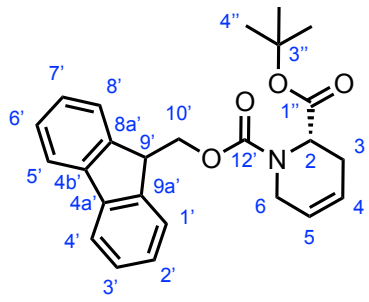
**(43) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-hydroxyhexanoate**




$PtO_2$ (232 mg, 1.02 mmol, 0.1 eq.) was added to a solution of compound **41** (4.32 g, 10.2 mmol, 1.0 eq.) in EtOAc (150 mL) under an argon atmosphere at rt. The flask was purged with hydrogen gas *via* 3 successive evacuate/fill cycles and the reaction mixture was stirred vigorously under a hydrogen atmosphere at rt for 20 h. The reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 12–100% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a yellow oil (3.24 g, 75%): $R_f$ 0.33 (EtOAc:petroleum ether 40–60 °C, 1:1); $[\alpha]_D^{25}$ = +3.0 (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 3340 (O-H, br, w), 2930 (C-H, w), 1727 (C=O, s), 1155 (C-O, s), 760 (s), 742 (s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.77 (2H, d, *J* 7.6, H-4', 5'), 7.60 (2H, d, *J* 7.6, H-1', 8'), 7.40 (2H, t, *J* 7.6, H-3', 6'), 7.32 (2H, t, *J* 7.6, H-2', 7'), 5.34 (1H, d, *J* 8.1, NH), 4.39 (2H, d, *J* 7.2, H-10'), 4.32–4.19 (2H, m, H-2, 9'), 3.70–3.58 (2H, m, H-6), 2.00–1.34 (15H, m, H-3, 4, 5, 2''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 171.8 (C-1),

156.1 (C-12'), 144.1, 144.0 (C-8a', 9a'), 141.5 (C-4a', 4b'), 127.8 (C-3', 6'), 127.2 (C-2', 7'), 125.3 (C-1', 8'), 120.13, 120.11 (C-4', 5'), 82.3 (C-1''), 67.1 (C-10'), 62.7 (C-6), 54.3 (C-2), 47.4 (C-9'), 32.8 (C-3), 32.3 (C-5), 28.2 (C-2''), 21.5 (C-4); HRMS *m/z* ($ESI^+$) [Found: 448.2092, $C_{25}H_{31}NO_5Na$ requires $[M+Na]^+$ 448.2094]; LRMS *m/z* ($ESI^+$) 448 ($[M+Na]^+$, 27%), 873 ($[2M+Na]^+$, 100%); HPLC retention time: 11.3 min, 96.1% (220 nm), 99.2% (254 nm), 99.1% (280 nm).

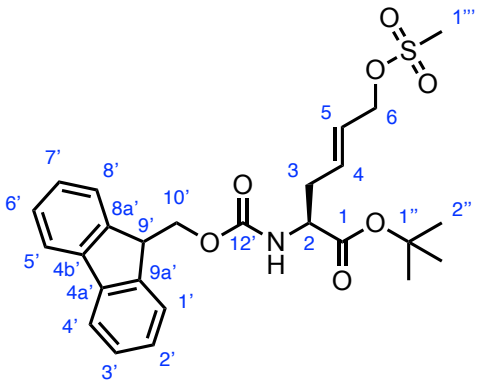
**(44) 1-((9'*H*-Fluoren-9'-yl)methyl) 2-(*tert*-butyl) (*S*)-3,6-dihydropyridine-1,2(2*H*)-dicarboxylate**




Tetrabromomethane (2.03 g, 6.13 mmol, 1.5 eq.) was added to a solution of compound **41** (1.73 g, 4.08 mmol, 1.0 eq.) and triphenylphosphine (1.61 g, 6.13 mmol, 1.5 eq.) in anhydrous $CH_2Cl_2$ (20 mL) at 0 °C. The reaction solution was warmed to rt and stirred for 2 h. The reaction solution was diluted with $Et_2O$ (50 mL), and the resulting mixture was filtered. The filtrate was concentrated *in vacuo* and the product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 5-40% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a colourless oil (1.05 g, 63%): $R_f$ 0.55 (EtOAc:petroleum ether 40–60 °C, 1:4); $[\alpha]_D^{25} = -16.2$ (*c* 0.5, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/$cm^{-1}$ 2978 (C-H, w), 1741 (C=O, s), 1712 (C=O, s), 1658 (C=C, m), 1155 (C-O, s), 758 (s), 740 (s); $^1H$ NMR (500 MHz, $D_6$-DMSO, T=393 K) $\delta_H$ 7.89–7.83 (2H, m, H-4', 5'), 7.67–7.59 (2H, m, H-1', 8'), 7.46–7.38 (2H, m, H-3', 6'), 7.37–7.29 (2H, m, H-2', 7'),

6.76–6.64 (1H, m, H-5), 4.94–4.84 (1H, m, H-4), 4.63–4.50 (2H, m, H-2, 10'$^{a}$), 4.45 (1H, dd, *J* 10.7, 6.4, H-10'$^{b}$), 4.30 (1H, t, *J* 6.4, H-9'), 2.26–2.16 (1H, m, H-6$^{a}$), 2.06–1.72 (3H, m, H-6$^{b}$, 3$^{a}$, 3$^{b}$), 1.40 (9H, s, H-4''); $^{13}$C NMR (126 MHz, $D_6$-DMSO, T=393 K) $\delta_C$ 168.6 (C-1), 152.1 (C-12'), 143.2, 143.1 (C-8a', 9a'), 140.27, 140.25 (C-4a', 4b'), 127.0 (C-3', 6'), 126.5 (C-2', 7'), 124.1 (C-1', 8'), 123.5 (C-5), 119.3 (C-4', 5'), 104.8 (C-4), 80.5 (C-3''), 66.8 (C-10'), 53.6 (C-2), 46.3 (C-9'), 27.1 (C-4''), 22.7 (C-6), 17.3 (C-3); HRMS *m/z* (ESI$^+$) [Found: 428.1836, $C_{25}H_{27}NO_4Na$ requires $[M+Na]^+$ 428.1832]; LRMS *m/z* (ESI$^+$) 428 ($[M+Na]^+$, 47%), 833 ($[2M+Na]^+$, 100%); HPLC retention time: 13.5 min, 96.1% (220 nm), 95.6% (254 nm), 95.8% (280 nm).

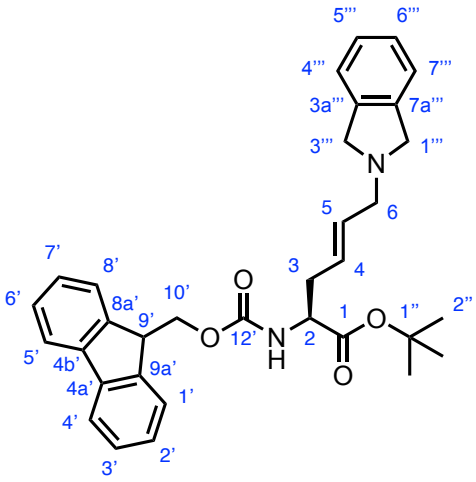
### (45) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-((methylsulfonyl)oxy)hex-4-enoate




Methanesuflonyl chloride (1.11 mL, 14.3 mmol, 1.5 eq.) was added dropwise to a solution of compound **41** (4.05 g, 9.56 mmol, 1.0 eq.) and $NEt_3$ (2.66 mL, 19.1 mmol, 2.0 eq.) in $CH_2Cl_2$ (30 mL) at 0 °C. The reaction mixture was warmed to rt and stirred for 1 h. The reaction mixture was diluted with $CHCl_3$ (50 mL) and washed with sat. aqueous $NaHCO_3$ (50 mL). The organic component was dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution using a Biotage® system with 12–60% EtOAc/petroleum ether 40–60 °C) to yield the title

compound as a colourless oil (3.96 g, 83%): $R_f$ 0.48 (EtOAc:petroleum ether 40–60 °C, 1:1); $[\alpha]_D^{25}$ = +20.0 (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/$cm^{-1}$ 2980 (C-H, w), 1726 (C=O, s), 1354 (S=O, s), 1173 (C-O, s), 1157 (C-O, s), 927 (s), 742 (s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.77 (2H, d, *J* 7.5, H-4', 5'), 7.63–7.56 (2H, m, H-1', 8'), 7.43–7.38 (2H, m, H-3', 6'), 7.35–7.29 (2H, m, H-2', 7'), 5.86–5.64 (2H, m, H-4, 5), 5.39 (2H, d, *J* 7.8, NH), 4.71–4.60 (2H, m, H-6), 4.47–4.30 (3H, m, H-2, 10'), 4.22 (1H, t, *J* 6.8, H-9'), 2.96 (3H, s, H-1'''), 2.70–2.46 (2H, m, H-3), 1.48 (9H, s, H-2''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 170.5 (C-1), 155.8 (C-12'), 144.0, 143.9 (C-8a', 9a'), 141.5 (C-4a', 4b'), 132.4 (C-5), 127.9 (C-3', 6'), 127.2 (C-2', 7'), 126.7 (C-4), 125.24, 125.20 (C-1', 8'), 120.16, 120.15 (C-4', 5'), 82.9 (C-1''), 69.8 (C-6), 67.1 (C-10'), 53.7 (C-2), 47.3 (C-9'), 38.3 (C-1'''), 35.6 (C-3), 28.2 (C-2''); HRMS *m/z* ($ESI^+$) [Found: 524.1713, $C_{26}H_{31}NO_7SNa$ requires $[M+Na]^+$ 524.1714]; LRMS *m/z* ($ESI^+$) 524 ($[M+Na]^+$, 69%), 1025 ($[2M+Na]^+$, 100%).

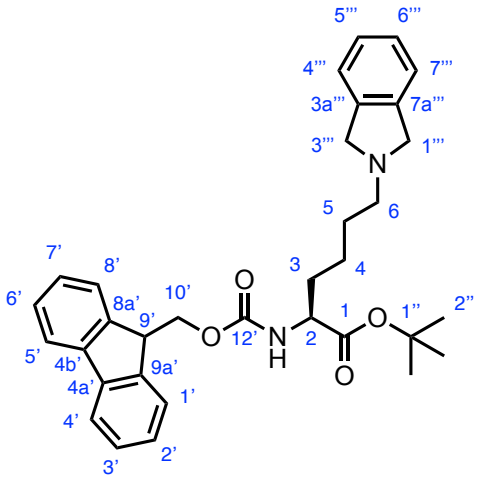
**(46) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-(isoindolin-2'''-yl)hex-4-enoate**




A solution of compound **45** (2.65 g, 5.28 mmol, 1.0 eq.) in anhydrous DMF (20 mL) was added dropwise to mixture of isoindoline hydrochloride (904 mg, 5.81 mmol, 1.1 eq.), $K_2CO_3$ (2.18 g, 15.8 mmol, 3.0 eq.), and KI (964 mg, 5.81 mmol, 1.1 eq.) in anhydrous

DMF (30 mL) at rt. The reaction mixture was stirred at rt for 1 h and quenched with $H_2O$ (10 mL). The reaction mixture was concentrated *in vacuo* and the residue was diluted with aqueous LiCl (0.5 M, 50 mL). The aqueous phase was extracted with EtOAc (3 × 50 mL) and the organic components were combined, washed with aqueous LiCl (0.5 M, 50 mL), brine (50 mL), dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 1–3% 2 M $NH_3$ in MeOH/$CH_2Cl_2$) to yield the title compound as a yellow oil (1.97 g, 71%): $R_f$ 0.63 (2 M $NH_3$ in MeOH:$CH_2Cl_2$, 1:9); $[\alpha]_D^{25}$ = +14.6 (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/$cm^{-1}$ 2932 (C-H, w), 1722 (C=O, s), 1156 (C-O, s), 743 (s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.76 (2H, d, *J* 7.6, H-4', 5'), 7.61 (2H, d, *J* 7.6, H-1', 8'), 7.39 (2H, t, *J*, 7.6, H-3', 6'), 7.31 (2H, t, *J* 7.6, H-2', 7'), 7.21–7.13 (4H, m, H-4''', 5''', 6''', 7'''), 5.74 (1H, dt, *J* 15.2, 6.4, H-5), 5.63 (1H, dt, *J* 15.2, 7.2, H-4), 5.40 (1H, d, *J* 8.0, NH), 4.43–4.32 (3H, m, H-2, 10'), 4.24 (1H, t, *J* 7.2, H-9'), 3.93 (4H, s, H-1''', 3'''), 3.35 (2H, d, *J* 6.4, H-6), 2.72–2.44 (2H, m, H-3), 1.49 (9H, s, H-2''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 171.0 (C-1), 155.8 (C-12'), 144.1, 144.0 (C-8a', 9a'), 141. 4 (C-4a', 4b'), 140.0 (C-3a''', 7a'''), 132.0 (C-5), 127.8 (C-3', 6'), 127.22 (C-2', 7'), 127.19 (C-4), 126.9 (C-5''', 6'''), 125.3 (C-1', 8'), 122.4 (C-4''', 7'''), 120.1 (C-4', 5'), 82.5 (C-1''), 67.2 (C-10'), 58.9 (C-1''', 3'''), 57.8 (C-6), 54.1 (C-2), 47.3 (C-9'), 35.8 (C-3), 28.2 (C-2''); HRMS *m/z* ($ESI^+$) [Found: 547.2572, $C_{33}H_{36}N_2O_4Na$ requires $[M+Na]^+$ 547.2567]; LRMS *m/z* ($ESI^+$) 525 ($[M+H]^+$, 100%); HPLC retention time: 9.7 min, 93.8% (220 nm), 96.3% (254 nm), 93.0% (280 nm).

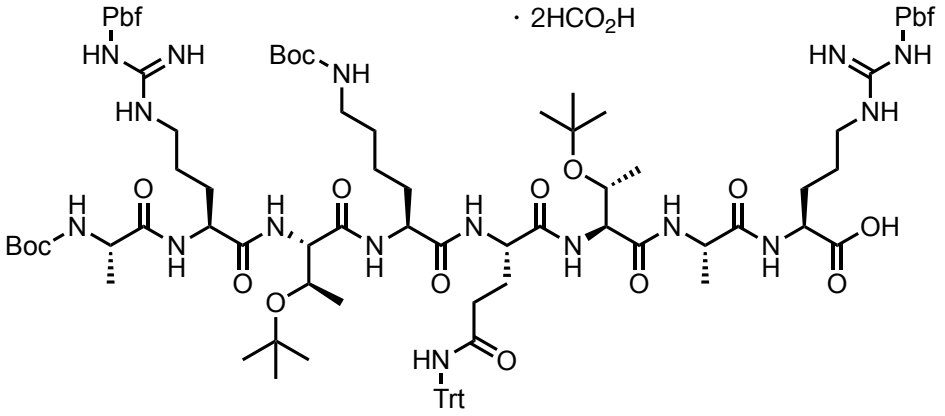
**(47) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-(isoindolin-2'''-yl)hexanoate**




$PtO_2$ (11.4 mg, 0.05 mmol, 0.1 eq.) was added to a solution of compound **46** (253 mg, 0.48 mmol, 1.0 eq.) in EtOAc (25 mL) under an argon atmosphere at rt. The flask was purged with hydrogen gas *via* 3 successive evacuate/fill cycles and the reaction mixture was stirred vigorously under a hydrogen atmosphere at rt for 6 h. The reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 1–3% 2 M $NH_3$ in MeOH/$CH_2Cl_2$) to yield the title compound as a yellow oil (209 mg, 83%): $R_f$ 0.60 (2 M $NH_3$ in MeOH:$CH_2Cl_2$, 1:9); $[\alpha]_D^{25}$ = +8.4 (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/$cm^{-1}$ 2936 (C-H, w), 1722 (C=O, s), 1156 (C-O, s), 762 (s), 744 (s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 7.76 (2H, d, *J* 7.6, H-4', 5'), 7.63–7.57 (2H, m, H-1', 8'), 7.42–7.37 (2H, m, H-3', 6'), 7.34–7.28 (2H, m, H-2', 7'), 7.22–7.13 (4H, m, H-4''', 5''', 6''', 7'''), 5.39 (1H, d, *J* 8.2, NH), 4.38 (2H, d, *J* 7.2, H-10'), 4.32–4.25 (1H, m, H-2), 4.22 (1H, t, *J* 7.2, H-9'), 3.95 (4H, s, H-1''', 3'''), 2.74 (2H, t, *J* 7.5, H-6), 1.94–1.84 (1H, m, H-3$^a$), 1.78–1.32 (14H, m, H-2'', 3$^b$, 4, 5); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 171.8 (C-1), 156.0 (C-12'), 144.1, 144.0 (C-8a', 9a'), 141.5 (C-4a', 4b'), 139.9 (C-3a''', 7a'''), 127.8 (C-3', 6'), 127.2 (C-2', 7'), 126.9 (C-5''',

6’’’), 125.3 (C-1’, 8’), 122.4 (C-4’’’, 7’’’), 120.12, 120.10 (C-4’, 5’), 82.2 (C-1’’), 67.1 (C-10’), 59.2 (C-1’’’, 3’’’), 55.9 (C-6), 54.4 (C-2), 47.4 (C-9’), 32.8 (C-3), 28.5 (C-5), 28.2 (C-2’’), 23.0 (C-4); HRMS *m/z* ($ESI^+$) [Found: 527.2904, $C_{33}H_{39}N_2O_4$ requires $[M+H]^+$ 527.2904]; LRMS *m/z* ($ESI^+$) 527 ($[M+H]^+$, 100%); HPLC retention time: 9.6 min, 92.1% (220 nm), 95.0% (254 nm), 94.1% (280 nm).

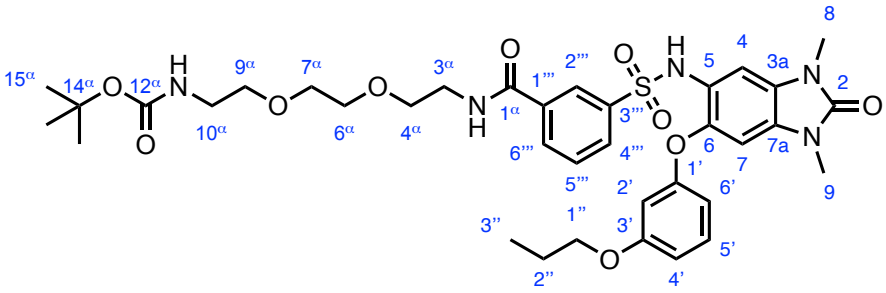
**(48)** ***N*²-(*N*-(*N*²-(*N*⁶-(*tert*-Butoxycarbonyl)-*N*²-(*N*-(*N*²-((*tert*-butoxycarbonyl)-*L*-alanyl)-*N*$^{\omega}$-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)-*L*-arginyl)-*O*-(*tert*-butyl)-*L*-threonyl)-*L*-lysyl)-*N*⁵-trityl-*L*-glutaminyl)-*O*-(*tert*-butyl)-*L*-threonyl-*L*-alanyl)-*N*$^{\omega}$-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)-*L*-arginine diformic acid**




The title compound was synthesised using SPPS on H-Arg(Pbf)-2-ClTrt resin (1.00 g, 0.30–0.90 mmol/g loading, Novabiochem®) using a Liberty Blue™ 2.0 Automated Microwave Peptide Synthesiser (CEM). Amino acids (5 eq.) were coupled in the following order: Fmoc-Ala-OH, Fmoc-Thr(*t*-Bu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Thr(*t*-Bu)-OH, Fmoc-Arg(Pbf)-OH, Boc-Ala-OH. Coupling reactions were carried out using an activator solution of DIC (7.8 mL/100 mL) in ultrapure DMF and an additive solution of Oxyma Pure (14.2 g/100 mL) and DIPEA (1.5% *v/v*) in ultrapure DMF. Fmoc deprotection steps were carried out using a solution of piperidine in ultrapure DMF

(20% *v/v*). After the final coupling step the resin was washed with $CH_2Cl_2$ (20 mL), MeOH (20 mL), and $CH_2Cl_2$ (20 mL). The peptide was cleaved from the resin with a solution of HFIP in $CH_2Cl_2$ (30% *v/v*) at rt for 1 h. The resin was filtered and washed with $CH_2Cl_2$; and the filtrate was concentrated *in vacuo*. The product was purified by semi-preparative reverse phase HPLC (gradient method 1, 40 mL $min^{-1}$, Agilent Zorbax PrepHT column [7 µm, 21.2 mm × 250 mm]) and lyophilised to yield the title compound as a colourless solid (482 mg): $R_f$ 0.22 (MeOH:$CH_2Cl_2$, 1:9); HRMS *m/z* ($ESI^+$) [Found: 1990.0480, $C_{100}H_{149}N_{16}O_{22}S_2$ requires $[M-2FA+H]^+$ 1990.0468]; LRMS *m/z* ($ESI^+$) 996 ($[M-2FA+2H]^{2+}$, 100%), 1990 ($[M-2FA+H]^+$, 40%); HPLC retention time: 13.6 min, 97.9% (220 nm), 96.7% (254 nm), 95.4% (280 nm).

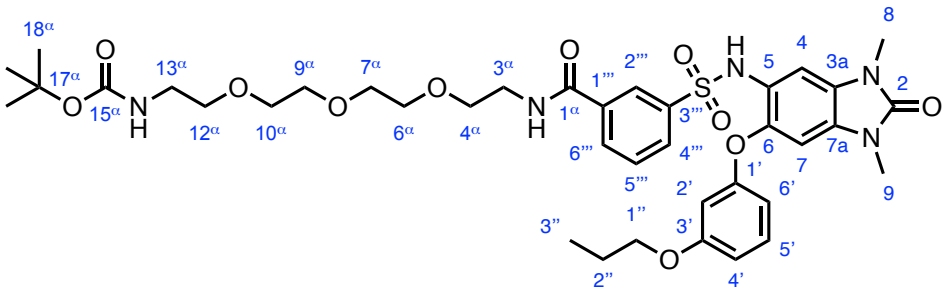
**(49) *tert*-Butyl (1$^{\alpha}$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-1$^{\alpha}$-oxo-5$^{\alpha}$,8$^{\alpha}$-dioxa-2$^{\alpha}$-azadecan-10$^{\alpha}$-yl)carbamate**



Carboxylic acid **30** (100 mg, 0.20 mmol, 1.0 eq.) and *tert*-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate (48.5 mg, 0.20 mmol, 1.0 eq.) were reacted according to general amide coupling procedure 1. The product was purified using flash silica column chromatography (gradient elution with 1–5% MeOH/$CH_2Cl_2$) to yield the title compound as a brown oil (98.3 mg, 66%): $R_f$ 0.28 (MeOH:$CH_2Cl_2$, 1:19); $^1H$ NMR (400 MHz, $CD_3OD$) $\delta_H$ 8.17 (1H, t, *J* 1.8, H-2'''), 7.96–7.90 (1H, m, H-6'''), 7.80–7.74

(1H, m, H-4’’’), 7.43–7.34 (2H, m, H-4, 5’’’), 7.01 (1H, t, *J* 8.5, H-5’), 6.60–6.52 (2H, m, H-4’, 7), 6.11–6.01 (2H, m, H-2’, 6’), 3.81 (2H, t, *J* 6.5, H-1’’), 3.70–3.46 (10H, m, H-3$^{\alpha}$, 4$^{\alpha}$, 6$^{\alpha}$, 7$^{\alpha}$, 9$^{\alpha}$), 3.44 (3H, s, H-8), 3.24 (3H, s, H-9), 3.20 (2H, t, *J* 5.7, H-10$^{\alpha}$), 1.83–1.69 (2H, m, H-2’’), 1.41 (9H, s, H-15$^{\alpha}$), 1.02 (3H, t, *J* 7.4, H-3’’); LRMS *m/z* (ESI$^{+}$) 742 ([M+H]$^{+}$, 39%), 764 ([M+Na]$^{+}$, 100%). The spectroscopic data are in good agreement with previous work in the Conway group.

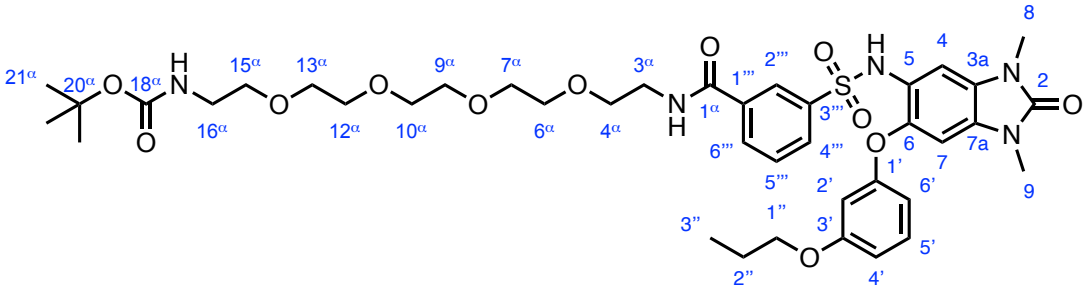
**(50) *tert*-Butyl (1$^{\alpha}$-(3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-1$^{\alpha}$-oxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$-trioxa-2$^{\alpha}$-azatridecan-13$^{\alpha}$-yl)carbamate**




Carboxylic acid **30** (100 mg, 0.20 mmol, 1.0 eq.) and *tert*-butyl (2-(2-(2-(2-aminoethoxy)ethoxy)ethoxy)ethyl)carbamate (58.5 mg, 0.20 mmol, 1.0 eq.) were reacted according to general amide coupling procedure 1. The product was purified using flash silica column chromatography (gradient elution with 1–5% MeOH/$CH_2Cl_2$) to yield the title compound as a brown oil (126 mg, 80%): $R_f$ 0.25 (MeOH:$CH_2Cl_2$, 1:19); $^1$H NMR (400 MHz, $CD_3OD$) $\delta_H$ 8.17 (1H, t, *J* 1.8, H-2’’’), 7.98–7.90 (1H, m, H-6’’’), 7.81–7.75 (1H, m, H-4’’’), 7.44–7.34 (2H, m, H-4, 5’’’), 7.02 (1H, t, *J* 8.5, H-5’), 6.60–6.52 (2H, m, H-4’, 7), 6.08–6.02 (2H, m, H-2’, 6’), 3.81 (2H, t, *J* 6.5, H-1’’), 3.66–3.50 (12H, m, H-3$^{\alpha}$, 4$^{\alpha}$, 6$^{\alpha}$, 7$^{\alpha}$, 9$^{\alpha}$, 10$^{\alpha}$), 3.48–3.42 (5H, m, H-8, 12$^{\alpha}$), 3.24 (3H, s, H-9), 3.18 (2H, t, *J* 5.6, H-13$^{\alpha}$), 1.82–1.70 (2H, m, H-2’’), 1.41 (9H, s, H-18$^{\alpha}$), 1.03 (3H, t, *J* 7.4, H-3’’); LRMS

*m/z* (ESI$^+$) 786 ([M+H]$^+$, 37%), 808 ([M+Na]$^+$, 100%). The spectroscopic data are in good agreement with previous work in the Conway group.

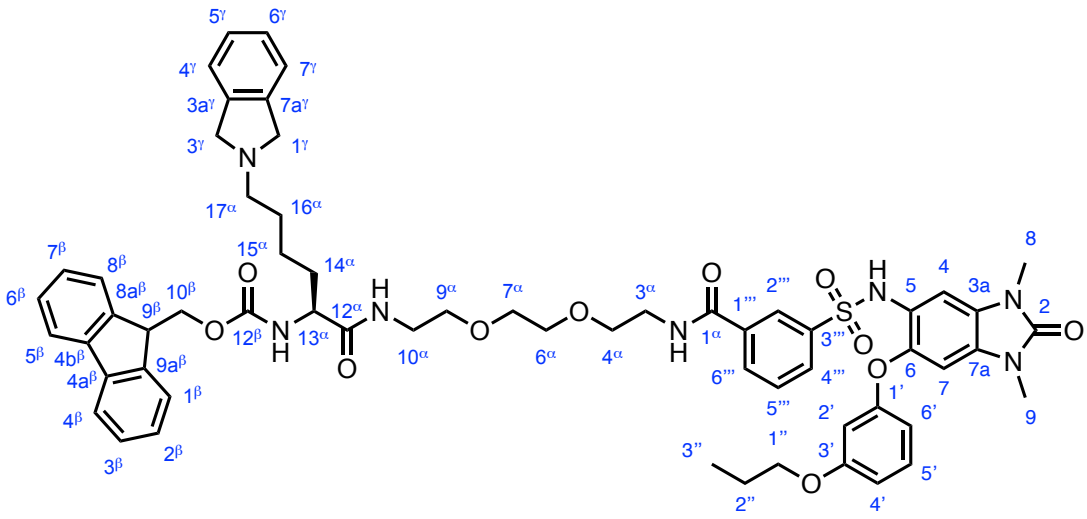
**(51) *tert*-Butyl (1$^\alpha$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-1$^\alpha$-oxo-5$^\alpha$,8$^\alpha$,11$^\alpha$,14$^\alpha$-tetraoxa-2$^\alpha$-azahexadecan-16$^\alpha$-yl)carbamate**



Carboxylic acid **30** (100 mg, 0.20 mmol, 1.0 eq.) and *tert*-butyl (14-amino-3,6,9,12-tetraoxatetradecyl)carbamate (67.3 mg, 0.20 mmol, 1.0 eq.) were reacted according to general amide coupling procedure 1. The product was purified using flash silica column chromatography (gradient elution with 2–6% MeOH/$CH_2Cl_2$) to yield the title compound as a brown oil (146 mg, 88%): $R_f$ 0.17 (MeOH:$CH_2Cl_2$, 1:19); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2875 (C-H, w), 1698 (C=O, s), 1508 (s), 1460 (m), 1256 (m), 1173 (s), 736 (m); $^1$H NMR (600 MHz, $CD_3OD$) $\delta_H$ 8.17 (1H, t, *J* 1.8, H-2'''), 7.96–7.92 (1H, m, H-6'''), 7.80–7.76 (1H, m, H-4'''), 7.40 (1H, t, *J* 7.9, H-5'''), 7.36 (1H, s, H-4), 7.02 (1H, t, *J* 8.6, H-5'), 6.59–6.53 (2H, m, H-4', 7), 6.08–6.03 (2H, m, H-2', 6'), 3.81 (2H, t, *J* 6.4, H-1''), 3.66–3.50 (16H, m, H-3$^\alpha$, 4$^\alpha$, 6$^\alpha$, 7$^\alpha$, 9$^\alpha$, 10$^\alpha$, 12$^\alpha$, 13$^\alpha$), 3.46 (2H, t, *J* 5.6, H-15$^\alpha$), 3.43 (3H, s, H-8), 3.24 (3H, s, H-9), 3.19 (2H, t, *J* 5.6, H-16$^\alpha$), 1.80–1.71 (2H, m, H-2''), 1.42 (9H, s, H-21$^\alpha$), 1.02 (3H, t, *J* 7.4, H-3''); $^{13}$C NMR (151 MHz, $CD_3OD$) $\delta_C$ 167.9 (C-1$^\alpha$), 161.7 (C-3'), 159.6 (C-1'), 158.4 (C-18$^\alpha$), 156.5 (C-2), 146.7 (C-6), 142.0 (C-3'''), 136.2 (C-1'''), 132.3 (C-6'''), 131.1 (C-4''', C-5'), 130.1 (C-5'''), 129.9 (C-3a), 127.28, 127.25 (C-2''', C-7a),

123.1 (C-5), 110.6 (C-6'), 110.3 (C-4'), 108.2 (C-4), 105.3 (C-2'), 100.8 (C-7), 80.1 (C-20$^{\alpha}$), 71.59, 71.54, 71.52, 71.51, 71.3, 71.2, 71.0, 70.6 (C-4$^{\alpha}$, 6$^{\alpha}$, 7$^{\alpha}$, 9$^{\alpha}$, 10$^{\alpha}$, 12$^{\alpha}$, 13$^{\alpha}$, 15$^{\alpha}$), 70.4 (C-1''), 41.3 (C-16$^{\alpha}$), 41.1 (C-3$^{\alpha}$), 28.8 (C-21$^{\alpha}$), 27.6, 27.5 (C-8, 9), 23.6 (C-2''), 10.9 (C-3''); HRMS *m/z* (ESI$^{+}$) [Found: 830.3673, $C_{40}H_{56}N_5O_{12}S$ requires $[M+H]^{+}$ 830.3641]; LRMS *m/z* (ESI$^{+}$) 830 ($[M+H]^{+}$, 79%), 852 ($[M+Na]^{+}$, 100%); HPLC retention time: 9.6 min, 99.3% (250 nm), 96.6% (280 nm), >99.9% (310 nm).

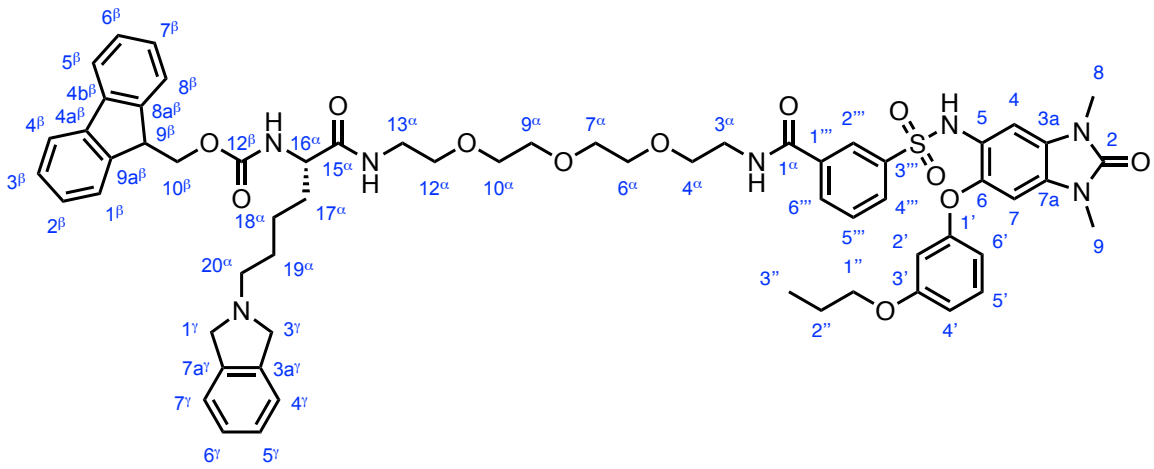
**(52) (9$^{\beta}$*H*-Fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-17$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,12$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$-dioxa-2$^{\alpha}$,11$^{\alpha}$-diazaheptadecan-13$^{\alpha}$-yl)carbamate**



Compounds **49** (67.5 mg, 0.09 mmol, 1.0 eq.) and **47** (47.9 mg, 0.09 mmol, 1.0 eq.) were deprotected according to the general Boc and $^{t}$Bu ester deprotection procedures. The crude amine and carboxylic acid were then reacted directly according to general amide coupling procedure 2. The product was purified using flash silica column chromatography (gradient elution with 2–6% MeOH/$CH_2Cl_2$) to yield the title compound as a brown oil (62.7 mg, 64%): $R_f$ 0.30 (MeOH:$CH_2Cl_2$, 1:11.5); $[\alpha]_D^{25}$ = +11.3 (*c* 0.1, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2935

(C-H, w), 1699 (C=O, s), 1653 (C=O, s), 1453 (s), 1151 (C-O, s), 846 (s), 741 (s); $^{1}$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 8.26–8.19 (1H, m, H-2'''), 7.96–7.90 (1H, m, H-6'''), 7.77–7.63 (3H, m, H-4''', $4^{\beta}$, $5^{\beta}$), 7.57–7.51 (2H, m, H-$1^{\beta}$, $8^{\beta}$), 7.42–7.21 (5H, m, H-5''', $2^{\beta}$, $3^{\beta}$, $6^{\beta}$, $7^{\beta}$), 7.20–7.12 (5H, m, H-4, $4^{\gamma}$, $5^{\gamma}$, $6^{\gamma}$, $7^{\gamma}$), 7.07–6.95 (1H, m, H-5'), 6.55–6.42 (1H, m, H-4'), 6.34 (1H, s, H-7), 6.10–5.95 (2H, m, H-2', 6'), 4.42–4.24 (2H, m, H-$10^{\beta}$), 4.19 (1H, t, *J* 7.1, H-$13^{\alpha}$), 4.13 (1H, t, *J* 7.3, H-$9^{\beta}$), 4.00 (4H, s, H-$1^{\gamma}$, $3^{\gamma}$), 3.77 (2H, t, *J* 6.3, H-1''), 3.67–3.30 (15H, m, H-$3^{\alpha}$, $4^{\alpha}$, $6^{\alpha}$, $7^{\alpha}$, $9^{\alpha}$, $10^{\alpha}$, 8), 3.19 (3H, s, H-9), 2.83 (2H, t, *J* 6.5, H-$17^{\alpha}$), 1.90–1.58 (6H, m, H-2'', $14^{\alpha}$, $16^{\alpha}$), 1.54–1.39 (2H, m, H-$15^{\alpha}$), 0.99 (3H, t, *J* 7.4, H-3''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 172.4, 172.3 (C-$12^{\alpha}$), 165.73, 165.66 (C-$1^{\alpha}$), 160.5 (C-3'), 158.2 (C-1'), 156.59, 156.55 (C-$12^{\beta}$), 154.9 (C-2), 143.92, 143.87 (C-6, $8a^{\beta}$, $9a^{\beta}$), 141.4, 141.3 (C-$4a^{\beta}$, $4b^{\beta}$), 139.9 (C-3'''), 138.6 (C-$3a^{\gamma}$, $7a^{\gamma}$), 135.4 (C-1'''), 132.0 (C-6'''), 130.3 (C-5'), 129.9 (C-4'''), 129.1 (C-5'''), 128.7 (C-5), 127.8 (C-$3^{\beta}$, $6^{\beta}$), 127.4, 127.2 (C-$2^{\beta}$, $5^{\gamma}$, $6^{\gamma}$, $7^{\beta}$), 126.5 (C-3a), 125.7 (C-2'''), 125.3, 125.2 (C-$1^{\beta}$, $8^{\beta}$), 122.5 (C-$4^{\gamma}$, $7^{\gamma}$), 121.5 (C-7a), 120.1 (C-$4^{\beta}$, $5^{\beta}$), 109.4 (C-6'), 109.1 (C-4'), 105.4 (C-4), 104.2 (C-2'), 99.4 (C-7), 70.3. 69.7 (C-1'', $4^{\alpha}$, $6^{\alpha}$, $7^{\alpha}$, $9^{\alpha}$), 67.1 (C-$10^{\beta}$), 59.1 (C-$1^{\gamma}$, $3^{\gamma}$), 55.7 (C-$17^{\alpha}$), 47.2 (C-$9^{\beta}$), 40.1, 40.0 (C-$3^{\alpha}$), 39.3, 39.2 (C-$10^{\alpha}$), 32.5 (C-$14^{\alpha}$), 27.7 (C-$16^{\alpha}$), 27.5, 27.4 (C-8, 9), 23.1 (C-$15^{\alpha}$), 22.6 (C-2''), 10.7 (C-3''); HRMS *m/z* ($ESI^+$) [Found: 1116.4519, $C_{60}H_{67}N_7O_{11}SNa$ requires $[M+Na]^+$ 1116.4512]; LRMS *m/z* ($ESI^+$) 1094 ($[M+H]^+$, 100%); HPLC retention time: 9.2 min, 83.8% (220 nm), 90.1% (254 nm), 92.2% (280 nm).

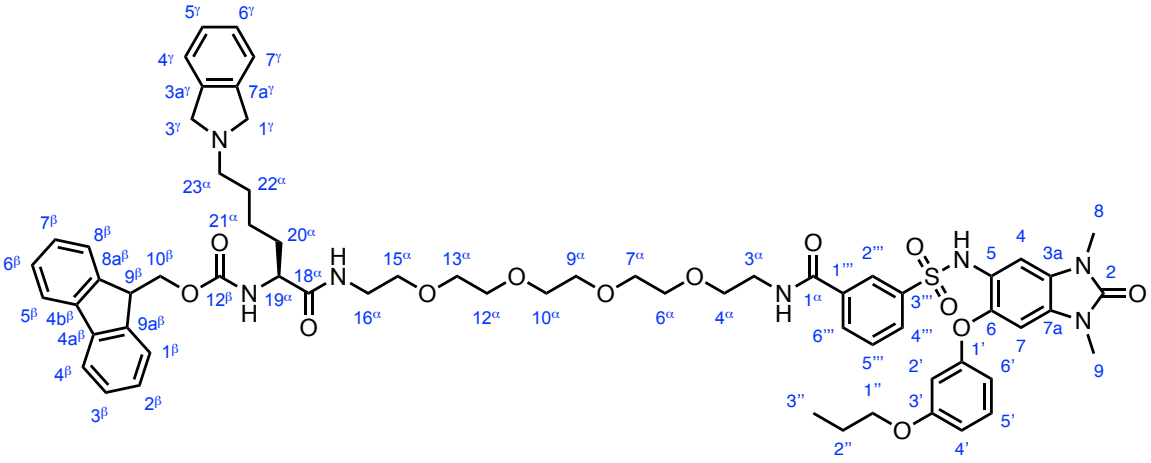
**(53) (9$^{\beta}$*H*-Fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-20$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,15$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$-trioxa-2$^{\alpha}$,14$^{\alpha}$-diazaicosan-16$^{\alpha}$-yl)carbamate**



Compounds **50** (71.1 mg, 0.09 mmol, 1.0 eq.) and **47** (47.6 mg, 0.09 mmol, 1.0 eq.) were deprotected according to the general Boc and $^{t}$Bu ester deprotection procedures. The crude amine and carboxylic acid were then reacted directly according to general amide coupling procedure 2. The product was purified using flash silica column chromatography (gradient elution with 2–7% MeOH/$CH_2Cl_2$) to yield the title compound as a brown oil (50.2 mg, 49%): $R_f$ 0.38 (MeOH:$CH_2Cl_2$, 1:9); $[\alpha]_D^{25}$ = −12.1 (*c* 0.1, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2943 (C-H, w), 1700 (C=O, s), 1654 (C=O, s), 1456 (s), 1152 (C-O, s), 847 (s), 744 (s); $^1$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 8.26–8.19 (1H, m, H-2'''), 7.99–7.92 (1H, m, H-6'''), 7.72 (2H, d, *J* 7.6, H-4$^{\beta}$, 5$^{\beta}$), 7.67 (1H, d, *J* 8.1, H-4'''), 7.57–7.52 (2H, m, H-1$^{\beta}$, 8$^{\beta}$), 7.41–7.30 (3H, m, H-5''', 3$^{\beta}$, 6$^{\beta}$), 7.29–7.23 (2H, m, H-2$^{\beta}$, 7$^{\beta}$), 7.22 (1H, s, H-4), 7.20–7.12 (4H, m, H-4$^{\gamma}$, 5$^{\gamma}$, 6$^{\gamma}$, 7$^{\gamma}$), 7.02 (1H, t, *J* 8.2, H-5'), 6.56–6.48 (1H, m, H-4'), 6.34 (1H, s, H-7), 6.10–6.00 (2H, m, H-2', 6'), 4.45–4.24 (2H, m, H-10$^{\beta}$), 4.20 (1H, t, *J* 6.4, H-16$^{\alpha}$), 4.15 (1H, t, *J* 7.3, H-9$^{\beta}$), 3.99 (4H, s, H-1$^{\gamma}$, 3$^{\gamma}$), 3.78 (2H, t, *J* 6.6, H-1''), 3.66–3.42 (14H, m, H-3$^{\alpha}$, 4$^{\alpha}$, 6$^{\alpha}$, 7$^{\alpha}$, 9$^{\alpha}$,

$10^{\alpha}$, $12^{\alpha}$), 3.38 (3H, s, H-8), 3.36–3.30 (2H, m, H-$13^{\alpha}$), 3.20 (3H, s, H-9), 2.79 (2H, t, *J* 7.3, H-$20^{\alpha}$), 1.91–1.54 (6H, m, H-2'', $17^{\alpha}$, $19^{\alpha}$), 1.50–1.39 (2H, m, H-$18^{\alpha}$), 1.00 (3H, t, *J* 7.4, H-3''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 172.2, 172.1 (C-$15^{\alpha}$), 165.6, 165.7 (C-$1^{\alpha}$), 160.5 (C-3'), 158.1 (C-1'), 156.5, 156.4 (C-$12^{\beta}$), 154.9 (C-2), 143.94, 143.89, 143.77 (C-6, $8a^{\beta}$, $9a^{\beta}$), 141.4 (C-$4a^{\beta}$, $4b^{\beta}$), 139.7 (C-3'''), 138.9 (C-$3a^{\gamma}$, $7a^{\gamma}$), 135.5 (C-1'''), 132.1 (C-6'''), 130.3 (C-5'), 129.9 (C-4'''), 129.1 (C-5'''), 128.7 (C-5), 127.8 (C-$3^{\beta}$, $6^{\beta}$), 127.23, 127.20 (C-$2^{\beta}$, $5^{\gamma}$, $6^{\gamma}$, $7^{\beta}$), 126.5 (C-3a), 125.7 (C-2'''), 125.24, 125.20 (C-$1^{\beta}$, $8^{\beta}$), 122.5 (C-$4^{\gamma}$, $7^{\gamma}$), 121.4 (C-7a), 120.1 (C-$4^{\beta}$, $5^{\beta}$), 109.4 (C-6'), 109.3 (C-4'), 105.1 (C-4), 104.2 (C-2'), 99.4 (C-7), 70.43, 70.38, 70.2, 70.1 (C-$4^{\alpha}$, $6^{\alpha}$, $7^{\alpha}$, $9^{\alpha}$, $10^{\alpha}$), 69.73 (C-$12^{\alpha}$), 69.71 (C-1''), 67.1 (C-$10^{\beta}$), 59.1 (C-$1^{\gamma}$, $3^{\gamma}$), 55.7 (C-$20^{\alpha}$), 54.9, 54.8 (C-$16^{\alpha}$), 47.2 (C-$9^{\beta}$), 40.1, 39.9 (C-$3^{\alpha}$), 39.4, 39.2 (C-$13^{\alpha}$), 32.8 (C-$17^{\alpha}$), 27.8 (C-$19^{\alpha}$), 27.5, 27.4 (C-8, 9), 23.0 (C-$18^{\alpha}$), 22.6 (C-2''), 10.7 (C-3''); HRMS *m/z* ($ESI^+$) [Found: 1138.4966, $C_{62}H_{72}N_7O_{12}S$ requires $[M+H]^+$ 1138.4954]; LRMS *m/z* ($ESI^+$) 1138 ($[M+H]^+$, 100%); HPLC retention time: 9.2 min, 90.5% (220 nm), 93.6% (254 nm), 93.6% (280 nm).

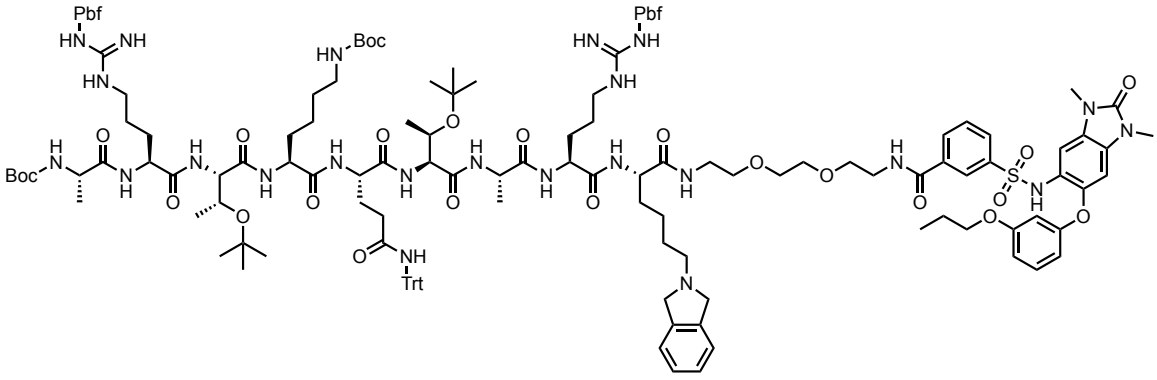
**(54) (9$^{\beta}$*H*-Fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-23$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,18$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$,14$^{\alpha}$-tetraoxa-2$^{\alpha}$,17$^{\alpha}$-diazatricosan-19$^{\alpha}$-yl)carbamate**



Compounds **51** (62.2 mg, 0.07 mmol, 1.0 eq.) and **47** (39.5 mg, 0.07 mmol, 1.0 eq.) were deprotected according to the general Boc and $^{t}$Bu ester deprotection procedures. The crude amine and carboxylic acid were then reacted directly according to general amide coupling procedure 2. The product was purified using flash silica column chromatography (gradient elution with 2-8% MeOH/$CH_2Cl_2$) to yield the title compound as a brown oil (63.7 mg, 77%): $R_f$ 0.42 (MeOH:$CH_2Cl_2$, 1:9); $[\alpha]_D^{25}$ = −3.6 (*c* 0.1, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2929 (C-H, w), 1703 (C=O, s), 1663 (C=O, s), 1462 (s), 1260 (s), 1132 (C-O, s), 741 (s);$^{1}$H NMR (600 MHz, $CDCl_3$) $\delta_H$ 8.28 (1H, m, H-2'''), 7.99–7.93 (1H, m, H-6'''), 7.73 (2H, d, *J* 7.6, H-4$^{\beta}$, 5$^{\beta}$), 7.69 (1H, d, *J* 8.0, H-4'''), 7.59–7.52 (2H, m, H-1$^{\beta}$, 8$^{\beta}$), 7.37 (2H, t, *J* 7.5, H-3$^{\beta}$, 6$^{\beta}$), 7.32 (1H, t, *J* 8.0, H-5'''), 7.30–7.24 (2H, m, H-2$^{\beta}$, 7$^{\beta}$), 7.23 (1H, s, H-4), 7.20–7.12 (4H, m, H-4$^{\gamma}$, 5$^{\gamma}$, 6$^{\gamma}$, 7$^{\gamma}$), 7.02 (1H, t, *J* 8.1, H-5'), 6.58–6.49 (1H, m, H-4'), 6.36 (1H, s, H-7), 6.10–6.00 (2H, m, H-2', 6'), 4.43–4.27 (2H, m, H-10$^{\beta}$), 4.26–4.20 (1H, m, H-19$^{\alpha}$), 4.17 (1H, t, *J* 7.3, H-9$^{\beta}$), 3.96 (4H, s, H-1$^{\gamma}$, 3$^{\gamma}$), 3.79 (2H, t, *J* 6.5, H-1''), 3.68–3.31 (23H,

m, H-3$^{\alpha}$, 4$^{\alpha}$, 6$^{\alpha}$, 7$^{\alpha}$, 9$^{\alpha}$, 10$^{\alpha}$, 12$^{\alpha}$, 13$^{\alpha}$, 15$^{\alpha}$, 16$^{\alpha}$, 8), 3.21 (3H, s, H-9), 2.80–2.71 (2H, m, H-23$^{\alpha}$), 1.92–1.52 (6H, m, H-2'', 20$^{\alpha}$, 22$^{\alpha}$), 1.50–1.38 (2H, m, H-21$^{\alpha}$), 1.00 (3H, t, *J* 7.4, H-3''); $^{13}$C NMR (151 MHz, $CDCl_3$) $\delta_C$ 172.0 (C-18$^{\alpha}$), 165.6 (C-1$^{\alpha}$), 160.5 (C-3'), 158.1 (C-1'), 156.3 (C-12$^{\beta}$), 154.9 (C-2), 144.0, 143.9, 143.7 (C-6, 8a$^{\beta}$, 9a$^{\beta}$), 141.4 (C-4a$^{\beta}$, 4b$^{\beta}$), 139.8 (C-3'''), 139.2 (C-3a$^{\gamma}$, 7a$^{\gamma}$), 135.6 (C-1'''), 132.1 (C-6'''), 130.3 (C-5'), 129.8 (C-4'''), 129.0 (C-5'''), 128.7 (C-5), 127.8 (C-3$^{\beta}$, 6$^{\beta}$), 127.18, 127.11 (C-2$^{\beta}$, 5$^{\gamma}$, 6$^{\gamma}$, 7$^{\beta}$), 126.6 (C-3a), 125.9 (C-2'''), 125.25, 125.22 (C-1$^{\beta}$, 8$^{\beta}$), 122.5 (C-4$^{\gamma}$, 7$^{\gamma}$), 121.5 (C-7a), 120.1 (C-4$^{\beta}$, 5$^{\beta}$), 109.4 (C-6'), 109.3 (C-4'), 105.1 (C-4), 104.2 (C-2'), 99.5 (C-7), 70.5, 70.3, 70.2, 69.8, 69.73, 69.71 (C-1'', 4$^{\alpha}$, 6$^{\alpha}$, 7$^{\alpha}$, 9$^{\alpha}$, 10$^{\alpha}$, 12$^{\alpha}$, 13$^{\alpha}$, 15$^{\alpha}$), 67.1 (C-10$^{\beta}$), 59.0 (C-1$^{\gamma}$, 3$^{\gamma}$), 55.8 (C-23$^{\alpha}$), 54.9 (C-19$^{\alpha}$), 47.3 (C-9$^{\beta}$), 40.1 (C-3$^{\alpha}$), 39.4 (C-16$^{\alpha}$), 33.1 (C-20$^{\alpha}$), 28.0 (C-22$^{\alpha}$), 27.5, 27.4 (C-8, 9), 23.1 (C-21$^{\alpha}$), 22.6 (C-2''), 10.7 (C-3''); HRMS *m/z* (ESI$^{+}$) [Found: 1182.5219, $C_{64}H_{76}N_7O_{13}S$ requires $[M+H]^+$ 1182.5216]; LRMS *m/z* (ESI$^{+}$) 1182 ($[M+H]^+$, 100%); HPLC retention time: 9.3 min, 91.2% (220 nm), 95.1% (254 nm), 96.4% (280 nm).

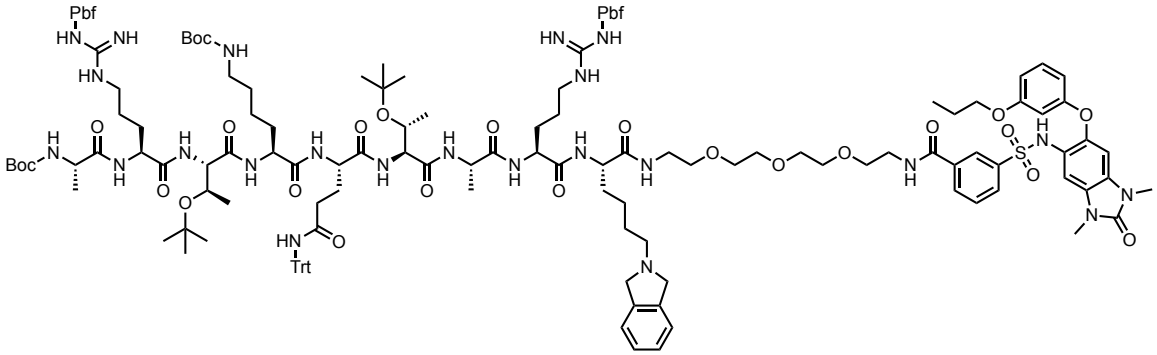
**(55) *tert*-Butyl ((13*S*,16*S*,19*S*,22*S*,25*S*,28*S*)-22-((*R*)-1-(*tert*-butoxy)ethyl)-28-((6*S*,9*S*,12*S*)-12-((*R*)-1-(*tert*-butoxy)ethyl)-2,2,6-trimethyl-4,7,10-trioxo-9-(3-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)propyl)-3-oxa-5,8,11-triazatridecan-13-amido)-1-(3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-13-(4-(isoindolin-2-yl)butyl)-19-methyl-1,12,15,18,21,24,27-heptaoxo-25-(3-oxo-3-(tritylamino)propyl)-16-(3-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)propyl)-5,8-dioxa-2,11,14,17,20,23,26-heptaazadotriacontan-32-yl)carbamate (PDC1)**



Compound **52** was reacted according to the general Fmoc deprotection procedure to yield the purified amine (diformate salt, 14.6 mg, 15.1 μmol, 1.0 eq.) which was reacted with compound **48** (31.4 mg, 15.1 μmol, 1.0 eq.) according to general amide coupling procedure 3. The product was purified using flash silica column chromatography (gradient elution with 2–6% (1% $NH_4OH$ in MeOH)/$CH_2Cl_2$ then gradient elution using a Biotage® system with 2–20% MeOH/$CH_2Cl_2$) to yield the title compound as a colourless oil (22.5 mg, 52%): $R_f$ 0.40 (1% $NH_4OH$ in MeOH:$CH_2Cl_2$, 1:9); HRMS *m/z* ($ESI^+$) [Found: 2843.4341, $C_{145}H_{204}N_{23}O_{30}S_3$ requires $[M+H]^+$ 2843.4301]; LRMS *m/z* ($ESI^+$) 1422 ($[M+2H]^{2+}$,

100%); 1433 ([M+H+Na]$^+$, 54%); HPLC retention time: 12.1 min, 95.4% (220 nm), 95.3% (254 nm), 96.5% (280 nm).

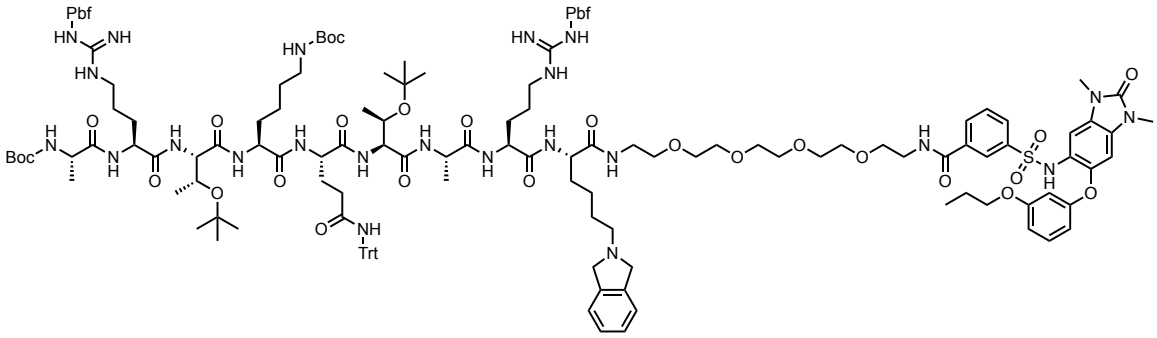
**(56) *tert*-Butyl ((16*S*,19*S*,22*S*,25*S*,28*S*,31*S*,34*S*,37*S*,40*S*)-25,34-bis((*R*)-1-(*tert*-butoxy)ethyl)-31-(4-((*tert*-butoxycarbonyl)amino)butyl)-1-(3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-16-(4-(isoindolin-2-yl)butyl)-22-methyl-1,15,18,21,24,27,30,33,36,39-decaoxo-28-(3-oxo-3-(tritylamino)propyl)-19,37-bis(3-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)propyl)-5,8,11-trioxa-2,14,17,20,23,26,29,32,35,38-decaazahentetracontan-40-yl)carbamate (PDC2)**



Compound **53** was reacted according to the general Fmoc deprotection procedure to yield the purified amine (diformate salt, 11.3 mg, 11.2 μmol, 1.0 eq.) which was reacted with compound **48** (23.3 mg, 11.2 μmol, 1.0 eq.) according to general amide coupling procedure 3. The product was purified using flash silica column chromatography (gradient elution with 2–6% (1% $NH_4OH$ in MeOH)/$CH_2Cl_2$ then gradient elution using a Biotage® system with 2–20% MeOH/$CH_2Cl_2$) to yield the title compound as a colourless oil (18.3 mg, 57%): $R_f$ 0.41 (1% $NH_4OH$ in MeOH:$CH_2Cl_2$, 1:9); HRMS *m/z* (ESI$^+$) [Found: 2887.4678, $C_{147}H_{208}N_{23}O_{31}S_3$ requires $[M+H]^+$ 2887.4563]; LRMS *m/z* (ESI$^+$) 1444 ($[M+2H]^{2+}$,

100%), 1455 ([M+H+Na]$^{2+}$, 34%); HPLC retention time: 12.2 min, 97.3% (220 nm), 95.7% (254 nm), 95.4% (280 nm).

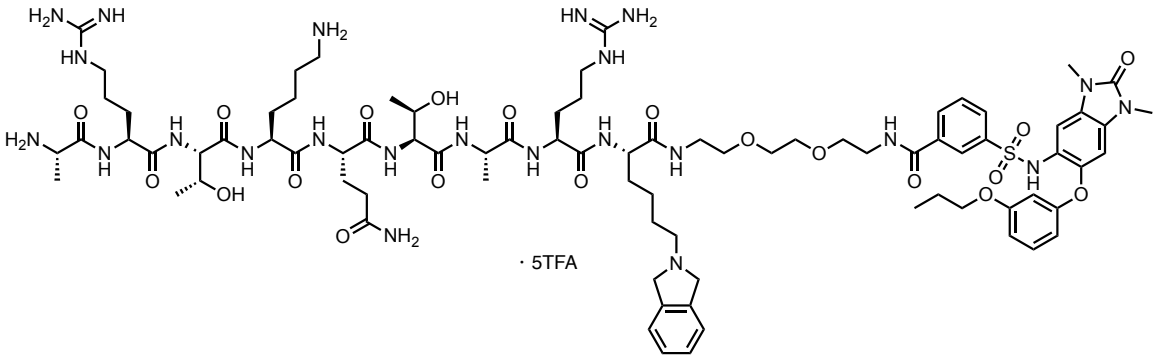
**(57) *tert*-Butyl ((19*S*,22*S*,25*S*,28*S*,31*S*,34*S*)-28-((*R*)-1-(*tert*-butoxy)ethyl)-34-((6*S*,9*S*,12*S*)-12-((*R*)-1-(*tert*-butoxy)ethyl)-2,2,6-trimethyl-4,7,10-trioxo-9-(3-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)propyl)-3-oxa-5,8,11-triazatridecan-13-amido)-1-(3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-19-(4-(isoindolin-2-yl)butyl)-25-methyl-1,18,21,24,27,30,33-heptaoxo-31-(3-oxo-3-(tritylamino)propyl)-22-(3-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)propyl)-5,8,11,14-tetraoxa-2,17,20,23,26,29,32-heptaazaoctatriacontan-38-yl)carbamate (PDC3)**



Compound **54** was reacted according to the general Fmoc deprotection procedure to yield the purified amine (diformate salt, 20.8 mg, 19.8 μmol, 1.0 eq.) which was reacted with compound **48** (41.2 mg, 19.8 μmol, 1.0 eq.) according to general amide coupling procedure 3. The product was purified using flash silica column chromatography (gradient elution with 2–7% (1% $NH_4OH$ in MeOH)/$CH_2Cl_2$ then gradient elution using a Biotage® system with 2–20% MeOH/$CH_2Cl_2$) to yield the title compound as a colourless oil (24.4 mg, 42%): $R_f$ 0.41 (1% $NH_4OH$ in MeOH:$CH_2Cl_2$, 1:9); HRMS *m/z* (ESI$^+$) [Found: 2931.4868,

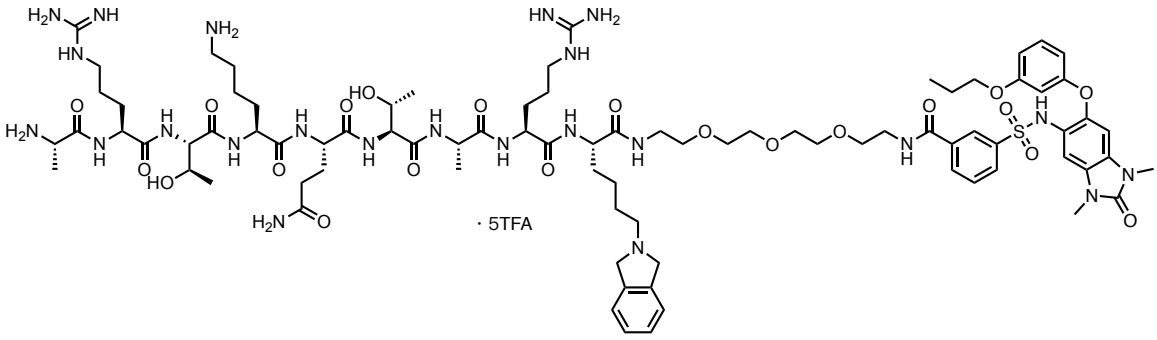
$C_{149}H_{212}N_{23}O_{32}S_3$ requires $[M+H]^+$ 2931.4826] ; LRMS *m/z* ($ESI^+$) 1466 ($[M+2H]^{2+}$, 22%), 1477 ($[M+H+Na]^{2+}$, 100%); HPLC retention time: 12.1 min, 95.2% (220 nm), 95.3% (254 nm), 95.0% (280 nm).

**(58) (*S*)-2-((*S*)-6-Amino-2-((2*S*,3*R*)-2-((*S*)-2-((*S*)-2-aminopropanamido)-5-guanidinopentanamido)-3-hydroxybutanamido)hexanamido)-*N*$^1$-((13*S*,16*S*,19*S*,22*S*,23*R*)-1-(3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-16-(3-guanidinopropyl)-23-hydroxy-13-(4-(isoindolin-2-yl)butyl)-19-methyl-1,12,15,18,21-pentaoxo-5,8-dioxa-2,11,14,17,20-pentaazatetracosan-22-yl)pentanediamide pentatrifluoroacetic acid (PDC4)**



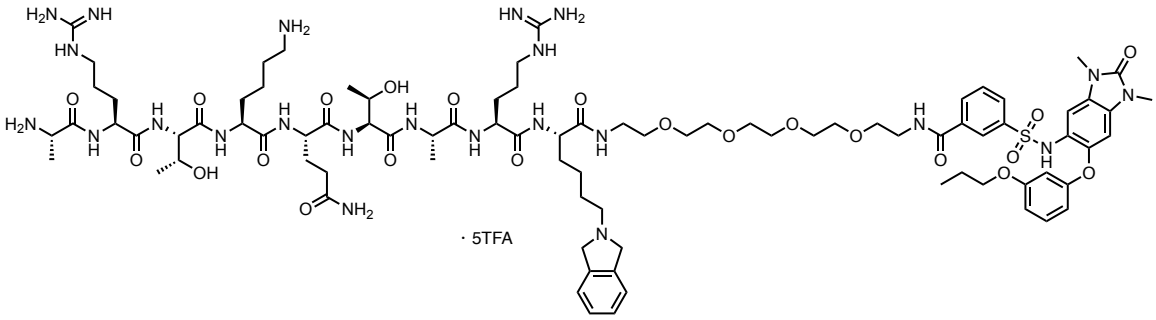

Compound **55** (22.5 mg, 7.91 µmol, 1.0 eq.) was reacted according to the global deprotection procedure. The product was purified using semi-preparative reverse phase HPLC (gradient method 3, 20 mL min$^{-1}$, Agilent Zorbax PrepHT column [5 µm, 21.2 mm × 50 mm]) and lyophilised to yield the title compound as a colourless solid (5.9 mg, 32%): HRMS *m/z* ($ESI^+$) [Found: 1784.9306, $C_{82}H_{126}N_{23}O_{20}S$ requires $[M-5TFA+H]^+$ 1784.9265]; LC-MS *m/*z ($ESI^+$) 5.0 min, 358.2 ($[M-5TFA+5H]^{5+}$), 447.4 ($[M-5TFA+4H]^{4+}$), 596.0 ($[M-5TFA+3H]^{3+}$), 893.3 ($[M-5TFA+2H]^{2+}$); HPLC retention time: 5.8 min, 95.5% (220 nm), 95.2% (254 nm), 95.2% (280 nm).

**(59) (*S*)-2-((*S*)-6-Amino-2-((2*S*,3*R*)-2-((*S*)-2-((*S*)-2-aminopropanamido)-5-guanidinopentanamido)-3-hydroxybutanamido)hexanamido)-*N*$^1$-((16*S*,19*S*,22*S*,25*S*,26*R*)-1-(3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-19-(3-guanidinopropyl)-26-hydroxy-16-(4-(isoindolin-2-yl)butyl)-22-methyl-1,15,18,21,24-pentaoxo-5,8,11-trioxa-2,14,17,20,23-pentaazaheptacosan-25-yl)pentanediamide pentatrifluoroacetic acid (PDC5)**



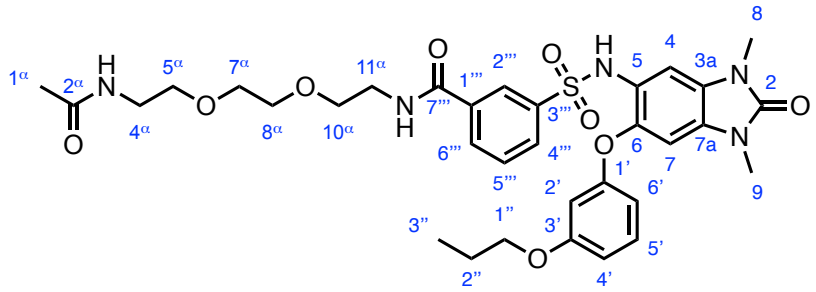

Compound **56** (18.3 mg, 6.34 μmol, 1.0 eq.) was reacted according to the global deprotection procedure. The product was purified using semi-preparative reverse phase HPLC (gradient method 3, 20 mL min$^{-1}$, Agilent Zorbax PrepHT column [5 μm, 21.2 mm × 50 mm]) and lyophilised to yield the title compound as a colourless solid (7.8 mg, 51%): HRMS *m/z* (ESI$^+$) [Found: 1828.9561, $C_{84}H_{130}N_{23}O_{21}S$ requires [M−5TFA+H]$^+$ 1828.9527]; LC-MS *m/z* (ESI$^+$) 5.0 min, 366.9 ([M−5TFA+5H]$^{5+}$), 458.4 ([M−5TFA+4H]$^{4+}$), 610.7 ([M−5TFA+3H]$^{3+}$), 915.4 ([M−5TFA+2H]$^{2+}$); HPLC retention time: 5.7 min, 97.3% (220 nm), 97.1% (254 nm), 96.2% (280 nm).

**(60) (*S*)-2-((*S*)-6-Amino-2-((2*S*,3*R*)-2-((*S*)-2-((*S*)-2-aminopropanamido)-5-guanidinopentanamido)-3-hydroxybutanamido)hexanamido)-*N*$^1$-((19*S*,22*S*,25*S*,28*S*,29*R*)-1-(3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-22-(3-guanidinopropyl)-29-hydroxy-19-(4-(isoindolin-2-yl)butyl)-25-methyl-1,18,21,24,27-pentaoxo-5,8,11,14-tetraoxa-2,17,20,23,26-pentaazatriacontan-28-yl)pentanediamide pentatrifluoroacetic acid (PDC6)**




Compound **57** (18.8 mg, 6.41 μmol, 1.0 eq.) was reacted according to the global deprotection procedure. The product was purified using semi-preparative reverse phase HPLC (gradient method 3, 20 mL min$^{-1}$, Agilent Zorbax PrepHT column [5 μm, 21.2 mm × 50 mm]) and lyophilised to yield the title compound as a colourless solid (7.8 mg, 50%): HRMS *m/z* (ESI$^+$) [Found: 1872.9797, $C_{86}H_{134}N_{23}O_{22}S$ requires $[M-5TFA+H]^+$ 1872.9789]; LC-MS *m/z* (ESI$^+$) 5.2 min, 375.7 ($[M-5TFA+5H]^{5+}$), 469.4 ($[M-5TFA+4H]^{4+}$), 625.3 ($[M-5TFA+3H]^{3+}$), 937.3 ($[M-5TFA+2H]^{2+}$); HPLC retention time: 5.9 min, 97.7% (220 nm), 97.6% (254 nm), 97.3% (280 nm).

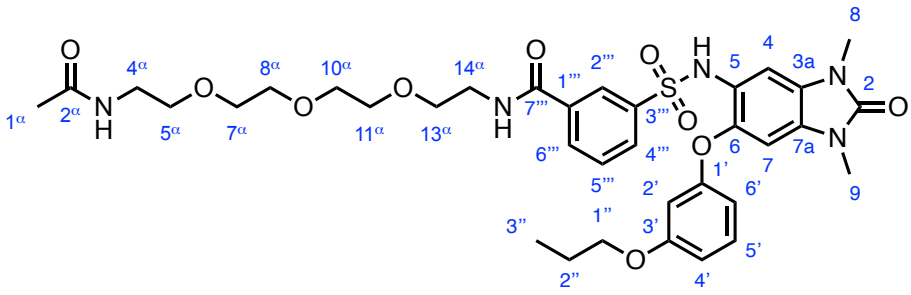
**(61) 3’’’-(*N*-(1,3-Dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$-dioxa-3$^{\alpha}$-azaundecan-11$^{\alpha}$-yl)benzamide**



Compound **49** (58.3 mg, 78.6 μmol, 1.0 eq.) was deprotected according to the general Boc deprotection procedure to yield the crude amine which was directly acetylated according to the general acetylation procedure. The product was purified using semi-preparative reverse phase HPLC (gradient method 2, 20 mL min$^{-1}$, Agilent Prep C18 column [5 μm, 21.2 mm × 50 mm]) and lyophilised to yield the title compound as a colourless oil (38.9 mg, 72%): $R_f$ 0.16 (MeOH:$CH_2Cl_2$, 1:19); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2938 (C-H, w), 2878 (C-H, w), 1693 (C=O, s), 1642 (C=O, s), 1453 (s), 1153 (C-O, s), 1132 (C-O, s); $^1$H NMR (600 MHz, $CD_3OD$) $\delta_H$ 8.20 (1H, dd, *J* 1.8, 1.7, H-2’’’), 7.97 (1H, ddd, *J* 7.8, 1.7, 1.1 H-6’’’), 7.81 (1H, ddd, *J* 7.8, 1.8, 1.1, H-4’’’), 7.43 (1H, dd, *J* 7.8, 7.8, H-5’’’), 7.39 (1H, s, H-4), 7.05 (1H, dd, *J* 8.5, 8.4, H-5’), 6.60 (1H, s, H-7), 6.58 (1H, ddd, *J* 8.4, 2.3, 1.0, H-4’), 6.11–6.07 (2H, m, H-2’, 6’), 3.84 (2H, t, *J* 6.5, H-1’’), 3.70–3.52 (10H, m, H-5$^{\alpha}$, 7$^{\alpha}$, 8$^{\alpha}$, 10$^{\alpha}$, 11$^{\alpha}$), 3.46 (3H, s, H-8), 3.35* (2H, m, H-4$^{\alpha}$), 3.27 (3H, s, H-9), 1.95 (3H, s, H-1$^{\alpha}$), 1.83–1.75 (2H, m, H-2’’), 1.06 (3H, t, *J* 7.4, H-3’’); $^{13}$C NMR (151 MHz, $CD_3OD$) $\delta_C$ 173.4 (C-2$^{\alpha}$), 167.9 (C-7’’’), 161.7 (C-3’), 159.6 (C-1’), 156.5 (C-2), 146.7 (C-6), 142.0 (C-3’’’), 136.2 (C-1’’’), 132.3 (C-6’’’), 131.12, 131.10 (C-4’’’, C-5’), 130.1 (C-5’’’), 129.9 (C-3a), 127.3, 127.2 (C-2’’’, C-7a), 123.0 (C-5), 110.6 (C-6’), 110.3 (C-4’), 108.2 (C-4), 105.3

(C-2'), 100.8 (C-7), 71.32, 71.32, 70.59, 70.56 (C-5$^{\alpha}$, 7$^{\alpha}$, 8$^{\alpha}$, 10$^{\alpha}$), 70.5 (C-1''), 41.0 (C-11$^{\alpha}$), 40.4 (C-4$^{\alpha}$), 27.6, 27.5 (C-8, 9), 23.6 (C-2''), 22.5 (C-1$^{\alpha}$), 10.9 (C-3''); HRMS *m/z* (ESI$^{+}$) [Found: 684.2706, $C_{33}H_{42}N_5O_9S$ requires $[M+H]^+$ 684.2698]; LRMS *m/z* (ESI$^{+}$) 706 ($[M+Na]^+$, 100%); HPLC retention time: 8.4 min, >99.9% (220 nm), >99.9% (254 nm), >99.9% (280 nm). *Peak obscured by NMR solvent peak and was assigned by HSQC.

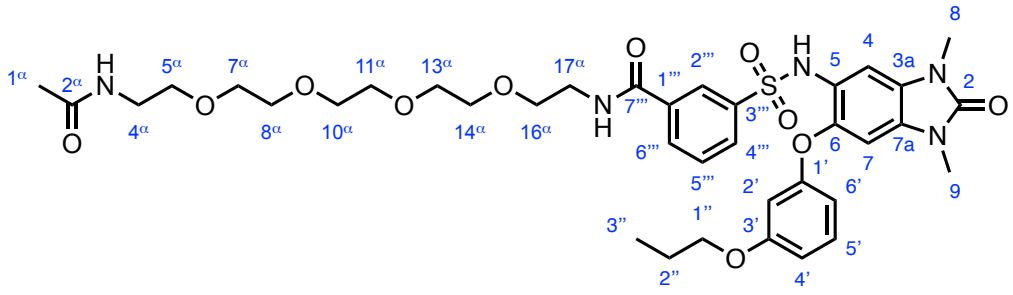
**(62) 3'''-(*N*-(1,3-Dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$,12$^{\alpha}$-trioxa-3$^{\alpha}$-azatetradecan-14$^{\alpha}$-yl)benzamide**



Compound **50** (79.8 mg, 102 μmol, 1.0 eq.) was deprotected according to the general Boc deprotection procedure to yield the crude amine which was directly acetylated according to the general acetylation procedure. The product was purified using semi-preparative reverse phase HPLC (gradient method 2, 20 mL min$^{-1}$, Agilent Prep C18 column [5 μm, 21.2 mm × 50 mm]) and lyophilised to yield the title compound as a colourless oil (46.7 mg, 63%): $R_f$ 0.19 (MeOH:$CH_2Cl_2$, 1:19); $\tilde{\nu}_{max}$ (neat)/cm$^{-1}$ 2955 (C-H, w), 2881 (C-H, w), 1695 (C=O, s), 1643 (C=O, s), 1461 (s), 1151 (C-O, s), 1134 (C-O, s); $^1$H NMR (600 MHz, $CD_3OD$) $\delta_H$ 8.20 (1H, dd, *J* 1.9, 1.9, H-2'''), 7.97 (1H, ddd, *J* 7.8, 1.9, 1.2, H-6'''), 7.81 (1H, ddd, *J* 7.8, 1.9, 1.2, H-4'''), 7.43 (1H, dd, *J* 7.8, 7.8, H-5'''), 7.39 (1H, s, H-4), 7.09–7.03 (1H, m, H-5'), 6.63–6.56 (2H, m, H-4', 7), 6.11–6.06 (2H, m, H-2', 6'),

3.84 (2H, t, *J* 6.5, H-1’’), 3.72–3.50 (14H, m, H-5$^{\alpha}$, 7$^{\alpha}$, 8$^{\alpha}$, 10$^{\alpha}$, 11$^{\alpha}$, 13$^{\alpha}$, 14$^{\alpha}$), 3.47 (3H, s, H-8), 3.34* (2H, m, H-4$^{\alpha}$), 3.27 (3H, s, H-9), 1.95 (3H, s, H-1$^{\alpha}$), 1.83–1.74 (2H, m, H-2’’), 1.06 (3H, t, *J* 7.4, H-3’’); $^{13}$C NMR (151 MHz, $CD_3OD$) $\delta_C$ 173.3 (C-2$^{\alpha}$), 167.9 (C-7’’’), 161.7 (C-3’), 159.6 (C-1’), 156.6 (C-2), 146.7 (C-6), 142.0 (C-3’’’), 136.2 (C-1’’’), 132.3 (C-6’’’), 131.1, 131.1 (C-4’’’, C-5’), 130.1 (C-5’’’), 130.0 (C-3a), 127.3, 127.2 (C-2’’’, C-7a), 123.1 (C-5), 110.6 (C-6’), 110.3 (C-4’), 108.3 (C-4), 105.3 (C-2’), 100.8 (C-7), 71.59, 71.57, 71.3, 71.2, 70.6, 70.5, 70.4 (C-1’’, 5$^{\alpha}$, 7$^{\alpha}$, 8$^{\alpha}$, 10$^{\alpha}$, 11$^{\alpha}$, 13$^{\alpha}$), 41.0 (C-14$^{\alpha}$), 40.4 (C-4$^{\alpha}$), 27.6, 27.5 (C-8, 9), 23.6 (C-2’’), 22.5 (C-1$^{\alpha}$), 10.9 (C-3’’); HRMS *m/z* (ESI$^+$) [Found: 728.2952, $C_{35}H_{46}N_5O_{10}S$ requires $[M+H]^+$ 728.2960]; LRMS *m/z* (ESI$^+$) 750 ($[M+Na]^+$, 100%); HPLC retention time: 8.4 min, 97.1% (220 nm), 96.5% (254 nm), 95.1% (280 nm). *Peak obscured by NMR solvent peak and was assigned by HSQC.

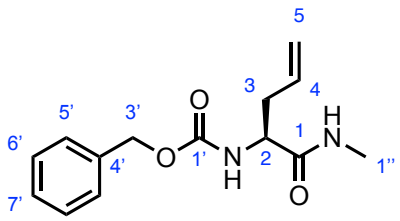
**(63) 3’’’-(*N*-(1,3-Dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$,12$^{\alpha}$,15$^{\alpha}$-tetraoxa-3$^{\alpha}$-azaheptadecan-17$^{\alpha}$-yl)benzamide**



Compound **51** (77.2 mg, 93.0 µmol, 1.0 eq.) was deprotected according to the general Boc deprotection procedure to yield the crude amine which was directly acetylated according to the general acetylation procedure. The product was purified using semi-preparative reverse phase HPLC (gradient method 2, 20 mL min$^{-1}$, Agilent Prep C18 column [5 µm,

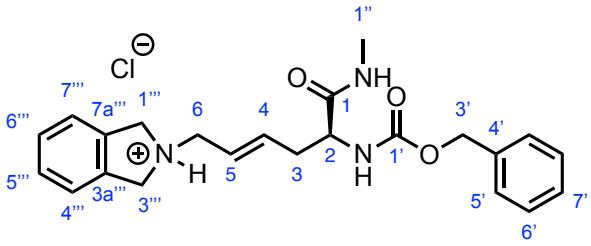
21.2 mm × 50 mm]) and lyophilised to yield the title compound as a colourless oil (31.1 mg, 43%): $R_f$ 0.12 (MeOH:$CH_2Cl_2$, 1:19); $\tilde{\nu}_{max}$ (neat)/$cm^{-1}$ 2935 (C-H, w), 2878 (C-H, w), 1697 (C=O, s), 1642 (C=O, s), 1462 (s), 1152 (C-O, s), 1134 (C-O, s); $^1$H NMR (600 MHz, $CD_3OD$) $\delta_H$ 8.21 (1H, dd, *J* 1.8, 1.8, H-2’’’), 7.97 (1H, ddd, *J* 7.9, 1.8, 1.2, H-6’’’), 7.82 (1H, ddd, *J* 7.9, 1.8, 1.2, H-4’’’), 7.44 (1H, dd, *J* 7.9, 7.9, H-5’’’), 7.39 (1H, s, H-4), 7.09–7.03 (1H, m, H-5’), 6.64–6.55 (2H, m, H-4’, 7), 6.12–6.06 (2H, m, H-2’, 6’), 3.84 (2H, t, *J* 6.5, H-1’’), 3.69–3.50 (18H, m, H-5$^\alpha$, 7$^\alpha$, 8$^\alpha$, 10$^\alpha$, 11$^\alpha$, 13$^\alpha$, 14$^\alpha$, 16$^\alpha$, 17$^\alpha$), 3.47 (3H, s, H-8), 3.35* (2H, m, H-4$^\alpha$), 3.27 (3H, s, H-9), 1.96 (3H, s, H-1$^\alpha$), 1.84–1.74 (2H, m, H-2’’), 1.06 (3H, t, *J* 7.4, H-3’’); $^{13}$C NMR (151 MHz, $CD_3OD$) $\delta_C$ 173.3 (C-2$^\alpha$), 167.9 (C-7’’’), 161.7 (C-3’), 159.6 (C-1’), 156.5 (C-2), 146.7 (C-6), 142.0 (C-3’’’), 136.2 (C-1’’’), 132.3 (C-6’’’), 131.1, 131.1 (C-4’’’, C-5’), 130.1 (C-5’’’), 130.0 (C-3a), 127.3, 127.2 (C-2’’’, C-7a), 123.1 (C-5), 110.6 (C-6’), 110.3 (C-4’), 108.3 (C-4), 105.3 (C-2’), 100.8 (C-7), 71.57, 71.53, 71.52, 71.50, 71.3, 71.2, 70.6, 70.51, 70.45 (C-1’’, 5$^\alpha$, 7$^\alpha$, 8$^\alpha$, 10$^\alpha$, 11$^\alpha$, 13$^\alpha$, 14$^\alpha$, 16$^\alpha$), 41.1 (C-16$^\alpha$), 40.5 (C-4$^\alpha$), 27.6, 27.5 (C-8, 9), 23.6 (C-2’’), 22.5 (C-1$^\alpha$), 10.9 (C-3’’); HRMS *m/z* ($ESI^+$) [Found: 772.3239, $C_{37}H_{50}N_5O_{11}S$ requires $[M+H]^+$ 772.3222]; LRMS *m/z* ($ESI^+$) 794 ($[M+Na]^+$, 100%); HPLC retention time: 8.4 min, >99.9% (220 nm), >99.9% (254 nm), >99.9% (280 nm). *Peak obscured by NMR solvent peak and was assigned by HSQC.

**(64) Benzyl (*S*)-(1-(methylamino)-1-oxopent-4-en-2-yl)carbamate**




Methyl amine (2 M in THF, 5 mL, 10.0 mmol, 1.0 eq.) was added to a solution of (*S*)-2-(((benzyloxy)carbonyl)amino)pent-4-enoic acid (2.5 g, 10.0 mmol, 1.0 eq.) and DIPEA (3.48 mL, 20.0 mmol, 2.0 eq.) in anhydrous DMF (30 mL) at rt. HATU (3.80 g, 10.0 mmol, 1.0 eq.) was added and the reaction solution was stirred at rt for 2 h. The reaction mixture was concentrated *in vacuo* and the residue was dissolved with EtOAc (50 mL). The organic component was washed with aqueous LiCl (0.5 M, 50 mL), aqueous citric acid (10% *w/v*, 50 mL), sat. aqueous $NaHCO_3$ (50 mL), brine (50 mL), dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (isocratic elution with 50% EtOAc/petroleum ether 40–60 °C) to yield the title compound as a white solid (2.47 g, 94%): $R_f$ 0.20 (EtOAc:petroleum ether 40–60 °C, 1:1); $[\alpha]_D^{25}$ = +3.6 (*c* 1.0, MeOH); m.p. 115–118 °C (from EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm$^{-1}$ 3297 (N-H, m), 1693 (C=O, s), 1650 (C=C, s), 1261 (C-O, s), 862 (s); $^1$H NMR (600 MHz, $CD_3OD$) $\delta_H$ 7.39–7.26 (5H, m, H-5', 6', 7'), 5.81–5.70 (1H, m, H-4), 5.18–5.02 (4H, m, H-3'$^a$, 3'$^b$, 5), 4.15–4.07 (1H, m, H-2), 2.72 (3H, s, H-1''), 2.56–2.46 (1H, m, H-3$^a$), 2.42–2.30 (1H, m, H-3$^b$); $^{13}$C NMR (151 MHz, $CD_3OD$) $\delta_C$ 174.7 (C-1), 158.3 (C-1'), 138.1 (C-4'), 134.7 (C-4), 129.5, 129.0, 128.9 (C-5', 6', 7'), 118.6 (C-5), 67.7 (C-3'), 56.2 (C-2), 37.6 (C-3), 26.3 (C-1''); HRMS *m/z* (ESI$^+$) [Found: 285.1212, $C_{14}H_{18}N_2O_3Na$ requires $[M+Na]^+$ 285.1210]; LRMS *m/z* (ESI$^+$) 263 ($[M+H]^+$, 100%); HPLC retention time: 7.8 min, 97.4% (220 nm), 98.6% (250 nm), 95.8% (254 nm).

**(65) Benzyl (*S*,*E*)-(6-(isoindolin-2'''-yl)-1-(methylamino)-1-oxohex-4-en-2-yl)carbamate hydrochloride**

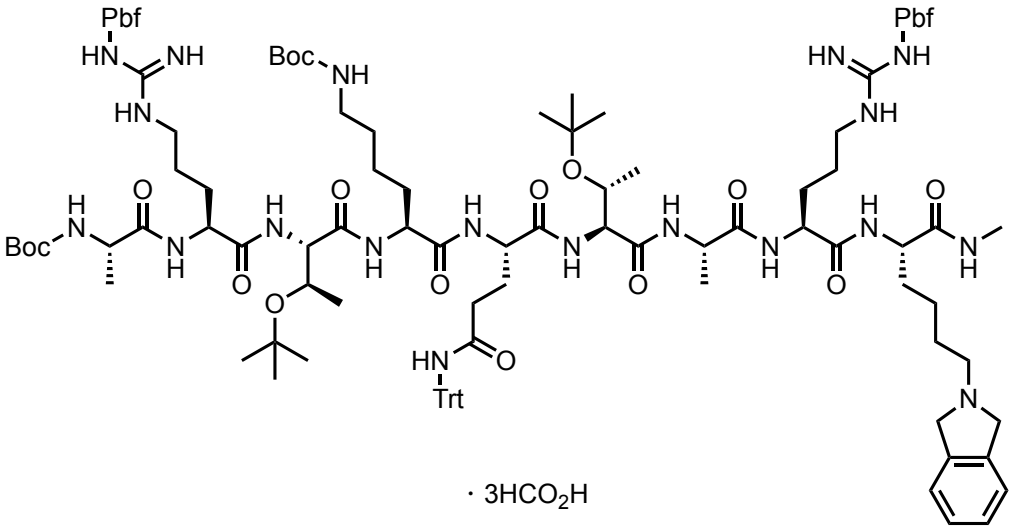



*cis*-Butene-1,4-diol (1.27 mL, 15.4 mmol, 3.00 eq.) and Grubbs 2$^{nd}$ generation catalyst (219 mg, 0.26 mmol, 0.05 eq.) were added to a solution of compound **64** (1.35 g, 5.15 mmol, 1.00 eq.) in anhydrous, degassed $CH_2Cl_2$ (50 mL) under an argon atmosphere at rt. The reaction solution was stirred at rt for 6 h. The reaction solution was filtered through Celite® and the filtrate was concentrated *in vacuo*. The product was purified using flash silica column chromatography (gradient elution with 2–5% MeOH/$CH_2Cl_2$) and then using semi-preparative reverse phase HPLC (gradient method 1, 40 mL min$^{-1}$, Agilent Zorbax PrepHT column [7 µm, 21.2 mm × 250 mm]) to yield the crude metathesis product as a colourless oil (988 mg crude yield, 66%): $R_f$ 0.38 (MeOH:$CH_2Cl_2$, 1:9); LRMS *m/z* (ESI$^+$) 315 ([M+Na]$^+$, 100%). The product was taken forward crude to next step due to the presence of an impurity that was inseparable by both normal and reverse phase column chromatography.

Methanesuflonyl chloride (392 µL, 5.07 mmol, 1.5 eq.) was added dropwise to a solution of the crude metathesis product (988 mg, 9.56 mmol, 1.0 eq.) and $NEt_3$ (944 µL, 6.77 mmol, 2.0 eq.) in $CH_2Cl_2$ (20 mL) at 0 °C. The reaction mixture was warmed to rt and stirred for 1 h. The reaction mixture was diluted with $CHCl_3$ (20 mL) and washed with sat. aqueous $NaHCO_3$ (50 mL), and brine (50 mL). The organic component was dried ($Na_2SO_4$), filtered, and concentrated *in vacuo* to yield the crude mesylate which was reacted

directly in the next step. Isoindoline (755 μL, 6.77 mmol, 2.0 eq.) was added dropwise to a solution of the crude mesylate in anhydrous DMF (20 mL) and the reaction solution was stirred at rt for 1 h. The reaction solution was diluted with aqueous LiCl (0.5 M, 50 mL) and the aqueous phase was extracted with EtOAc (3 × 50 mL). The organic components were combined, washed with brine (50 mL), dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The product was purified using flash silica column chromatography (1–7% (1% $NH_4OH$ in MeOH)/$CH_2Cl_2$). The purified product was dissolved in $Et_2O$ (25 mL) and washed with aqueous HCl (6 M, 3 × 25 mL). The aqueous components were combined and concentrated *in vacuo* to yield the title compound as a dark green solid (797 mg, 36% over 3 steps): $R_f$ 0.30 (free base, 1% $NH_4OH$ in MeOH:$CH_2Cl_2$, 1:9); $[\alpha]_D^{25}$ = +12.9 (*c* 0.1, MeOH); m.p. 96–98 °C (from $CH_2Cl_2$); $\tilde{\nu}_{max}$ (thin film)/$cm^{-1}$ 3271 (N-H, m), 2941 (C-H, w), 1713 (C=O, s), 1663 (C=C, s), 1252 (C-O, s), 750 (s); $^1$H NMR (700 MHz, $CD_3OD$) $\delta_H$ 7.46–7.26 (10H, m, H-4’’’, 5’’’, 6’’’, 7’’’, 5’, 6’, 7’), 6.12–6.00 (1H, dt, *J* 15.5, 6.7, H-5), 5.74 (1H, dt, *J* 15.5, 7.2, H-4), 5.12 (1H, d, *J* 12.5, H-3’$^a$), 5.07 (1H, d, *J* 12.5, H-3’$^b$), 4.79–4.72 (2H, m, H-1’’’$^a$, 3’’’$^a$), 4.53–4.45 (2H, m, H-1’’’$^b$, 3’’’$^b$), 4.21 (1H, dd, *J* 8.6, 5.4, H-2), 3.94 (2H, d, *J* 6.7, H-6), 2.78–2.62 (4H, m, H-1’’, 3$^a$), 2.50–2.41 (1H, m, H-3$^b$); $^{13}$C NMR (176 MHz, $CD_3OD$) $\delta_C$ 174.1 (C-1), 158.3 (C-1’), 139.1 (C-5), 138.1 (C-4’), 134.86, 134.83 (C-3a’’’, 7a’’’), 130.2 (C-4’’’, 7’’’), 129.6, 129.2, 129.0 (C-5’, 6’, 7’), 124.1 (C-5’’’, 6’’’), 123.3 (C-4), 67.9 (C-3’), 59.0, 58.9 (C-1’’’, 3’’’), 57.1 (C-6), 55.8 (C-2), 36.3 (C-3), 26.4 (C-1’’); HRMS *m/z* (ESI$^+$) [Found: 394.2137; $C_{23}H_{28}N_3O_3$ requires [M−HCl+H]$^+$ 394.2125]; LRMS *m/z* (ESI$^+$) 394 ([M−HCl+H]$^+$, 100%); HPLC retention time: 6.5 min, 98.3% (220 nm), 98.2% (250 nm), 96.1% (254 nm).

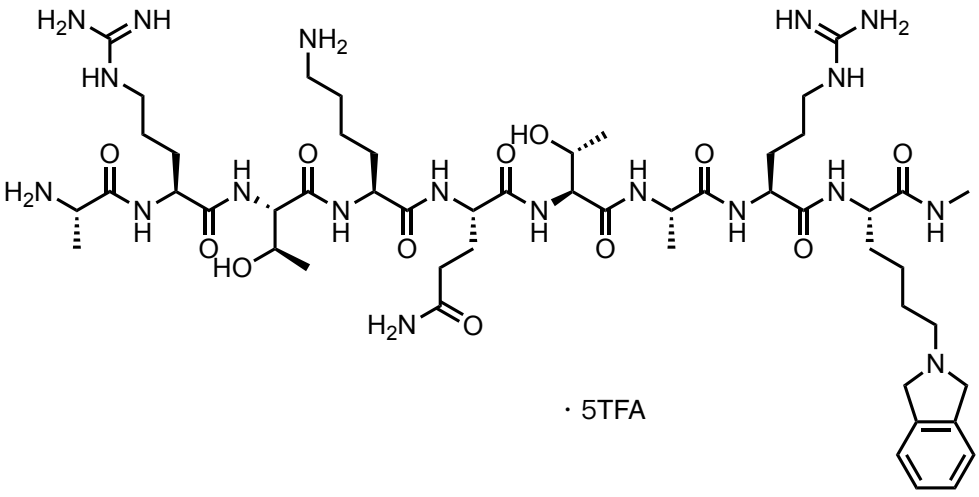
**(66)** ***tert*-Butyl ((4*S*,7*S*,10*S*,13*S*,16*S*,19*S*)-13-((*R*)-1-(*tert*-butoxy)ethyl)-19-((6*S*,9*S*,12*S*)-12-((*R*)-1-(*tert*-butoxy)ethyl)-2,2,6-trimethyl-4,7,10-trioxo-9-(3-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)propyl)-3-oxa-5,8,11-triazatridecan-13-amido)-4-(4-(isoindolin-2-yl)butyl)-10-methyl-3,6,9,12,15,18-hexaoxo-16-(3-oxo-3-(tritylamino)propyl)-7-(3-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)propyl)-2,5,8,11,14,17-hexaazatricosan-23-yl)carbamate triformic acid**




10% Pd/C (11.4 mg, 10.7 μmol, 0.2 eq.) was added to a solution of compound **65** (46.0 mg, 107 μmol, 2.0 eq.) in MeOH (10 mL) under an argon atmosphere at rt. The flask was purged with hydrogen gas *via* 3 successive evacuate/fill cycles and the reaction mixture was stirred vigorously under a hydrogen atmosphere at rt for 6 h. The reaction mixture was filtered through Celite® and the filtrate was concentrated *in vacuo* to yield the crude amine which was used directly in the next step without further purification. HATU (20.3 mg, 53.3 μmol, 1.0 eq.) was added to a solution of the crude amine, compound **48** (111 mg, 53.3 μmol, 1.0 eq.), and DIPEA (46.5 μL, 267 μmol, 5.0 eq.) in anhydrous $CH_2Cl_2$ (5 mL)

at rt. The reaction solution was stirred at rt for 6 h. The reaction solution was concentrated *in vacuo*. The product was purified using semi-preparative reverse phase HPLC (gradient method 4, 20 mL $min^{-1}$, Agilent Zorbax PrepHT column [5 μm, 21.2 mm × 50 mm]) and lyophilised to yield the title compound as a colourless solid (35.1 mg, 28%): HRMS *m/z* ($ESI^+$) [Found: 2233.2301, $C_{115}H_{170}N_{19}O_{22}S_2$ requires $[M-3FA+H]^+$ 2233.2204], [Found: 2255.2088, $C_{115}H_{169}N_{19}O_{22}S_2Na$ requires $[M-3FA+Na]^+$ 2255.2023]; LRMS *m/z* ($ESI^+$) 1117 ($[M-3FA+2H]^{2+}$, 100%), 1128 ($[M-3FA+H+Na]^{2+}$, 66%); HPLC retention time: 11.5 min, 97.3% (220 nm), 97.9% (250 nm), 97.7% (254 nm).

**(67) (*S*)-2-((*S*)-6-amino-2-((2*S*,3*R*)-2-((*S*)-2-((*S*)-2-aminopropanamido)-5-guanidinopentanamido)-3-hydroxybutanamido)hexanamido)-*N*[1]-((4*S*,7*S*,10*S*,13*S*,14*R*)-7-(3-guanidinopropyl)-14-hydroxy-4-(4-(isoindolin-2-yl)butyl)-10-methyl-3,6,9,12-tetraoxo-2,5,8,11-tetraazapentadecan-13-yl)pentanediamide pentatrifluoroacetic acid**




Compound **66** (8.00 mg, 3.37 μmol, 1.0 eq.) was dissolved in a solution of TFA/TIPS/$H_2O$ (38:1:1, 1 mL) and stirred at rt for 4 h. The reaction solution was concentrated *in vacuo* and the residue was redissolved in $H_2O$ (1 mL). The aqueous phase was washed with $CHCl_3$

(3 × 2 mL) and concentrated *in vacuo*. The product was purified by semi-preparative reverse phase HPLC (gradient method 3, 20 mL $min^{-1}$, Agilent Zorbax PrepHT column [5 μm, 21.2 mm × 50 mm]) and lyophilised to yield the title compound as a colourless solid (4.13 mg, 70%): HRMS *m/z* ($ESI^+$) [Found: 392.2438, $C_{52}H_{94}O_{12}N_{19}$ requires $[M-5TFA+3H]^{3+}$ 392.2438]; LC-MS *m/z* ($ESI^+$) 0.8 min, 294.6 ($[M-5TFA+4H]^{4+}$), 392.5 ($[M-5TFA+3H]^{3+}$), 588.0 ($[M-5TFA+2H]^{2+}$), 1174.8 ($[M-5TFA+H]^+$), 97.6% (220 nm), 98.1% (254 nm), 97.3% (280 nm).

## 6.4 General Biology Experimental

**Purified (Milli-Q) Water** was obtained from a Millipore Elix® Reverse Osmosis system and was further purified by a Millipore Milli-Q® Synthesis system with a 0.22 μm filter on the outlet.

**Peptides** were purchased from GenScript at >98% purity. A terminal tyrosine was included for accurate quantification of peptide concentration using absorbance at 280 nm.

**Table E5:** Unbiotinylated peptide (>98% purity, Genscript) amino acid sequences.

| Peptide | Amino Acid Sequence (N–C) |
|---|---|
| $H3_{(1–27)}K18Ac$ | ARTKQTARKSTGGKAPRK(Ac)QLATKAARKY |
| $H3_{(1–27)}K9Me_3$ | ARTKQTARK($Me_3$)STGGKAPRKQLATKAARKY |
| $H3_{(1–27)}K9Me_3K18Ac$ | ARTKQTARK($Me_3$)STGGKAPRK(Ac)QLATKAARKY |
| $H3_{(1–27)}K9Me_0$ | ARTKQTARKSTGGKAPRKQLATKAARKY |
| $H3_{(1–27)}K9C$ | ARTKQTARCSTGGKAPRKQLATKAARKY |
| $H3_{(1-15)}K4Me_3$ | ARTK($Me_3$)QTARKSTTGKA |
| $H3_{(1–8)}$ | ARTKQTAR |
| Ac-$H3_{(1–9)}K9Me_3$ | Ac-ARTKQTARK($Me_3$) |
| $H3_{(1–27)}K9Me_3K27$(FITC) | ARTKQTARK($Me_3$)STGGKAPRKQLATKAARK(FITC) |

**Table E6:** Biotinylated peptide (>98% purity, Genscript) amino acid sequences (b=biotin).

| Peptide | Amino Acid Sequence (N–C) |
|---|---|
| $H3_{(1–27)}K9Me_3$ (B1) | ARTKQTARK($Me_3$)STGGKAPRKQLATKAARK(b)Y |
| $H3_{(1–27)}K18Ac$ (B2) | ARTKQTARKSTGGKAPRK(Ac)QLATKAARK(b)Y |
| $H3_{(1–27)}K9Me_3K18Ac$ (B3) | ARTKQTARK($Me_3$)STGGKAPRK(Ac)QLATKAARK(b)Y |

**Protein/Peptide Absorbance** at 280 nm ($A_{280}$) was measured using a NanoDrop™ Lite (Thermo Fisher Scientific) using the Protein A280 setting.

**Protein/Peptide Concentrations** were calculated using the Beer-Lambert law (Equation 1) with the path length set to 1 cm. H3-mimicking peptide extinction coefficients ($\varepsilon$) were approximated as being equal to that of tyrosine (1490 $M^{-1}$ $cm^{-1}$). Protein molecular weights and molar extinction coefficients were calculated using ProtParam on the ExPASy Bioinformatics Resource Portal.

$$c = \frac{A_{280}}{\varepsilon} \qquad (1)$$

**pH Measurements** were carried out using a PH 550 Benchtop pH Meter Kit (Oakton). Calibration was performed with standard solutions at pH 4.0 (solution of potassium acid phthalate, cetylpyridinium chloride and methyl red), pH 7.0 (solution of potassium dihydrogen phosphate, cetylpyridinium chloride, sodium hydrogen phosphate and methyl orange), and pH 10.0 (solution of sodium carbonate, sodium bicarbonate and thymol blue). The pH probe was stored in 3 M KCl.

**Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)** was carried out using precast Invitrogen NuPAGE™ 4–12% Bis-Tris precast gels (Thermo Fisher Scientific) using a mini-PROTEAN Tetra Electrophoresis System (Bio-Rad).

Protein samples were mixed with 4× Laemmli sample loading buffer in a 3:1 ratio (sample:buffer) and incubated at 100 ºC for 3 min. Denatured samples were loaded onto the gel and the gels were run in 1× MES buffer at 180 V for 50 min. Following electrophoresis, gels were stained with InstantBlue™ Coomassie for 15 min and rinsed with Milli-Q water, then destained multiple times in Milli-Q water overnight.

**Lyophilisation** was carried out using a Christ Alpha 2-4 LSCBasic Laboratory freeze dryer. Samples were flash frozen in liquid nitrogen, before being lyophilised overnight.

**Liquid Chromatography-Mass Spectrometry (LC-MS)** was carried out using a Waters Xevo G2-S QTOF mass spectrometer, equipped with a Waters Acquity UPLC or an Agilent 1260 Infinity II system with a Quaternary LC Pump, DADWR, and InfinityLab LC/MSD. For histone-mimicking peptide characterisation, an ACE® Equivalence™ C18 [3 µm, 2.1 × 50 mm] column (ACE) was used with a constant flow rate of 0.3 mL/min and gradient method of 7 min from $H_2O$:MeCN (99:1) to $H_2O$:MeCN (50:50). Mass spectra were deconvoluted using Waters MassLynx™ software (MaxEnt1 algorithm, 3–5 kDa range, 0.01 Da resolution). Theoretical peptide molecular weights were calculated using ChemDraw 20.0. For the cell permeability assay, a Poroshell 120 EC-C18 [4 µm, 4.6 × 100 mm] column was used with a constant flow rate of 1.0 mL/min and a gradient method of 10 min from $H_2O$:MeCN (95:5) to $H_2O$:MeCN (5:95) with a 5 min hold. All samples were run with 0.1% formic acid in the elution solvents.

**NanoLC-MS/MS** was carried out by Vaishnavi Ravikumar, after enzymatic digestion with endoproteinase LysC (1 µg enzyme/100 µg peptide sample for 2 h at 37°C ), using an UltiMate™ 3000 RSLCnano System (Thermo Scientific) coupled in line with a Q Exactive™ mass spectrometer, equipped with an Easy-Spray source (Thermo Scientific). Peptide samples were trapped onto a C18 PepMac100 precolumn (300 µm × 5 mm, 100 Å,

Thermo Scientific) using Solvent A (0.1% formic acid, HPLC grade water). The peptides were further separated onto an Easy-Spray RSLC C18 column (75 um × 50 cm, Thermo Scientific) using a 15 min linear gradient [15–35% solvent B (0.1% formic acid in MeCN)] at a flow rate 200 nL/min. The raw data were acquired on the mass spectrometer in data-dependent acquisition mode (DDA). Full-scan MS spectra were acquired in the Orbitrap (Scan range 350–1500 *m/z*; resolution: 70000; AGC target: $3\times10^6$; maximum injection time: 50 ms). The 5 most intense peaks were selected for higher-energy collision dissociation (HCD) fragmentation at 30% of normalized collision energy. HCD spectra were acquired using the Orbitrap (resolution: 17500; AGC target: $5\times10^4$; maximum injection time: 120 ms with fixed mass at 180 *m/z*). Charge exclusion was selected for unassigned and $1^+$ ions. The dynamic exclusion was set to 5 s. The data were processed by Marjorie Fournier using the PEAKS software and results were filtered using a False Discovery Rate (FDR) < 1%.

**MALDI-TOF Mass Spectrometry** was carried out using a Bruker Autoflex™ Speed MALDI-TOF mass spectrometer. Sample plates were prepared using the dried droplet method. Peptide samples were diluted in MeCN:$H_2O$ (30:70) (0.1% TFA) to a final concentration of ~1 μM and premixed in a 1:1 ratio with a 10 mg/mL matrix solution of $\alpha$-cyano-4-hydroxycinnamic acid (CHCA) in MeCN:0.1% TFA in $H_2O$ (70:30); before being spotted (0.5 μL) onto an MTP 384 ground steel target plate (Bruker Daltonics). Mass spectra were acquired by manual operation in positive reflectron mode with a laser intensity of 38%. Spectra were processed using Bruker flexAnalysis software.

**Intact Protein Mass Spectrometry** was carried out using a Waters Xevo G2-XS QTOF mass spectrometer, equipped with a Waters Acquity UPLC. A ProSwift™ RP-2H monolithic 4.6 mm × 50 mm column (Thermo Scientific™) was used with a constant flow

rate of 0.4 mL/min and a gradient method of 10 min from $H_2O$:MeCN (95:5) to H2O:MeCN (5:95) with a 5 min hold. All samples were run with 0.1% formic acid in the elution solvents. Mass spectra were deconvoluted using Waters MassLynx™ software (MaxEnt1 algorithm, 20–30 kDa range, 1.0 Da resolution).

**Analytical HPLC** was carried out using a PerkinElmer Flexar system with a Binary LC Pump, and UV/VIS LC Detector set at 220 nm. For purification of peptides a Dionex Acclaim® 120 C18 [5 µm, 120 Å, 4.6 mm × 150 mm] reverse phase column was used with a constant flow rate of 1.0 mL/min and a gradient method of 25 min from $H_2O$:MeCN (95:5) to $H_2O$:MeCN (35:65). All samples were run with 0.1% TFA in the elution solvents. Peaks corresponding to monoalkylated peptides were collected and rerun with an identical gradient for purity determination. HPLC data were processed using Chromera 3.4.4 software.

## 6.5 Histone-mimicking Peptide Alkylations

### 6.5.1 Alkylation Buffer

**Table E7:** Alkylation Buffer (pH 7.8) composition.

| Reagent | Mass required per 100 mL (g) | Concentration (M) |
|---|---|---|
| HEPES | 6.46 | 0.27 |
| HEPES sodium salt | 18.98 | 0.73 |
| D/L-Methionine | 0.15 | 0.01 |
| Guanidine hydrochloride | 38.21 | 4.00 |

HEPES, HEPES sodium salt, and D/L-methionine were dissolved with Milli-Q® water (50 mL) in a 100 mL Duran bottle. Guanidine hydrochloride was added portionwise over 5 min with stirring. After all solids had dissolved, the pH was adjusted to pH 7.8 as required. The resulting solution was transferred to a 100 mL measuring cylinder and topped

up to 100 mL with Milli-Q® water. The solution was then filtered through 0.20 μm nylon filter paper under vacuum and stored at 4 °C. Alkylation buffer was degassed for 30 min prior to use.

### 6.5.2 Test Reaction

Dithiothreitol (1 M, 0.2 mL) was added to a solution of *N*-acetyl-L-cysteine methyl ester (50 mg, 0.28 mmol, 1.0 eq.) in alkylation buffer (9.8 mL). The solution was stirred and heated at 37 °C for 1 h. The reaction solution was cooled to rt and (2-bromoethyl)-diethylammonium bromide (369 mg, 1.41 mmol, 5.0 eq.) was added. The reaction solution was stirred at rt for 5 h, before being quenched with β-mercaptoethanol (100 μL). The reaction solution was basified with sat. aqueous $NaHCO_3$ (5 mL) and the aqueous phase was extracted with EtOAc (3 × 10 mL). The organic components were combined, dried ($Na_2SO_4$), filtered, and concentrated *in vacuo*. The products were analysed by $^1$H NMR and LC-MS.

### 6.5.3 Peptide Alkylation Reactions

Alkylation reactions were carried out using adapted versions of the protocols reported by Simon *et al.*[5] (2012) and Pieters *et al*.[6] (2019). $H3_{(1\text{-}27)}$K9C peptides (5 mg) were dissolved in 1.5 mL Eppendorf tubes with alkylation buffer (980 μL) and mixed by gentle inversion. Dithiothreitol (1 M, 20 μL) was added, and the solutions were incubated at 37 °C with shaking at 350 rpm for 1 h. Alkylating agents were then added directly to the reduced peptides in the specified molar ratios (Table E8). For compounds **P10**, **P11**, **P13**, and **P14**; reaction mixtures were supplemented with DMSO (20 μL) to improve the solubility of the reagents. Reaction mixtures were flicked gently to dissolve the alkylating agents, covered with aluminium foil, and incubated at the specified temperatures (Table E8) with shaking

at 350 rpm for 2.5 h. Dithiothreitol (1 M, 10 μL) was added and the reactions were allowed to proceed for a further 2.5 h. Reactions were quenched by the addition of β-mercaptoethanol (50 μL) and incubated at rt for 30 min. Peptides were desalted using Pierce™ polyacrylamide desalting columns (1.8K MWCO) (Thermo Fisher Scientific), and eluted with an aqueous solution of β-mercaptoethanol (2 mM). Samples (10 μL) from each fraction were diluted in 90 μL of β-mercaptoethanol (2 mM) and analysed using LC-MS (Table E9; Appendix 3, Figure A3.1) and nanoLC-MS/MS (Appendix 3, Figure A3.2) with Lys-C digest. Peptide-containing fractions were combined, purified using analytical HPLC (Table E9; Appendix 3, Figure A3.3), lyophilised, and stored at −80 °C. After purification peptides were characterised using MALDI-TOF MS (Appendix 3, Figure A3.4).

**Table E8:** Optimised alkylation reaction conditions for peptides **P1**-**P16**.

| Peptide | Alkylating Reagent | Temperature | Alkylating Agent Equivalents |
|---|---|---|---|
| **P1** | (2-bromoethyl)-trimethylammonium bromide | 50 °C | 300 |
| **P2** | (2-bromoethyl)-dimethylammonium bromide | rt | 50 |
| **P3** | (2-bromoethyl)-diethylammonium bromide | rt | 50 |
| **P4** | *N*-(2-bromoethyl)-*N*-isopropylpropan-2-amine hydrobromide **(5)** | rt | 50 |
| **P5** | 1-(2-bromoethyl)pyrrolidine hydrobromide **(6)** | rt | 50 |
| **P6** | 1-(2-bromoethyl)piperidine hydrobromide **(7)** | rt | 50 |
| **P7** | 4-(2-bromoethyl)morpholine hydrobromide **(8)** | rt | 50 |

**Table E8 (cont.):** Optimised alkylation reaction conditions for peptides **P1**-**P16**.

| Peptide | Alkylating Reagent | Temperature | Alkylating Agent Equivalents |
|---|---|---|---|
| **P8** | 1-(2-bromoethyl)piperazine ditrifluoroacetic acid **(12)** | rt | 50 |
| **P9** | 1-(2-bromoethyl)-4-methylpiperazine dihydrobromide **(9)** | rt | 50 |
| **P10** | 2-(2-bromoethyl)isoindoline hydrobromide **(3)** | rt | 50 |
| **P11** | 2-(2-bromoethyl)-1,2,3,4-tetrahydroisoquinoline hydrobromide **(4)** | rt | 50 |
| **P12** | 1-(2-bromoethyl)pyridinium bromide **(13)** | 50 °C | 300 |
| **P13** | 2-(2-bromoethyl)pyridine **(15)** | 50 °C | 300 |
| **P14** | 1-(2-bromoethyl)-1*H*-pyrazole **(14)** | 50 °C | 300 |
| **P15** | 7a-(chloromethyl)hexahydro-1*H*-pyrrolizine hydrochloride **(21)** | 50 ºC | 300 |
| **P16** | (2*R*,3*R*,4*S*,5*R*,6*R*)-2-(2-bromoethoxy)-6-(hydroxymethyl)tetrahydro-2*H*-pyran-3,4,5-triol **(23)** | 50 ºC | 300 |

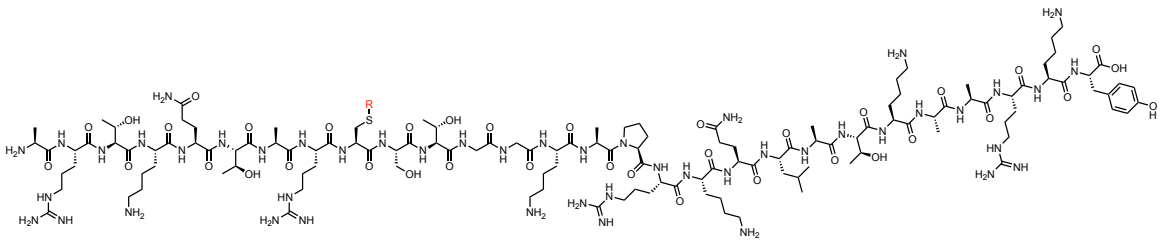
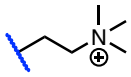
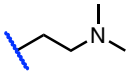
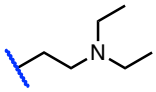
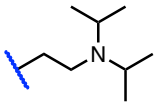
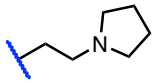
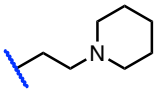
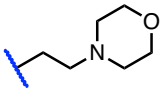
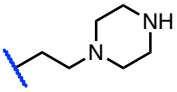
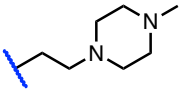
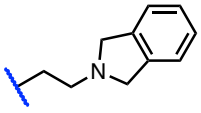
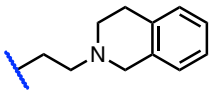
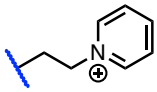
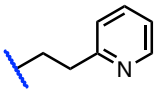
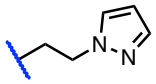
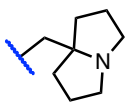
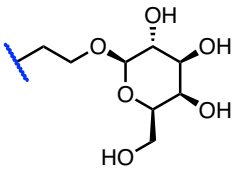


**Figure E1:** Structure of the H3K9Me$_3$-mimicking peptides (R = Modification).

**Table E9:** Peptide retention times, purities, and molecular weights after deconvolution ($ESI^+$). Theoretical molecular weights were calculated in ChemDraw (Appendix 3, Figures A3.1–3.4).

| Peptide | Modification (R) | Retention Time (min) | HPLC Purity (%) | Theoretical MW (Da) | Observed MW (Da) |
|---|---|---|---|---|---|
| **P1** |  | 9.2 | 98.3 | 3134.73 | 3134.76 |
| **P2** |  | 9.8 | >99.9 | 3119.70 | 3119.93 |
| **P3** |  | 9.5 | 97.4 | 3147.75 | 3147.90 |
| **P4** |  | 10.9 | 95.6 | 3175.81 | 3175.86 |
| **P5** |  | 9.5 | 99.8 | 3145.74 | 3145.82 |
| **P6** |  | 9.8 | 99.2 | 3159.76 | 3159.89 |
| **P7** |  | 9.9 | >99.9 | 3161.74 | 3161.93 |
| **P8** |  | 9.2 | 98.4 | 3160.75 | 3161.00 |
| **P9** |  | 8.7 | 97.4 | 3174.78 | 3175.10 |
| **P10** |  | 9.8 | 96.5 | 3193.78 | 3193.75 |
| **P11** |  | 10.1 | 99.2 | 3207.81 | 3207.89 |
| **P12** |  | 10.4 | 96.0 | 3154.72 | 3154.13 |
| **P13** |  | 9.1 | 97.7 | 3153.72 | 3154.03 |
| **P14** |  | 9.9 | 96.3 | 3142.69 | 3142.94 |
| **P15** |  | 9.9 | >99.9 | 3171.78 | 3171.60 |
| **P16** |  | 10.1 | >99.9 | 3254.77 | 3254.90 |

## 6.6 Protein Production

### 6.6.1 Constructs, Plasmids, and Cell Lines

His-tagged human TRIM24 PHD-BRD WT residues N825–E1007 (Appendix 1, Table A1.1) was expressed from the plasmid pNIC28-Bsa4 (Appendix 1, Figure A1.1), provided by the Structural Genomics Consortium, in competent *Escherichia coli* BL21-Gold (DE3) (Agilent). $His_6$-SUMO-tagged TRIM24 PHD-BRD WT residues N825–E1007 (Appendix 1, Table A1.1) was expressed from the plasmid pETM11-SUMO3GFP (Appendix 1, Figure A1.3), provided by the Structural Genomics Consortium, in competent *E. coli* Mach1-$T1^R$ (Thermo Fisher Scientific). GST-tagged human TRIM24 PHD-BRD WT residues N825–E1007 (Appendix 1, Table A1.1) was expressed from the plasmid pGTvL1-SGC (Appendix 1, Figure A1.2) and purified by Dr. Angelina Sekirnik. $His_6$-TRIM24 PHD-BRD W828A mutant residues N825–E1007 (Appendix 1, Table A1.1) was expressed from the plasmid pNIC28-Bsa4 (Appendix 1, Figure A1.1) and purified by Dr. Angelina Sekirnik.[1] Protein molecular weights and molar extinction coefficients were calculated using ProtParam on the ExPASy Bioinformatics Resource Portal.

### 6.6.2 Bacterial Transformation

All media and equipment were sterilised by autoclaving at 121 °C for 20 min prior to use, and standard aseptic technique was used throughout. Competent *E. coli* (40 µL) were thawed on ice and the TRIM24 plasmid (1 µL) was added. The tube was flicked 10 times and incubated on ice for 30 min. The bacteria were heat shocked in a 42 °C water bath for 45 s and then returned to ice for 5 min. Super optimal broth with catabolite repression (SOC) outgrowth medium (100 µL) was added and the bacteria were incubated at 37 °C for 90 min. Autoclaved 2× YT medium (Table E10) plates containing bacteriological agar (15 g/L) and kanamycin (50 µg/mL) were streaked with bacteria (140 µL) under aseptic conditions. Plates were inverted, incubated overnight at 37 °C, and then stored at 4 °C.

**Table E10:** 2× YT medium composition (1 L).

| Media | Reagent | Amount required (g/L) |
|---|---|---|
| 2× YT | Tryptone | 16.0 |
| | Yeast extract | 10.0 |
| | NaCl | 5.0 |

### 6.6.3 Protein Expression

Autoclaved 2× YT media (100 mL), containing kanamycin (50 µg/mL), was inoculated with one colony from an agar plate; and incubated overnight at 37 °C with shaking at 250 rpm. 6 × 2 L PYREX® narrow-mouth baffled Tunair™ flasks, containing autoclaved 2× YT media (1 L) and kanamycin (50 µg/mL), were inoculated with 1% by volume (10 mL) of the overnight starter culture; and incubated at 37 °C with shaking at 180 rpm until an $OD_{600}$ of 0.6–0.8 was reached. Optical density measurements were carried out at regular intervals using a Novaspec® II spectrophotometer at 600 nm, with 2× YT media (supplemented with 50 µg/mL kanamycin) used as the reference. After reaching the desired optical density, cultures were equilibrated to 18 °C, supplemented with aqueous $ZnCl_2$ (0.1 mM), induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) (0.25 mM), and incubated overnight at 18 °C with shaking at 180 rpm. 20–24 h after induction, bacteria were harvested by centrifugation using an Avanti™ JXN-26 centrifuge (6227 ×*g*, 4 °C for 20 min). Cell pellets were transferred to 50 mL Falcon tubes and stored at −80 °C.

### 6.6.4 $His_6$-TRIM24 Purification

Bacterial pellets were thawed, resuspended in extraction buffer [1 Sigma*FAST*™ protease inhibitor cocktail tablet EDTA free (Sigma Aldrich) dissolved in 100 mL binding buffer per 20 g pellet], and sonicated (35% amplitude) on ice for 15 min *via* a 5 s burst/10 s pause

cycle (5 min total sonication time). The lysate was supplemented with polyethylenimine (PEI) (0.15% *w/v*) and incubated on ice for 15 min. The cell lysate was clarified by centrifugation (48298 ×*g*, 4 °C for 30 min) and the supernatant was filtered through a 0.45 μm syringe filter.

The protein was purified using nickel immobilised metal affinity chromatography (Ni-IMAC) at 4 °C on an ÄKTA pure™ system (Cytiva) using a HisTrap™ FF 5 mL column (Cytiva) preloaded with $Ni^{2+}$. The column was washed sequentially with Milli-Q® water (25 mL, 5 CV) and elution buffer (25 mL, 5 CV); and equilibrated with binding buffer (50 mL, 10 CV). The cell lysate was loaded onto the column with a 1.0 mL/min flow rate and purified protein was eluted with a 1.5 mL/min flow rate in 5 mL fractions using step gradient method 1 (Table E12; Appendix 1, Figure A1.4).

The protein was further purified using size exclusion chromatography (SEC) at 4 °C on an ÄKTA pure™ system (Cytiva) using an XK 16/600 gel filtration column (Cytiva) packed with 120 mL Superdex 75 resin. The column was washed with degassed Milli-Q® water (150 mL) and equilibrated with degassed gel filtration buffer (150 mL). IMAC fractions containing purified protein were combined, concentrated (10K MWCO spin concentrator, ~4 mL), and clarified by centrifugation (15000 rpm, 4 °C for 10 min). Concentrated protein (2 mL) was loaded onto the column and the protein was eluted in 5 mL fractions with degassed gel filtration buffer (150 mL) using a flow rate of 1 mL/min (Appendix 1, Figure A1.5). Fractions containing the purified protein (determined from SDS-PAGE and SEC chromatogram) were combined and concentrated. Purified protein was aliquoted, frozen in liquid nitrogen, and stored at −80 °C. Proteins were characterised by intact protein MS and purity was evaluated using SDS-PAGE (Appendix 1, Figures A1.6–A1.8).

**Table E11:** IMAC and SEC buffer composition for $His_6$-TRIM24 purification. All buffers were filtered through 0.2 µm filter paper under vacuum prior to use. Gel filtration buffer was degassed for 30 min prior to use. Buffers were stored at 4 ºC.

| Buffer (pH 7.6) | [HEPES] (mM) | [NaCl] (mM) | [Imidazole] (mM) | [Glycerol] (% *v*/*v*) |
|---|---|---|---|---|
| Binding | 50 | 500 | 5 | 10 |
| Elution | 50 | 500 | 500 | 10 |
| Gel Filtration | 50 | 500 | – | – |

**Table E12:** IMAC step gradient elution method 1 for the purification of $His_6$-TRIM24.

| Eluent | % Elution buffer (in binding buffer) | Volume (mL) | [Imidazole] (mM) |
|---|---|---|---|
| 1 | 0 | 50 (10 CV) | 5 |
| 2 | 5 | 50 (10 CV) | 30 |
| 3 | 11 | 50 (10 CV) | 60 |
| 4 | 60 | 20 (4 CV) | 300 |
| 5 | 100 | 10 (2 CV) | 500 |

### 6.6.5 $His_6$SUMO-TRIM24 Purification and Tag Cleavage

Pellets were thawed, resuspended in binding buffer (200 mL), and supplemented with benzonase nuclease (40 µL) and protease inhibitor cocktail set III (200 µL). Bacteria were sonicated (35% amplitude) on ice for 15 min *via* a 5 s burst/10 s pause cycle (5 min total sonication time) and the lysate was clarified by centrifugation (53248 ×*g*, 4 °C for 1 h).

The protein was purified using manual Ni-IMAC at 4 ºC using Ni Sepharose™ 6 FF resin (Cytiva). The supernatant was added to the resin (2 mL, pre-equilibrated with 40 mL binding buffer) and incubated for 2 h before loading onto a gravity flow column. The flowthrough was collected and reapplied to the column. The flowthrough was recollected and purified protein was eluted using a step gradient method 2 (Table E14; Appendix 1,

Figure A1.9). Fractions containing purified protein were combined, supplemented with $His_6$-SENP1 protease (1:300, protease:protein), and dialysed into SUMO cleavage buffer overnight using a Slide-A-Lyzer™ G3 dialysis cassette (3.5K MWCO, Thermo Fisher Scientific). Fractions containing the cleaved protein were combined and repurified using Ni-IMAC as previously described using step gradient method 3 (Table E15; Appendix 1, Figure A1.9).

**Table E13:** IMAC and SEC buffer composition for $His_6$-SUMO-TRIM24 purification and SUMO tag cleavage. All buffers were filtered through 0.2 μm filter paper under vacuum before use. Gel filtration buffer was degassed for 30 min before use. Buffers were stored at 4 ºC.

| Buffer (pH 8.0) | [Tris base] (mM) | [NaCl] (mM) | [Imidazole] (mM) | [DTT] (mM) |
|---|---|---|---|---|
| Binding | 20 | 500 | 20 | – |
| Wash I | 20 | 500 | 45 | – |
| Wash II | 20 | 500 | 60 | – |
| Elution I | 20 | 500 | 300 | – |
| Elution II | 20 | 500 | 500 | – |
| SUMO cleavage | 20 | 150 | – | 1 |
| Gel Filtration | 10 | 100 | – | 5 |

**Table E14:** IMAC step gradient elution method 2 for the purification of $His_6$-SUMO-TRIM24.

| Eluent | Buffer | Volume (mL) | [Imidazole] (mM) |
|---|---|---|---|
| 1 | Binding | 40 (20 CV) | 20 |
| 2 | Wash I | 20 (10 CV) | 45 |
| 3 | Wash II | 20 (10 CV) | 60 |
| 4 | Elution I | 7 (3.5 CV) | 300 |
| 5 | Elution I | 7 (3.5 CV) | 300 |
| 6 | Elution I | 7 (3.5 CV) | 300 |
| 7 | Elution II | 7 (3.5 CV) | 500 |

**Table E15:** IMAC step gradient elution method 3 for the purification of untagged TRIM24.

| Eluent | Buffer | Volume (mL) | [Imidazole] (mM) |
|---|---|---|---|
| 1 | Binding | 7 (3.5 CV) | 20 |
| 2 | Binding | 7 (3.5 CV) | 20 |
| 3 | Wash I | 7 (3.5 CV) | 45 |
| 4 | Wash II | 7 (3.5 CV) | 60 |
| 5 | Elution I | 7 (3.5 CV) | 300 |
| 6 | Elution II | 7 (3.5 CV) | 500 |

The protein was further purified by size exclusion chromatography (SEC) at 4 °C on an ÄKTA pure™ system (Cytiva) using an XK 16/600 gel filtration column (Cytiva) packed with 120 mL Superdex 75 resin. The column was washed with degassed Milli-Q® water (150 mL) and equilibrated with degassed gel filtration buffer (150 mL). IMAC fractions containing the purified cleaved protein were combined, concentrated (10K MWCO spin concentrator, ~5 mL), and clarified by centrifugation (15000 rpm, 4 °C for 10 min). Concentrated protein (2–3 mL) was loaded onto the column and the protein was eluted in 1 mL fractions with degassed gel filtration buffer (150 mL) using a flow rate of 1 mL/min (Appendix 1, Figure A1.10). Fractions containing the purified protein (determined from SDS-PAGE and SEC chromatogram) were combined and concentrated. Purified protein was aliquoted, frozen in liquid nitrogen, and stored at −80 °C. Proteins were characterised by intact protein MS and purity was evaluated using SDS-PAGE (Appendix 1, Figure A1.11–A1.13)

## 6.7 Protein Crystallography and X-ray Diffraction

### 6.7.1 TRIM24 Crystallisation in the *apo* State

Purified untagged TRIM24 was concentrated to 21.8 mg/mL and crystallisation was carried out using the sitting-drop vapour-diffusion technique. Crystallisation plates were prepared

using a mosquito® HTS Nanolitre Liquid Handler with a drop size of 150 nL and drop ratios of 2:1, 1:1, and 1:2 (protein:reservoir). Crystals were grown at 20 °C using a fine screen based on previous crystallisation conditions.[2]

### 6.7.2 TRIM24 Co-crystallisation

Compounds (40 mM DMSO stock) were mixed with purified untagged TRIM24 (21.8 mg/mL) in a 1:15 ratio (compound : protein). Crystallisation was carried out using the sitting-drop vapour-diffusion technique. Protein compound mixtures were incubated on ice for 1 h, before being clarified by centrifugation (15000 rpm, 4 °C for 10 min). Crystallisation was carried out using the sitting-drop vapour-diffusion technique with a reservoir solution of 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.3), and 5% *v/v* PEG-400. Crystallisation plates were prepared using a mosquito® HTS Nanolitre Liquid Handler with a drop size of 150 nL and drop ratios of 2:1, 1:1, and 1:2 (protein:reservoir) with crystal seeding (20 µL) carried out as required. Crystals were grown for 2 months at 20 °C.

### 6.7.3 TRIM24 Crystal Soaking

Purified untagged TRIM24 was concentrated to 21.8 mg/mL and crystallisation was carried out using the sitting-drop vapour-diffusion technique with a reservoir solution of 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.3), and 5% *v/v* PEG-400. Crystallisation plates were prepared using a mosquito® HTS Nanolitre Liquid Handler with a drop size of 150 nL and drop ratios of 2:1, 1:1, and 1:2 (protein:reservoir). Crystals were left to grow for 7 days at 20 °C and then soaked for 48 h in reservoir solution containing ~ 100 µM compound of interest (0.5 µL).

### 6.7.4 Diffraction Data Collection and Analysis

Crystals of interest were mounted and vitrified in liquid nitrogen. Diffraction data were collected at Diamond Light Source (DLS) beamline I03. Diffraction data were processed at DLS. Crystal structures were solved by molecular replacement using the published *apo* crystal structure (PDB ID: 3O33) and refined using the REFMAC5 and Coot programmes within the CCP4 Cloud software.

## 6.8 AlphaScreen®

### 6.8.1 General AlphaScreen® Experimental

**Table E16:** AlphaScreen® buffer (1×, pH 7.6) composition.

| Reagent | Mass required per 50 mL (mg) | Concentration |
|---|---|---|
| HEPES | 298 | 25 mM |
| NaCl | 292 | 100 mM |
| CHAPS | 25 | 0.05% *w/v* |
| BSA | 50 | 0.1% *w/v* |

AlphaScreen® buffer (Table E16) was prepared as a 5× solution in Milli-Q® water without bovine serum albumin (BSA) and stored at 4 °C until required. Prior to carrying out the assay, an aliquot of buffer was adjusted to pH 7.6 using aqueous NaOH (1 M), supplemented with BSA (0.1% *w/v*), diluted to a 1× solution, and filtered through a 0.22 μm syringe filter.

An electronic multi-channel pipette (Thermo Fisher) was used to dispense the protein, peptides, and compounds into an AlphaPlate™-384SW (Perkin Elmer) microplate. A PHERAstar FS (BMG LABTECH) microplate reader was used to read the plate (Excitation: 680 nm, 0.30 s; Emission: 570 nm, 0.60 s; Temperature: 25 °C; Gain: 3000).

$Ni^{2+}$ chelate/Anti-GST acceptor and streptavidin donor beads (Perkin Elmer) were used for the signal response and dose response assays. $Ni^{2+}$ chelate acceptor and glutathione donor beads (Perkin Elmer) were used for the ternary complex assay. Incubation steps involve sealing the plate with an aluminium plate seal and shaking on a plate oscillator at 300 rpm in the dark for a specified duration at rt. Dose response curves were obtained in triplicate by fitting the data to a four-parameter equation (Equation 2) in GraphPad Prism.

$$Emission = Bottom + \frac{Top - Bottom}{1 + 10^{[(Hill\ slope)(log\ IC_{50} - log\ x)]}} \qquad (2)$$

$$x = [compound]$$

### 6.8.2 Signal-response Assay

Biotinylated peptides (5 μL) were dispensed into wells as 1:2 serial dilutions in AlphaScreen® buffer. $His_6$-TRIM24 or GST-TRIM24 (7 μL)was dispensed into wells to give the desired concentrations and plates were incubated for 1 h. Assay beads (8 μL, 10 μg/mL FAC) were subsequently added under dark conditions to attain a final assay volume of 20 μL. Plates were incubated in the plate reader at 25 °C for 1 h before being read (Appendix 4, Figures A4.1 and A4.2).

### 6.8.3 Dose-response Assay

Compounds (5 μL) were dispensed into wells as 1:2 serial dilutions in AlphaScreen® buffer supplemented with DMSO (1% FAC). $His_6$-TRIM24–biotinylated peptide mixes (7 μL, 1% DMSO, Table E17) were dispensed into wells and plates were incubated for 1 h. Assay beads (8 μL, 10 μg/mL FAC, 1% DMSO) were subsequently added under dark conditions to attain a final assay volume of 20 μL. Plates were incubated in the plate reader at 25 °C for 1 h before being read.

**Table E17:** FACs for the TRIM24-biotinylated peptide mixes determined from signal response assays.

| Biotinylated Peptide | Biotinylated Peptide FAC (nM) | TRIM24 FAC (nM) |
|---|---|---|
| $H3_{(1–27)}K9Me_3$ (B1) | 30 | 50 |
| $H3_{(1–27)}K18Ac$ (B2) | 8 | 25 |
| $H3_{(1–27)}K9Me_3K18Ac$ (B3) | 8 | 25 |

### 6.8.4 Ternary Complex Detection Assay

Compounds (5 μL) were dispensed into wells as 1:2 serial dilutions in AlphaScreen® buffer supplemented with DMSO (1% FAC). TRIM24 protein mixes (7 μL, 25 nM $His_6$-TRIM24, 25 nM GST-TRIM24, 1% DMSO) were dispensed into wells and plates were incubated for 1 h. Assay beads (8 μL, 10 μg/mL FAC, 1% DMSO) were subsequently added under dark conditions to attain a final assay volume of 20 μL. Plates were incubated in the plate reader at 25 °C for 1 h before being read. An $His_6$-GST construct was used as a technical positive control. 7 μL of buffer, supplemented with 1% DMSO, was used in place of the TRIM24 protein mix for the technical positive control.

### 6.8.5 TruHits Assay

Compounds (5 μL) were dispensed into wells as 1:2 serial dilutions in AlphaScreen® buffer supplemented with DMSO (1% FAC). TruHits assay beads (15 μL, 4.7 μg/mL FAC, 1% DMSO) were subsequently added under dark conditions to attain a final assay volume of 20 μL. Plates were incubated in the plate reader at 25 °C for 30 min before being read.

## 6.9 Fluorescence Anisotropy

### 6.9.1 General Fluorescence Anisotropy Experimental

**Table E18:** Fluorescence anisotropy buffer (1×, pH 6.5) composition.

| Reagent | Mass/volume required per L | Concentration (mM) |
|---|---|---|
| $NH_4OAc$ | 3.85 g | 50 |
| Aniline | 913 µL | 10 |

Fluorescence Anisotropy buffer (Table E18) was prepared as a 5× stock in Milli-Q® water and stored at 4 °C until required. Prior to carrying out the assay, an aliquot of buffer was diluted to a 1× solution with Milli-Q® water. The buffer was adjusted to pH 6.5 using aqueous AcOH, before being filtered through a 0.22 µm syringe filter.

An electronic multi-channel pipette (Thermo Fisher Scientific) was used to dispense the protein, peptides, and compounds into a black polystyrene 384 well, flat bottom, low flange, non-binding microplate (Corning). A PHERAstar FS (BMG LABTECH) microplate reader was used to read the plate (Excitation: 485 nm, 0.20 s; Emission A (∥): 520 nm, Emission B (⊥): 520 nm; Temperature: 25 °C). A well containing the free FITC-tracer (100 nM FAC) in buffer was used for automatic gain and focus adjustment.

For the tracer titration assay fluorescence intensities were background subtracted (buffer only) and anisotropy values were calculated for each well (Equation 3). For the direct titration assay fluorescence intensities were background subtracted (buffer + protein) and anisotropy values were calculated for each well. Average anisotropy was plotted against log[protein] using GraphPad Prism; and a logistic curve (Equation 4) was fitted to give $r_{max}$ and $r_{min}$. Anisotropy was converted to fraction ligand bound (Equation 5) and multiplied by the tracer concentration. The data were then plotted against the protein

concentration and fitted to an equation (Equation 6) in GraphPad Prism in order to obtain $K_d$ values. For the competition assay fluorescence intensities were background subtracted (buffer + protein + anchor peptide). Average anisotropy was plotted against the log[anchor peptide] and the data was fitted to a four-parameter equation (Equation 7) in GraphPad Prism in order to obtain $IC_{50}$ values. For the dynamic ligation screen fluorescence intensities were background subtracted (buffer + protein + anchor peptide + aldehyde fragment). Duplicate anisotropy values were averaged for each sample and normalised relative to the negative buffer control.

$$r = \frac{(I_{\parallel} - I_{\perp})}{(I_{\parallel} + 2I_{\perp})} \qquad (3)$$

$$r = r_{min} + \frac{(r_{max} - r_{min})}{(1 + 10^{((Hill\ slope)(A - log x)}} \qquad (4)$$

$$x = [protein]$$

$$F_b = \frac{(r - r_{min})}{\lambda(r_{max} - r) + r - r_{min}} \qquad (5)$$

$$y = F_b \times L^* = \frac{((K_d + x + L^*) - \sqrt{((K_d + x + L^*)^2 - 4xL^*)})}{2} \qquad (6)$$

$$x = [protein]$$

$$y = bottom + \frac{top - bottom}{1 + 10^{[(\text{Hill slope})(\log IC_{50} - \log x)]}} \qquad (7)$$

$$x = [anchor\ peptide]$$

($r$ = anisotropy, $I_{\parallel}$ = parallel intensity, $I_{\perp}$ = perpendicular intensity, $A$ = constant, $F_b$ = fraction ligand bound, $\lambda = \frac{I_{bound}}{I_{unbound}} = 1$, $K_d$ = dissociation constant, $L^*$ = fluorescent tracer concentration)

### 6.9.2 Stability Tests

A solution of a $H3_{(1–8)}$ peptide (100 μM) and benzaldehyde (500 μM) in fluorescence anisotropy buffer, containing 1% DMSO, was shaken at 600 rpm for 24 h. Samples (10 μL) were taken at 0, 1, 4, and 24 h time points; and diluted in Milli-Q® water (90 μL) for LC-MS analysis (Appendix 5, Figure A5.1). A solution of $His_6$-TRIM24 (10 μM) and benzaldehyde (500 μM) in fluorescence anisotropy buffer, containing 1% DMSO, was shaken at 600 rpm for 24 h. Samples (10 μL) were taken at 0, 1, 4, and 24 h time points; and diluted in Milli-Q® water (90 μL) for LC-MS analysis (Appendix 5, Figure A5.2).

### 6.9.3 Tracer Titration Assay

Serial dilutions of the $H3_{(1-27)}K9Me_3K27$(FITC) tracer and free FITC fluorophore in fluorescence anisotropy buffer (40 μL) were prepared in triplicate. Background wells, containing buffer (40 μL) were prepared in triplicate. Plates were covered with an aluminium plate seal and incubated at rt for 30 mins before being read (Appendix 5, Figure A5.3)

### 6.9.4 Direct Titration Assay

Serial dilutions of $His_6$-TRIM24 in fluorescence anisotropy buffer (20 μL) were prepared in triplicate. A solution of the FITC-tracer in buffer (20 μL, 100 nM FAC) was added to each well to give a final assay volume of 40 μL. Background wells, containing serial dilutions of $His_6$-TRIM24 in buffer (40 uL) were prepared in triplicate. Plates were covered with an aluminium plate seal and incubated at rt for 2 h before being read.

### 6.9.5 Competition Assay

Serial dilutions anchor peptide Ac-c001 (synthesised by Yue Li, University of Birmingham; Appendix 5, A5.7) in fluorescence anisotropy buffer (20 uL) were prepared

in triplicate. $His_6$-TRIM24 (20 µL, 500 nM FAC) and the FITC tracer (20 µL, 100 nM FAC) in buffer were added to each well to give a final assay volume of 60 µL. Background wells containing buffer (20 µL), $His_6$-TRIM24 (20 µL, 500 nM FAC) and serial dilutions of the anchor peptide in buffer (20 µL) were prepared in triplicate. Plates were covered with an aluminium plate seal and incubated at rt for 2 h before being read.

### 6.9.6 Acyl Hydrazone Dynamic Ligation Screen

Screening was performed in duplicate using an adapted version of a previously reported procedure.[134,135] Aldehyde libraries were supplied by the lab of Prof. Andrew Wilson (University of Birmingham) as DMSO stocks and diluted to the desired stock concentration with fluorescence anisotropy buffer. Each aldehyde (10 µL, 375 µM) was added to four wells in a black 384-well plate (2 active wells, 2 blank wells). Anchor peptide Ac-c001 (20 µL, 62.5 µM FAC) and $His_6$-TRIM24 (20 µL, 500 nM FAC) were added to each well. Plates were sealed and incubated at rt for 24 h. The FITC tracer (20 µL, 100 nM FAC) was then added to experiment wells and buffer (20 µL) was added to blank wells to give a final assay volume of 60 µL. Plates were incubated at rt for 2 h before being read. Control wells (60 µL final assay volume) were also prepared (Table E19) and were supplemented with DMSO to the same concentration as the active assay wells.

**Table E19:** Control well compositions. All controls contained $His_6$-TRIM24 (500 nM FAC) and the tracer (100 nM FAC). FACs are shown in parentheses.

| Control | Composition |
|---|---|
| (−) Buffer | – |
| (–) Ac-c001 | Ac-c001 (62.5 µM) |
| (–) Ac-c001 + Benzaldehyde | Ac-c001 (62.5 µM) + benzaldehyde (375 µM) |
| (+) Ac-$H3_{(1–9)}K9Me_3$ | Ac-$H3_{(1–9)}K9Me_3$ (62.5 µM) |

## 6.10 BROMO*scan*

This assay was carried out by Eurofins DiscoverX Corporation (San Diego, USA) using the following protocol:

T7 phage strains displaying proteins of interest were grown in parallel in 24-well blocks in an *E. coli* host derived from the BL21 strain. *E. coli* were grown to log-phase and infected with T7 phage from a frozen stock (multiplicity of infection = 0.4) and incubated with shaking at 32 °C until lysis (90–150 min). The lysates were centrifuged (5000 ×*g*) and filtered (0.2 μm) to remove cell debris. Streptavidin-coated magnetic beads were treated with biotinylated small molecule or acetylated peptide ligands for 30 min at room temperature to generate affinity resins for bromodomain assays. The liganded beads were blocked with excess biotin and washed with blocking buffer (SeaBlock (Pierce), 1 % BSA, 0.05 % Tween 20, 1 mM DTT) to remove unbound ligand and to reduce non-specific phage binding. Binding reactions were assembled by combining bromodomains, liganded affinity beads, and test compounds in 1× binding buffer (17% SeaBlock, 0.33× PBS, 0.04% Tween 20, 0.02% BSA, 0.004% Sodium azide, 7.4 mM DTT). Test compounds were prepared as 1000× stocks in 100% DMSO. $K_d$ values were determined using an 11-point 3-fold compound dilution series with one DMSO control point. All compounds for $K_d$ measurements are distributed by acoustic transfer (non-contact dispensing) in 100% DMSO. The compounds were then diluted directly into the assays such that the final concentration of DMSO was 0.09%. All reactions were performed in polypropylene 384-well plates with a final assay volume of 20 μL. The assay plates were incubated at room temperature with shaking for 1 h and the affinity beads were washed with wash buffer (1× PBS, 0.05% Tween 20). The beads were then re-suspended in elution buffer (1× PBS, 0.05% Tween 20, 2 μM non-biotinylated affinity ligand) and incubated at room temperature

with shaking for 30 min. The protein concentration in the eluates was measured using qPCR. Binding constants ($K_d$) were calculated with a standard dose-response curve using the Hill equation (Equation 8):

$$Response = Background + \frac{Signal - Background}{1 + (\frac{K_d}{Dose})^{Hill\ slope}} \qquad (8)$$

The Hill Slope was set to −1. Curves were fitted using a non-linear least square fit with the Levenberg-Marquardt algorithm.

## 6.11 Surface Plasmon Resonance (SPR)

### 6.11.1 General SPR Experimental

SPR was carried out using a Biacore™ T200 sensor with a C1 or CM5 sensor chip (Cytiva) and HBS-EP (pH 7.4, Table E20) as the running/analyte binding buffer. Protein (50 µg/mL) was covalently immobilised onto flow cell (FC) surfaces in aqueous NaOAc (10 mM, pH 4.0) at 25 ºC and a flow rate of 10 µL/min. Immobilisation was achieved by activating carboxymethyl groups on the chip surface using EDC (400 mM) : NHS (100 mM) (1:1) for 600 s, followed by covalent capture of the protein on the surface for 12–100 s, and capping of excess activated carboxylates with ethanolamine-HCl (1 M, pH 8.5) for 600 s. A reference FC was activated with EDC/NHS and directly capped with ethanolamine. Each analyte sample was injected continuously for a specified amount of time and allowed to flow over the chip surface (association phase). Injection was then stopped and replaced with a continuous flow of running buffer (dissociation phase). Lastly a conditioning step was carried out using an injection of running buffer before the next analyte injection. Sensogram report points were collected at 10 Hz. For all experiments, the reference FC and the previous conditioning step were subtracted from the active FC sensograms to obtain the bound ligand concentration, expressed in response units (RU), as a function of time.

**Table E20:** HBS-EP buffer (pH 7.4, Cytiva) composition. The buffer was filtered through 0.2 μm filter paper under vacuum and degassed prior to use.

| Reagent | Concentration |
|---|---|
| HEPES | 10 mM |
| NaCl | 150 mM |
| EDTA | 3 mM |
| P20 surfactant | 0.005% (*v*/*v*) |

### 6.11.2 Binding Saturation Assay (Equilibrium Analysis)

Two-fold serial dilutions of the analyte were prepared in HBS-EP buffer from an aqueous stock solution. $His_6$-TRIM24 WT or W828A mutant was immobilised to high protein loadings (>1000 RU) onto the surface of 3 flow cells. Samples were injected and allowed to reach equilibrium before dissociation ($H3_{(1\text{-}27)}K9Me_3K18Ac$ - 30 s association; 120 s dissociation; 50 μL/min; 15 ºC). Binding saturation curves were obtained by plotting the equilibrium response (curve plateau) at constant time against the analyte concentration and fitting the data using a Langmuir (1:1) one-site total binding model (Equation 9) in GraphPad Prism.

$$Response = \left(\frac{B_{max}x}{K_d + x}\right) + x(NS) + Background \qquad (9)$$

($x$ = [compound] , $B_{max}$ = maximum signal, $NS$ = non-specific binding)

### 6.11.3 Multi-cycle Kinetics Assay (Kinetic Analysis)

Two or three-fold serial dilutions (as indicated) of the analyte were prepared in HBS-EP buffer from an aqueous stock solution. $His_6$-TRIM24 WT or W828 mutant was immobilised to low protein loadings (500–1000 RU) onto the surface of three flow cells, followed by sample injection and dissociation [(Compound **63** - 30 s association; 300 s

dissociation; 75 µL/min; 15 ºC), (Compound **60** - 90 s association; 900 s dissociation; 75 µL/min; 15 ºC)]. One-phase association (Equation 10) and dissociation (Equation 11) models were used to fit the data in GraphPad Prism to obtain values for $k_{off}$ ($s^{-1}$) and $k_{obs}$ ($s^{-1}$) for each concentration. These values were then substituted into Equation 12 to obtain $k_{on}$ ($M^{-1}\ s^{-1}$) for each concentration. Mean $k_{off}$ and $k_{on}$ values were obtained by taking the values from the fitted curves for each concentration and then averaging across the replicates. The law of mass action (Equation 13) was then used to calculate a $K_d$ (M).

$$Response\ =\ (R_0\ -\ NS)exp(-k_{off}t)\ +\ NS \qquad (10)$$

$$Response = R_0\ +\ (R_{max} - R_0)(1 - exp(-k_{obs}t)) \qquad (11)$$

$$k_{on}\ =\ \frac{k_{obs} - k_{off}}{[compound]} \qquad (12)$$

$$K_d = \frac{k_{off}}{k_{on}} \qquad (13)$$

($R_0$ = response at $t_0$, $R_{max}$ = plateau response, $NS$ = non-specific binding)

## 6.12 Cell Permeability Assay

MCF-7 cells were cultured by Dr. Antoine Wallabrègue (University of Oxford) in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% FBS and 1% PS, at 37°C and 5% $CO_2$ in a humidified incubator. Cells were routinely mycoplasma tested using a HEK-Blue™ detection kit (Invivogen).

**Table E21:** Description and origin of cell lines used in the cell permeability assay.

| Cell Line | Description | Origin |
|---|---|---|
| **MCF-7** | Human breast adenocarcinoma, luminal A, ER positive | ATCC® (ATCC-HB-22) |

10000 cells were seeded in cell culture glass dishes (60 × 15 mm) and allowed to adhere for 24 h, before changing the medium to DMEM. Cells were treated with the compound of interest (20 μM) or an equal volume of the DMSO control, and incubated for 6 h. Cells were washed three times with HBS buffer supplemented with a protease inhibitor cocktail (pH 7.4, Table E22) (washes were collected as a control). Cells were collected by scraping and centrifugation at 15000 rpm for 15 min (the supernatant was collected as a control). The remaining cell pellet was stored at −80 °C overnight.

**Table E22:** 1× HBS buffer (pH 7.4) containing a Sigma*FAST*™ protease inhibitor cocktail tablet EDTA free (Sigma Aldrich). The buffer was filtered through 0.2 μm filter paper under vacuum before use.

| Reagent | Concentration |
|---|---|
| HEPES | 10 mM |
| NaCl | 150 mM |
| Sigma*FAST*™ protease inhibitor cocktail tablet EDTA free (Sigma Aldrich) | 1 tablet per 400 mL |

Cell pellets were resuspended in MeCN/$H_2O$ (1:1, 200 uL) and transferred to 15 mL Falcon tubes, before being sonicated (35% amplitude) on ice for 15 min *via* a 5 s burst/10 s pause cycle (5 min total sonication time). The cell lysates were transferred to 1.5 mL Eppendorf tubes and clarified by centrifugation at 15000 rpm for 15 min. The supernatant was then analysed using LC-MS (50 μL injection). Blank subtraction was carried out using the DMSO control sample.

## 6.13 *in vitro* Cellular Assays

All cell cultures and assays were carried out by Louise Martin (University of Oxford), unless otherwise stated.

### 6.13.1 Cell Lines

**Table E23:** Description and origin of cell lines used in the cellular assays.

| Cell Line | Description | Origin |
|---|---|---|
| **CAL-51** | Human Breast carcinoma, triple negative | Prof. Eileen Parkes, University of Oxford |
| **MCF-7** | Human Breast adenocarcinoma, luminal A, ER positive | Prof. Katherine Vallis, University of Oxford |
| **BT-474** | Human Breast adenocarcinoma, triple positive | Prof. Adrian Harris, University of Oxford |

### 6.13.2 Cell Culture

Cells were grown in high glucose DMEM medium (Life Technologies Ltd., 41965062) supplemented with 10% fetal bovine serum (FBS, Merck Life Sciences U.K. Ltd., F7524). Unless otherwise stated, cells were cultured in a humidified incubator at 37 °C and 5% $CO_2$. Cell lines were passaged by trypsinisation (Trypsin-EDTA solution 1×, Merck Life Sciences U.K. Ltd., T3924), and by diluting the cells into a culture flask at the desired density in complete media. Cell lines were routinely mycoplasma tested (MycoAlert Mycoplasma Detection Kit, Lonza, LT07-418; MycoAlert Assay Control Set, Lonza, LT07-518) and were found to be negative.

### 6.13.3 MTT Assay

Cells were seeded and allowed to adhere for 6 h in plastic clear-bottomed 96-well plates in a total volume of 200 µL/well complete medium. Optimal seeding numbers for each cell line were established (BT-474 8000 cells/well, CAL-51 4000 cells/well). The media was replaced with 200 µL of fresh media containing the compound of interest at 10 µM concentration, and the plates were incubated at 37 °C for 16 h. At the end of treatment, without removing the media, 0.5 mg/mL MTT reagent (Life technologies Ltd., M6494) was added, and incubated for 3 h at 37 °C, protected from light. The resulting formazan (purple) crystals were solubilised by replacing the media with 100 µL of DMSO (Merck

Life Sciences U.K. Ltd., D8418), and leaving for 15 min at 37 °C, protected from light. Absorbance at 570 nm was immediately measured using a POLARstar Omega plate reader (BMG LABTECH). Data are expressed as percentage cell viability relative to the DMSO control.

### 6.13.4 Cell Proliferation Assay

Cells were seeded and allowed to adhere for 6 h in 24-well plates (CAL-51 4000 cell/well, MCF-7 10000 cells/well). The media was replaced with 300 µL of fresh media containing the compound of interest (10 µM), and the plates were incubated at 37 °C. For single treatment assays, counting was repeated each day in triplicate without media replacement. For multiple treatment assays, counting was repeated each day in triplicate before replacing the media with 300 µL of fresh media containing the compound of interest (10 µM).

### 6.13.5 Clonogenic Assay

Pretreated cells from the proliferation assays were seeded at low density (CAL51 250 cells/well, MCF7 500 cells/well) in 6 well plates and incubated at 37 °C until colony formation. Colonies were stained with 2% crystal violet diluted in 50% methanol and 20% ethanol, and counted manually. Plating efficiency was determined by dividing the total number of colonies by the seeding number. The surviving fraction was then calculated by dividing the plating efficiency for treatment by the plating efficiency for the DMSO control.

# References


(1) Sekirnik, A.; Reynolds, J.; See, L.; Bluck, J.; Scorah, A.; Tallant, C.; Lee, B.; Leszczynska, K.; Grimley, R.; Storer, I.; Malattia, M.; Crespillo, S.; Caria, S.; Duclos, S.; Hammond, E.; Knapp, S.; Morris, G.; Duarte, F.; Biggin, P.; Conway, S. *ACS Chem. Biol.* **2022**, 17, 2753–2768.

(2) Tsai, W. W.; Wang, Z.; Yiu, T. T.; Akdemir, K. C.; Xia, W.; Winter, S.; Tsai, C. Y.; Shi, X.; Schwarzer, D.; Plunkett, W.; Aronow, B.; Gozani, O.; Fischle, W.; Hung, M. C.; Patel, D. J.; Barton, M. C. *Nature* **2010**, 468, 927–932.

(3) Bennett, J.; Fedorov, O.; Tallant, C.;Monteiro, O.; Meier, J.; Gamble, V.;Savitsky, P.; Nunez-Alonso, G.; Haendler, B.; Rogers, C.; Brennan, P. E.; Müller, S.; Knapp, S. *J. Med. Chem.* **2016**, 59, 1642–1647.

(4) Palmer, W. S.; Poncet-Montange, G.; Liu, G.; Petrocchi, A.; Reyna, N.; Subramanian, G.; Theroff, J.; Yau, A.; Kost-Alimova, M.; Bardenhagen, J. P.; Leo, E.; Shepard, H. E.; Tieu, T. N.; Shi, X.; Zhan, Y.; Zhao, Shuping; Barton, M. C.; Draetta, G.; Toniatti, C.; Jones, P.; Geck Do, M.; Andersen, J. N. *J. Med. Chem.* **2016**, 59, 1440–1454.

(5) Simon, M. D.; Shokat, K. M. A. *Methods in Enzymology,* Academic Press Inc. **2012**, 512, 57–69.

(6) Pieters, B. J. G. E.; Hintzen, J. C. J.; Grobben, Y.; Temimi, A. H. K. Al; Kamps, J. J. A. G.; Mecinović, J. *Bioconjug. Chem.* **2019**, 30, 952–958.

(7) Simon, M. D.; Chu, F.; Racki, L. R.; Cruz, C. C. de la; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. *Cell* **2007**, 128, 1003–1012.

(8) Arrowsmith, C. H.; Bountra, C.; Fish, P. V.; Lee, K.; Schapira, M. *Nat. Rev. Drug Discov.* **2012**, 11, 384–400.

(9) Sharma, S.; Kelly, T. K.; Jones, P. A. *Carcinogenesis* **2010**, 31, 27–36.

(10) Martin, G. M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 10413–10414.

(11) Srinageshwar, B.; Maiti, P.; Dunbar, G. L.; Rossignol, J. *Int. J. Mol. Sci.* **2016**, 17, 199.

(12) Tammen, S. A.; Friso, S.; Choi, S. W. *Mol. Aspects Med.* **2013**, 34, 753–764.

(13) Quina, A. S.; Buschbeck, M.; Di Croce, L. *Biochem. Pharmacol.* **2006**, 72, 1563–1569.

(14) Luger, K.; Mäder, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J. *Nature* **1997**, 389, 251–260.

(15) Huisinga, K. L.; Brower-Toland, B.; Elgin, S. C. R. *Chromosoma* **2006**, 115, 110–122.

(16) Clarke, T. L.; Mostoslavsky, R. *Mol. Oncol.* **2022**, 16, 3352–3379.

(17) Liu, R.; Wu, J.; Guo, H.; Yao, W.; Li, S.; Lu, Y.; Jia, Y.; Liang, X.; Tang, J.; Zhang, H. *MedComm* **2023**, 4, e292.

(18) Fenley, A. T.; Anandakrishnan, R.; Kidane, Y. H.; Onufriev, A. V. *Epigenetics & Chromatin* **2018**, 11, 11.

(19) Dhalluin, C.; Carlson, J. E.; Zeng, L.; He, C.; Aggarwal, A. K.; Zhou, M. M. *Nature* **1999**, 399, 491–496.

(20) Platt, M. *Development of Hypoxia-Activated PROTAC Degraders of BRD4*, Unpublished Masters Thesis, University of Oxford, **2020**.

(21) Biswas, S.; Rao, C. M. *Eur. J. Pharmacol.* **2018**, 837, 8–24.

(22) Biswas, S.; Rao, C. M. *Pharmacol. Ther.* **2017**, 173, 118–134.

(23) Jennings, L. E.; Measures, A. R.; Wilson, B. G.; Conway, S. J. *Future Med. Chem.* **2014**, 6, 179–204.

(24) Zaware, N.; Zhou, M. M. *Nat. Struct. Mol. Biol.* **2019**, 26, 870–879.

(25) Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.; Lambert, J. P.; Barsyte-Lovejoy, D.; Felletar, I.; Volkmer, R.; Müller, S.; Pawson, T.; Gingras, A. C.; Arrowsmith, C. H.; Knapp, S. *Cell* **2012**, 149, 214–231.

(26) Meslamani, J.; Smith, S. G.; Sanchez, R.; Zhou, M. M. *Drug Discov. Today Technol.* **2016**, 19, 3–15.

(27) Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E. *Nature* **2010**, 468, 1067–1073.

(28) Zengerle, M.; Chan, K. H.; Ciulli, A. *ACS Chem. Biol.* **2015**, 10, 1770–1777.

(29) Moustakim, M.; Clark, P. G. K.; Hay, D. A.; Dixon, D. J.; Brennan, P. E. *Med. Chem. Commun.* **2016**, 7, 2246–2264.

(30) Pan, Z.; Zhao, Y.; Wang, X.; Xie, X.; Liu, M.; Zhang, K.; Wang, L.; Bai, D.; Foster, L. J.; Shu, R.; He, G. *Mol. Biomed.* **2023**, 4, 13.

(31) Umehara, T.; Nakamura, Y.; Jang, M. K.; Nakano, K.; Tanaka, A.; Ozato, K.; Padmanabhan, B.; Yokoyama, S. *J. Biol. Chem.* **2010**, 285, 7610–7618.

(32) Welti, J.; Sharp, A.; Brooks, N.; Yuan, W.; McNair, C.; Chand, S. N.; Pal, A.; Figueiredo, I.; Riisnaes, R.; Gurel, B.; Rekowski, J.; Bogdan, D.; West, W.; Young, B.; Raja, M.; Prosser, A.; Lane, J.; Thomson, S.; Worthington, J.; Onions, S.; Shannon, J.; Paoletta, S.; Brown, R.; Smyth, D.; Harbottle, G. W.; Gil, V. S.; Miranda, S.; Crespo, M.; Ferreira, A.; Pereira, R.; Tunariu, N.; Carreira, S.; Neeb,

A. J.; Ning, J.; Swain, A.; Taddei, D.; Schiewer, M. J.; Knudsen, K. E.; Pegg, N.; de Bono, J. S. *Cancer Discov.* **2021**, 11, 1118–1137.

(33) Bamborough, P.; Barnett, H. A.; Becher, I.; Bird, M. J.; Chung, C. W.; Craggs, P. D.; Demont, E. H.; Diallo, H.; Fallon, D. J.; Gordon, L. J.; Grandi, P.; Hobbs, C. I.; Hooper-Greenhill, E.; Jones, E. J.; Law, R. P.; Le Gall, A.; Lugo, D.; Michon, A. M.; Mitchell, D. J.; Prinjha, R. K.; Sheppard, R. J.; Watson, A. J. B.; Watson, R. J. *ACS Med. Chem. Lett.* **2016**, 7, 552–557.

(34) Moustakim, M.; Clark, P. G. K.; Trulli, L.; Fuentes de Arriba, A. L.; Ehebauer, M. T.; Chaikuad, A.; Murphy, E. J.; Mendez-Johnson, J.; Daniels, D.; Hou, C. F. D.; Lin, Y. H.; Walker, J. R.; Hui, R.; Yang, H.; Dorrell, L.; Rogers, C. M.; Monteiro, O. P.; Fedorov, O.; Huber, K. V. M.; Knapp, S.; Heer, J.; Dixon, D. J.; Brennan, P. E. *Angew. Chem. Int. Ed.* **2017**, 56, 827–831.

(35) Musselman, C. A.; Kutateladze, T. G. *Mol. Interventions* **2009**, 9, 314–323.

(36) Li, H.; Ilin, S.; Wang, W.; Duncan, E. M.; Wysocka, J.; Allis, C. D.; Patel, D. J. *Nature* **2006**, 442, 91–95.

(37) Sanchez, R.; Zhou, M. M. *Trends Biochem. Sci.* **2011**, 36, 364–372.

(38) Musselman, C. A.; Khorasanizadeh, S.; Kutateladze, T. G. *Biochim. Biophys. Acta - Gene Regulatory Mechanisms* **2014**, 1839, 686–693.

(39) Morrison, E. A.; Musselman, C. A. *Chromatin Signaling and Diseases*; Academic Press Inc. **2016**, 127–147.

(40) Milosevich, N.; Hof, F. *Biochem.* **2016**, 55, 1570–1583.

(41) Santiago, C.; Nguyen, K.; Schapira, M. *J. Comput. Aided Mol. Des.* **2011**, 25, 1171–1178.

(42) Wagner, E. K.; Nath, N.; Flemming, R.; Feltenberger, J. B.; Denu, J. M. *Biochem.* **2012**, 51, 8293–8306.

(43) Ali, F.; Yamaguchi, K.; Fukuoka, M.; Elhelaly, A. E.; Kuwata, K. *Cancer Sci.* **2016**, 107, 1321–1328.

(44) Liu, W. H.; Miner, R. E.; Albaugh, B. N.; Ananiev, G. E.; Wildman, S. A.; Denu, J. M. *Biochem.* **2022**, 61, 354–366.

(45) Engelberg, I. A.; Liu, J.; Norris-Drouin, J. L.; Cholensky, S. H.; Ottavi, S. A.; Frye, S. V.; Kutateladze, T. G.; James, L. I. *J. Med. Chem.* **2021**, 64, 8510–8522.

(46) Zhang, M. Y.; Yang, H.; Ortiz, G.; Trnka, M. J.; Petronikolou, N.; Burlingame, A. L.; DeGrado, W. F.; Fujimori, D. G. *Chem. Sci.* **2022**, 13, 6599–6609.

(47) Ortiz, G.; Kutateladze, T. G.; Fujimori, D. G. *Curr. Opin. Chem. Biol.* **2023**, 74, 102286.

(48) Musselman, C. A.; Kutateladze, T. G. *Nucleic Acids Res.* **2011**, 39, 9061–9071.

(49) Ruthenburg, A. J.; Li, H.; Patel, D. J.; David Allis, C. *Nat. Rev. Mol. Cell Biol.* **2007**, 8, 983–994.

(50) Jacobson, R. H.; Ladurner, A. G.; King, D. S.; Tjian, R. *Science* **2000**, 288, 1422–1425.

(51) Fischle, W.; Boo, S. T.; Dormann, H. L.; Ueberheide, B. M.; Garcia, B. A.; Shabanowitz, J.; Hunt, D. F.; Funabiki, H.; Allis, C. D. *Nature* **2005**, 438, 1116–1122.

(52) Wang, Z.; Patel, D. J. *J. Biol. Chem.* **2011**, 286, 18363–18368.

(53) Waring, M. J.; Chen, H.; Rabow, A.; Walker, G.; Bobby, R.; Boiko, S.; Bradbury, R. H.; Callis, R.; Clark, E.; Dale, I.; Daniels, D. L.; Dulak, A.; Flavell, L.; Holdgate, G.; Jowitt, T.; Kikhney, A.; Mcalister, M.; Méndez, J.; Ogg, D.; Patel, J.; Petteruti, P.; Robb, G.; Robers, M. B.; Saif, S.; Stratton, N.; Svergun, D. I.; Wang, W.; Whittaker, D.; Wilson, D. M.; Yao, Y. *Nat. Chem. Biol.* **2016**, 12, 1097–1104.

(54) Bamborough, P.; Chung, C.; Goodwin, N. C.; Mitchell, D. J.; Neipp, C. E.; Phillipou, A.; Preston, A.; Prinjha, R. K.; Soden, P. E.; Watson, R. J.; Demont, E. H. *ACS Med. Chem. Lett.* **2023**, 14, 1231–1236

(55) Tanaka, M. *Nat. Chem. Biol.* **2016**, 12, 1089–1096.

(56) Suh, J. L.; Watts, B.; Stuckey, J. I.; Norris-Drouin, J. L.; Cholensky, S. H.; Dickson, B. M.; An, Y.; Mathea, S.; Salah, E.; Knapp, S.; Khan, A.; Adams, A. T.; Strahl, B. D.; Sagum, C. A.; Bedford, M. T.; James, L. I.; Kireev, D. B.; Frye, S. V. *Biochem.* **2018**, 57, 2140–2149.

(57) D'Amico, F.; Mukhopadhyay, R.; Ovaa, H.; Mulder, M. P. C. *ChemBioChem* **2021**, 22, 2011–2031.

(58) Ozato, K.; Shin, D.-M.; Chang, T.-H.; Morse, H. C. *Nat. Rev. Immunol.* **2008**, 8, 849–860.

(59) Vunjak, M.; Versteeg, G. A. **2019**, 29, R42–R44.

(60) Yang, W.; Gu, Z.; Zhang, H.; Hu, H. *Front. Immunol.* **2020**, 11.

(61) Koepke, L.; Gack, M. U.; Sparrer, K. M. *Curr. Opin. Microbiol.* **2021**, 59, 50–57.

(62) Watanabe, M.; Hatakeyama, S. *J. Biochem.* **2017**, 161, 135–144.

(63) Hatakeyama, S. *Nat. Rev. Cancer* **2011**, 11, 792–804.

(64) McAvera, R. M.; Crawford, L. J. *Cancers* **2020**, 12, 2094.

(65) Khetchoumian, K.; Teletin, M.; Mark, M.; Lerouge, T.; Cerviñ, M.; Oulad-Abdelghani, M.; Chambon, P.; Losson, R. G. *J. Biol. Chem.* **2004**, 46, 48329–48341.

(66) Yao, Y.; Zhou, S.; Yan, Y.; Fu, K.; Xiao, S. *Cell Biosci.* **2024**, 14, 103.

(67) Xi, Q.; Wang, Z.; Zaromytidou, A-I.; Zhang, X. H-F.; Chow-Tsang, L-F.; Liu, J. X.; Kim, H.; Barlas, A.; Manova-Todorova, K.; Kaartinen, V.; Studer, L.; Mark, W.; Patel, D. J.; Massagué, J. *Cell* **2011**, 147, 1511–1524.

(68) Jain, A. K.; Allton, K.; Duncan, A. D.; Barton, M. C. *Mol. Cell Biol.* **2014**, 34, 2695–2709.

(69) Allton, K.; Jain, A. K.; Herz, H. M.; Tsai, W. W.; Sung, Y. J.; Qin, J.; Bergmann, A.; Johnson, R. L.; Barton, M. C. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 11612–11616.

(70) Isbel, L.; Iskar, M.; Durdu, S.; Weiss, J.; Grand, R. S.; Hietter-Pfeiffer, E.; Kozicka, Z.; Michael, A. K.; Burger, L.; Thomä, N. H.; Schübeler, D. *Nat. Struct. Mol. Biol.* **2023**, 30, 948–957.

(71) Vom Baur, E.; Zechel, C.; Heery, D.; Heine, M. J. S.; Garnier, J. M.; Vivat, V.; Le Douarin, B.; Gronemeyer, H.; Chambon, P.; Losson, R. *EMBO J.* **1996**, 15, 110–124.

(72) Le Douarin, B.; Nielsen, A. L.; Garnier, J. M.; Ichinose, H.; Jeanmougin, F.; Losson, R.; Chambon, P. *EMBO J.* **1996**, 15, 6701–6715.

(73) Kikuchi, M.; Okumura, F.; Tsukiyama, T.; Watanabe, M.; Miyajima, N.; Tanaka, J.; Imamura, M.; Hatakeyama, S. *Biochim. Biophys. Acta – Mol. Cell Res.* **2009**, 1793, 1828–1836.

(74) Carrier, M.; Lutzing, R.; Gaouar, S.; Rochette-Egly, C. *FEBS Lett.* **2023**, 597, 1176–1176.

(75) Appikonda, S.; Thakkar, K. N.; Shah, P. K.; Dent, S. Y. R.; Andersen, J. N.; Barton, M. C. *J. Biol. Chem.* **2018**, 293, 7476–7485.

(76) Wang, Y.; Yao, Y.; Wei, Q.; Long, S.; Chen, Y.; Xie, J.; Tan, R.; Jiang, W.; Zhang, Q.; Wu, D.; Xiao, S.; Wan, F.; Fu, K. *Oncogene* **2022**, 42, 586–600.

(77) Horvath, R. M.; Brumme, Z. L.; Sadowski, I. *Sci. Rep.* **2023**, 13, 556.

(78) Horvath, R. M.; Dahabieh, M.; Malcolm, T.; Sadowski, I. *Commun. Biol.* **2023**, 6, 86.

(79) Zhu, Q.; Yu, T.; Gan, S.; Wang, Y.; Pei, Y.; Zhao, Q.; Pei, S.; Hao, S.; Yuan, J.; Xu, J.; Hou, F.; Wu, X.; Peng, C.; Wu, P.; Qin, J.; Xiao, Y. *J. Exp. Med.* **2020**, 217, e20192083.

(80) Gandini, D.; De Angeli, C.; Aguiari, G.; Manzati, E.; Lanza, F.; Pandolfi, P. P.; Cuneo, A.; Castoldi, G. L.; Del Senno, L. *Leukemia* **2002**, 16, 886–893.

(81) Li, C.; Xin, H.; Shi, Y.; Mu, J. *Hum. Exp. Toxicol.* **2020**, 39, 1725–1736.

(82) Zhang, L. H.; Yin, A. A.; Cheng, J. X.; Huang, H. Y.; Li, X. M.; Zhang, Y. Q.; Han, N.; Zhang, X. *Oncogene* **2015**, 34, 600–610.

(83) Tian, H.; Zhao, H.; Qu, B.; Chu, X.; Xin, X.; Zhang, Q.; Li, W.; Yang, S. *Am. J. Transl. Res.* **2022**, 14, 831–848.

(84) Groner, A. C.; Cato, L.; de Tribolet-Hardy, J.; Bernasocchi, T.; Janouskova, H.; Melchers, D.; Houtman, R.; Cato, A. C. B.; Tschopp, P.; Gu, L.; Corsinotti, A.; Zhong, Q.; Fankhauser, C.; Fritz, C.; Poyet, C.; Wagner, U.; Guo, T.; Aebersold, R.; Garraway, L. A.; Wild, P. J.; Theurillat, J. P.; Brown, M. *Cancer Cell* **2016**, 29, 846–858.

(85) Offermann, A.; Roth, D.; Hupe, M. C.; Hohensteiner, S.; Becker, F.; Joerg, V.; Carlsson, J.; Kuempers, C.; Ribbat-Idel, J.; Tharun, L.; Sailer, V.; Kirfel, J.; Svensson, M.; Andren, O.; Lubczyk, V.; Kuefer, R.; Merseburger, A. S.; Perner, S. *Urologic Oncology: Seminars and Original Investigations* **2019**, 37, 576.e1-576.e10.

(86) Zhang, L.; Chen, H.; Ding, B.; Jiang, W. *J. Ovarian Res.* **2022**, 15, 19.

(87) Lin, L.; Zhao, W.; Sun, B.; Wang, X.; Liu, Q. *Am. J. Transl. Res.* **2017**, 9, 620–628.

(88) Ma, L.; Yuan, L.; An, J.; Barton, M. C.; Zhang, Q.; Liu, Z. *Tumor Biol.* **2016**, 37, 14803–14812.

(89) Palmer, W. S. *Drug Discov. Today Technol.* **2016**, 19, 65–71.

(90) Hu, Q.; Wang, C.; Xiang, Q.; Wang, R.; Zhang, C.; Zhang, M.; Xue, X.; Luo, G.; Liu, X.; Wu, X.; Zhang, Y.; Wu, D.; Xu, Y. *Bioorg. Chem.* **2020**, 94, 103424.

(91) Xiang, Q.; Luo, G.; Zhang, C.; Hu, Q.; Wang, C.; Wu, T.; Xu, H.; Hu, J.; Zhuang, X.; Zhang, M.; Wu, S.; Xu, J.; Zhang, Y.; Liu, J.; Xu, Y. *Eur. J. Med. Chem.* **2022**, 236, 114311.

(92) Gechijian, L. N.; Buckley, D. L.; Lawlor, M. A.; Reyes, J. M.; Paulk, J.; Ott, C. J.; Winter, G. E.; Erb, M. A.; Scott, T. G.; Xu, M.; Seo, H. S.; Dhe-Paganon, S.; Kwiatkowski, N. P.; Perry, J. A.; Qi, J.; Gray, N. S.; Bradner, J. E. *Nat. Chem. Biol.* **2018**, 14, 405–412.

(93) Dean, T. T.; Jelu-́Reyes, J.; Allen, A. C.; Moore, T. W. *J. Med. Chem.* **2024**, 67, 1641–1661.

(94) Cartwright, O. C.; Beekman, A. M.; Cominetti, M. M. D.; Russell, D. A.; Searcey, M. *Bioconjug. Chem.* **2020**, 31, 1745–1749.

(95) Sakamoto, K. M.; Kim, K. B.; Kumagai, A.; Mercurio, F.; Crews, C. M.; Deshaies, R. J. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 8554–8559.

(96) Tsuji, K.; Hymel, D.; Ma, B.; Tamamura, H.; Nussinov, R.; Burke, T. R. *RSC Chem. Biol*. **2022**, 3, 1111–1120.

(97) Lamers, C. *Future Drug Discov.* **2022**, 4, FDD75.

(98) Ullman, E. F.; Kirakossian, H.; Singh, S.; Wu, Z. P.; Irvin, B. R.; Pease, J. S.; Switchenko, A. C.; Irvine, J. D.; Dafforn, A.; Skold, C. N.; Wagner, D. B. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, 91, 5426–5430.

(99) Peppard, J.; Glickman, F.; He, Y.; Hu, S. I.; Doughty, J.; Goldberg, R. *J. Biomol. Screen* **2003**, 8, 149–156.

(100) Eglen, R. M.; Reisine, T.; Roby, P.; Rouleau, N.; Illy, C.; Bossé, R.; Bielefeld, M. *Curr. Chem. Genomics* **2008**, 1, 2–10.

(101) Schorpp, K.; Rothenaigner, I.; Salmina, E.; Reinshagen, J.; Low, T.; Brenke, J. K.; Gopalakrishnan, J.; Tetko, I. V.; Gul, S.; Hadian, K. *J. Biomol. Screen* **2014**, 19, 715–726.

(102) Perrin, F. *J. Phys. Radium*, **1926**, 7, 390–401.

(103) Moerke, N. J. *Curr. Protoc. Chem. Biol.* **2009**, 1, 1–15.

(104) Lea, W. A.; Simeonov, A. *Expert Opin. Drug Discov.* **2011**, 6, 17–32.

(105) Nguyen, H. H.; Park, J.; Kang, S.; Kim, M. *Sensors* **2015**, 15, 10481–10510.

(106) Patching, S. G. *Biochim. Biophys. Acta - Biomembranes* **2014**, 1838, 43–55.

(107) Frutiger, A.; Tanno, A.; Hwu, S.; Tiefenauer, R. F.; Vörös, J.; Nakatsuka, N. *Chem. Rev.* **2021**, 121, 8095–8160.

(108) Damborský, P.; Švitel, J.; Katrlík, J. *Essays Biochem.* **2016**, 60, 91–100.

(109) Sparks, R. P.; Jenkins, J. L.; Fratti, R. SNAREs. *Methods Mol. Biol.* Human Press, New York, NY, **2019**, 1860, 199–210.

(110) Voit, E. O.; Martens, H. A.; Omholt, S. W. *PLoS Comput. Biol.* **2015**, 11.

(111) Wu, Q.; Heidenreich, D.; Zhou, S.; Ackloo, S.; Krämer, A.; Nakka, K.; Lima-Fernandes, E.; Deblois, G.; Duan, S.; Vellanki, R. N.; Li, F.; Vedadi, M.; Dilworth, J.; Lupien, M.; Brennan, P. E.; Arrowsmith, C. H.; Müller, S.; Fedorov, O.; Filippakopoulos, P.; Knapp, S. *Nat. Commun.* **2019**, 10, 1915.

(112) Fabian, M. A.; Biggs, W. H.; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lélias, J. M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. *Nat. Biotechnol.* **2005**, 23, 329–336.

(113) Sekirnik, A. R.; Hewings, D. S.; Theodoulou, N. H.; Jursins, L.; Lewendon, K. R.; Jennings, L. E.; Rooney, T. P. C.; Heightman, T. D.; Conway, S. J. *Angew. Chem. Int. Ed.* **2016**, 55, 8353–8357.

(114) Hintzen, J. C. J.; Poater, J.; Kumar, K.; Temimi, A. H. K. Al; Pieters, B. J. G. E.; Paton, R. S.; Bickelhaupt, F. M.; Mecinović, J. *Molecules* **2020**, 25, 1918.

(115) Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2012**, 51, 1835–1839.

(116) Shogren-Knaak, M. A. *ACS Chem. Biol*. **2007**, 2, 225–227.

(117) Chen, Z.; Notti, R. Q.; Ueberheide, B.; Ruthenburg, A. J. *Biochem.* **2018**, 57, 300–304.

(118) Simhadri, C.; Daze, K. D.; Douglas, S. F.; Milosevich, N.; Monjas, L.; Dev, A.; Brown, T. M.; Hirsch, A. K. H.; Wulff, J. E.; Hof, F. *ChemMedChem* **2019**, 14, 1444–1456.

(119) Bhushan, B.; Erdmann, A.; Zhang, Y.; Belle, R.; Johannson, C.; Oppermann, U.; Hopkinson, R. J.; Schofield, C. J.; Kawamura, A. *Bioorg. Med. Chem*. **2018**, 26, 2984–2991.

(120) Lamb, K. N.; Bsteh, D.; Dishman, S. N.; Moussa, H. F.; Fan, H.; Stuckey, J. I.; Norris, J. L.; Cholensky, S. H.; Li, D.; Wang, J.; Sagum, C.; Stanton, B. Z.; Bedford, M. T.; Pearce, K. H.; Kenakin, T. P.; Kireev, D. B.; Wang, G. G.; James, L. I.; Bell, O.; Frye, S. V. *Cell Chem. Biol.* **2019**, 26, 1365-1379.

(121) Wang, S.; Alpsoy, A.; Sood, S.; Ordonez-Rubiano, S. C.; Dhiman, A.; Sun, Y.; Jiao, G.; Krusemark, C. J.; Dykhuizen, E. *ChemBioChem* **2021**, 22, 2335–2344.

(122) Senisterra, G.; Zhu, H. Y.; Luo, X.; Zhang, H.; Xun, G.; Lu, C.; Xiao, W.; Hajian, T.; Loppnau, P.; Chau, I.; Li, F.; Allali-Hassani, A.; Atadja, P.; Oyang, C.; Li, E.; Brown, P. J.; Arrowsmith, C. H.; Zhao, K.; Yu, Z.; Vedadi, M. *SLAS Discov.* **2018**, 23, 930–940.

(123) Wang, S.; Denton, K. E.; Hobbs, K. F.; Weaver, T.; McFarlane, J. M. B.; Connelly, K. E.; Gignac, M. C.; Milosevich, N.; Hof, F.; Paci, I.; Musselman, C. A.; Dykhuizen, E. C.; Krusemark, C. J. *ACS Chem. Biol.* **2020**, 15, 112–131.

(124) James, L. I.; Korboukh, V. K.; Krichevsky, L.; Baughman, B. M.; Herold, J. M.; Norris, J. L.; Jin, J.; Kireev, D. B.; Janzen, W. P.; Arrowsmith, C. H.; Frye, S. V. *J. Med. Chem.* **2013**, 56, 7358–7371.

(125) Pan, X.; Wang, H.; Li, C.; Zhang, J. Z. H.; Ji, C. *J. Chem. Inf. Model* **2021**, 61, 3159–3165.

(126) Kiessling, L. L.; Diehl, R. C. *ACS Chem. Biol.* **2021**, 16, 1884–1893.

(127) Bhagwat, S.; Tiu, R. V.; Wu, W. *Combination of ERK1/2 inhibitor compound with gemcitabine or with gemcitabine and nab-paclitaxel for use in treatment of pancreatic cancer*. WO2018/005234 A1, **2018**.

(128) López, P.; Seipelt, C. G.; Merkling, P.; Sturz, L.; Álvarez, J.; Dölle, A.; Zeidler, M. D.; Cerdán, S.; Ballesteros, P. *Bioorg. Med. Chem.* **1999**, 7, 517–527.

(129) Mambwe, D.; Kumar, M.; Ferger, R.; Taylor, D.; Njoroge, M.; Coertzen, D.; Reader, J.; Van Der Watt, M.; Birkholtz, L. M.; Chibale, K. *ACS Med. Chem. Lett.* **2021**, 12, 1333–1341.

(130) Oka, M.; Baba, K.; Nakamura, K.; Dong, L.; Hamajima, H.; Unno, R.; Matsumoto, Y. *J. Heterocycl. Chem.* **2003**, 40, 177–180.

(131) Marx, M. A.; Christensen, J. G.; Smith, C. R.; Fischer, J. P.; Burns, A. C. *KRAS G12C Inhibitors.* WO2020/146613 A1, **2020**.

(132) Marx, M. A.; Christensen, J. G.; Smith, C. R.; Fischer, J. P.; Burns, A. C. *KRAS G12C Inhibitors.* WO2020/146613 A1, **2020**.

(133) Hwu, J. R.; Hsu, C. I.; Hsu, M. H.; Liang, Y. C.; Huang, R. C. C.; Lee, Y. C. *Bioorg. Med. Chem. Lett.* **2011**, 21, 380–382.

(134) Srdanović, S.; Hegedüs, Z.; Warriner, S. L.; Wilson, A. J. *RSC Chem. Biol.* **2022**, 3, 546–550.

(135) Hegedüs, Z.; Hóbor, F.; Shoemark, D. K.; Celis, S.; Lian, L. Y.; Trinh, C. H.; Sessions, R. B.; Edwards, T. A.; Wilson, A. J. *Chem. Sci*. **2021**, 12, 2286–2293.

(136) Hua, L.; Wang, D.; Wang, K.; Wang, Y.; Gu, J.; Zhang, Q.; You, Q.; Wang, L. *J. Med. Chem.* **2023**, 66, 10934–10958.

(137) Edman, P. *Acta Chem. Scand.* **1950**, 4, 283–293.

(138) Gechijian, L. N. *The Design and Characterization of Bifunctional Degraders of TRIM24, a Novel Transcriptional Dependency in Hematologic Malignancies* Published PhD Thesis, Harvard University, **2018**.

(139) Plettner, E.; Gong, Y.; Gries, R. *Methods and compositions for control of gypsy moths, lymanria dispar.* US2010/0190865 A1, **2010**.

(140) Stratton, O. *Synthesis of a Novel PROTAC to Recruit TRIM24 as an E3 Ligase*. Unpublished Masters Thesis, University of Oxford, **2019**.

(141) Reynolds, J. *Development of Chemical Tools to Probe the Functions of TRIM Proteins*, Unpublished PhD Thesis, University of Oxford, **2020**.

(142) Kurono, M.; Kondo, Y.; Matsumoto, Y.; Oka, M.; Sawai, K. *1-Azoniabicyclo(3.3.0)oct-1(5)-ene salt derivatives and process for the preparation of the same.* US005110945, **1992**.

(143) Cloke, J. B. *J. Am. Chem. Soc.* **1929**, 51, 1174–1187.

(144) Nazeer, U.; Mushtaq, A.; Zahoor, A. F.; Hafeez, F.; Shahzadi, I.; Akhtar, R. *RSC Adv.* **2023**, 13, 35695–35732.

(145) Miyano, S.; Yamashita, O.; Sumoto, K.; Shima, K.; Hayashimatsu, M.; Satoh, F. *J. Heterocycl. Chem.* **1987**, 24, 271–274.

(146) Nascimento, D. L.; Reim, I.; Foscato, M.; Jensen, V. R.; Fogg, D. E. *ACS Catal*. **2020**, 10, 11623–11633.

(147) Woodward, C. P.; Spiccia, N. D.; Jackson, W. R.; Robinson, A. J. *Chem. Commun.* **2010**, 47, 779–781.

(148) Richelle Gaston, J. J.; Streefkerk Dieuwertje, E.; Van Maarseveen, J. H.; Timmerman, P. *Multicyclic peptides and methods for their preparation*. WO2018/106112 A1, **2018**.

(149) Grieco, P. A.; Gilman, S.; Nishizawa, M. *J. Org. Chem.* **1976**, 41, 1485–1486.

(150) Oka, M.; Baba, K.; Suzuki, T.; Matsumoto', *Heterocycles* **1997**, 45, 2317.

(151) Mancuso, A. J.; Huang, S. L.; Swern, D. *J. Org. Chem.* **1978**, 43, 2480–2482.

(152) Maercker, A. *Org. Reactions* **2011**, 270–490.

(153) Thomas, A.; Gopalan, B.; Lingam Prasada Rao, V. S.; Shah, D. M. *Novel dipeptidyl peptidase iv inhibitors, pharmaceutical compositions containing them, and process for their preparation.* WO2006/040625 A1, **2006**.

(154) Baldwin, J. E.; Killin, S. J.; Adlington, R. M.; Spiegel, U. *Tetrahedron* **1988**, 44, 2633–2636.

(155) Zhang, H.; Ginn, J.; Zhan, W.; Liu, Y. J.; Leung, A.; Toita, A.; Okamoto, R.; Wong, T. T.; Imaeda, T.; Hara, R.; Yukawa, T.; Michino, M.; Vendome, J.; Beuming, T.; Sato, K.; Aso, K.; Meinke, P. T.; Nathan, C. F.; Kirkman, L. A.; Lin, G. *J. Med. Chem.* **2022**, 65, 9350–9375.

(156) De Luca, L.; Giacomelli, G.; Porcheddu, A. *Org. Lett.* **2001**, 3, 3041–3043.

(157) Fanning, K. N.; Sutherland, A. *Tetrahedron Lett.* **2007**, 48, 8479–8481.

(158) Appel, R. *Angew. Chem. Int. Ed.* **1975**, 14, 801–811.

(159) Subirós-Funosas, R.; Prohens, R.; Barbas, R.; El-Faham, A.; Albericio, F. *Chem. Eur. J.* **2009**, 15, 9394–9403.

(160) Barlos, K.; Gatos, D.; Kallitsis, J.; Papaphotiu, G.; Sotiriu, P.; Wenqing, Y.; Schäfer, W. Darstellung Geschützter *Tetrahedron Lett.* **1989**, 30, 3943–3946.

(161) Bollhagen, R.; Schmiedberger, M.; Barlos, K.; Grell, E. *J. Chem. Soc. Chem. Commun.* **1994**, 22, 2559–2560.

(162) Sekirnik, A. *Probing Tandem Epigenetic Reader Domains*, Published PhD Thesis, University of Oxford, **2015**.

(163) Douglass, E. F.; Miller, C. J.; Sparer, G.; Shapiro, H.; Spiegel, D. A. A *J. Am. Chem. Soc.* **2013**, 135, 6092–6099.

(164) Vezzoli, A.; Bonadies, N.; Allen, M. D.; Freund, S. M. V.; Santiveri, C. M.; Kvinlaug, B. T.; Huntly, B. J. P.; Göttgens, B.; Bycroft, M. *Nat. Struct. Mol. Biol.* **2010**, 17, 617–619.

(165) Klein, B. J.; Cox, K. L.; Jang, S. M.; Côté, J.; Poirier, M. G.; Kutateladze, T. G. *Structure* **2020**, 28, 105-110.

(166) Obi, J. O.; Lubula, M. Y.; Cornilescu, G.; Henrickson, A.; McGuire, K.; Evans, C. M.; Phillips, M.; Boyson, S. P.; Demeler, B.; Markley, J. L.; Glass, K. C. *Curr. Res. Struct. Biol.* **2020**, 2, 104–115.

(167) Quinn, E.; Wodicka, L.; Ciceri, P.; Pallares, G.; Pickle, E.; Torrey, A.; Floyd, M.; Hunt, J.; Treiber, D. *Cancer Res.* **2013**, 73, 4238–4238.

(168) Skwarska, A.; Calder, E. D. D.; Sneddon, D.; Bolland, H.; Odyniec, M. L.; Mistry, I. N.; Martin, J.; Folkes, L. K.; Conway, S. J.; Hammond, E. M. *Cell Chem. Biol.* **2021**, 28, 1258-1270.

(169) O'Connor, L. J.; Cazares-Körner, C.; Saha, J.; Evans, C. N. G.; Stratford, M. R. L.; Hammond, E. M.; Conway, S. J. *Nat. Protoc.* **2016**, 11, 781–794.

(170) Hamamoto, K.; Kida, Y.; Zhang, Y.; Shimizu, T.; Kuwano, K. *Microbiol. Immunol.* **2002**, 46, 741–749.

(171) Werle, M.; Bernkop-Schnürch, A. *Amino Acids* **2006**, 30, 351–367.

(172) Hong, S. Y.; Oh, J. E.; Lee, K. H. *Biochem. Pharmacol*. **1999**, 58, 1775–1780.

(173) Chatterjee, J.; Rechenmacher, F.; Kessler, H. *Angew. Chem. Int. Ed.* **2013**, 52, 254–269.

(174) Li, X.; Wang, N.; Liu, Y.; Li, W.; Bai, X.; Liu, P.; He, C. Y. *Bioorg. Chem.* **2023**, 141, 106892.

(175) Scorah, A. *Chemical tools to probe the function of TRIM33.* Unpublished PhD Thesis, University of Oxford, **2019**.

(176) Sanchez, R.; Zhou, M. M. *Curr. Opin. Drug Discov. Devel.* **2009**, 12, 659–665.

(177) Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. *Organometallics* **2010**, 29, 2176–2179.

(178) Hou, D. R.; Hsieh, Y. D.; Hsieh, Y. W. *Tetrahedron Lett*. **2005**, 46, 5927–5929.

(179) Xie, W.; Gong, B.; Ning, S.; Liu, N.; Zhang, Z.; Che, X.; Zheng, L.; Xiang, J. *Synlett* **2019**, 30, 2077–2080.

(180) Hansen, B.; Zahradník, R.; Dušek, K.; Klaban, J. *Acta. Chem. Scand.* **1962**, 16, 1945–1955.

(181) Lakhan, R.; Rai, B. J. *J. Chem. Eng. Data.* **2002**, 32, 384–386.

(182) Lakhan, R.; Singh, O. P. *Arch. Pharm.* **1985**, 318, 228–238.

(183) Faust, G.; Fiedler, W. *J. Prakt. Chem.* **1963**, 21, 113–130.

(184) Tse, E. G.; Houston, S. D.; Williams, C. M.; Savage, G. P.; Rendina, L. M.; Hallyburton, I.; Anderson, M.; Sharma, R.; Walker, G. S.; Obach, R. S.; Todd, M. H. *J. Med. Chem.* **2020**, 63, 11585–11601.

(185) Xue, C. B.; Roderick, J.; Jackson, S.; Rafalski, M.; Rockwell, A.; Mousa, S.; Olson, R. E.; DeGrado, W. F. *Bioorg. Med. Chem.* **1997**, 5, 693–705.

(186) Baraldi, P. G.; Romagnoli, R.; Nuñez, M. D. C.; Perretti, M.; Paul-Clark, M. J.; Ferrario, M.; Govoni, M.; Benedini, F.; Ongini, E. *J. Med. Chem.* **2003**, 47, 711–719.

(187) R. Katritzky, A.; J. Mokrosz, M. *Heterocycles* **1984**, 22, 505–512.

(188) Blaszykowski, C.; Aktoudianakis, E.; Bressy, C.; Alberico, D. and; Lautens, M. *Org. Lett.* **2006**, 8, 2043–2045.

(189) Oka, M.; Matsumoto, Y.; Hirooka, K.; Suzuki, T. *Chem. Pharm. Bull.* **2000**, 48, 1121–1124.

(190) Geng, J.; Mantovani, G.; Tao, L.; Nicolas, J.; Chen, G.; Wallis, R.; Mitchell, D. A.; Johnson, B. R. G.; Evans, S. D.; Haddleton, D. M. *J. Am. Chem. Soc.* **2007**, 129, 15156–15163.

(191) Gillam, T. A.; Caorale, C.; Brooks, R. D.; Bader, C. A.; Sorvina, A.; Werrett, M .V.; Wright, P. J.; Morrison, J .L.; Massi, M.; Brooks, D. A.; Zacchini, S.; Hickey, S. M.; Stagni, S.; Plush, S. E. *Inorg. Chem.* **2021**, 60, 10173–10185.

(192) Ryu, K. E.; Kim, B. R.; Sung, G. H.; Yoon, H. J.; Yoon, Y. J. *Synlett* **2015**, 26, 1985–1990.

(193) Smolyar, N. N.; Pankina, O. Y.; Bondarenko, A. I.; Borodkin, Y. S.; Khizhan, A. I. *Russ. J. Org. Chem.* **2011**, 47, 1190–1193.

(194) Conconi, M. T.; Marzaro, G.; Urbani, L.; Zanusso, I.; Di Liddo, R.; Castagliuolo, I.; Brun, P.; Tonus, F.; Ferrarese, A.; Guiotto, A.; Chilin, A. *Eur. J. Med. Chem.* **2013**, 67, 373–383.

(195) Kolodyazhnaya, S. N.; Simonov, A. M. *Chem. Heterocyclic Compounds* **1969**, 5, 529–532.

(196) Carpino, L. A.; Nasr, K.; Abdel-Maksoud, A. A.; El-Faham, A.; Ionescu, D.; Henklein, P.; Wenschuh, H.; Beyermann, M.; Krause, E.; Bienert, M. *Org. Lett.* **2009**, 11, 3718–3721.

(197) Jakopin, Ž.; Gobec, M.; Kodela, J.; Hazdovac, T.; Mlinarič-Raščan, I.; Sollner Dolenc, M. *Eur. J. Med. Chem.* **2013**, 69, 232–243.

(198) Xiong, S.; Wang, N.; Liu, C.; Shen, H.; Qu, Z.; Zhu, L.; Bai, X.; Hu, H. gang; Cong, W.; Zhao, L. *Bioorg. Med. Chem. Lett.* **2021**, 37, 127831.

(199) Jakopin, Ž.; Gobec, M.; Kodela, J.; Hazdovac, T.; Mlinarič-Raščan, I.; Sollner Dolenc, M. *Eur. J. Med. Chem.* **2013**, 69, 232–243.

(200) Miyano, S.; Sumoto, K.; Satoh, F.; Shima, K.; Hayashimatsu, M.; Morita, M.; Aisaka, K.; Noguchi, T. *J. Med. Chem.* **1985**, 28, 714–717.

(201) Allevi, P.; Galligani, M.; Anastasia, M. *Tetrahedron Asymmetry* **2002**, 13, 1901–1910.

(202) Zhang, H.; Ginn, J.; Zhan, W.; Liu, Y. J.; Leung, A.; Toita, A.; Okamoto, R.; Wong, T. T.; Imaeda, T.; Hara, R.; Yukawa, T.; Michino, M.; Vendome, J.; Beuming, T.; Sato, K.; Aso, K.; Meinke, P. T.; Nathan, C. F.; Kirkman, L. A.; Lin, G. *J. Med. Chem.* **2022**, 65, 9350–9375.

# Appendices

## Appendix 1 Protein Purification and Characterisation

**Table A1.1:** TRIM24 construct information.

| | |
|---|---|
| **Gene** | $His_6$-TRIM24 (PHD-BRD) WT |
| **Uniprot ID** | O15164 |
| **Residues** | N825–E1007 |
| **DNA sequence** | ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGG<br>GTACCGAGAACCTGTACTTCCAATCCATGAATGAGGACT<br>GGTGTGCAGTTTGTCAAAACGGAGGGGAACTCCTCTGCT<br>GTGAAAAGTGCCCCAAAGTATTCCATCTTTCTTGTCATGT<br>GCCCACATTGACAAATTTTCCAAGTGGAGAGTGGATTTG<br>CACTTTCTGCCGAGACTTATCTAAACCAGAAGTTGAATA<br>TGATTGTGATGCTCCCAGTCACAACTCAGAAAAAAAGAA<br>AACTGAAGGCCTTGTTAAGTTAACACCTATAGATAAAAG<br>GAAGTGTGAGCGCCTACTTTTATTTCTTTACTGCCATGAA<br>ATGAGCCTGGCTTTTCAAGACCCTGTTCCTCTAACTGTGC<br>CTGATTATTACAAAATAATTAAAAATCCAATGGATTTGT<br>CAACCATCAAGAAAAGACTACAAGAAGATTATTCCATGT<br>ACTCAAAACCTGAAGATTTTGTAGCTGATTTTAGATTGA<br>TCTTTCAAAACTGTGCTGAATTCAATGAGCCTGATTCAG<br>AAGTAGCCAATGCTGGTATAAAACTTGAAAATTATTTTG<br>AAGAACTTCTAAAGAACCTCTATCCAGAATGA |
| **Amino Acid Sequence** | MHHHHHHSSGVDLGTENLYFQSMNEDWCAVCQNGGELL<br>CCEKCPKVFHLSCHVPTLTNFPSGEWICTFCRDLSKPEVEY<br>DCDAPSHNSEKKKTEGLVKLTPIDKRKCERLLLFLYCHEMS<br>LAFQDPVPLTVPDYYKIIKNPMDLSTIKKRLQEDYSMYSKP<br>EDFVADFRLIFQNCAEFNEPDSEVANAGIKLENYFEELLKN<br>LYPE |
| **MW (Da)** | 23922 |
| **$\varepsilon$ ($M^{-1}$ $cm^{-1}$)** | 25160 |

| **Gene** | $His_6$-TRIM24 (PHD-BRD) W828A mutant |
|---|---|
| **Uniprot ID** | O15164 |
| **Residues** | N825–E1007 |
| **DNA sequence** | ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGG<br>GTACCGAGAACCTGTACTTCCAATCCATGAATGAGGACG<br>CGTGTGCAGTTTGTCAAAACGGAGGGGAACTCCTCTGCT<br>GTGAAAAGTGCCCCAAAGTATTCCATCTTTCTTGTCATGT<br>GCCCACATTGACAAATTTTCCAAGTGGAGAGTGGATTTG<br>CACTTTCTGCCGAGACTTATCTAAACCAGAAGTTGAATA<br>TGATTGTGATGCTCCCAGTCACAACTCAGAAAAAAAGAA<br>AACTGAAGGCCTTGTTAAGTTAACACCTATAGATAAAAG<br>GAAGTGTGAGCGCCTACTTTTATTTCTTTACTGCCATGAA<br>ATGAGCCTGGCTTTTCAAGACCCTGTTCCTCTAACTGTGC<br>CTGATTATTACAAAATAATTAAAAATCCAATGGATTTGT<br>CAACCATCAAGAAAAGACTACAAGAAGATTATTCCATGT<br>ACTCAAAACCTGAAGATTTTGTAGCTGATTTTAGATTGA<br>TCTTTCAAAACTGTGCTGAATTCAATGAGCCTGATTCAG<br>AAGTAGCCAATGCTGGTATAAAACTTGAAAATTATTTTG<br>AAGAACTTCTAAAGAACCTCTATCCAGAATGA |
| **Amino Acid Sequence** | MHHHHHHSSGVDLGTENLYFQSMNEDACAVCQNGGELLC<br>CEKCPKVFHLSCHVPTLTNFPSGEWICTFCRDLSKPEVEYD<br>CDAPSHNSEKKKTEGLVKLTPIDKRKCERLLLFLYCHEMSL<br>AFQDPVPLTVPDYYKIIKNPMDLSTIKKRLQEDYSMYSKPE<br>DFVADFRLIFQNCAEFNEPDSEVANAGIKLENYFEELLKNL<br>YPE |
| **MW (Da)** | 23807 |
| **$\varepsilon$ ($M^{-1}$ $cm^{-1}$)** | 19660 |

| **Gene** | GST-TRIM24 (PHD-BRD) WT |
|---|---|
| **Uniprot ID** | O15164 |
| **Residues** | N825–E1007 |
| **DNA sequence** | AAATTAACACAGTCTATGGCCATCATACGTTATATAGCT<br>GACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGCGT<br>GCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATT<br>AGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTT |

| | |
|---|---|
| | GAAACTCTCAAAGTTGATTTTCTTAGCAAGCTANNNGAA<br>ATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAAACA<br>TATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGT<br>TGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAAT<br>GTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAAAA<br>ACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAA<br>ATCCAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCA<br>AGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGAG<br>CTCAGAGAACCTGTACTTCCAATCCATGAATGAGGACTG<br>GTGTGCAGTTTGTCAAAACGGAGGGGAACTCCTCTGCTG<br>TGAAAAGTGCCCCAAAGTATTCCATCTTTCTTGTCATGTG<br>CCCACATTGACAAATTTTCCAAGTGGAGAGTGGATTTGC<br>ACTTTCTGCCGAGACTTATCTAAACCAGAAGTTGAATAT<br>GATTGTGATGCTCCCAGTCACAACTCAGAAAAAAAGAA<br>AACTGAAGGCCTTGTTAAGTTAACACCTATAGATAAAAG<br>GAAGTGTGAGCGCCTACTTTTATTTCTTTACTGCCATGAA<br>ATGAGCCTGGCTTTTCAAGACCCTGTTCCTCTAACTGTGC<br>CTGATTATTACAAAATAATTAAAAATCCAATGGATTTGT<br>CAACCATCAAGAAAAGACTACAAGAAGATTATTCCATGT<br>ACTCAAAACCTGAAGATTTTGTAGCTGATTTTAGATTGA<br>TCTTTCAAAACTGTGCTGAATTCAATGAGCCTGATTCAG<br>AAGTAGCCAATGCTGGTATAAAACTTGAAAATTATTTTG<br>AAGAACTTCTAAAGAACCTCTATCCAGAATGA |
| **Amino Acid Sequence** | MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDK<br>WRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIADKHNM<br>LGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDF<br>LSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDV<br>VLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWP<br>LQGWQATFGGGDHPPKSSSENLYFQSMNEDWCAVCQNGG<br>ELLCCEKCPKVFHLSCHVPTLTNFPSGEWICTFCRDLSKPEV<br>EYDCDAPSHNSEKKKTEGLVKLTPIDKRKCERLLLFLYCHE<br>MSLAFQDPVPLTVPDYYKIIKNPMDLSTIKKRLQEDYSMYS<br>KPEDFVADFRLIFQNCAEFNEPDSEVANAGIKLENYFEELLK<br>NLYPE |
| **MW (Da)** | 47993 |
| **$\varepsilon$ ($M^{-1}$ $cm^{-1}$)** | 68270 |

| | |
|---|---|
| **Gene** | $His_6$-SUMO-TRIM24 (PHD-BRD) WT |
| **Uniprot ID** | O15164 |
| **Residues** | N825–E1007 |
| **DNA sequence** | ATGAAACATCACCATCACCATCACCCCATGAGCGATTAC<br>GACATCCCCACTACTGAGAATCTTTATTTTCAGGGCGCC<br>ATGGGCAACGATCACATTAACCTGAAAGTGGCCGGTCAA<br>GACGGTAGCGTAGTCCAGTTTAAAATCAAACGCCACACC<br>CCTCTGTCGAAACTGATGAAAGCCTATTGTGAACGCCAA<br>GGTCTGTCTATGCGTCAGATCCGTTTTCGCTTCGATGGAC<br>AGCCGATTAACGAAACCGACACTCCAGCACAGCTGGAA<br>ATGGAAGATGAGGACACCATTGACGTGTTCCAGCAACA<br>GACCGGTGGATCCATGAATGAGGACTGGTGTGCAGTTTG<br>TCAAAACGGAGGGGAACTCCTCTGCTGTGAAAAGTGCCC<br>CAAAGTATTCCATCTTTCTTGTCATGTGCCCACATTGACA<br>AATTTTCCAAGTGGAGAGTGGATTTGCACTTTCTGCCGA<br>GACTTATCTAAACCAGAAGTTGAATATGATTGTGATGCT<br>CCCAGTCACAACTCAGAAAAAAAGAAAACTGAAGGCCT<br>TGTTAAGTTAACACCTATAGATAAAAGGAAGTGTGAGCG<br>CCTACTTTTATTTCTTTACTGCCATGAAATGAGCCTGGCT<br>TTTCAAGACCCTGTTCCTCTAACTGTGCCTGATTATTACA<br>AAATAATTAAAAATCCAATGGATTTGTCAACCATCAAGA<br>AAAGACTACAAGAAGATTATTCCATGTACTCAAAACCTG<br>AAGATTTTGTAGCTGATTTTAGATTGATCTTTCAAAACTG<br>TGCTGAATTCAATGAGCCTGATTCAGAAGTAGCCAATGC<br>TGGTATAAAACTTGAAAATTATTTTGAAGAACTTCTAAA<br>GAACCTCTATCCAGAATGA |
| **Amino Acid Sequence ( \| = cleavage site)** | MKHHHHHHPMSDYDIPTTENLYFQGAMGNDHINLKVAGQ<br>DGSVVQFKIKRHTPLSKLMKAYCERQGLSMRQIRFRFDGQ<br>PINETDTPAQLEMEDEDTIDVFQQQTGG\|SMNEDWCAVCQ<br>NGGELLCCEKCPKVFHLSCHVPTLTNFPSGEWICTFCRDLS<br>KPEVEYDCDAPSHNSEKKKTEGLVKLTPIDKRKCERLLLFL<br>YCHEMSLAFQDPVPLTVPDYYKIIKNPMDLSTIKKRLQEDY<br>SMYSKPEDFVADFRLIFQNCAEFNEPDSEVANAGIKLENYF<br>EELLKNLYPE |
| **MW (Da)** | 33804; 21457 (cleaved) |
| **$\varepsilon$ ($M^{-1}$ $cm^{-1}$)** | 28140; 23670 (cleaved) |

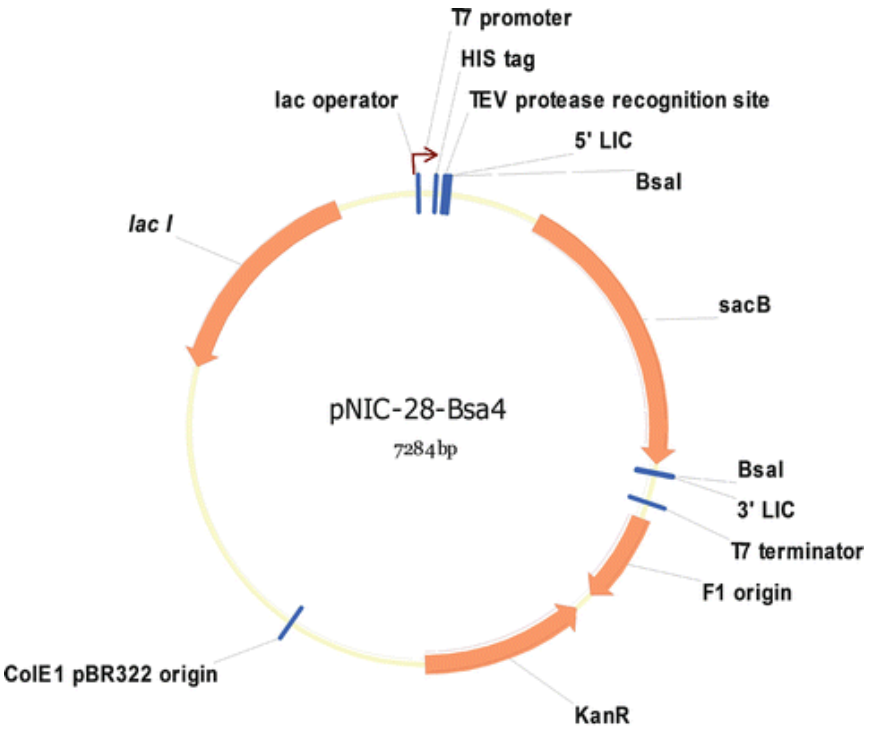



**Figure A1.1:** pNIC28-Bsa4 plasmid map.

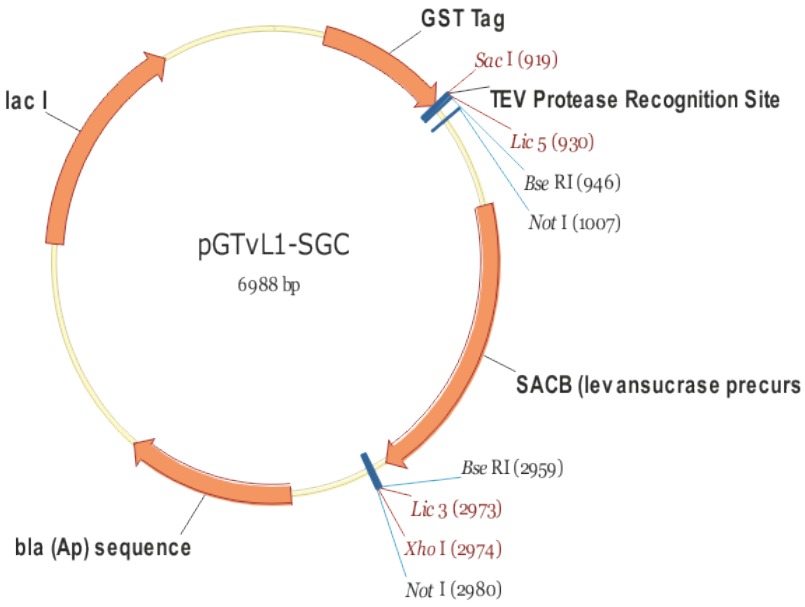



**Figure A1.2:** pGTvL1-SGC plasmid map.

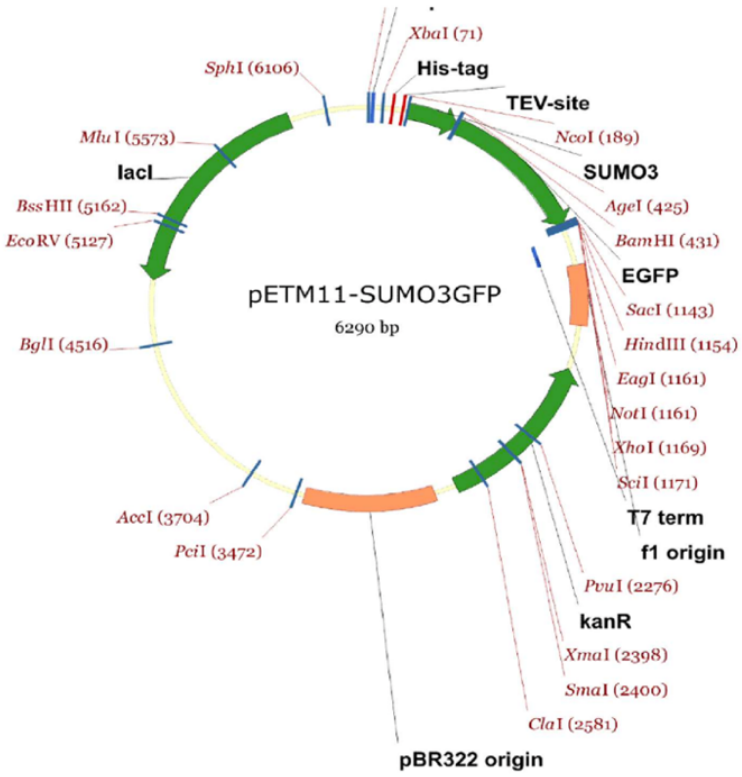



**Figure A1.3:** pETM11-SUMO3GFP plasmid map.

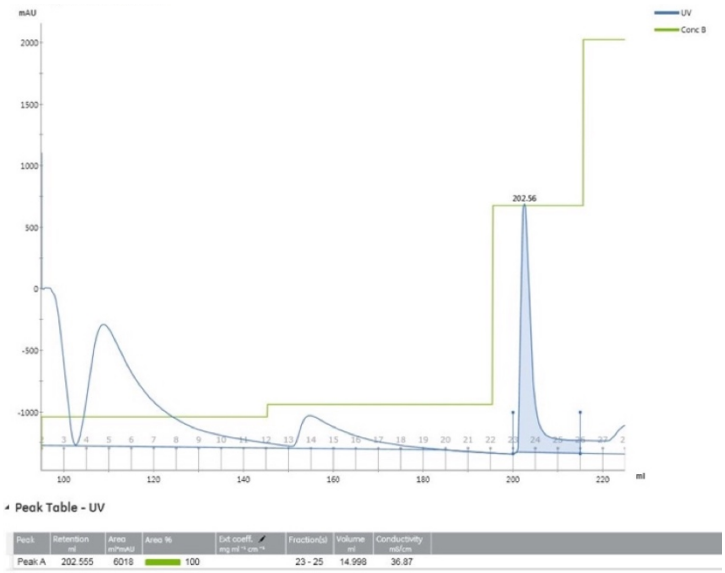



**Figure A1.4:** $His_6$-TRIM24 WT purification IMAC chromatogram. The peak corresponding to elution of TRIM24 is highlighted in blue.

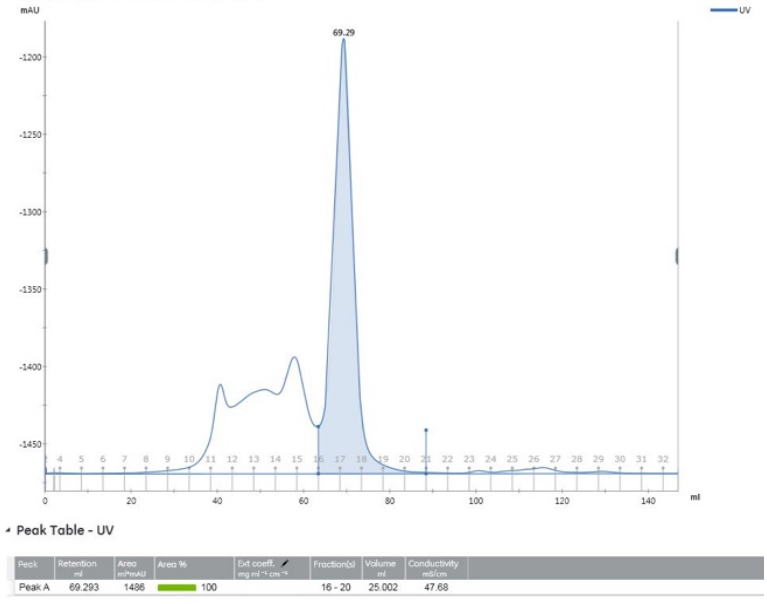



**Figure A1.5:** $His_6$-TRIM24 WT purification SEC chromatogram. The peak corresponding to elution of TRIM24 is highlighted in blue.

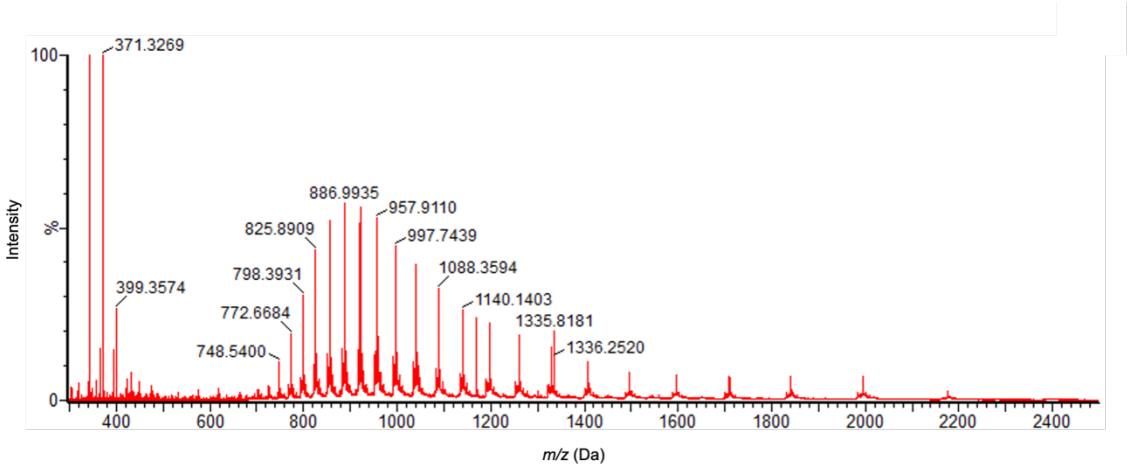



**Figure A1.6:** $His_6$-TRIM24 WT $ESI^+$ mass spectrum.

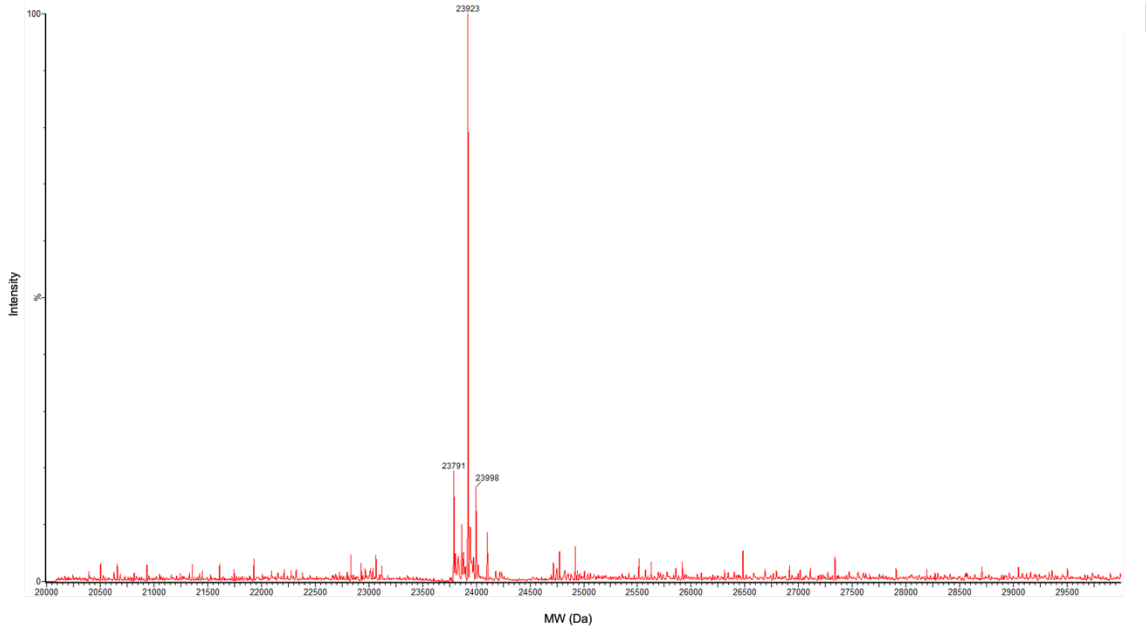



**Figure A1.7:** $His_6$-TRIM24 WT deconvoluted mass spectrum (20-30 kDa, 1 Da resolution). Observed MW = 23923 Da, Theoretical MW = 23922 Da.

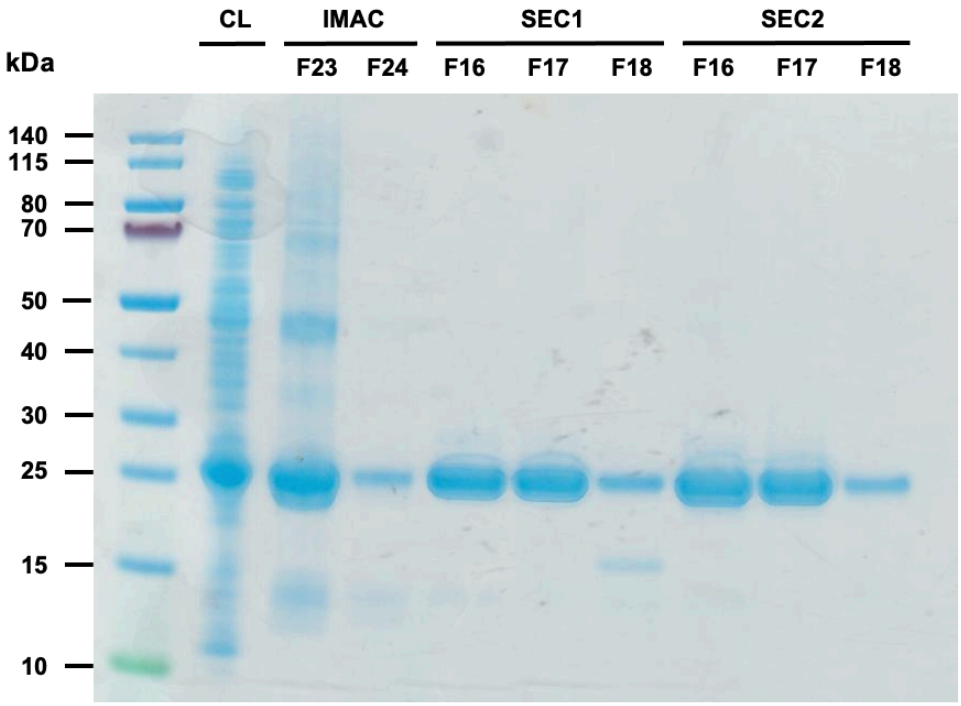



**Figure A1.8:** SDS-PAGE gel for the purification of $His_6$-TRIM24 WT. (CL, cell lysate; F, fraction).

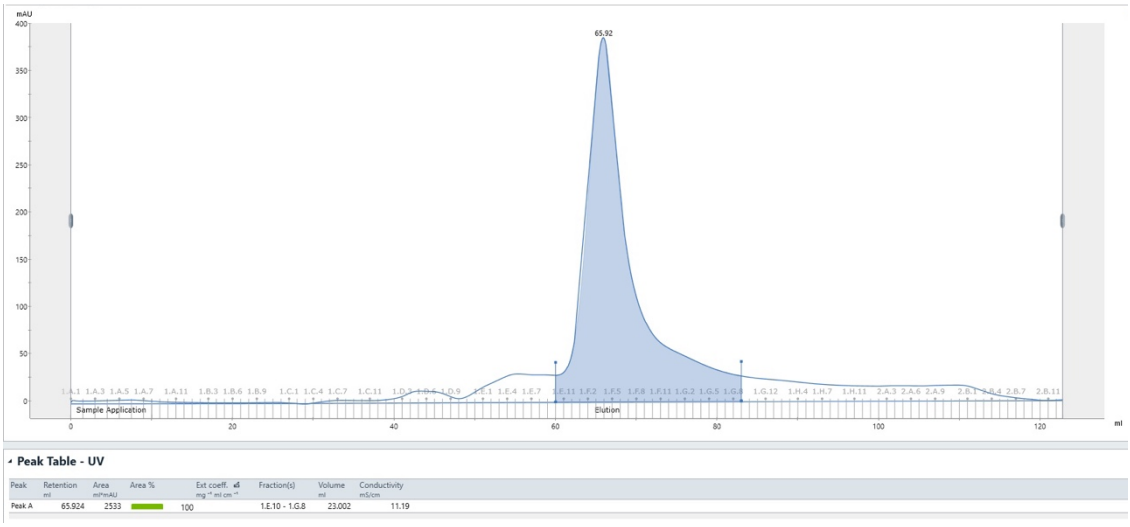



**Figure A1.9:** Untagged TRIM24 WT purification SEC chromatogram. The peak corresponding to elution of TRIM24 is highlighted in blue.

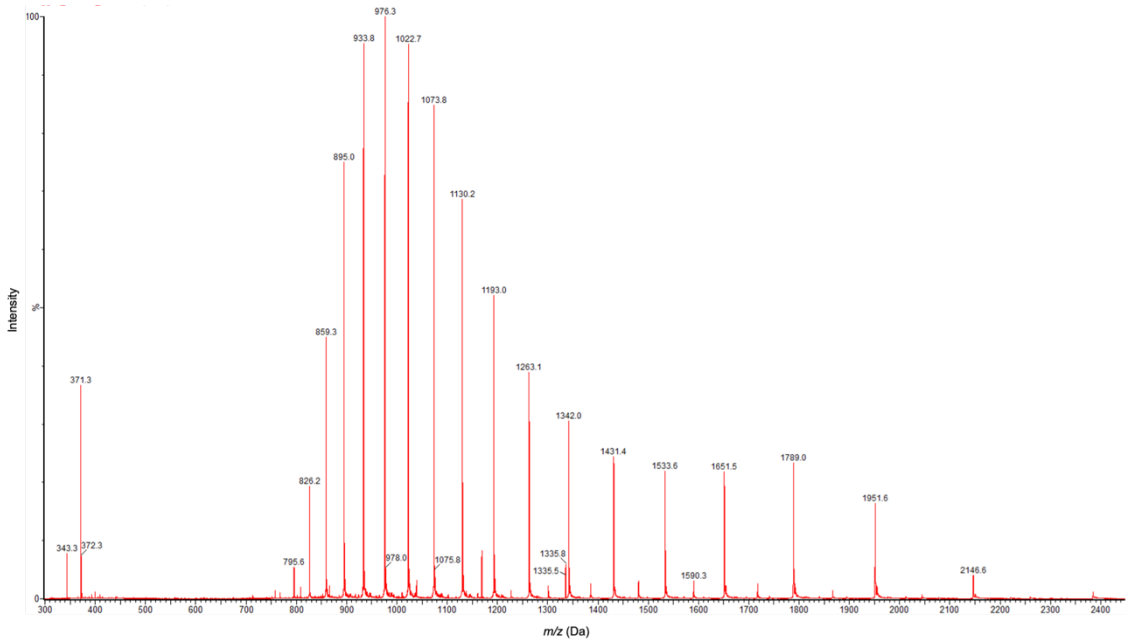



**Figure A1.10:** untagged TRIM24 WT $ESI^+$ mass spectrum.

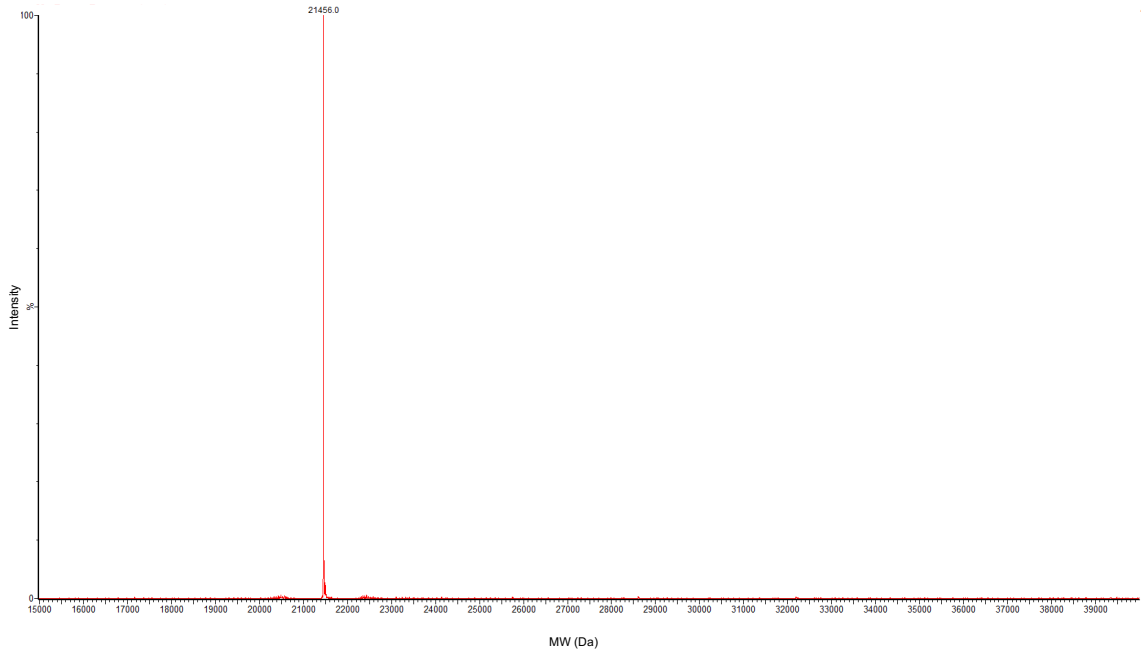



**Figure A1.11:** untagged TRIM24 WT deconvoluted mass spectrum (15-40 kDa, 1 Da resolution). Observed MW = 21456 Da, Theoretical MW = 21457 Da.

## Appendix 2 Protein Crystallography

**Table A2.1:** Data collection and refinement statistics for TRIM24 (PHD-BRD) crystal structures. Data for the highest resolution shell is shown in parentheses.

| **Crystal** | ***apo* TRIM24 (PHD-BRD) residues 825–1007** | **TRIM24 (PHD-BRD) residues 825–1007 in complex with ligand 61** |
|---|---|---|
| PDB ID | 9GD5 | 9GDG |
| Beamline | I03 (Diamond Light Source) | I03 (Diamond Light Source) |
| Wavelength (Å) | 0.9763 | 0.9763 |
| Space group | C121 | C121 |
| Unit Cell | | |
| a, b, c (Å) | 88.90, 36.91, 129.36 | 90.94, 36.65, 65.46 |
| α, β, γ (°) | 90.00, 110.04, 90.00 | 90.00, 111.33, 90.00 |
| Resolution (Å) | 44.45–1.68 (1.71–1.68) | 60.97–1.46 (1.64–1.46) |
| No. of reflections | 282838 (6530) | 153510 (5178) |
| No. of unique reflections | 43525 (1485) | 23849 (1192) |
| I/σ (I) | 18.1 (1.2) | 18.5 (2.0) |
| Completeness (%) | 95.7 (66.1) | 84.9 (32.6) |
| Multiplicity | 6.5 (4.4) | 6.4 (4.3) |
| $R_{meas}$ | 0.144 (2.469) | 0.051 (0.623) |
| $R_{factor}/R_{free}$ | 0.21/0.24 | 0.15/0.21 |
| Number of atoms (non-H) | | |
| Protein (Chain A) | 1439 | 1431 |
| Protein (Chain B) | 1439 | – |
| Water | 130 | 119 |
| Zinc | 4 | 2 |
| Ligand | – | 48 |
| Average B-factors (Å$^2$) | | |
| Protein (Chain A) | 25.596 | 25.127 |
| Protein (Chain B) | 26.667 | – |
| Water | 27.278 | 29.831 |
| Zinc | 21.690 | 17.093 |
| Ligand | – | 53.996 |
| RMS deviations | | |
| Bond lengths (Å) | 0.0090 | 0.0139 |
| Bond Angles (°) | 1.6020 | 2.1750 |
| Ramachandran favoured (%) | 99.4 | 98.8 |
| Ramachandran allowed (%) | 0.6 | 1.2 |
| Ramachandran outliers (%) | 0.0 | 0.0 |

**Figure A2.1:** PDB validation reports (p. 1–2) showing the overall model quality for PDB entries 9GD5 and 9GDG.

For Manuscript Review

WORLDWIDE PDB PROTEIN DATA BANK

# Full wwPDB X-ray Structure Validation Report (i)

Aug 7, 2024 – 10:07 am BST

| | | |
|---|---|---|
| PDB ID | : | 9GD5 |
| Title | : | Crystal structure of apo TRIM24 PHD-BRD in C121 space group |
| Deposited on | : | 2024-08-05 |
| Resolution | : | 1.68 Å(reported) |

**This wwPDB validation report is for manuscript review**

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

---

The following versions of software and data (see references (i)) were used in the production of this report:

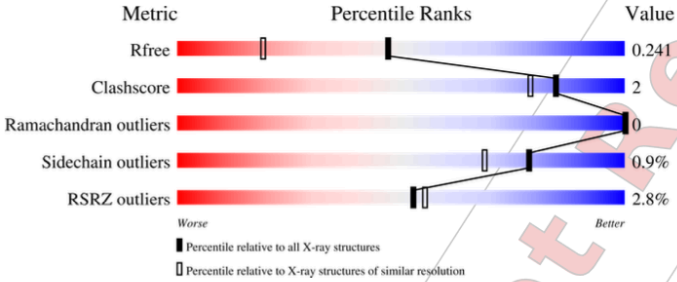
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.37.1 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.37.1 |

# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.68 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 6780 (1.70-1.66) |
| Clashscore | 141614 | 7310 (1.70-1.66) |
| Ramachandran outliers | 138981 | 7173 (1.70-1.66) |
| Sidechain outliers | 138945 | 7172 (1.70-1.66) |
| RSRZ outliers | 127900 | 6661 (1.70-1.66) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 185 | 3% 88% 7% |
| 1 | B | 185 | 3% 92% |

For Manuscript Review

# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2024 – 10:23 am BST

| | | |
|---|---|---|
| PDB ID | : | 9GDG |
| Title | : | Crystal structure of TRIM24 PHD-BRD in complex with N-(2-(2-(2-acetamidoethoxy)ethoxy)ethyl)-3-(N-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1H-benzo[d]imidazol-5-yl)sulfamoyl)benzamide (PEG linker unresolved) |
| Deposited on | : | 2024-08-05 |
| Resolution | : | 1.46 Å(reported) |

**This wwPDB validation report is for manuscript review**

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

---

The following versions of software and data (see references ⓘ) were used in the production of this report:

| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.4, CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.37.1 |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |

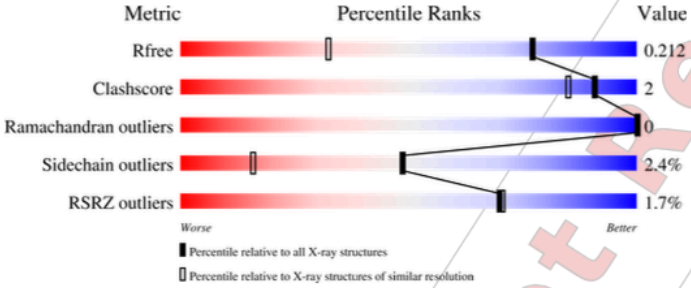
For Manuscript Review

# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.46 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 1156 (1.46-1.46) |
| Clashscore | 141614 | 1202 (1.46-1.46) |
| Ramachandran outliers | 138981 | 1178 (1.46-1.46) |
| Sidechain outliers | 138945 | 1178 (1.46-1.46) |
| RSRZ outliers | 127900 | 1139 (1.46-1.46) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

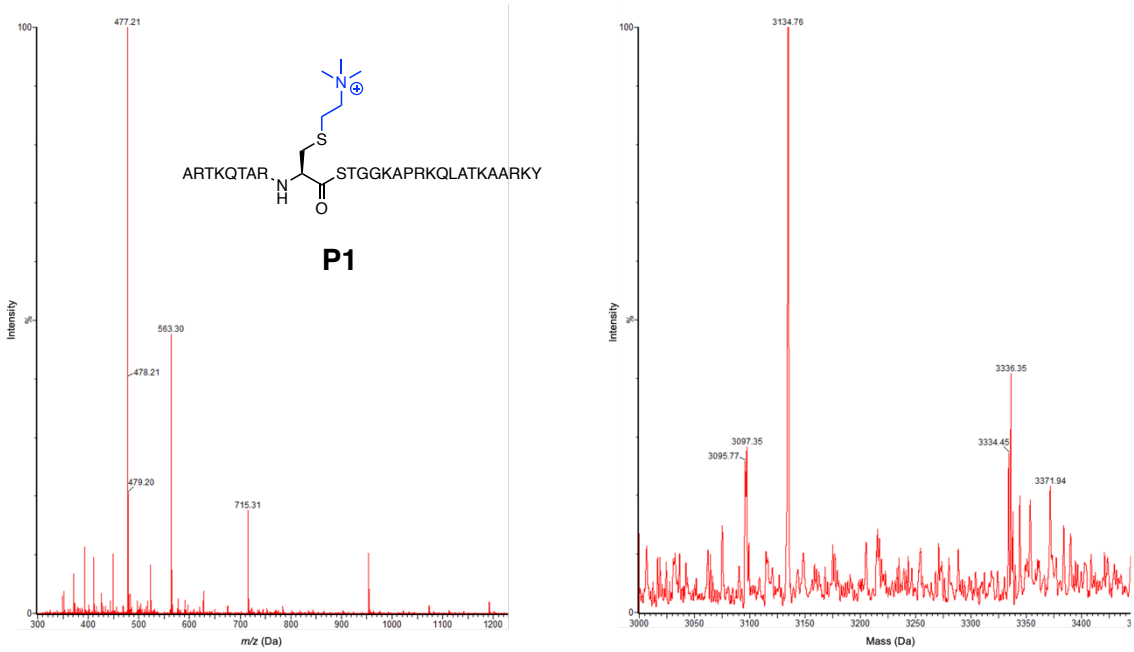
| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | B | 185 | 2% 86% 8% 5% |

| | | |
|---|---|---|
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.37.1 |

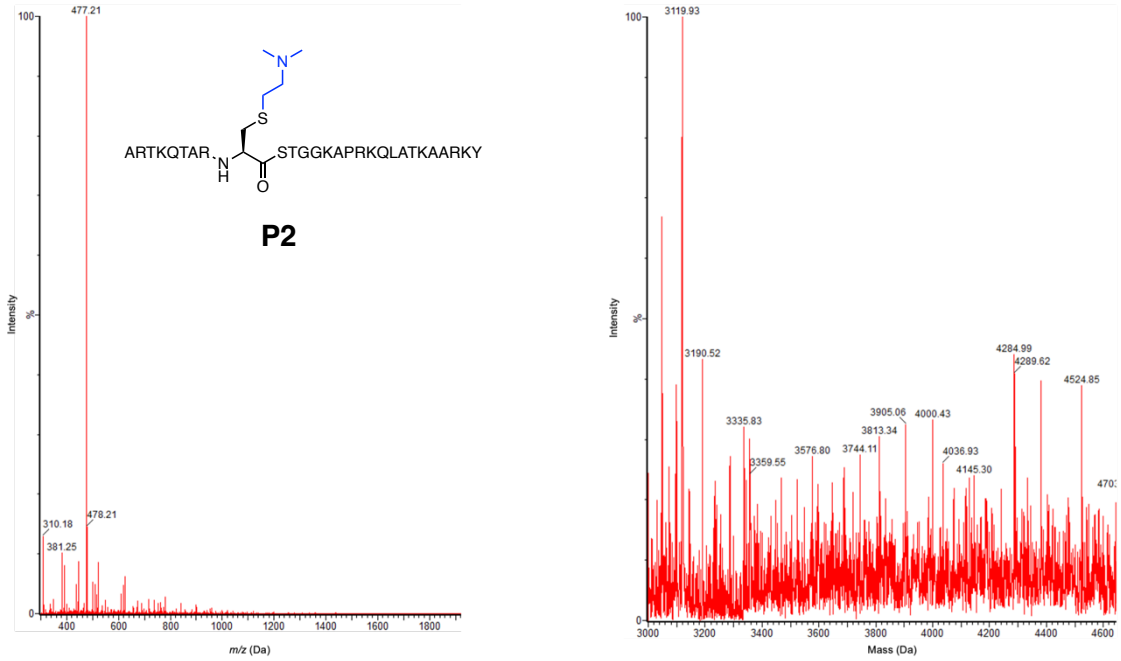
For Manuscript Review

## Appendix 3 Histone-mimicking Peptide Characterisation

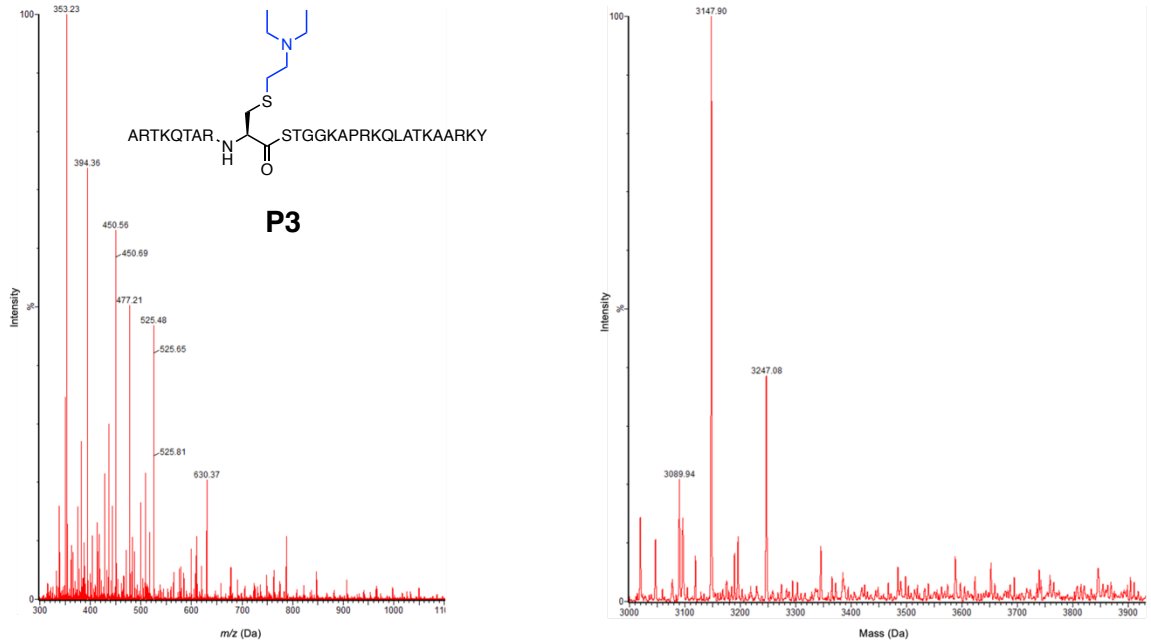
**Figure A3.1:** ESI$^+$ and deconvoluted mass spectra for histone-mimicking peptides **P1**–**P16** before purification. Structures of the peptides are shown with single letter amino acid codes. The modification is shown in blue. Theoretical molecular weights were calculated in ChemDraw.



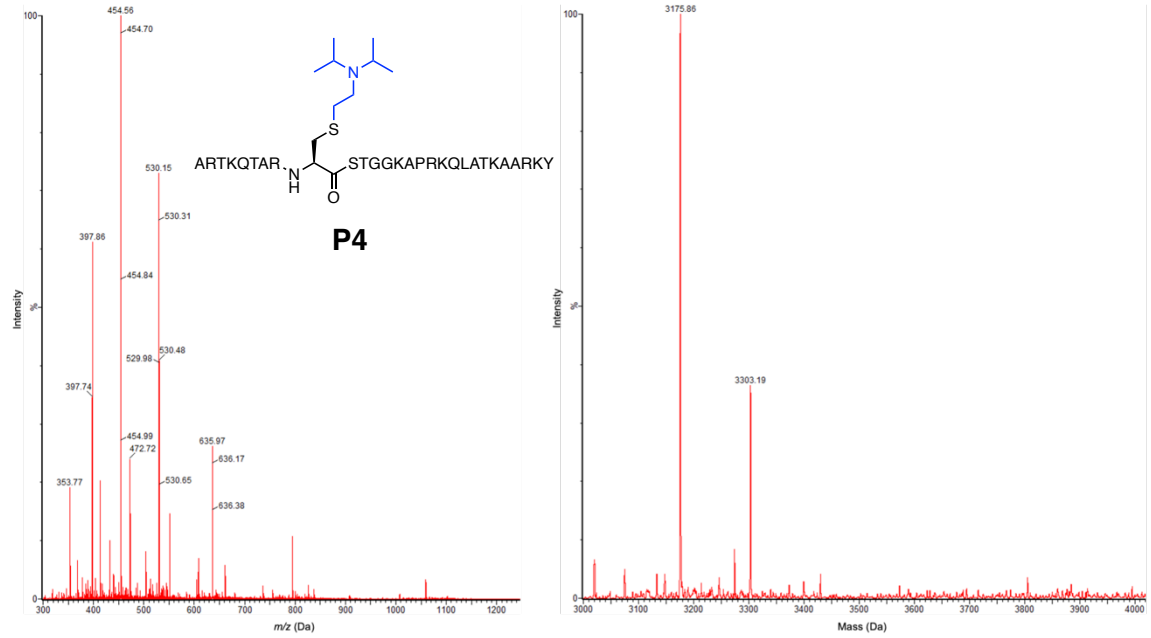

Observed MW: 3134.76 Da; Theoretical MW: 3134.73 Da



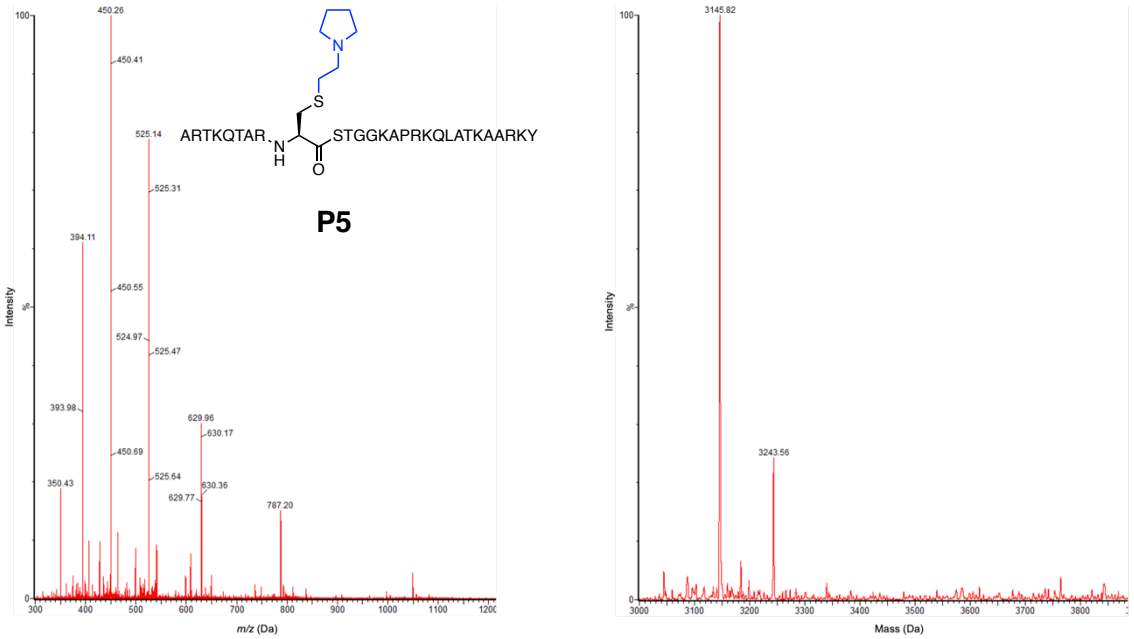

Observed MW: 3119.93 Da; Theoretical MW: 3119.70 Da



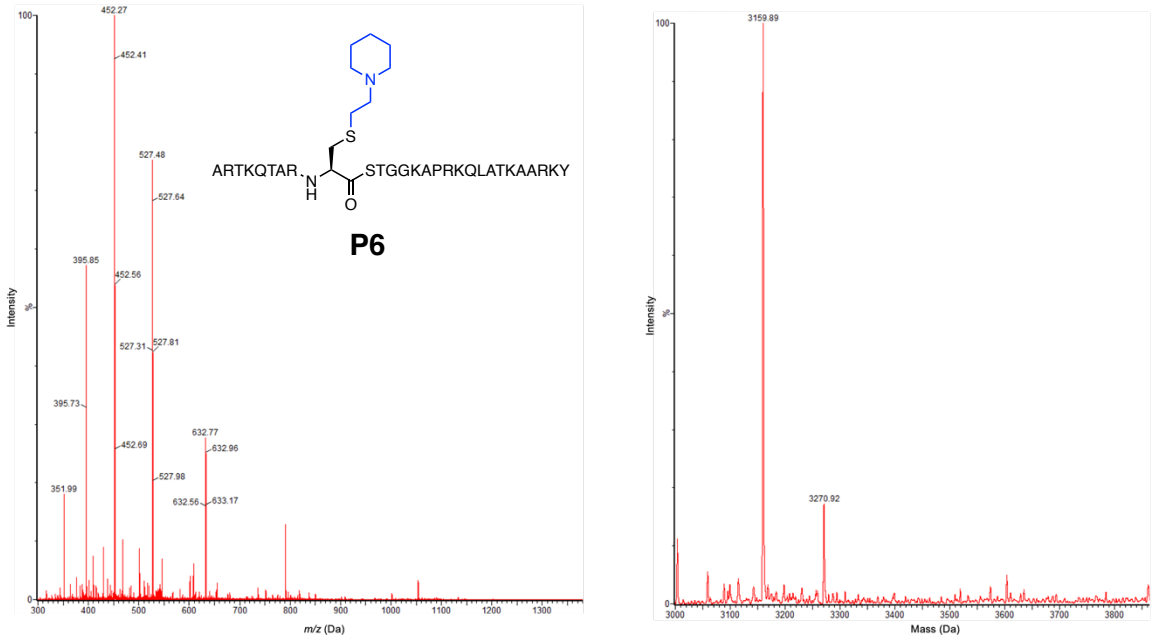

Observed MW: 3147.90 Da; Theoretical MW: 3147.75 Da



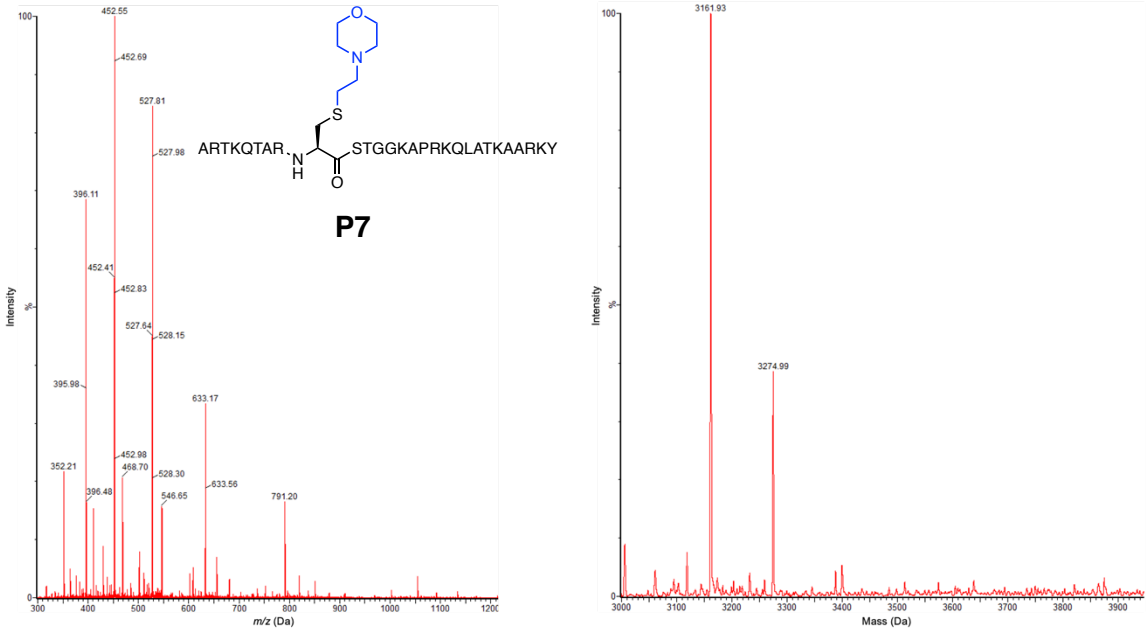

Observed MW: 3175.86 Da; Theoretical MW: 3175.81 Da



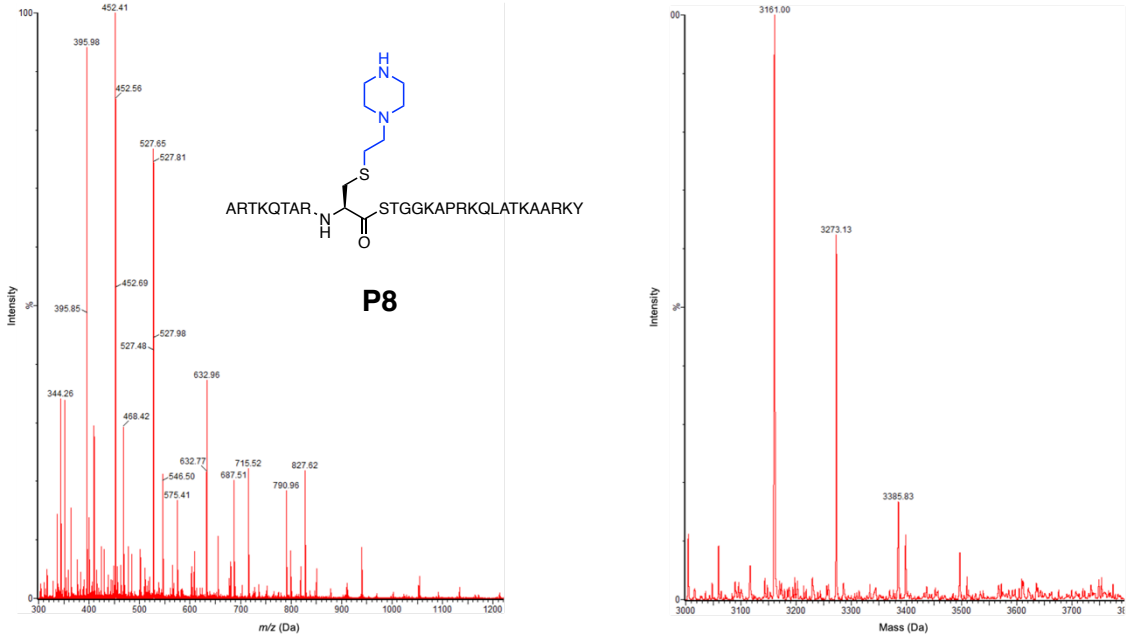

Observed MW: 3145.82 Da; Theoretical MW: 3145.74 Da



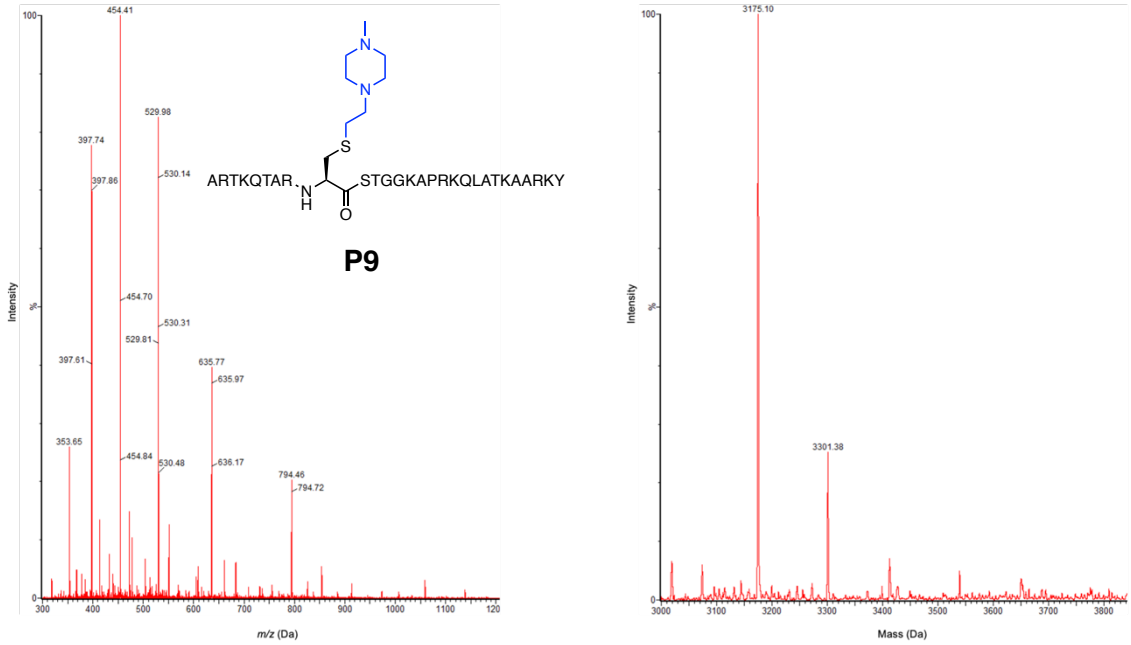

Observed MW: 3159.89 Da; Theoretical MW: 3159.76 Da



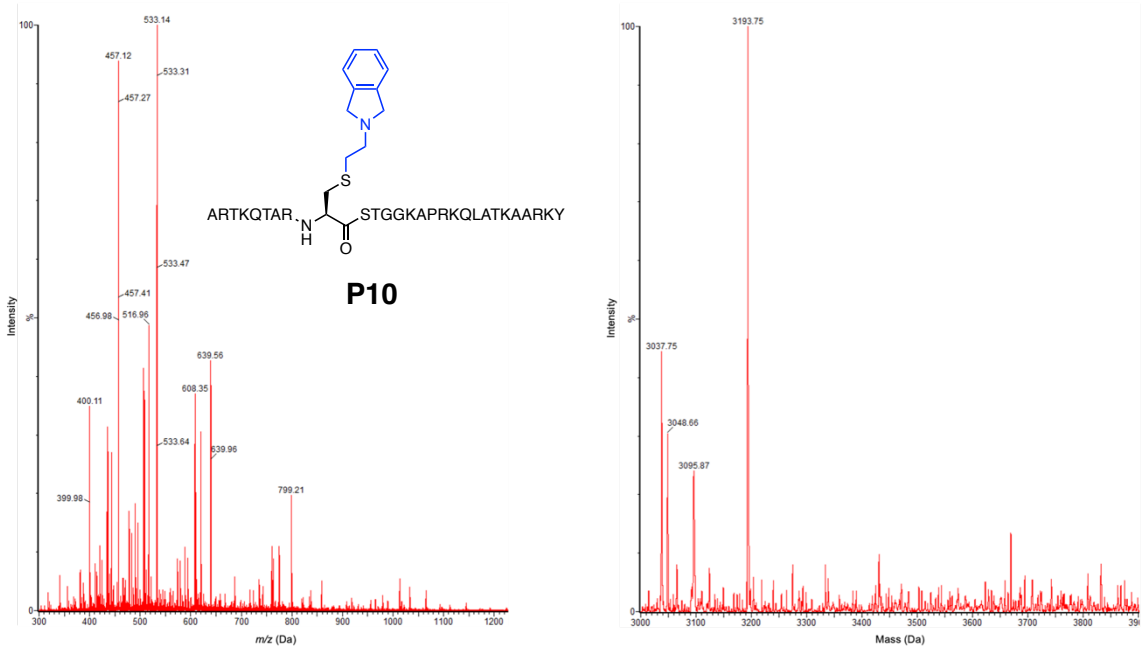

Observed MW: 3161.93 Da; Theoretical MW: 3161.74 Da



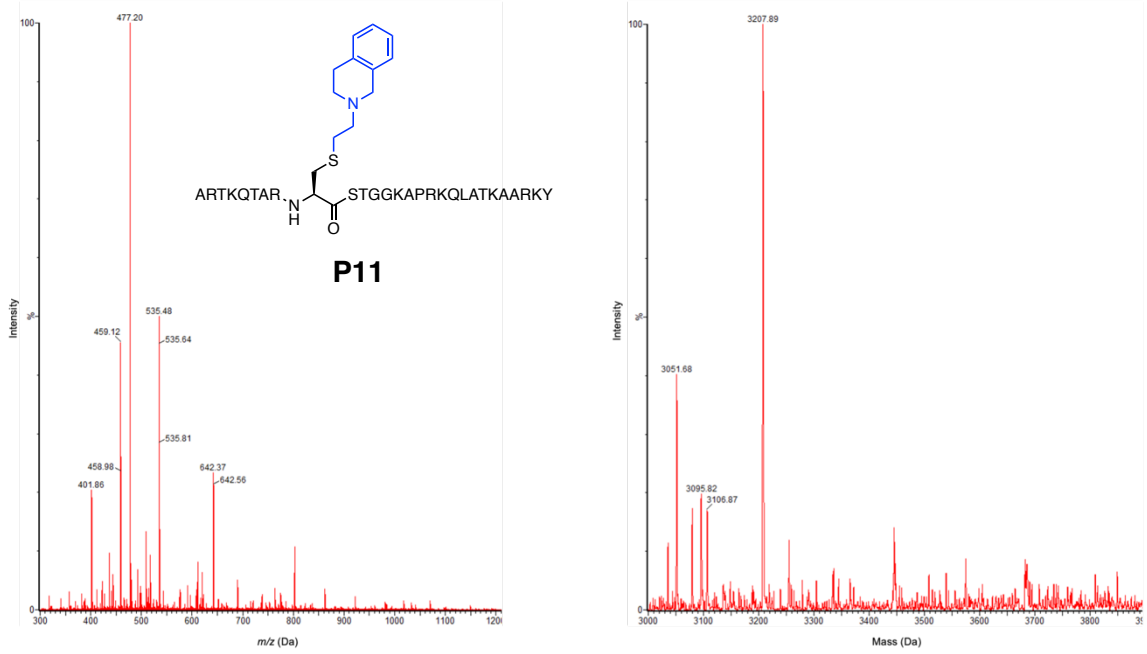

Observed MW: 3161.00 Da; Theoretical MW: 3160.75 Da



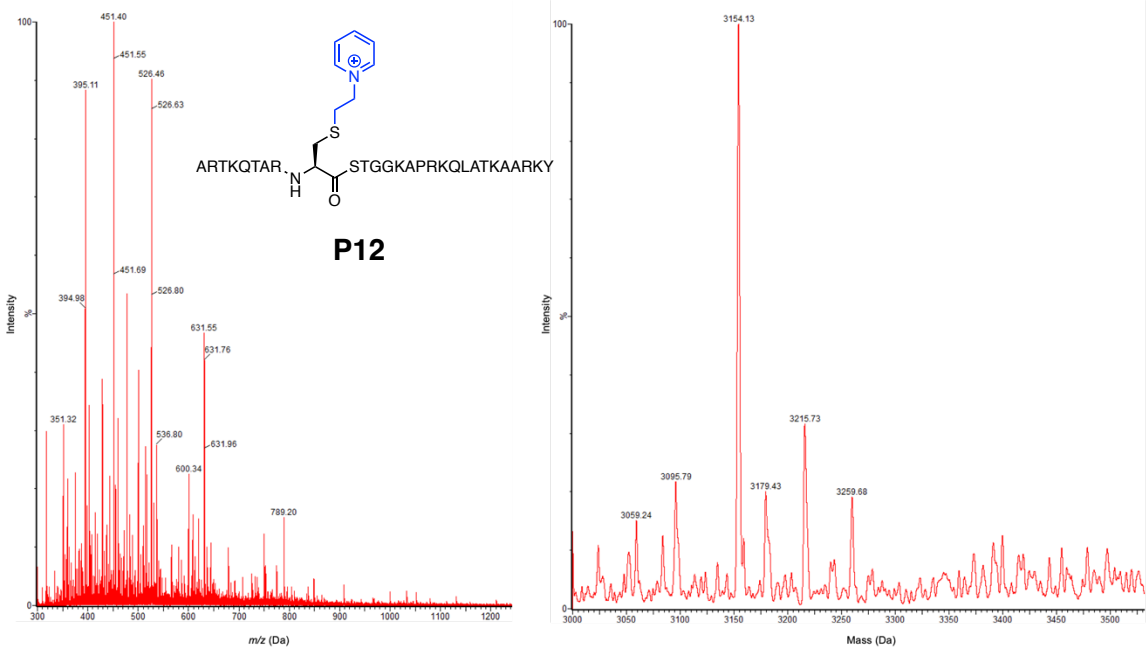

Observed MW: 3175.10 Da; Theoretical MW: 3174.78 Da



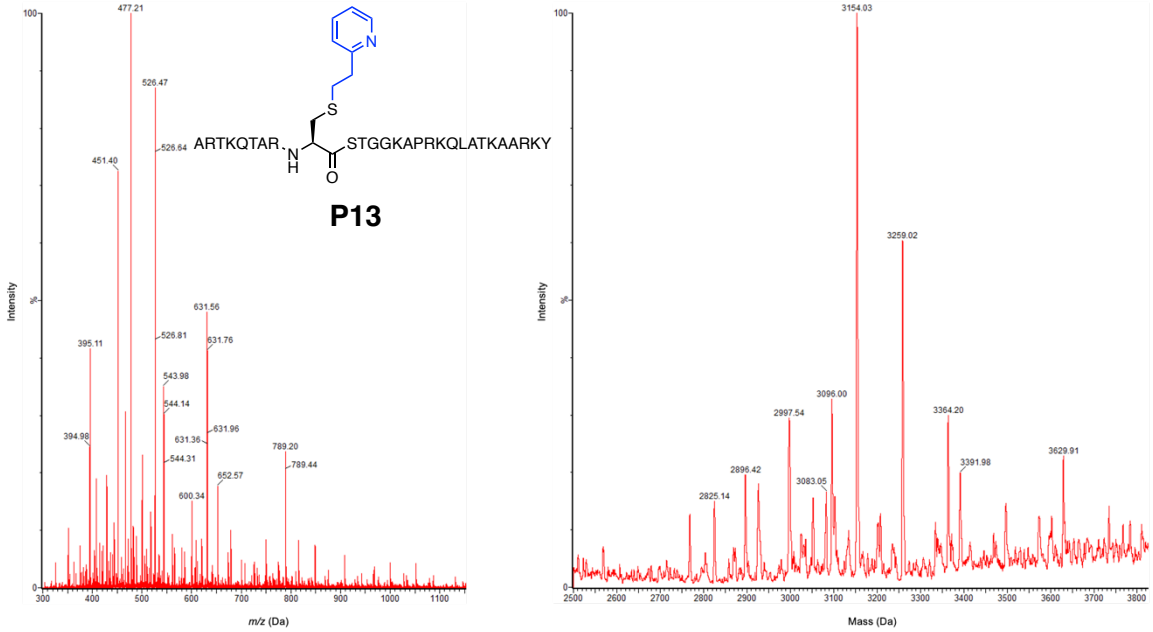

Observed MW: 3193.75 Da; Theoretical MW: 3193.78 Da



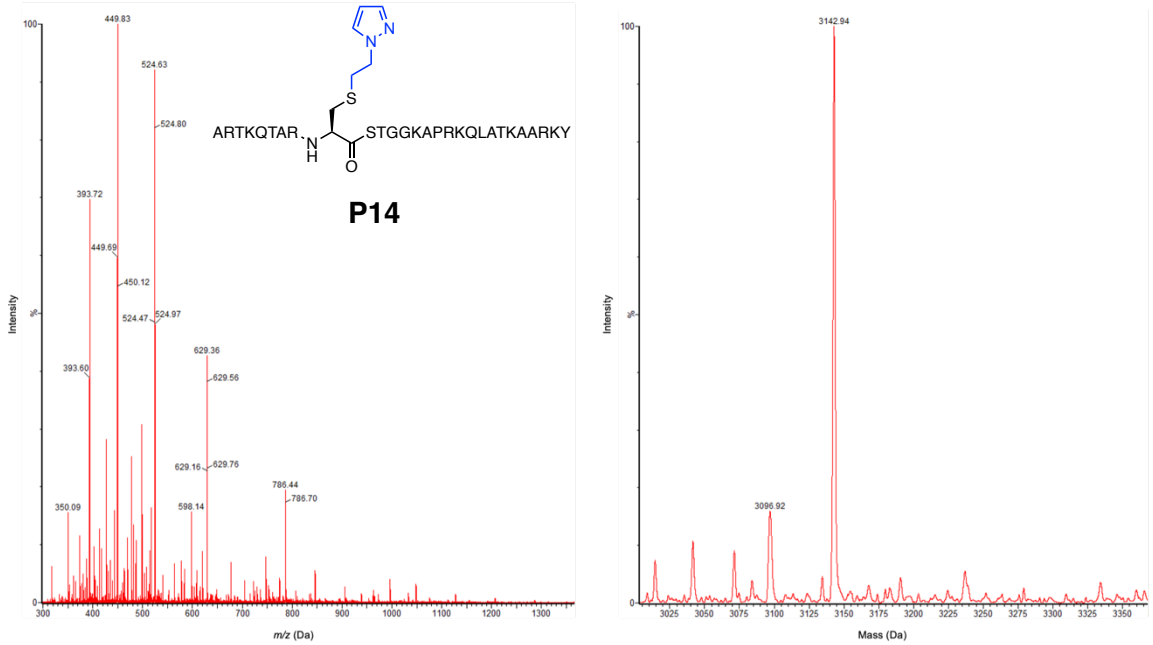

Observed MW: 3207.89 Da; Theoretical MW: 3207.81 Da



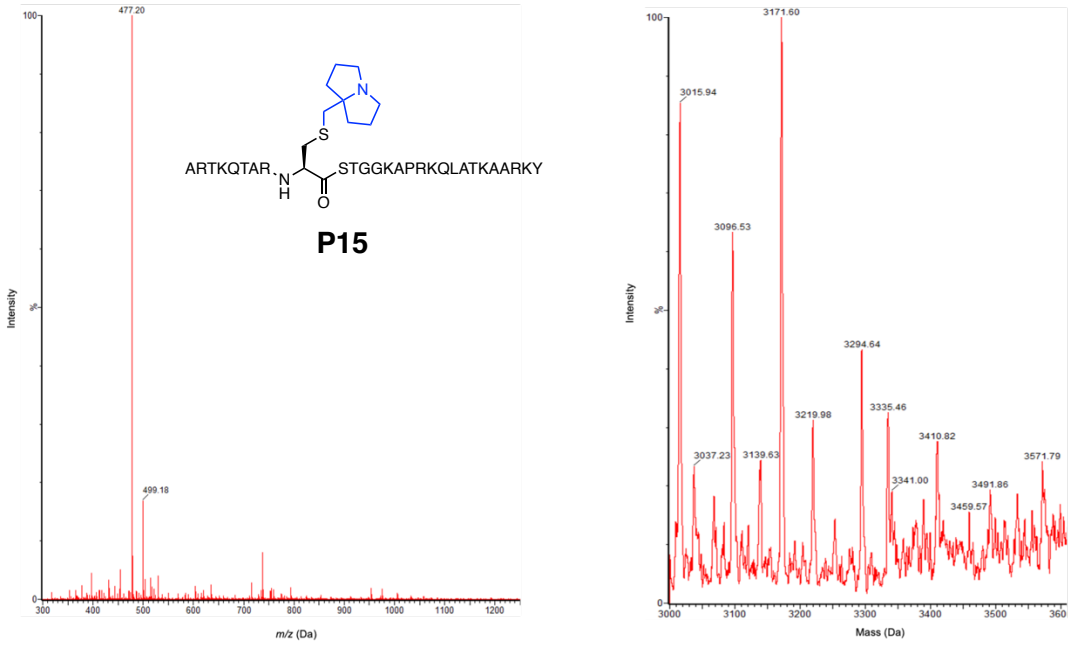

Observed MW: 3154.13 Da; Theoretical MW: 3154.72 Da



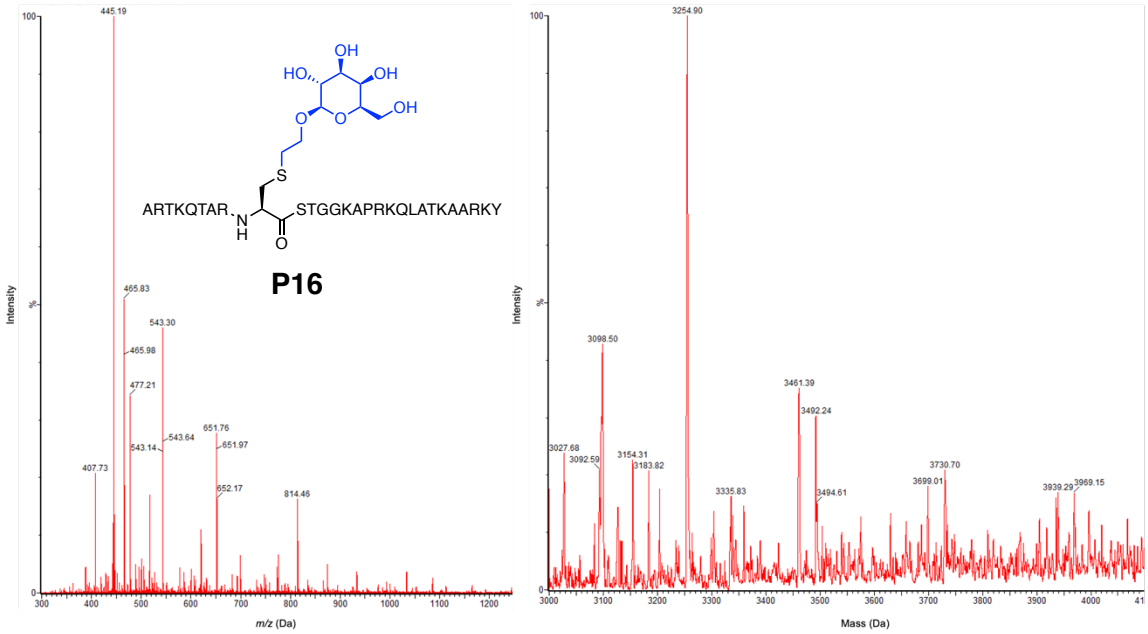

Observed MW: 3154.03 Da; Theoretical MW: 3153.72 Da




Observed MW: 3142.94 Da; Theoretical MW: 3142.69 Da




Observed MW: 3171.60 Da; Theoretical MW: 3171.78 Da




Observed MW: 3254.90 Da; Theoretical MW: 3254.77 Da

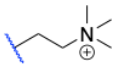
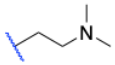
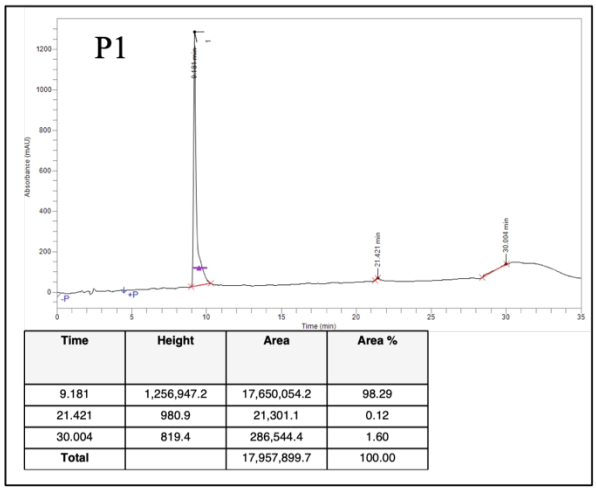
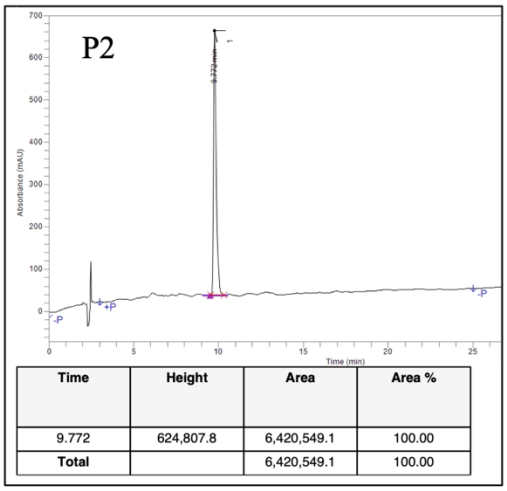
| Peptide and Modification | Fragment Peptide Amino Acid Sequences | *m/z* | Charge (*z*) | Theoretical Mass (Da) | Error (ppm) | -10log(*p*-value) | Modification sites |
|---|---|---|---|---|---|---|---|
| P1  | QTARC(+86.10)STGGK | 365.5308 | 3 | 1093.5789 | -5.5 | 42.68 | C9 |
| | APRKQLATK | 506.8168 | 2 | 1011.6189 | 2.3 | 65.33 | – |
| | QLATKAARKY | 575.3401 | 2 | 1148.6665 | 1.4 | 69.04 | – |
| | QLATKAARKY | 383.8959 | 3 | 1148.6665 | 1.6 | 68.09 | – |
| | QTARCSTGGK | 504.7487 | 2 | 1007.4818 | 3.1 | 44.98 | – |
| P2  | QLATKAARKY | 575.3378 | 2 | 1148.6665 | -4.8 | 96.44 | – |
| | QLATKAARKY | 383.8938 | 3 | 1148.6665 | -6.0 | 63.75 | – |
| | ARTKQTARC(+71.07)STGGK | 384.7150 | 4 | 1534.8362 | -3.5 | 70.29 | C9 |
| | ARTKQTARC(+71.07)STGGK | 512.6183 | 3 | 1534.8362 | -2.1 | 66.64 | C9 |
| | QTARCSTGGK | 504.7491 | 2 | 1007.4818 | 1.7 | 60.67 | – |
| | QTARC(+71.07)STGGKAPRK | 383.7159 | 4 | 1530.8413 | -4.3 | 57.17 | C9 |
| | QTARC(+71.07)STGGKAPRK | 766.4248 | 2 | 1530.8413 | -4.1 | 49.14 | C9 |
| | QTARC(+71.07)STGGK | 540.2843 | 2 | 1078.5553 | -1.2 | 56.96 | C9 |
| | QTARC(+71.07)STGGK | 360.5251 | 3 | 1078.5553 | -1.8 | 52.59 | C9 |
| | A(+71.07)(+71.07)RTKQTARC(+71.07)STGGK | 420.2516 | 4 | 1676.9832 | -3.4 | 50.82 | A1 (N-terminus) ×2, C9 |
| | APRKQLATK | 506.8165 | 2 | 1011.6189 | -0.4 | 46.74 | – |
| | APRKQLATK(+71.07) | 361.9046 | 3 | 1082.6924 | -0.4 | 46.61 | K23 |
| | QLATK(+71.07)AARKY | 610.8752 | 2 | 1219.7400 | -3.3 | 39.84 | K23 |
| | QTARC(+71.07)S(+71.07)T(+71.07)GGK | 407.9077 | 3 | 1220.7023 | -0.8 | 45.20 | C9, S10, T11 |
| | QLAT(+71.07)KAARKY | 610.8772 | 2 | 1219.7400 | -0.1 | 45.07 | T22 |
| | QLAT(+71.07)K(+71.07)AARKY | 431.2800 | 3 | 1290.8136 | 3.5 | 41.3 | T22, K23 |
| | APRKQLAT(+71.07)K(+71.07) | 385.5951 | 3 | 1153.7659 | -2.1 | 38.84 | T22, K23 |
| | QTARC(+71.07)S(+71.07)T(+71.07)GGKAPRK | 419.2528 | 4 | 1672.9883 | -3.7 | 36.88 | C9, S10, T11 |
| | QTARC(+71.07)S(+71.07)TGGK | 575.8159 | 2 | 1149.6288 | -10.1 | 33.86 | C9, S10 |

**Figure A3.2:** NanoLC-MS/MS data for peptides **P1** and **P2**. Data were processed by Marjorie Fournier uisng PEAKS software. The amino acid sequence for the unalkylated $H3_{(1\text{-}27)}$K9C peptide is ARTKQTARCSTGGKAPRKQLATKAARKY. Sequence coverage: 86% (**P1**), 100% (**P2**).

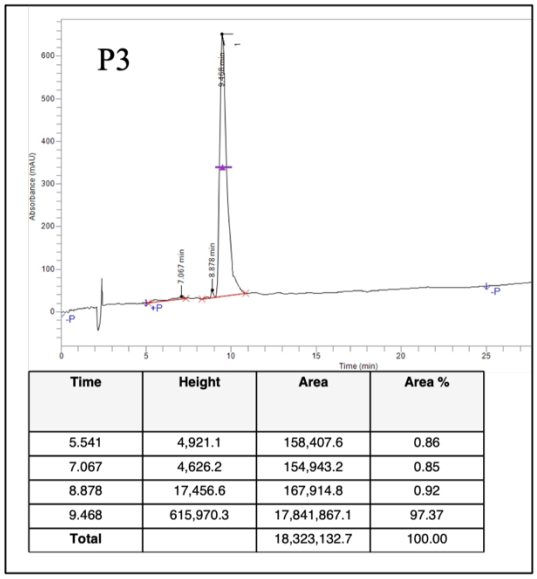
**Figure A3.3:** Analytical HPLC chromatograms for peptides **P1**–**P16** [5–65% MeCN/$H_2O$ (+0.1% TFA), 25 min, 220 nm].



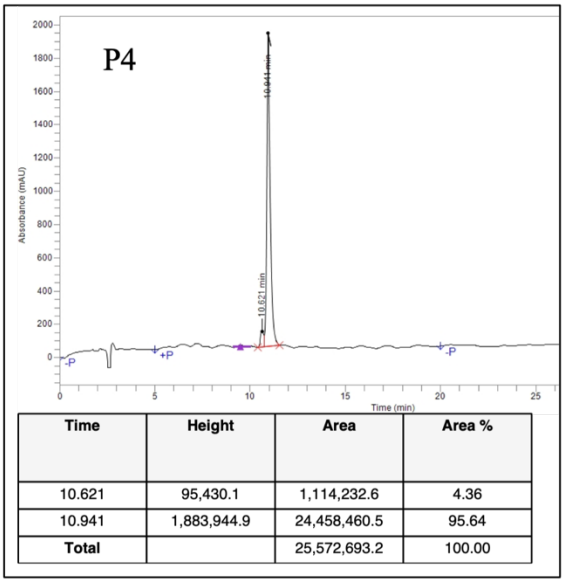

| Time | Height | Area | Area % |
|---|---|---|---|
| 9.181 | 1,256,947.2 | 17,650,054.2 | 98.29 |
| 21.421 | 980.9 | 21,301.1 | 0.12 |
| 30.004 | 819.4 | 286,544.4 | 1.60 |
| **Total** | | 17,957,899.7 | 100.00 |



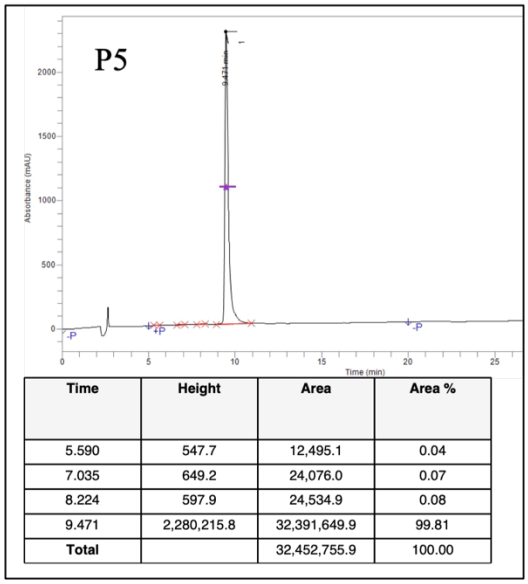

| Time | Height | Area | Area % |
|---|---|---|---|
| 9.772 | 624,807.8 | 6,420,549.1 | 100.00 |
| **Total** | | 6,420,549.1 | 100.00 |



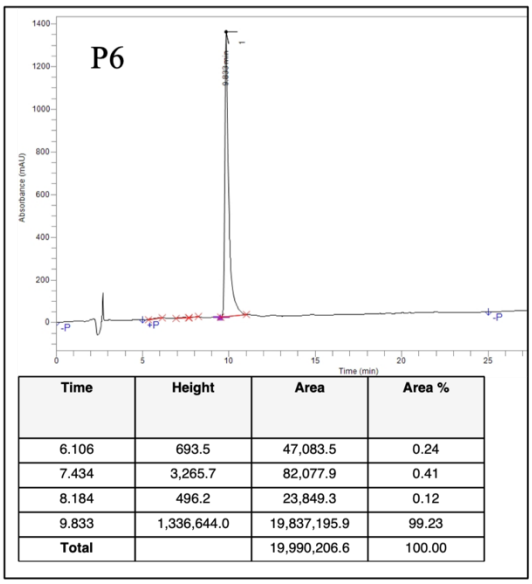

| Time | Height | Area | Area % |
|---|---|---|---|
| 5.541 | 4,921.1 | 158,407.6 | 0.86 |
| 7.067 | 4,626.2 | 154,943.2 | 0.85 |
| 8.878 | 17,456.6 | 167,914.8 | 0.92 |
| 9.468 | 615,970.3 | 17,841,867.1 | 97.37 |
| **Total** | | 18,323,132.7 | 100.00 |




| Time | Height | Area | Area % |
|---|---|---|---|
| 10.621 | 95,430.1 | 1,114,232.6 | 4.36 |
| 10.941 | 1,883,944.9 | 24,458,460.5 | 95.64 |
| **Total** | | 25,572,693.2 | 100.00 |

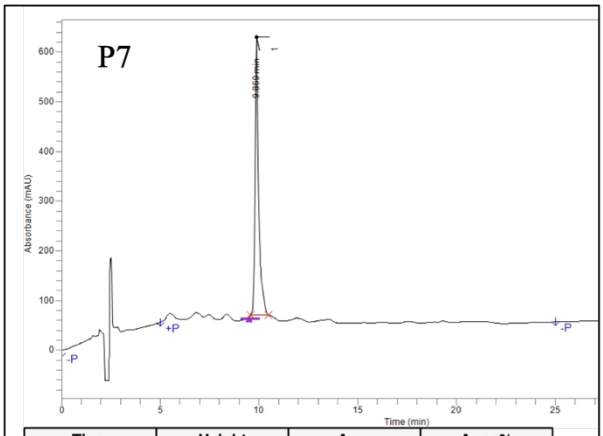



| Time | Height | Area | Area % |
|---|---|---|---|
| 5.590 | 547.7 | 12,495.1 | 0.04 |
| 7.035 | 649.2 | 24,076.0 | 0.07 |
| 8.224 | 597.9 | 24,534.9 | 0.08 |
| 9.471 | 2,280,215.8 | 32,391,649.9 | 99.81 |
| **Total** | | 32,452,755.9 | 100.00 |



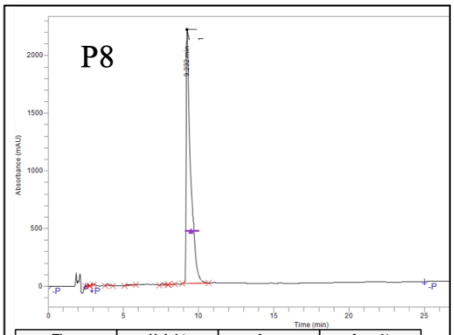

| Time | Height | Area | Area % |
|---|---|---|---|
| 6.106 | 693.5 | 47,083.5 | 0.24 |
| 7.434 | 3,265.7 | 82,077.9 | 0.41 |
| 8.184 | 496.2 | 23,849.3 | 0.12 |
| 9.833 | 1,336,644.0 | 19,837,195.9 | 99.23 |
| **Total** | | 19,990,206.6 | 100.00 |

P7



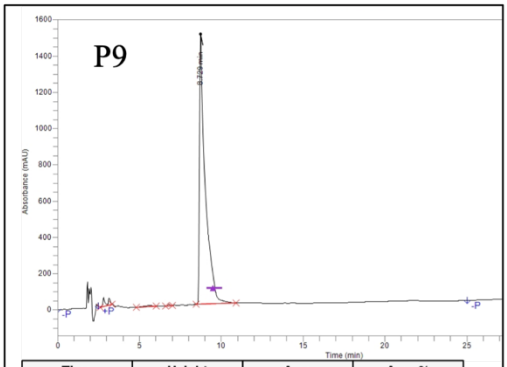

| Time | Height | Area | Area % |
|---|---|---|---|
| 9.869 | 560,452.2 | 6,591,956.1 | 100.00 |
| **Total** | | 6,591,956.1 | 100.00 |

P8



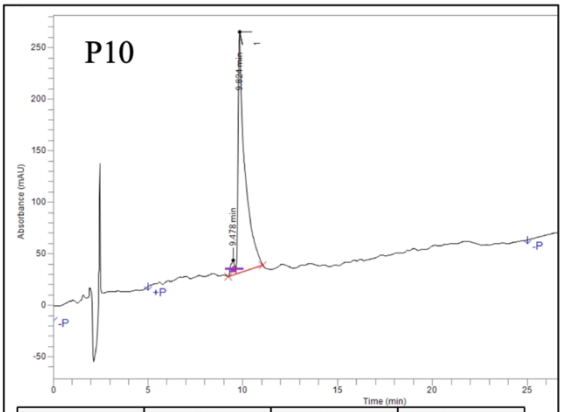

| Time | Height | Area | Area % |
|---|---|---|---|
| 2.672 | 9,039.5 | 52,107.3 | 0.11 |
| 2.940 | 11,324.4 | 60,222.9 | 0.13 |
| 3.874 | 13,471.5 | 198,045.0 | 0.43 |
| 5.228 | 7,737.0 | 182,846.8 | 0.40 |
| 7.576 | 16,818.4 | 220,087.2 | 0.48 |
| 8.282 | 5,316.0 | 49,451.8 | 0.11 |
| 9.232 | 2,201,790.5 | 45,419,643.5 | 98.35 |
| **Total** | | 46,182,404.5 | 100.00 |

P9



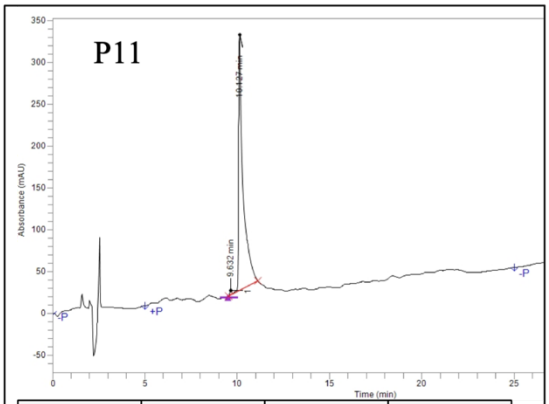

| Time | Height | Area | Area % |
|---|---|---|---|
| 2.816 | 47,132.1 | 358,806.4 | 0.93 |
| 3.156 | 35,299.4 | 287,543.7 | 0.75 |
| 5.650 | 6,711.2 | 258,025.8 | 0.67 |
| 6.803 | 10,701.6 | 91,811.1 | 0.24 |
| 8.729 | 1,490,732.7 | 37,490,307.3 | 97.41 |
| **Total** | | 38,486,494.2 | 100.00 |

P10



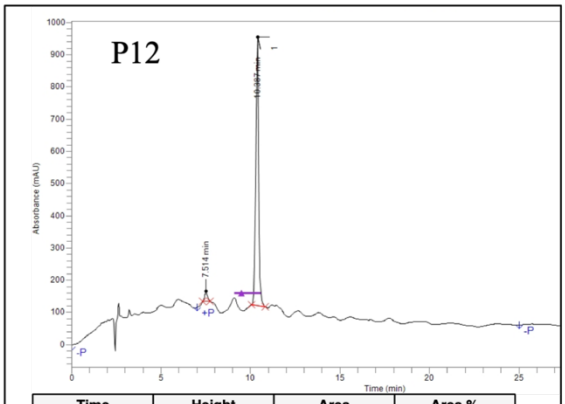

| Time | Height | Area | Area % |
|---|---|---|---|
| 9.478 | 14,476.5 | 202,959.9 | 3.50 |
| 9.824 | 234,152.7 | 5,594,026.3 | 96.50 |
| **Total** | | 5,796,986.2 | 100.00 |

P11

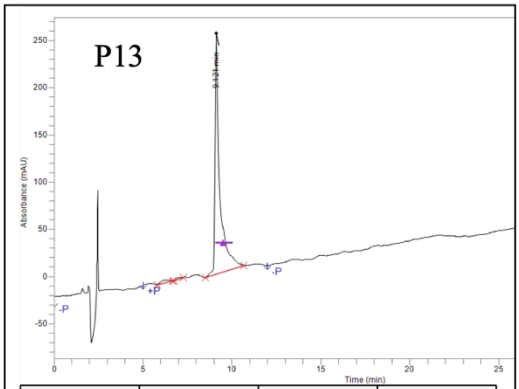



| Time | Height | Area | Area % |
|---|---|---|---|
| 9.632 | 6,125.8 | 43,353.7 | 0.83 |
| 10.127 | 305,888.9 | 5,157,467.5 | 99.17 |
| **Total** | | 5,200,821.1 | 100.00 |

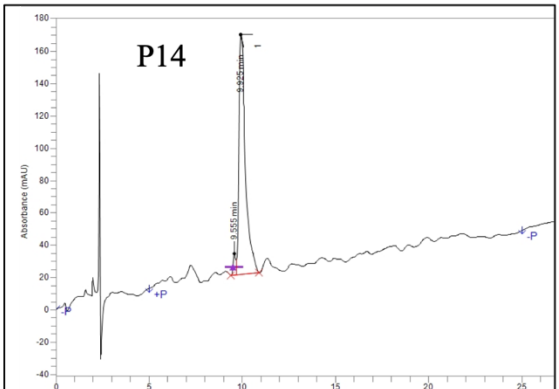
P12



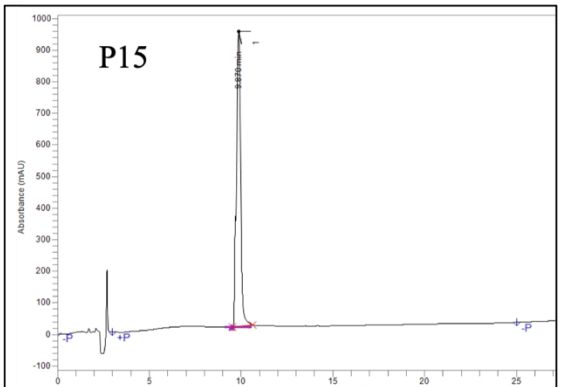

| Time | Height | Area | Area % |
|---|---|---|---|
| 7.514 | 31,089.7 | 361,834.1 | 4.05 |
| 10.387 | 835,331.8 | 8,565,068.8 | 95.95 |
| **Total** | | 8,926,902.9 | 100.00 |



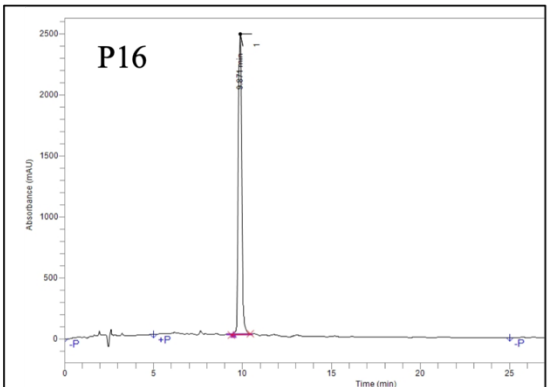

| Time | Height | Area | Area % |
|---|---|---|---|
| 6.456 | 1,963.3 | 88,594.8 | 1.75 |
| 7.032 | 1,621.6 | 27,938.1 | 0.55 |
| 9.121 | 255,656.3 | 4,939,819.9 | 97.70 |
| **Total** | | 5,056,352.9 | 100.00 |




| Time | Height | Area | Area % |
|---|---|---|---|
| 9.555 | 13,332.1 | 135,959.7 | 3.67 |
| 9.925 | 148,701.8 | 3,568,829.3 | 96.33 |
| **Total** | | 3,704,788.9 | 100.00 |

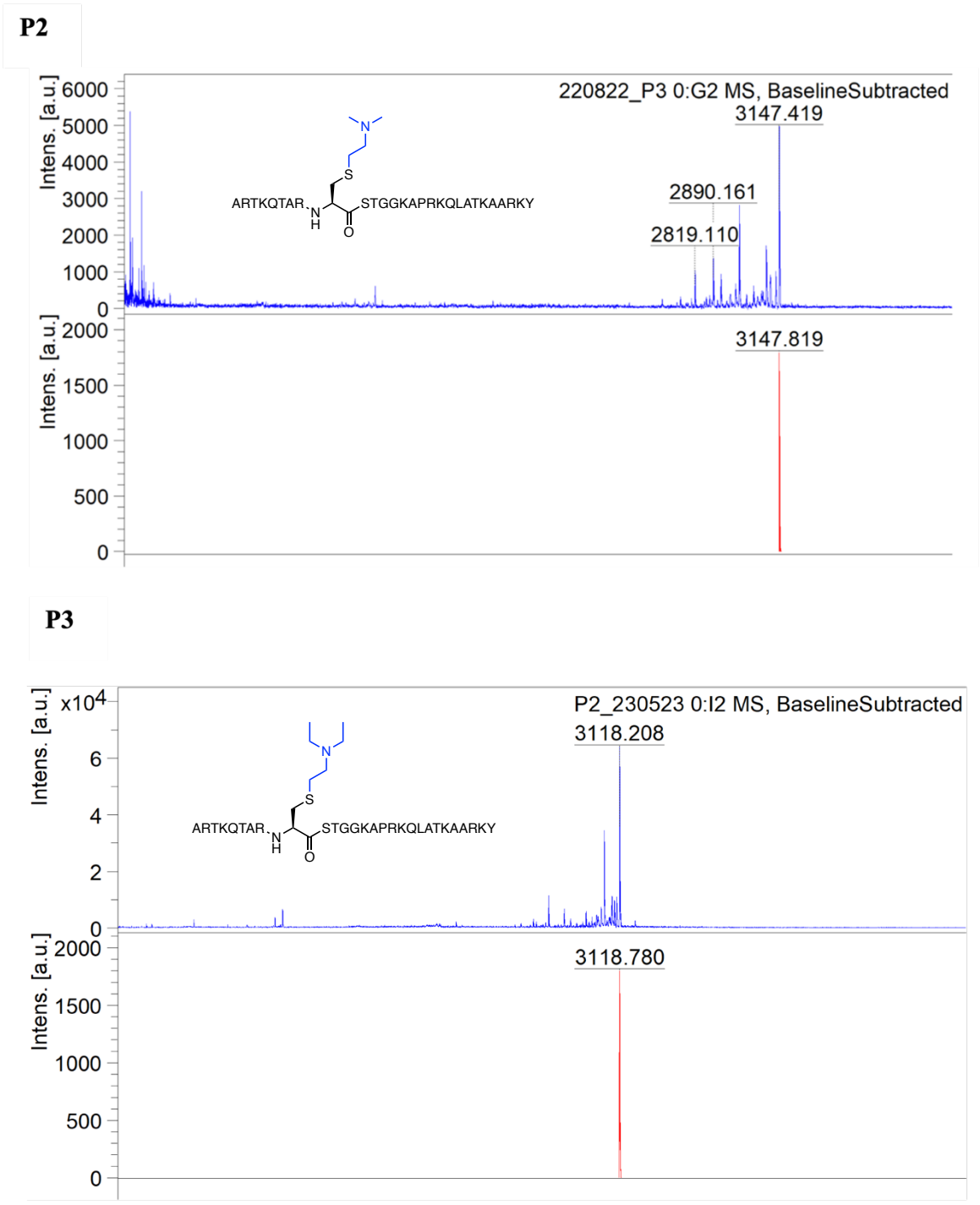



| Time | Height | Area | Area % |
|---|---|---|---|
| 9.870 | 934,087.5 | 14,600,122.2 | 100.00 |
| **Total** | | 14,600,122.2 | 100.00 |

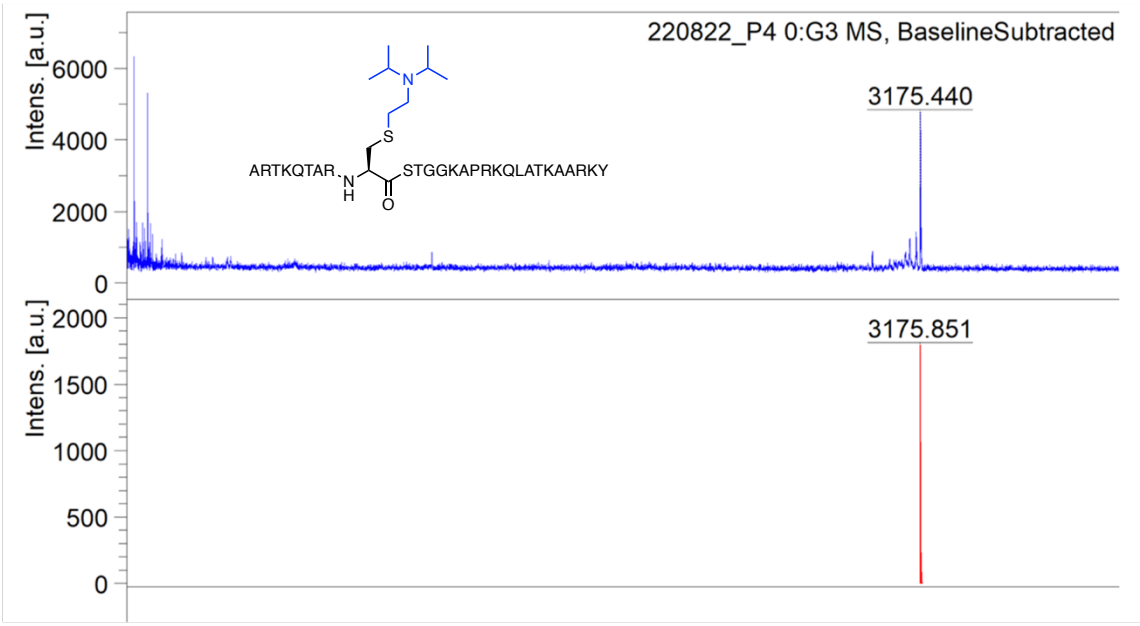



| Time | Height | Area | Area % |
|---|---|---|---|
| 9.871 | 2,462,217.8 | 31,848,986.0 | 100.00 |
| **Total** | | 31,848,986.0 | 100.00 |

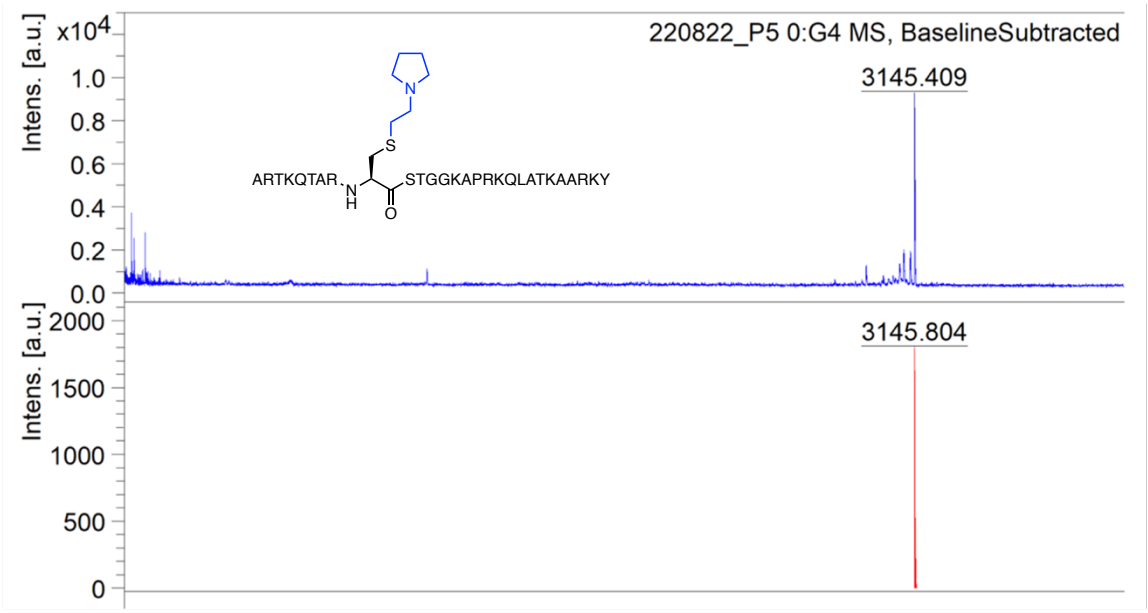
**Figure A3.4:** MALDI-TOF MS (blue) and simulated spectra (red, $[M+H]^+$) for the histone-mimicking peptides after HPLC purification. Spectra were simulated in Bruker flexAnalysis software. **P1** and **P12** masses could not be observed by MALDI-TOF MS. Structures of the peptides are shown with single letter amino acid codes. The modification is shown in blue.



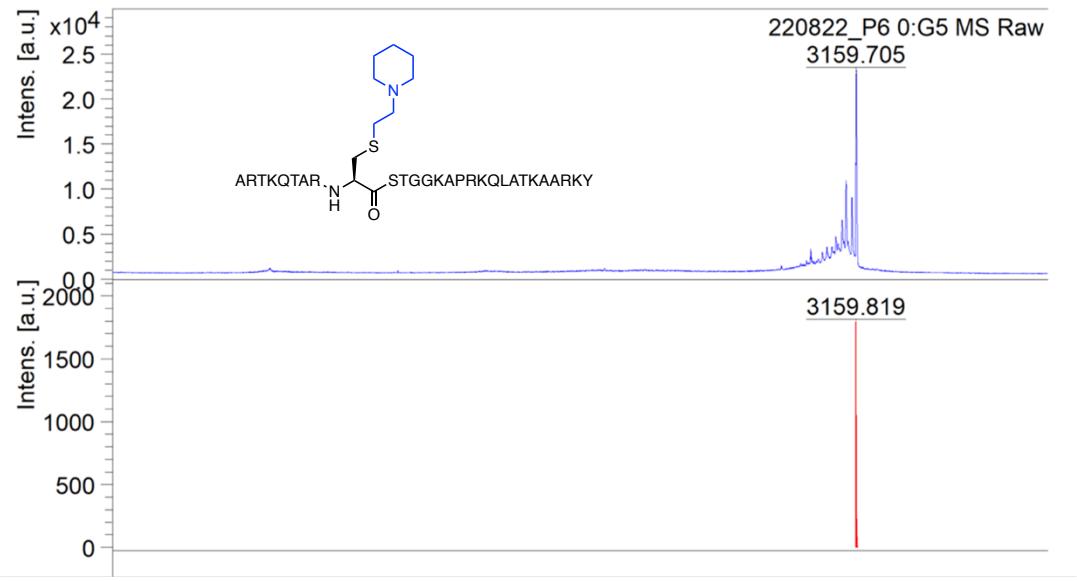
**P4**


220822_P4 0:G3 MS, BaselineSubtracted
Intens. [a.u.]
6000
4000
2000
0
3175.440
ARTKQTAR
STGGKAPRKQLATKAARKY
Intens. [a.u.]
2000
1500
1000
500
0
3175.851

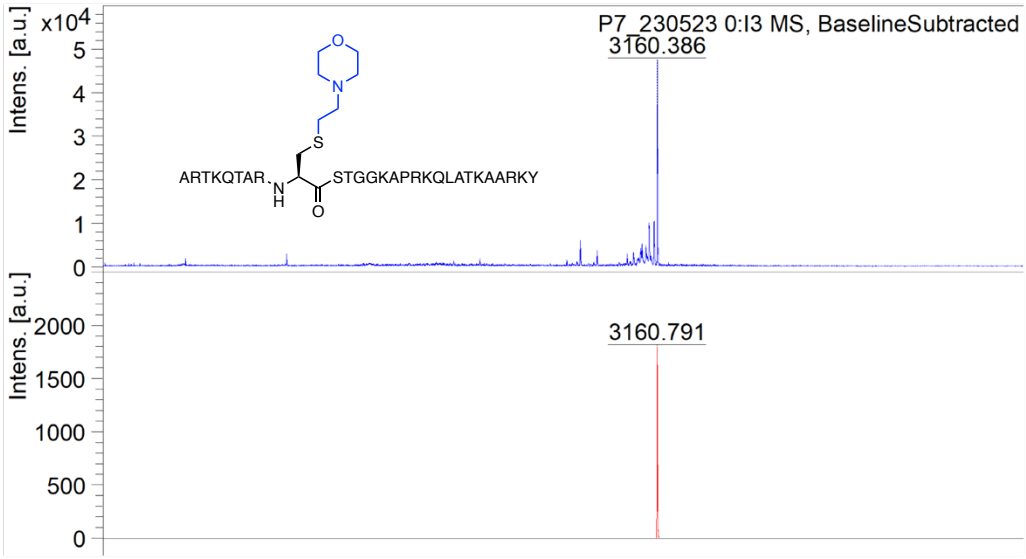
**P5**


220822_P5 0:G4 MS, BaselineSubtracted
Intens. [a.u.]
x10^4
1.0
0.8
0.6
0.4
0.2
0.0
3145.409
ARTKQTAR
STGGKAPRKQLATKAARKY
Intens. [a.u.]
2000
1500
1000
500
0
3145.804

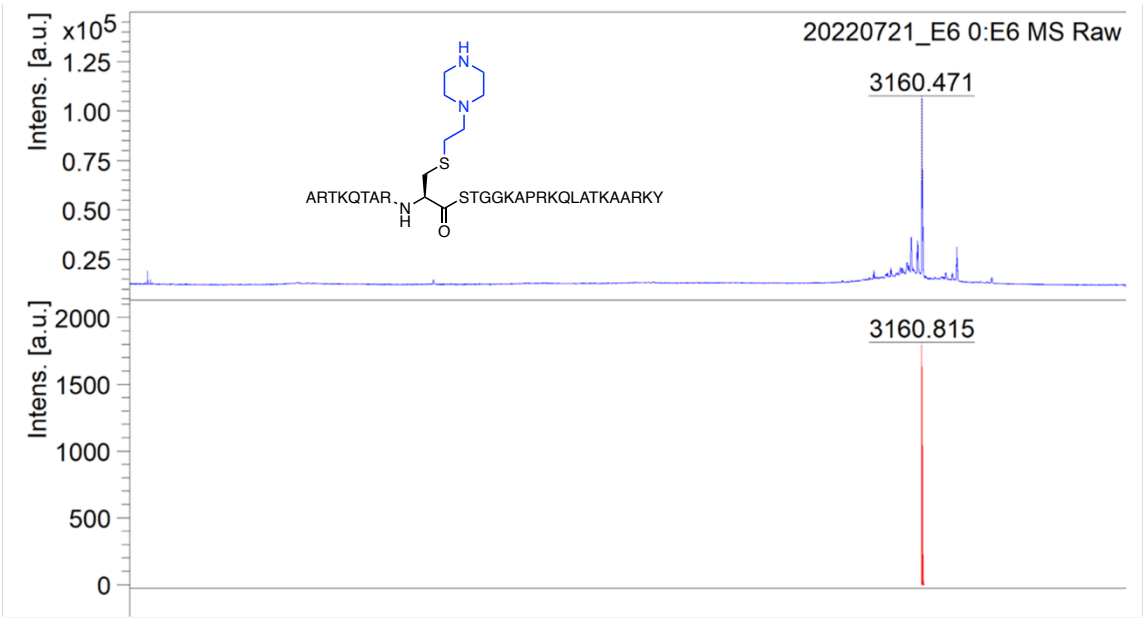
**P6**


220822_P6 0:G5 MS Raw
3159.705
3159.819
Intens. [a.u.]
x10^4
ARTKQTAR
STGGKAPRKQLATKAARKY

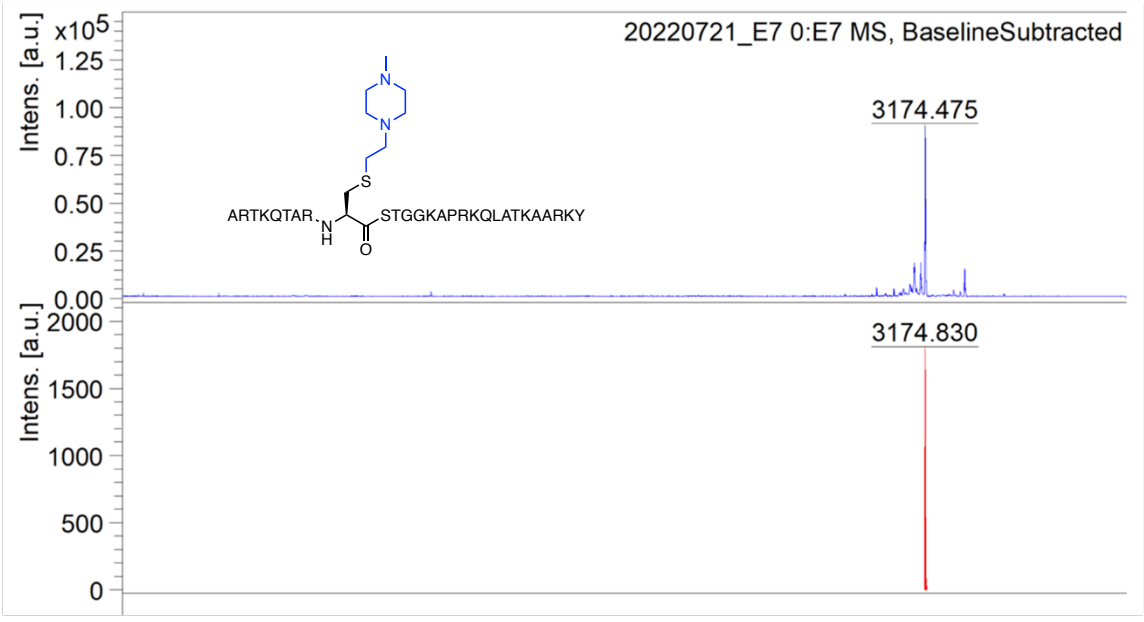
**P7**


P7_230523 0:I3 MS, BaselineSubtracted
3160.386
3160.791
Intens. [a.u.]
x10^4
ARTKQTAR
STGGKAPRKQLATKAARKY

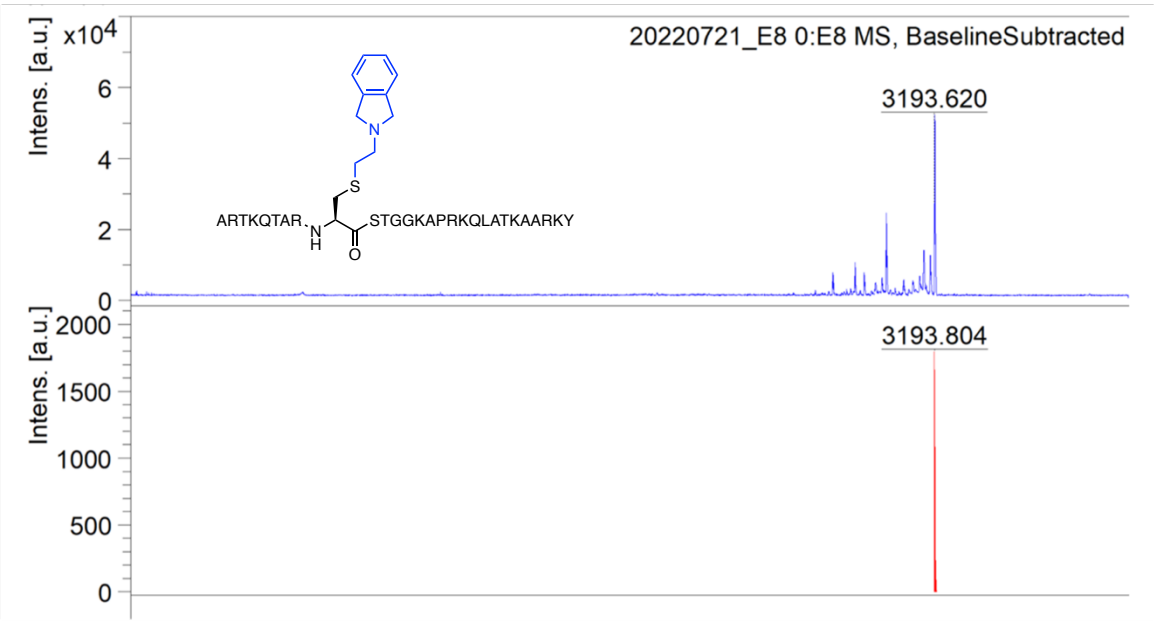
**P8**


20220721_E6 0:E6 MS Raw
Intens. [a.u.]
x10^5
1.25
1.00
0.75
0.50
0.25
3160.471
ARTKQTAR
STGGKAPRKQLATKAARKY
Intens. [a.u.]
2000
1500
1000
500
0
3160.815

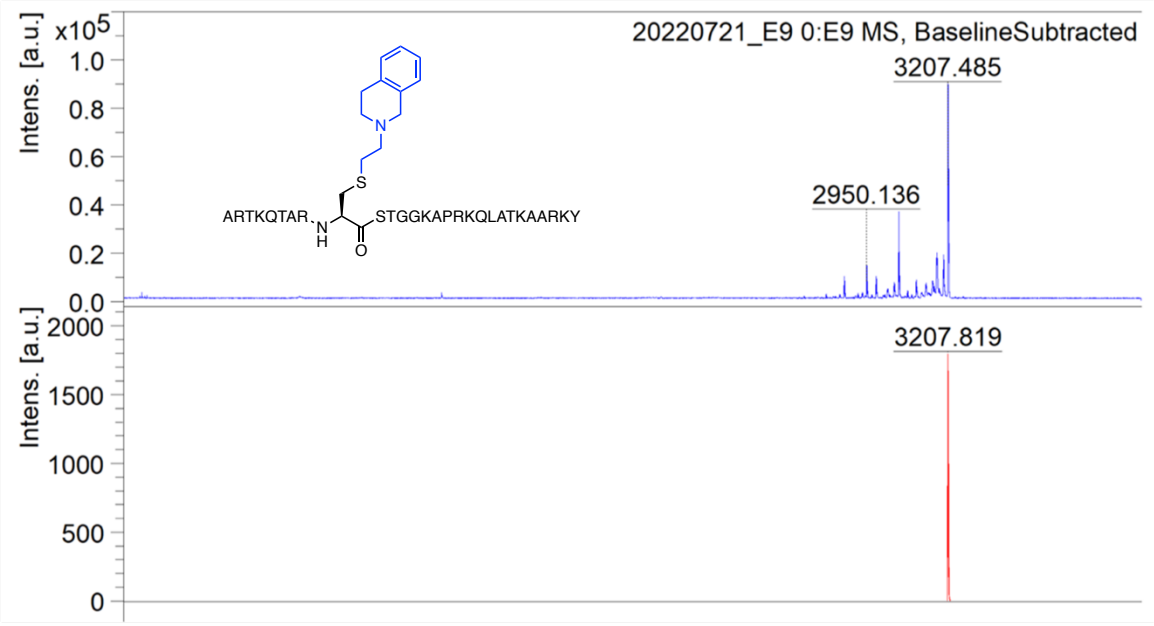
**P9**


20220721_E7 0:E7 MS, BaselineSubtracted
Intens. [a.u.]
x10^5
1.25
1.00
0.75
0.50
0.25
0.00
3174.475
ARTKQTAR
STGGKAPRKQLATKAARKY
Intens. [a.u.]
2000
1500
1000
500
0
3174.830

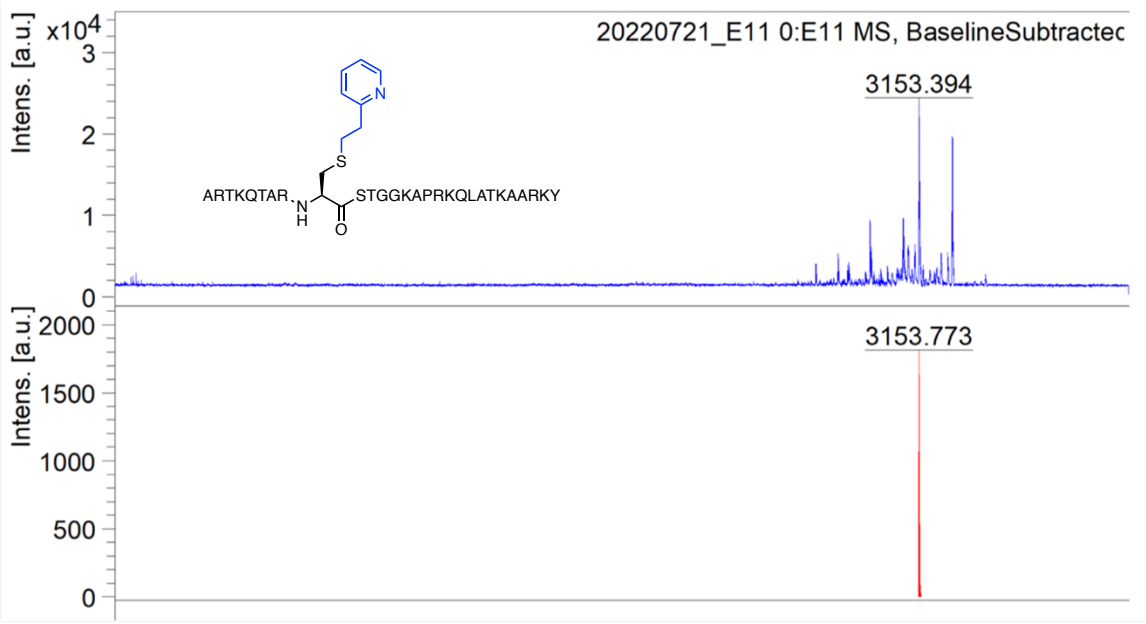
**P10**


20220721_E8 0:E8 MS, BaselineSubtracted
Intens. [a.u.]
x10^4
6
4
2
0
3193.620
ARTKQTAR
STGGKAPRKQLATKAARKY
Intens. [a.u.]
2000
1500
1000
500
0
3193.804

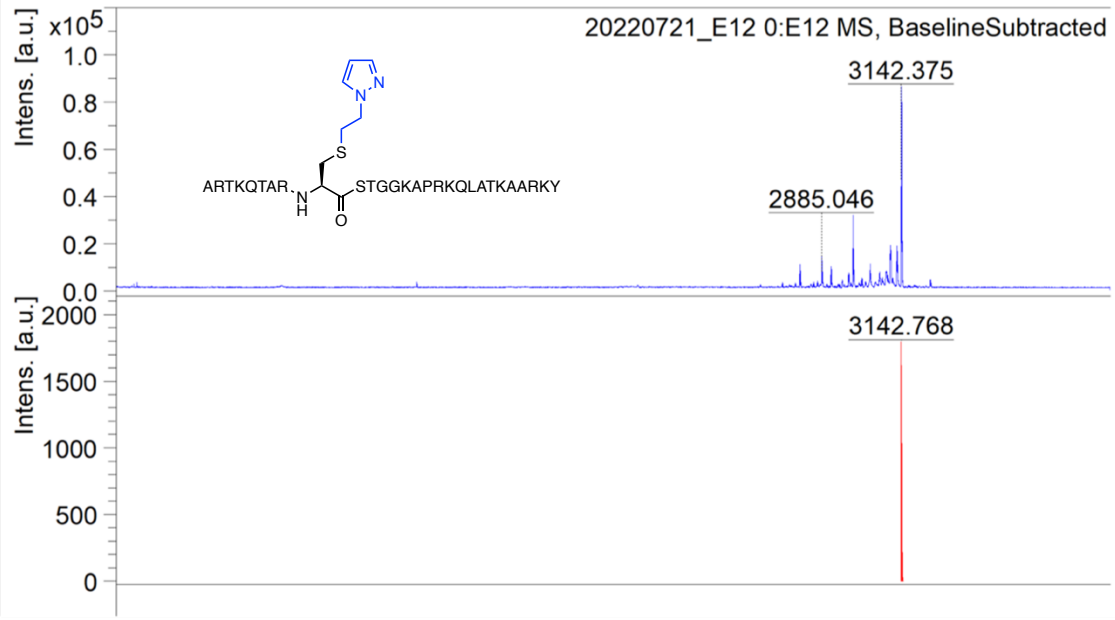
**P11**


20220721_E9 0:E9 MS, BaselineSubtracted
Intens. [a.u.]
x10^5
1.0
0.8
0.6
0.4
0.2
0.0
3207.485
2950.136
ARTKQTAR
STGGKAPRKQLATKAARKY
Intens. [a.u.]
2000
1500
1000
500
0
3207.819

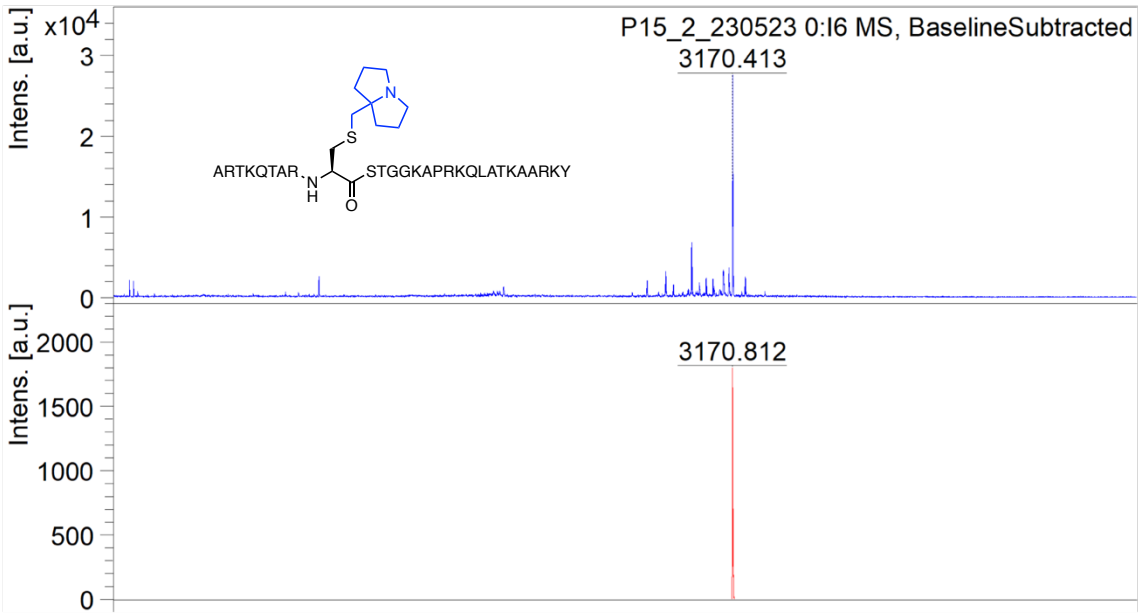
**P13**


20220721_E11 0:E11 MS, BaselineSubtractec
Intens. [a.u.]
x10^4
3
2
1
0
3153.394
ARTKQTAR
N
H
S
STGGKAPRKQLATKAARKY
O
Intens. [a.u.]
2000
1500
1000
500
0
3153.773

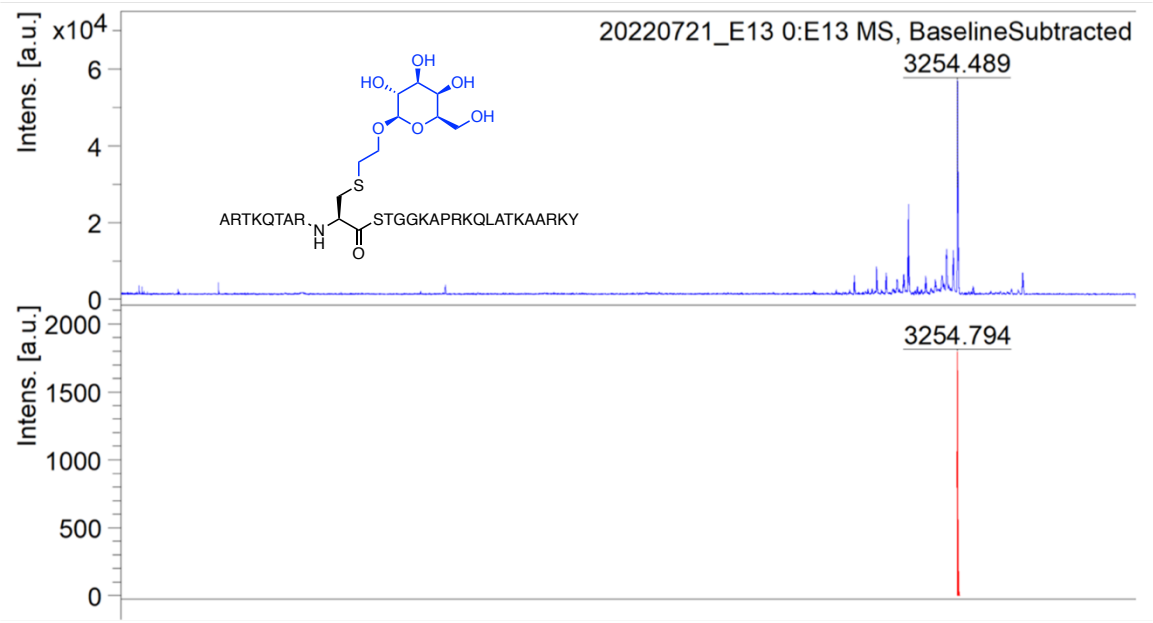
**P14**


20220721_E12 0:E12 MS, BaselineSubtracted
Intens. [a.u.]
x10^5
1.0
0.8
0.6
0.4
0.2
0.0
3142.375
2885.046
ARTKQTAR
N
H
S
STGGKAPRKQLATKAARKY
O
Intens. [a.u.]
2000
1500
1000
500
0
3142.768

**P15**


P15_2_230523 0:I6 MS, BaselineSubtracted
3170.413
3170.812
Intens. [a.u.]
x10^4
ARTKQTAR
STGGKAPRKQLATKAARKY

**P16**


20220721_E13 0:E13 MS, BaselineSubtracted
3254.489
3254.794
Intens. [a.u.]
x10^4
ARTKQTAR
STGGKAPRKQLATKAARKY

## Appendix 4 AlphaScreen®

**Figure A4.1:** AlphaScreen® signal response curves for biotinylated histone-mimicking peptides binding to $His_6$-TRIM24. Dose-response curves were obtained in triplicate with error bars indicating the standard deviation.

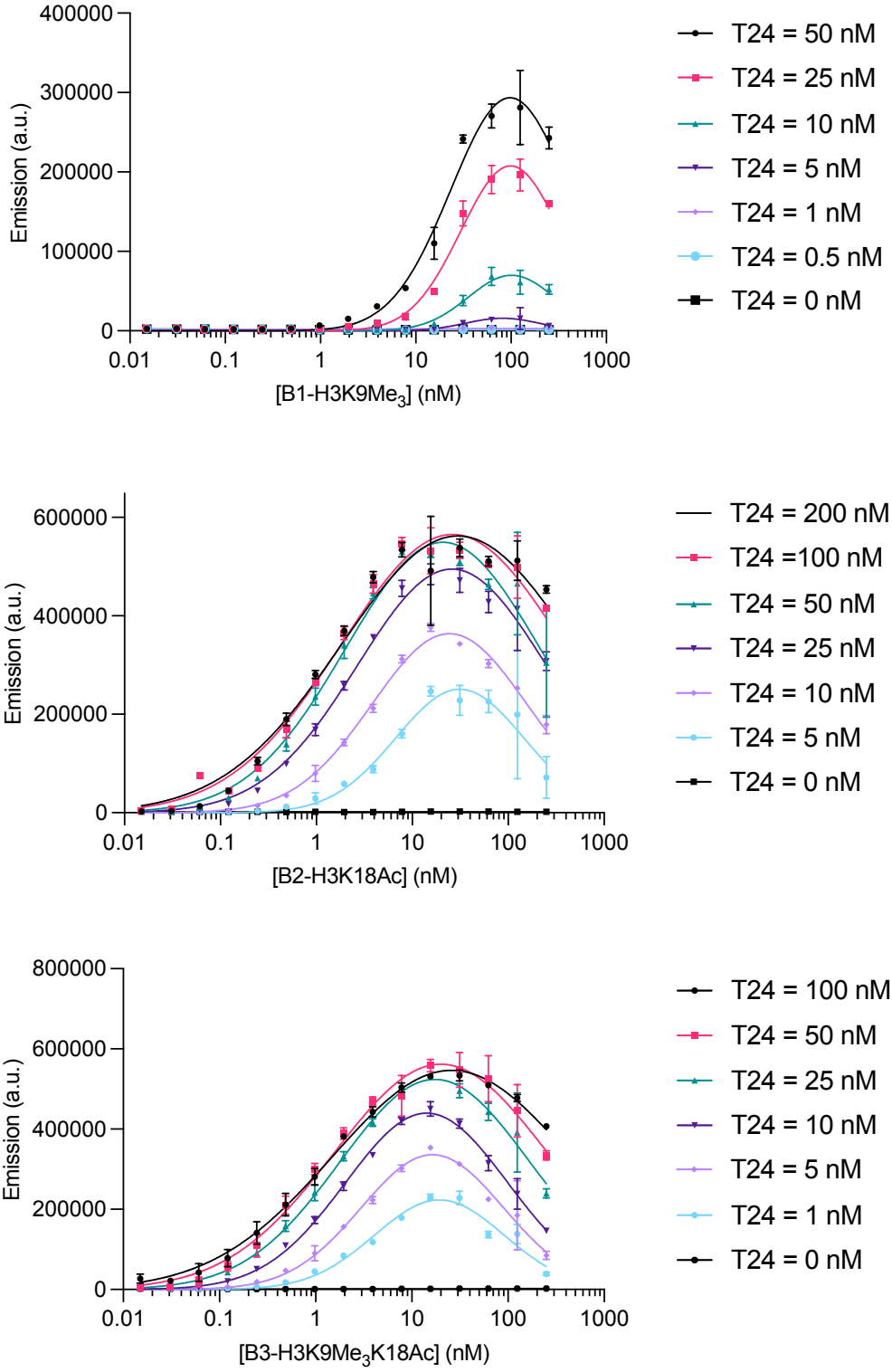


**Figure A4.2:** AlphaScreen® signal response curves for biotinylated histone-mimicking peptides binding to GST-TRIM24. Dose-response curves were obtained in triplicate with error bars indicating the standard deviation.

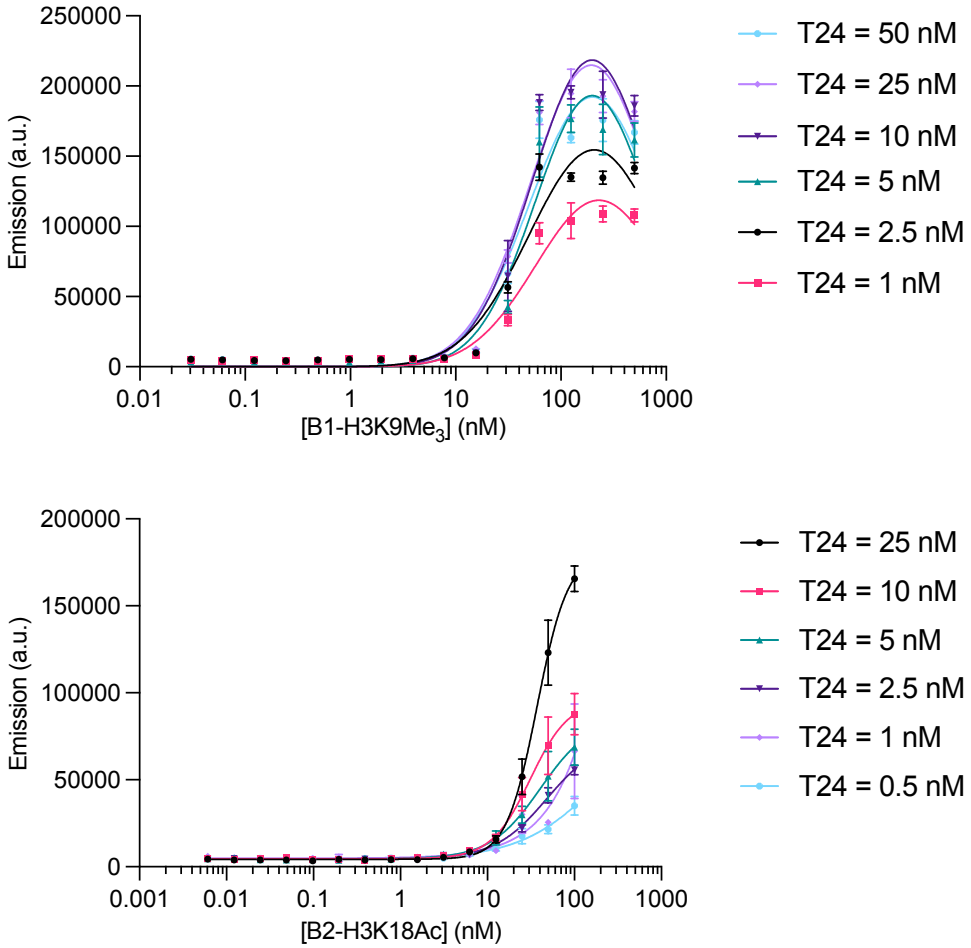



**Figure A4.3:** AlphaScreen® TruHits assay dose-response curves. Dose-response curves were obtained in triplicate with error bars indicating the standard deviation.

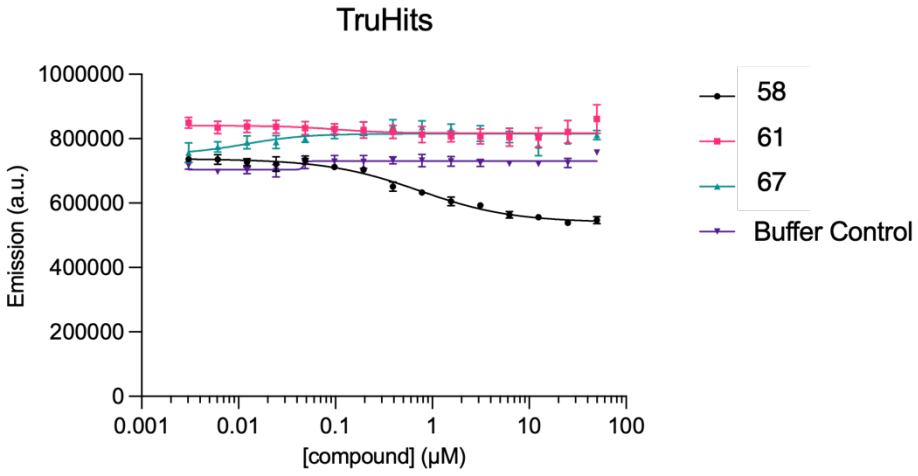
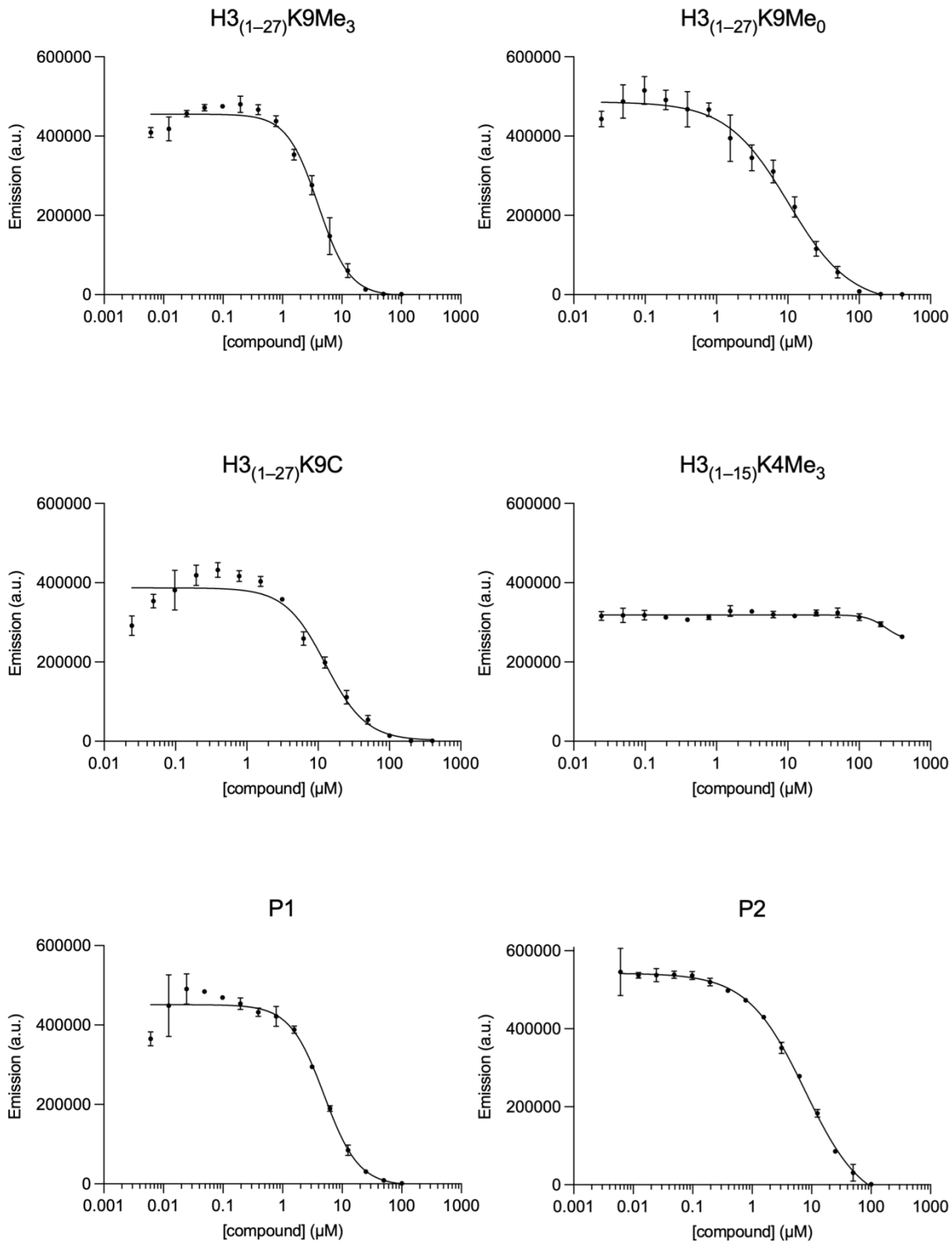
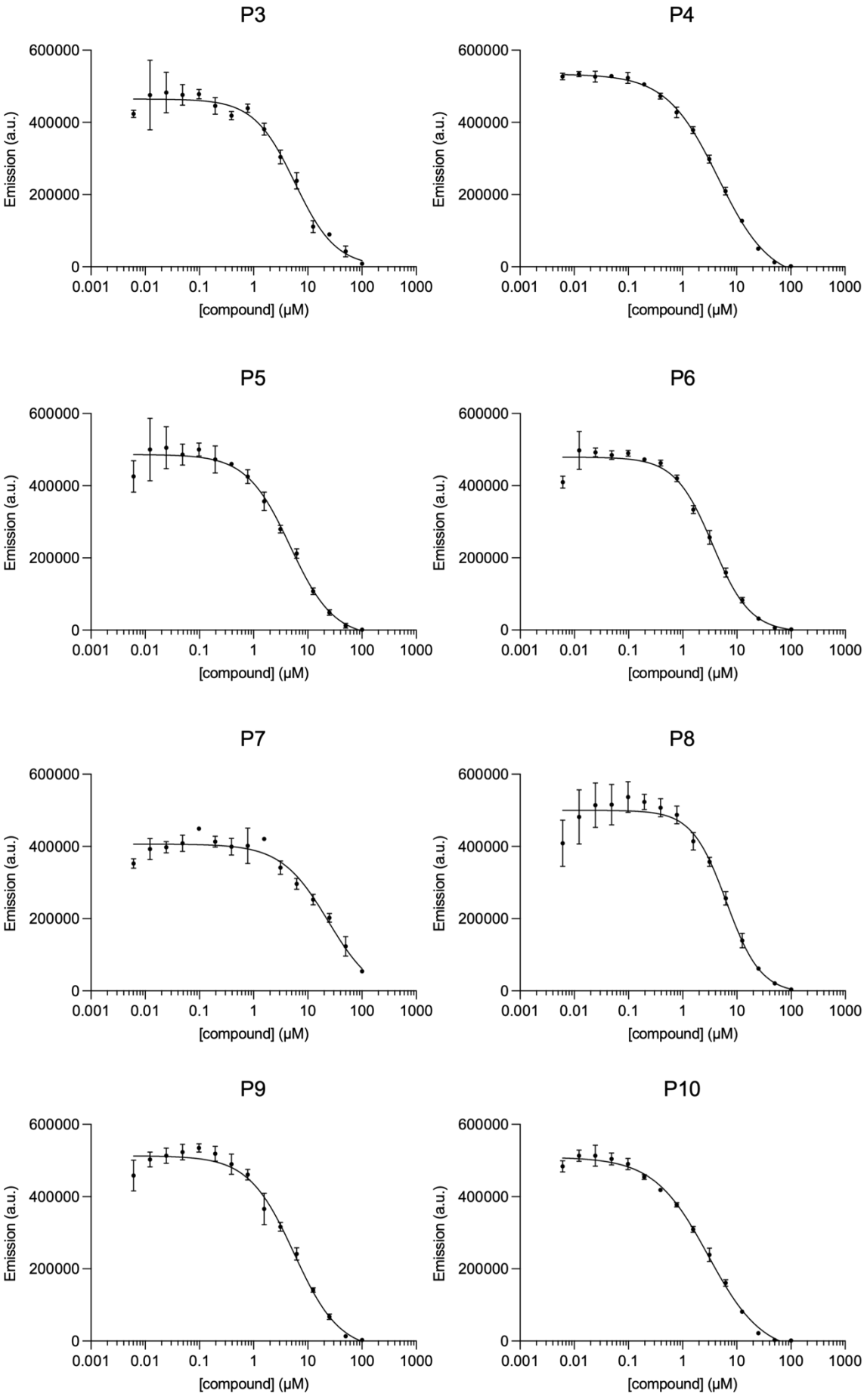
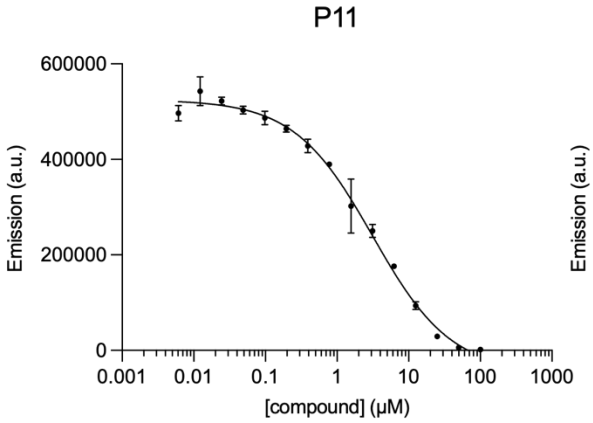
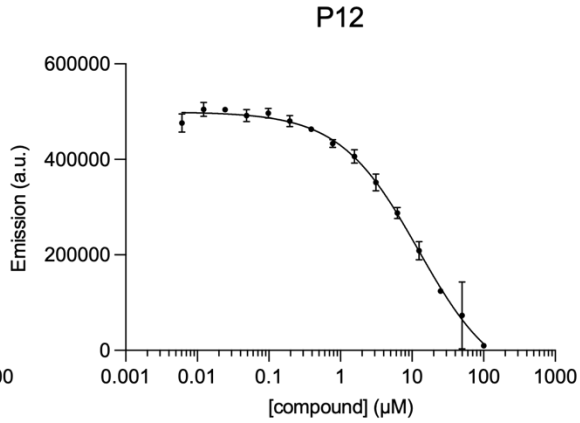


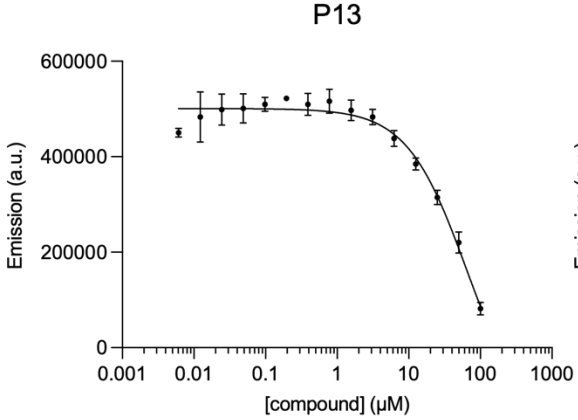
**Figure A4.4:** AlphaScreen® dose-response curves for the inhibition of TRIM24-H3K9Me$_3$ (B1) binding by H3-mimicking peptides **P1**–**P16**. Unbiotinylated H3$_{(1–27)}$K9Me$_3$ and H3$_{(1–15)}$K4Me$_3$ are the positive and negative controls, respectively. Dose-response curves were obtained in triplicate with error bars indicating the standard deviation.

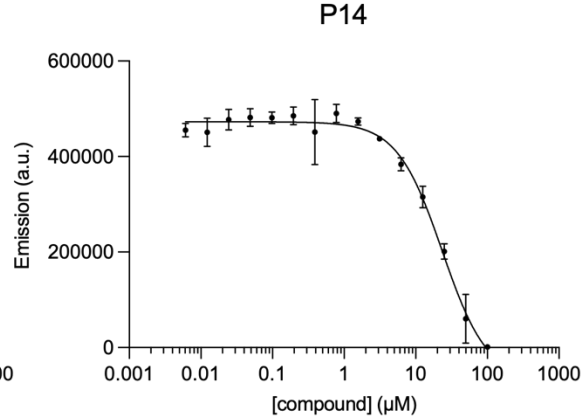


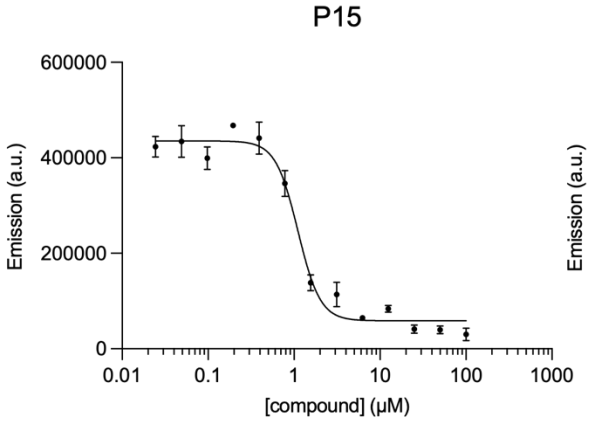

P3
P4
P5
P6
P7
P8
P9
P10
Emission (a.u.)
600000
400000
200000
0
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)


P11
600000
400000
200000
0
Emission (a.u.)
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)


P12
600000
400000
200000
0
Emission (a.u.)
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)


P13
600000
400000
200000
0
Emission (a.u.)
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)


P14
600000
400000
200000
0
Emission (a.u.)
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)


P15
600000
400000
200000
0
Emission (a.u.)
0.01
0.1
1
10
100
1000
[compound] (μM)

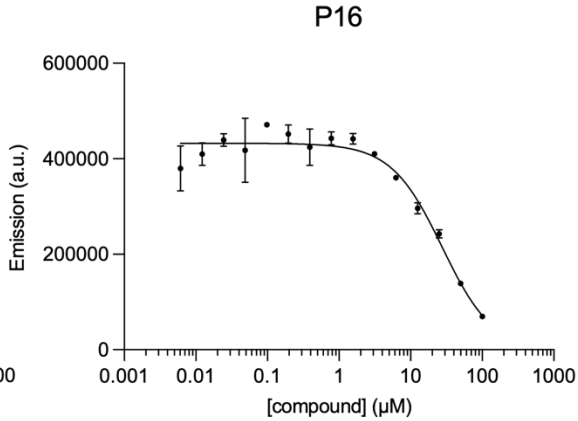

P16
600000
400000
200000
0
Emission (a.u.)
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)

**Figure A4.5:** AlphaScreen® dose-response curves for the inhibition of TRIM24-H3K9Me$_3$ (B1) binding by the PDCs and control compounds. Unbiotinylated H3$_{(1–27)}$K9Me$_3$ and H3$_{(1–15)}$K4Me$_3$ are the positive and negative controls, respectively. Dose-response curves were obtained in triplicate with error bars indicating the standard deviation.

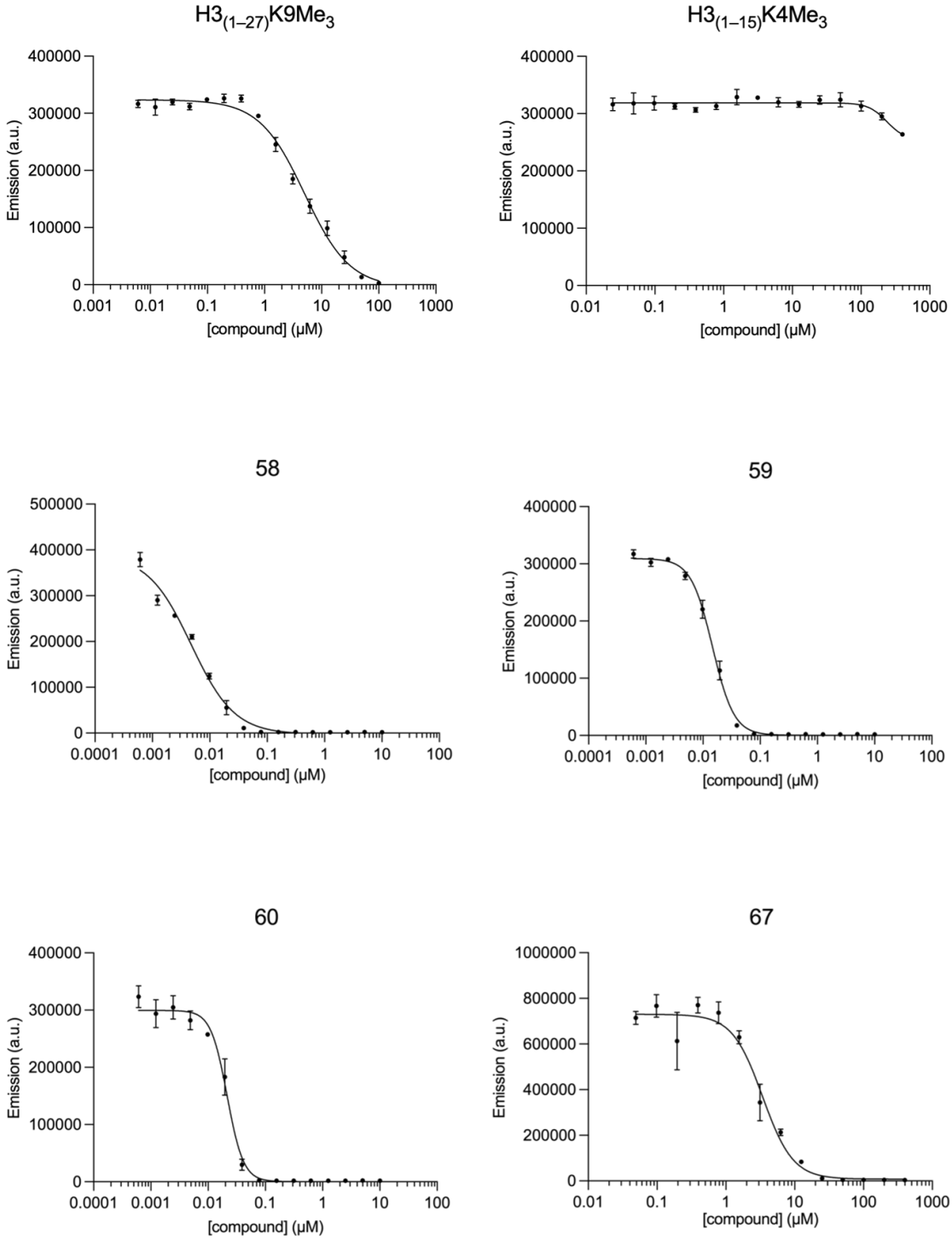
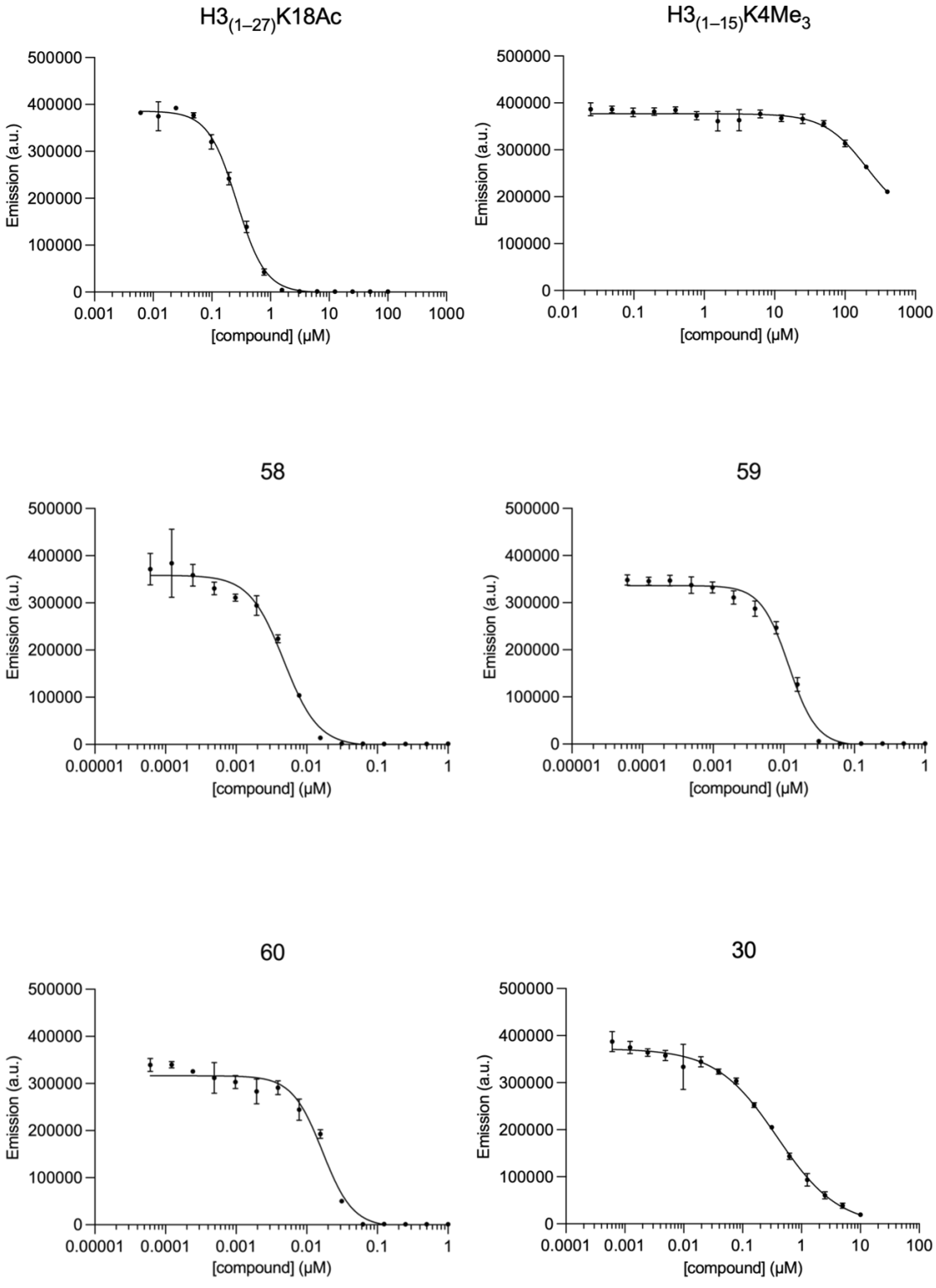
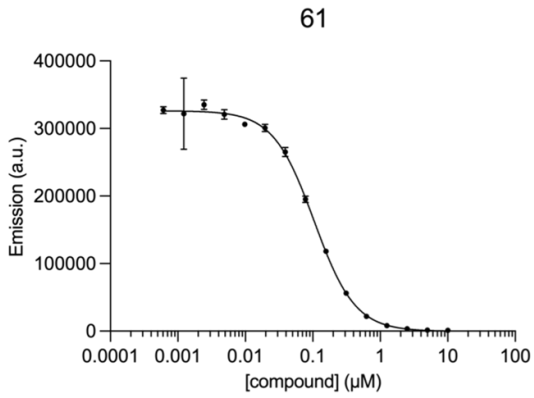
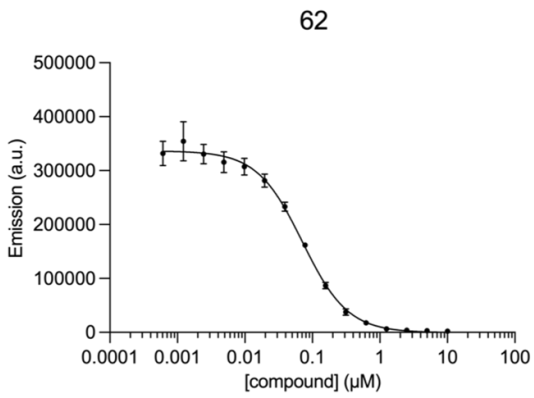
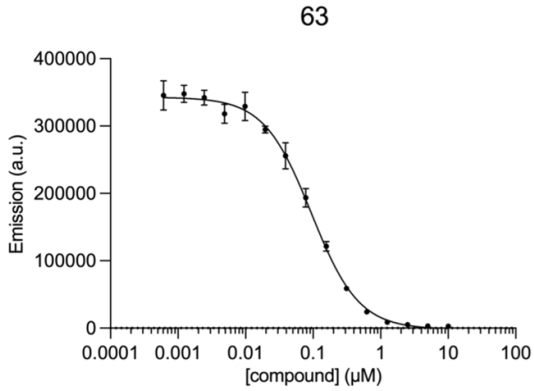


**Figure A4.6:** AlphaScreen® dose-response curves for the inhibition of TRIM24-H3K18Ac (B2) binding by the PDCs and control compounds. Unbiotinylated $H3_{(1–27)}K18Ac$ and $H3_{(1–15)}K4Me_3$ are the positive and negative controls, respectively. Dose-response curves were obtained in triplicate with error bars indicating the standard deviation.




61
Emission (a.u.)
400000
300000
200000
100000
0
0.0001
0.001
0.01
0.1
1
10
100
[compound] (μM)


62
Emission (a.u.)
500000
400000
300000
200000
100000
0
0.0001
0.001
0.01
0.1
1
10
100
[compound] (μM)


63
Emission (a.u.)
400000
300000
200000
100000
0
0.0001
0.001
0.01
0.1
1
10
100
[compound] (μM)

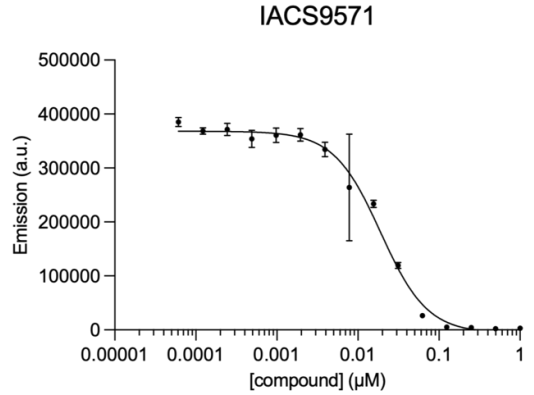
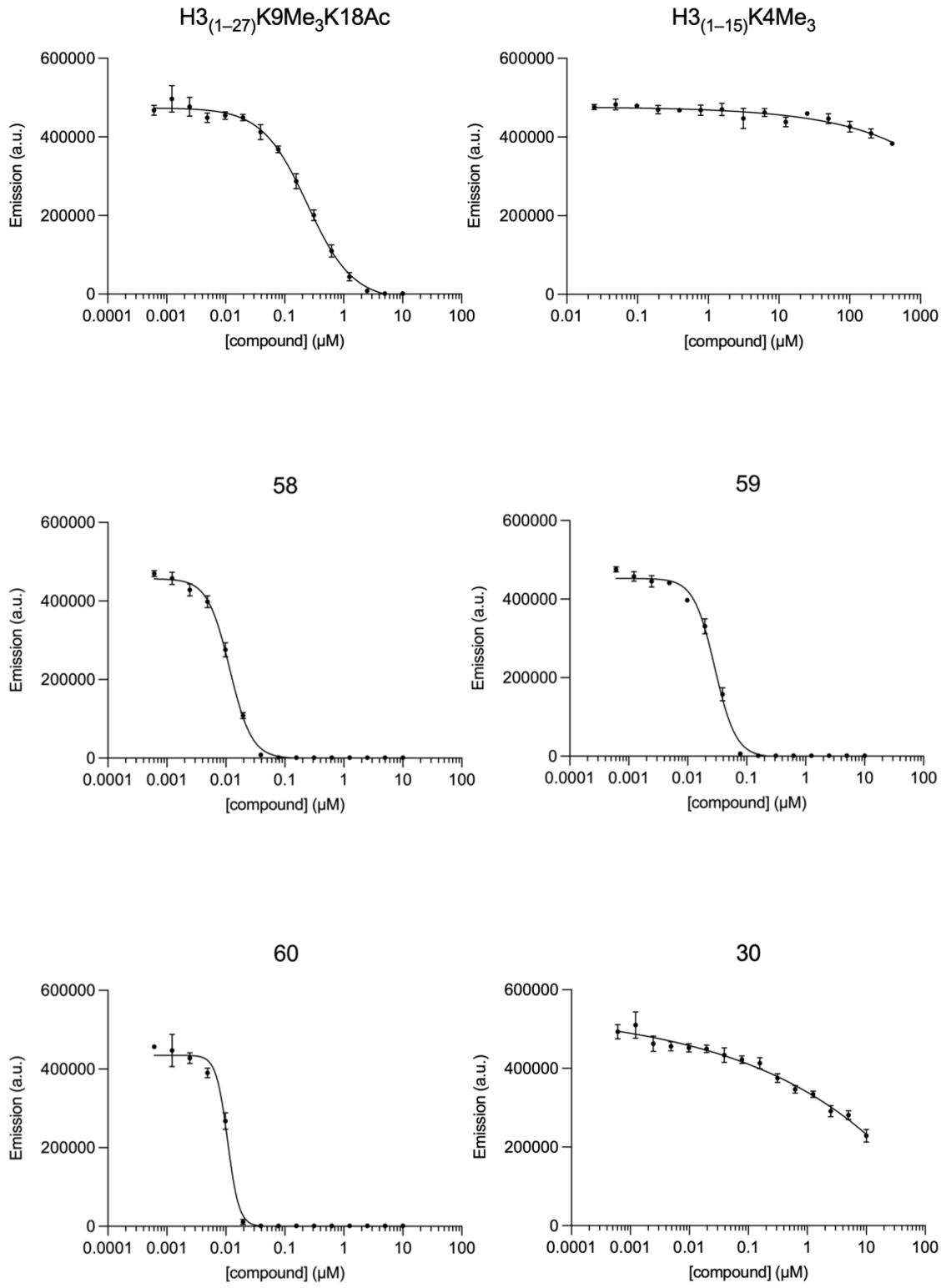
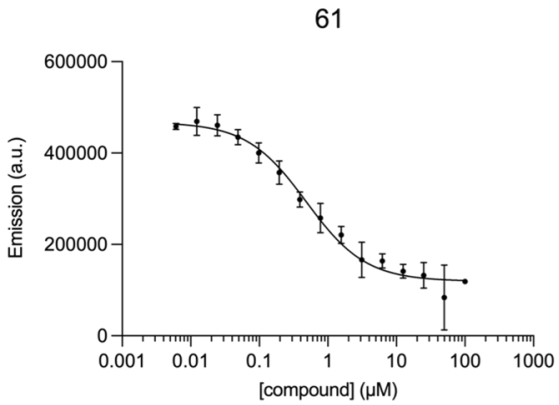
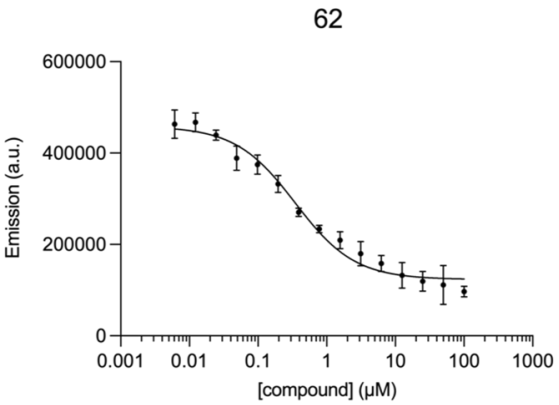
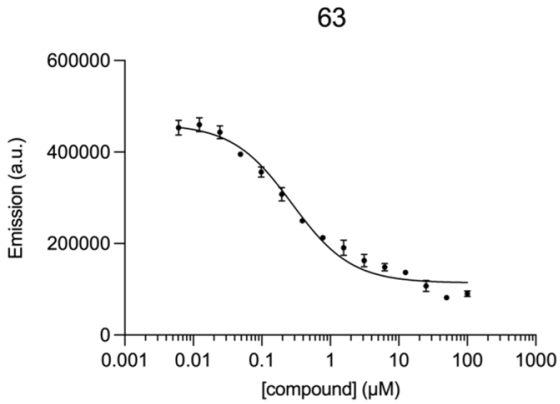

IACS9571
Emission (a.u.)
500000
400000
300000
200000
100000
0
0.00001
0.0001
0.001
0.01
0.1
1
[compound] (μM)

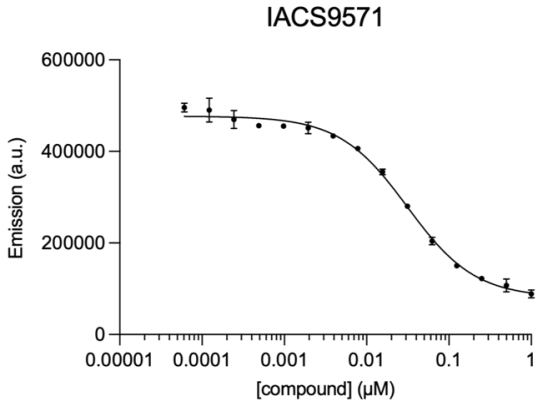
**Figure A4.7:** AlphaScreen® dose-response curves for the inhibition of TRIM24-H3K9Me$_3$K18Ac (B3) binding by the PDCs and control compounds. Unbiotinylated H3$_{(1–27)}$K9Me$_3$K18Ac and H3$_{(1–15)}$K4Me$_3$ are the positive and negative controls, respectively. Dose-response curves were obtained in triplicate with error bars indicating the standard deviation.

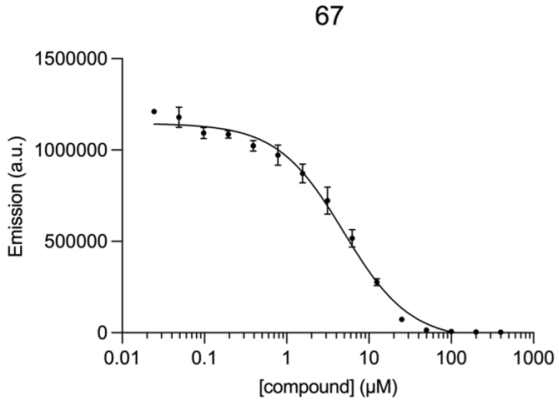


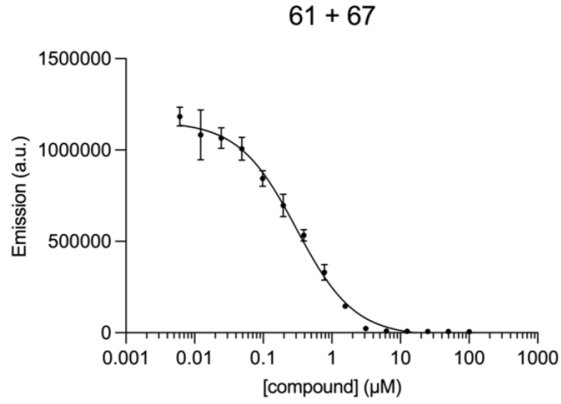

61
Emission (a.u.)
600000
400000
200000
0
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)

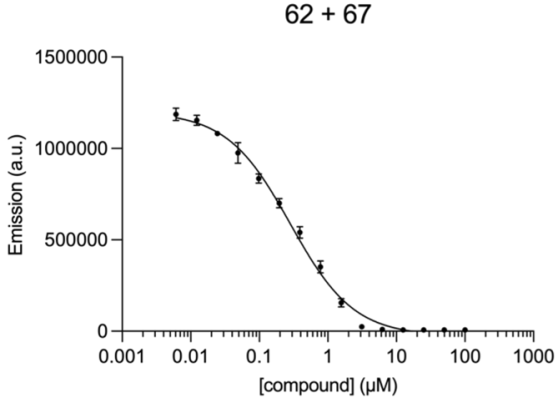

62
Emission (a.u.)
600000
400000
200000
0
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)


63
Emission (a.u.)
600000
400000
200000
0
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)


IACS9571
Emission (a.u.)
600000
400000
200000
0
0.00001
0.0001
0.001
0.01
0.1
1
[compound] (μM)


67
Emission (a.u.)
1500000
1000000
500000
0
0.01
0.1
1
10
100
1000
[compound] (μM)


61 + 67
Emission (a.u.)
1500000
1000000
500000
0
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)


62 + 67
Emission (a.u.)
1500000
1000000
500000
0
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)

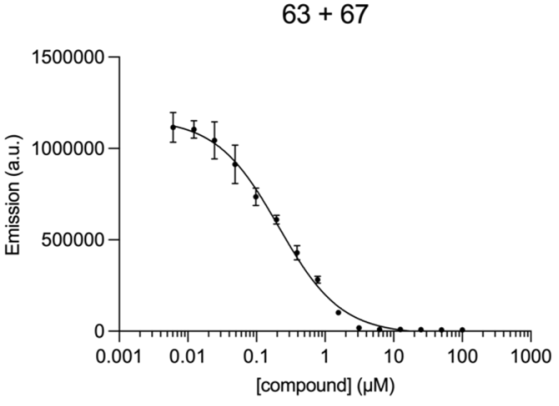
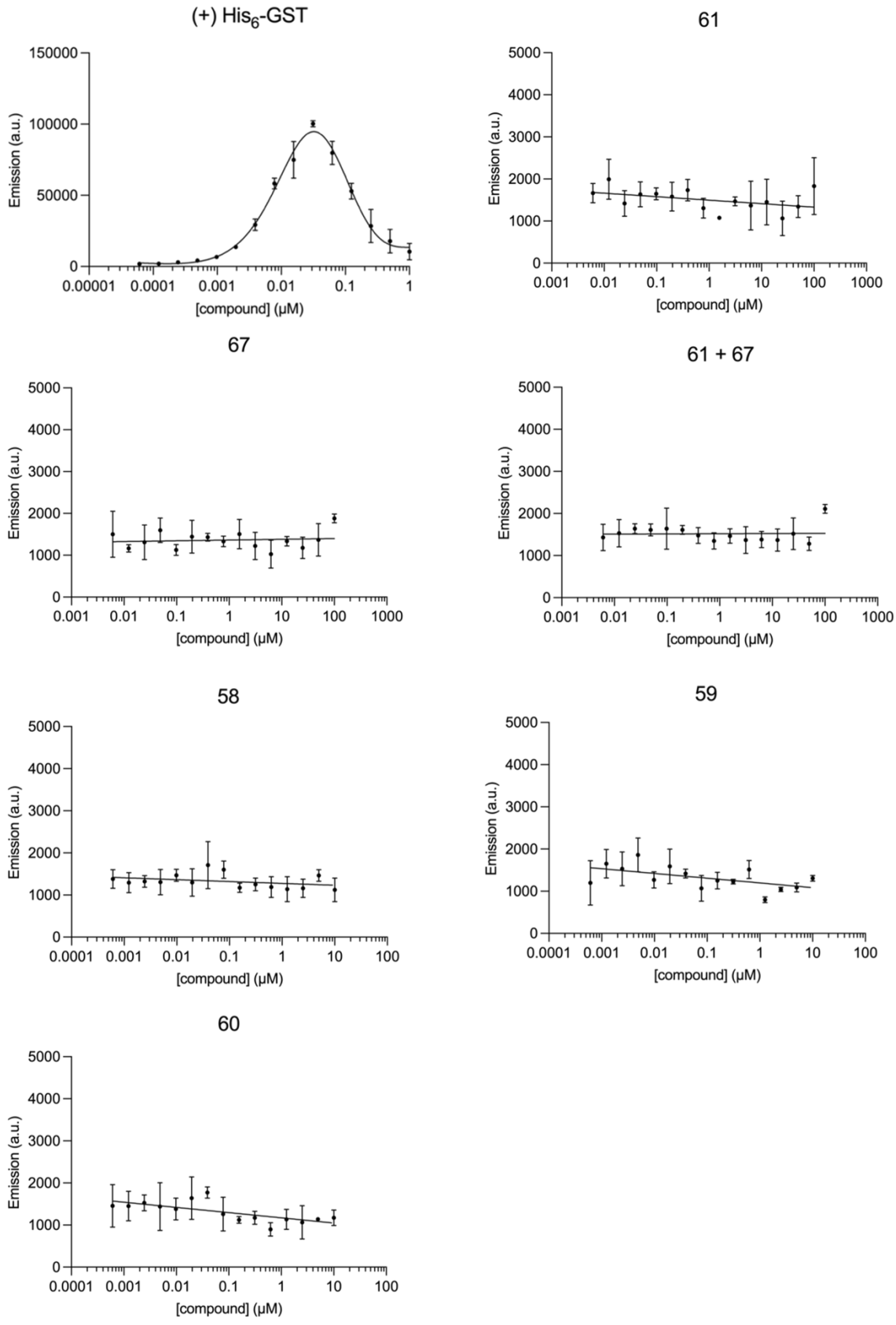

63 + 67
Emission (a.u.)
1500000
1000000
500000
0
0.001
0.01
0.1
1
10
100
1000
[compound] (μM)

**Figure A4.8:** AlphaScreen® ternary complex detection assay. Dose-response curves were obtained in triplicate with error bars indicating the standard deviation. $His_6$-GST is a technical positive control.



## Appendix 5 Fluorescence Anisotropy

**Figure A5.1:** $ESI^+$ mass spectra for the $H3_{(1–8)}$ peptide (Exact Mass = 930.54) stability test with benzaldehyde in fluorescence anisotropy buffer.

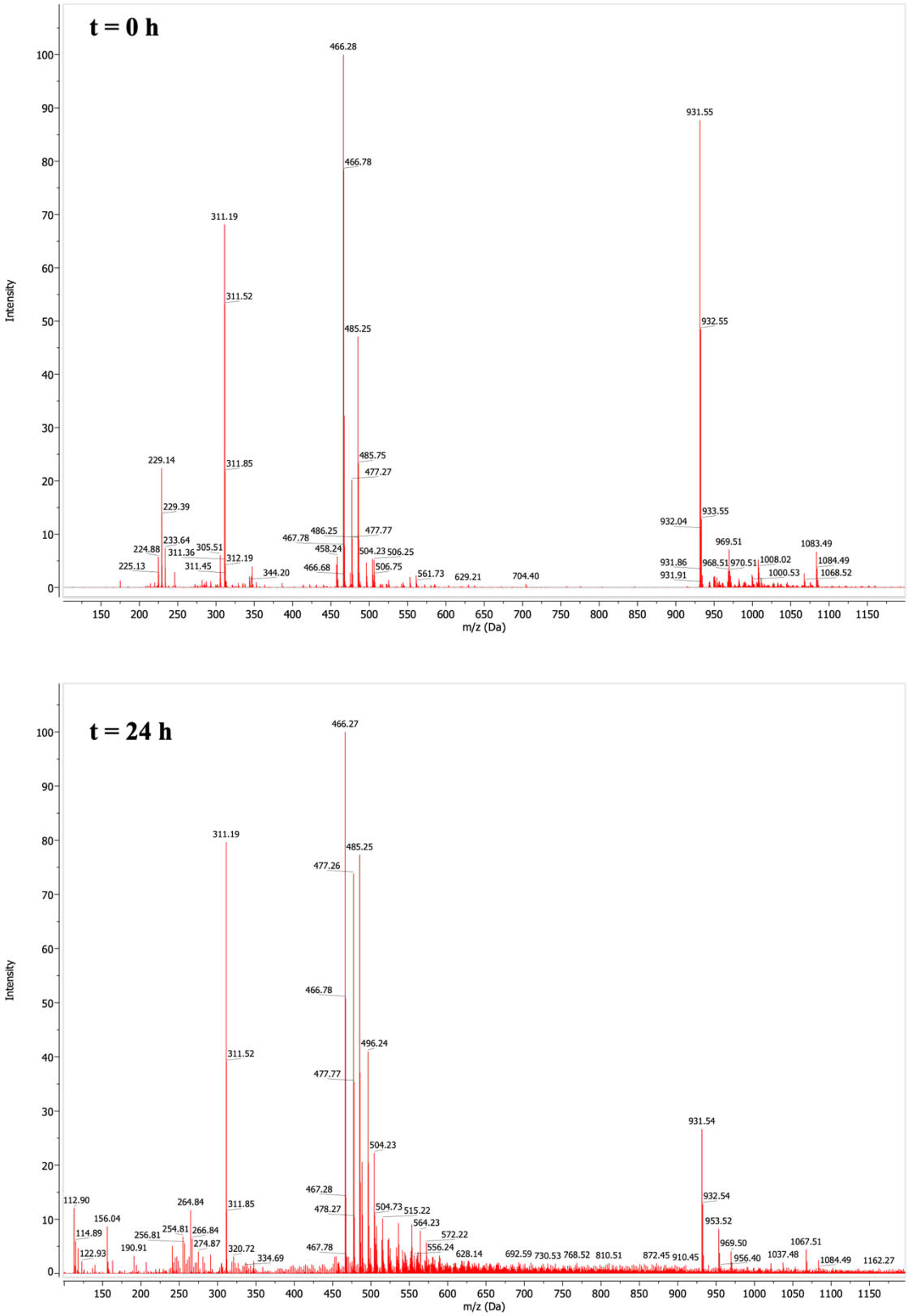
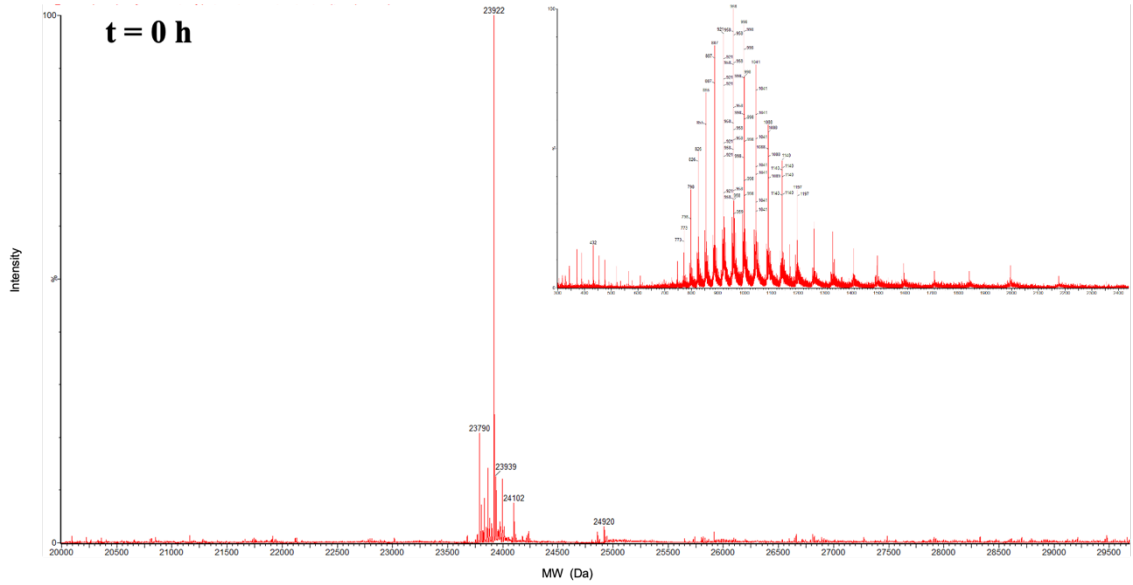
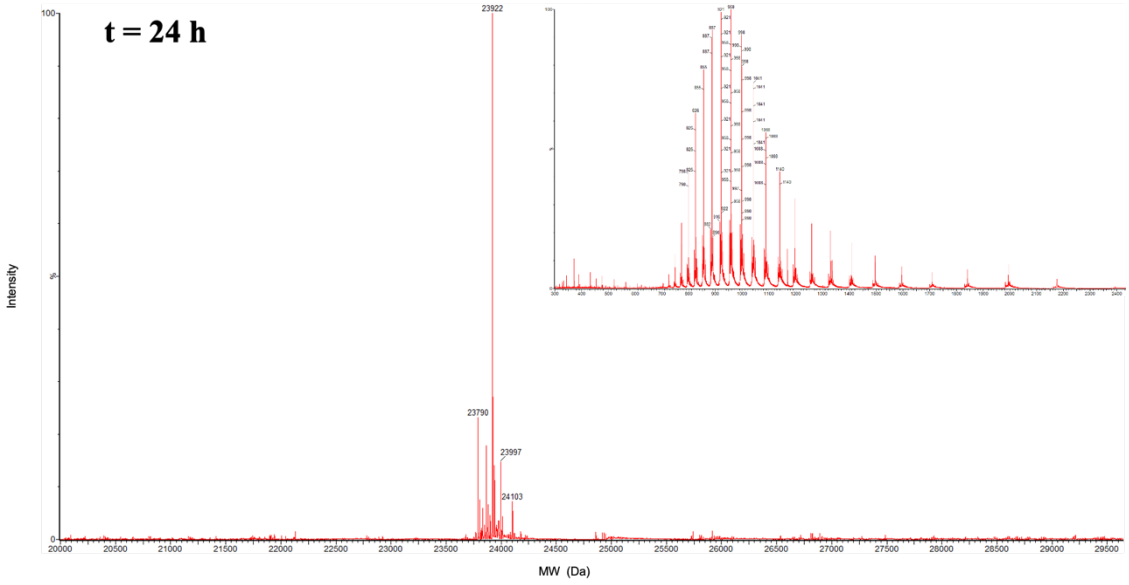


**Figure A5.2:** $ESI^+$ mass spectra and deconvoluted mass spectra for the $His_6$-TRIM24 WT (MW = 23922) stability test with benzaldehyde in fluorescence anisotropy buffer.



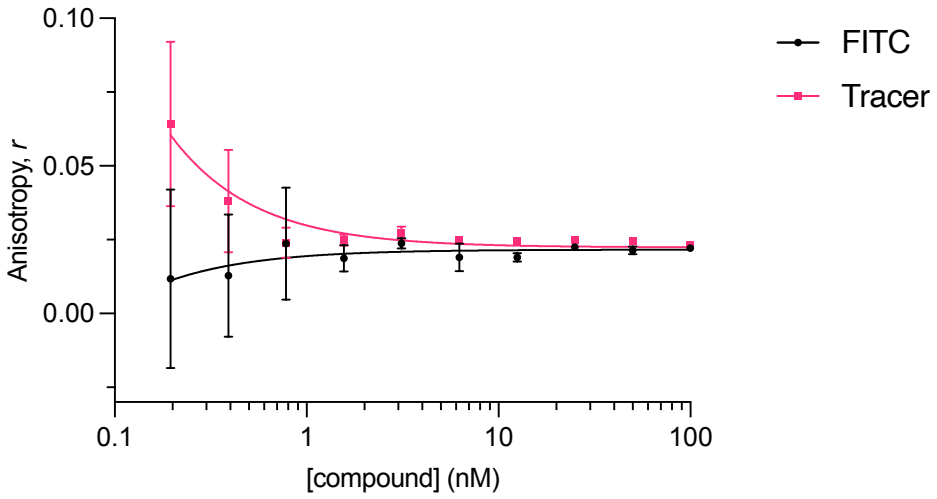
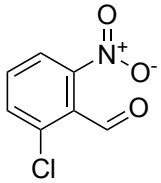
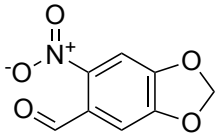
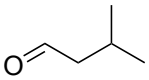
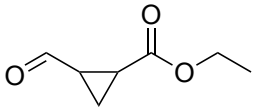
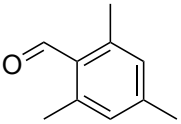
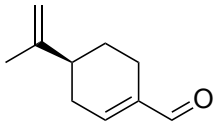
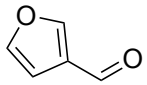
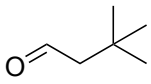
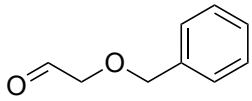
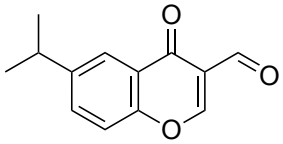
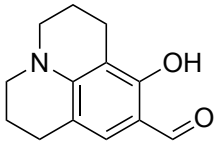
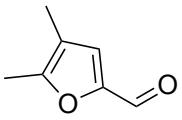
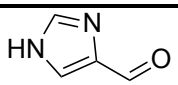
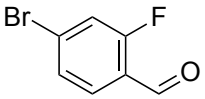
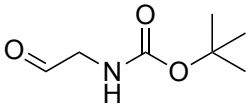
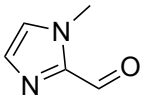
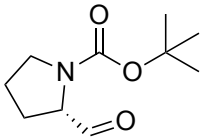
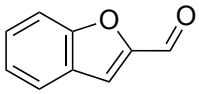
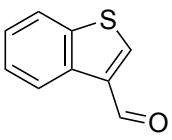
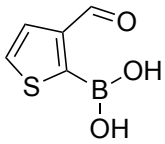
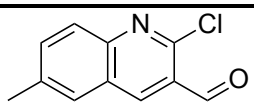
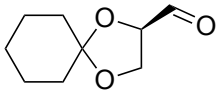
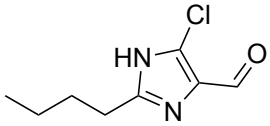
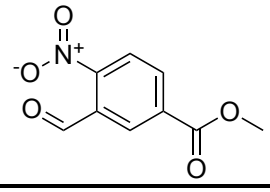
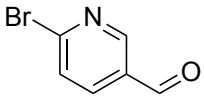
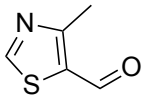
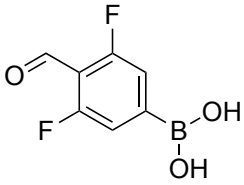
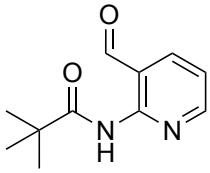
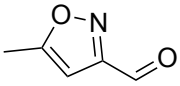
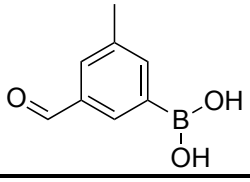
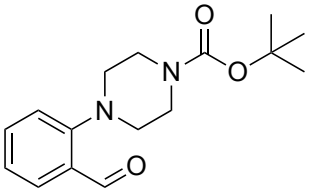


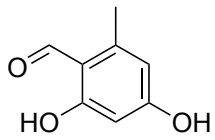
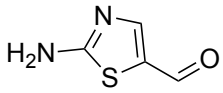
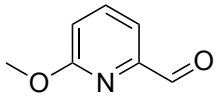
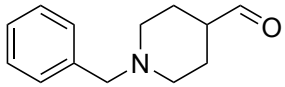
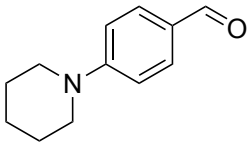
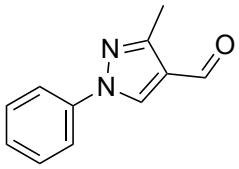
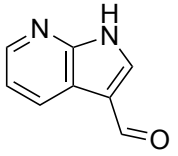
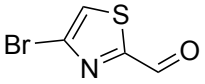
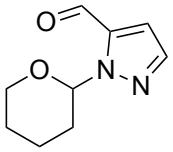
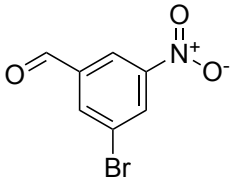
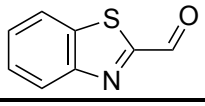
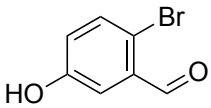
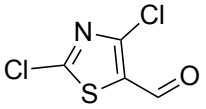

**Figure A5.3:** Graph comparing the fluorescence anisotropy of FITC tracer 2 and free FITC fluorophore at different concentrations. Error bars indicate the standard deviation of triplicate data.

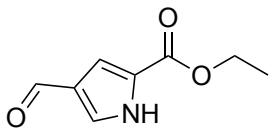
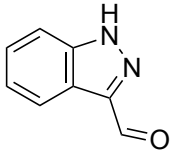
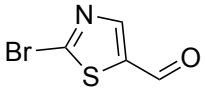
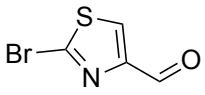
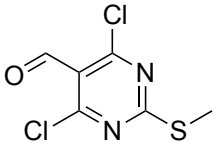
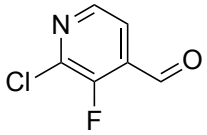
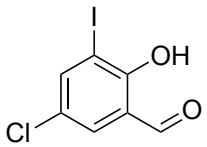
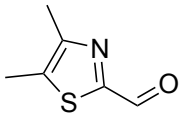
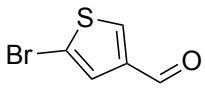
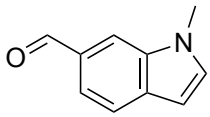
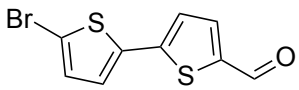
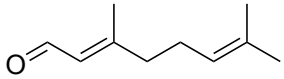
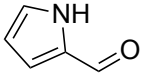
**Table A5.1:** Structures of compounds in the aldehyde fragment libraries provided by the lab of Prof. Andrew Wilson (University of Birmingham).

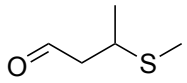
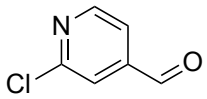
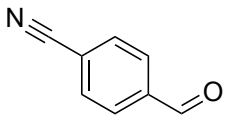
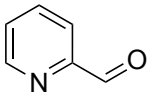
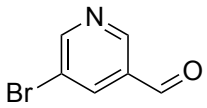
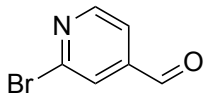
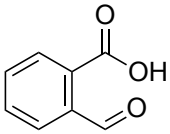
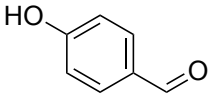
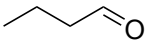
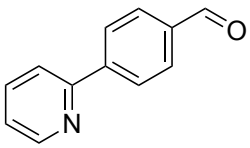
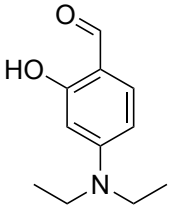
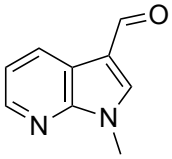
| Aldehyde | Structure |
|---|---|
| **A001** |  |
| **A002** |  |
| **A003** |  |
| **A004** |  |
| **A005** |  |

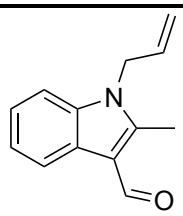
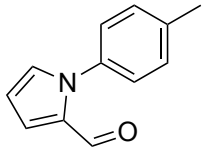
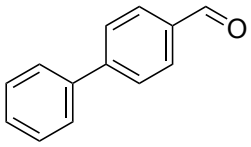
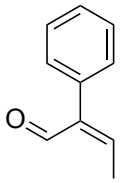
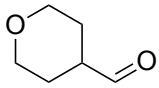
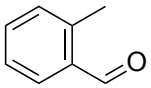
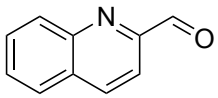
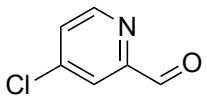
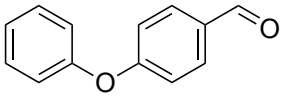
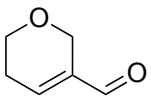
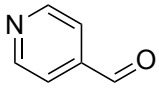
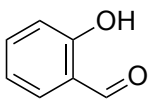
| | |
|---|---|
| **A006** |  |
| **A007** |  |
| **A008** |  |
| **A009** |  |
| **A010** |  |
| **A011** |  |
| **A012** |  |
| **A013** |  |
| **A014** |  |
| **A015** |  |
| **A016** |  |
| **A017** |  |
| **A018** |  |
| **A019** |  |

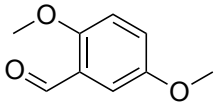
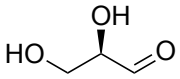
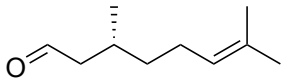
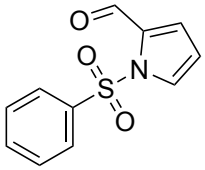
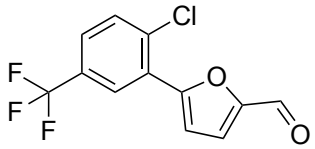
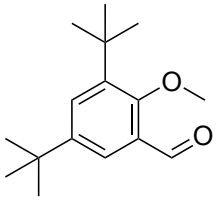
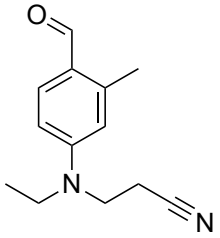
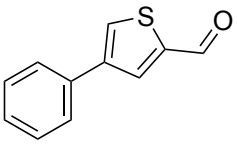
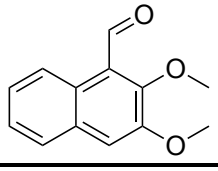
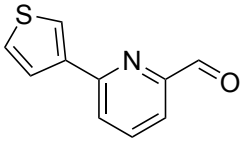
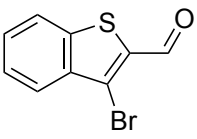
| | |
|---|---|
| **A020** |  |
| **A021** |  |
| **A022** |  |
| **A023** |  |
| **A024** |  |
| **A025** |  |
| **A026** |  |
| **A027** |  |
| **A028** |  |
| **A029** |  |
| **A030** |  |
| **A031** |  |

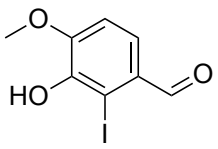
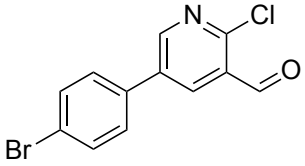
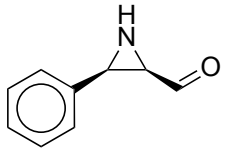
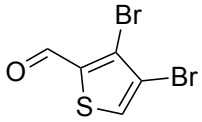
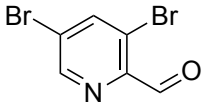
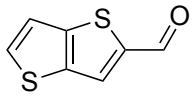
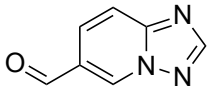
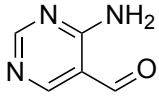
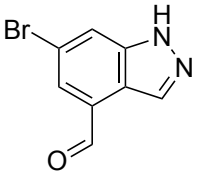
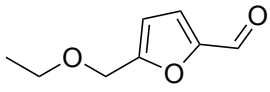
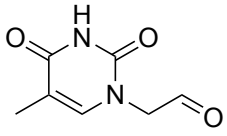
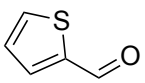
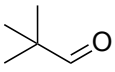
| | |
|---|---|
| **A032** |  |
| **A033** |  |
| **A034** |  |
| **A035** |  |
| **A036** |  |
| **A037** |  |
| **A038** |  |
| **A039** |  |
| **A040** |  |
| **A041** |  |
| **A042** |  |
| **A043** |  |
| **A044** |  |

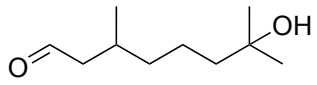
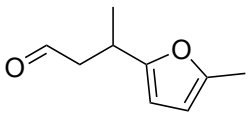
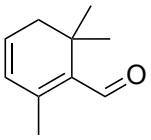
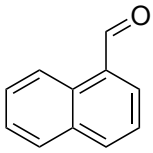
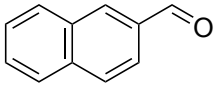
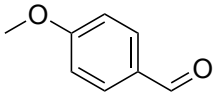
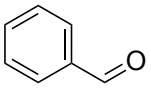
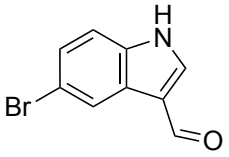
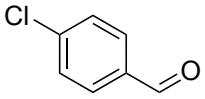
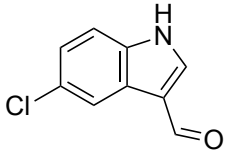
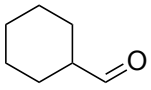
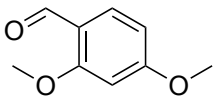
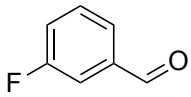
| | |
|---|---|
| **A045** |  |
| **A046** |  |
| **A047** |  |
| **A048** |  |
| **A049** |  |
| **A050** |  |
| **A051** |  |
| **A052** |  |
| **A053** |  |
| **A054** |  |
| **A055** |  |
| **A056** |  |
| **A057** |  |

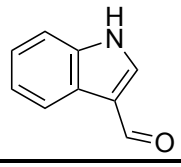
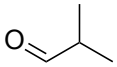
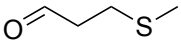
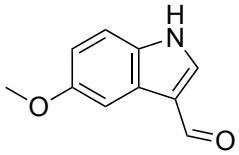
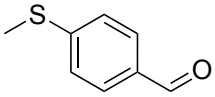
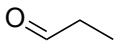
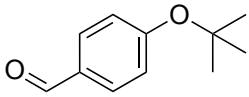
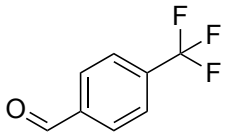
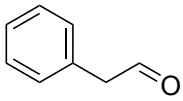
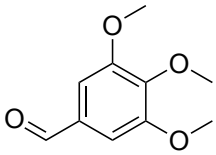
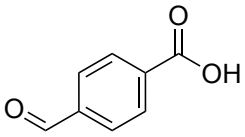
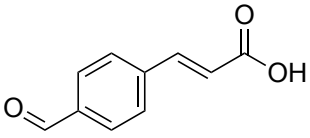
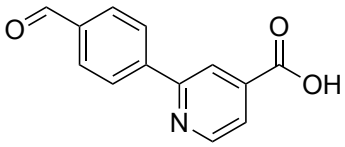
| | |
|---|---|
| **A058** |  |
| **A059** |  |
| **A060** |  |
| **A061** |  |
| **A062** |  |
| **A063** |  |
| **A064** |  |
| **A065** |  |
| **A066** |  |
| **A067** |  |
| **A068** |  |
| **A069** |  |

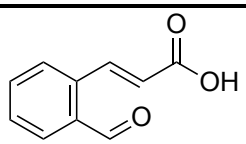
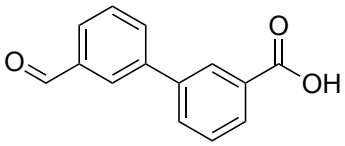
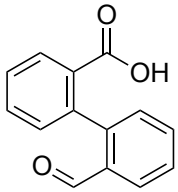
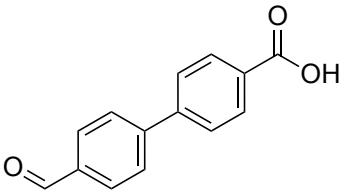
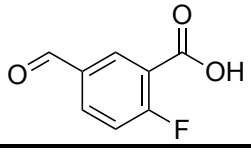
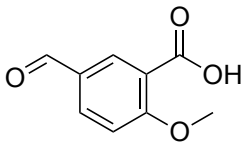
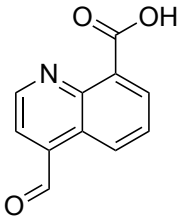
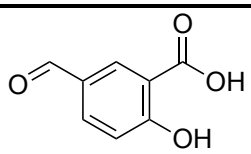
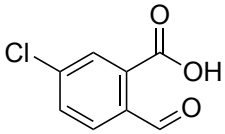
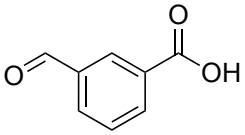
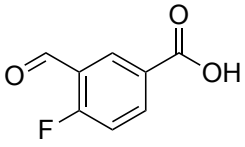
| | |
|---|---|
| **A070** |  |
| **A071** |  |
| **A072** |  |
| **A073** |  |
| **A074** |  |
| **A075** |  |
| **A076** |  |
| **A077** |  |
| **A078** |  |
| **A079** |  |
| **A080** |  |
| **A081** |  |

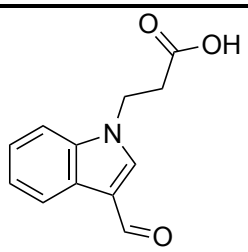
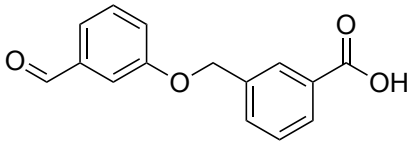
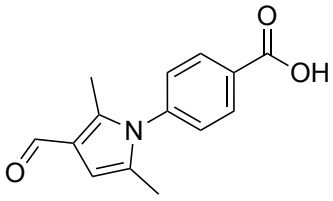
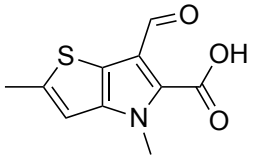
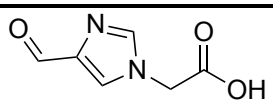
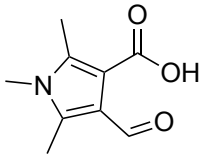
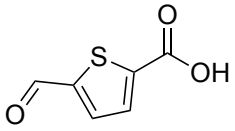
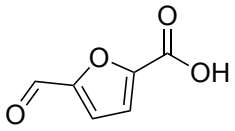
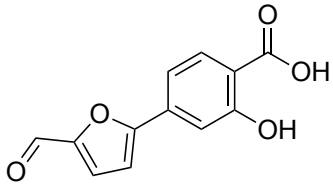
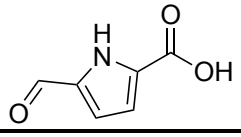
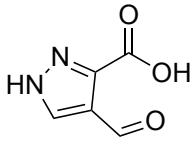
| | |
|---|---|
| **A082** |  |
| **A083** |  |
| **A084** |  |
| **A085** |  |
| **A086** |  |
| **A087** |  |
| **A088** |  |
| **A089** |  |
| **A090** |  |
| **A091** |  |
| **A092** |  |

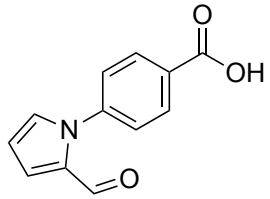
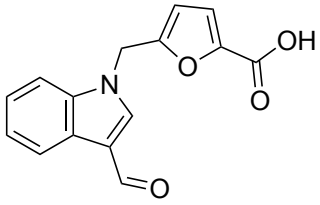
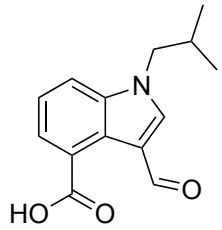
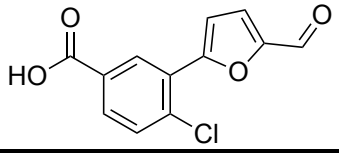
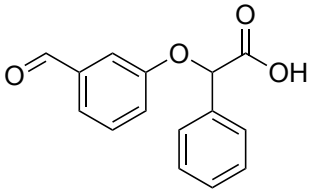
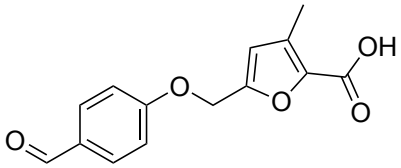
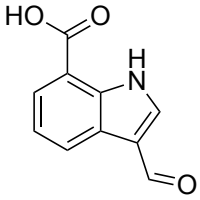
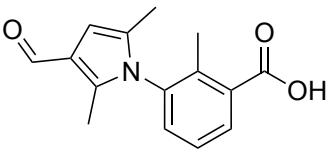
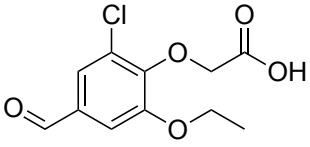
| | |
|---|---|
| **A093** |  |
| **A094** |  |
| **A095** |  |
| **A096** |  |
| **A097** |  |
| **A098** |  |
| **A099** |  |
| **A100** |  |
| **A101** |  |
| **A102** |  |
| **A103** |  |
| **A104** |  |
| **A105** |  |

| | |
|---|---|
| **A106** |  |
| **A107** |  |
| **A108** |  |
| **A109** |  |
| **A110** |  |
| **A111** |  |
| **A112** |  |
| **A113** |  |
| **A114** |  |
| **A115** |  |
| **A116** |  |
| **A117** |  |
| **A118** |  |

| | |
|---|---|
| **A119** |  |
| **A120** |  |
| **A121** |  |
| **A122** |  |
| **A123** |  |
| **A124** |  |
| **A125** |  |
| **A126** |  |
| **A127** |  |
| **A128** |  |
| **A129** |  |
| **A130** |  |
| **A131** |  |

| | |
|---|---|
| **A132** |  |
| **A133** |  |
| **A134** |  |
| **A135** |  |
| **A136** |  |
| **A137** |  |
| **A138** |  |
| **A139** |  |
| **A140** |  |
| **A141** |  |
| **A142** |  |

| | |
|---|---|
| **A143** |  |
| **A144** |  |
| **A145** |  |
| **A146** |  |
| **A147** |  |
| **A148** |  |
| **A149** |  |
| **A150** |  |
| **A151** |  |
| **A152** |  |
| **A153** |  |

| | |
|---|---|
| **A154** |  |
| **A155** |  |
| **A156** |  |
| **A157** |  |
| **A158** |  |
| **A159** |  |
| **A160** |  |
| **A161** |  |
| **A162** |  |

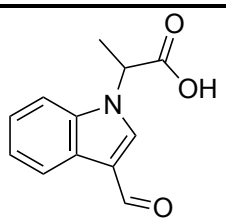
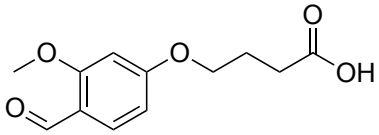
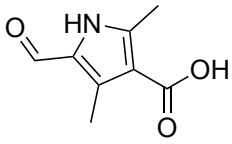
| | |
|---|---|
| **A163** |  |
| **A164** |  |
| **A165** |  |

**Plate 1 Layout**

| Plate 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | A001 | | A002 | | A003 | | A004 | | A005 | | A006 | | A007 | | A008 | | A009 | | A010 | | A011 | | A012 | |
| B | | | | | | | | | | | | | | | | | | | | | | | | |
| C | A013 | | A014 | | A015 | | A016 | | A017 | | A018 | | A019 | | A020 | | A021 | | A022 | | A023 | | A024 | |
| D | | | | | | | | | | | | | | | | | | | | | | | | |
| E | A025 | | A026 | | A027 | | A028 | | A029 | | A030 | | A031 | | A032 | | A033 | | A034 | | A035 | | A036 | |
| F | | | | | | | | | | | | | | | | | | | | | | | | |
| G | A037 | | A038 | | A039 | | A040 | | A041 | | A042 | | A043 | | A044 | | A045 | | A046 | | A047 | | A048 | |
| H | | | | | | | | | | | | | | | | | | | | | | | | |
| I | A049 | | A050 | | A051 | | A052 | | A053 | | A054 | | A055 | | A056 | | A057 | | A058 | | A059 | | A060 | |
| J | | | | | | | | | | | | | | | | | | | | | | | | |
| K | A061 | | A062 | | A063 | | A064 | | A065 | | A066 | | A067 | | A068 | | A069 | | A070 | | A071 | | A072 | |
| L | | | | | | | | | | | | | | | | | | | | | | | | |
| M | A073 | | A074 | | A075 | | A076 | | A077 | | A078 | | A079 | | A080 | | A081 | | A082 | | | | | |
| N | | | | | | | | | | | | | | | | | | | | | | | | |
| O | | | | | | | | | | | | | | | | | | | | | | | | |
| P | | | | | | | | | | | | | | | | | | | | | | | | |

**P (Perpendicular)**

| Plate 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 16569 | 13913 | 16474 | 13827 | 16507 | 14370 | 15402 | 13818 | 15218 | 13677 | 15621 | 12527 | 15113 | 12848 | 15236 | 12773 | 14339 | 12666 | 14658 | 12598 | 13426 | 11970 | 14542 | 12292 |
| B | 245 | 287 | 216 | 223 | 202 | 195 | 188 | 197 | 193 | 205 | 208 | 205 | 348 | 344 | 187 | 198 | 186 | 195 | 361 | 433 | 468 | 461 | 211 | 226 |
| C | 16628 | 12634 | 16246 | 13216 | 15678 | 13736 | 15204 | 13455 | 14868 | 13246 | 15169 | 12642 | 14266 | 12148 | 15345 | 12358 | 14043 | 12475 | 14768 | 12507 | 13659 | 11866 | 13459 | 12380 |
| D | 243 | 255 | 221 | 242 | 222 | 241 | 214 | 216 | 204 | 214 | 216 | 218 | 218 | 217 | 222 | 225 | 208 | 202 | 190 | 165 | 199 | 201 | 202 | 221 |
| E | 15901 | 13472 | 15225 | 13006 | 14364 | 12625 | 15543 | 13911 | 14791 | 13427 | 15191 | 12636 | 14758 | 12433 | 13804 | 11429 | 15869 | 13407 | 14437 | 12474 | 14920 | 12763 | 13902 | 12249 |
| F | 219 | 290 | 203 | 198 | 218 | 223 | 231 | 215 | 185 | 213 | 208 | 201 | 211 | 206 | 230 | 233 | 283 | 356 | 175 | 164 | 591 | 566 | 229 | 229 |
| G | 16556 | 14368 | 15406 | 13488 | 16156 | 13479 | 15091 | 13393 | 14586 | 13399 | 15595 | 13192 | 13922 | 12079 | 14426 | 11603 | 14594 | 12412 | 14700 | 12487 | 14986 | 12267 | 13792 | 12441 |
| H | 240 | 267 | 230 | 252 | 231 | 228 | 217 | 208 | 218 | 213 | 231 | 228 | 220 | 230 | 291 | 274 | 292 | 295 | 214 | 213 | 241 | 242 | 182 | 266 |
| I | 12541 | 10315 | 14645 | 12608 | 18240 | 15003 | 15101 | 13002 | 14862 | 12877 | 15122 | 12642 | 14501 | 12460 | 15126 | 12283 | 13560 | 12093 | 15016 | 13720 | 15256 | 11715 | 14138 | 11982 |
| J | 350 | 347 | 192 | 198 | 381 | 594 | 199 | 223 | 193 | 202 | 222 | 225 | 225 | 222 | 223 | 224 | 176 | 186 | 190 | 175 | 195 | 195 | 201 | 194 |
| K | 16109 | 13095 | 15344 | 12898 | 14937 | 12473 | 15035 | 13300 | 15532 | 11493 | 15372 | 12600 | 4842 | 12524 | 14798 | 12396 | 13698 | 11831 | 15417 | 13088 | 15393 | 12635 | 14422 | 11503 |
| L | 222 | 211 | 193 | 190 | 188 | 193 | 197 | 231 | 218 | 198 | 168 | 153 | 192 | 204 | 233 | 237 | 155 | 154 | 198 | 217 | 199 | 177 | 153 | 184 |
| M | 15706 | 13631 | 14955 | 14051 | 15226 | 13698 | 16630 | 14417 | 15603 | 13125 | 14933 | 12698 | 15160 | 12332 | 14734 | 12191 | 14035 | 12505 | 14586 | 12765 | | | | |
| N | 207 | 218 | 193 | 185 | 181 | 241 | 804 | 925 | 393 | 393 | 191 | 177 | 218 | 216 | 177 | 190 | 174 | 185 | 173 | 167 | | | | |
| O | | | | | | | | | | | | | | | | | | | | | | | | |
| P | | | | | | | | | | | | | | | | | | | | | | | | |

**S (Parallel)**

| Plate 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 19170 | 16157 | 19113 | 16029 | 19128 | 16620 | 17991 | 16011 | 17681 | 15819 | 17702 | 14219 | 17735 | 15028 | 17683 | 14654 | 16671 | 14647 | 17021 | 14596 | 15536 | 13858 | 16606 | 14004 |
| B | 248 | 298 | 221 | 228 | 208 | 199 | 195 | 204 | 204 | 210 | 218 | 214 | 501 | 486 | 190 | 204 | 192 | 199 | 400 | 463 | 480 | 475 | 220 | 236 |
| C | 19299 | 14779 | 18935 | 15437 | 18310 | 15954 | 17864 | 15696 | 17418 | 15422 | 17703 | 14696 | 16746 | 14245 | 17919 | 14451 | 16456 | 14556 | 17268 | 14602 | 16081 | 13958 | 15761 | 14351 |
| D | 248 | 259 | 226 | 245 | 227 | 246 | 220 | 221 | 210 | 225 | 228 | 233 | 233 | 235 | 231 | 234 | 212 | 207 | 195 | 172 | 215 | 218 | 206 | 225 |
| E | 18510 | 15743 | 17803 | 15204 | 17137 | 14963 | 18195 | 16180 | 17240 | 15552 | 17777 | 14791 | 17282 | 14530 | 16163 | 13404 | 17875 | 15143 | 16916 | 14568 | 17326 | 14793 | 16337 | 14275 |
| F | 234 | 303 | 207 | 204 | 227 | 230 | 239 | 226 | 192 | 216 | 215 | 208 | 216 | 213 | 243 | 247 | 321 | 399 | 185 | 173 | 667 | 633 | 284 | 280 |
| G | 19168 | 16698 | 18059 | 15754 | 18815 | 15723 | 17705 | 15635 | 17102 | 15633 | 18133 | 15378 | 16386 | 14146 | 16790 | 13538 | 17242 | 14637 | 17133 | 14565 | 17321 | 14182 | 16104 | 14319 |
| H | 245 | 288 | 235 | 258 | 245 | 241 | 222 | 214 | 224 | 224 | 247 | 244 | 235 | 245 | 387 | 353 | 421 | 429 | 225 | 227 | 279 | 281 | 192 | 277 |
| I | 14802 | 12242 | 17128 | 14706 | 20871 | 17278 | 17735 | 15245 | 17387 | 15064 | 17724 | 14804 | 16999 | 14565 | 17263 | 14058 | 16007 | 14214 | 17281 | 15745 | 17758 | 13723 | 16579 | 13945 |
| J | 541 | 536 | 198 | 207 | 543 | 976 | 208 | 238 | 198 | 210 | 247 | 255 | 236 | 233 | 251 | 251 | 181 | 191 | 196 | 181 | 202 | 201 | 209 | 203 |
| K | 18704 | 15314 | 17924 | 15076 | 17439 | 14564 | 17609 | 15476 | 18162 | 13509 | 18005 | 14703 | 5482 | 14650 | 17287 | 14490 | 16103 | 13907 | 17915 | 15191 | 17926 | 14746 | 16937 | 13440 |
| L | 230 | 220 | 196 | 195 | 193 | 200 | 203 | 235 | 224 | 205 | 175 | 157 | 210 | 222 | 261 | 262 | 161 | 160 | 230 | 249 | 208 | 187 | 169 | 196 |
| M | 17978 | 15568 | 17321 | 16191 | 17841 | 15915 | 19687 | 17044 | 17990 | 15174 | 17484 | 14768 | 17276 | 14098 | 17200 | 14247 | 16465 | 14631 | 17036 | 14886 | | | | |
| N | 217 | 229 | 198 | 193 | 187 | 246 | 1442 | 1743 | 433 | 428 | 196 | 183 | 229 | 229 | 183 | 193 | 181 | 190 | 180 | 172 | | | | |
| O | | | | | | | | | | | | | | | | | | | | | | | | |
| P | | | | | | | | | | | | | | | | | | | | | | | | |

**Figure A5.4:** Layout of 384-well plate 1 and raw fluorescence intensity data for the parallel and perpendicular channels in the dynamic ligation screen assay. The tracer was added to rows A, C, E, G, I, K, M, O and buffer was added to the remaining rows for blank subtraction.

**Plate 2 Layout**

| Plate 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | A109 | | A110 | | A111 | | A112 | | A113 | | A114 | | A115 | | A116 | | A117 | | A118 | | A119 | | A120 | |
| B | | | | | | | | | | | | | | | | | | | | | | | | |
| C | A121 | | A122 | | A123 | | A124 | | A125 | | A126 | | A127 | | A128 | | A129 | | A083 | | A084 | | A085 | |
| D | | | | | | | | | | | | | | | | | | | | | | | | |
| E | A086 | | A087 | | A088 | | A089 | | A090 | | A091 | | A092 | | A093 | | A094 | | A095 | | A096 | | A097 | |
| F | | | | | | | | | | | | | | | | | | | | | | | | |
| G | A098 | | A099 | | A100 | | A101 | | A102 | | A103 | | A104 | | A105 | | A106 | | A107 | | A108 | | A130 | |
| H | | | | | | | | | | | | | | | | | | | | | | | | |
| I | A131 | | A132 | | A133 | | A134 | | A135 | | A136 | | A137 | | A138 | | A139 | | A140 | | A141 | | A142 | |
| J | | | | | | | | | | | | | | | | | | | | | | | | |
| K | A143 | | A144 | | A145 | | A146 | | A147 | | A148 | | A149 | | A150 | | A151 | | A152 | | A153 | | A154 | |
| L | | | | | | | | | | | | | | | | | | | | | | | | |
| M | A155 | | A156 | | A157 | | A158 | | A159 | | A160 | | A161 | | A162 | | A163 | | A164 | | A165 | | | |
| N | | | | | | | | | | | | | | | | | | | | | | | | |
| O | buffer | | Ac-c001 | | Ac-c001 + Benzaldehyde | | Ac-H3(1-9)K9Me3 | | | | | | | | | | | | | | | | | |
| P | | | | | | | | | | | | | | | | | | | | | | | | |

**P (Perpendicular)**

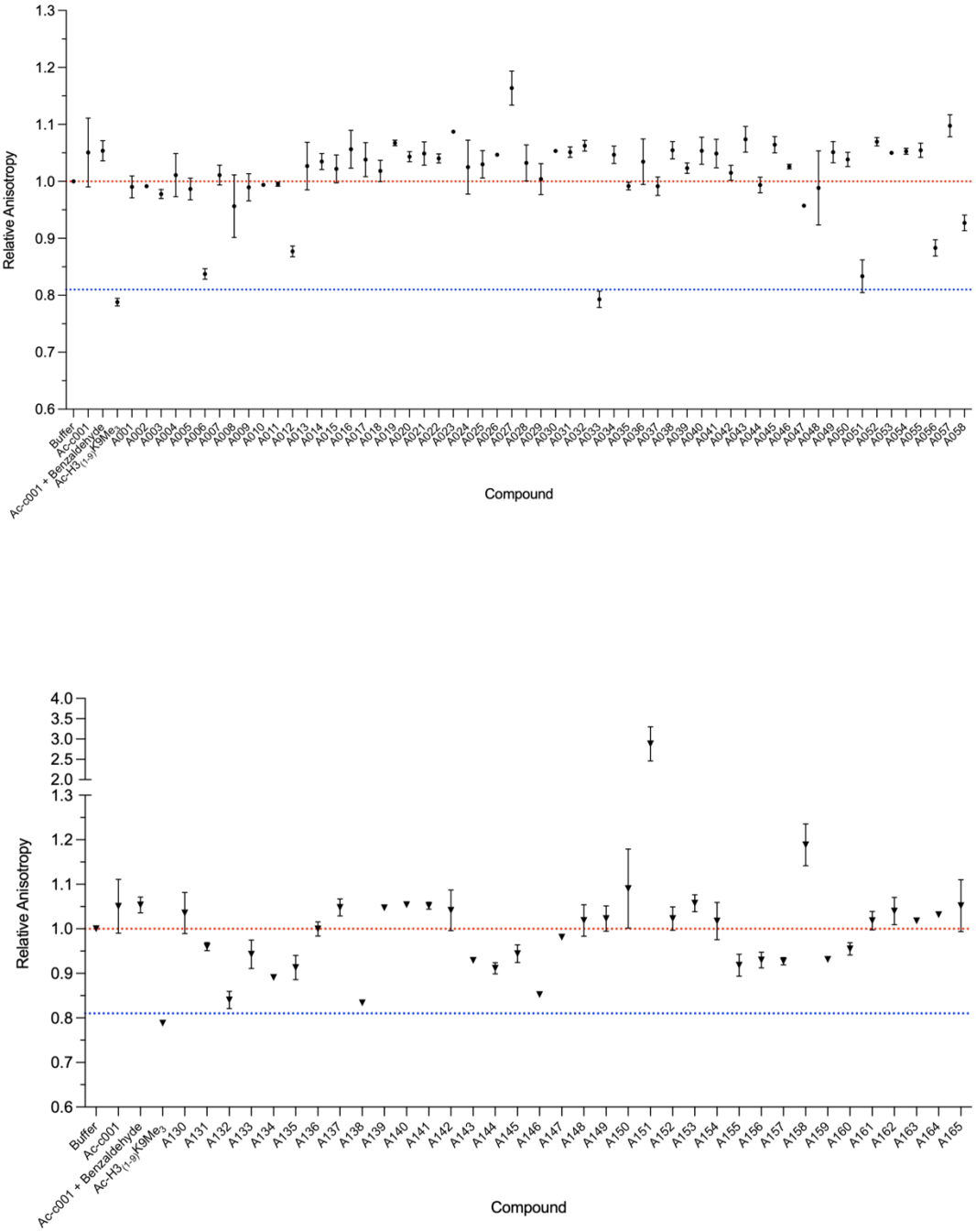
| Plate 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 13442 | 10434 | 14366 | 11893 | 14044 | 11376 | 13769 | 10636 | 12496 | 9411 | 11555 | 10285 | 10896 | 8472 | 10348 | 7926 | 10393 | 8835 | 9043 | 7257 | 7430 | 6468 | 5660 | 5048 |
| B | 227 | 227 | 243 | 239 | 225 | 215 | 212 | 216 | 222 | 229 | 221 | 219 | 298 | 295 | 308 | 301 | 276 | 249 | 231 | 207 | 218 | 208 | 207 | 230 |
| C | 14585 | 11041 | 13741 | 11585 | 13650 | 11788 | 14047 | 11063 | 12658 | 8682 | 11336 | 10165 | 10825 | 8510 | 10590 | 8726 | 10214 | 8433 | 8939 | 7547 | 9031 | 7851 | 7065 | 6511 |
| D | 226 | 227 | 249 | 242 | 224 | 225 | 223 | 223 | 227 | 228 | 367 | 278 | 263 | 235 | 245 | 240 | 233 | 185 | 219 | 203 | 210 | 218 | 250 | 255 |
| E | 15573 | 11922 | 14758 | 12596 | 12907 | 11219 | 15182 | 12076 | 11652 | 9564 | 11033 | 9240 | 11472 | 8976 | 9219 | 7688 | 10856 | 8862 | 9618 | 7830 | 8717 | 7675 | 6288 | 5665 |
| F | 223 | 227 | 212 | 210 | 256 | 254 | 220 | 225 | 209 | 211 | 228 | 232 | 232 | 234 | 225 | 226 | 228 | 171 | 258 | 248 | 229 | 222 | 212 | 213 |
| G | 15378 | 12320 | 15014 | 12793 | 14570 | 12539 | 14361 | 11745 | 11509 | 9171 | 12146 | 9704 | 9531 | 7518 | 9502 | 7208 | 10412 | 8257 | 8479 | 6825 | 9109 | 7501 | 6375 | 6122 |
| H | 235 | 238 | 223 | 223 | 229 | 235 | 231 | 230 | 320 | 340 | 219 | 229 | 233 | 236 | 228 | 209 | 194 | 196 | 441 | 405 | 243 | 235 | 209 | 225 |
| I | 14273 | 11133 | 14925 | 11850 | 12488 | 11037 | 13599 | 11261 | 11148 | 7696 | 11390 | 9305 | 10708 | 8423 | 10545 | 8747 | 9192 | 7340 | 8770 | 6294 | 8330 | 6631 | 6571 | 5702 |
| J | 208 | 214 | 245 | 246 | 192 | 205 | 201 | 213 | 194 | 230 | 213 | 213 | 223 | 225 | 242 | 232 | 204 | 184 | 193 | 178 | 168 | 177 | 191 | 217 |
| K | 14229 | 11401 | 13834 | 11447 | 7503 | 6071 | 4242 | 2825 | 11295 | 8923 | 11078 | 8894 | 10171 | 8433 | 9028 | 5264 | 3443 | 3249 | 8211 | 6545 | 8289 | 6343 | 3213 | 2626 |
| L | 221 | 237 | 193 | 188 | 236 | 260 | 255 | 253 | 175 | 175 | 216 | 214 | 201 | 212 | 184 | 211 | 3115 | 3028 | 183 | 184 | 187 | 187 | 237 | 280 |
| M | 14135 | 11334 | 13975 | 11769 | 10213 | 8641 | 9006 | 7766 | 11613 | 9194 | 10785 | 9277 | 9617 | 8322 | 7547 | 6229 | 8701 | 7079 | 9088 | 6463 | 7996 | 7090 | | |
| N | 238 | 242 | 257 | 246 | 322 | 343 | 163 | 172 | 205 | 231 | 244 | 235 | 223 | 223 | 221 | 187 | 199 | 182 | 168 | 144 | 807 | 819 | | |
| O | 14425 | 12246 | 13103 | 10574 | 13117 | 10278 | 12865 | 11244 | | | | | | | | | | | | | | | | |
| P | 48 | 42 | 228 | 249 | 330 | 42 | 47 | 44 | | | | | | | | | | | | | | | | |

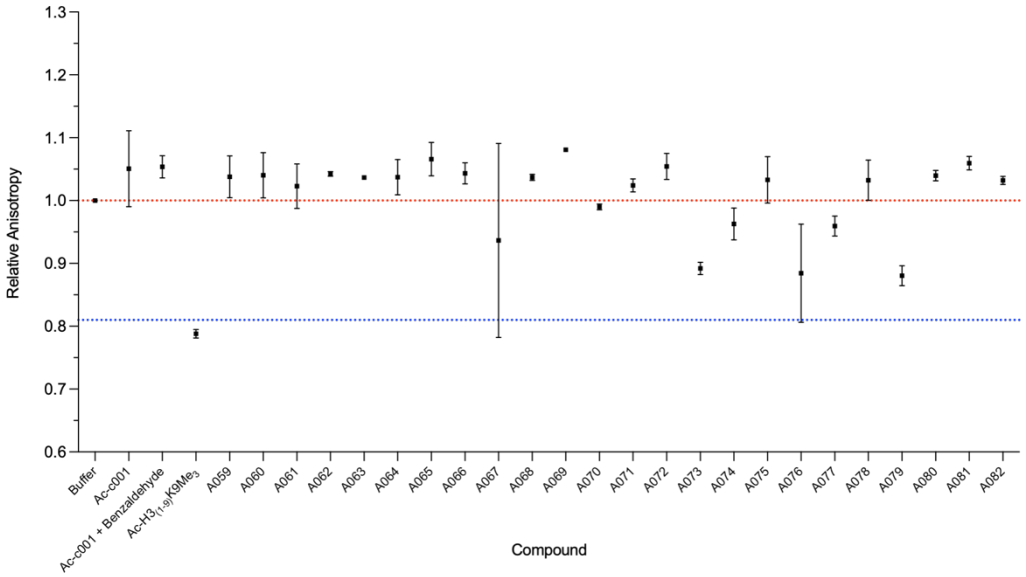
**S (Parallel)**

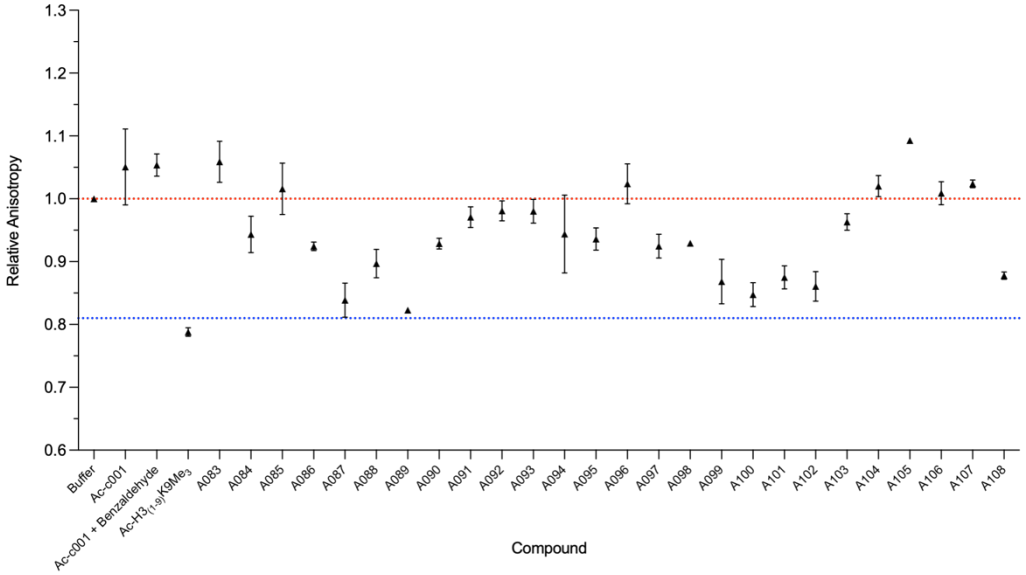
| Plate 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 15535 | 12050 | 16465 | 13607 | 16170 | 12939 | 15688 | 12081 | 14444 | 10829 | 13334 | 11813 | 12638 | 9801 | 11872 | 9097 | 12021 | 10209 | 10459 | 8418 | 8677 | 7482 | 6559 | 5797 |
| B | 233 | 233 | 253 | 248 | 229 | 222 | 218 | 221 | 233 | 241 | 224 | 224 | 311 | 308 | 318 | 311 | 286 | 260 | 237 | 214 | 231 | 219 | 214 | 234 |
| C | 16758 | 12669 | 15796 | 13244 | 15656 | 13422 | 16015 | 12573 | 14479 | 10004 | 13126 | 11736 | 12341 | 9749 | 12296 | 10114 | 11869 | 9777 | 10491 | 8796 | 10429 | 9009 | 8272 | 7559 |
| D | 232 | 230 | 281 | 269 | 229 | 233 | 228 | 226 | 232 | 232 | 386 | 284 | 278 | 246 | 252 | 244 | 239 | 193 | 224 | 208 | 223 | 231 | 292 | 296 |
| E | 17875 | 13696 | 16788 | 14245 | 14837 | 12837 | 17177 | 13659 | 13373 | 10992 | 12769 | 10652 | 13309 | 10378 | 10684 | 8868 | 12573 | 10142 | 11123 | 9024 | 10184 | 8907 | 7231 | 6487 |
| F | 231 | 234 | 221 | 219 | 310 | 306 | 229 | 229 | 214 | 217 | 233 | 237 | 250 | 255 | 232 | 231 | 233 | 181 | 318 | 304 | 243 | 236 | 228 | 228 |
| G | 17694 | 14157 | 17170 | 14518 | 16579 | 14207 | 16399 | 13358 | 13140 | 10537 | 14005 | 11208 | 11116 | 8732 | 11167 | 8467 | 12079 | 9611 | 10108 | 8170 | 10394 | 8548 | 7465 | 7092 |
| H | 255 | 256 | 235 | 233 | 238 | 243 | 234 | 235 | 423 | 450 | 226 | 231 | 249 | 251 | 230 | 214 | 201 | 204 | 731 | 671 | 260 | 253 | 220 | 230 |
| I | 16498 | 12834 | 16914 | 13477 | 14432 | 12668 | 15553 | 12854 | 12753 | 8841 | 13196 | 10811 | 12539 | 9820 | 11942 | 9916 | 10744 | 8570 | 10249 | 7364 | 9753 | 7747 | 7704 | 6613 |
| J | 218 | 220 | 272 | 268 | 202 | 216 | 210 | 217 | 210 | 243 | 217 | 217 | 231 | 235 | 256 | 247 | 211 | 192 | 199 | 188 | 178 | 188 | 201 | 224 |
| K | 16383 | 13133 | 15890 | 13112 | 8690 | 7002 | 4856 | 3238 | 13084 | 10327 | 12851 | 10388 | 11882 | 9793 | 10526 | 6250 | 5659 | 5266 | 9610 | 7607 | 9686 | 7432 | 3782 | 3069 |
| L | 261 | 281 | 209 | 203 | 289 | 312 | 317 | 312 | 183 | 181 | 228 | 226 | 210 | 226 | 197 | 238 | 5138 | 4943 | 214 | 203 | 200 | 196 | 296 | 345 |
| M | 16307 | 13015 | 16173 | 13575 | 11903 | 10075 | 10795 | 9217 | 13376 | 10566 | 12460 | 10686 | 11212 | 9657 | 8849 | 7243 | 10119 | 8238 | 10603 | 7536 | 9475 | 8315 | | |
| N | 295 | 301 | 351 | 334 | 511 | 534 | 167 | 179 | 231 | 253 | 262 | 253 | 231 | 231 | 249 | 196 | 206 | 192 | 179 | 151 | 997 | 1009 | | |
| O | 16767 | 14258 | 15414 | 12274 | 15387 | 12030 | 14513 | 12668 | | | | | | | | | | | | | | | | |
| 2 | 59 | 49 | 230 | 250 | 367 | 51 | 55 | 53 | | | | | | | | | | | | | | | | |

**Figure A5.5:** Layout of 384-well plate 1 and raw fluorescence intensity data for the parallel and perpendicular channels in the dynamic ligation screen assay. The tracer was added to rows A, C, E, G, I, K, M, O and buffer was added to the remaining rows for blank subtraction.

**Figure A5.6:** Results of the dynamic ligation screen. The red dashed line represents the buffer control and the blue dashed line represents the hit threshold (relative anisotropy of positive control + [3 × standard deviation of the positive control]). Error bars indicate the standard deviation of duplicate data.




Relative Anisotropy
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
Buffer
Ac-c001
Ac-c001 + Benzaldehyde
Ac-H3(1-9)K9Me3
A059
A060
A061
A062
A063
A064
A065
A066
A067
A068
A069
A070
A071
A072
A073
A074
A075
A076
A077
A078
A079
A080
A081
A082
Compound


Relative Anisotropy
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
Buffer
Ac-c001
Ac-c001 + Benzaldehyde
Ac-H3(1-9)K9Me3
A083
A084
A085
A086
A087
A088
A089
A090
A091
A092
A093
A094
A095
A096
A097
A098
A099
A100
A101
A102
A103
A104
A105
A106
A107
A108
Compound

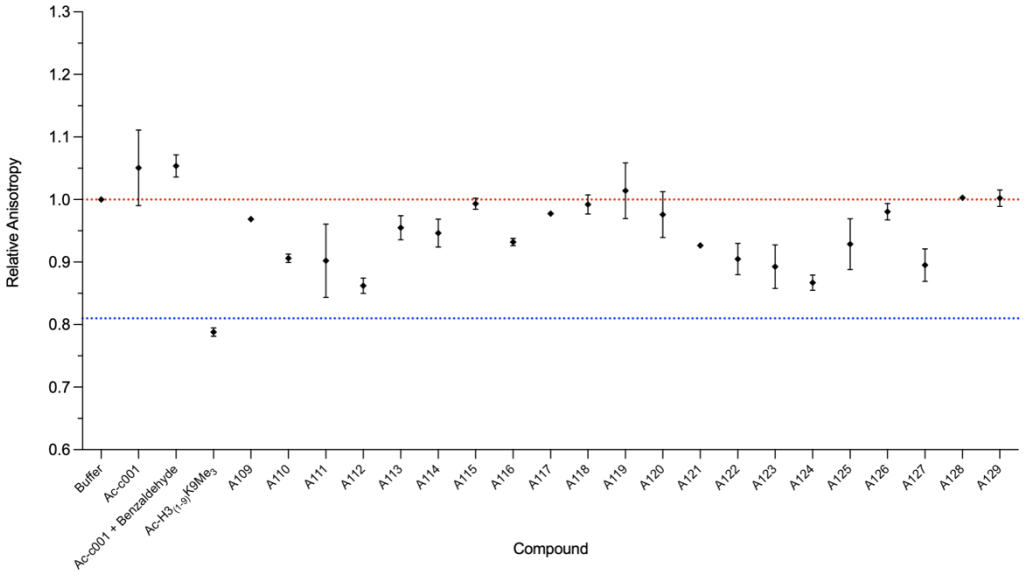
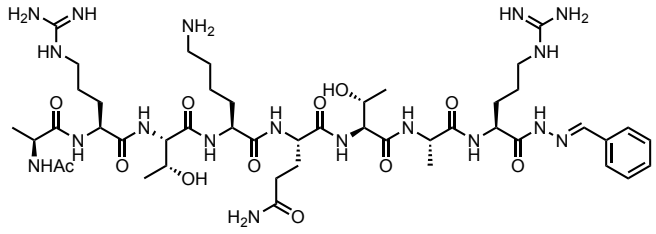
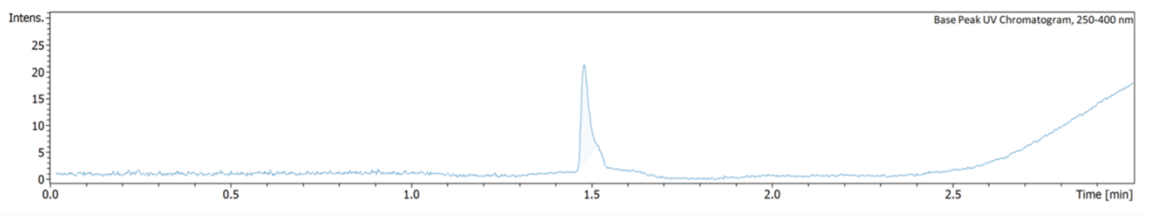
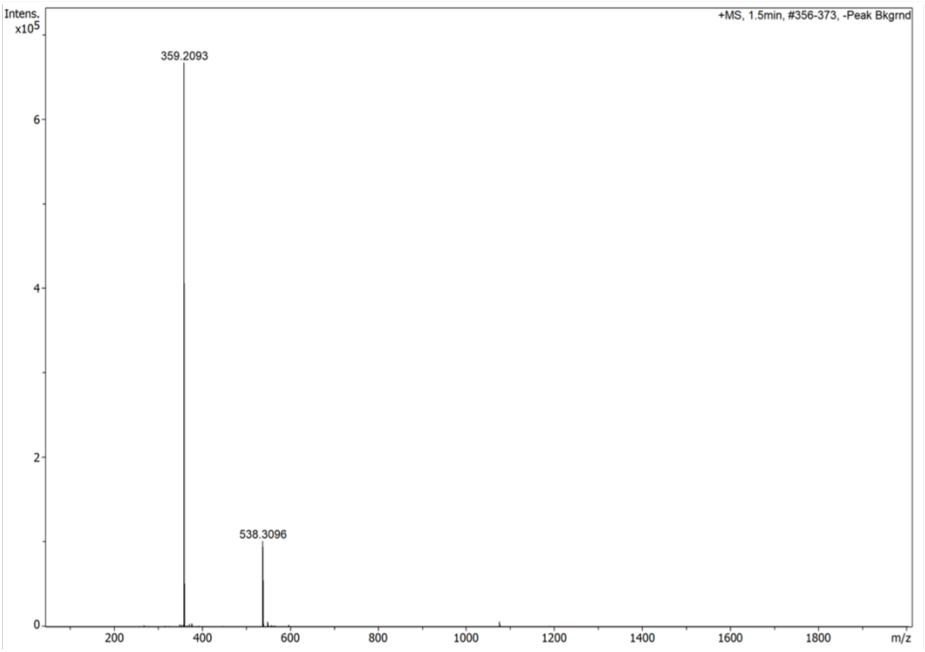
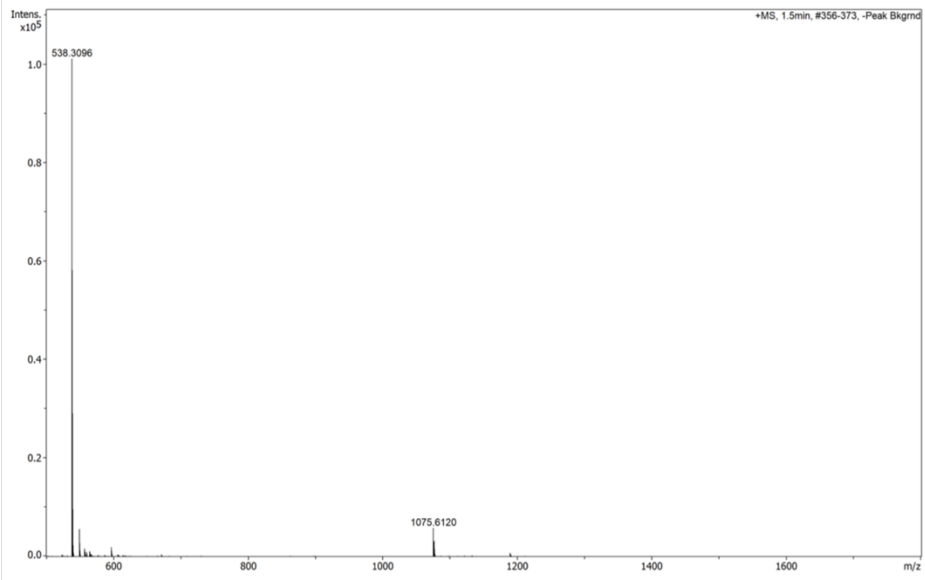

Relative Anisotropy
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
Buffer
Ac-c001
Ac-c001 + Benzaldehyde
Ac-H3(1-9)K9Me3
A109
A110
A111
A112
A113
A114
A115
A116
A117
A118
A119
A120
A121
A122
A123
A124
A125
A126
A127
A128
A129
Compound

**Figure A5.7:** LC-MS chromatograms and $ESI^+$ mass spectra for anchor peptide Ac-c001 (synthesised by Yue Li, University of Birmingham).



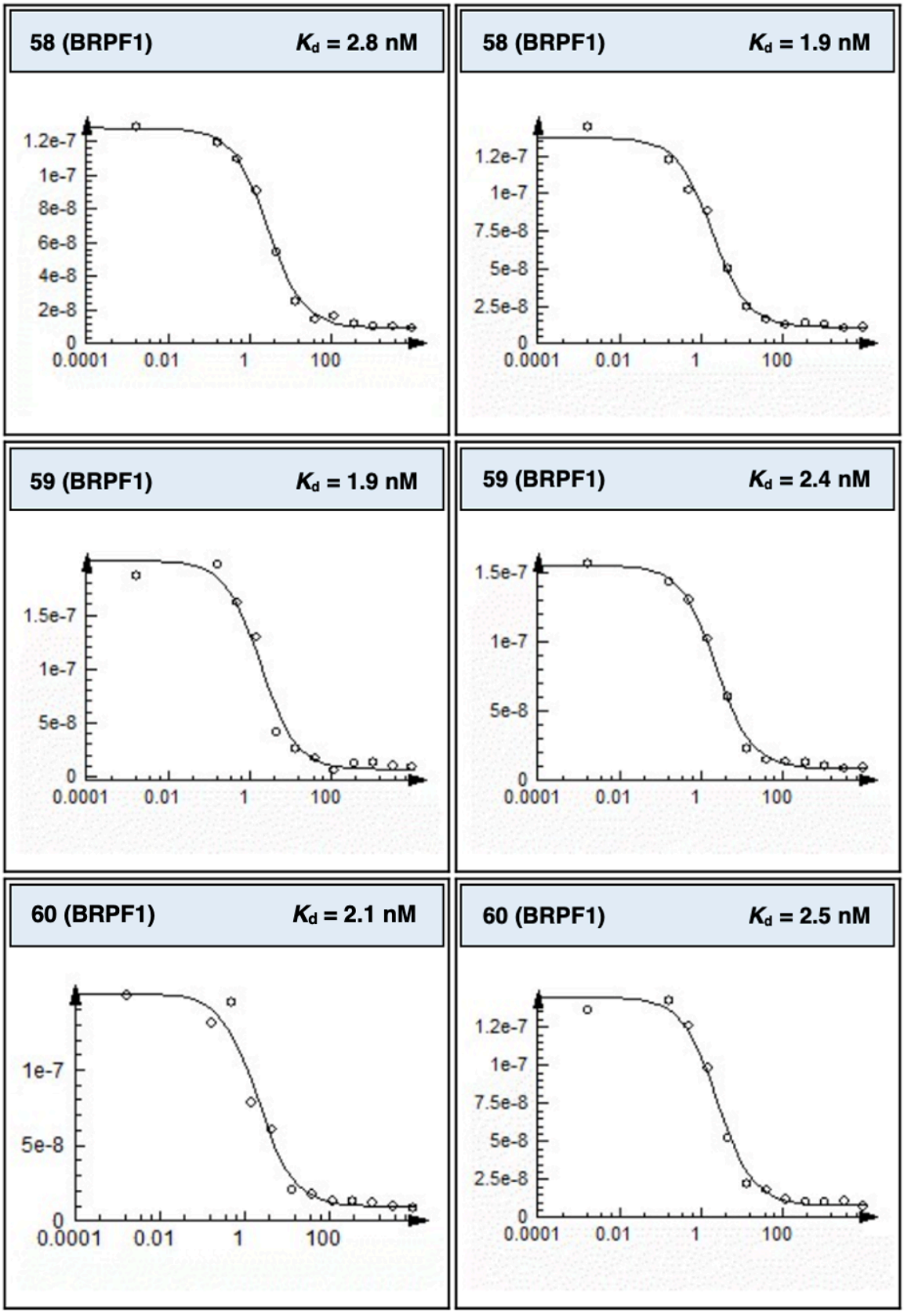



## Appendix 6 BROMO*scan*

**Figure A6.1:** BROMO*scan* dose-response curves for compounds **58**, **59**, **60**, **63** and **67**; against BRPF1 (BRD). Graphs produced by Eurofins DiscoverX Corporation (San Diego, USA).



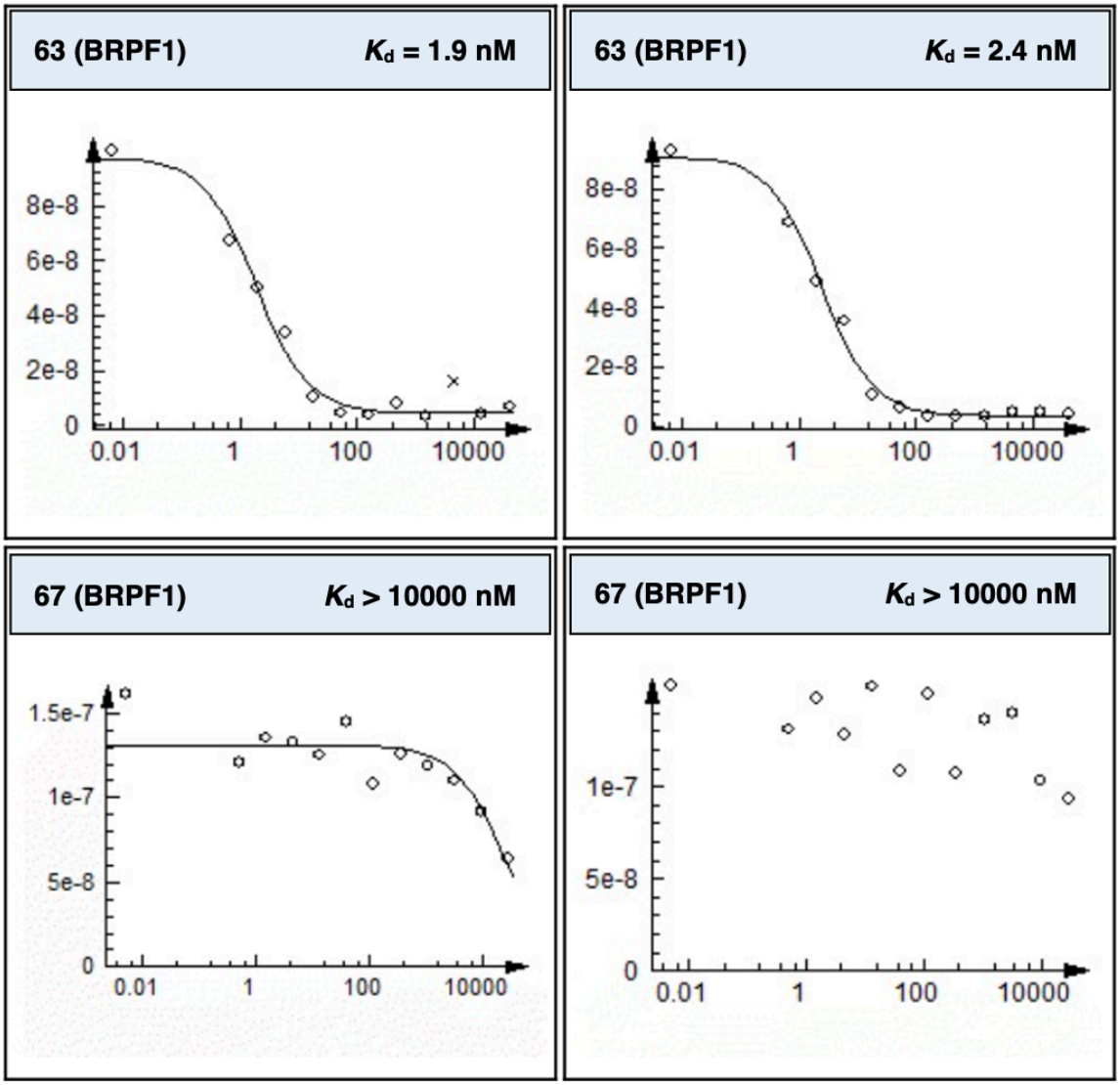
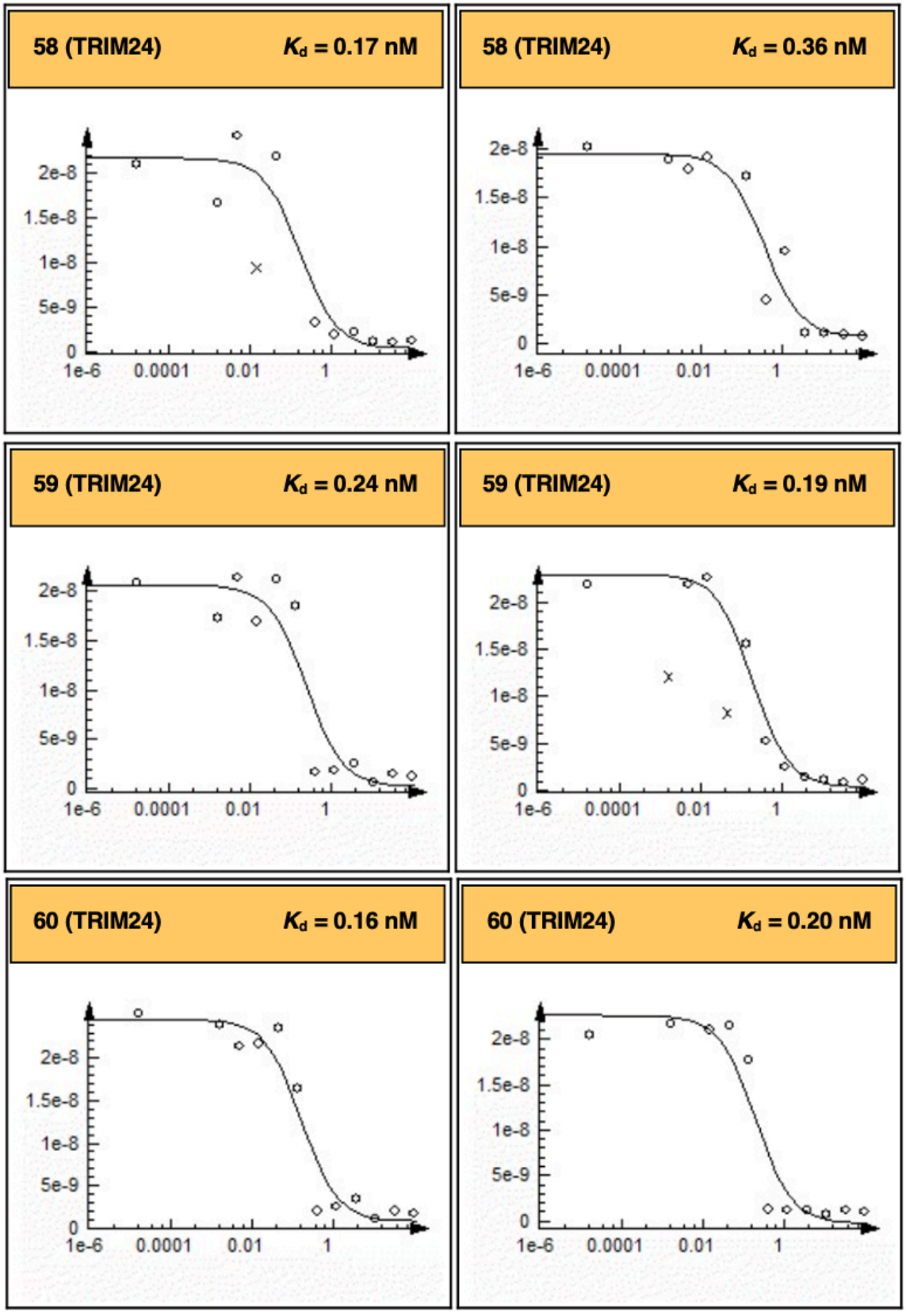
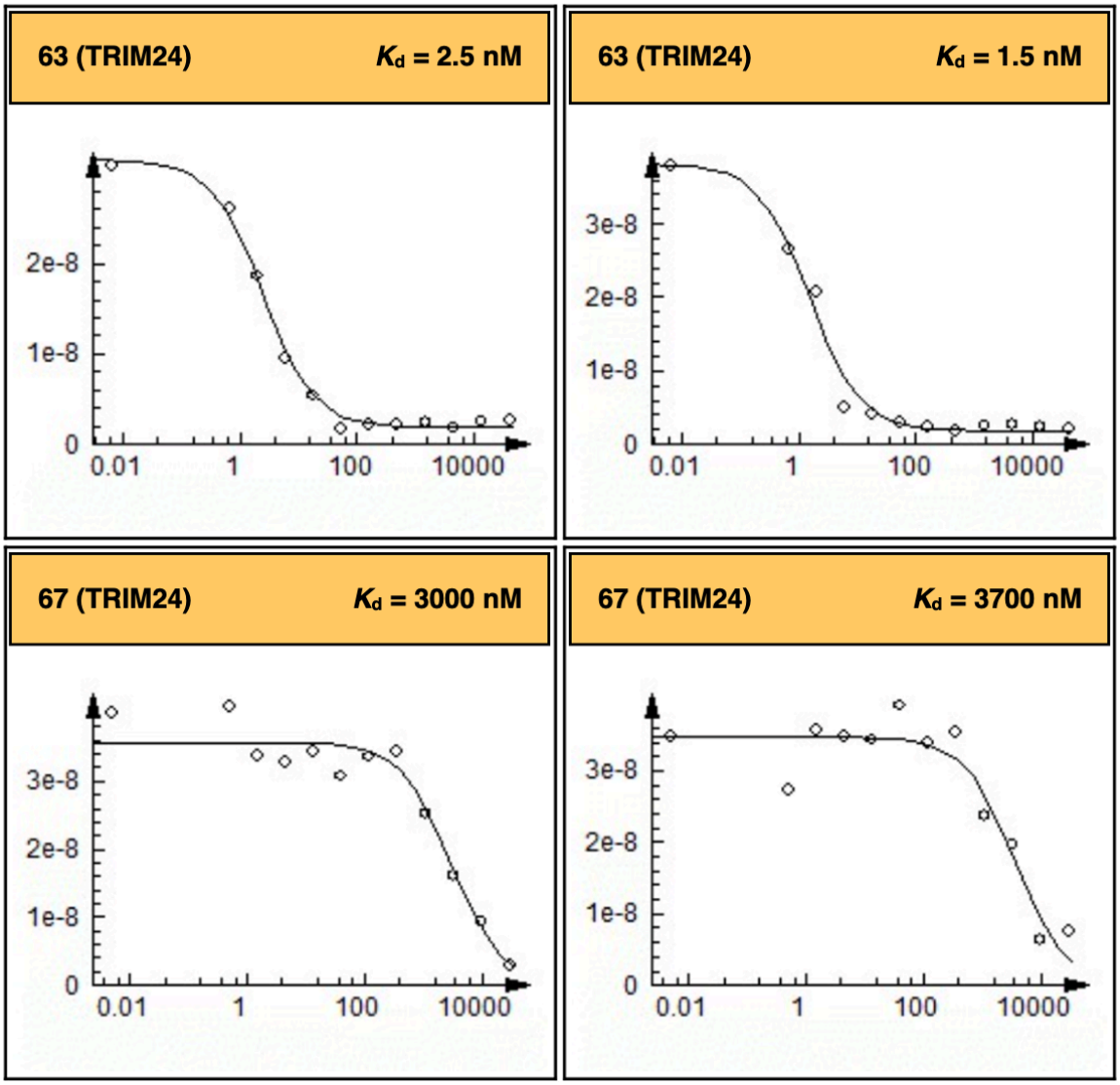

63 (BRPF1)
$K_d$ = 1.9 nM
8e-8
6e-8
4e-8
2e-8
0
0.01
1
100
10000
63 (BRPF1)
$K_d$ = 2.4 nM
8e-8
6e-8
4e-8
2e-8
0
0.01
1
100
10000
67 (BRPF1)
$K_d$ > 10000 nM
1.5e-7
1e-7
5e-8
0
0.01
1
100
10000
67 (BRPF1)
$K_d$ > 10000 nM
1e-7
5e-8
0
0.01
1
100
10000

**Figure A6.2:** BROMO*scan* dose-response curves for compounds **58**, **59**, **60**, **63** and **67**; against TRIM24 (PHD-BRD). Graphs produced by Eurofins DiscoverX Corporation (San Diego, USA).




63 (TRIM24)
$K_d$ = 2.5 nM
2e-8
1e-8
0
0.01
1
100
10000
63 (TRIM24)
$K_d$ = 1.5 nM
3e-8
2e-8
1e-8
0
0.01
1
100
10000
67 (TRIM24)
$K_d$ = 3000 nM
3e-8
2e-8
1e-8
0
0.01
1
100
10000
67 (TRIM24)
$K_d$ = 3700 nM
3e-8
2e-8
1e-8
0
0.01
1
100
10000

## Appendix 7 Surface Plasmon Resonance (SPR)

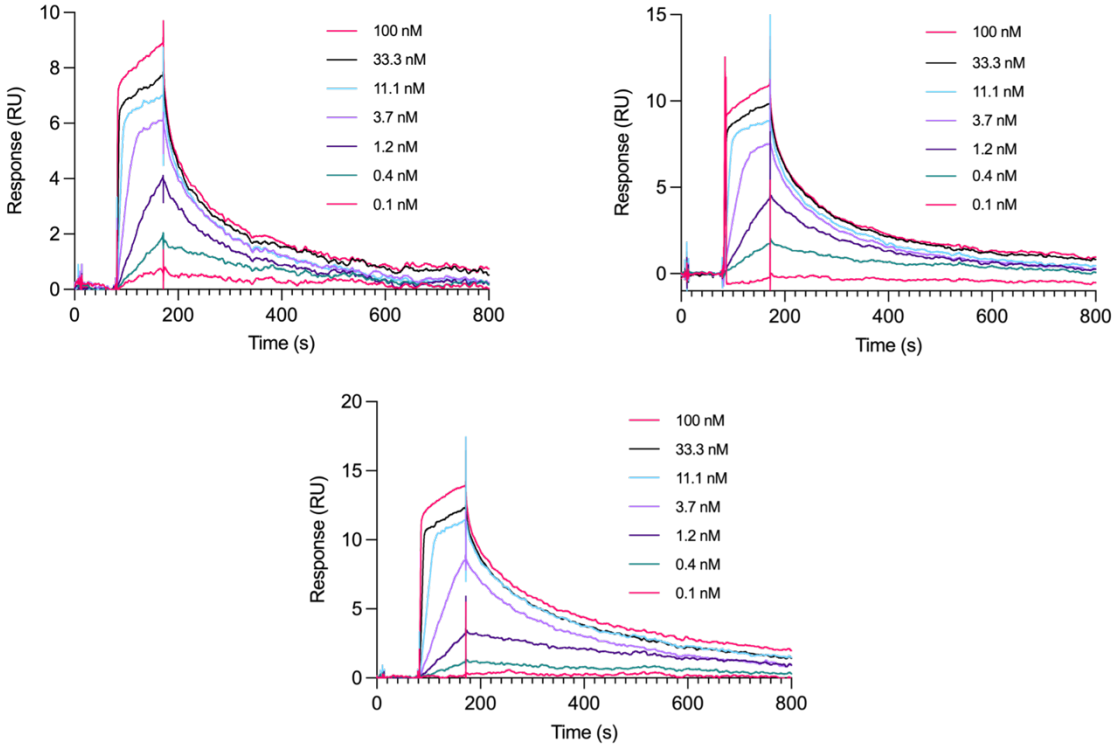



**Figure A7.1:** Raw sensogram data (replotted in GraphPad Prism) for PDC6 (**60**) binding to TRIM24 (PHD-BRD) WT. Sensograms are reference cell subtracted.

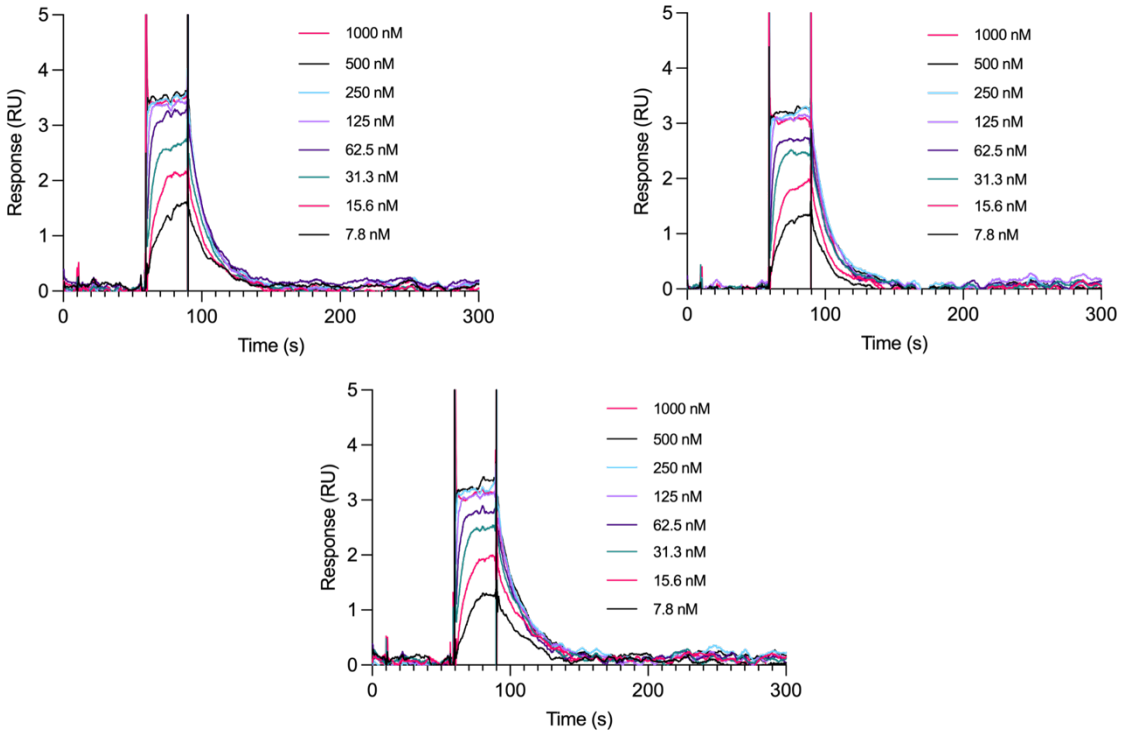



**Figure A7.2:** Raw sensogram data (replotted in GraphPad Prism) for TRIM24 BRD ligand **63** binding to TRIM24 (PHD-BRD) WT. Sensograms are reference cell subtracted.

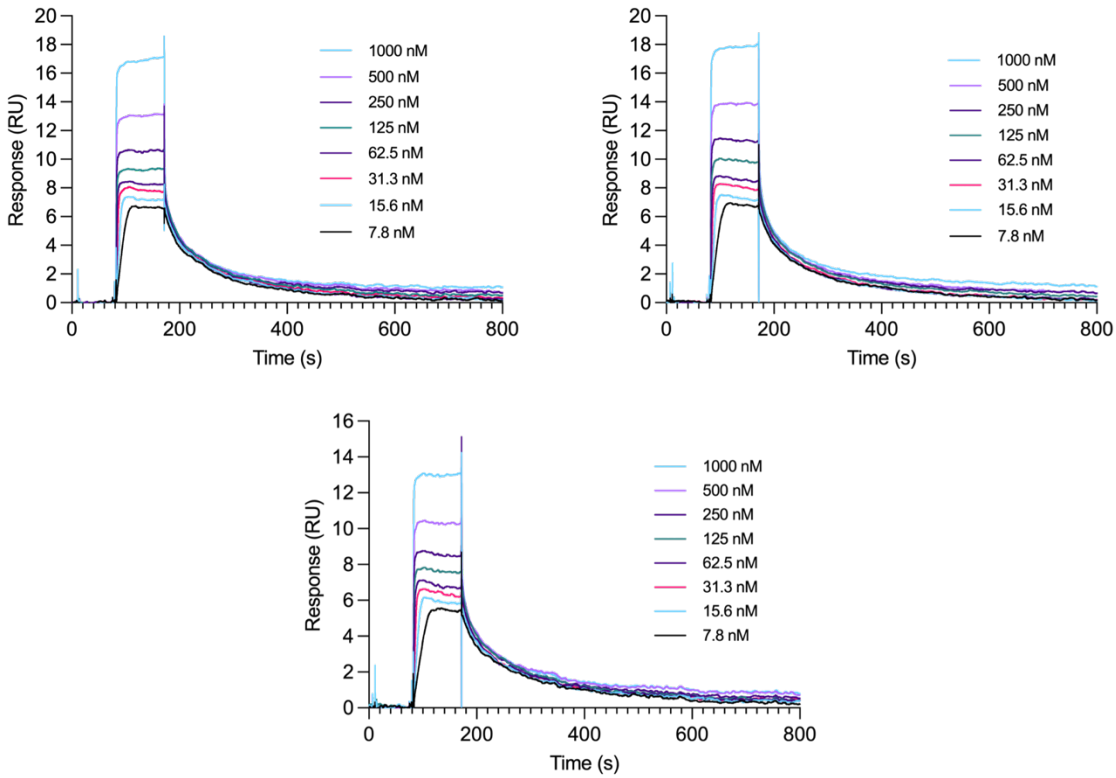



**Figure A7.3:** Raw sensogram data (replotted in GraphPad Prism) for PDC6 (**60**) binding to TRIM24 (PHD-BRD) W828A mutant. Sensograms are reference cell subtracted.

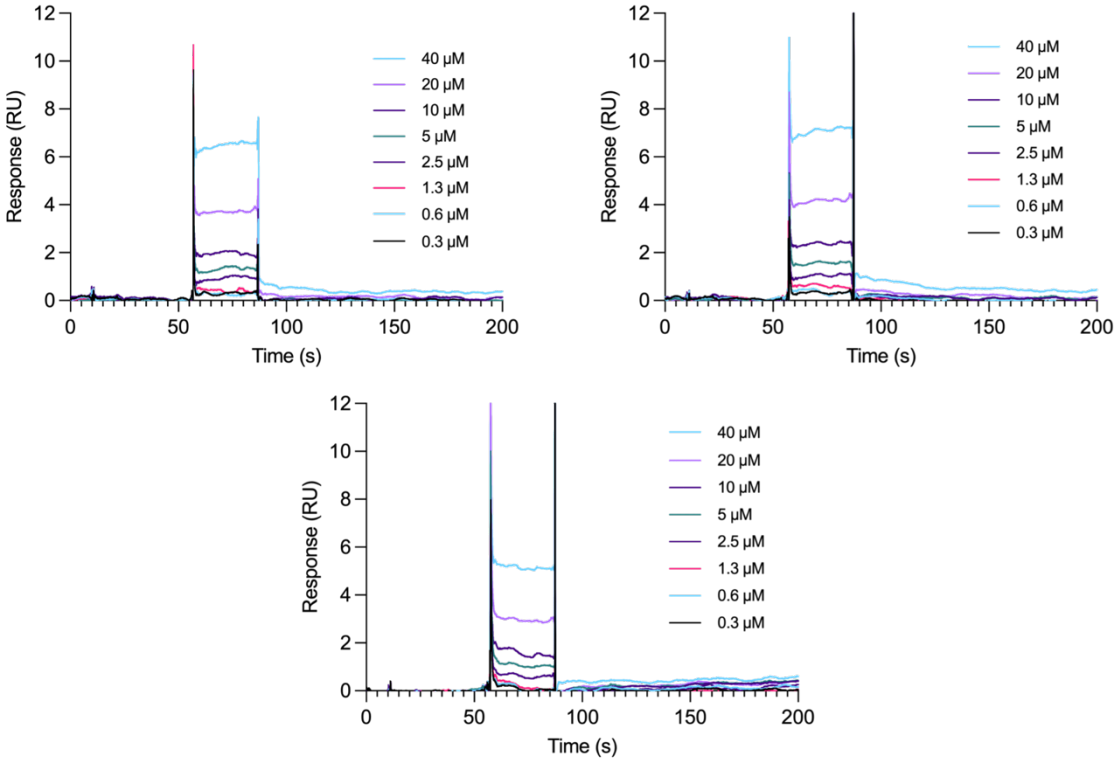



**Figure A7.4:** Raw sensogram data (replotted in GraphPad Prism) for $H3_{(1\text{-}27)}K9Me_3K18Ac$ binding to TRIM24 (PHD-BRD) W828A mutant. Sensograms are reference cell subtracted.

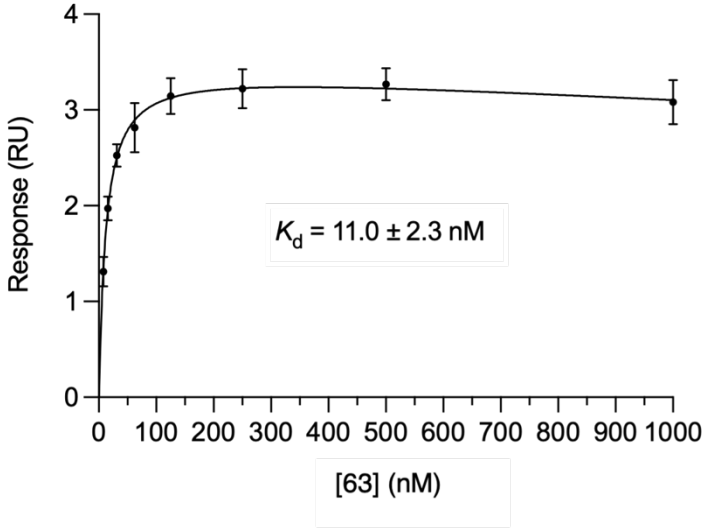



**Figure A7.5:** Binding saturation curve for TRIM24 BRD ligand **63** binding to wild type TRIM24 (PHD-BRD) WT. $K_d$ values are quoted as the mean of triplicate data ± standard error of the mean.

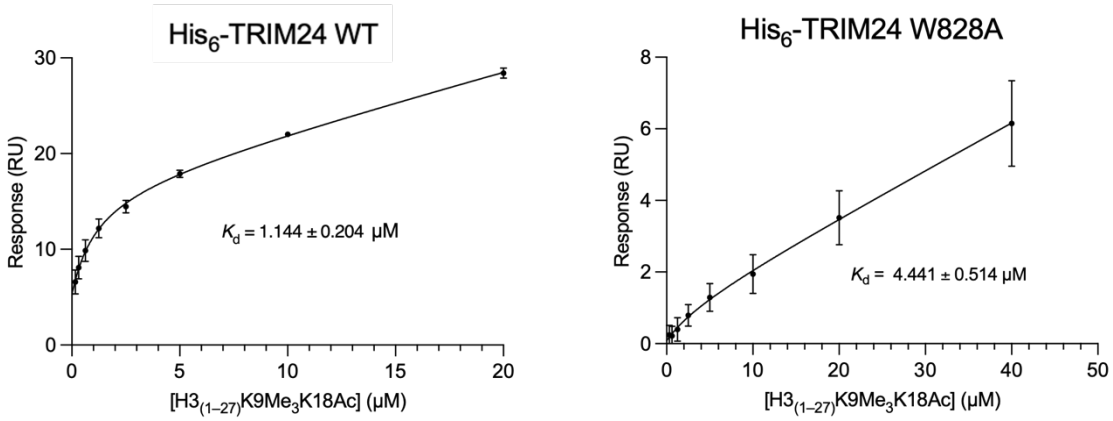



**Figure A7.6:** Binding saturation curves for $H3_{(1-27)}K9Me_3K18Ac$ binding to wild type and W828A mutant $His_6$-TRIM24. $K_d$ values are quoted as the mean of triplicate data ± standard error of the mean.

## Appendix 8 Cell Permeability Assay

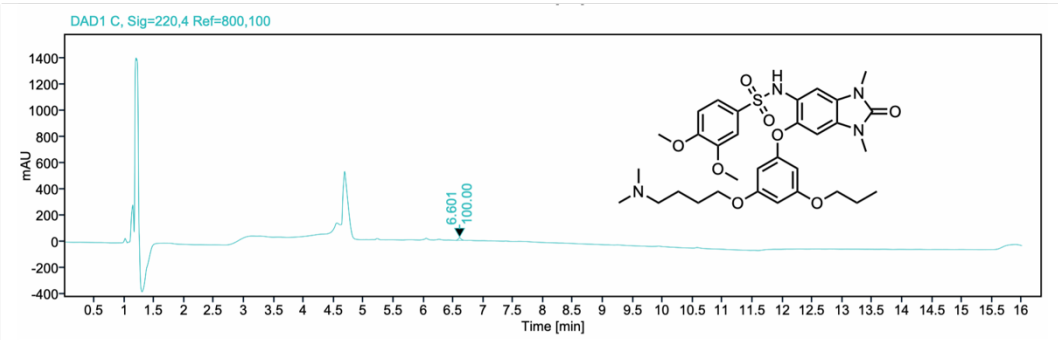



**Figure A8.1:** LC-MS chromatogram of IACS9571 ($t_R$ = 6.60 min) in HBS buffer supplemented with protease inhibitors.

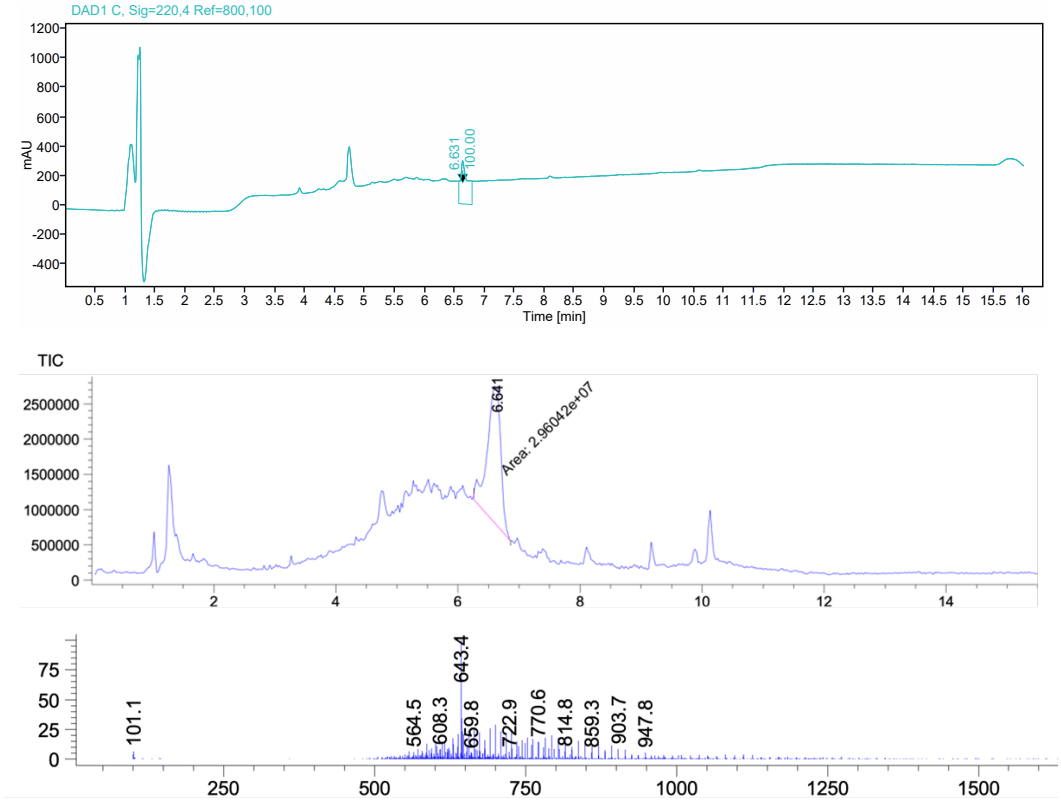



**Figure A8.2:** LC-MS chromatograms and ESI$^+$ mass spectrum ($t_R$ = 6.641 min) from the cell permeability assay for IACS9571 ($n$ = 1). IACS9571 exact mass = 642.27 (observed mass = 643.4, $[M+H]^+$).

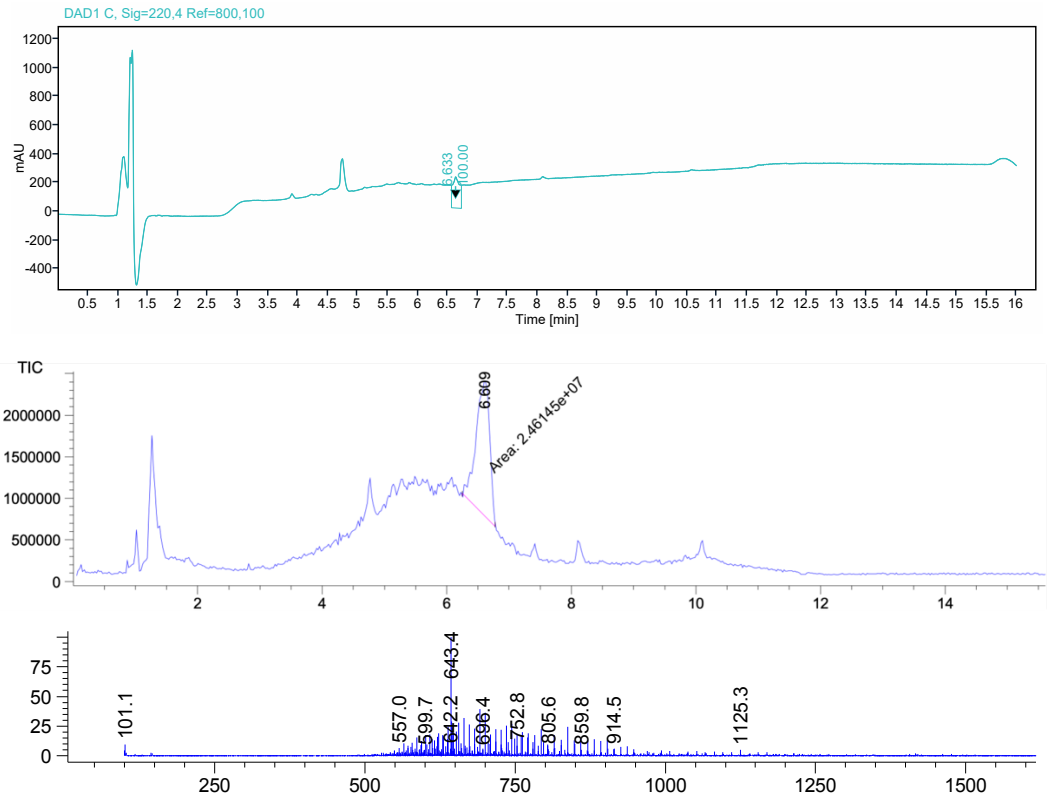



**Figure A8.3:** LC-MS chromatograms and ESI$^+$ mass spectrum ($t_R$ = 6.609 min) from the cell permeability assay for IACS9571 ($n$ = 2). IACS9571 exact mass = 642.27 (observed mass = 643.4, $[M+H]^+$).

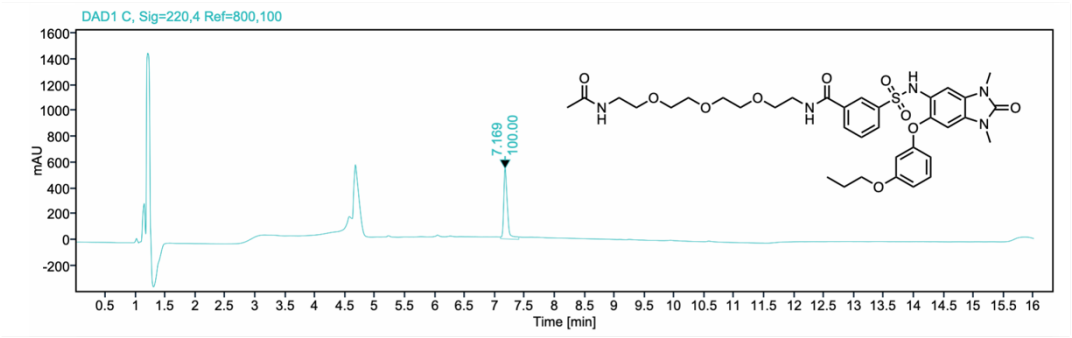



**Figure A8.4:** LC-MS chromatogram of compound **62** ($t_R$ = 7.17 min) in HBS buffer supplemented with protease inhibitors.

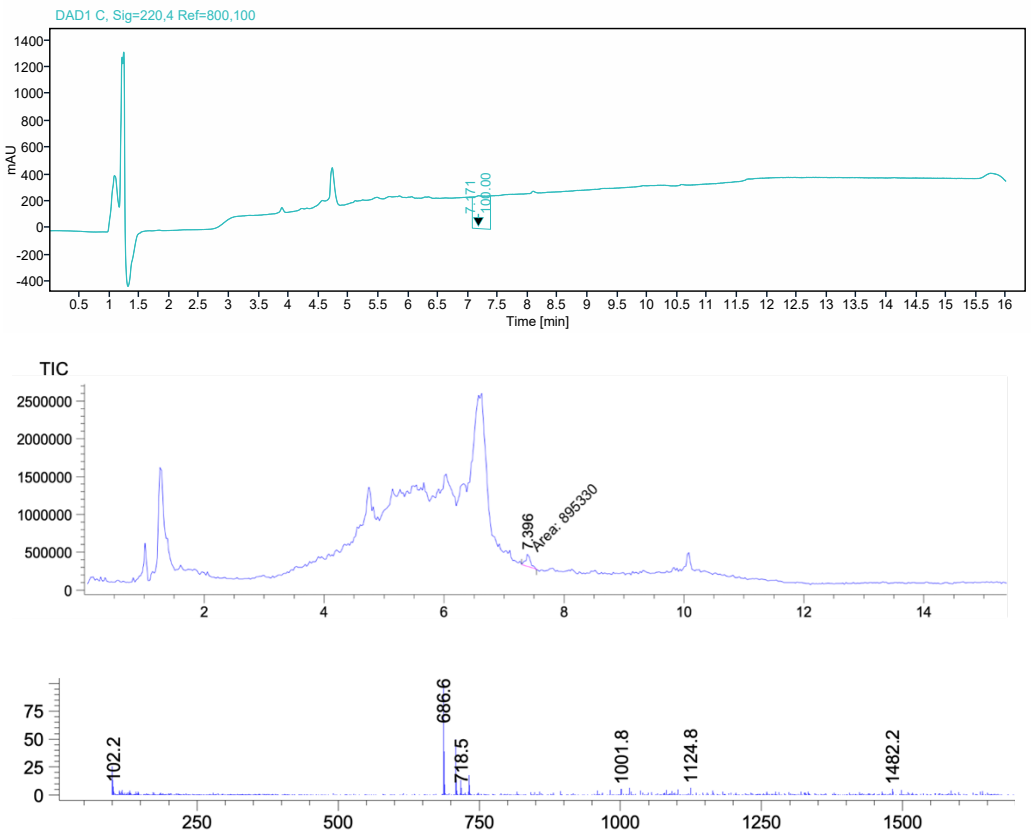



**Figure A8.5:** LC-MS chromatograms and ESI$^+$ mass spectrum ($t_R$ = 7.396 min) from the cell permeability assay for compound **62** ($n$ = 1). Compound **62** exact mass = 727.29 (observed mass = 686.6, $[M-Ac+2H]^+$).

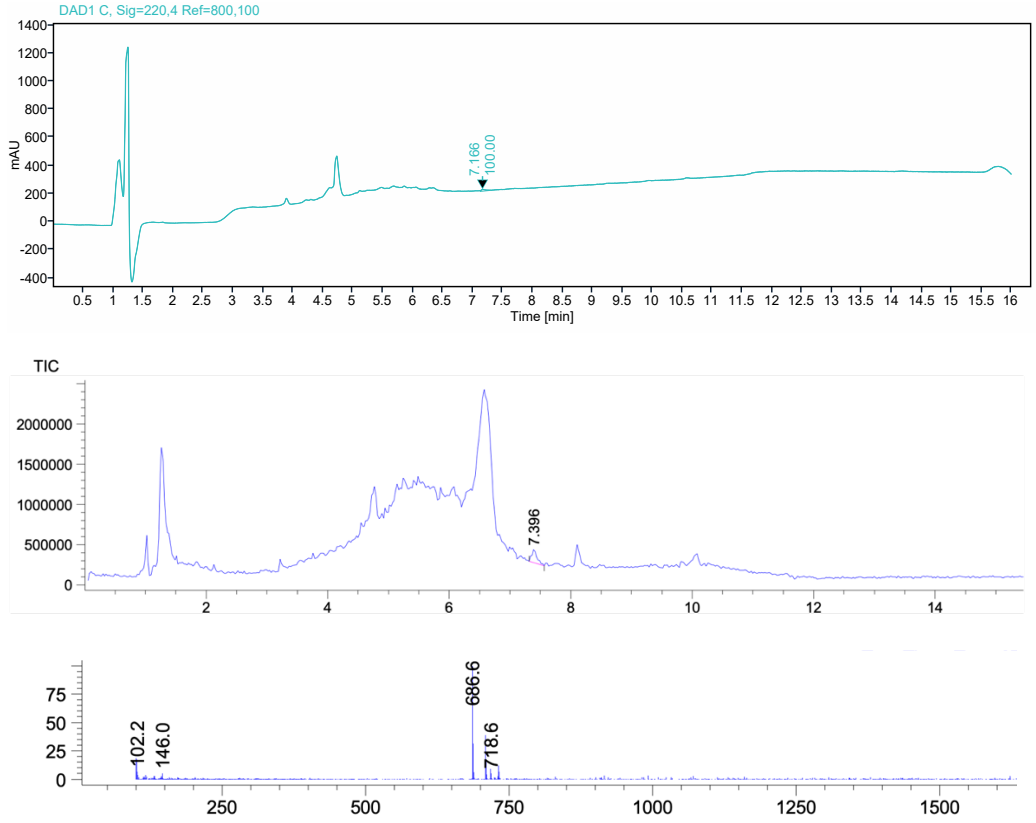



**Figure A8.6:** LC-MS chromatograms and ESI$^+$ mass spectrum ($t_R$ = 7.396 min) from the cell permeability assay for compound **62** ($n$ = 2). Compound **62** exact mass = 727.29 (observed mass = 686.6, $[M-Ac+2H]^+$).

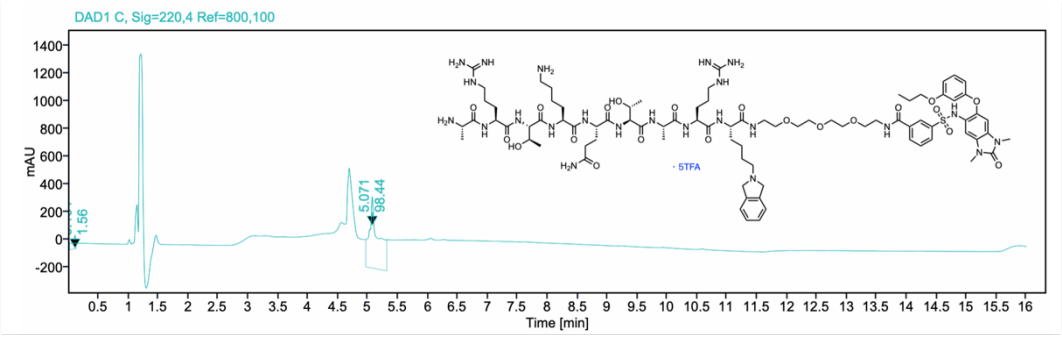



**Figure A8.7:** LC-MS chromatogram of PDC5 (compound **59**) ($t_R$ = 5.07 min) in HBS buffer supplemented with protease inhibitors.

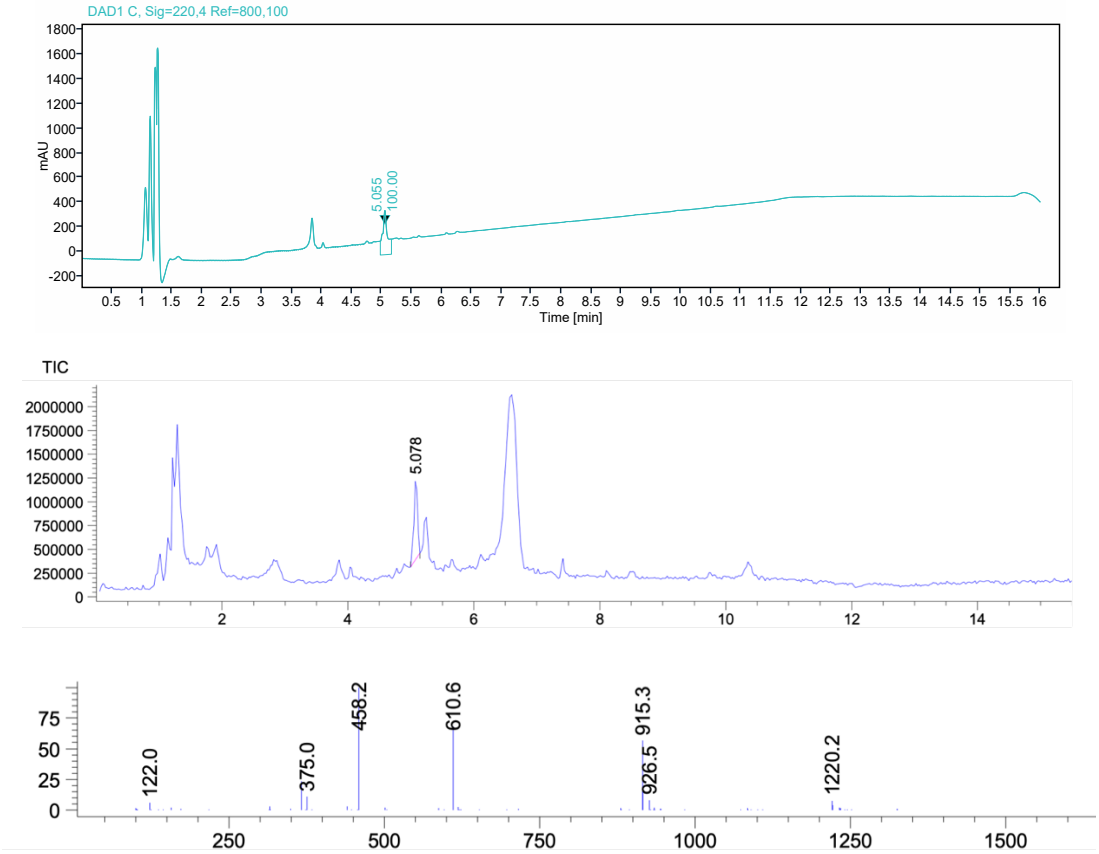



**Figure A8.8:** LC-MS chromatograms and ESI$^+$ mass spectrum ($t_R$ = 5.078 min) from the cell permeability assay for PDC5 (compound **59**) ($n$ = 1). PDC5 exact mass = 1827.95 (observed masses = 915.3, $[M-5TFA+2H]^{2+}$; 610.6, $[M-5TFA+3H]^{3+}$; 458.2, $[M-5TFA+4H]^{4+}$).

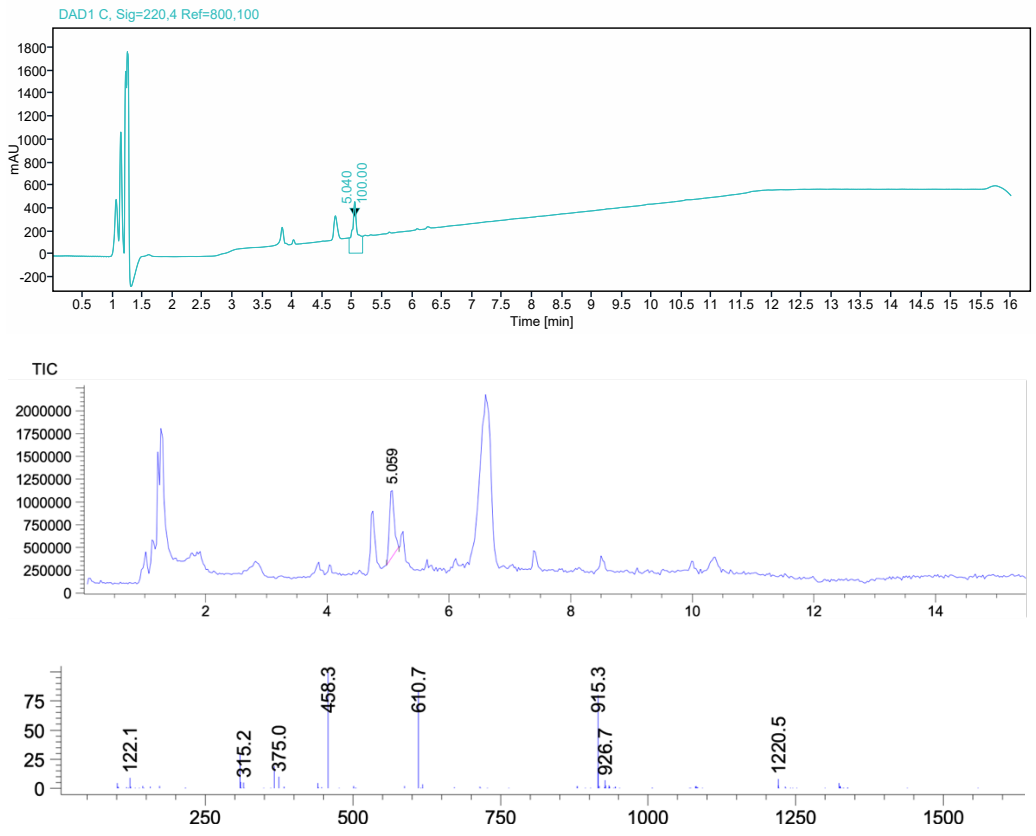
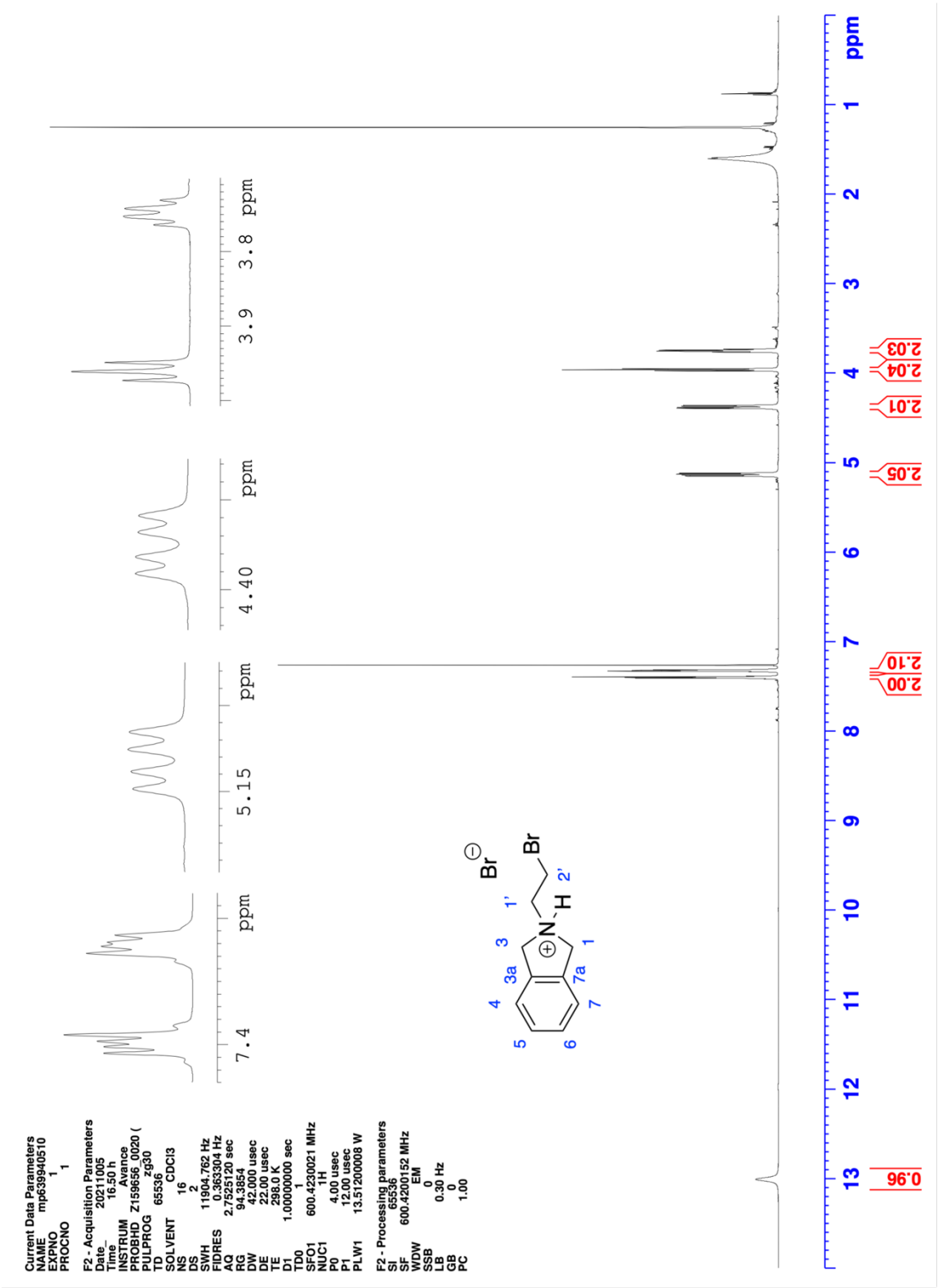


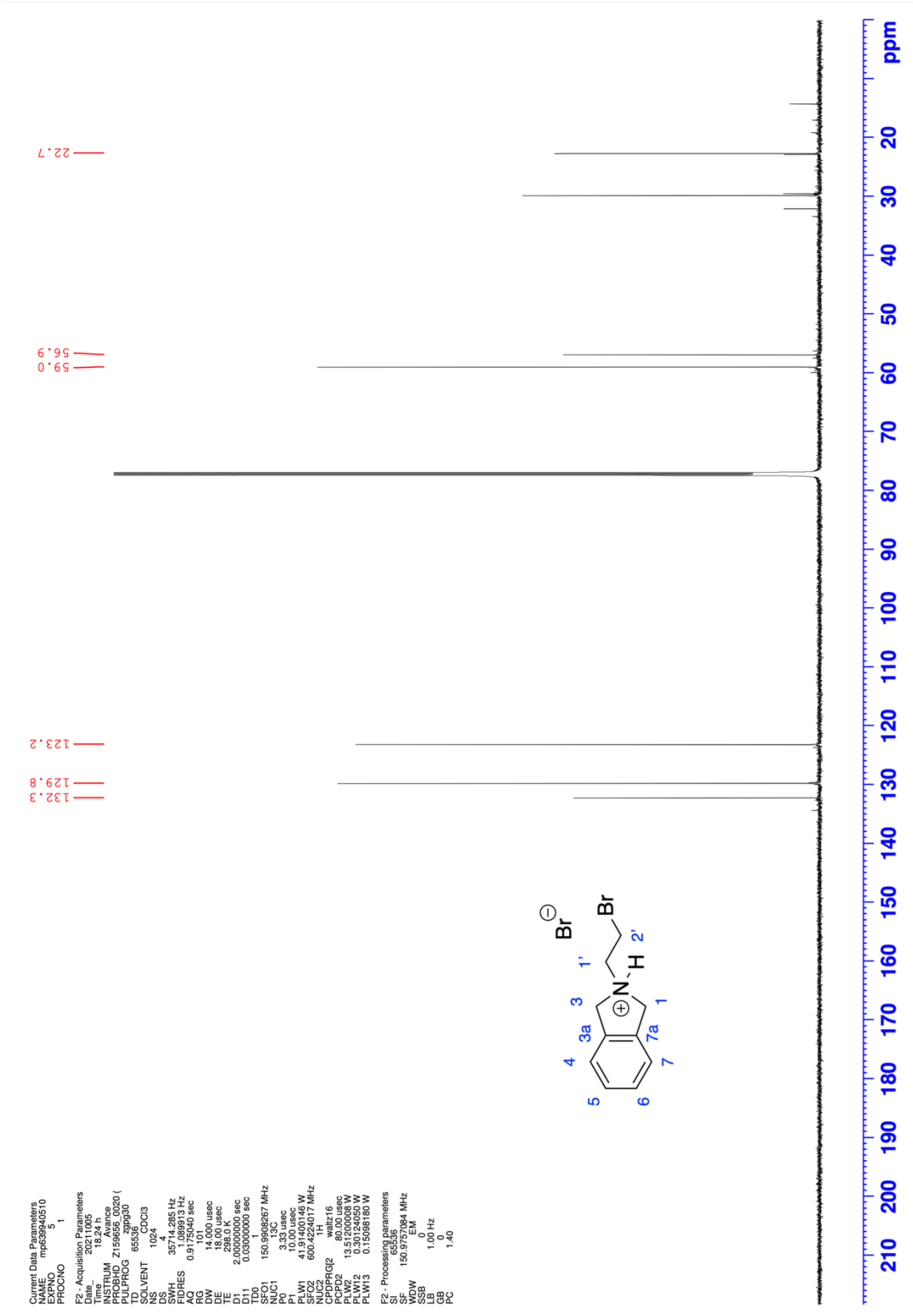

**Figure A8.9:** LC-MS chromatograms and ESI$^+$ mass spectrum ($t_R$ = 5.059 min) from the cell permeability assay for PDC5 (compound **59**) ($n$ = 2). PDC5 exact mass = 1827.95 (observed masses = 915.3, $[M-5TFA+2H]^{2+}$; 610.7, $[M-5TFA+3H]^{3+}$; 458.3, $[M-5TFA+4H]^{4+}$).

## Appendix 9 NMR Spectra of Novel Compounds

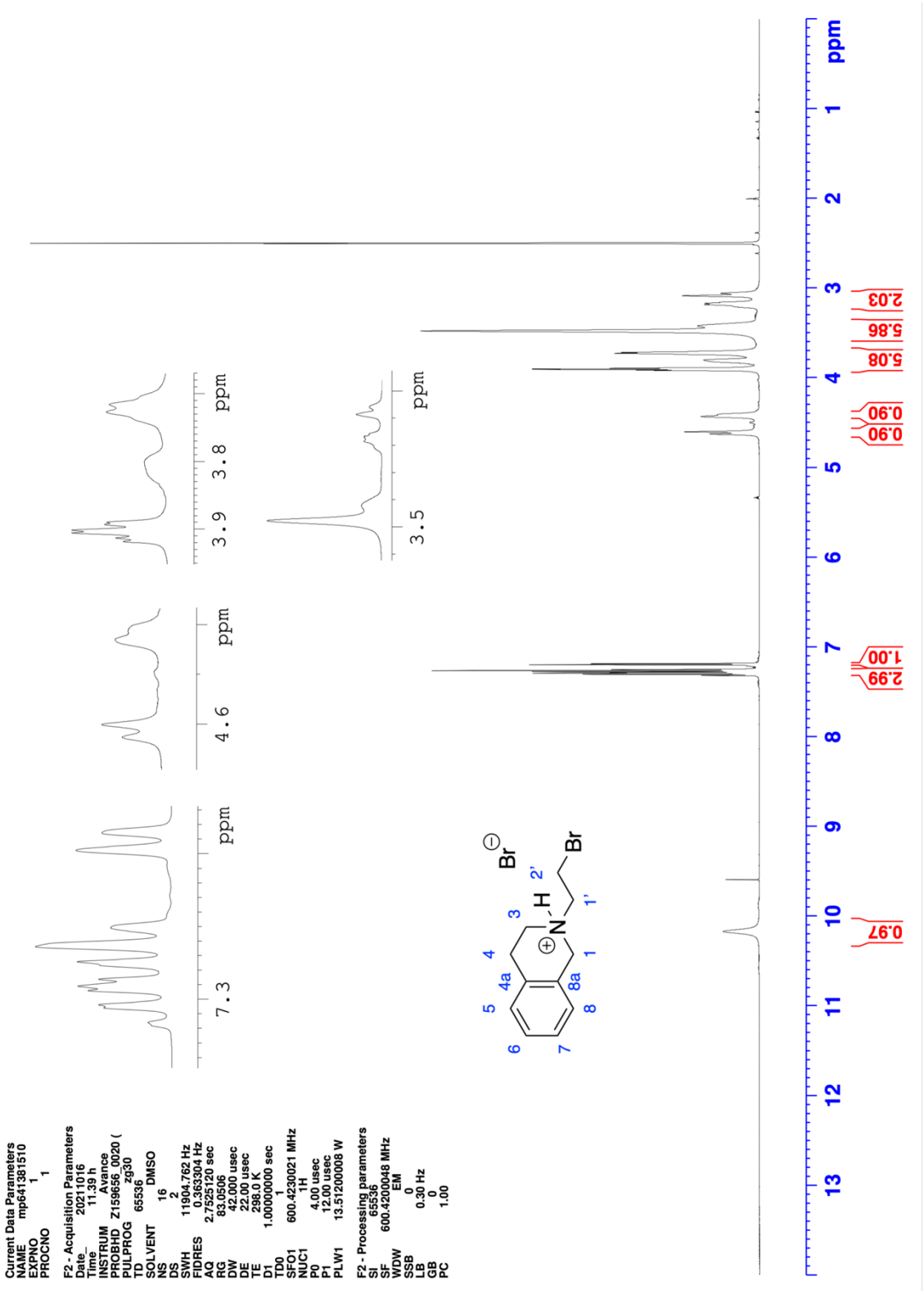
### (3) 2-(2'-Bromoethyl)isoindoline hydrobromide $^{1}H$ NMR



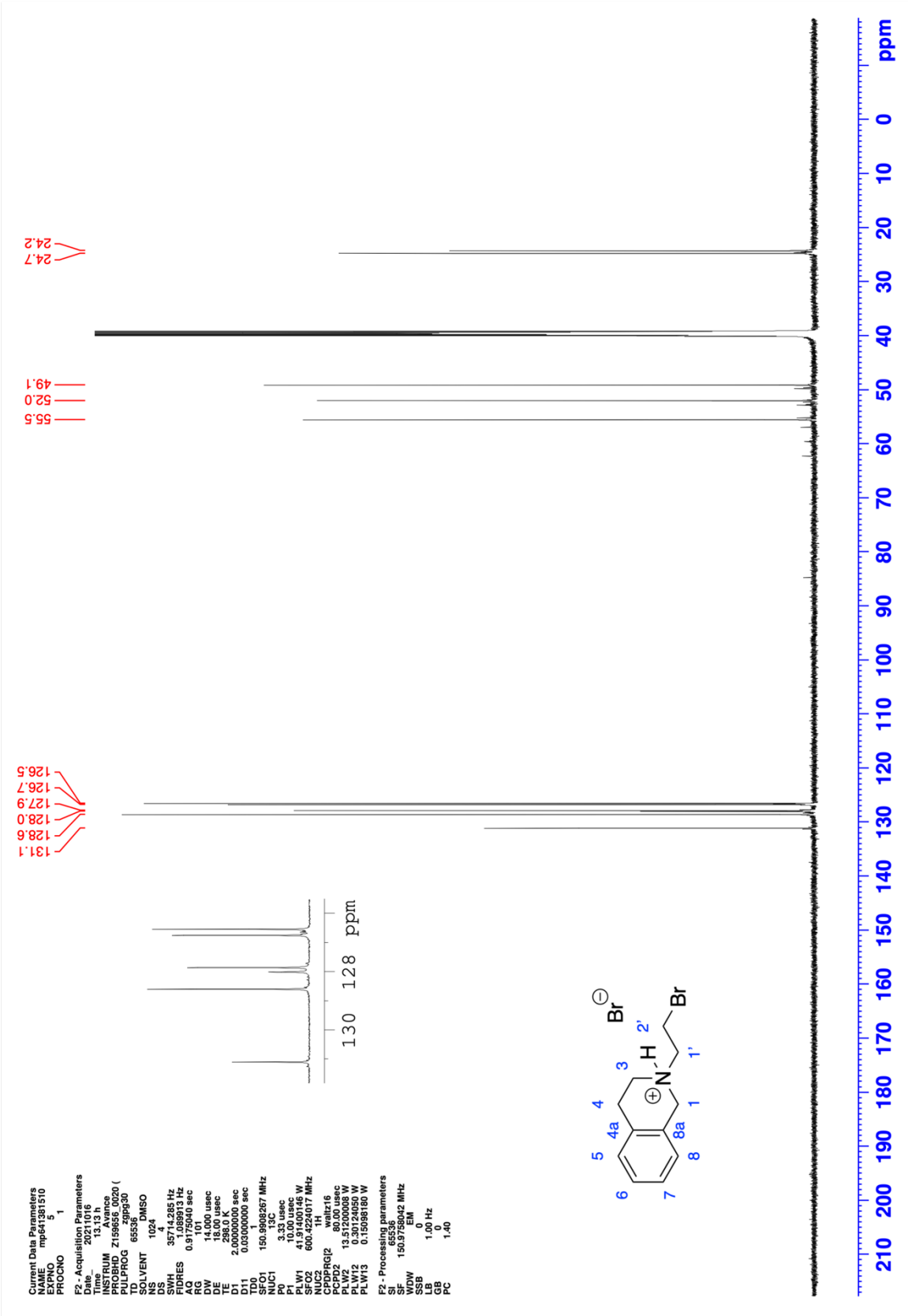
## (3) 2-(2'-Bromoethyl)isoindoline hydrobromide $^{13}$C NMR



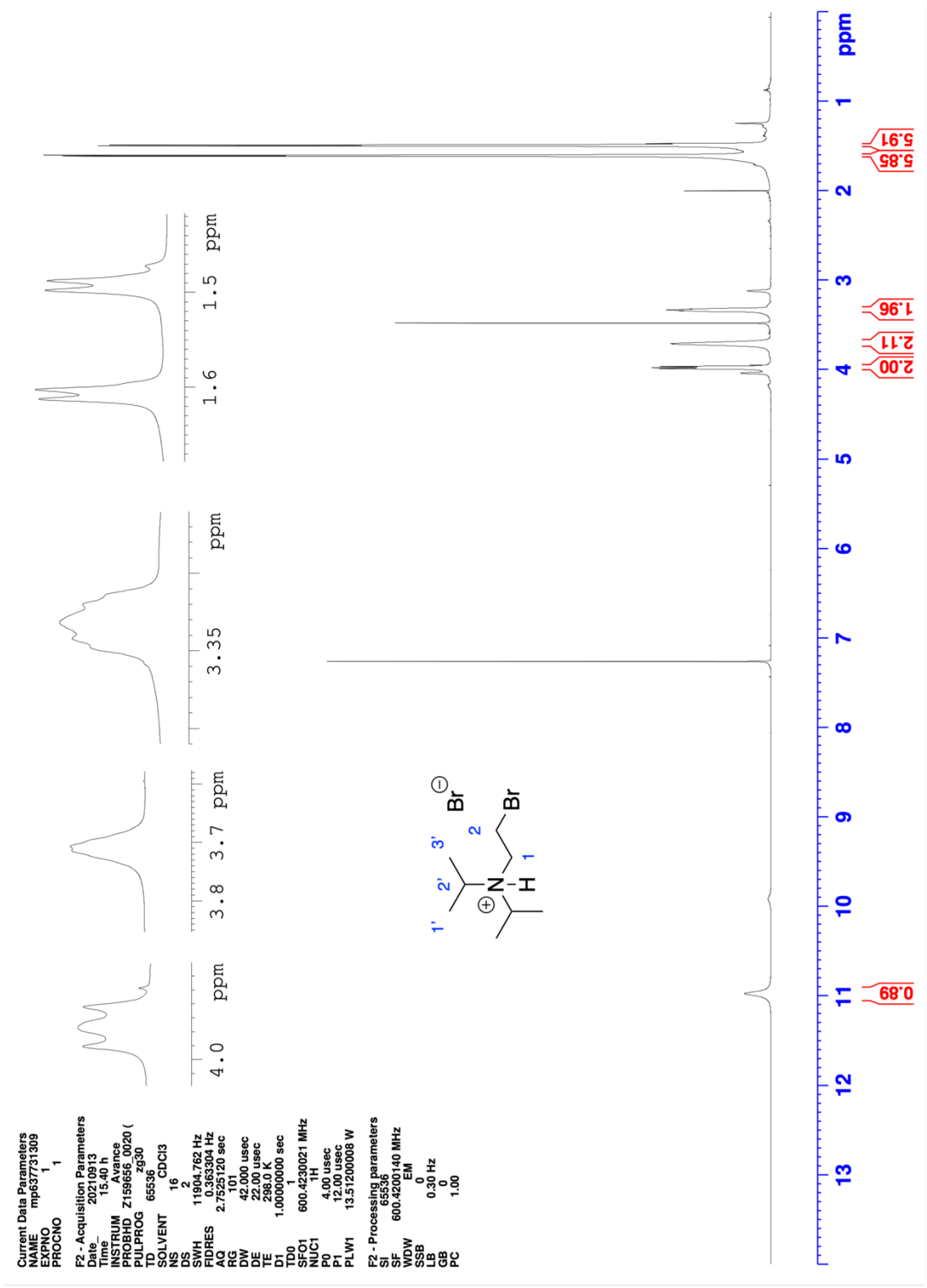
## (4) 2-(2'-Bromoethyl)-1,2,3,4-tetrahydroisoquinoline hydrobromide $^{1}$H NMR



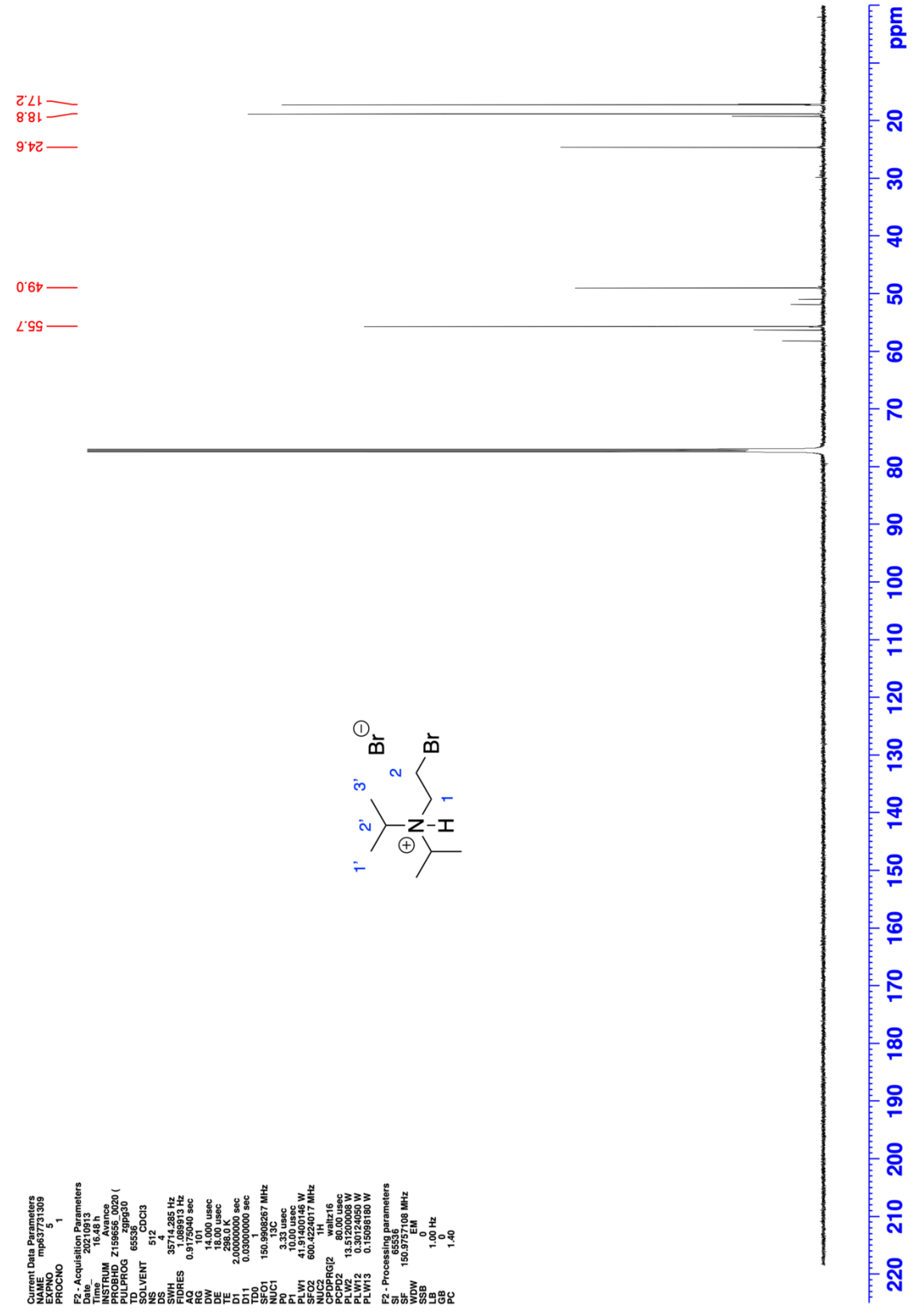
## (4) 2-(2'-Bromoethyl)-1,2,3,4-tetrahydroisoquinoline hydrobromide $^{13}$C NMR



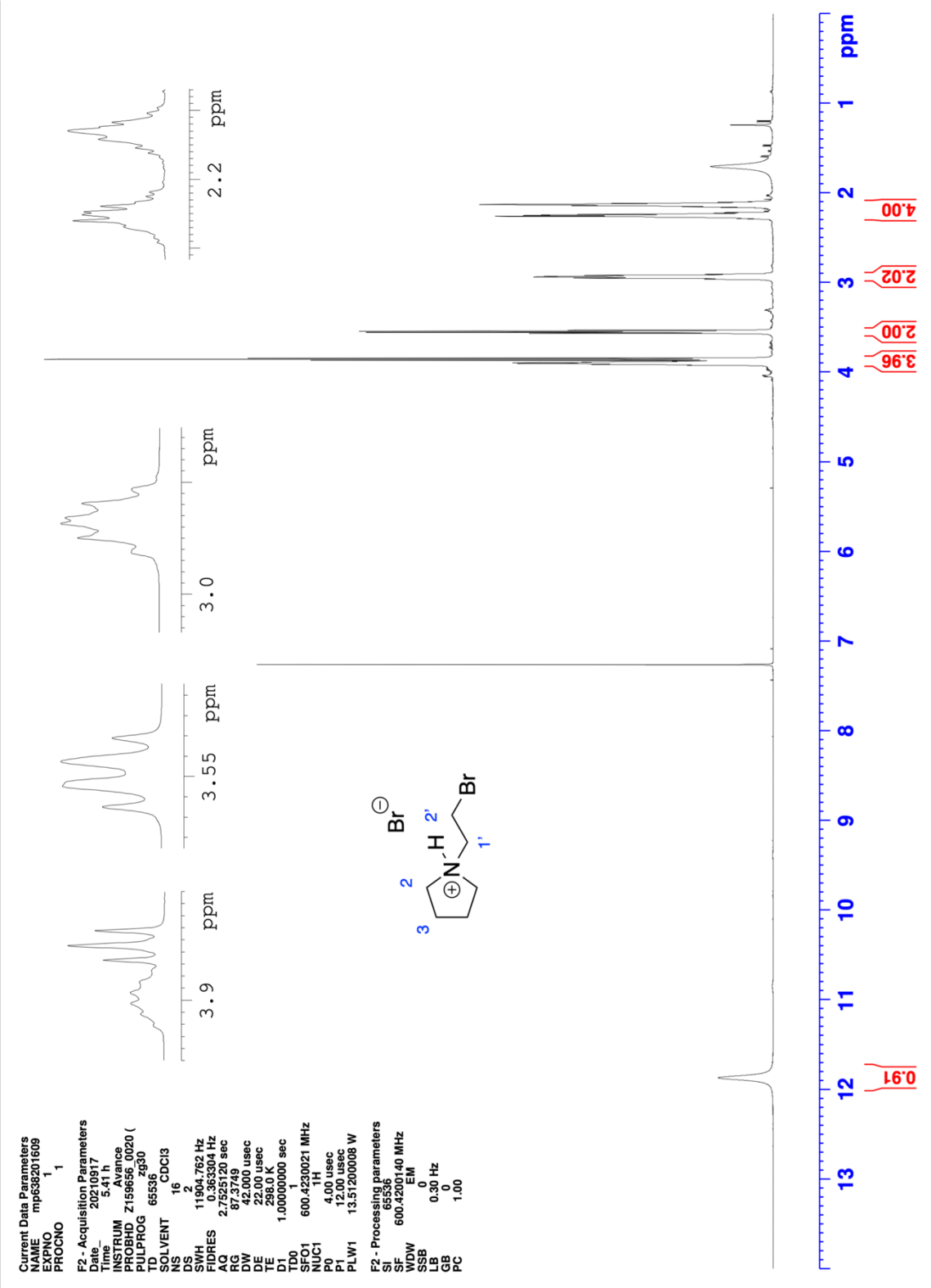
## (5) *N*-(2-Bromoethyl)-*N*-isopropylpropan-2'-amine hydrobromide $^{1}$H NMR



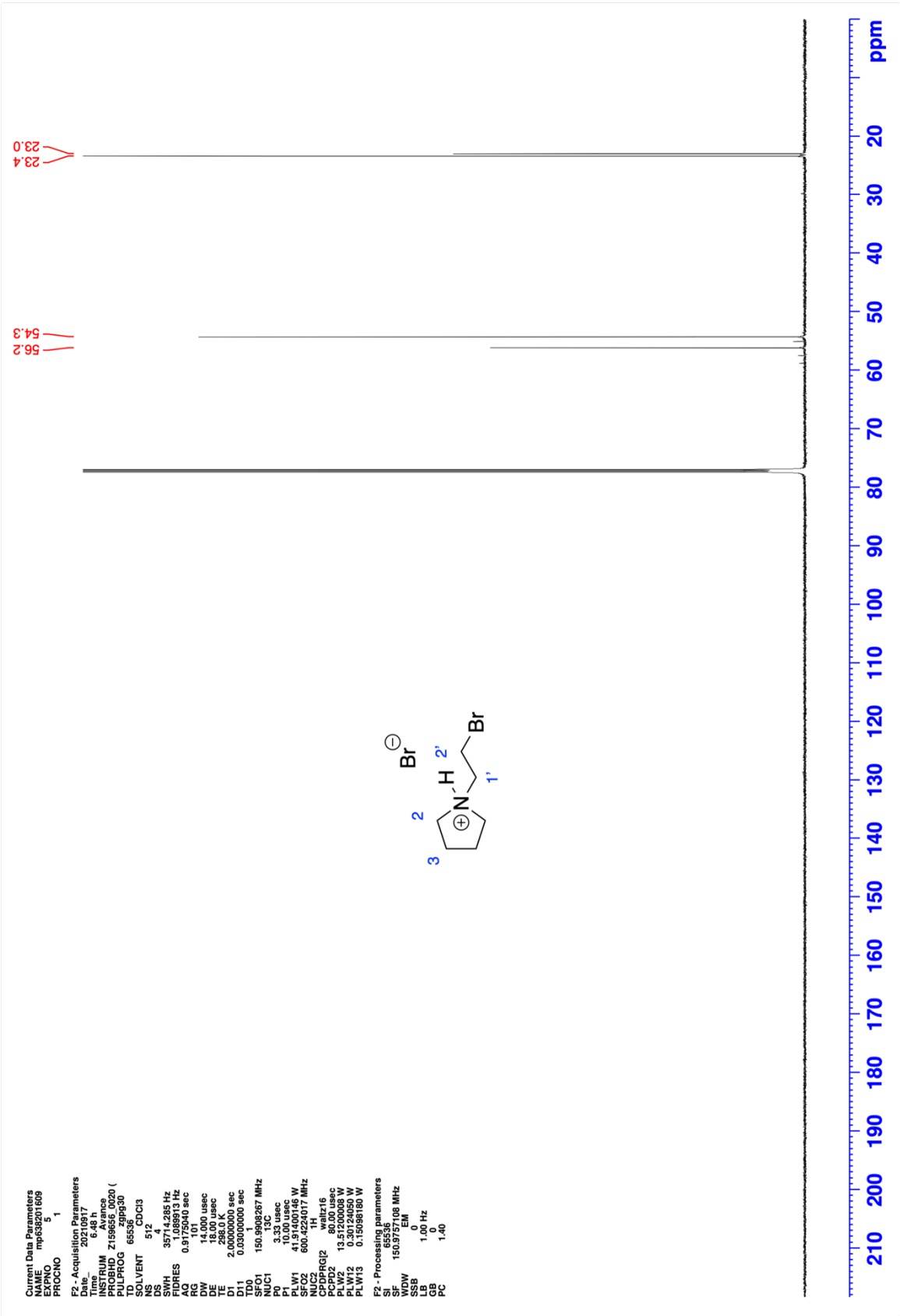
## (5) *N*-(2-Bromoethyl)-*N*-isopropylpropan-2'-amine hydrobromide $^{13}$C NMR



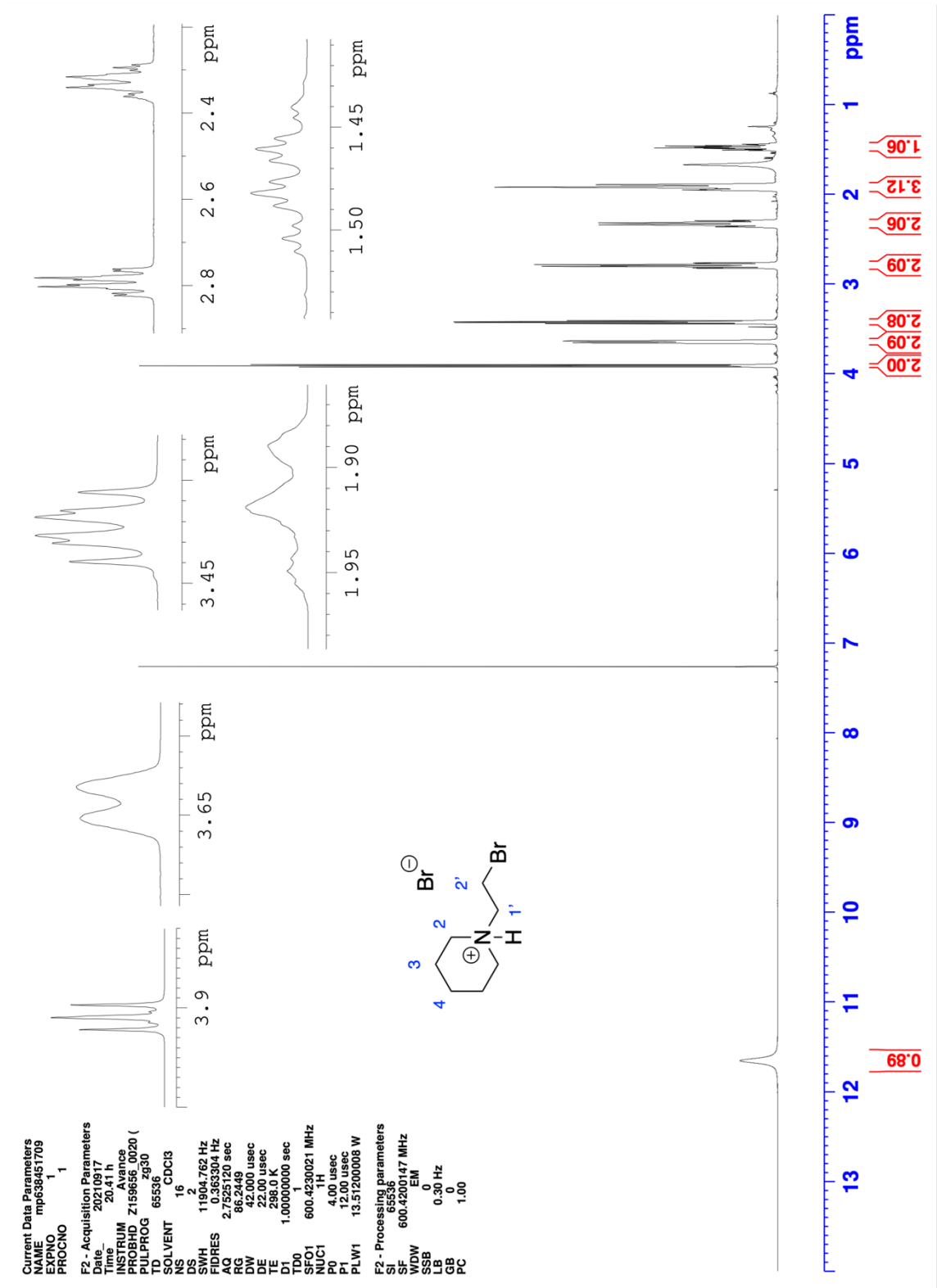
## (6) 1-(2'-Bromoethyl)pyrrolidine hydrobromide $^{1}$H NMR



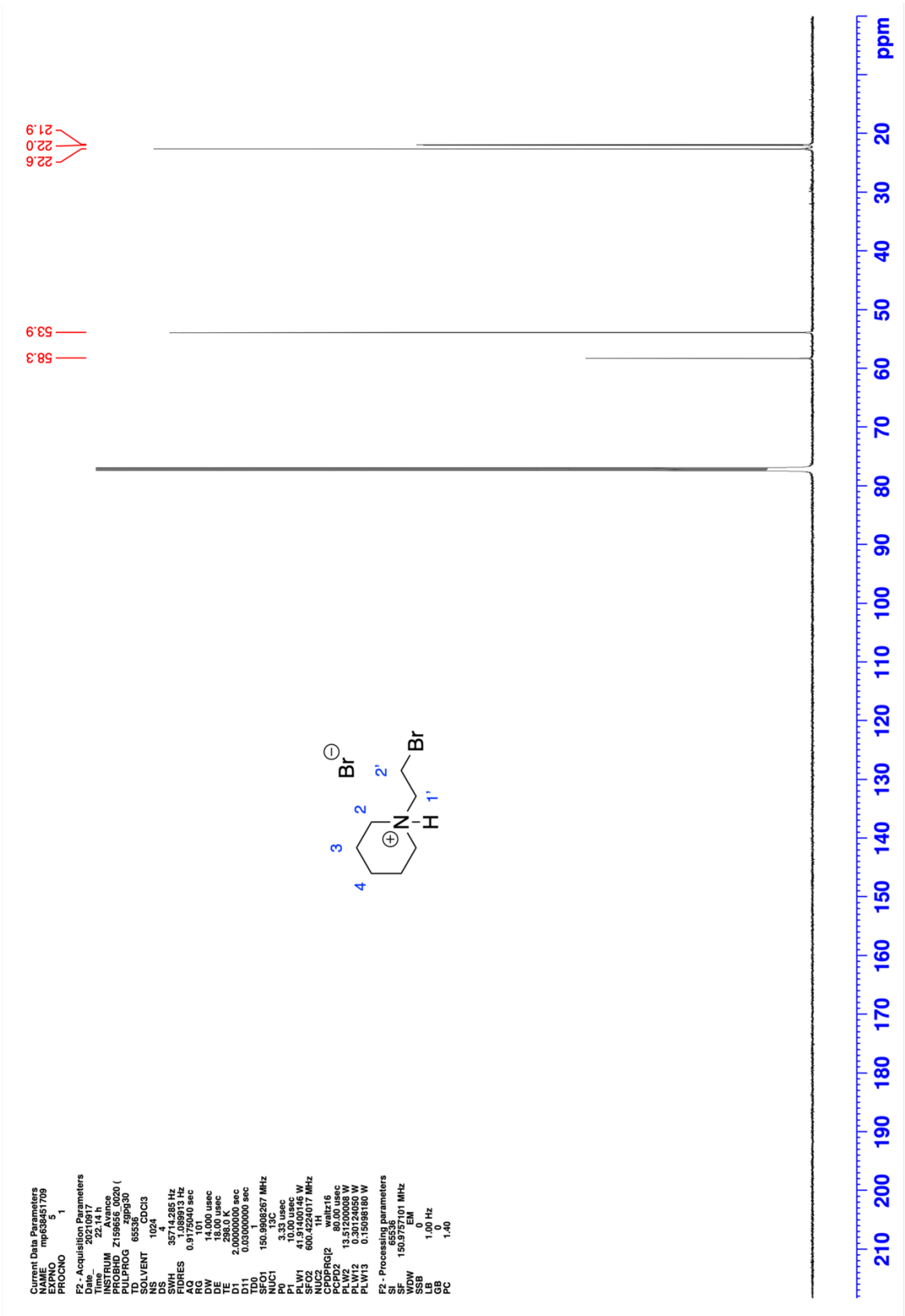
## (6) 1-(2'-Bromoethyl)pyrrolidine hydrobromide $^{13}$C NMR



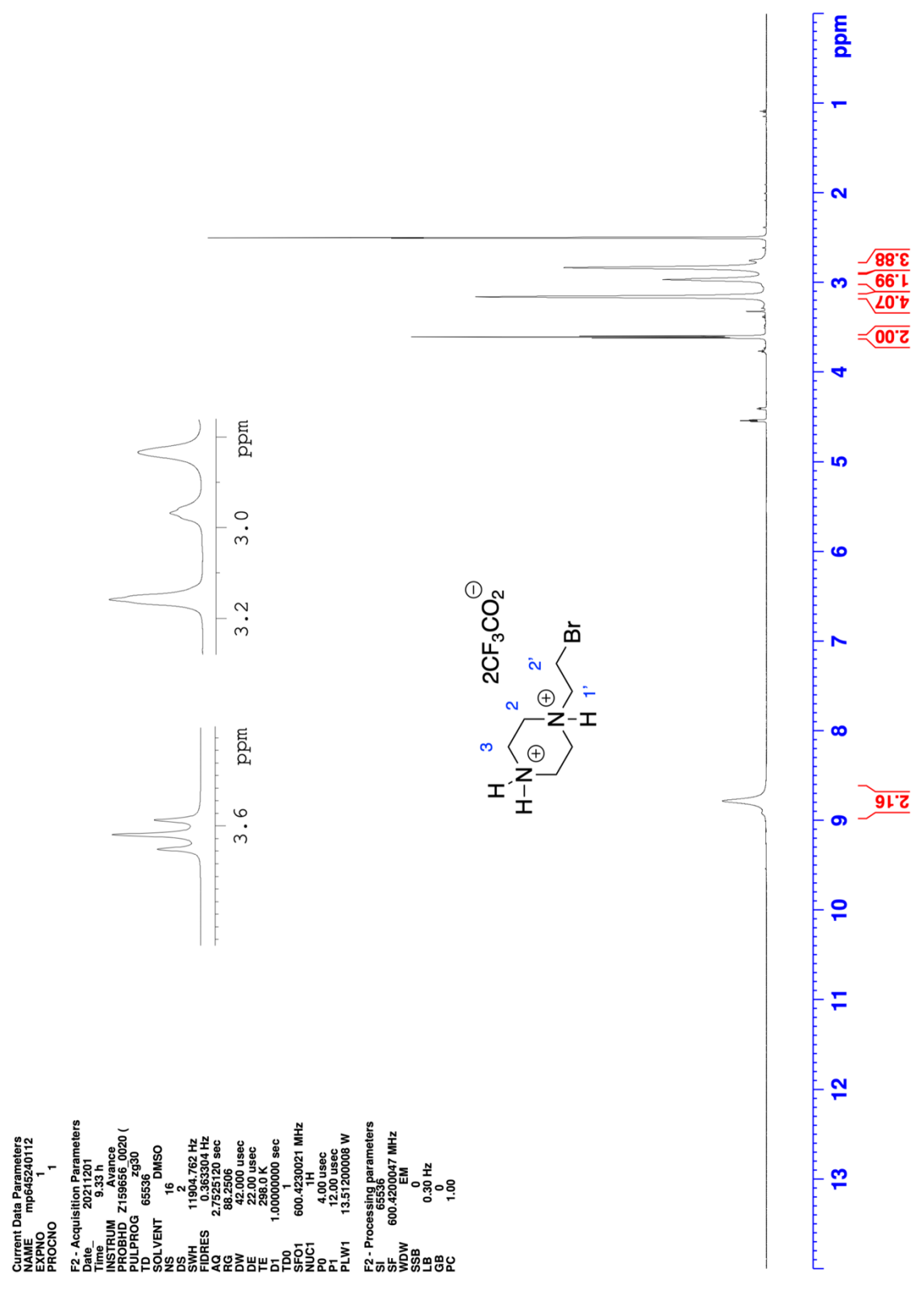
## (7) 1-(2'-Bromoethyl)piperidine hydrobromide $^{1}$H NMR



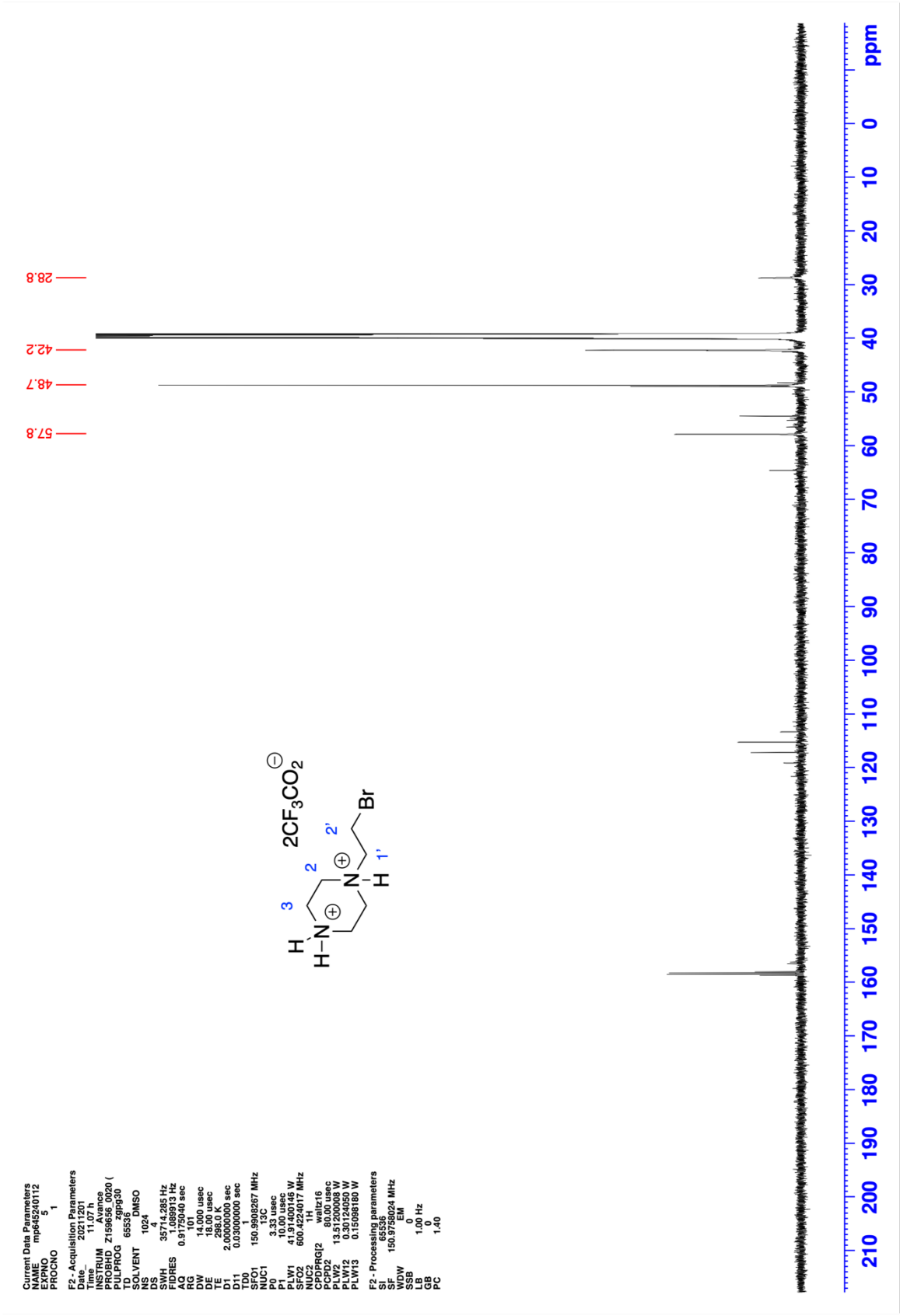
## (7) 1-(2'-Bromoethyl)piperidine hydrobromide $^{13}$C NMR



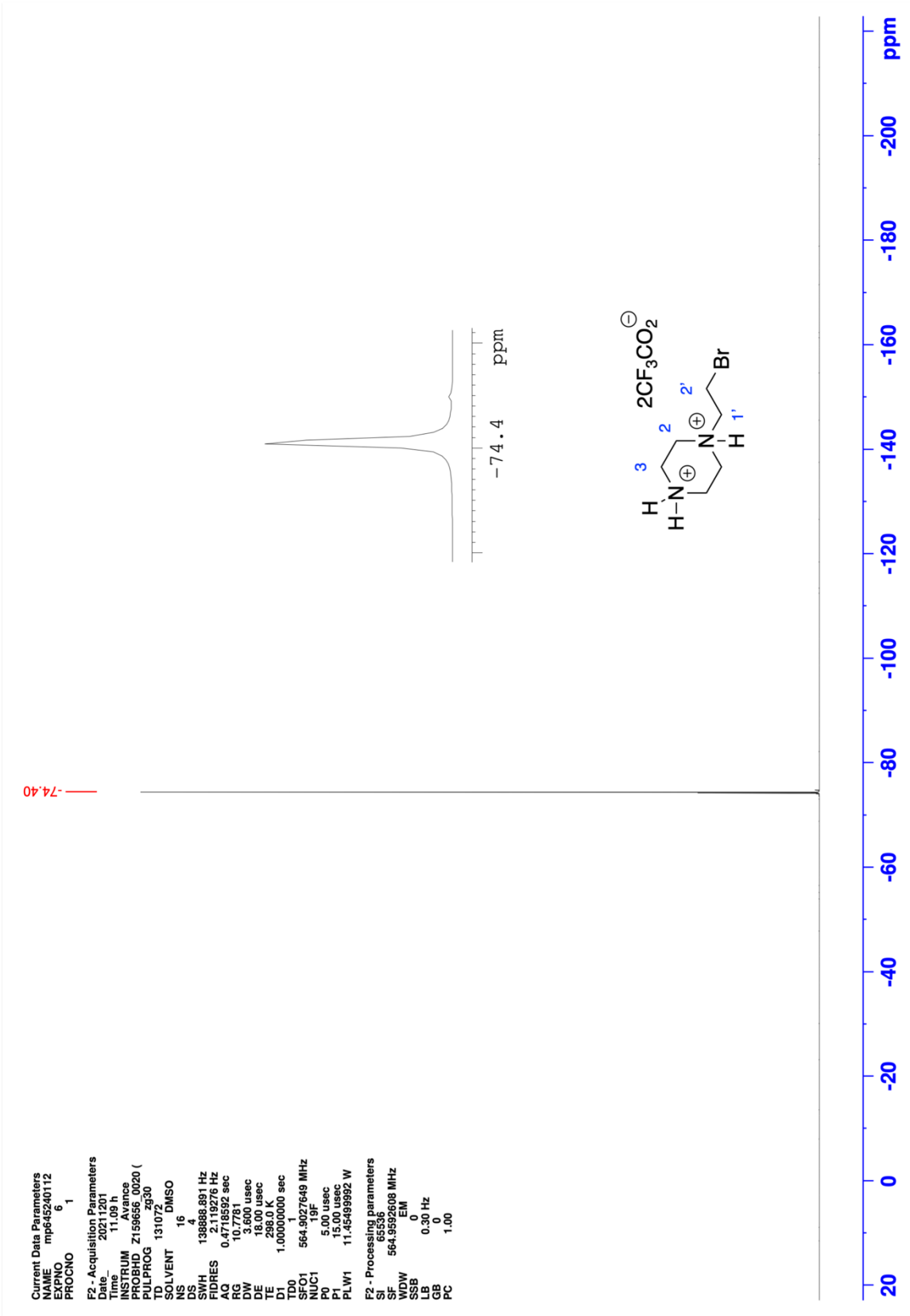
## (12) 1-(2'-Bromoethyl)piperazine ditrifluoroacetic acid $^{1}$H NMR



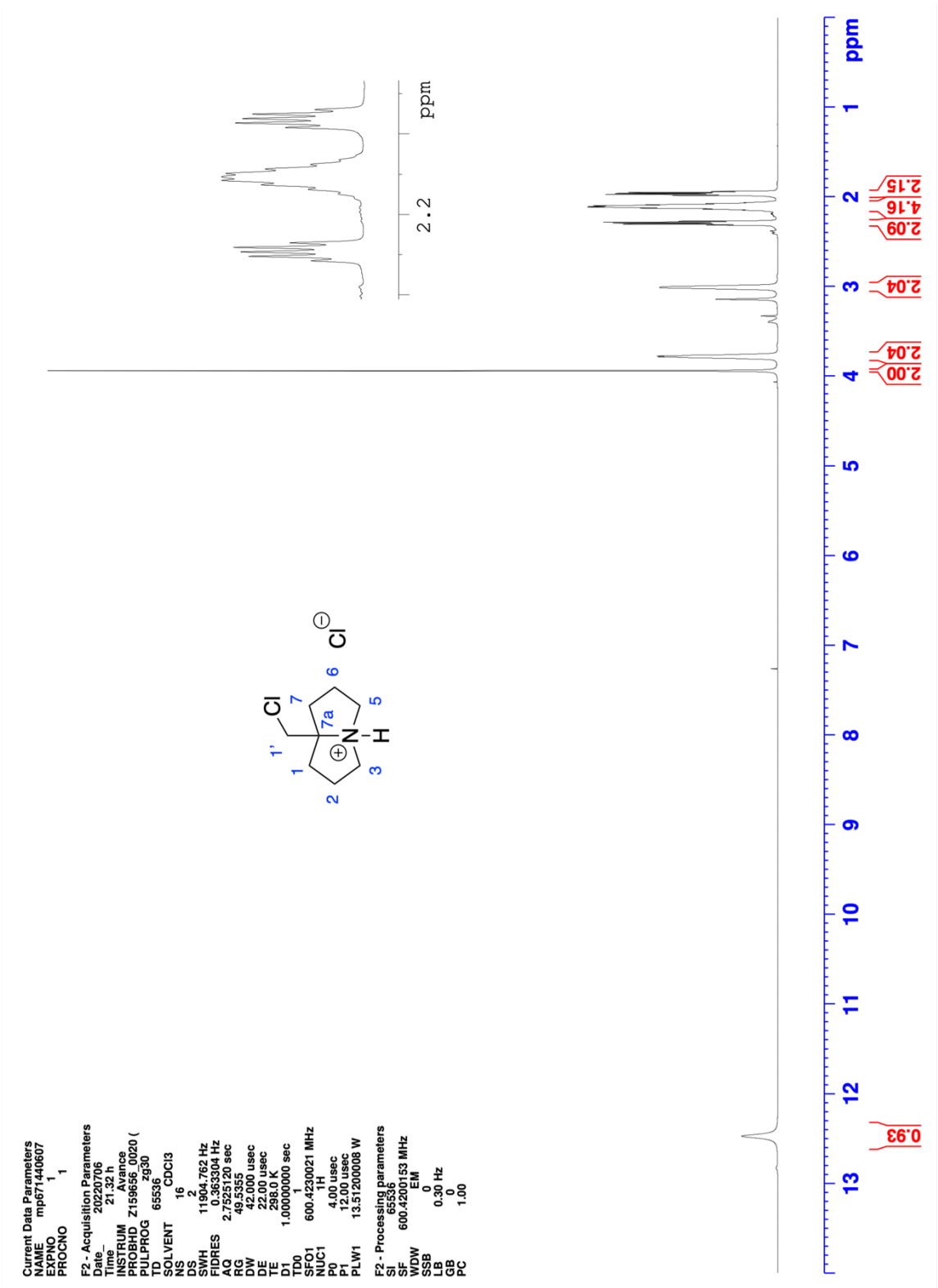
## (12) 1-(2'-Bromoethyl)piperazine ditrifluoroacetic acid $^{13}C$ NMR



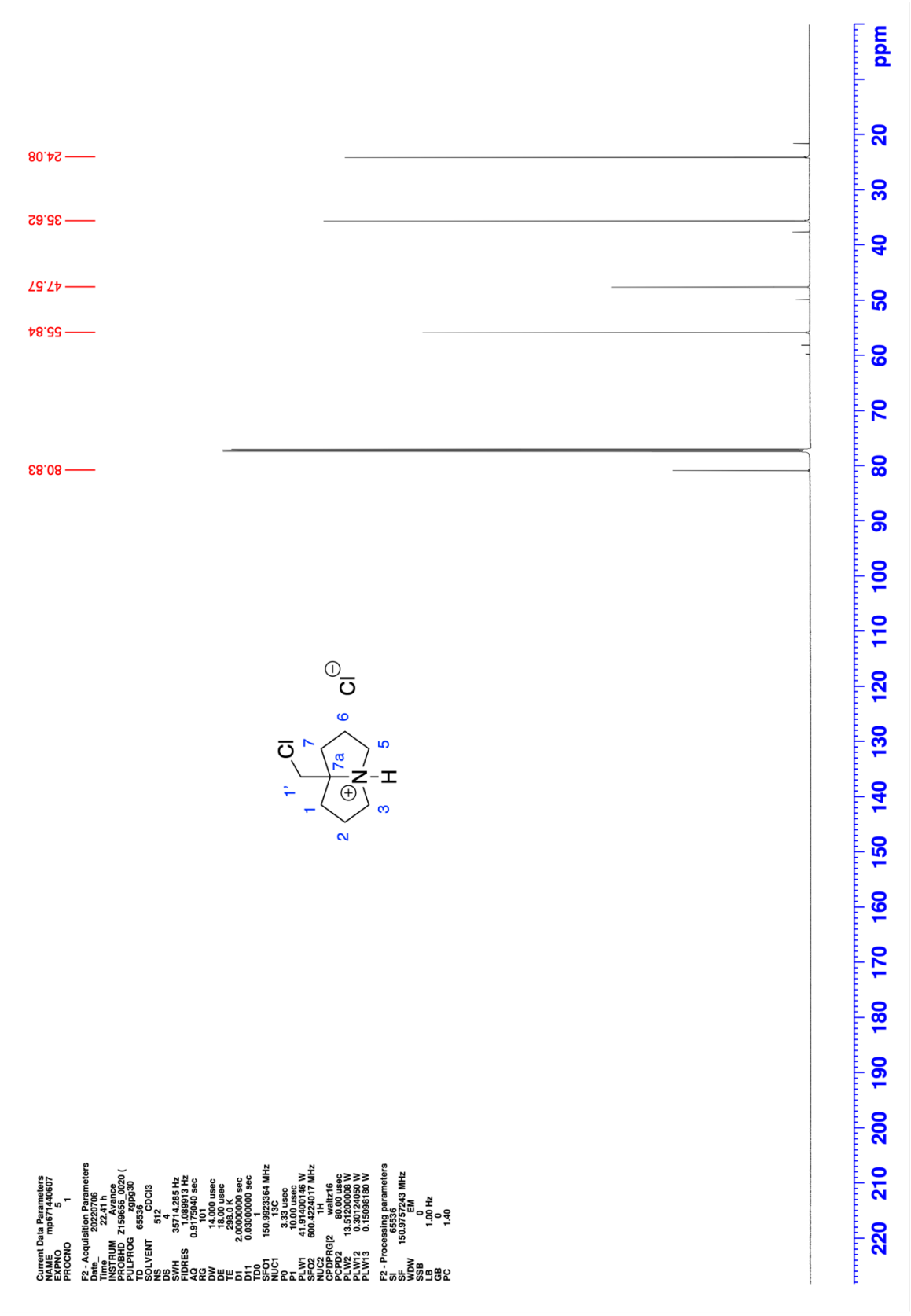
## (12) 1-(2'-Bromoethyl)piperazine ditrifluoroacetic acid $^{19}F$ NMR



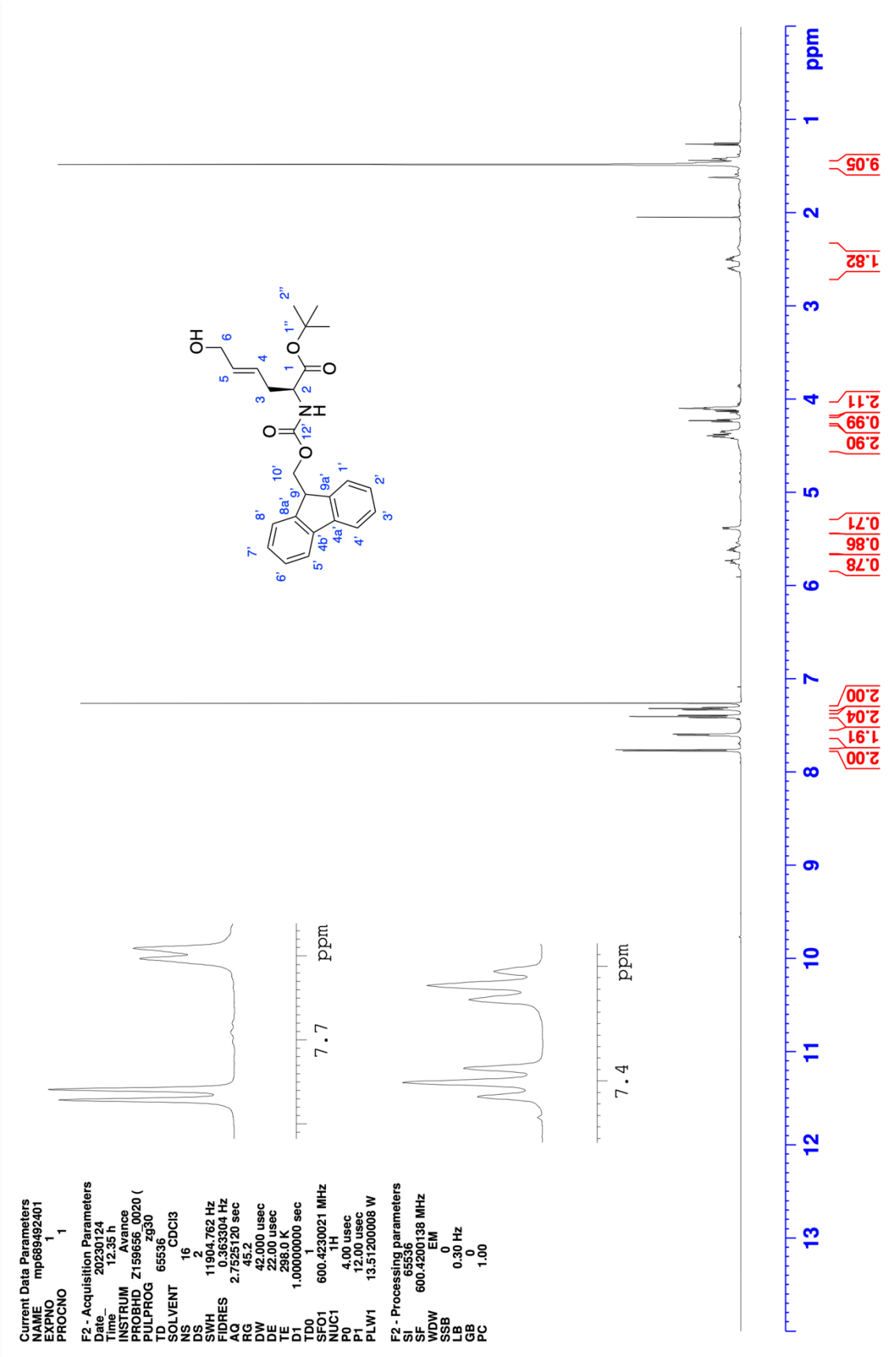
## (21) 7a-(Chloromethyl)hexahydro-1*H*-pyrrolizine hydrochloride $^{1}$H NMR



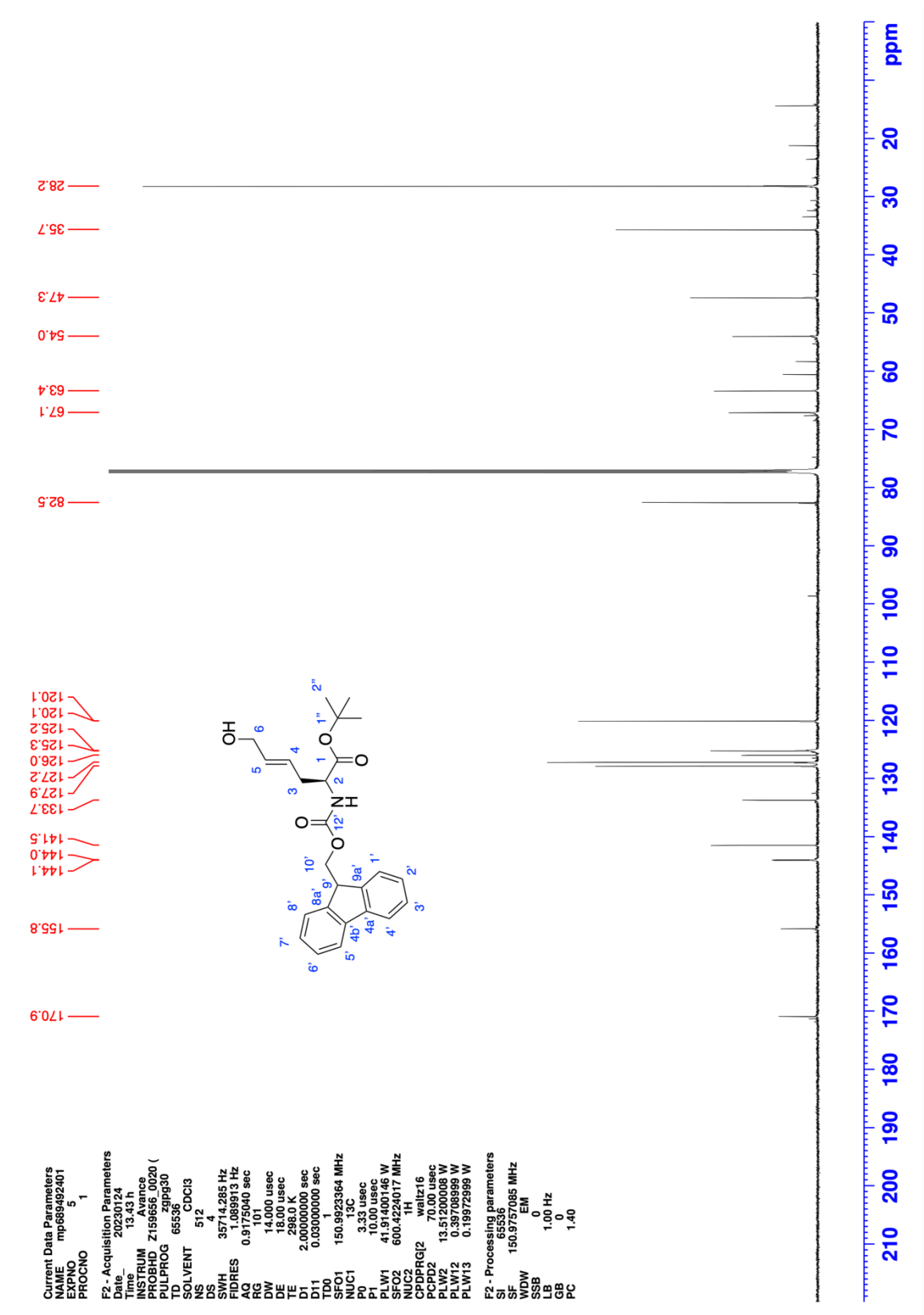
## (21) 7a-(Chloromethyl)hexahydro-1*H*-pyrrolizine hydrochloride $^{13}$C NMR



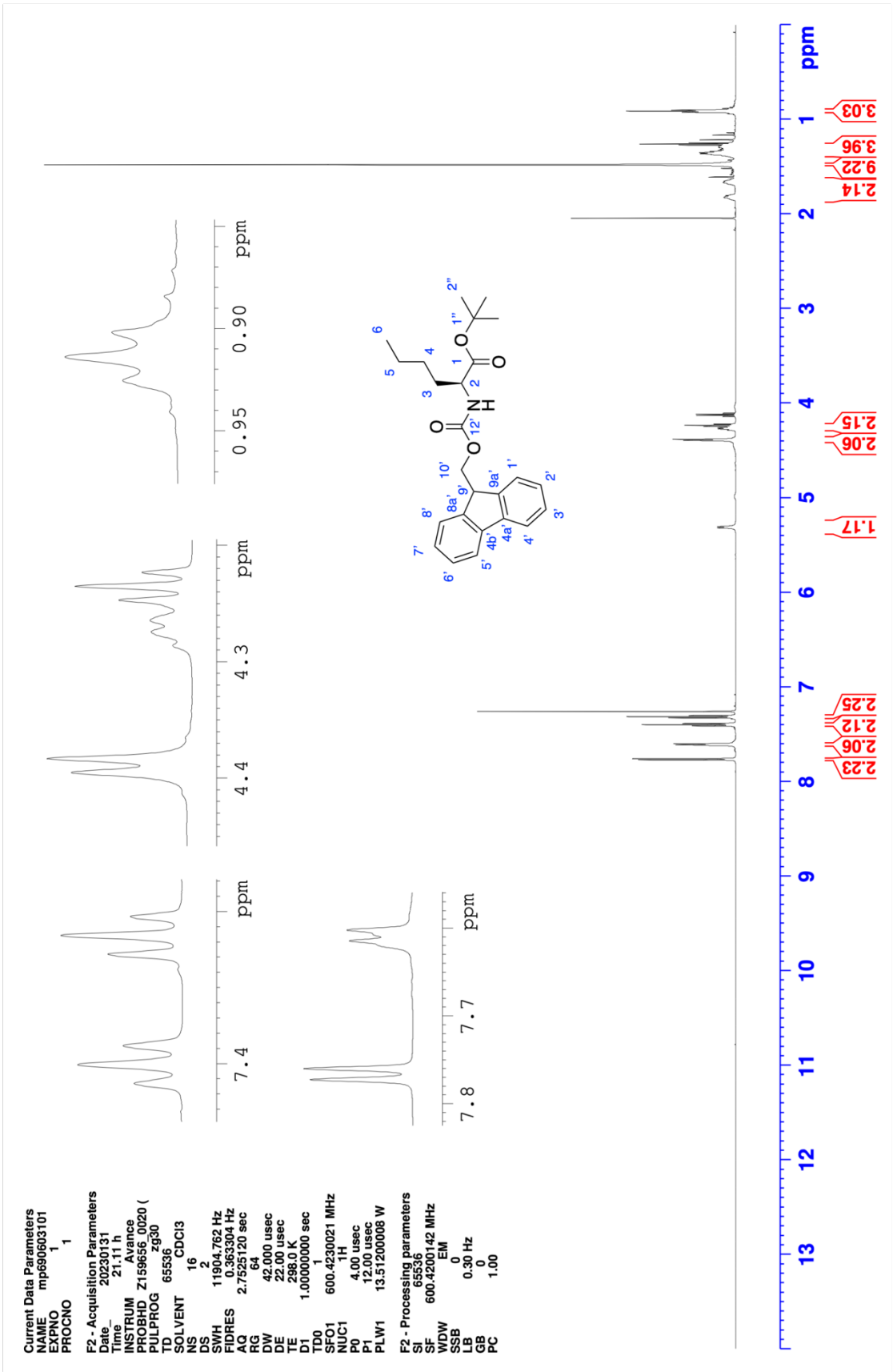
**(41) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-hydroxyhex-4-enoate $^{1}$H NMR**



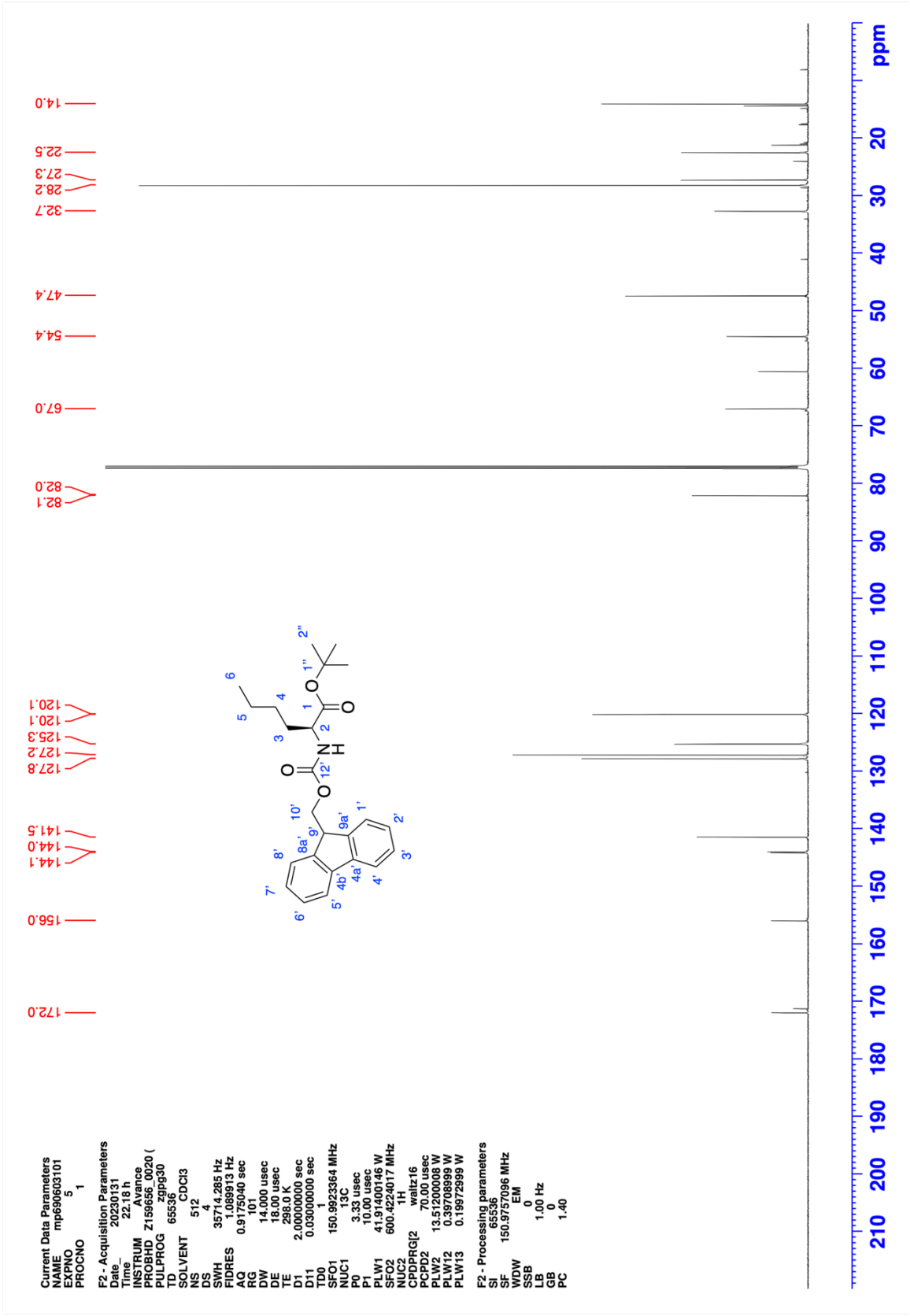
**(41) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-hydroxyhex-4-enoate $^{13}$C NMR**



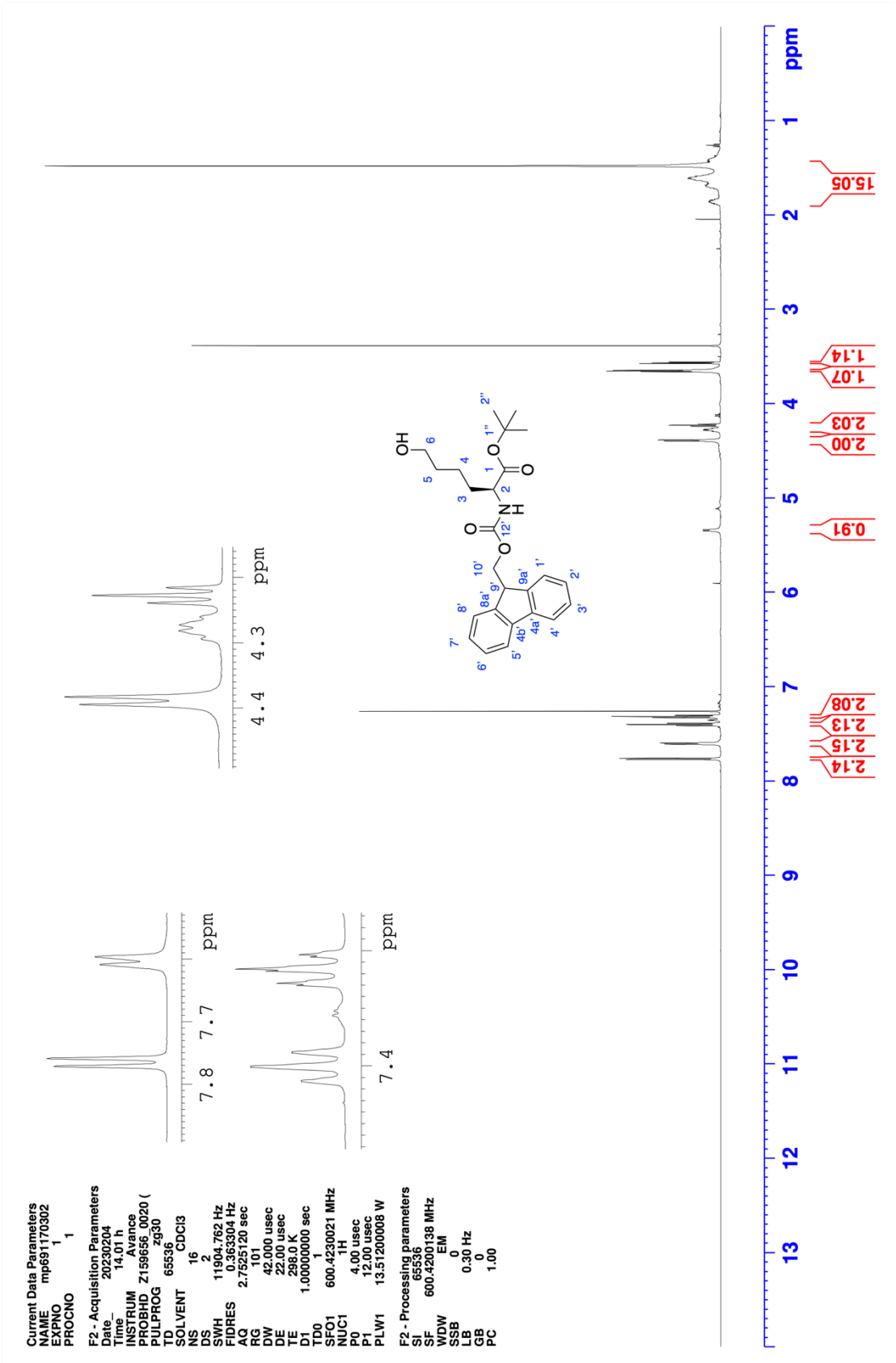
**(42) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)hexanoate**
**$^{1}$H NMR**



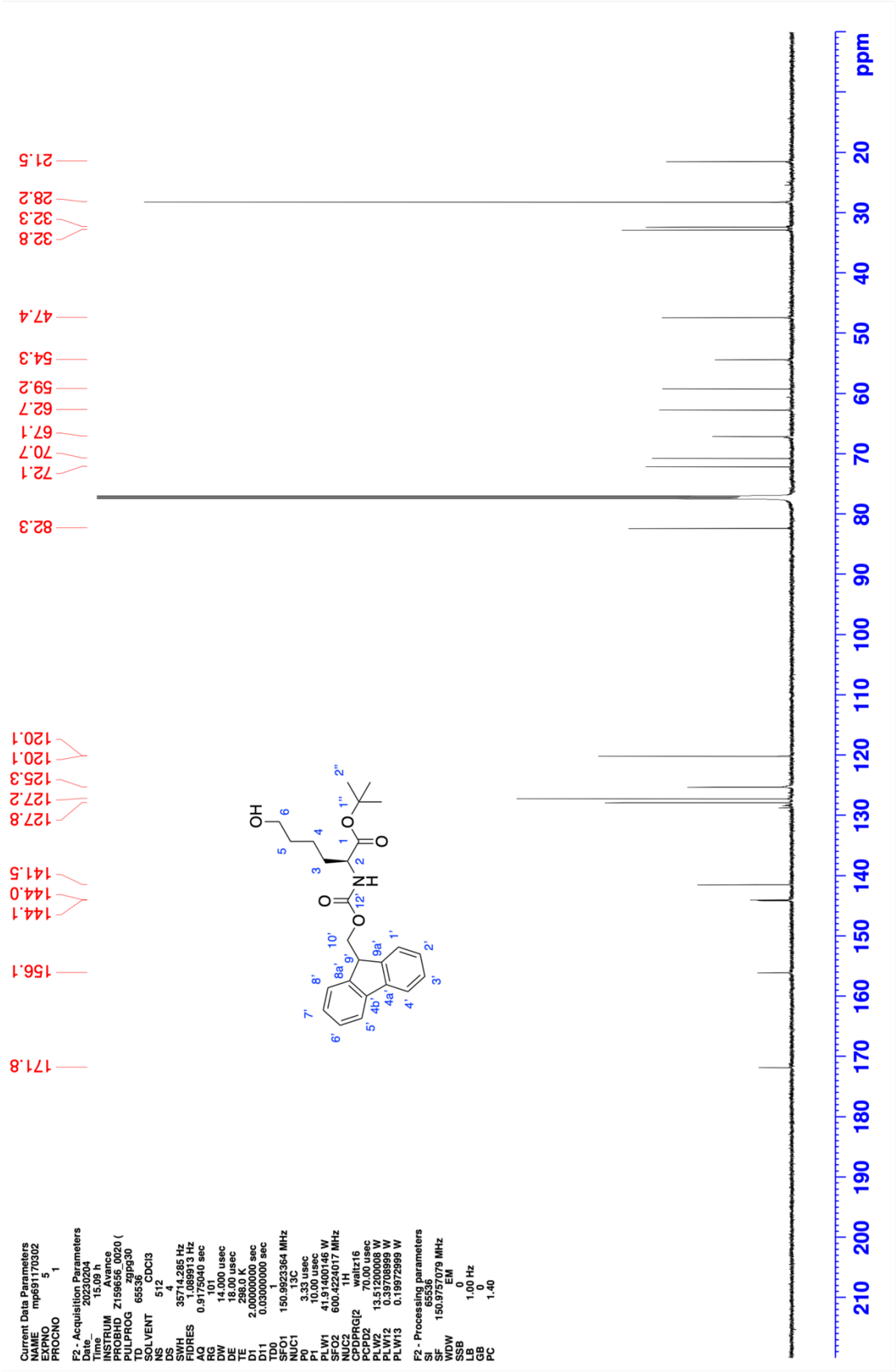
## (42) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)hexanoate $^{13}$C NMR



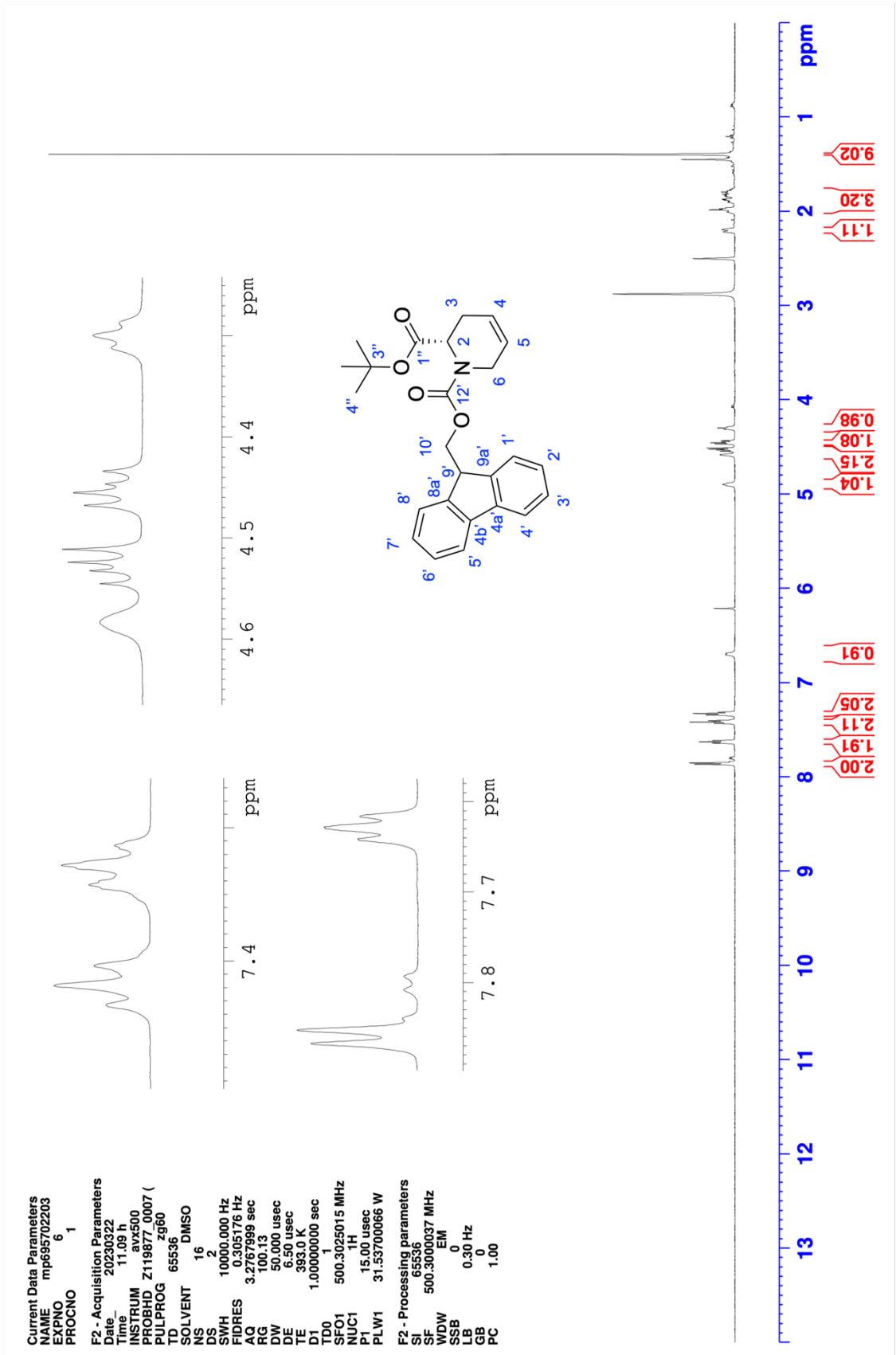
**(43) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-hydroxyhexanoate $^{1}$H NMR**



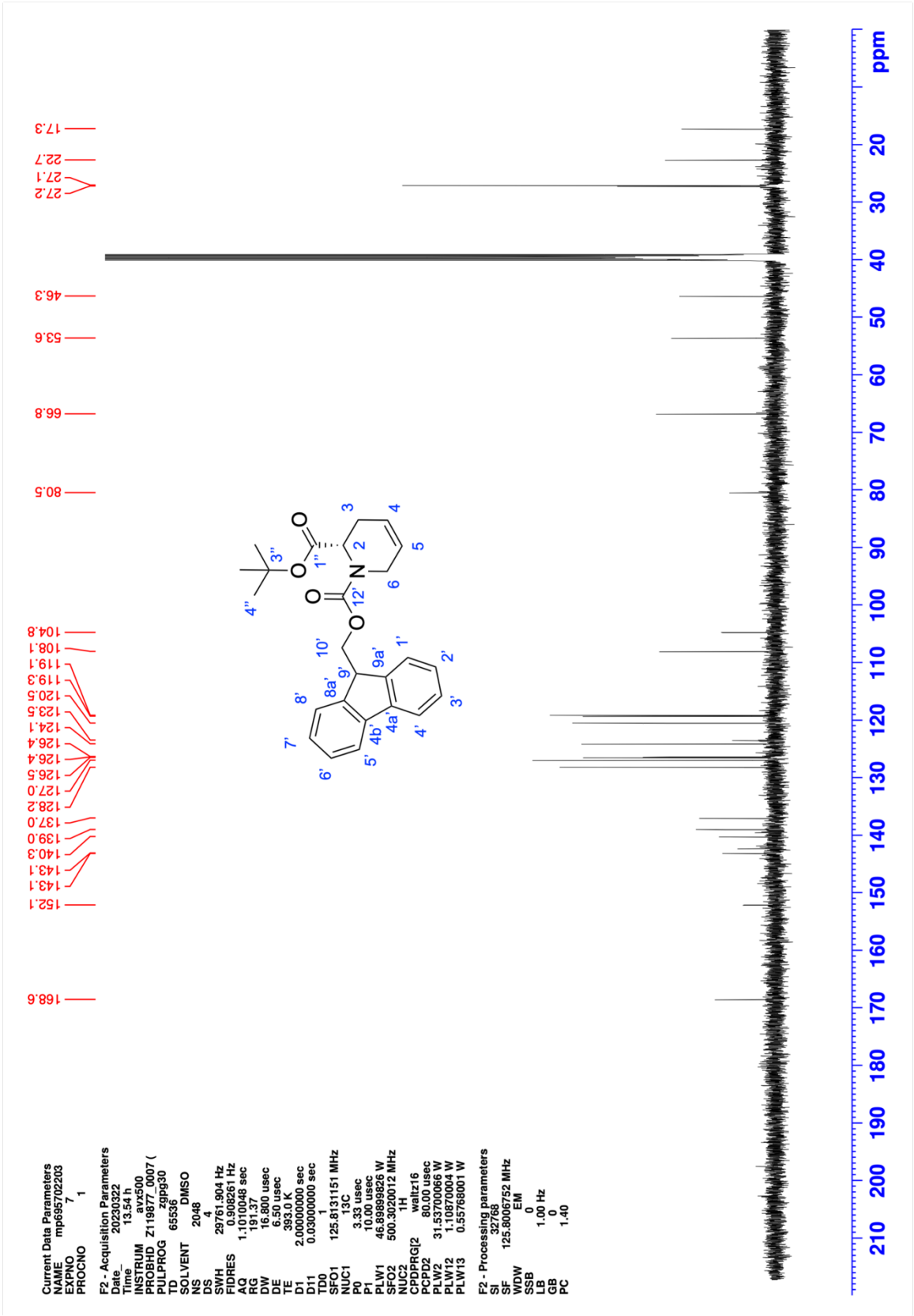
**(43) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-hydroxyhexanoate $^{13}$C NMR**



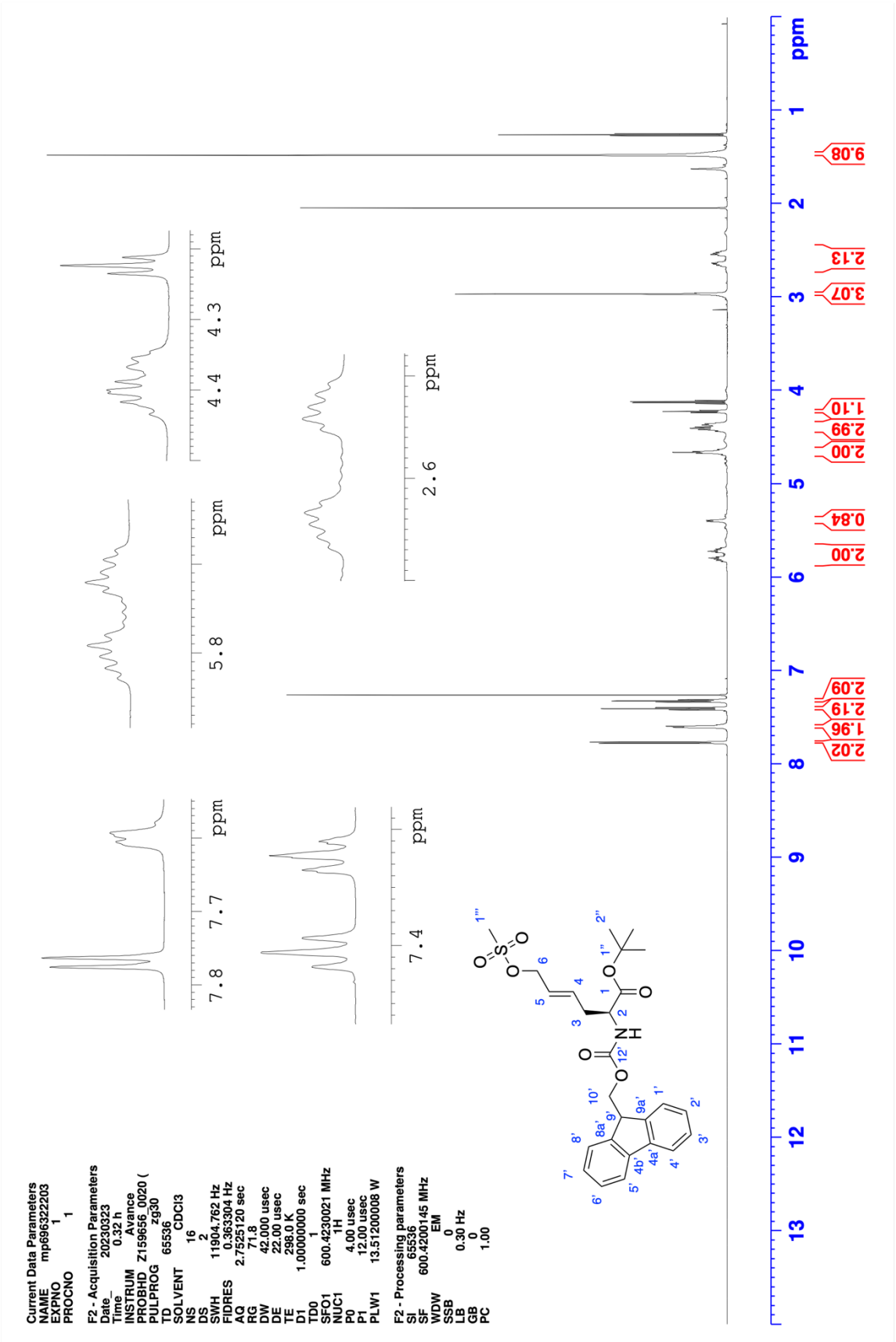
**(44) 1-((9'*H*-Fluoren-9'-yl)methyl) 2-(*tert*-butyl) (*S*)-3,6-dihydropyridine-1,2(2*H*)-dicarboxylate $^{1}$H NMR**



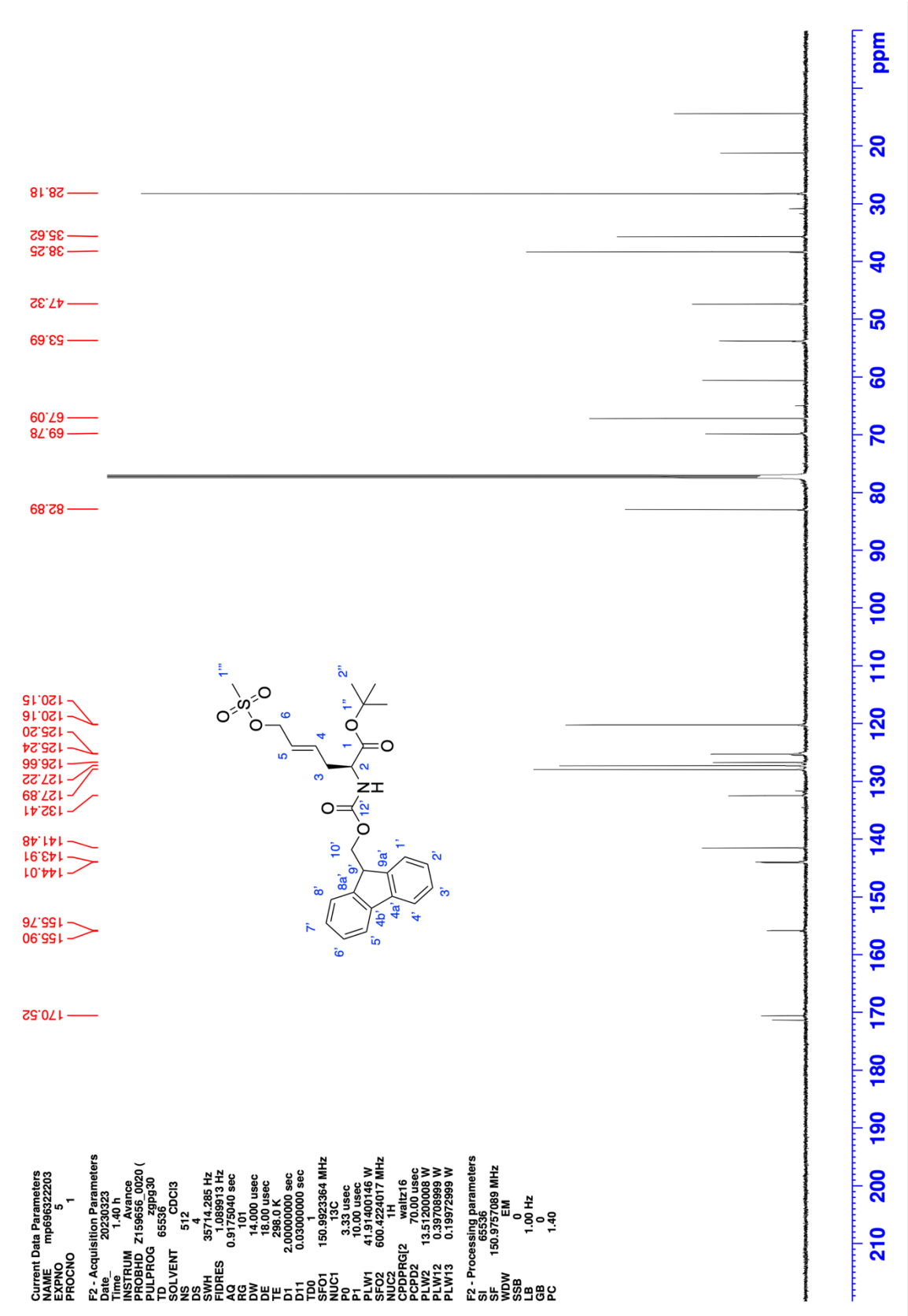
**(44) 1-((9'*H*-Fluoren-9'-yl)methyl) 2-(*tert*-butyl) (*S*)-3,6-dihydropyridine-1,2(2*H*)-dicarboxylate $^{13}$C NMR**



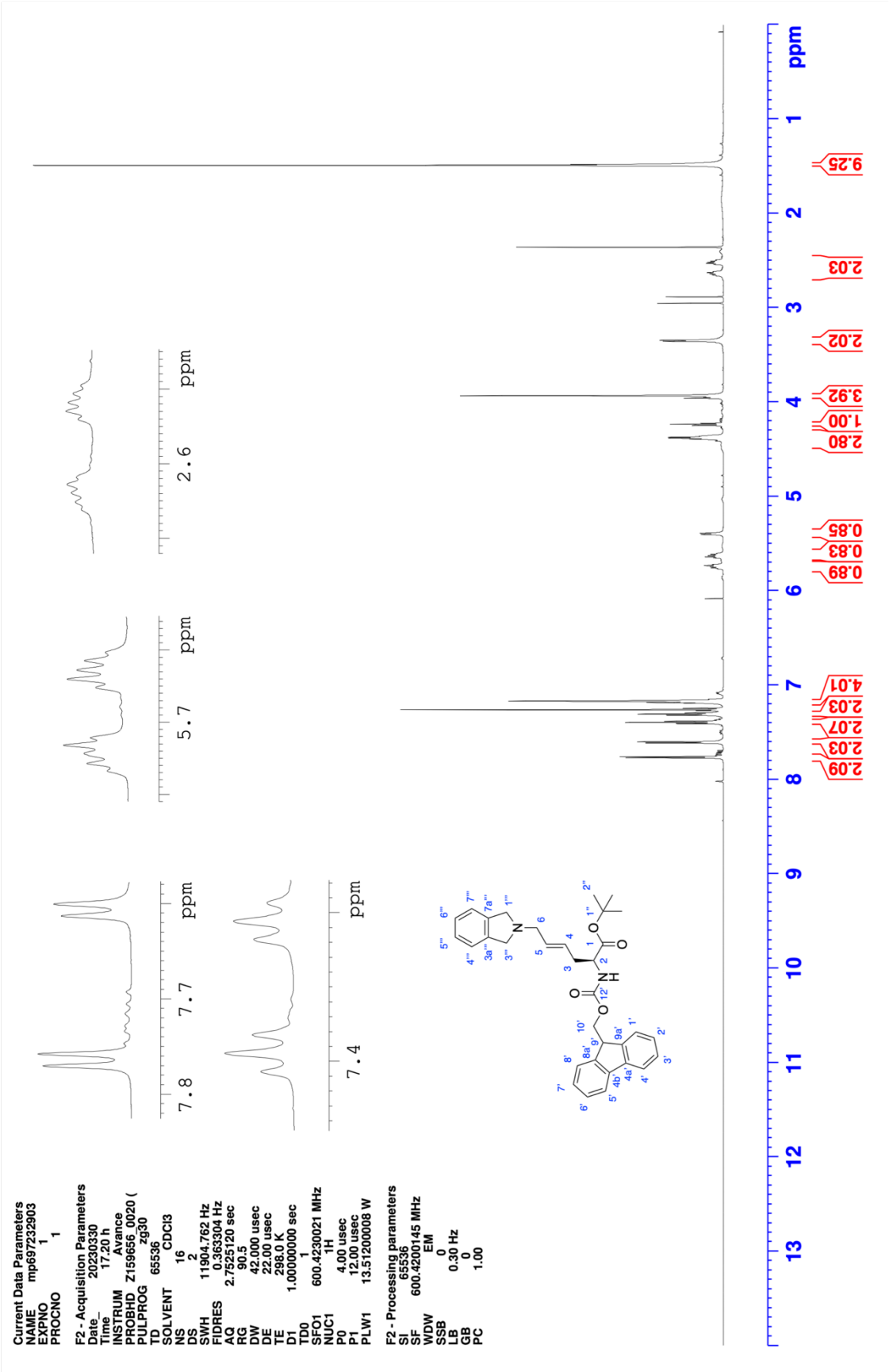
**(45) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-((methylsulfonyl)oxy)hex-4-enoate $^{1}$H NMR**



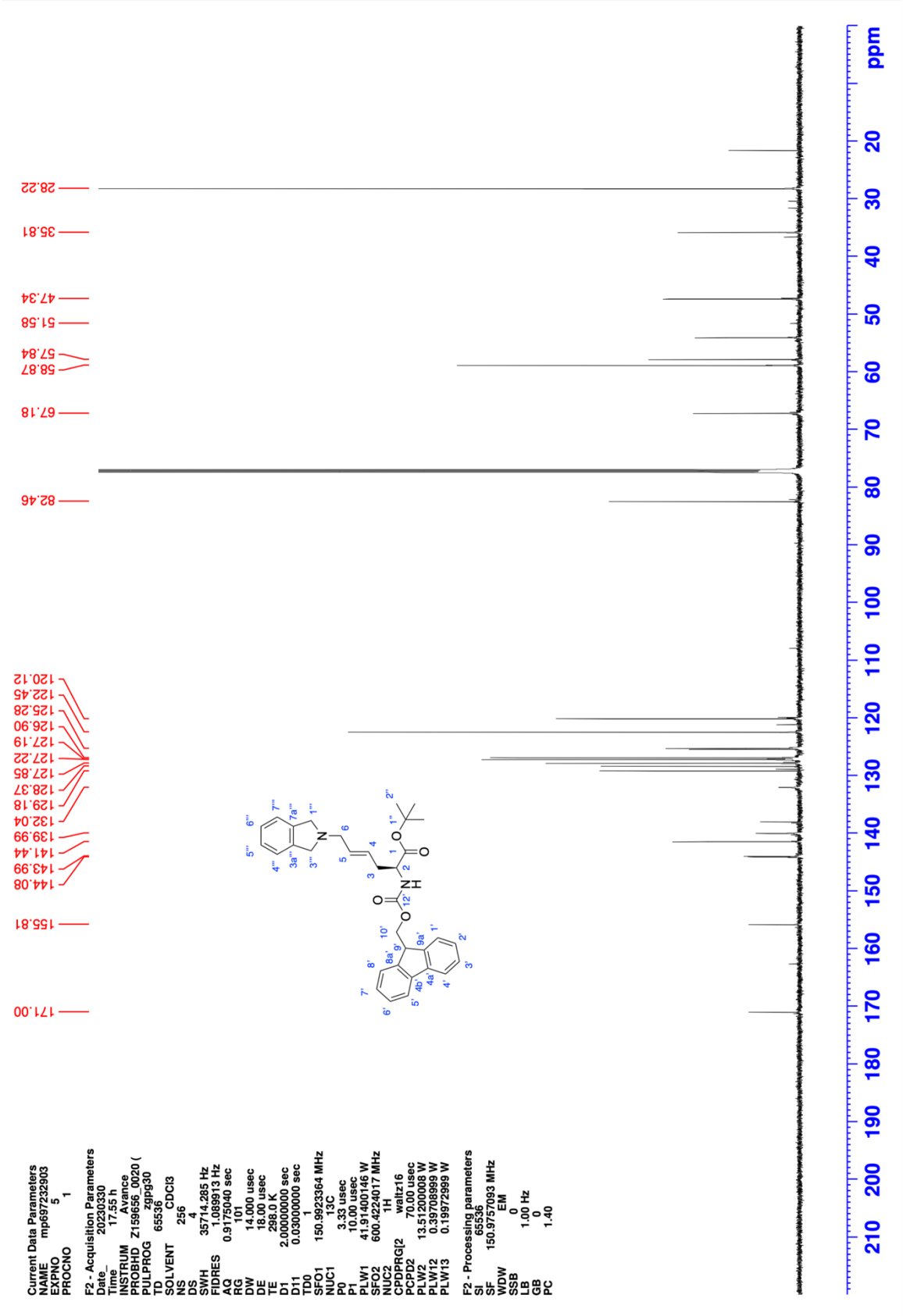
**(45) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-((methylsulfonyl)oxy)hex-4-enoate $^{13}$C NMR**



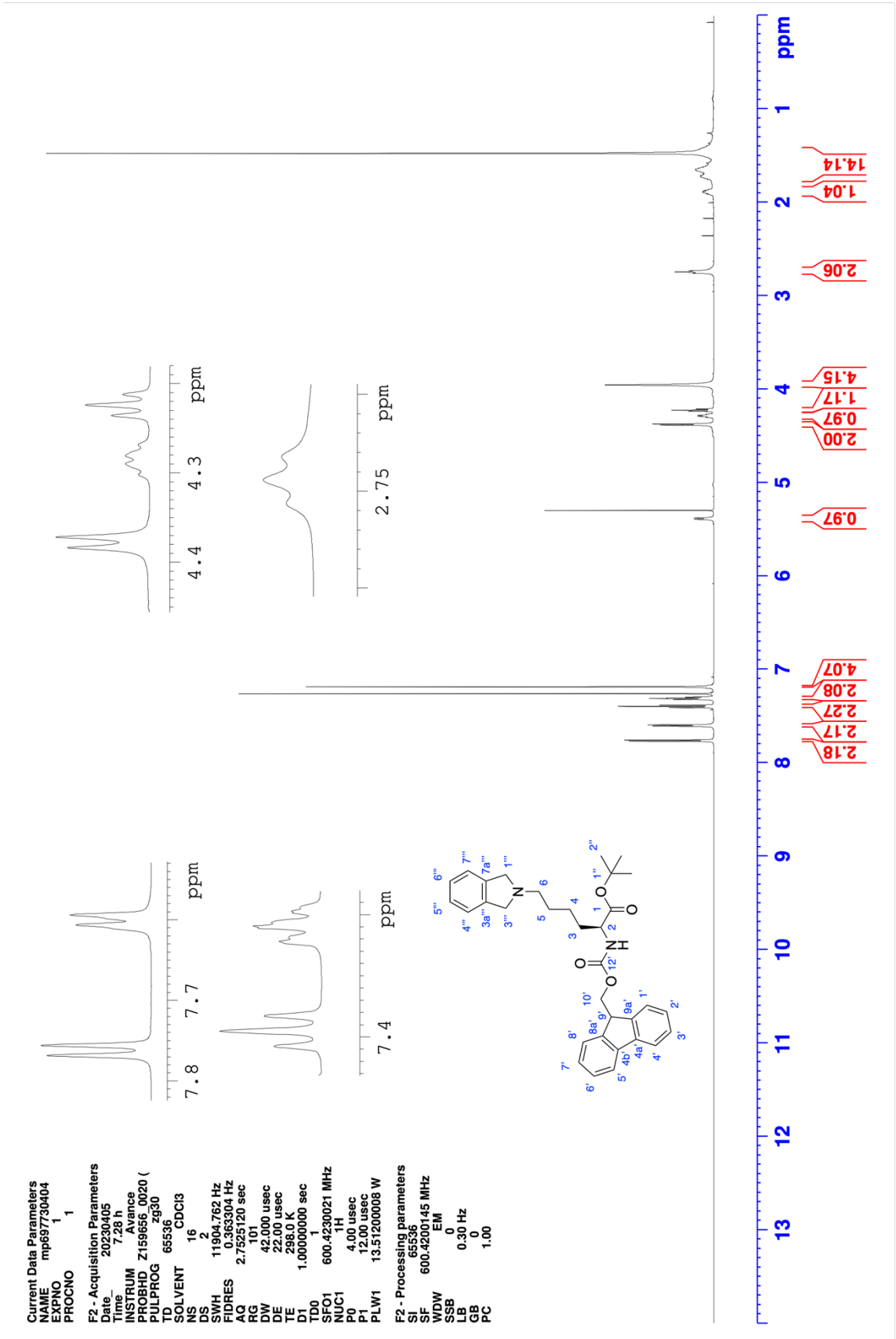
**(46) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-(isoindolin-2'''-yl)hex-4-enoate $^{1}$H NMR**



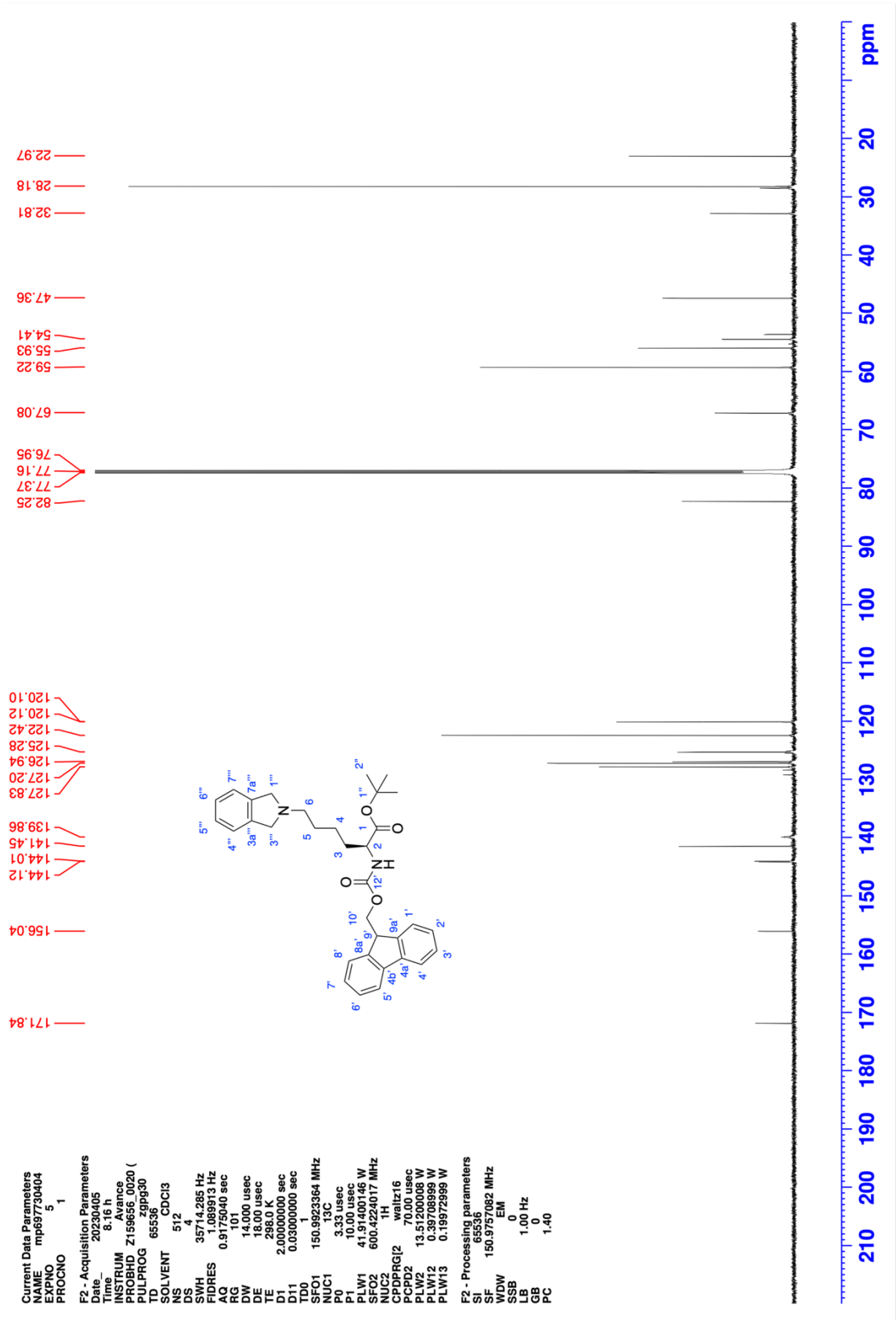
**(46) *tert*-Butyl (*S*,*E*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-(isoindolin-2'''-yl)hex-4-enoate $^{13}$C NMR**



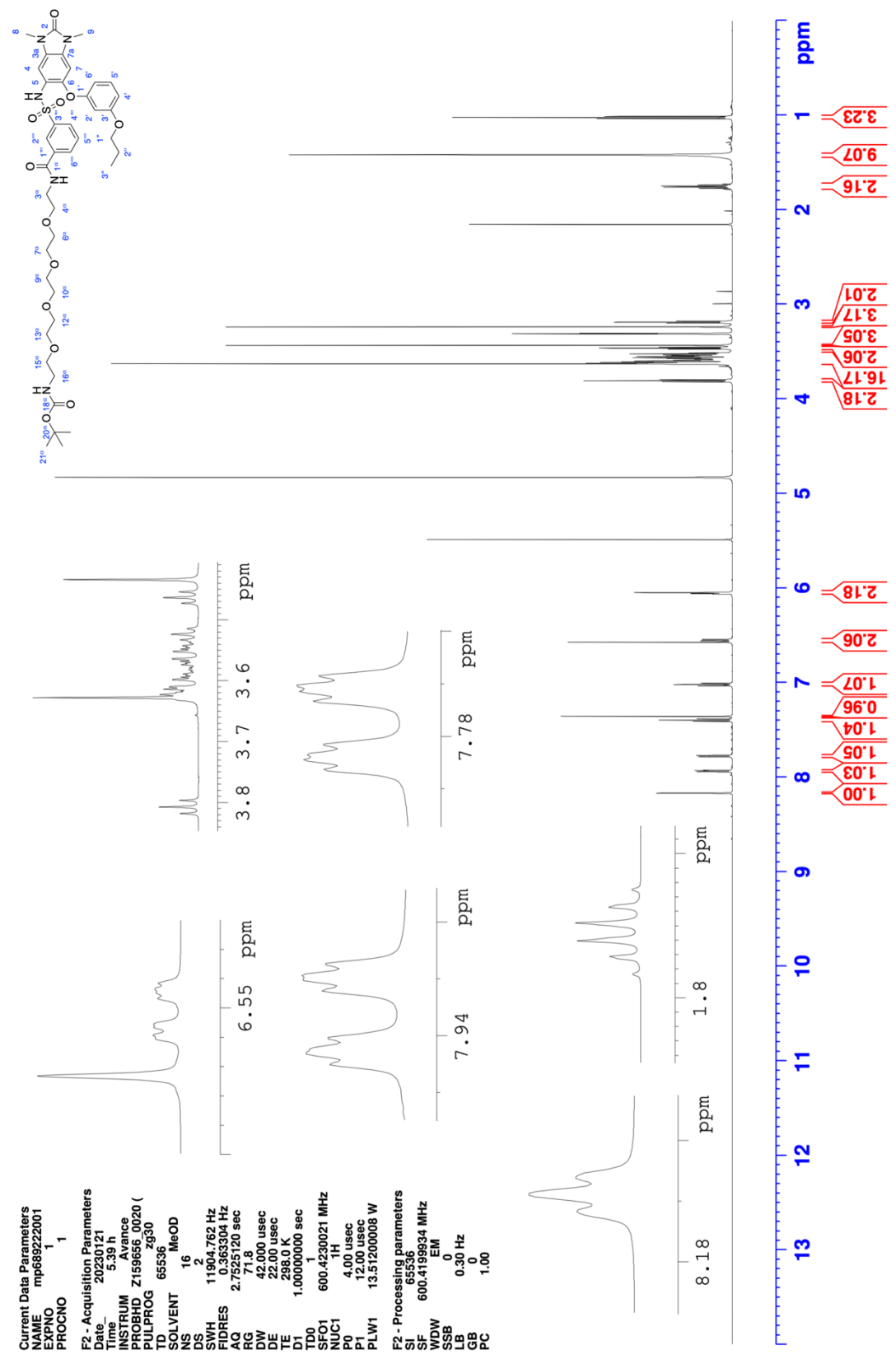
**(47) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-(isoindolin-2'''-yl)hexanoate $^{1}$H NMR**



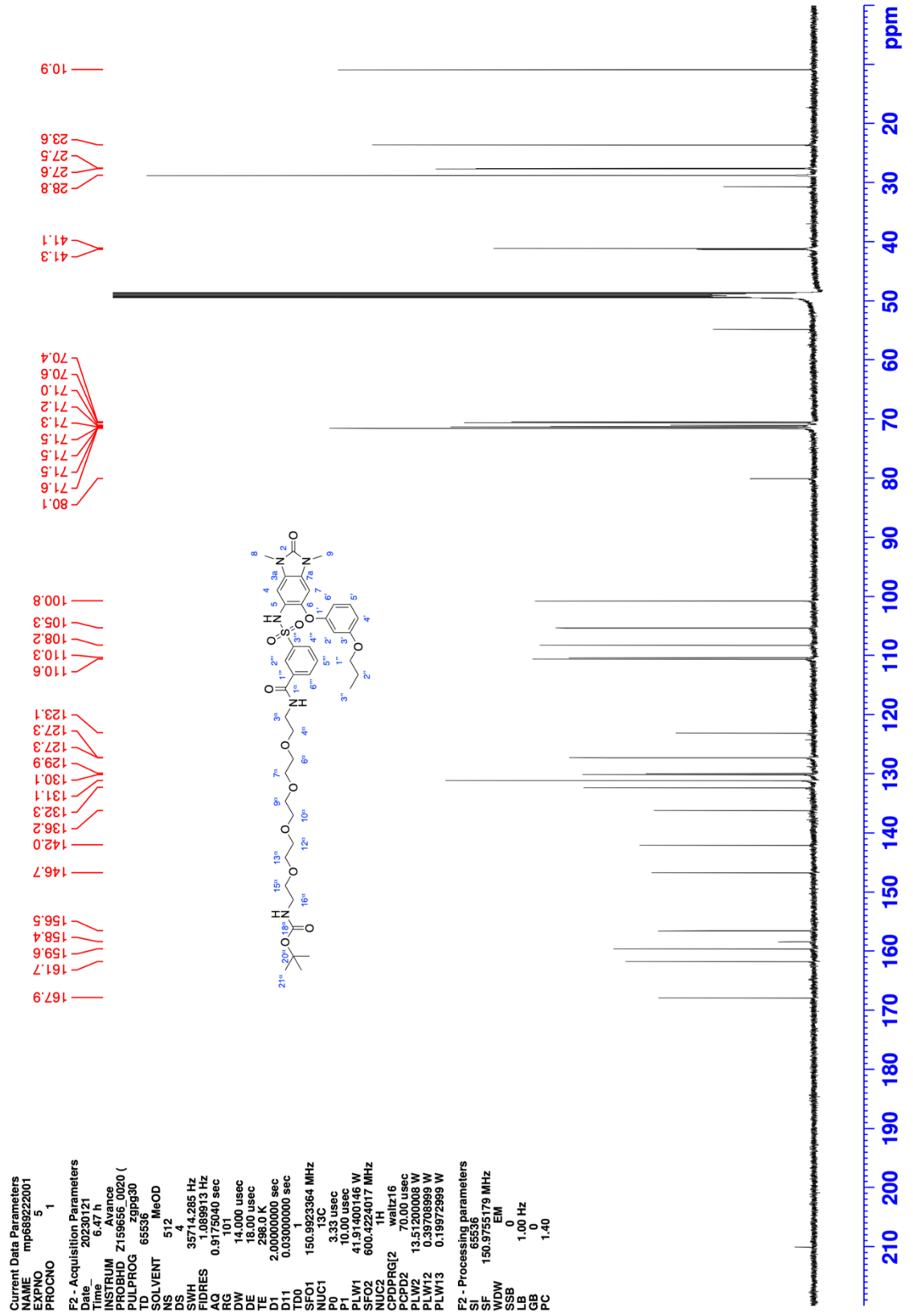
**(47) *tert*-Butyl (*S*)-2-((((9'*H*-fluoren-9'-yl)methoxy)carbonyl)amino)-6-(isoindolin-2'''-yl)hexanoate $^{13}$C NMR**



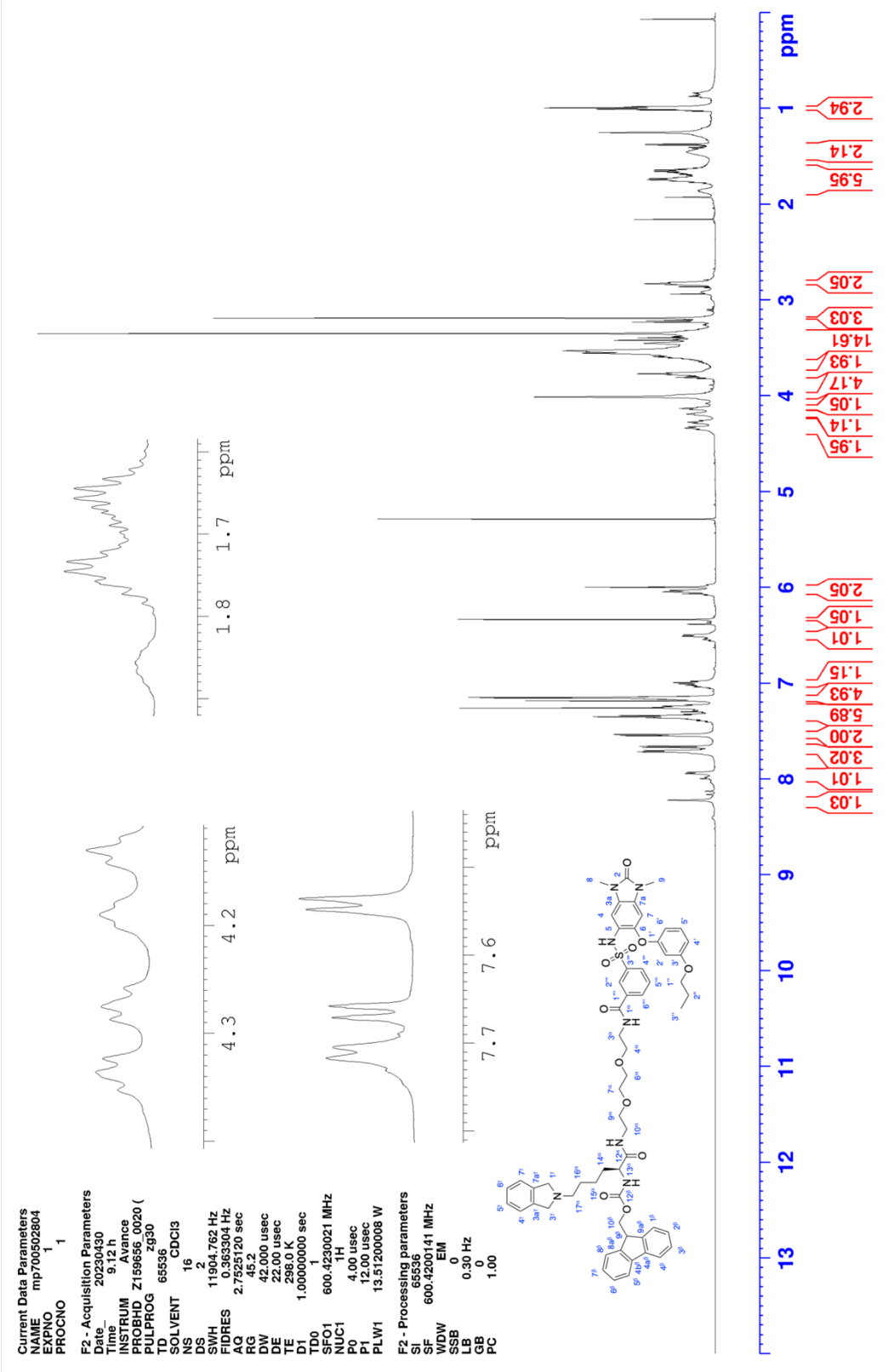
**(51) *tert*-Butyl (1$^{\alpha}$-(3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-1$^{\alpha}$-oxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$,14$^{\alpha}$-tetraoxa-2$^{\alpha}$-azahexadecan-16$^{\alpha}$-yl)carbamate $^{1}$H NMR**



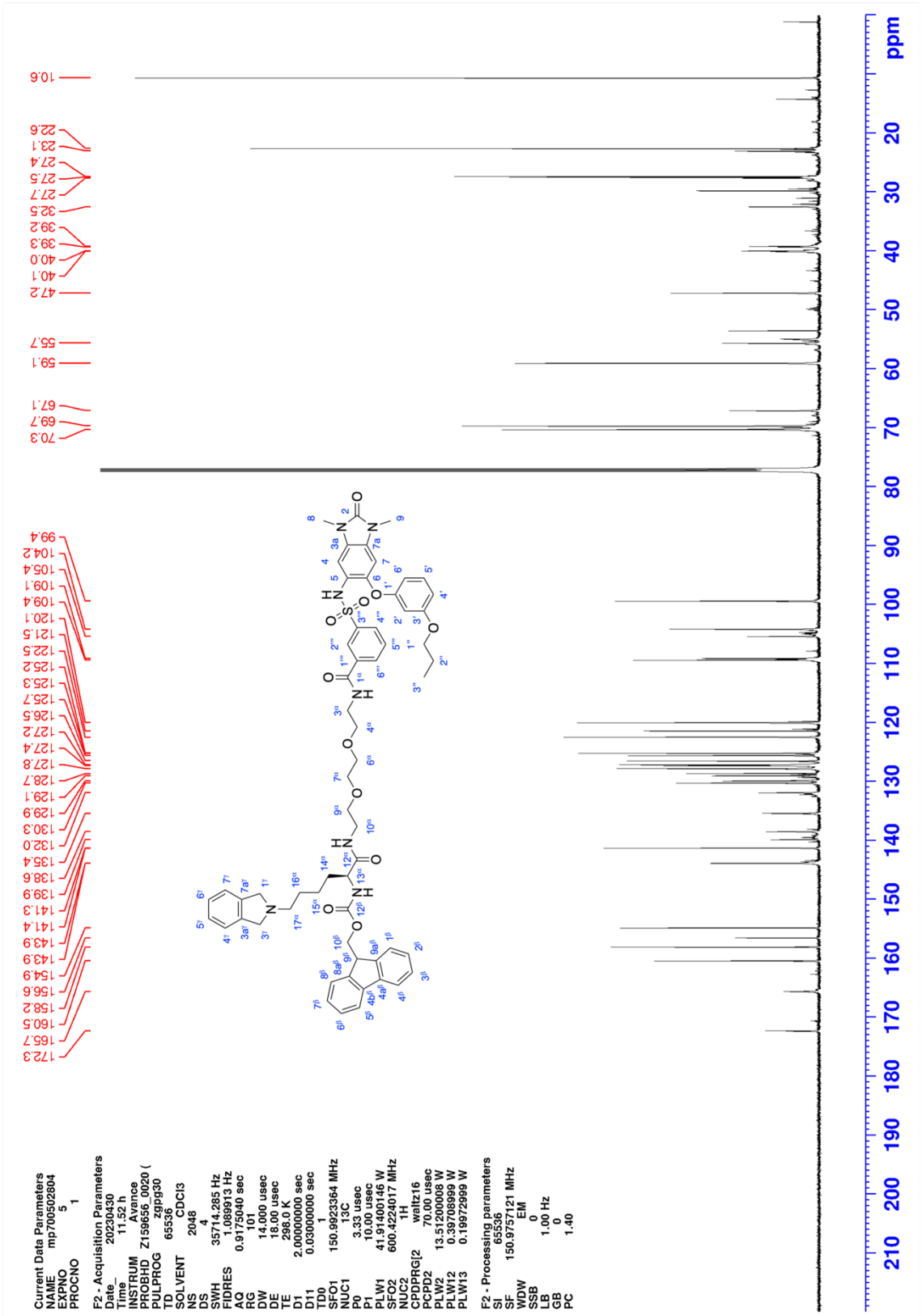
**(51)** ***tert*-Butyl (1$^{\alpha}$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-1$^{\alpha}$-oxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$,14$^{\alpha}$-tetraoxa-2$^{\alpha}$-azahexadecan-16$^{\alpha}$-yl)carbamate $^{13}$C NMR**



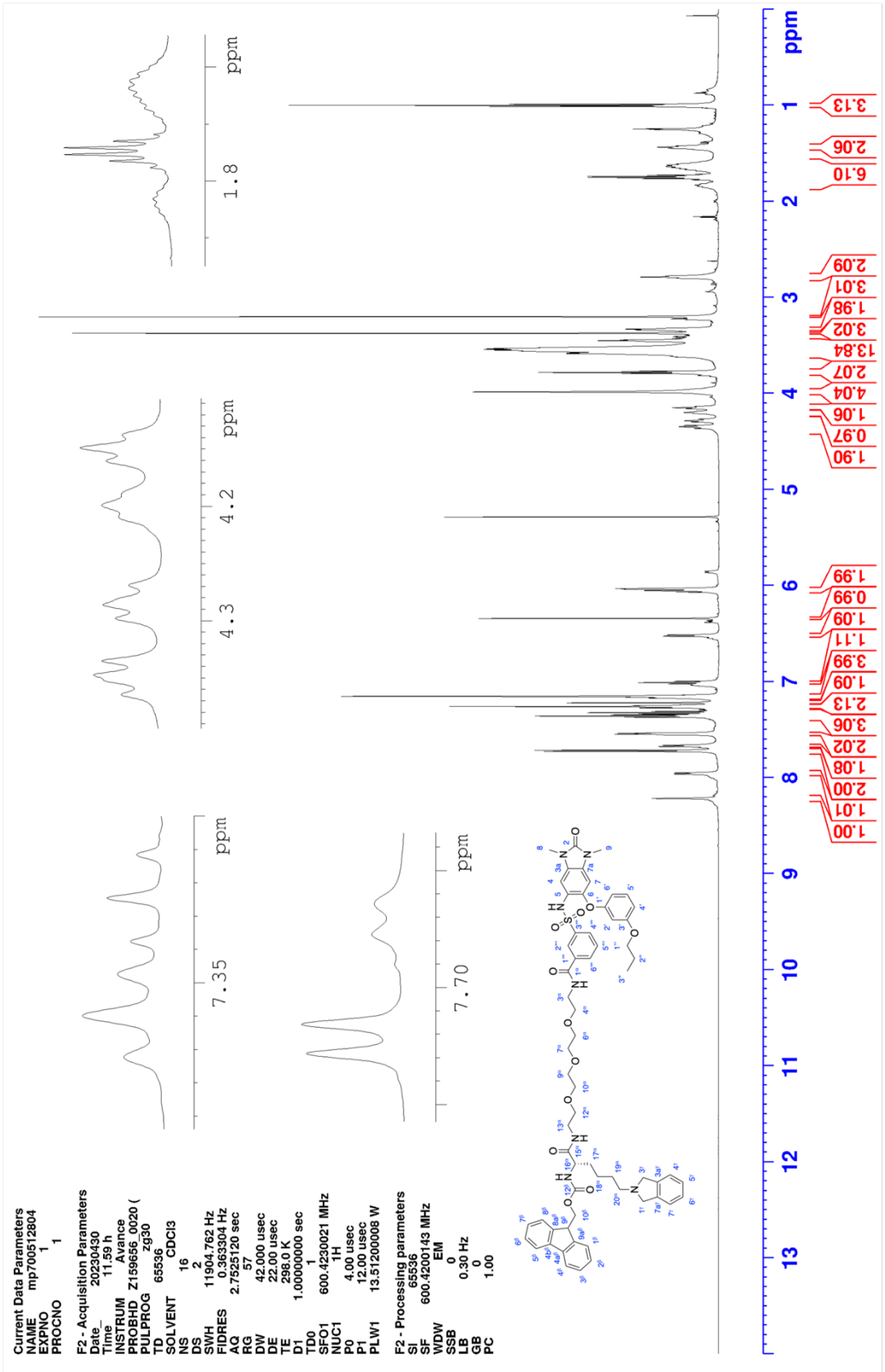
**(52) (9$^{\beta}$*H*-Fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-17$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,12$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$-dioxa-2$^{\alpha}$,11$^{\alpha}$-diazaheptadecan-13$^{\alpha}$-yl)carbamate $^{1}$H NMR**



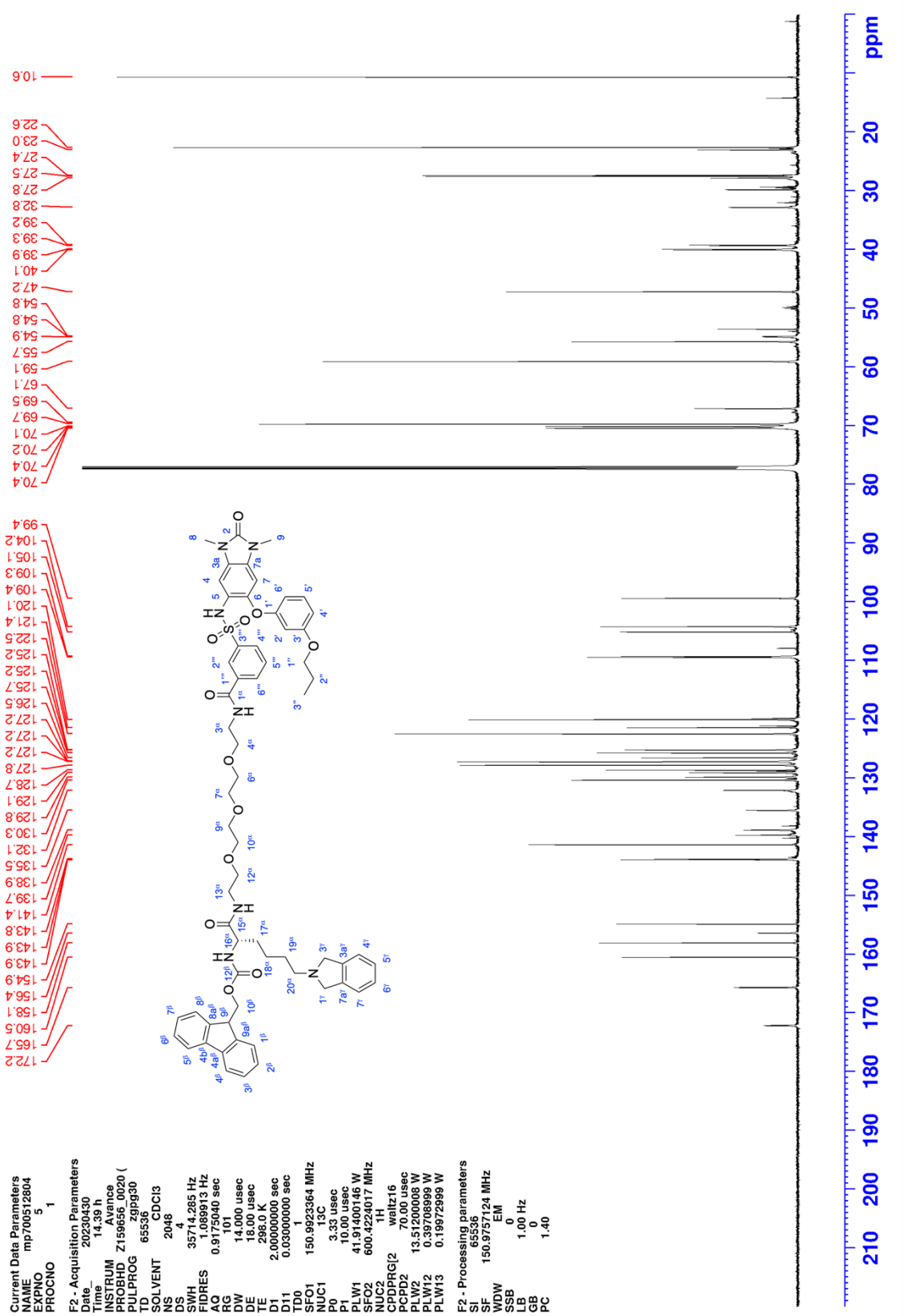
**(52) (9$^{\beta}$*H*-Fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-17$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,12$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$-dioxa-2$^{\alpha}$,11$^{\alpha}$-diazaheptadecan-13$^{\alpha}$-yl)carbamate $^{13}$C NMR**



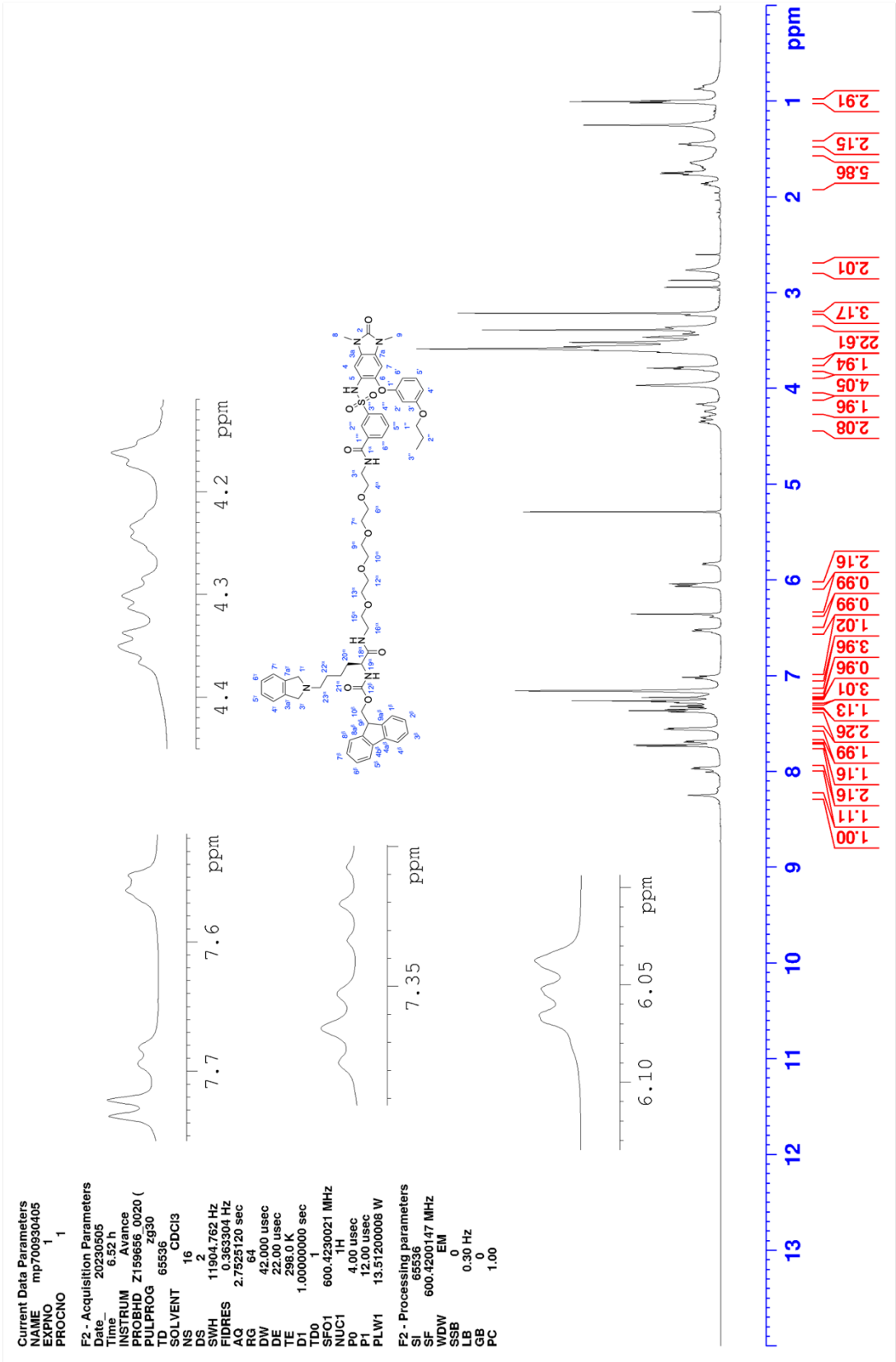
**(53)** **(9$^{\beta}$*H*-Fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-20$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,15$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$-trioxa-2$^{\alpha}$,14$^{\alpha}$-diazaicosan-16$^{\alpha}$-yl)carbamate $^{1}$H NMR**



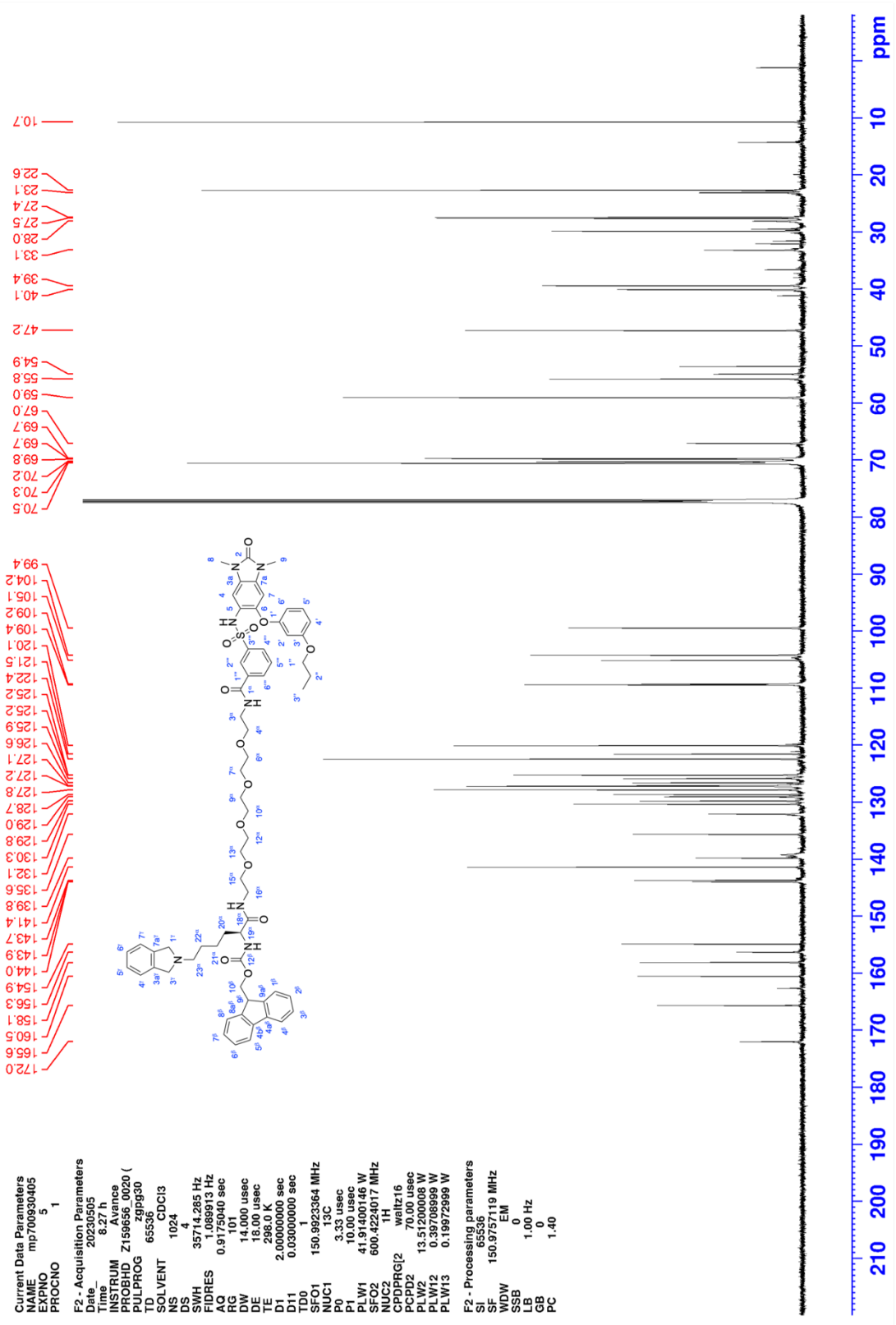
**(53) (9$^{\beta}$*H*-fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-20$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,15$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$-trioxa-2$^{\alpha}$,14$^{\alpha}$-diazaicosan-16$^{\alpha}$-yl)carbamate $^{13}$C NMR**



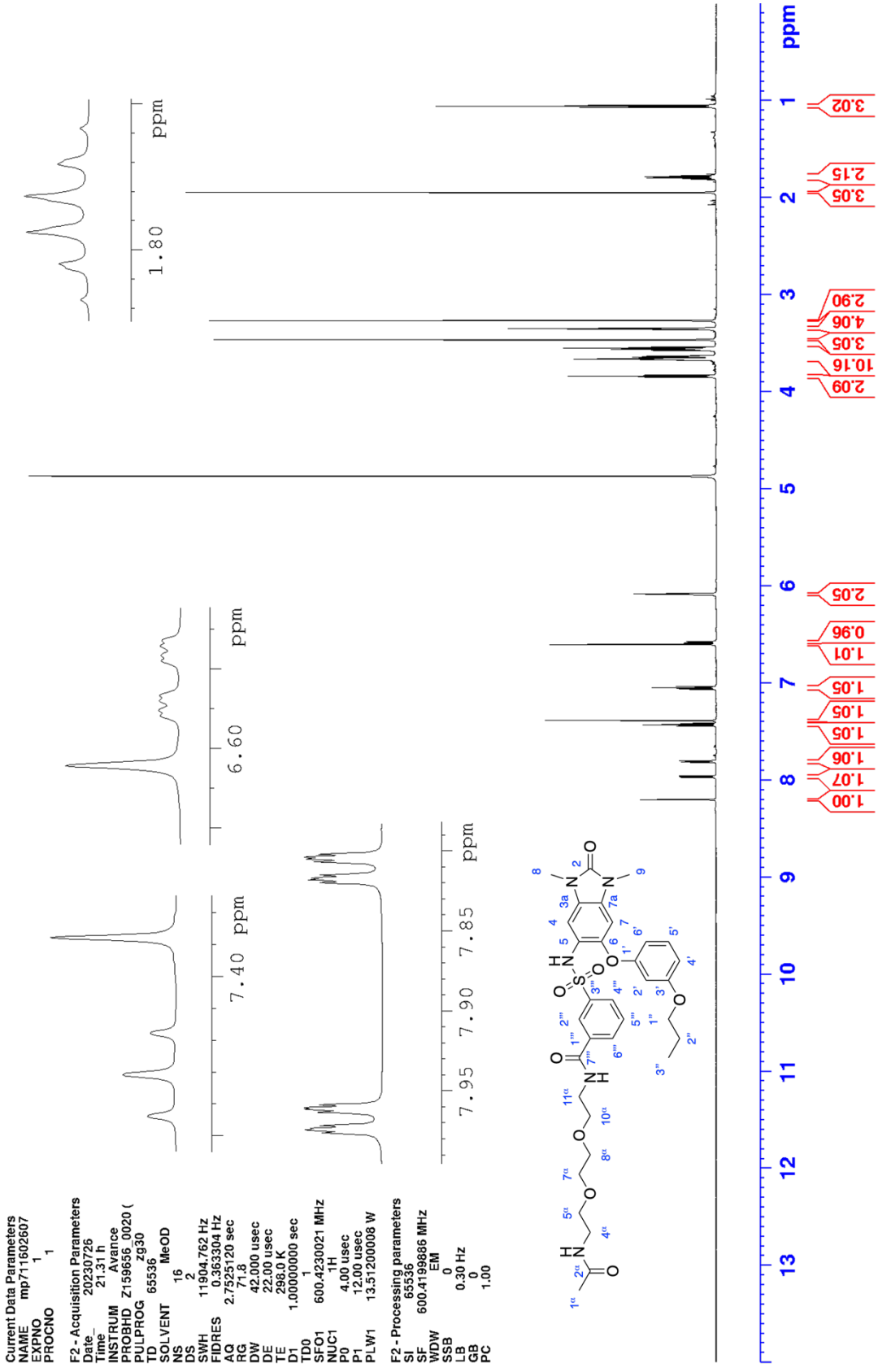
**(54) (9$^{\beta}$*H*-Fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-23$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,18$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$,14$^{\alpha}$-tetraoxa-2$^{\alpha}$,17$^{\alpha}$-diazatricosan-19$^{\alpha}$-yl)carbamate $^{1}$H NMR**



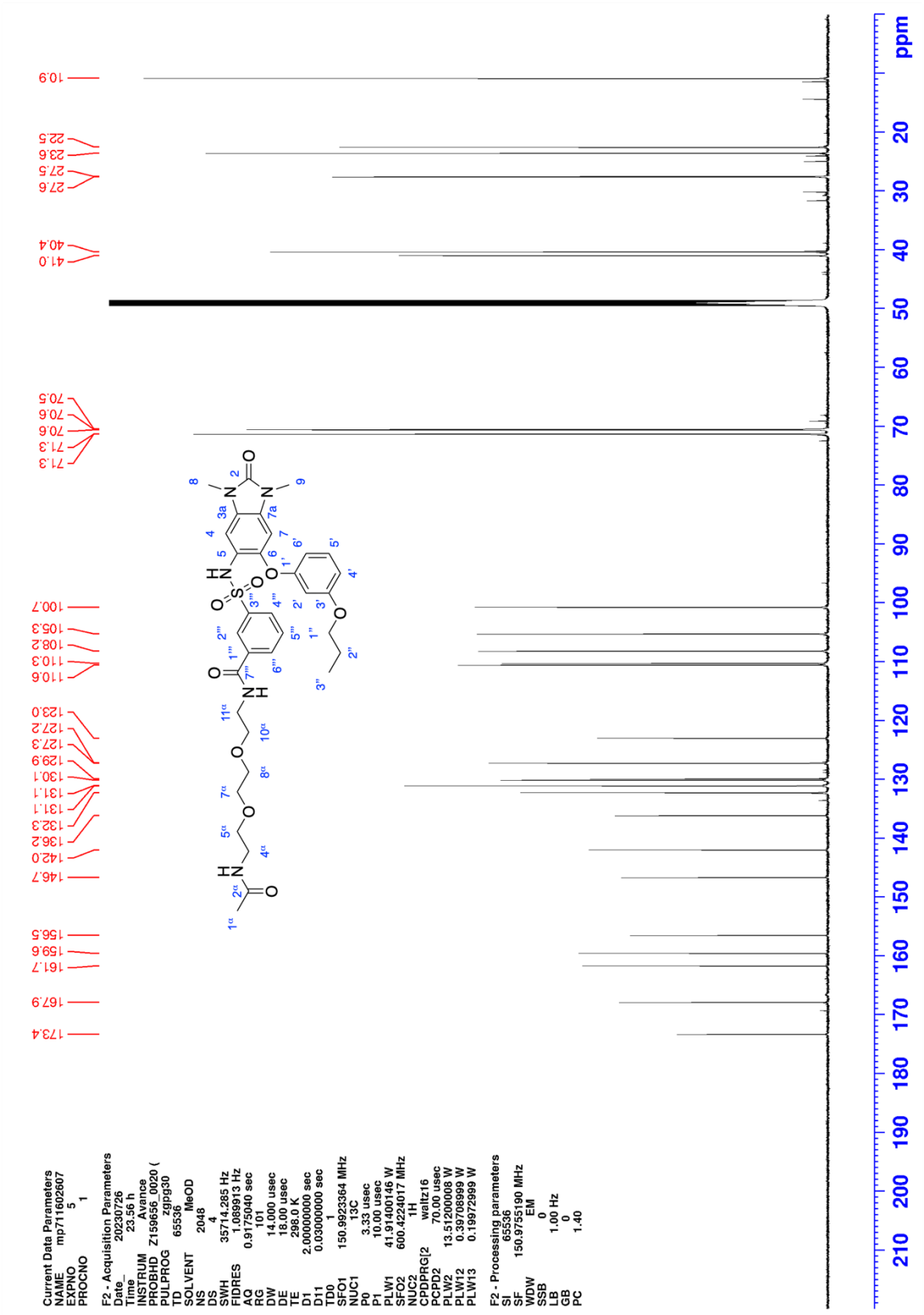
**(54) (9$^{\beta}$*H*-Fluoren-9$^{\beta}$-yl)methyl (*S*)-(1$^{\alpha}$-(3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)phenyl)-23$^{\alpha}$-(isoindolin-2$^{\gamma}$-yl)-1$^{\alpha}$,18$^{\alpha}$-dioxo-5$^{\alpha}$,8$^{\alpha}$,11$^{\alpha}$,14$^{\alpha}$-tetraoxa-2$^{\alpha}$,17$^{\alpha}$-diazatricosan-19$^{\alpha}$-yl)carbamate $^{13}$C NMR**



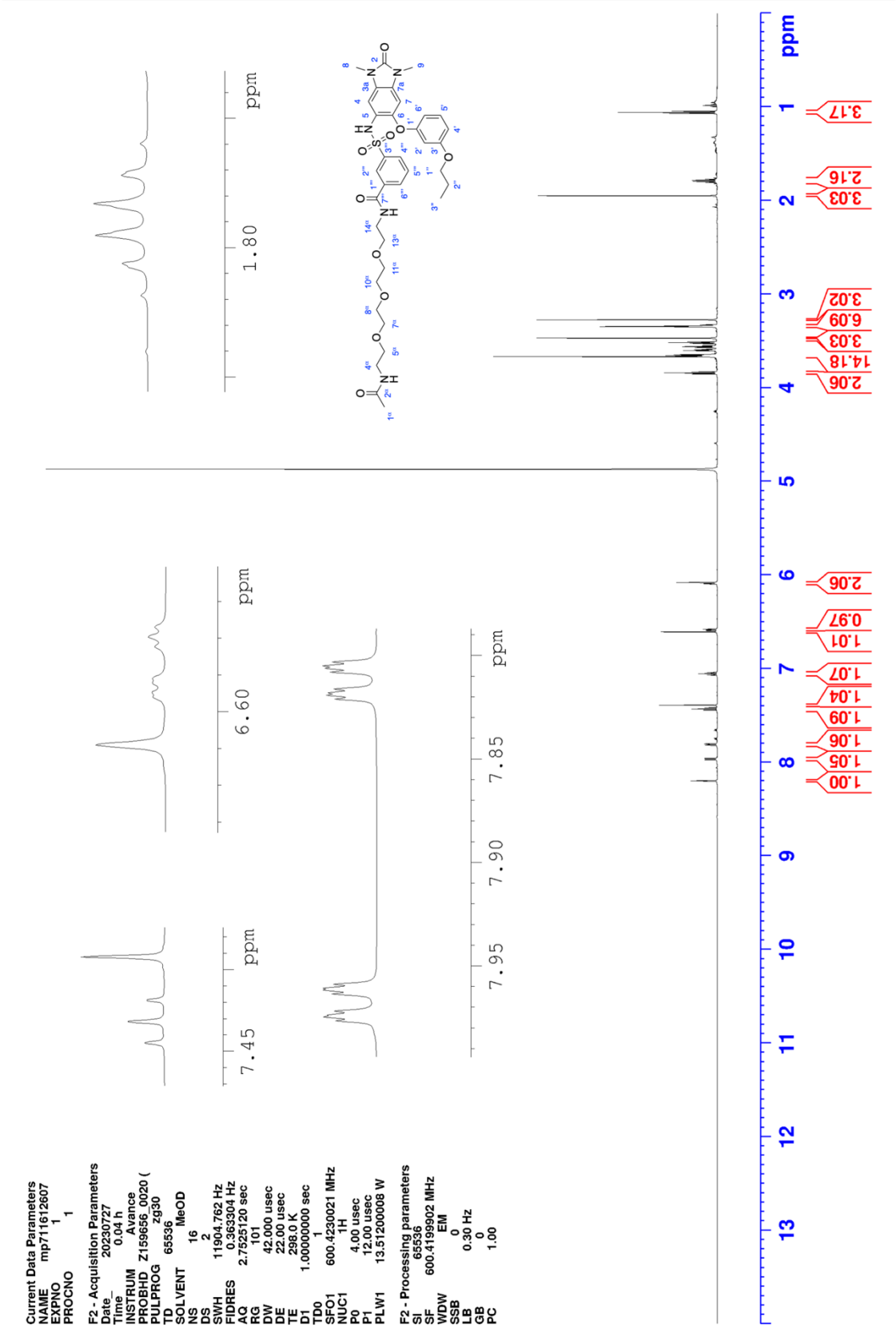
**(61) 3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$-dioxa-3$^{\alpha}$-azaundecan-11$^{\alpha}$-yl)benzamide $^{1}$H NMR**



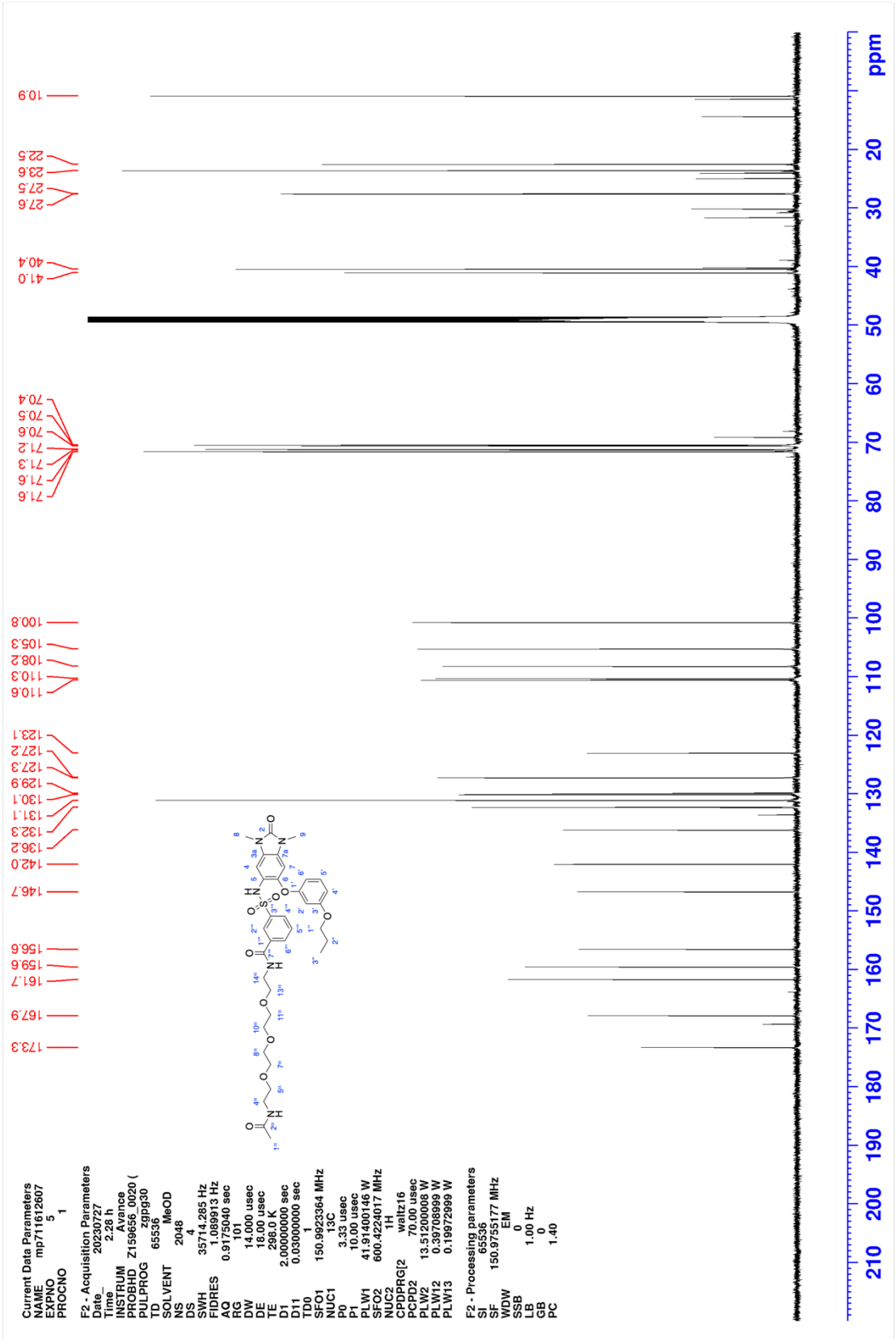
**(61) 3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$-dioxa-3$^{\alpha}$-azaundecan-11$^{\alpha}$-yl)benzamide $^{13}$C NMR**



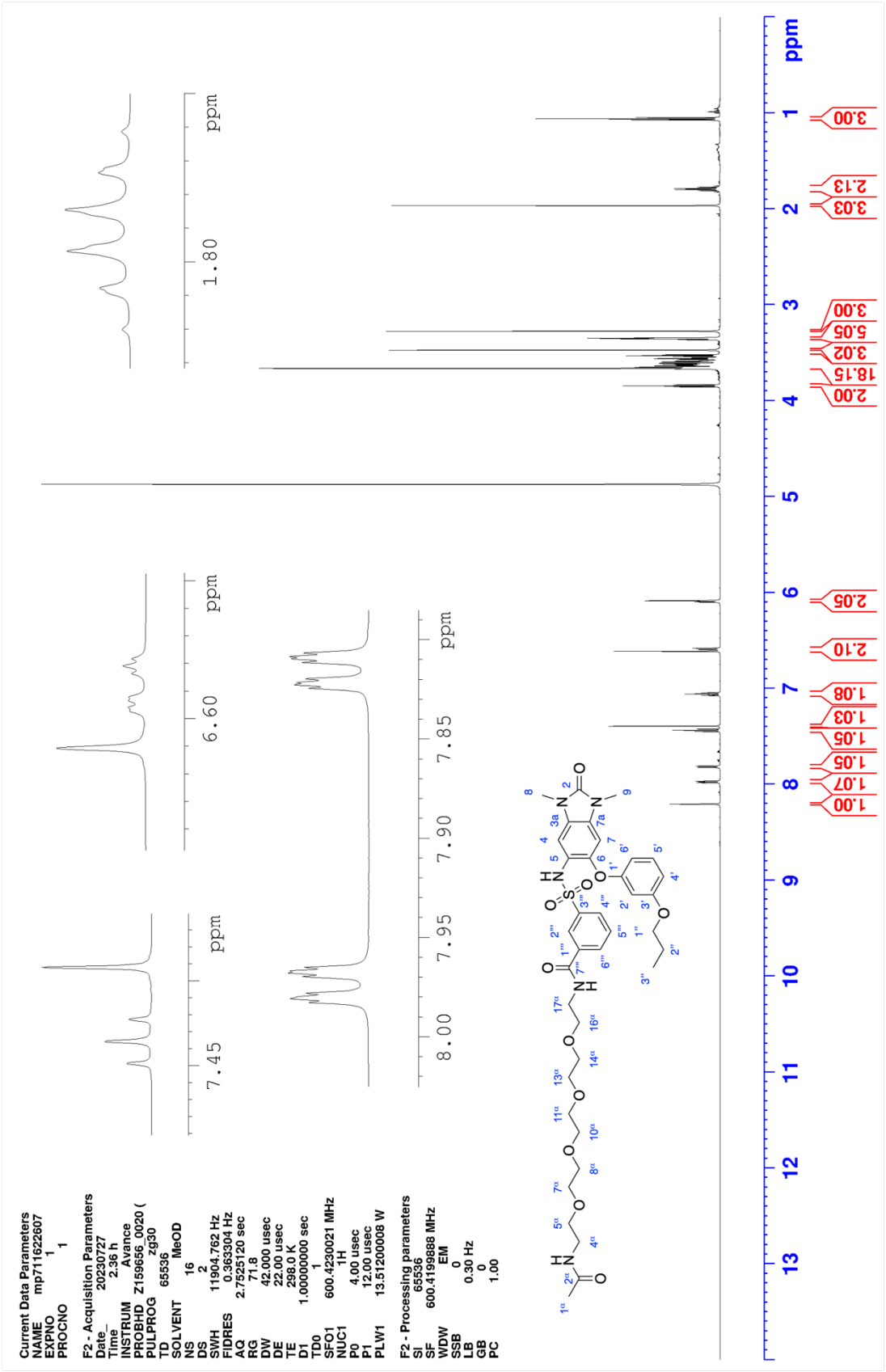
**(62) 3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$,12$^{\alpha}$-trioxa-3$^{\alpha}$-azatetradecan-14$^{\alpha}$-yl)benzamide $^{1}$H NMR**



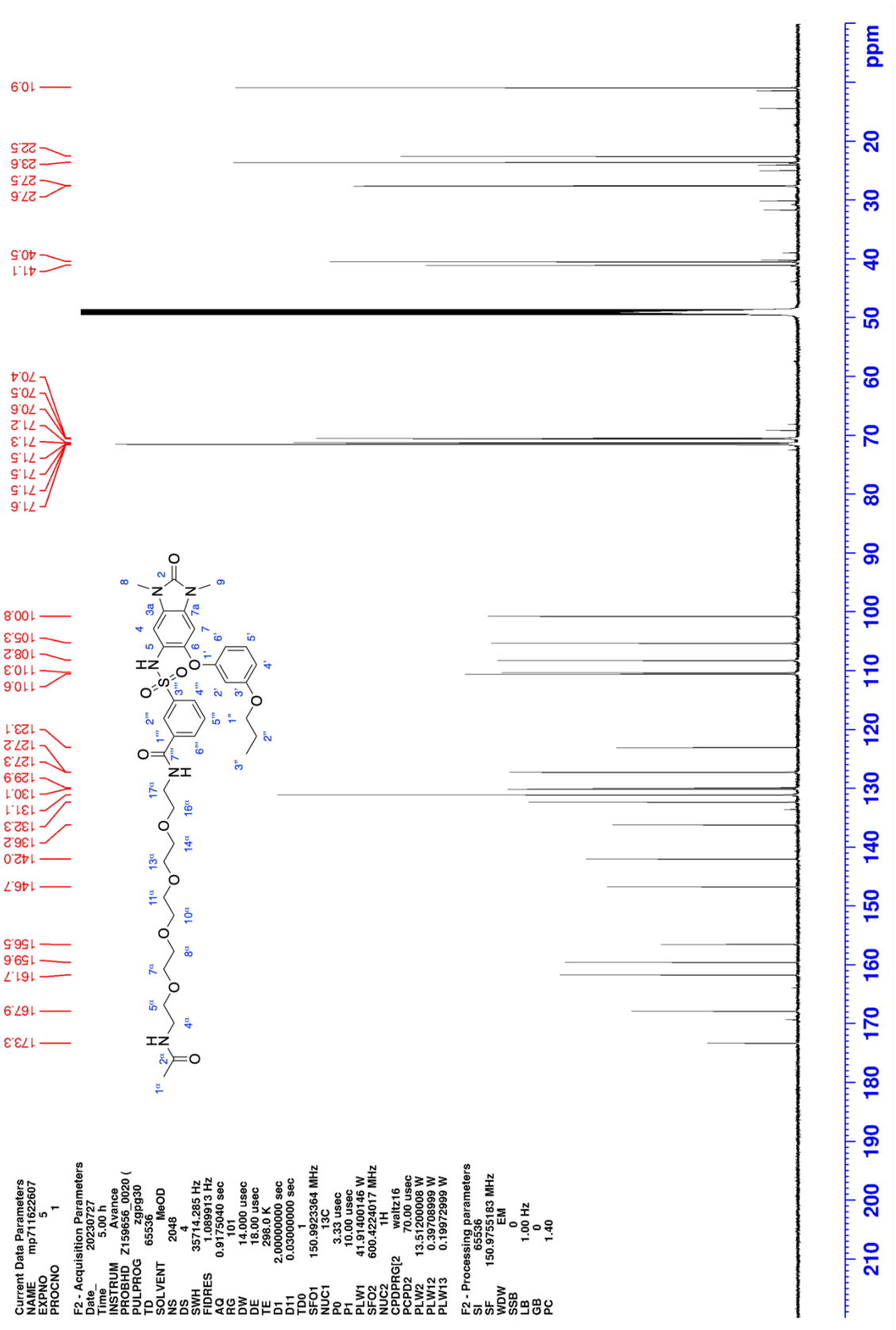
**(62) 3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$,12$^{\alpha}$-trioxa-3$^{\alpha}$-azatetradecan-14$^{\alpha}$-yl)benzamide $^{13}$C NMR**



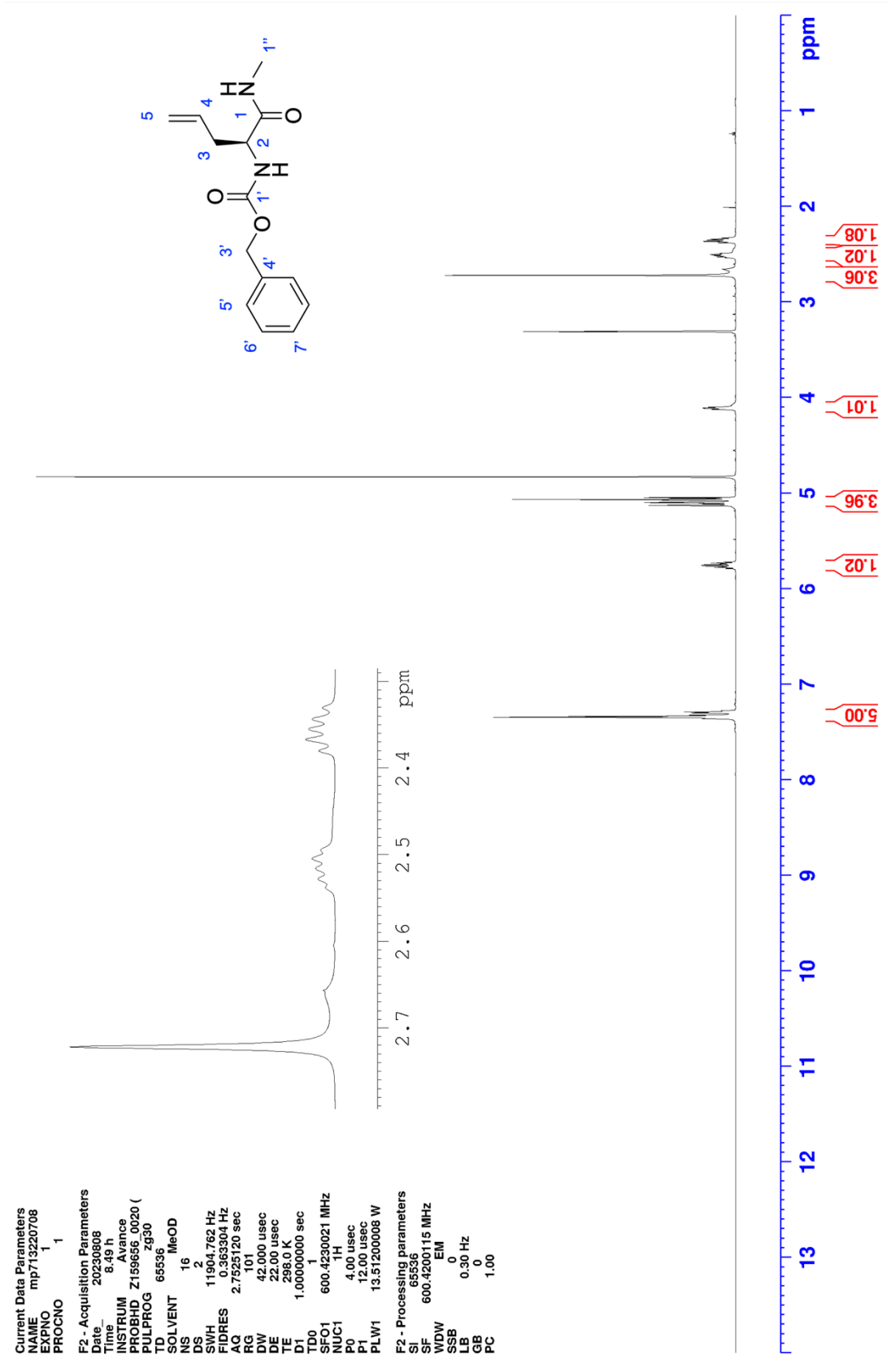
**(63) 3’’’-(*N*-(1,3-dimethyl-2-oxo-6-(3’-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$,12$^{\alpha}$,15$^{\alpha}$-tetraoxa-3$^{\alpha}$-azaheptadecan-17$^{\alpha}$-yl)benzamide $^{1}$H NMR**



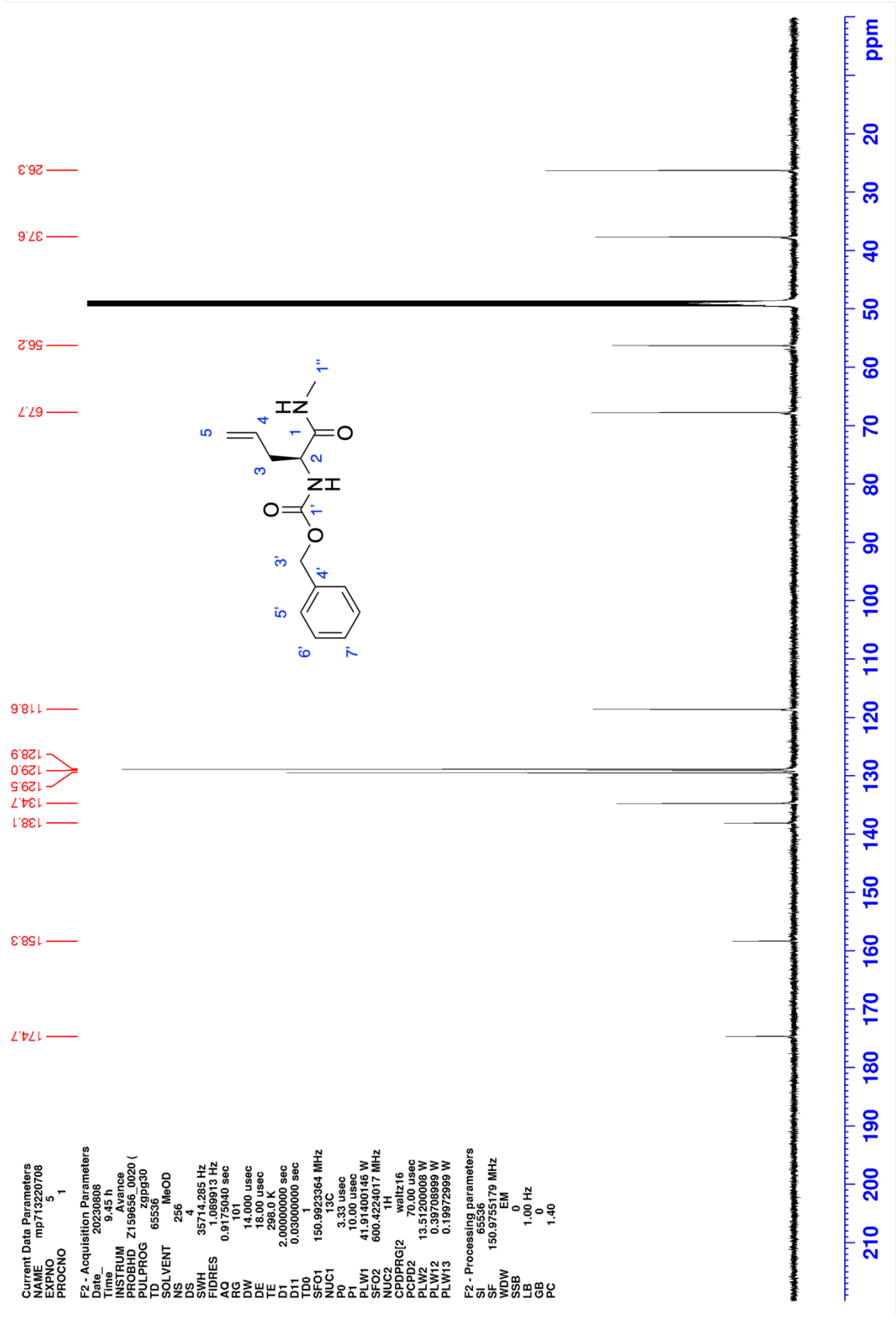
**(63) 3'''-(*N*-(1,3-dimethyl-2-oxo-6-(3'-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2$^{\alpha}$-oxo-6$^{\alpha}$,9$^{\alpha}$,12$^{\alpha}$,15$^{\alpha}$-tetraoxa-3$^{\alpha}$-azaheptadecan-17$^{\alpha}$-yl)benzamide $^{13}$C NMR**



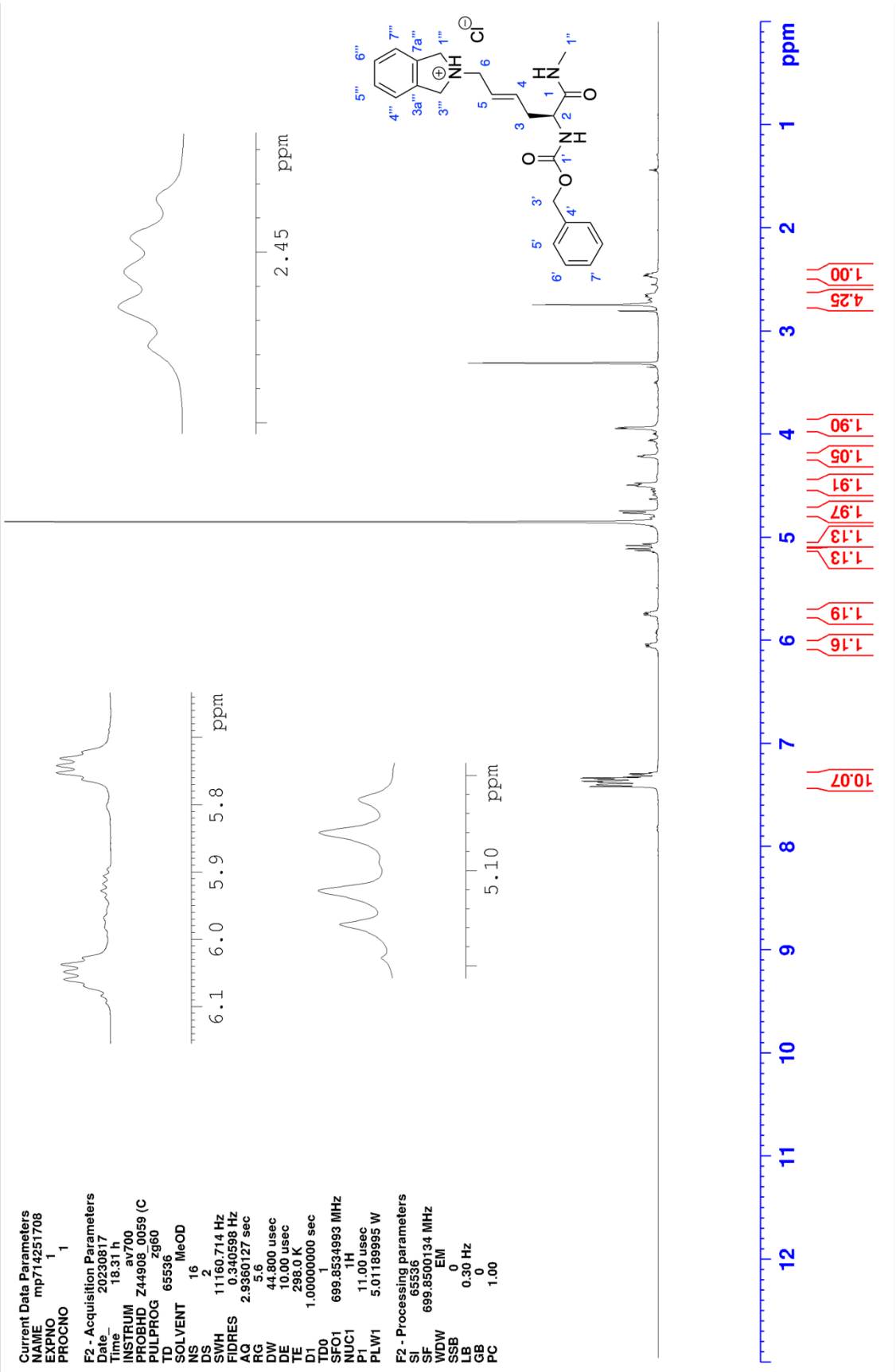
## (64) Benzyl (*S*)-(1-(methylamino)-1-oxopent-4-en-2-yl)carbamate $^{1}$H NMR



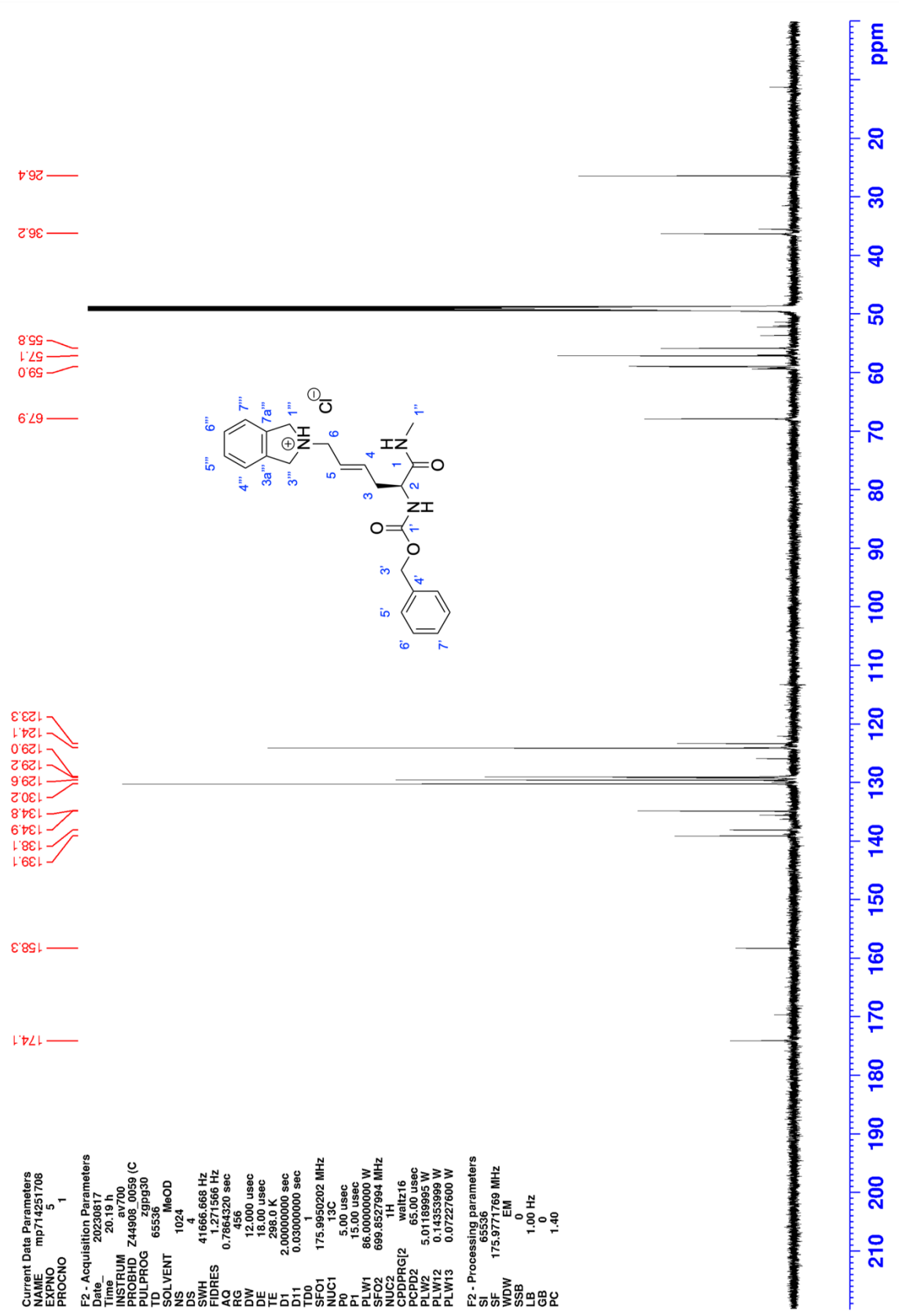
## (64) Benzyl (*S*)-(1-(methylamino)-1-oxopent-4-en-2-yl)carbamate $^{13}$C NMR



**(65) Benzyl (*S*,*E*)-(6-(isoindolin-2’’’-yl)-1-(methylamino)-1-oxohex-4-en-2-yl)carbamate hydrochloride $^{1}$H NMR**

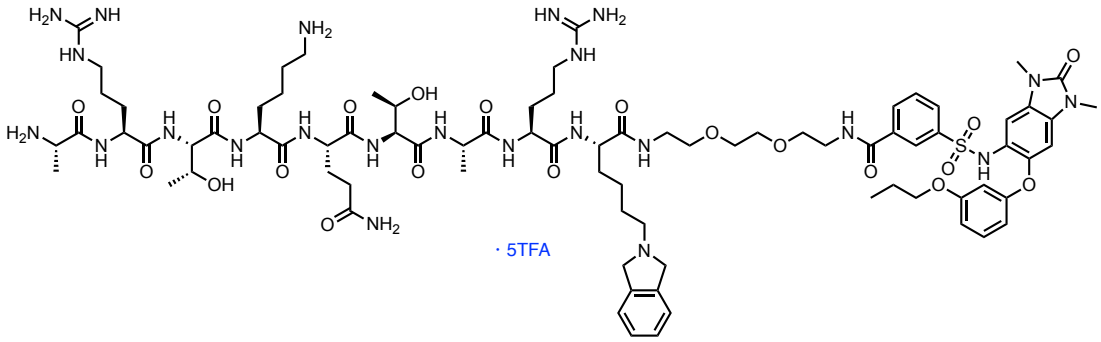


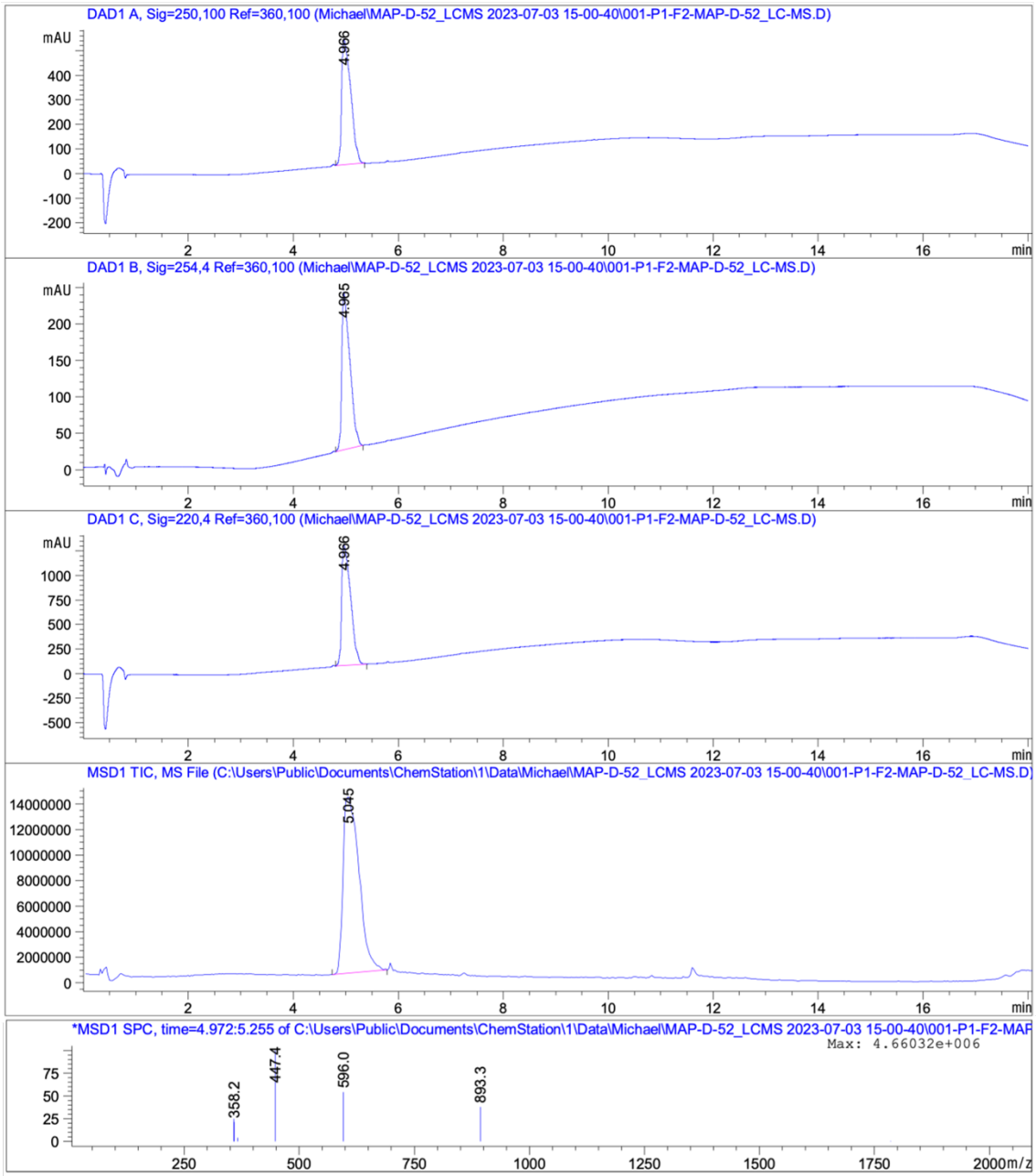
**(65) Benzyl (*S*,*E*)-(6-(isoindolin-2’’’-yl)-1-(methylamino)-1-oxohex-4-en-2-yl)carbamate hydrochloride $^{13}$C NMR**



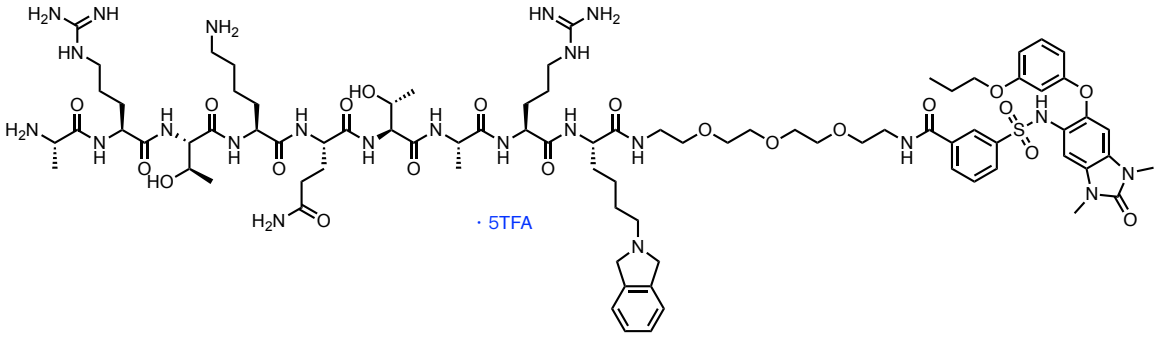
## Appendix 10 LC-MS Chromatograms for Compounds 58–60 and 67

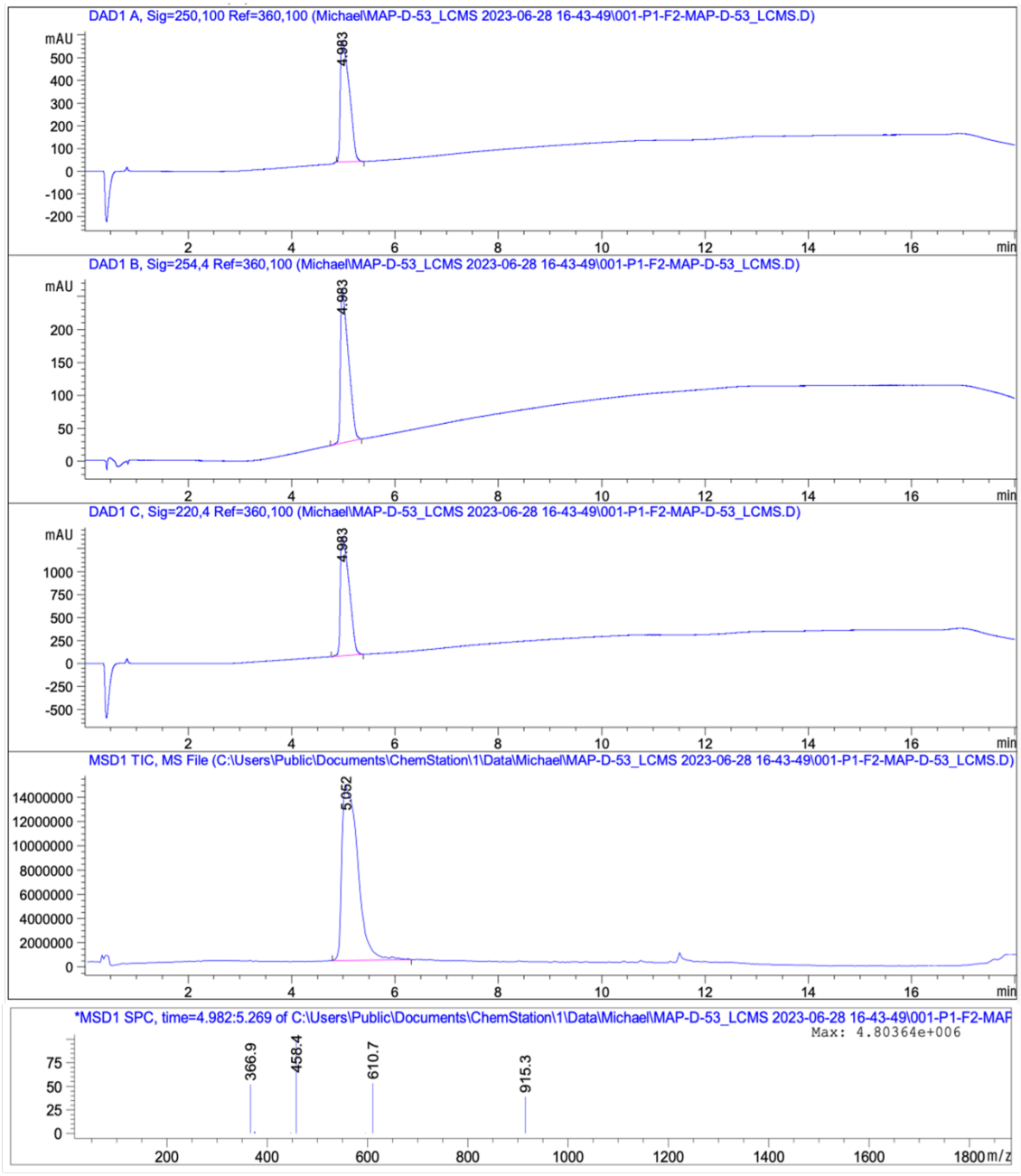
### (58) PDC4



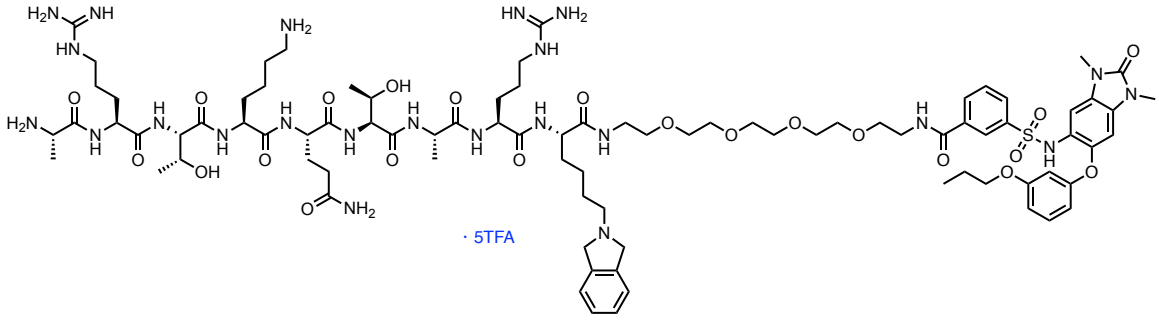



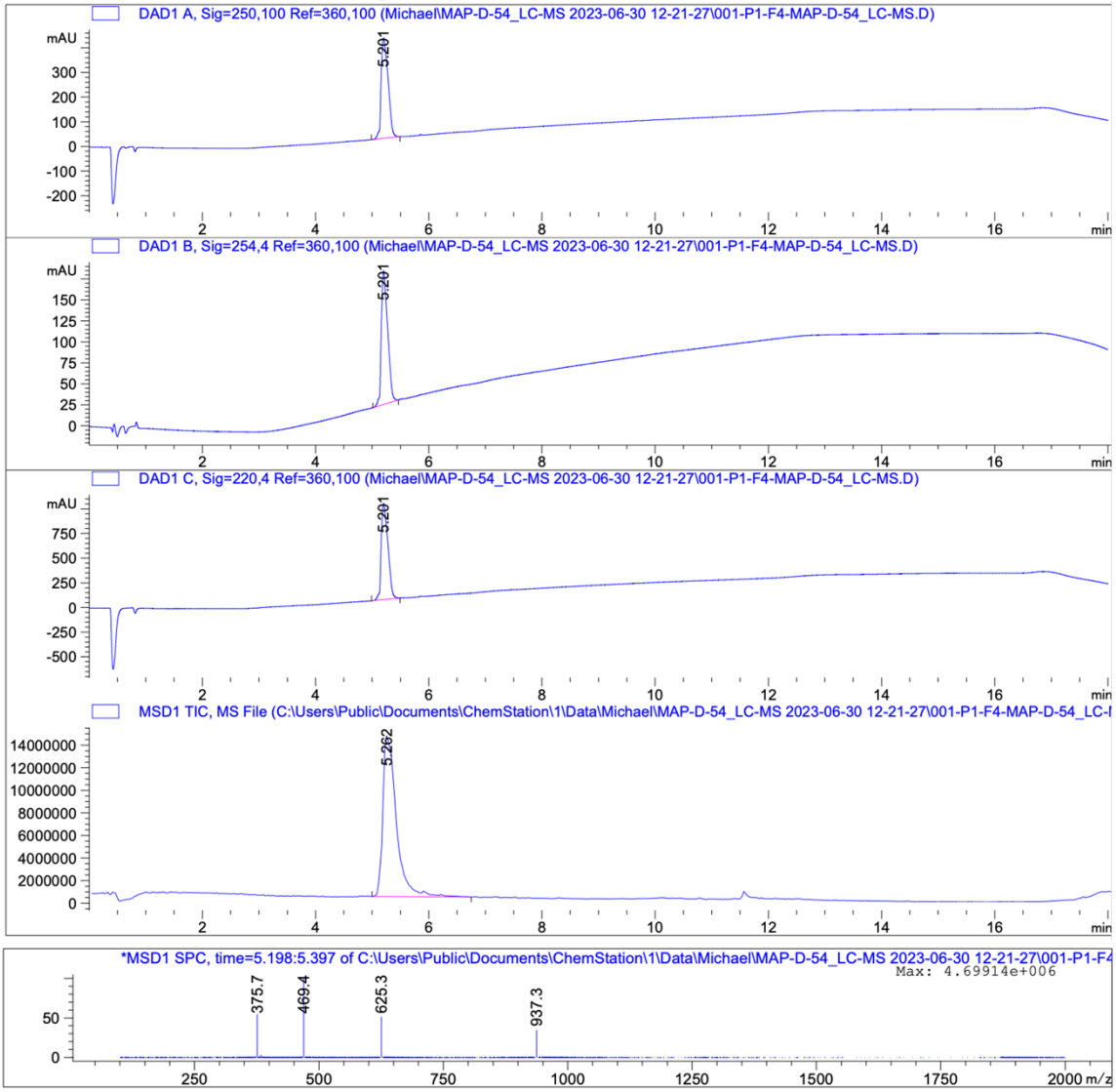
## (59) PDC5


· 5TFA


DAD1 A, Sig=250,100 Ref=360,100 (Michael\MAP-D-53_LCMS 2023-06-28 16-43-49\001-P1-F2-MAP-D-53_LCMS.D)
mAU
4.983
min
DAD1 B, Sig=254,4 Ref=360,100 (Michael\MAP-D-53_LCMS 2023-06-28 16-43-49\001-P1-F2-MAP-D-53_LCMS.D)
mAU
4.983
min
DAD1 C, Sig=220,4 Ref=360,100 (Michael\MAP-D-53_LCMS 2023-06-28 16-43-49\001-P1-F2-MAP-D-53_LCMS.D)
mAU
4.983
min
MSD1 TIC, MS File (C:\Users\Public\Documents\ChemStation\1\Data\Michael\MAP-D-53_LCMS 2023-06-28 16-43-49\001-P1-F2-MAP-D-53_LCMS.D)
5.052
min
*MSD1 SPC, time=4.982:5.269 of C:\Users\Public\Documents\ChemStation\1\Data\Michael\MAP-D-53_LCMS 2023-06-28 16-43-49\001-P1-F2-MAP
Max: 4.80364e+006
366.9
458.4
610.7
915.3
m/z

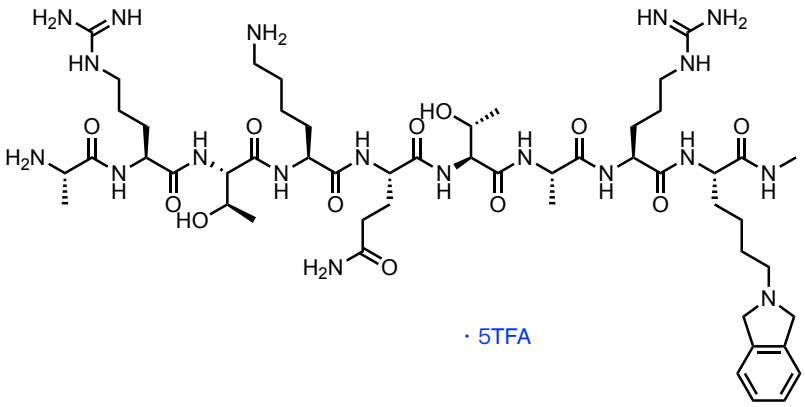
## (60) PDC6

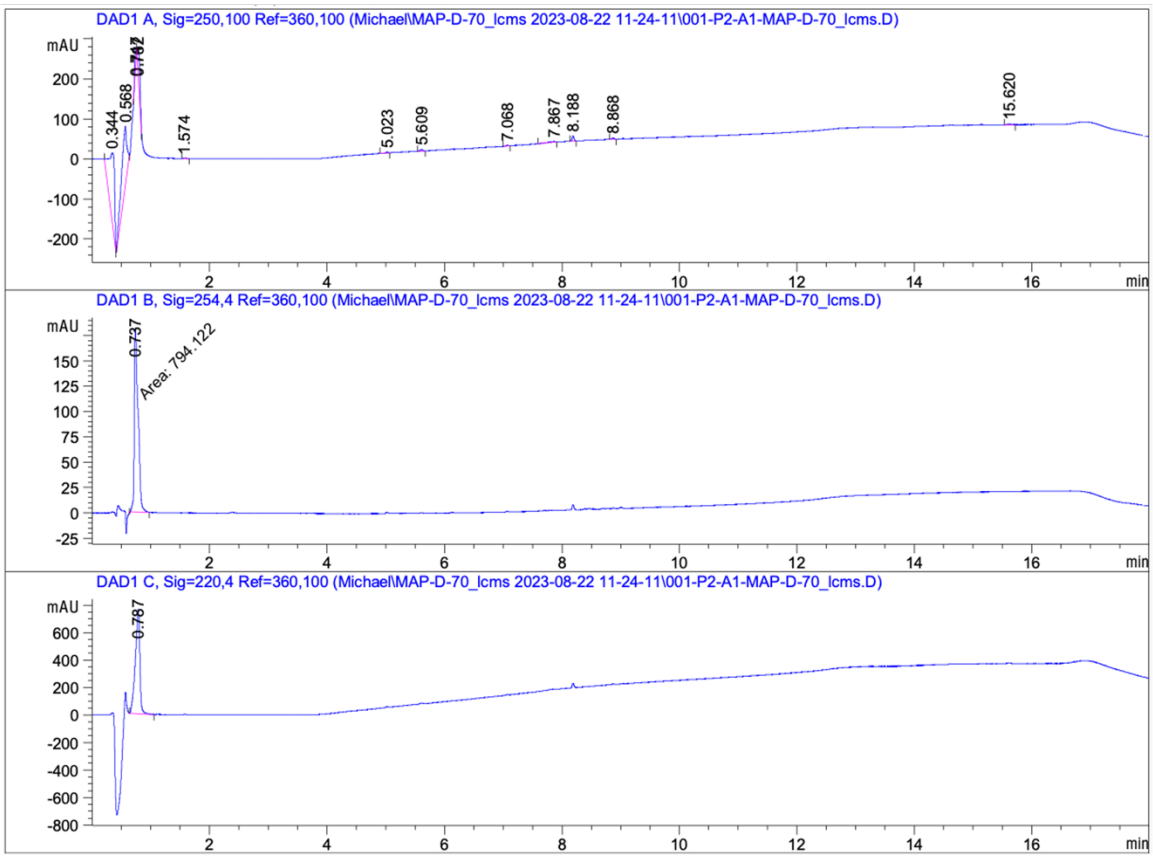

· 5TFA


DAD1 A, Sig=250,100 Ref=360,100 (Michael\MAP-D-54_LC-MS 2023-06-30 12-21-27\001-P1-F4-MAP-D-54_LC-MS.D)
mAU
5.201
min
DAD1 B, Sig=254,4 Ref=360,100 (Michael\MAP-D-54_LC-MS 2023-06-30 12-21-27\001-P1-F4-MAP-D-54_LC-MS.D)
mAU
5.201
min
DAD1 C, Sig=220,4 Ref=360,100 (Michael\MAP-D-54_LC-MS 2023-06-30 12-21-27\001-P1-F4-MAP-D-54_LC-MS.D)
mAU
5.201
min
MSD1 TIC, MS File (C:\Users\Public\Documents\ChemStation\1\Data\Michael\MAP-D-54_LC-MS 2023-06-30 12-21-27\001-P1-F4-MAP-D-54_LC-
5.262
min


*MSD1 SPC, time=5.198:5.397 of C:\Users\Public\Documents\ChemStation\1\Data\Michael\MAP-D-54_LC-MS 2023-06-30 12-21-27\001-P1-F4
Max: 4.69914e+006
375.7
469.4
625.3
937.3
m/z

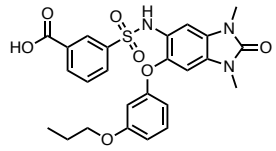
## (67) PHD Control Peptide

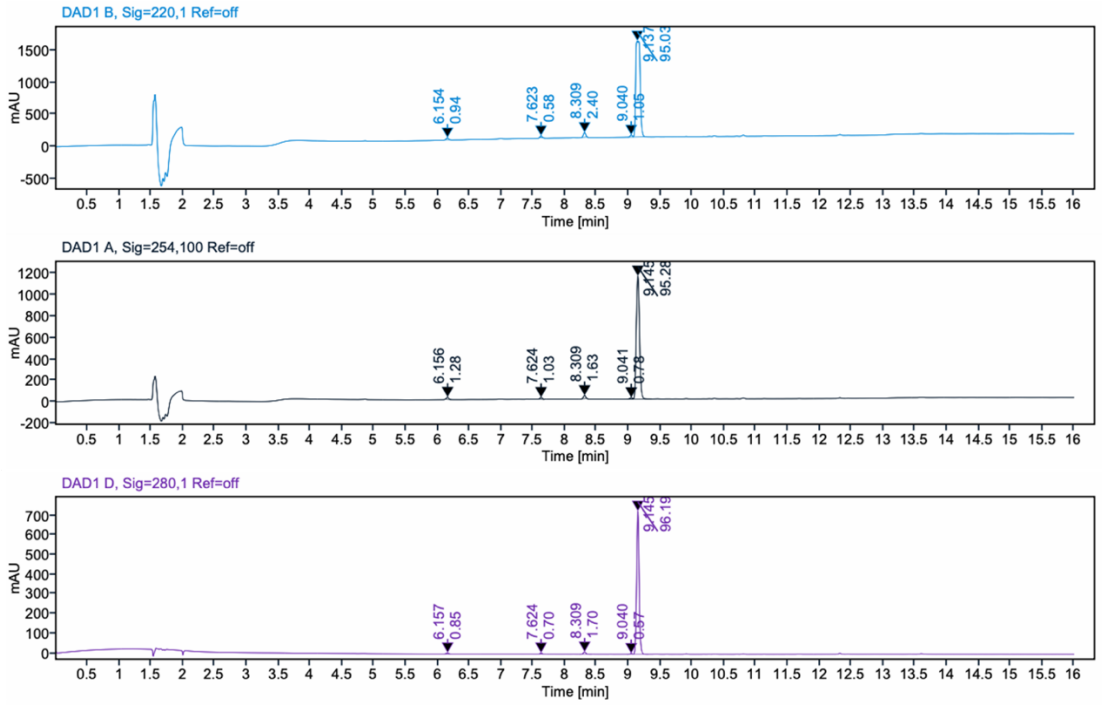


## Appendix 11 HPLC Chromatograms for Compounds 30, 58–63, and 67

### (30) 3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol5-yl)sulfamoyl)benzoic acid



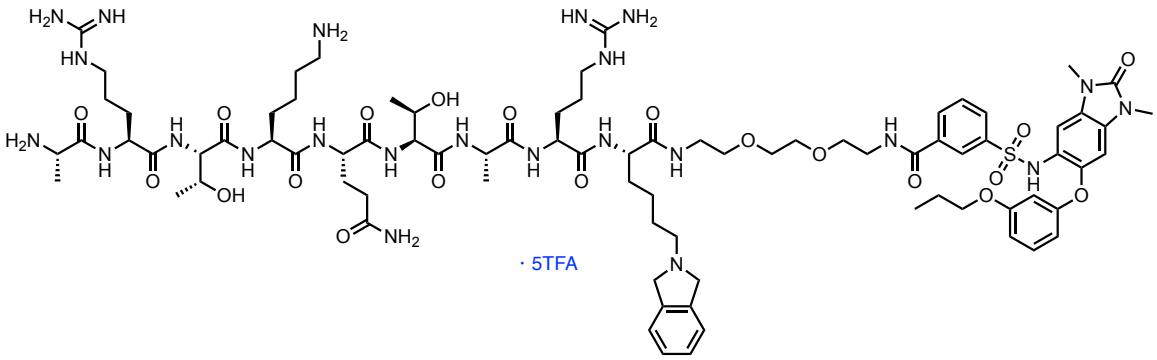


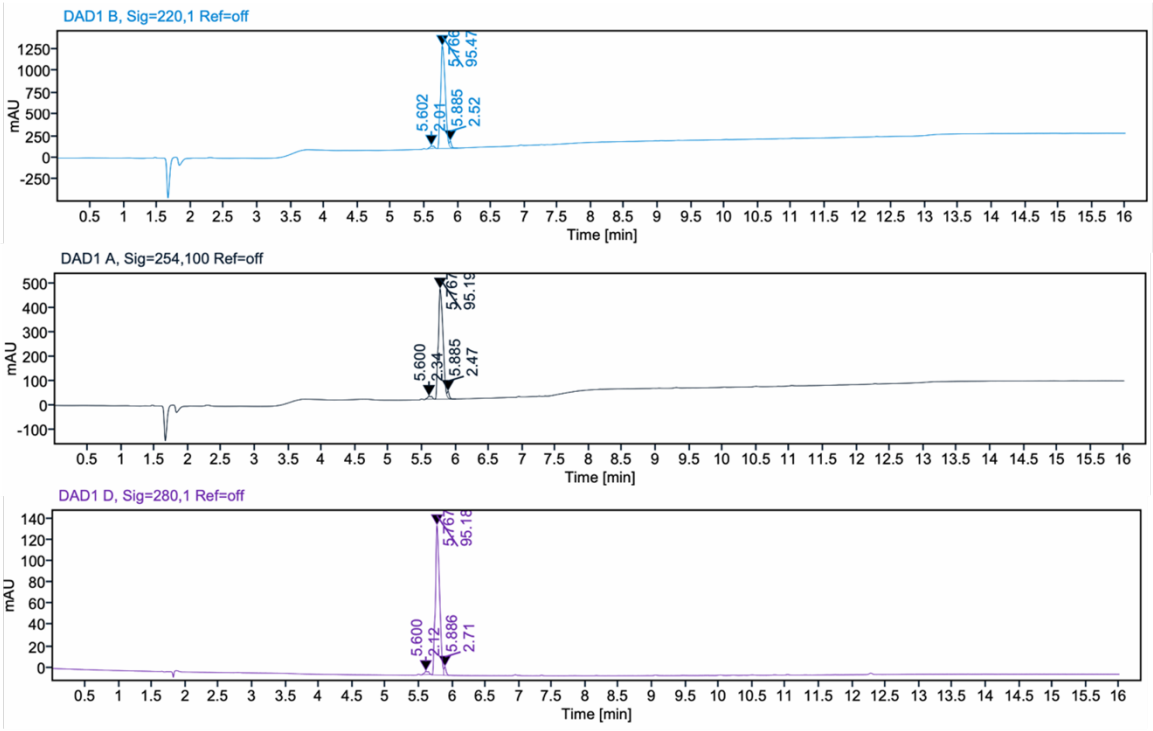

| Signal: | DAD1 B, Sig=220,1 Ref=off | | | | |
|---|---|---|---|---|---|
| RT [min] | Type | Width [min] | Area | Height | Area% |
| 6.154 | MM | 0.0406 | 70.7560 | 29.0731 | 0.9388 |
| 7.623 | MM | 0.0285 | 44.0723 | 25.7582 | 0.5847 |
| 8.309 | MM | 0.0412 | 180.7276 | 73.1441 | 2.3979 |
| 9.040 | MF | 0.0412 | 78.9560 | 31.9489 | 1.0476 |
| 9.137 | FM | 0.0804 | 7162.4551 | 1484.1761 | 95.0310 |
| | | Sum | 7536.9670 | | |

| Signal: | DAD1 A, Sig=254,100 Ref=off | | | | |
|---|---|---|---|---|---|
| RT [min] | Type | Width [min] | Area | Height | Area% |
| 6.156 | MM | 0.0437 | 53.6648 | 20.4808 | 1.2820 |
| 7.624 | MM | 0.0423 | 42.9388 | 16.9129 | 1.0257 |
| 8.309 | MM | 0.0395 | 68.3300 | 28.8032 | 1.6323 |
| 9.041 | BV E | 0.0384 | 32.6434 | 13.1872 | 0.7798 |
| 9.145 | VB R | 0.0577 | 3988.5232 | 1142.4351 | 95.2802 |
| | | Sum | 4186.1003 | | |

| Signal: | DAD1 D, Sig=280,1 Ref=off | | | | |
|---|---|---|---|---|---|
| RT [min] | Type | Width [min] | Area | Height | Area% |
| 6.157 | MM | 0.0486 | 17.3865 | 5.9681 | 0.8485 |
| 7.624 | MM | 0.0377 | 14.3446 | 6.3383 | 0.7001 |
| 8.309 | BB | 0.0445 | 34.8025 | 11.9448 | 1.6985 |
| 9.040 | MF | 0.0425 | 11.5964 | 4.5505 | 0.5660 |
| 9.145 | FM | 0.0452 | 1970.8463 | 727.3156 | 96.1869 |
| | | Sum | 2048.9763 | | |

## (58) PDC4




**Signal:** DAD1 B, Sig=220,1 Ref=off

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.602 | MM | 0.0657 | 120.3803 | 30.5181 | 2.0114 |
| 5.766 | BV R | 0.0643 | 5713.6802 | 1173.7549 | 95.4690 |
| 5.885 | VV E | 0.0312 | 150.7956 | 69.7947 | 2.5196 |
| | | **Sum** | 5984.8561 | | |

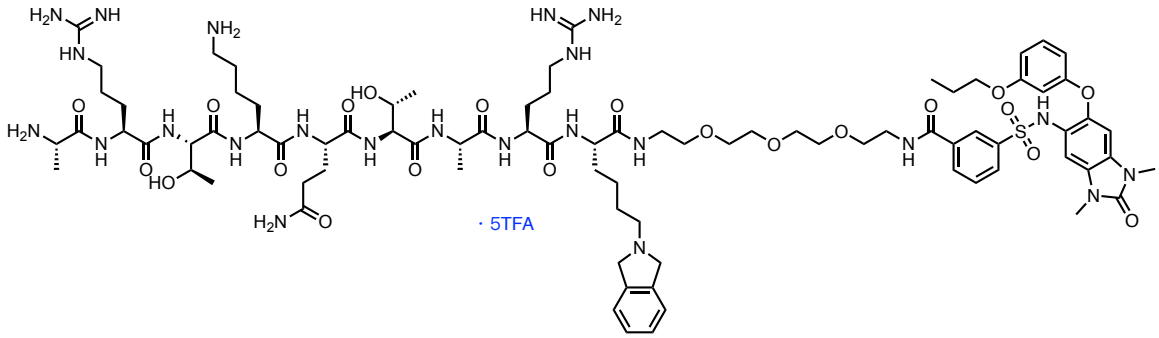
**Signal:** DAD1 A, Sig=254,100 Ref=off

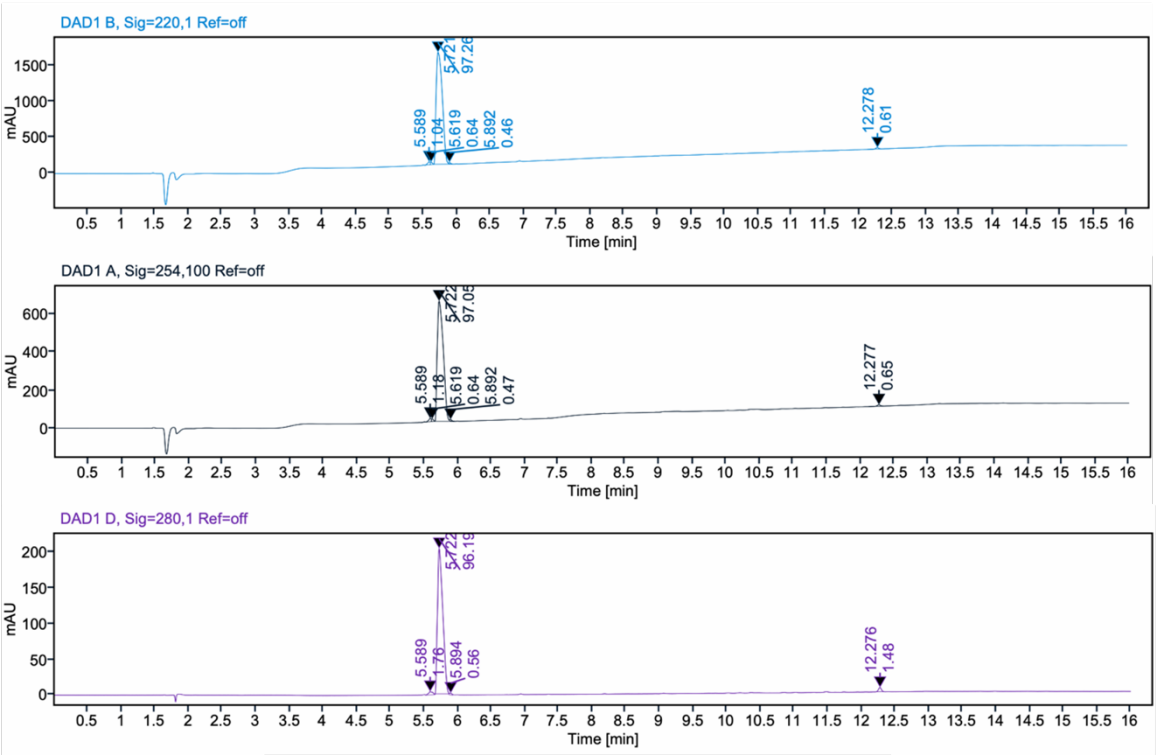
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.600 | MM | 0.0691 | 54.6659 | 13.1805 | 2.3436 |
| 5.767 | BV R | 0.0715 | 2220.3271 | 454.4994 | 95.1893 |
| 5.885 | VB E | 0.0310 | 57.5465 | 27.4085 | 2.4671 |
| | | **Sum** | 2332.5395 | | |

**Signal:** DAD1 D, Sig=280,1 Ref=off

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.600 | MM | 0.0695 | 13.8088 | 3.3111 | 2.1174 |
| 5.767 | MF | 0.0737 | 620.7062 | 140.3202 | 95.1756 |
| 5.886 | FM | 0.0363 | 17.6545 | 8.1163 | 2.7070 |
| | | **Sum** | 652.1695 | | |

## (59) PDC5




**Signal:** DAD1 B, Sig=220,1 Ref=off

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.589 | BV | 0.0345 | 105.0308 | 43.7189 | 1.0352 |
| 5.619 | VB | 0.0280 | 64.4301 | 34.9734 | 0.6350 |
| 5.721 | BV R | 0.0865 | 9867.9688 | 1546.1902 | 97.2595 |
| 5.892 | VB E | 0.0307 | 47.1364 | 22.2657 | 0.4646 |
| 12.278 | BB | 0.0438 | 61.4499 | 18.2534 | 0.6057 |
| | | **Sum** | 10146.0158 | | |

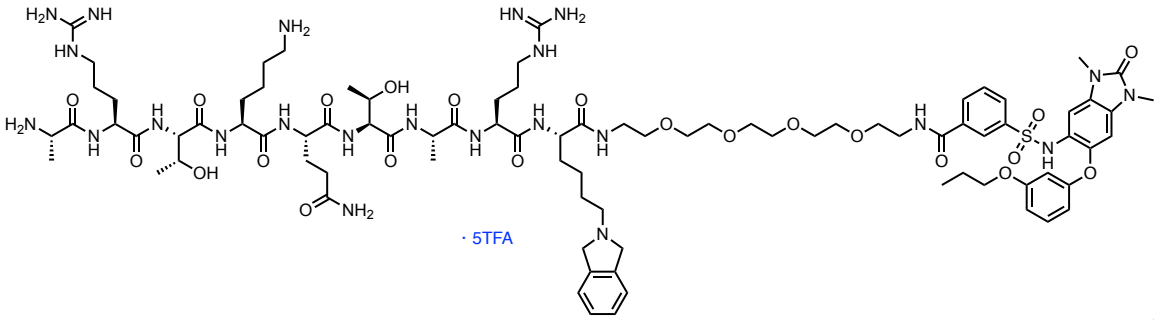
**Signal:** DAD1 A, Sig=254,100 Ref=off

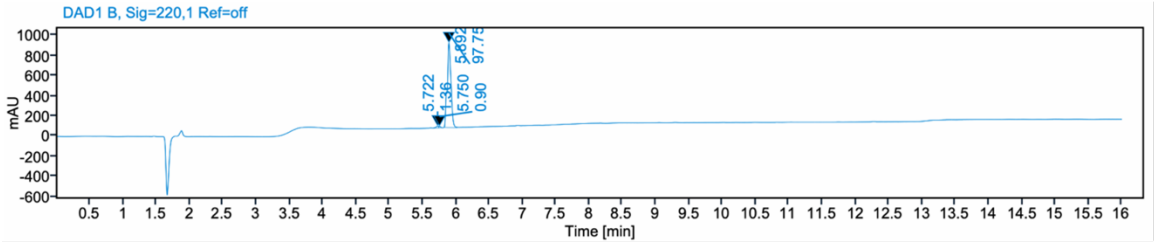
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.589 | BV | 0.0370 | 47.9056 | 18.0283 | 1.1808 |
| 5.619 | VB | 0.0277 | 26.0846 | 14.3393 | 0.6430 |
| 5.722 | BV R | 0.0843 | 3937.4248 | 635.6541 | 97.0543 |
| 5.892 | VB E | 0.0305 | 19.1095 | 9.0831 | 0.4710 |
| 12.277 | BB | 0.0485 | 26.4047 | 7.9953 | 0.6509 |

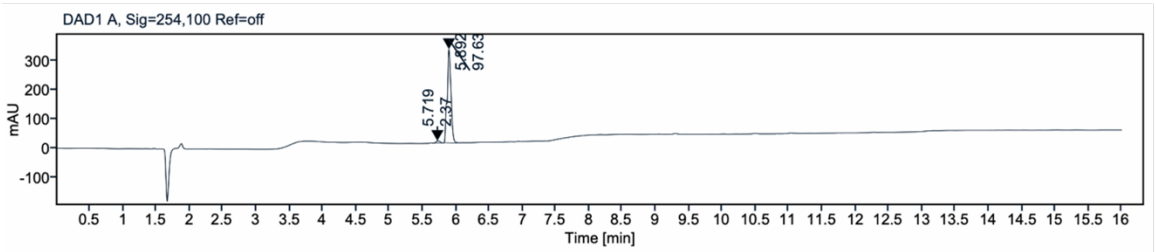
**Signal:** DAD1 D, Sig=280,1 Ref=off

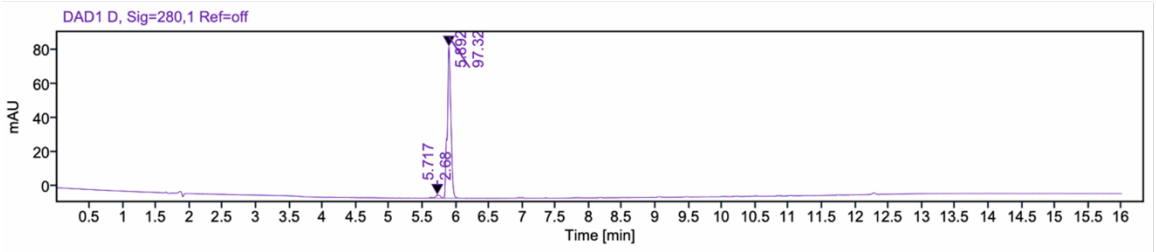
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.589 | MM | 0.0721 | 20.2840 | 4.6920 | 1.7611 |
| 5.722 | MM | 0.0911 | 1107.9709 | 202.6207 | 96.1943 |
| 5.894 | MM | 0.0413 | 6.4897 | 2.6173 | 0.5634 |
| 12.276 | MM | 0.0482 | 17.0600 | 5.8986 | 1.4812 |
| | | **Sum** | 1151.8047 | | |

## (60) PDC6




**Signal:** DAD1 B, Sig=220,1 Ref=off

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.722 | BV | 0.0302 | 45.4820 | 18.4662 | 1.3582 |
| 5.750 | VB | 0.0244 | 30.0065 | 15.3082 | 0.8961 |
| 5.892 | MM | 0.0650 | 3273.1433 | 838.9249 | 97.7457 |
| | | **Sum** | 3348.6317 | | |

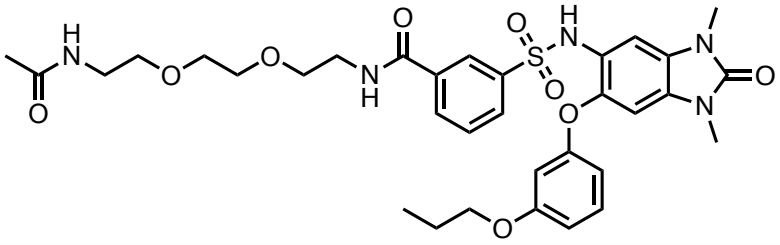
**Signal:** DAD1 A, Sig=254,100 Ref=off

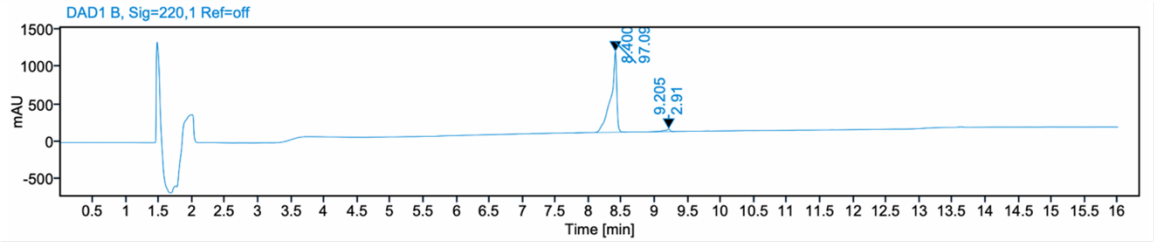
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.719 | BV | 0.0561 | 31.2783 | 7.5136 | 2.3734 |
| 5.892 | MM | 0.0661 | 1286.6101 | 324.3882 | 97.6266 |
| | | **Sum** | 1317.8884 | | |

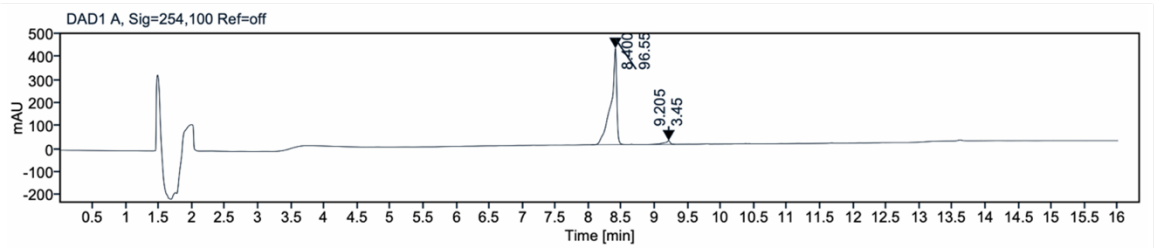
**Signal:** DAD1 D, Sig=280,1 Ref=off

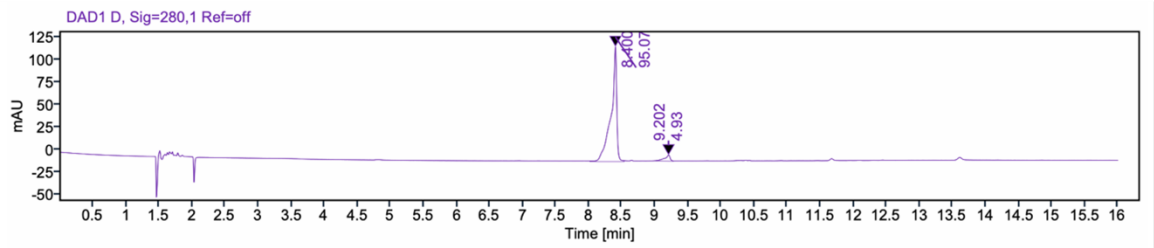
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 5.717 | MM | 0.0789 | 9.2368 | 1.9501 | 2.6835 |
| 5.892 | MM | 0.0628 | 334.9674 | 88.9557 | 97.3165 |
| | | **Sum** | 344.2042 | | |

**(61) 3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2-oxo-6,9-dioxa-3-azaundecan-11-yl)benzamide**






**Signal:** DAD1 B, Sig=220,1 Ref=off

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 8.400 | VB R | 0.0845 | 6792.0425 | 1070.2261 | 97.0871 |
| 9.205 | BB | 0.0803 | 203.7801 | 32.6387 | 2.9129 |
| | | **Sum** | 6995.8226 | | |

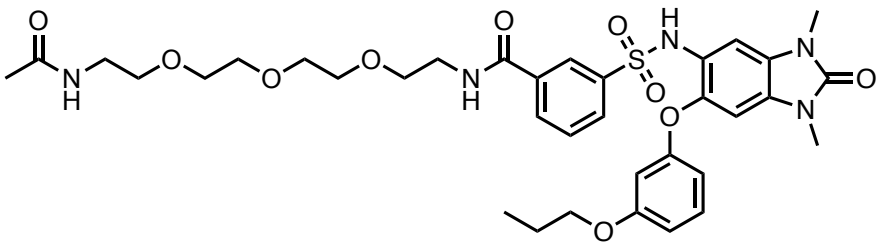
**Signal:** DAD1 A, Sig=254,100 Ref=off

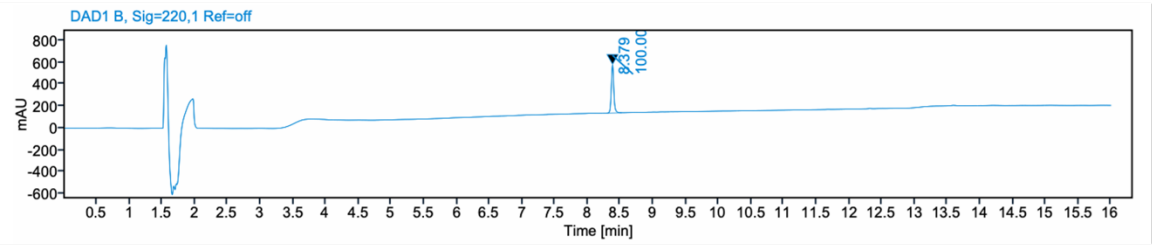
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 8.400 | BB | 0.0827 | 2586.3210 | 417.8310 | 96.5452 |
| 9.205 | BB | 0.0807 | 92.5486 | 15.4814 | 3.4548 |
| | | **Sum** | 2678.8696 | | |

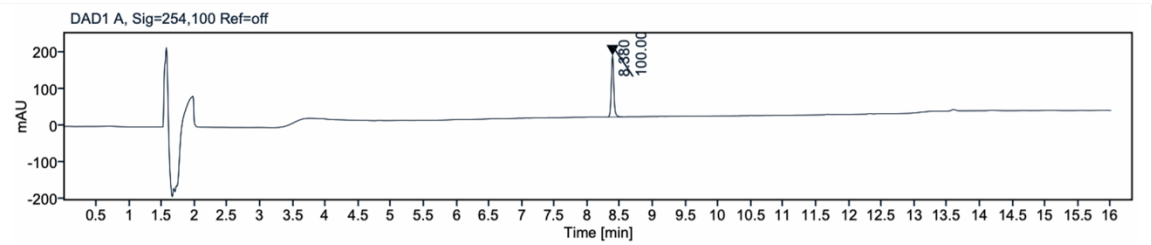
**Signal:** DAD1 D, Sig=280,1 Ref=off

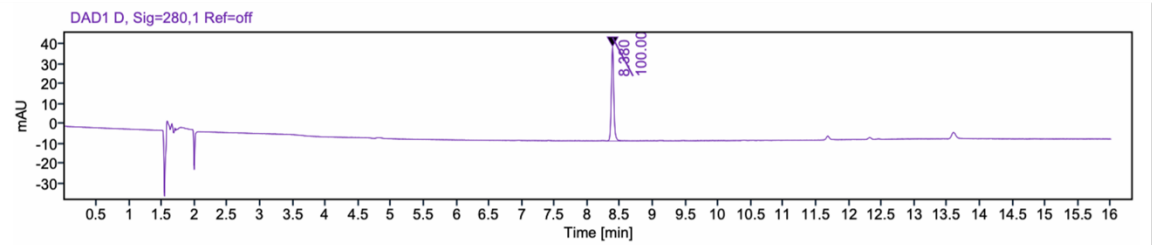
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 8.400 | MM | 0.0995 | 765.6229 | 128.2305 | 95.0712 |
| 9.202 | MM | 0.0974 | 39.6920 | 6.7944 | 4.9288 |
| | | **Sum** | 805.3148 | | |

**(62) 3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2-oxo-6,9,12-trioxa-3-azatetradecan-14-yl)benzamide**






**Signal:** DAD1 B, Sig=220,1 Ref=off

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 8.379 | BB | 0.0403 | 1159.1935 | 438.8969 | 100.0000 |
| | | **Sum** | 1159.1935 | | |

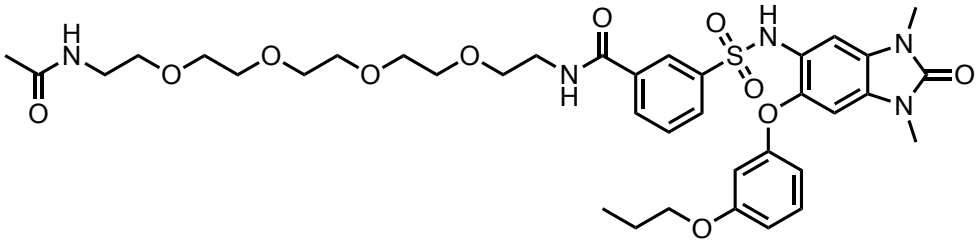
**Signal:** DAD1 A, Sig=254,100 Ref=off

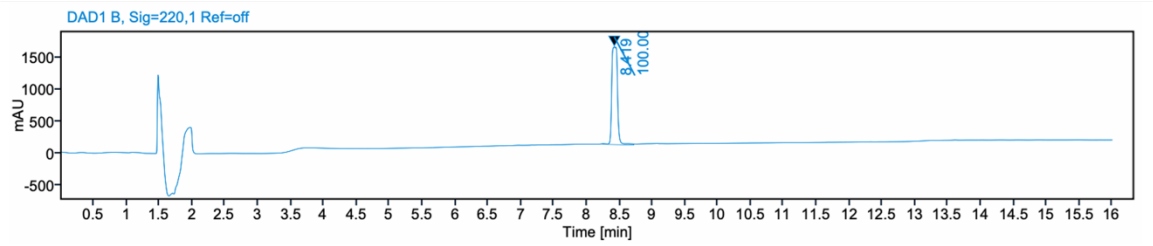
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 8.380 | BB | 0.0403 | 441.4357 | 167.2781 | 100.0000 |
| | | **Sum** | 441.4357 | | |

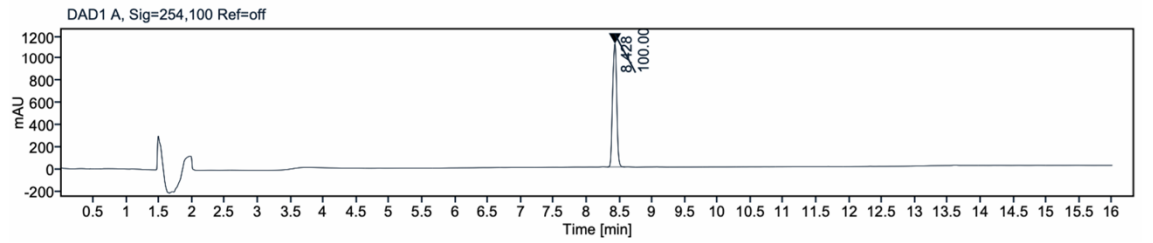
**Signal:** DAD1 D, Sig=280,1 Ref=off

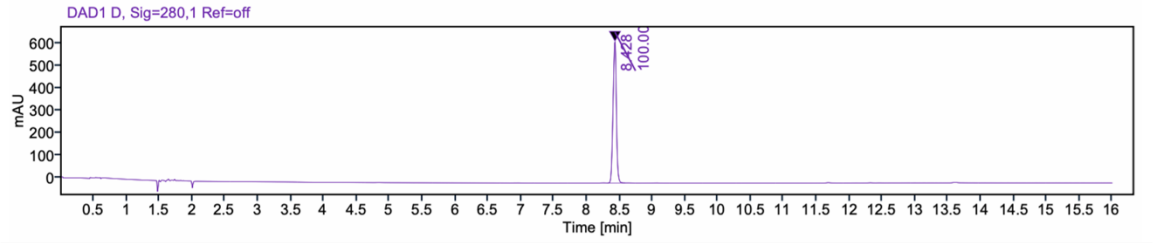
| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 8.380 | BB | 0.0406 | 123.2187 | 46.9504 | 100.0000 |
| | | **Sum** | 123.2187 | | |

**(63) 3-(*N*-(1,3-dimethyl-2-oxo-6-(3-propoxyphenoxy)-2,3-dihydro-1*H*-benzo[*d*]imidazol-5-yl)sulfamoyl)-*N*-(2-oxo-6,9,12,15-tetraoxa-3-azaheptadecan-17-yl)benzamide**





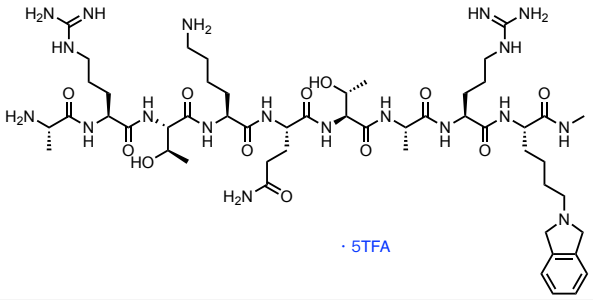



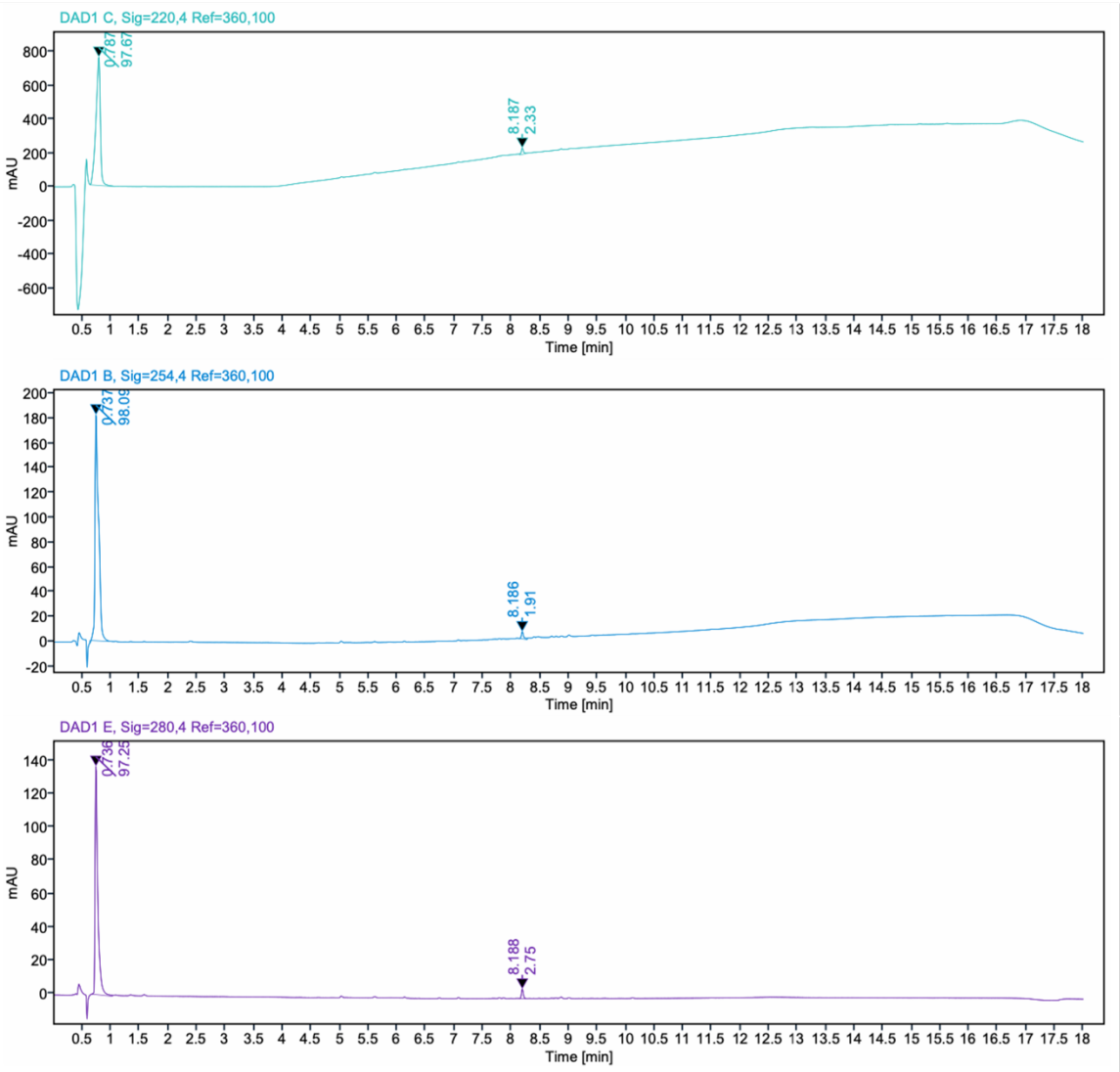

| Signal: | DAD1 B, Sig=220,1 Ref=off | | | | |
|---|---|---|---|---|---|
| RT [min] | Type | Width [min] | Area | Height | Area% |
| 8.419 | MM | 0.0999 | 9196.1250 | 1534.9658 | 100.0000 |
| | | Sum | 9196.1250 | | |

| Signal: | DAD1 A, Sig=254,100 Ref=off | | | | |
|---|---|---|---|---|---|
| RT [min] | Type | Width [min] | Area | Height | Area% |
| 8.428 | BB | 0.0740 | 4885.5239 | 1116.9390 | 100.0000 |
| | | Sum | 4885.5239 | | |

| Signal: | DAD1 D, Sig=280,1 Ref=off | | | | |
|---|---|---|---|---|---|
| RT [min] | Type | Width [min] | Area | Height | Area% |
| 8.428 | BB | 0.0560 | 2204.2925 | 626.8787 | 100.0000 |
| | | Sum | 2204.2925 | | |

## (67) PHD Control Peptide




**Signal:** DAD1 C, Sig=220,4 Ref=360,100

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 0.787 | BB | 0.0757 | 4056.1399 | 761.1058 | 97.6717 |
| 8.187 | MM | 0.0429 | 96.6887 | 37.5530 | 2.3283 |
| | | **Sum** | 4152.8286 | | |

**Signal:** DAD1 B, Sig=254,4 Ref=360,100

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 0.737 | MM | 0.0727 | 794.1224 | 181.9515 | 98.0888 |
| 8.186 | MM | 0.0455 | 15.4734 | 5.6685 | 1.9112 |
| | | **Sum** | 809.5958 | | |

**Signal:** DAD1 E, Sig=280,4 Ref=360,100

| RT [min] | Type | Width [min] | Area | Height | Area% |
|---|---|---|---|---|---|
| 0.736 | MM | 0.0556 | 456.8456 | 136.9389 | 97.2533 |
| 8.188 | MM | 0.0371 | 12.9026 | 5.7904 | 2.7467 |
| | | **Sum** | 469.7482 | | |

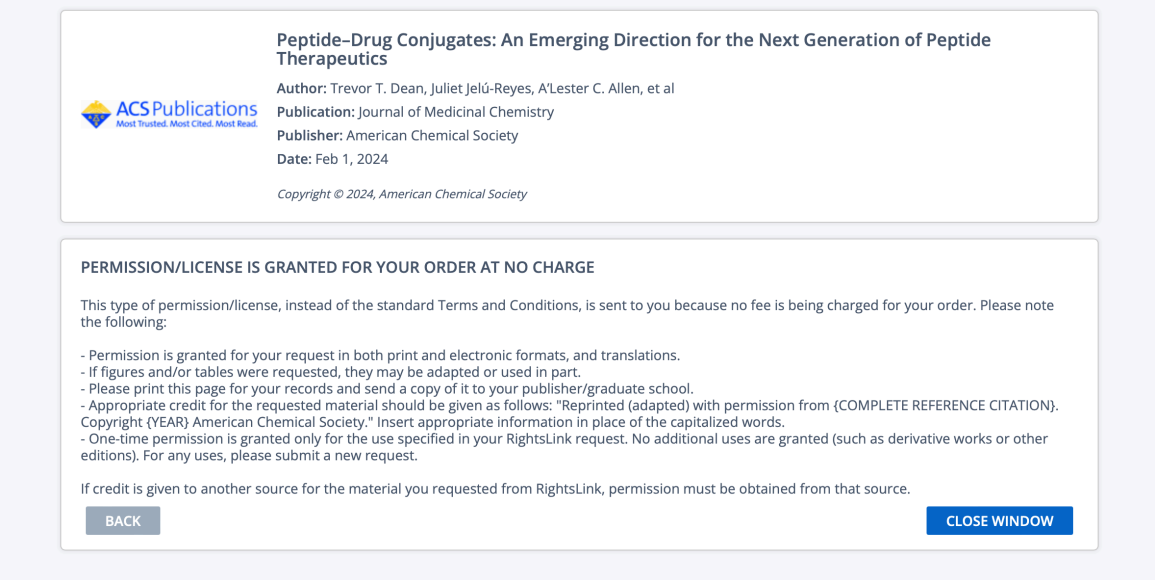
## Appendix 12 Image Permissions

| Figure | License |
|---|---|
| 1.1 | Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/) |
| 1.2 | No license required (see below) |
| 1.3 | No license required (see below) |
| 1.5 | Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 United States License (https://creativecommons.org/licenses/by-nc-nd/3.0/) |
| 1.6 | License No. 5873601288425 (Elsevier and Copyright Clearance Centre) |
| 1.8 | License No. 5873070598570 (Elsevier and Copyright Clearance Centre) |
| 1.9 | Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/) |
| 1.11 | License No. 5873730882225 (Elsevier and Copyright Clearance Centre) |
| 1.12 | Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/) |
| 1.15 | License No. 5876021407542 (Springer Nature and Copyright Clearance Centre) |
| 1.17 | No license required (see below) |
| 1.20 | License No. 1533678-1 (Portland Press and Copyright Clearance Centre) https://marketplace.copyright.com/rs-ui-web/mp/license/c934a7ea-26f5-4e95-bfa5-aeaf60148397/4108daf3-e40b-4eda-a71f-630107b1c9a6 |
| 2.10 | Creative Commons Attribution 3.0 Unported License (https://creativecommons.org/licenses/by/3.0/) |

### Figure 1.2 & 1.3



## Figure 1.17



Peptide–Drug Conjugates: An Emerging Direction for the Next Generation of Peptide Therapeutics

**Author:** Trevor T. Dean, Juliet Jelú-Reyes, A'Lester C. Allen, et al
**Publication:** Journal of Medicinal Chemistry
**Publisher:** American Chemical Society
**Date:** Feb 1, 2024

*Copyright © 2024, American Chemical Society*

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK CLOSE WINDOW