

## Durham E-Theses

---

# *Investigation of a Histidine-Based Probe for the Exploration of Proteomes*

MICHAELA LOUISE BUERDSELL

### How to cite:

---

BUERDSELL, MICHAELA LOUISE (2022) Investigation of a Histidine-Based Probe for the Exploration of Proteomes. Doctoral thesis, Durham University.

### Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a <https://etheses.durham.ac.uk/id/eprint/14389/> is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

# Investigation of a Histidine-Based Probe for the Exploration of Proteomes



*A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy  
from Durham University*

Michaela Louise Buerdsell

Department of Chemistry

November 2021

## Abstract

Leishmaniasis is a neglected tropical disease which affects 0.7-1 million people per year. Current chemotherapies for leishmaniasis are toxic with long treatment times and reports of increasing resistance, which stresses the importance of this research area. Inositol phosphorylceramide synthase is a membrane bound enzyme that has no direct human homologue, which converts ceramide to inositol phosphorylceramide through the action of a highly conserved HHD catalytic triad. An ideal method to study this enzyme further would be through activity-based protein profiling, however, there are currently no activity-based probes reported that reacts with this type of active site. Therefore, an activity-based probe was designed based on the structure of diethyl pyrocarbonate, a compound known to bind covalently to active site histidine residues. The synthesised activity-based probe was shown to inhibit *Leishmania major* inositol phosphorylceramide synthase in a simple assay. In addition, the probe was shown to selectively bind to the active site histidine residue in two pure enzyme models; one of which has the same catalytic triad as inositol phosphorylceramide synthase, and the other was an acid base active site histidine residue. Further, this activity-based probe was able to isolate an overexpressed enzyme in the lysate of *Escherichia coli* as well as bind to intrinsic proteins. Following the function validation of the activity-based probe, preliminary work was started in *Leishmania* to isolate proteins identify expressed enzymes.

## Acknowledgements

For the duration of my PhD, I have received the support of many individuals who have helped me to achieve my goal. Therefore, it is my pleasure to take the time and express my gratitude. Firstly, I would like to thank my primary supervisor Professor Patrick Steel for his continued guidance throughout my PhD and remaining patient when I was initially learning organic chemistry. I would also like to express gratitude to my secondary supervisor Dr Paul Denny who provided a friendly accommodating environment in the biology lab CG229. I would also like to thank The Royal Society for funding my research which without, I would be unable to complete my dream of achieving a doctorate.

I would also like to thank some individuals who have had a direct impact to my thesis. Firstly, Dr Adrian Brown who I have worked closely with in my last few years of my PhD. You have provided amazing mass spectrometry analyses and gave great advice for increasing the purity of my samples and increasing the strength of my analyses. Secondly, to Dr John Mina who helped me during my first year of my PhD with the analysis of *Leishmania major* IPCS assay in addition to providing the microsomes used to generate these results.

I thank Dr Abraham Alvarez who has continuously supported me throughout my PhD. During my time in the PGS group, I have met many individuals past and present who have brought joy to the laboratory and made working in CG001 a pleasure. I would like to individually thank Dr Katherine Norman, Dr Vanessa Lyne, Dr Exequiel Porta, Dr Jonathan Reuven and Victor Agostino. You have all provided great scientific feedback as well as continued support, encouragement and provided so many great memories.

Finally, I would like to thank my family. Whilst you lived far and seeing you was difficult, you have always kept in regular contact and made Christmas so special.

## Declaration

The work presented in this thesis was carried out in the Department of Chemistry at Durham University between September 2017 and November 2021. All work is the author's own except for collaborative which is acknowledged where appropriate. No part of this work has been submitted for any other degree at this, or any other University.

## Statement of Copyright

The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.

## Table of Contents

1	Introduction .....	1
1.1	Leishmaniasis.....	1
1.1.1	Clinical manifestations .....	1
1.1.2	Epidemiology of leishmaniasis .....	3
1.1.3	Lifecycle.....	5
1.1.4	Treatment.....	6
1.2	Sphingolipids .....	11
1.2.1	Sphingolipid synthesis.....	11
1.2.2	Inositol phosphorylceramide synthase (IPCS).....	13
1.3	Concluding remarks.....	16
1.4	Activity based protein profiling (ABPP) .....	16
1.4.1	Warhead.....	17
1.4.2	Spacer and binding group .....	19
1.4.3	Reporter tag .....	20
1.4.4	Fishing for targets .....	22
1.5	Application of ABPP .....	23
1.5.1	Serine hydrolases .....	24
1.5.2	Cysteine proteases .....	25
1.5.3	Lysine.....	26
1.6	Histidine activity in enzymes .....	26
1.6.1	Non-specific acid phosphatase .....	27
1.6.2	Ribonuclease A.....	30
1.6.3	$\alpha$ -amino- $\beta$ -Carboxymuconate- $\epsilon$ -Semialdehyde Decarboxylase .....	31
1.7	Histidine modifying reagents .....	33
1.7.1	Diethyl pyrocarbonate (DEPC) .....	34
1.8	Project aims .....	35
2	Design of a histidine specific probe .....	36
2.1.1	Synopsis of chapter .....	36
2.2	Initial probe design.....	36
2.3	A new probe design.....	41
2.3.1	BHDC imidazole reactivity.....	42

2.4	Stability of the carbamate .....	46
2.4.1	Copper-catalyzed azide-alkyne cycloaddition (CuAAC) .....	47
2.5	Alternative dipeptide synthesis.....	49
2.6	Fluorescent marker and binding group.....	52
2.6.1	Simple rhodamine fluorophore.....	53
2.6.2	Tetrafunctional reactive linker.....	54
3	Biological analysis of ABP in a model system.....	57
3.1.1	Synopsis of chapter .....	57
3.2	<i>Lmj</i> IPCS inhibition assay .....	57
3.2.1	Initial test for inhibition .....	60
3.2.2	Assessing the role of lipophilicity for inhibition.....	61
3.3	Non-specific acid phosphatase inhibition .....	63
3.3.2	Non-specific acid phosphatase activity.....	73
3.3.3	BHDC inhibition of pNP production .....	74
3.4	PhoN mass spectrometry analysis.....	74
3.5	Fluorescent labelling of PhoN .....	79
3.5.1	Pure protein labelling.....	80
3.5.2	Lysate protein labelling .....	82
3.5.3	Competitive inhibition .....	83
3.6	Reactivity in an acid-base active site.....	85
3.7	Pulldown purification .....	91
3.7.1	Pulldown protein composition.....	100
3.7.2	Summary .....	108
3.8	Chapter summary .....	108
4	Biological testing of ABP in <i>Leishmania</i> .....	109
4.1.1	Synopsis of chapter .....	109
4.2	Probe application to <i>L. mexicana</i> lysate.....	109
4.3	Promastigote inhibition assay .....	112
4.4	Exploration of the <i>Leishmania</i> proteome.....	113
4.4.1	Pulldown experiment.....	115
4.5	Chapter summary .....	116
5	Conclusions and Future work.....	117

5.1	Initial hypothesis .....	117
5.2	Conclusion .....	117
5.3	Future work .....	118
5.4	Broader application .....	121
6	Biology methodology .....	122
6.1	General equipment .....	122
6.2	Buffers and Medias.....	122
6.3	Microsomal assay materials .....	125
6.3.1	<i>Lmj</i> IPCS inhibition assay methodology .....	125
6.4	Promastigote inhibition assay materials .....	126
6.4.1	Promastigote inhibition assay method .....	126
6.5	Recombinant <i>E. coli</i> materials .....	127
6.6	Production of protein from <i>E. coli</i> .....	129
6.6.1	Transformation .....	129
6.6.2	Ligation into pOPINF .....	129
6.6.3	Sanger sequence .....	130
6.6.4	Overexpression and lysis.....	130
6.6.5	Protein purification .....	130
6.7	Non-specific acid phosphatase activity <i>materials</i> .....	131
6.7.1	NSAP assay method .....	131
6.8	Fluorescent and pull-down analysis .....	131
6.8.1	Materials .....	131
6.8.2	Fluorescent analysis methods.....	132
6.8.3	Pulldown assay .....	133
6.9	Proteolytic digestion.....	134
7	Chemical Experimental .....	135
7.1	Equipment .....	135
7.2	General procedure A: Formation of Dicarbonates.....	135
7.3	General procedure B: Carbamate Preparation .....	136
7.4	General procedure C: Formation of 1,2,3-triazole.....	136
7.5	General procedure D: Chalcone Preparation .....	136
8	Appendix 1 .....	156

9	Appendix 2 .....	158
10	Appendix 3 .....	162
11	References.....	193

## Abbreviations

**βME:** β-mercaptoethanol

**ABP:** Activity based probe

**ABPP:** Activity-based protein profiling

**ACMS:** α-amino-β-carboxymuconate-ε-semialdehyde

**ACMSD:** α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase

**ADCL:** anergic diffuse cutaneous leishmaniasis

**AIDS:** Acquired immunodeficiency virus

**ASAP:** Atmospheric solid analysis probe

**ATP:** Adenosine triphosphate

**BCA:** Bicinchoninic acid

**BHDC:** Bis(hex-5-yn-1-yl) dicarbonate

**BSA:** Bovine serum albumin

**CDPK1:** Calcium-dependent protein kinase 1

**CERT:** Ceramide transfer protein

**CHAPS:** 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate.

**CL:** Cutaneous leishmaniasis

**CuAAC:** Copper-Catalyzed Azide-Alkyne Cycloaddition

**DAG:** Diacylglycerol

**DALY:** disability adjusted life years

**DBU:** 1,8-Diazabicyclo[5.4.0]undec-7-ene

**DCL:** Diffuse cutaneous leishmaniasis

**DCM:** Dichloromethane

**DEC:** Diethyl carbonate

**DEPC:** Diethyl pyrocarbonate

**DHSph:** Dihydrosphingosine

**DIPEA:** N,N-Diisopropylethylamine

**DMAP:** 4-Dimethylaminopyridine

**DNA:** deoxyribonucleic acid

**DPC:** Dipropargyl carbonate

**DPPC:** Dipropargyl pyrocarbonate

**DTT:** Dithiothreitol

***E. coli:*** *Escherichia coli*

**EDC:** 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

**EDTA:** Ethylenediaminetetraacetic acid

**ER:** Endoplasmic reticulum

**FP:** Fluorophosphonates

**GC-MS:** Gas chromatography mass spectrometry

**HEPES:** 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

**HIV:** Human immunodeficiency virus

**HK:** Histidine kinase

**HPTLC:** High performance thin layer chromatography

**Hz:** Hertz

**IPC:** Inositol phosphorylceramide

**IPCS:** Inositol phosphorylceramide synthase

**IPTG:** Isopropyl  $\beta$ -D-1-thiogalactopyranoside

**KCl:** Potassium chloride

**LC-MS:** Liquid chromatography mass spectrometry

**LADH:** Liver alcohol dehydrogenase

***L. major:*** *Leishmania major*

***Lmj*IPCS:** *Leishmania* inositol phosphorylceramide synthase

**MALDI:** Matrix-assisted laser desorption/ionization

**MBP:** Membrane bound protein

**MCL:** Mucocutaneous leishmaniasis

**Na asc:** Sodium ascorbate

**NBD:** nitrobenzoxadiazole

**NEt<sub>3</sub>:** Triethylamine

**NGO:** Non-government organisation

**NHS:** N-hydroxysuccinimide

**NMR:** Nuclear magnetic resonance

**NSAP:** non-specific acid phosphatase

**NTD:** Neglected tropical disease

**PBS:** Phosphate-buffered saline

**PC:** Phosphatidylcholine

**PEG:** Polyethylene glycol

**PI:** Phosphatidylinositol

**PI3K:** Phosphoinositide 3-kinase

**PKDL:** Post-kala-azar dermal leishmaniasis

**pNP:** Para-nitrophenyl

**pNPP:** Para-nitrophenyl phosphate

**PPM:** Parts per million

**RNA:** Ribonucleic acid

**RNase:** Ribonuclease

**RR:** Response regulator

**RT:** Room temperature

**SAR:** Structure-activity relationship

**SDS-PAGE:** Sodium dodecyl sulfate polyacrylamide gel electrophoresis

**SL:** Sphingolipid

**SM:** Sphingomyelin

**SMS:** Sphingomyelin synthase

**SPT:** Serine palmitotransferase

**TBAB:** Tetrabutylammonium bromide

**TBDD:** Target based drug discovery

**TBTA:** Tris(benzyltriazolylmethyl)amine

**TCS:** Two component system

**TFA:** Trifluoroacetic acid

**THF:** Tetrahydrofuran

**THPTA:** tris-  
hydroxypropyltriazolylmethylamine

**TLC:** Thin layer chromatography

**UV:** Ultraviolet

**VL:** Visceral leishmaniasis

**WHO:** World health organisation

## 1 Introduction

The aim of this project was to develop a novel activity-based probe that targets active site histidine residues. Following the synthesis and validation of specificity, the probe was to then be implemented into *Leishmania* to identify enzymes with a related function for characterisation.

For the remainder of this chapter, leishmaniasis will be discussed with a focus on the sphingolipid pathway. Then, how activity-based protein profiling can be utilised for the advancement of proteome knowledge. Chapter 2 discusses the synthetic development of the activity-based probe, followed by Chapter 3 which validates the probes activity on enzyme models. Chapter 4 evaluates the probes implementation in *Leishmania* with Chapter 5 and Chapter 6 detailing the details the chemical and biological methodology respectively.

### 1.1 Leishmaniasis

#### 1.1.1 Clinical manifestations

Leishmaniasis is a neglected tropical disease caused by over 20 different species of leishmanial protozoans (Figure 1). <sup>1</sup> Approximately 700,000 to 1.2 million new cases of leishmaniasis are diagnosed annually and leaves 350 million people at risk. <sup>2,3</sup> Using the measure of disease burden, leishmaniasis was calculated to have 2.4 million disability adjusted life years (DALYs), which is the second highest of the parasitic diseases. <sup>4</sup> There are several different clinical manifestations of leishmaniasis: cutaneous leishmaniasis (CL), anergic diffuse cutaneous leishmaniasis (ADCL), disseminated leishmaniasis (DL), mucocutaneous leishmaniasis (MCL), visceral leishmaniasis (VL) and post-kala-azar dermal leishmaniasis (PKDL). <sup>5</sup> The form of leishmaniasis and outcome depend on the species of *Leishmania* the individual is infected with, and the immune status of the individual. <sup>5,6</sup>

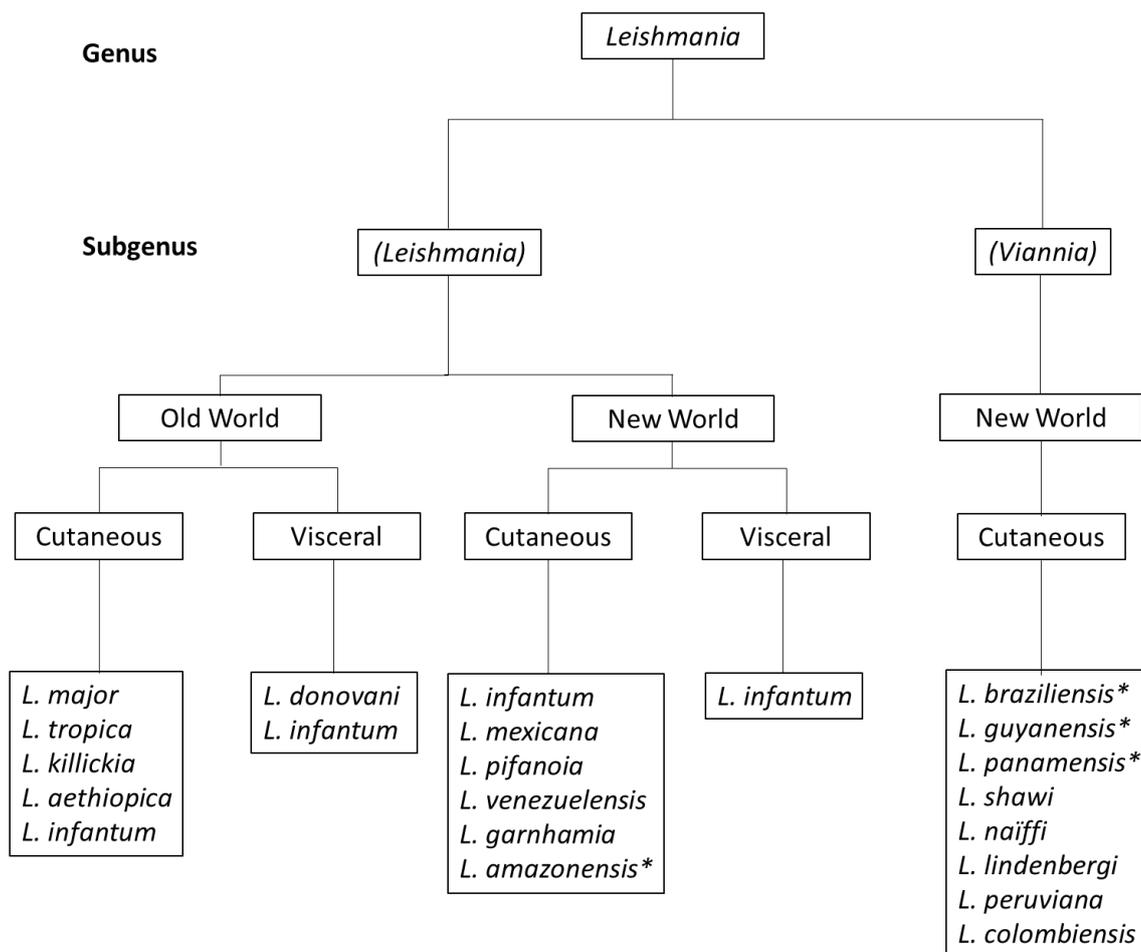


Figure 1 The Genus and Species of *Leishmania*  
Highlighted with \* are the species that can also develop into MCL<sup>7</sup>

CL is the most common form of the disease, with around 0.7- 1 million cases annually, and a further 350 million at risk of developing the disease.<sup>8</sup> CL appears as a lesion(s) on exposed skin, normally appearing as a localised round lesion limited to the cutaneous layer.<sup>9</sup> These lesions can heal spontaneously, however complications can occur, such as secondary bacterial infection within the open lesion, and the potential of the infection reoccurring as ADCL, DL or MCL.<sup>10,11</sup>

ADCL and DL are similar in which they both are an uncommon manifestation in which non ulcerative lesions proliferate to cover the body.<sup>12</sup> In the case of ADCL, lesions are typically nodular or popular in which multiple satellite lesions form around the initial lesion, which can develop into a plaque.<sup>13</sup> This is typically caused by *Leishmania* (*Leishmania*) subspecies such as *L. amazonensis*, *L. mexicana*, *L. aethiopica* and *L. major* with the exception of *L. (V.) pifanoi*.<sup>13</sup> In contrast, DL is normally caused by *Leishmania* (*Viannia*) with the most

common species being *L. braziliensis*, however *L. panamanensis*, *L. guyanensis* and *L. (L.) amazonensis* are also reported. <sup>13</sup> DL presents as a high number of pleomorphic lesions, which rapidly spread from the initial lesion. <sup>14</sup>

MCL is the rarest form of leishmaniasis. This form is often associated with the New World subspecies *Viannia* following CL infection. <sup>9</sup> MCL involves tissue necrosis in the nasal, pharyngeal and laryngeal mucosa, with the most frequent initial symptom being nasal congestion, which progresses to visible destruction. <sup>5,15</sup> Whilst the damage in the discussed tegumentary forms of leishmaniasis initially appears superficial, studies have shown increased depression and suicide in afflicted individuals, especially when the lesion is on exposed regions, such as the face. <sup>16</sup>

Visceral leishmaniasis (VL) is the most severe form the disease, resulting in death when left untreated, and 10-20 % fatality with treatment. <sup>8</sup> VL is caused by either the Old World species *L. donovani* or the New World species *L. infantum*, cumulating to less than 100,000 cases annually. <sup>2</sup> The symptoms of VL typically involve hepatosplenomegaly, fever, weight loss, and anaemia among others, with little differentiation of symptoms between the causative species. <sup>17,18</sup> Diagnosis of leishmaniasis with a coinfection can be difficult due to varied presentation, significantly in regards to HIV. <sup>6</sup>

A complication that can occur post treatment of VL caused by *L. donovani* is post-kala-azar dermal leishmaniasis (PKDL). <sup>19,20</sup> This manifestation first appears on the face in the form of hypopigmentation, papules or nodules. <sup>21,22</sup> Besides the morphological features, the patients are otherwise healthy. <sup>22</sup> PKDL is rarely fatal, however it is still significant because the infected individuals serve as a reservoir, infecting the sandfly vector and allowing the lifecycle to continue. <sup>20,23</sup>

### 1.1.2 Epidemiology of leishmaniasis

*Leishmania* is currently found in 98 countries, however the epidemiology has shown signs of change in recent years. <sup>24</sup> The control strategies for leishmaniasis rely on the control of the vector and reservoir species. <sup>25</sup> A major risk factor for contracting leishmaniasis appears to be poverty, with poor housing, limited sanitation access, and spacial clustering all associated with higher infection rates. <sup>26</sup> Of the endemic countries currently affected by CL,

10 countries contribute to 70-75 % of all CL cases reported (Figure 2), and 90 % of VL cases disproportionately affects just 7 countries (Figure 3).<sup>8,24</sup>

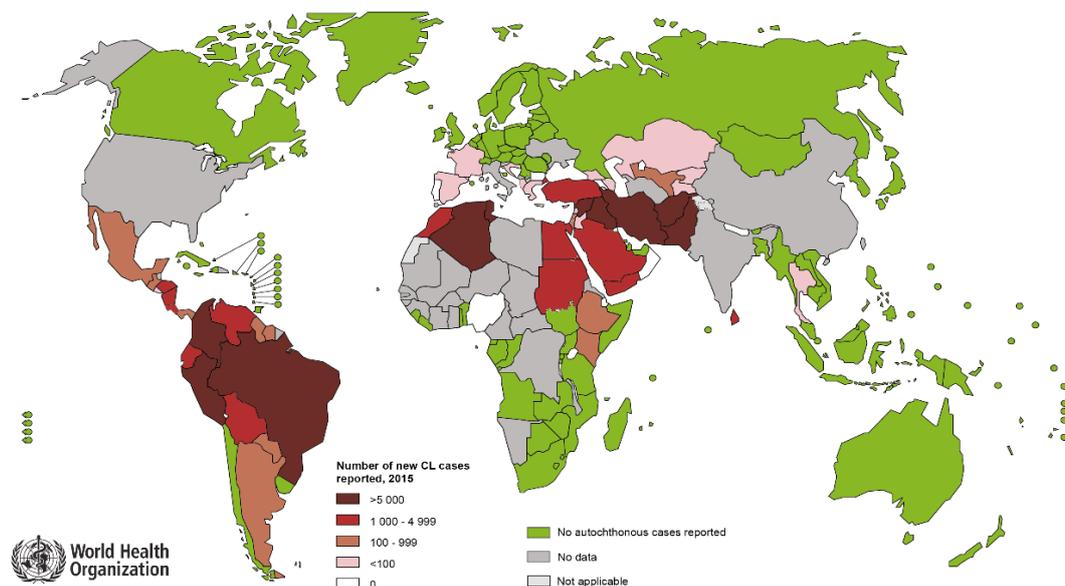


Figure 2 Global distribution of cutaneous leishmaniasis obtained from WHO website

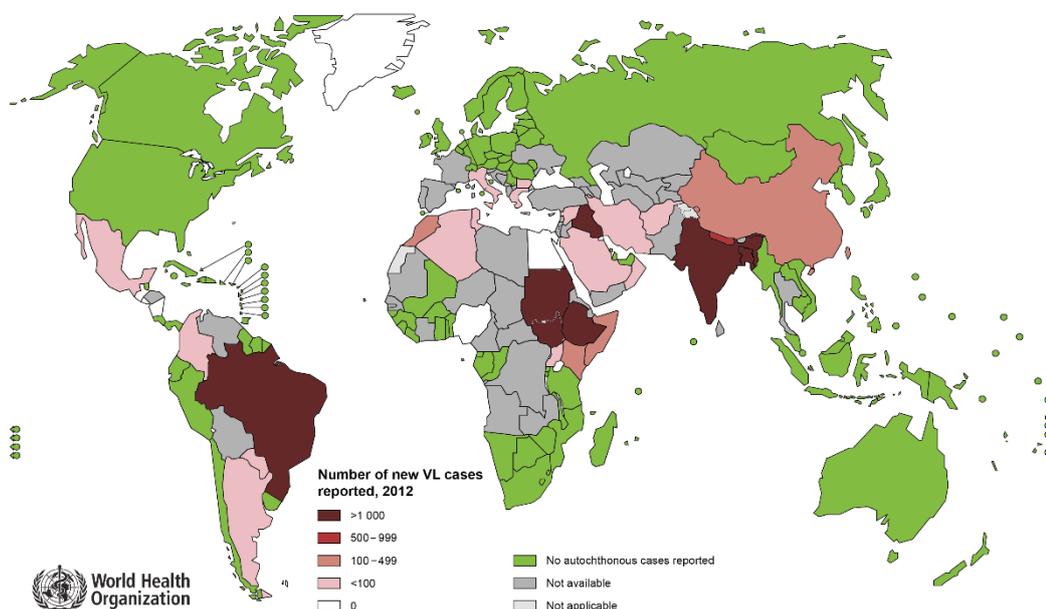


Figure 3 Global distribution of visceral leishmaniasis obtained from WHO

## 1.1.3 Lifecycle

*Leishmania* has a heteroxenous lifecycle which have two distinct morphological forms in the vertebrate and invertebrate hosts. <sup>27</sup> In brief, the parasite lifecycle begins in the gut of a female sand fly following an infected bloodmeal, from which the parasite differentiates into the extracellular highly replicative flagellated procyclic promastigote. <sup>28</sup> In the invertebrate gut, these procyclic promastigotes undergo metacyclogenesis to differentiate as metacyclic promastigotes, which appear with a slender body and long flagella and are infective to bitten vertebrates. <sup>27</sup> When the sandfly is infected, the feeding frequency of the vector increases, thus perpetuating the prevalence of the disease. <sup>29</sup>

Following the full differentiation of metacyclic promastigotes, the sand fly inoculates a vertebrate host either passively or by regurgitation of the parasites following the bloodmeal. <sup>29</sup> The wound attracts phagocytes which phagocytose the parasites which forms a vacuole known as a phagolysosome in which the parasite resides and differentiates into a non-motile amastigote. <sup>30</sup> The lifecycle is repeated when the amastigote infected phagocytes are taken up when the sandfly receives a bloodmeal from an infected vertebrate (Figure 4). <sup>30</sup>

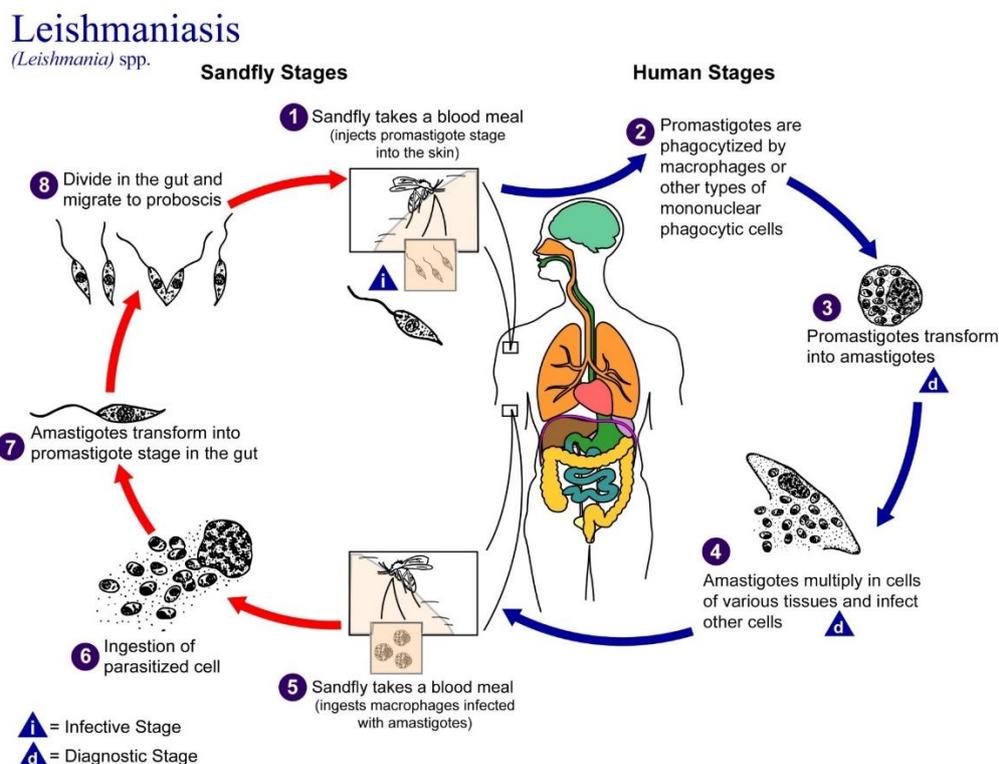


Figure 4 The lifecycle of *Leishmania* depicted by the CDC

Many wild and domesticated vertebrate species have been implicated as reservoir species, which provides additional challenges for the control of this disease.<sup>2</sup> Although the transmission of *Leishmania* is predominantly through the vector bite, other methods of transmission include: organ transplants, blood transfusions and drug use.<sup>25</sup>

#### 1.1.4 Treatment

Vaccination represents the ideal solution, however there are currently no human vaccines available.<sup>31</sup> Chemotherapy is the recommended treatment for leishmaniasis except in the cases of CL involving a single, or few small lesions, since this can heal without chemotherapeutic input thereby mitigating toxicity issues.<sup>32,33</sup> There are limited studies on MCL and DCL responses to current chemotherapy, with most cases of DCL resulting in relapses when treatment is halted.<sup>34,35</sup>

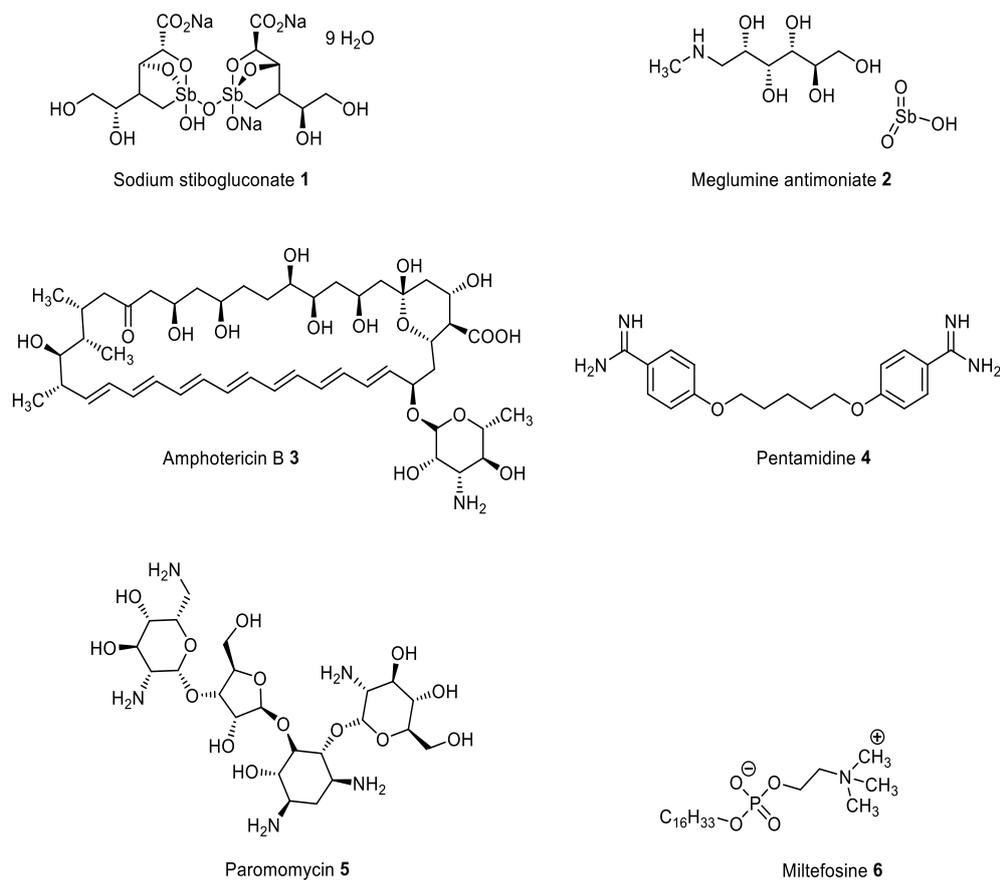


Figure 5 Chemotherapies currently available to treat leishmaniasis

#### 1.1.4.1 Pentavalent antimonials

In use since the 1920's, the first available chemotherapy for leishmaniasis were the pentavalent antimonials (**1** and **2**), which are effective against all forms of the disease.<sup>36,37</sup> These are still considered to be the first line of treatment in many countries despite issues with resistance and toxicity involving hepatotoxicity and cardiotoxicity.<sup>36,38</sup> In the Bihar region of India, resistance of VL to antimonials is around 60%.<sup>39</sup> These drugs also fail to treat DL in 75% of cases despite being a first choice therapy for this form.<sup>14</sup> Whilst the pentavalent antimonials are a low cost option, they require administration intramuscularly daily for at least 3 weeks.<sup>39,40</sup>

Although the mechanism of action for these drugs (**1** and **2**) have not yet been fully elucidated, several hypotheses have been explored over the years. For example, the antimonials (**1** and **2**) were proposed to be prodrugs, with Sb<sup>V</sup> being biologically reduced to the active trivalent form (Sb<sup>III</sup>).<sup>40</sup> Some reports also suggest that starting Sb<sup>V</sup> does have activity, but using a different route of entry to the parasitophorous vacuole.<sup>40,41</sup> Experimental evidence and field isolates have shown multiple modes of resistance, adding to the challenge of identifying a mechanism of action.<sup>42</sup>

#### 1.1.4.2 Amphotericin B

Amphotericin B (**3**) is an intravenous drug which was primarily developed as an antifungal, however it was found to be effective for treating cases that were unresponsive to other treatments.<sup>43</sup> Lipid formulations were developed to decrease toxicity, which improved the therapeutic index.<sup>44</sup> However, patients taking **3** still suffer with side effects such as fever, headaches, vomiting and nephrotoxicity.<sup>45</sup> The side effects are postulated to be due to non-specific binding to mammalian cell membranes.<sup>46</sup> **3** has been shown to be an effective drug, however it is not considered first line for VL in many countries because of the expense, especially for the affected demographic with the drug costing approximately \$260 per patient.<sup>47</sup>

Amphotericin B (**3**) works by binding to membrane ergosterol forming pores, which causes a loss in cell regulation causing cell death.<sup>48</sup> A study by Purkait *et al.* showed that resistance to **3** was conferred in a laboratory strain of *L. donovani* by replacing the target ergosterol with an alternative precursor.<sup>49</sup>

#### 1.1.4.3 Pentamidine

Pentamidine (**4**) is used for the treatment of *Pneumocystis pneumonia* when other treatment options have failed, as well as for trypanosomiasis and leishmaniasis.<sup>50</sup> Pentamidine is administered intramuscularly, with some common side effects being: nausea, vomiting, myalgia, syncope, and a sterile abscess around the injection site.<sup>51</sup> Pentamidine (**4**) also suffers from growing resistance, with the efficacy reportedly decreased from nearly 100 % to 70 % in Bihar, India.<sup>52</sup> This, in combination with the observed toxicity, has made **4** unfavourable as a first line drug.

Pentamidine (**4**) is believed to involve the inhibition of mitochondrial functions.<sup>53</sup> More specifically, what is seen in the mitochondria is a loss of membrane potential and cytochrome C accumulation.<sup>54</sup> However, other proposed hypotheses have been put forward including non-specific binding to tRNA, resulting in anticodon loop distortion which disrupts protein translation.<sup>55</sup>

#### 1.1.4.4 Paromomycin

Paromomycin (**5**) is an aminoglycoside antibiotic which is mainly administered intramuscularly. However, it is unique due to the potential for topical application to treat tegumentary forms of leishmaniasis.<sup>56</sup> Whilst paromomycin has been reported to be one of the safer anti-leishmanials, side effects from administration of **5** include nephrotoxicity, ototoxicity, and hepatotoxicity.<sup>56,57</sup> Resistance in the clinic to paromomycin has not been currently reported although it has been generated *in vitro*. This resistance was annotated with a mutation with calcium-dependent protein kinase 1 (CDPK1), a kinase which is involved in protein translation.<sup>58</sup>

The mechanism of action for **5** is not fully established for *Leishmania*. However, in bacteria, this antibiotic is known to form misfolded proteins, through binding to the A site of 16S rRNA, which induces codon misreading, and inhibits translocation.<sup>59,60</sup> In *Leishmania*, a similar mechanism has been proposed alongside other mechanisms including the alteration of mitochondrial membrane fluidity, and inhibition of respiration.<sup>61,62</sup>

#### 1.1.4.5 Miltefosine

The latest drug discovered for the treatment of leishmaniasis is miltefosine (**6**), which is a phosphatidylcholine analogue primarily designed to treat cancer, but showed efficacy

against *Leishmania*.<sup>63,64</sup> In contrast to the other discussed chemotherapies, **6** is orally available.<sup>64</sup> However, there are many side effects resulting from **6**, often affecting the gastrointestinal tract and include nausea, vomiting and anorexia but also it is thought to be teratogenic.<sup>65</sup>

In a small study, it was found that ADCL responded positively to miltefosine (**6**), however when the treatment was stopped on day 190 in which 94 % of patients showed no detectable *Leishmania*, 15/16 patients relapsed.<sup>35</sup> In addition, of those treated, some of the lesions reappeared during treatment, to which the authors concluded that the prolonged treatment may have led to drug resistance in these individuals.<sup>35</sup>

The mechanism of action is currently unknown, with several theories currently considered. Resistance to **6** in promastigotes are characterised by a decrease in unsaturated fatty acids, a higher percentage of shorter chains, and more cholesterol in the membrane.<sup>66</sup> This suggests that sterol metabolism is targeted.<sup>67</sup> Additionally, cytochrome C release was inhibited by **6** which could also suggest mitochondrial interaction.<sup>68</sup> This diverse set of observations could suggest a polypharmacology of **6**.<sup>67</sup>

#### 1.1.4.6 Concluding remarks on current treatment

As discussed, these drugs all have severe issues with toxicity, cost, increasing resistance, and variable efficacy in different regions and between species.<sup>69,70</sup> In addition, most of these drugs (**1**, **2**, **3**, **4** and **5**) are administered intravenous or intramuscularly, requiring regular hospital visits.<sup>71</sup> As leishmaniasis primarily affects the economically disadvantaged, the expense of most anti-leishmanial drugs are a considerable issue, though efforts by pharmaceutical companies along with WHO or NGOs which can aid with the cost.<sup>72</sup> In addition, most of the chemotherapies have unknown mechanisms, which means that structure activity relationship (SAR) cannot be utilised to increase efficacy and limit toxicity. These issues highlight the importance of new drug development in leishmaniasis (Table 1).

Collectively these issues emphasise the need for new drug development. In work undertaken in the Denny/Steel group directed towards this aim has identified IPCS as a potential drug target. This represents a pivotal enzyme in the sphingolipid synthesis pathway, which is a divergent step in eukaryotic cells.<sup>73</sup>

**Table 1** A summary of some of the current chemotherapeutic agents available to treat leishmaniasis. All the options available have toxicity issues associated with them and requires prolonged treatment, which reduces the adherence to the treatment. <sup>33,74–76</sup> SSG: sodium stibogluconate; MA: meglumine antimoniate.

<b>Drug</b>	<b>SSG 1 + MA 2</b>	<b>Amphotericin B 3</b>	<b>Pentamidine 4</b>	<b>Paromomycin 5</b>	<b>Miltefosine 6</b>
<b>Administrative route</b>	Intravenous, intramuscular, intralymphatic	Intravenous	Intramuscular	Topical, intramuscular	Oral
<b>Side effects</b>	Cardiotoxicity, nephrotoxicity, pancreatitis, hepatotoxicity	Nephrotoxicity, thrombophlebitis, hypokalaemia	Hyperglycaemia, myocarditis, hypotension, tachycardia	Ototoxicity, nephrotoxicity, hepatotoxicity	Hepatotoxicity, nephrotoxicity, teratogenicity
<b>Cost (USD)/ treatment</b>	56-60	126-280	120	15-21	65-150
<b>Treatment length (days)</b>	30	Until healing	Daily or alternating until healing	21	28 days
<b>Resistance</b>	Common	Not documented	Not documented	Lab strains	Lab strains, Few reports in India

## 1.2 Sphingolipids

Sphingolipids (SLs) are an important component of eukaryotic membranes, forming lipid rafts which have been postulated to be involved with membrane structure, and to act as signalling platforms. <sup>77,78</sup> Sphingolipids (SLs) have a general structure composed of a long chain base ( $R^1$ ), a fatty acid ( $R^2$ ), and a hydrophilic head group ( $R^3$ ) (Figure 6). <sup>79</sup>

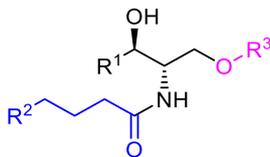


Figure 6 General structure of sphingolipids  
 $R^1$ : long chain base;  $R^2$ : fatty acid;  $R^3$ : hydrophilic head group

### 1.2.1 Sphingolipid synthesis

The pathway is conserved across all eukaryotic cells pathway for the synthesis of dihydrosphingosine (DHSph, **15**), starting with L-serine (**12**) (Figure 7). <sup>80</sup>

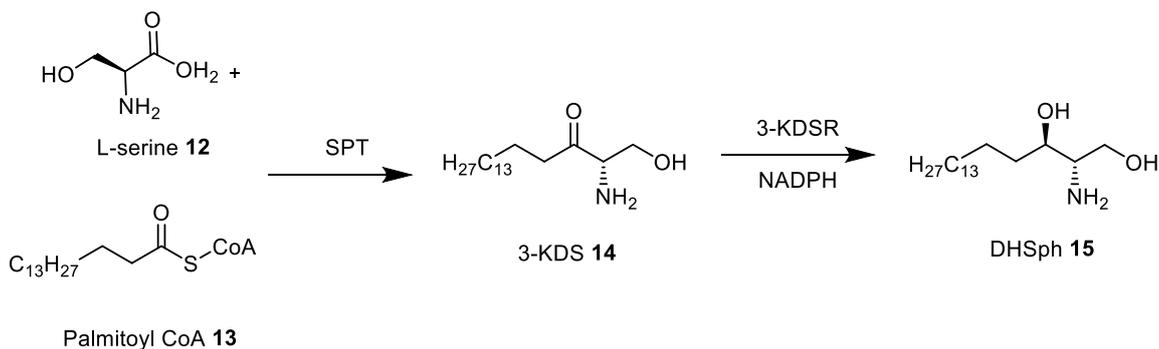


Figure 7 Conserved synthesis of dihydrosphingosine <sup>80</sup>

The further metabolism of **15** is divergent in different organisms, shown in Figure 8. <sup>77</sup>

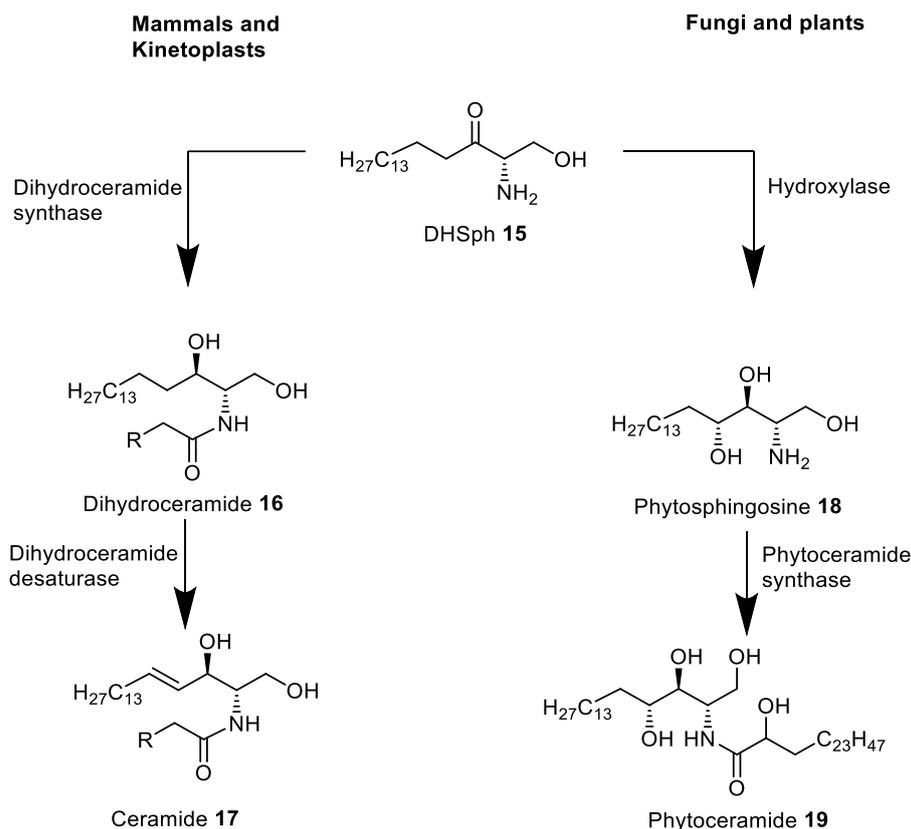


Figure 8 Divergent sphingolipid pathways<sup>81–83</sup>

In mammals and kinetoplasts, DHSph (**15**) is acylated to form the amide (**16**), then desaturated in the endoplasmic reticulum membrane or lumen to form the ceramide (**17**).<sup>79</sup> **17** is not soluble in the cytoplasm, therefore it is transported to the Golgi apparatus in vesicles, or by binding to ceramide transfer protein (CERT).<sup>83</sup> Contrary to this, in plants and fungi a reduction is carried out on DHSph (**15**), leading to **18** followed by amide formation to form phytoceramide (**19**).

Following the synthesis of ceramide (**17**), inositol phosphorylceramide synthase (IPCS) in kinetoplastids use **17** and phosphatidylinositol (PI, **21**) to synthesise inositol phosphorylceramide (IPC, **23**).<sup>84</sup> In mammals, **17** and phosphatidylcholine (PC, **24**) is converted to sphingomyelin (SM, **25**), using sphingomyelin synthase (SMS) (Figure 9).<sup>77,85</sup> This divergence is important to consider when identifying potential drug targets, since this path is the only divergent step in sphingolipid synthesis between kinetoplastids and mammals.<sup>77</sup>

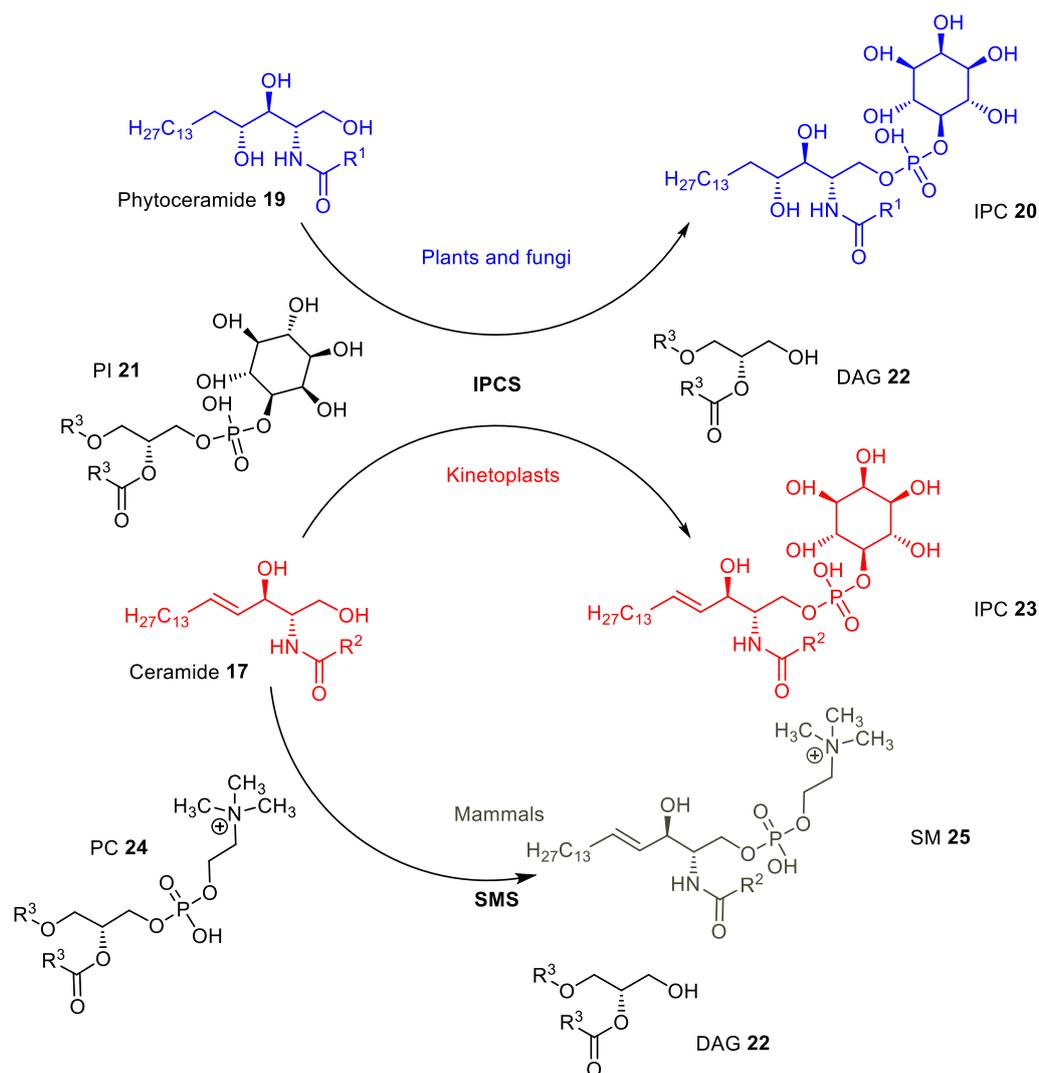


Figure 9 Synthesis of IPC and SM

IPC: inositol phosphorylceramide; IPCS: inositol phosphorylceramide synthase, PI: phosphatidylinositol, DAG: diacylglycerol, PC: phosphatidylcholine, SMS: sphingomyelin synthase, SM: sphingomyelin. The plants and fungi IPC production is highlighted in blue; the kinetoplastid IPC production is highlighted in red; the mammal SM production is black.

It is speculated that parasitic IPC (**23**) could contribute to membrane permeability and fluidity in the parasite.<sup>86</sup> Ceramide (**17**) is considered to be a marker for apoptotic like cell death, and DAG (**22**) which is released when IPC (**23**) is synthesised, is mitogenic. Therefore, it is hypothesised that disrupting this would cause the equilibrium to shift towards an accumulation of **17**.<sup>84,87</sup>

### 1.2.2 Inositol phosphorylceramide synthase (IPCS)

As discussed above, IPSCS represents the only divergent step in SL synthesis in eukaryotic cells. In addition, inhibition is hypothesised to cause a build-up of ceramide (**17**) which

causes apoptotic like cell death.<sup>73,87</sup> Understanding the structure of IPCS further could allow for structure activity relationship (SAR) drug optimisation of current identified inhibitors to occur.<sup>88</sup>

IPCS is a membrane bound protein (MBP) found in the Golgi apparatus, which is postulated to have 6 or 7 transmembrane helices, with the position of the catalytic triad is inside the lumen (Figure 10).<sup>89</sup> The exact structure of IPCS is unknown which could be due to the additional challenges of studying MBPs. These challenges can be attributed to the large amount of hydrophobic residues which causes the protein to aggregate in aqueous solutions, limited extraction success resulting in poor purification yield.<sup>90</sup> Less than 1 % of MBPs have a resolved crystal structure despite 60 % of current drugs targeting MBPs, with a large proportion of these targets being G-protein coupled receptors.<sup>91</sup>

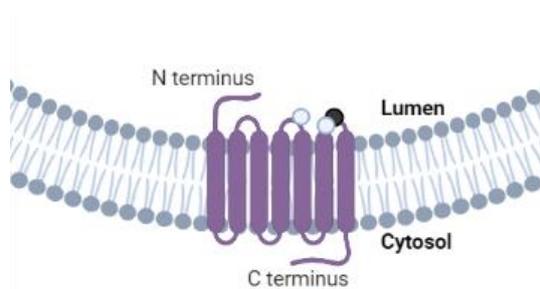


Figure 10 Hypothesised transmembrane structure of IPCS  
White circle: rough position of active site histidine; black circle: rough position of active site aspartic acid. Membrane is representative of the Golgi apparatus. Image created in biorender.com.<sup>73</sup>

IPCS is conserved in the different species of *Leishmania* as well as other *Kinetoplastidae* like *Trypanosoma* (Figure 11). Moreover, between IPCS and the mammalian orthologue SMS, there was only 29 % identity, although the catalytic triad residues are conserved suggesting a common catalytic triad mechanism.<sup>73</sup>

```

HsSMS1  117 DWEAQLRRIMKLIAGGGLSITGSHNMCGLDYLKSGHTVMLTLTYLFIKEYSPRRL-WWYHW
HsSMS2   62 DSQAKVQRILRLISGGGLSITGSHILCGDFLFSGHTVTLTLTYLFIKEYSPRHF-WWYHL
TbIPCS   38 --NPVKNVILTLVLTAGGGSITGSHNMCGLDYLKSGHTVMLTLTYLFIKEYSPRHF-WWYHL
TcIPCS   12 --NPVKNVIMTLVTLGSGSITGSHNMCGLDYLKSGHTVMLTLTYLFIKEYSPRHF-WWYHL
LmjIPCS  75 --HPVNLVILTLVTLGSGATH-----CGDLMFSGHTMILSLAFILAWDYSPFLHPWAVRV
LdIPCS   75 --HPVNLVILTLVTLSSGATH-----CGDLMFSGHTMILSLAFILAWDYSPFLHPWAVRV

HsSMS1  310 ICWLLSVVGIFCILLAHDRHYTVDVVVAYITTRLFWWYHTMANQQVLKEASQMNLLARVW
HsSMS2  254 ICWLLSAAGIICILVAHEHYTIDVVIAYITTRLFWWYHSMANEKLNKLVSSQTNFLSRAW
TbSLS   253 VVTVVAIFSYYCIVASRSHYTDVVLVAIYLTIAATF----IAVGHNADGAPWQLQLFIR-W
TcSLS   227 ASVLLVLLSFYSIIASRSHYTDVVLVAIYLTIAATF----IAVGHNADGAPWQLQLLIG-W
LmjIPCS 246 WVSVLLPISYYCIVASRSHYTDVVLVAIYLTIAATF----IAVGHNADGAPWQLQLLIR-W
LdIPCS  246 WVSALLPISYYCIVASRSHYTDVVLVAIYLTIAATF----IAVGHNADGAPWQLQLLIR-W

```

Figure 11 Alignment of inositol phosphorylceramide synthase (IPCS) peptide sequences in different species. The featured in the alignment is in ClustalW format. LmjIPCS: *Leishmania major* IPCS (accession E9AFX2); LdIPCS: *Leishmania donovi* IPCS (strain BPK282A1, accession E9BT22); HsSMS1: *Homo sapiens* sphingomyelin synthase 1 (accession Q86VZ5); HsSMS2: *Homo sapiens* sphingomyelin synthase 2 (accession Q8NHU3); TbSLS: *Trypanosoma brucei* sphingolipid synthase (accession B3A0L9); *Trypanosoma cruzi* sphingolipid synthase (accession PBJ73426). Highlighted in yellow are the conserved residues in the catalytic triad; highlighted in black are the conserved residues between all the species.

An assay using *Leishmania major* IPCS (LmjIPCS) enriched microsomal membrane extract obtained from complemented *Saccharomyces cerevisiae* was developed by Mina *et al.* (2010).<sup>92</sup> Briefly, *aur1* auxotrophic *S. cerevisiae* was complemented with the *Lmjipcs* gene to generate the expression system YPH499-HIS-GAL-AUR1 Prs246 LmjIPCS.<sup>73</sup> This resulted in *S. cerevisiae* expressing *L. major* IPCS (LmjIPCS) instead of the intrinsic yeast orthologue AUR1. The LmjIPCS enriched Golgi apparatus was then isolated using differential centrifugation, initially described by Fischl *et al.* (1999) for the isolation of intrinsic *S. cerevisiae* IPCS, and then adapted by Mina *et al.* (2010) for LmjIPCS.<sup>92,93</sup> The isolated Golgi apparatus was washed with CHAPS detergent to remove some of the contaminating proteins also present without affecting the enzymes activity.<sup>92</sup>

Using this assay, an SAR analysis using fluorescent derivatives of the substrate ceramide generated the hypothesised active site mechanism.<sup>84</sup> This is known as the ping-pong bi-bi mechanism and involves His-264 initially acting as a nucleophile on PI (**21**) to form the phosphohistidine intermediate. Then, His-220 deprotonates the ceramide alcohol which subsequently attacks the phosphohistidine, forming IPC (**23**, Figure 12).

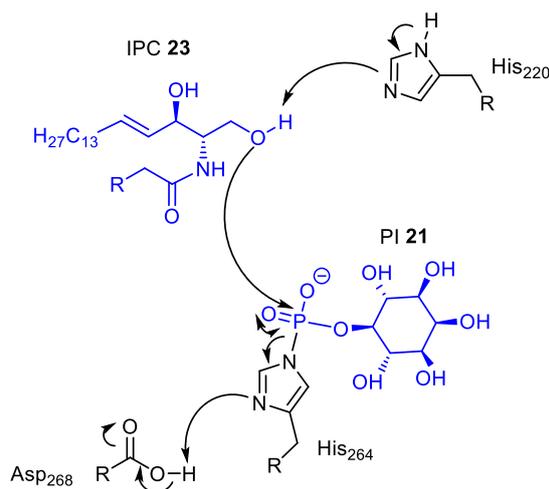


Figure 12 The catalytic mechanism of IPCS in kinetoplastids<sup>84</sup>  
 Blue: enzyme substrates; black: key amino acids for the catalytic mechanism. Shown is the intermediate phosphohistidine after a nucleophilic attack on the phosphate.

Whilst the evidence to date supports this process, not enough is known about this enzyme which can hinder a target-based drug discovery project. Currently, it is a challenge to study MBPs, however the nucleophilic IPCS histidine residue could allow an electrophilic probe to be designed that selectively binds to this histidine, taking advantage of the increased reactivity. This could then, potentially, lead to a better understanding and characterisation of this enzyme, aiding drug discovery efforts for *Leishmania* through TBDD.<sup>73</sup>

### 1.3 Concluding remarks

Leishmaniasis is a neglected tropical disease with major issues with the chemotherapeutic options available.<sup>39</sup> IPCS is an essential enzyme identified in *Leishmania* which is present in both lifecycle stages of the parasite with no direct human orthologue.<sup>73</sup> To explore this enzyme further, activity-based protein profiling can be utilised however, the active site of IPCS contains a HHD catalytic triad which is not currently targetable using this approach.<sup>94</sup>

### 1.4 Activity based protein profiling (ABPP)

In biomedical research, greater understanding of cellular function in cells and organisms can lead to more opportunities to target disease states.<sup>95</sup> Whilst progress have been made to address this with the increasing availability of whole genome sequences of organisms, a large number of discovered genes are uncharacterised, with qualitative information such as post translational modifications being lost.<sup>96</sup> A complementary technique to genomic analysis is activity-based protein profiling (ABPP) which can allow for the understanding of

the protein within the cellular context. <sup>97</sup> ABPP allows for the visualisation of expressed proteins, quantify dynamic levels of expression, modification states and cellular location which enhances knowledge of proteomics. <sup>98</sup>

ABPP uses active site-directed chemical probes tuned to covalently react with defined enzymes in a biological sample. <sup>99</sup> This labelling is determined by the active catalytic residues in the target enzymes reacting with the 'warhead' of the probe, making this technique specific to enzymes with related functions. <sup>97</sup> When the probe is not directly attached to the reporter group, bioorthogonal functional groups within the probe can allow for the attachment of reporter groups such as fluorophores and biotin to be added to allow for the visualisation and isolation of bound proteins. <sup>100</sup> To achieve this, ABPP can be broken down into 3 basic components: a warhead, a bioorthogonal linker group or binding group and a reporter group. <sup>100</sup> These are discussed in turn below.

#### 1.4.1 Warhead

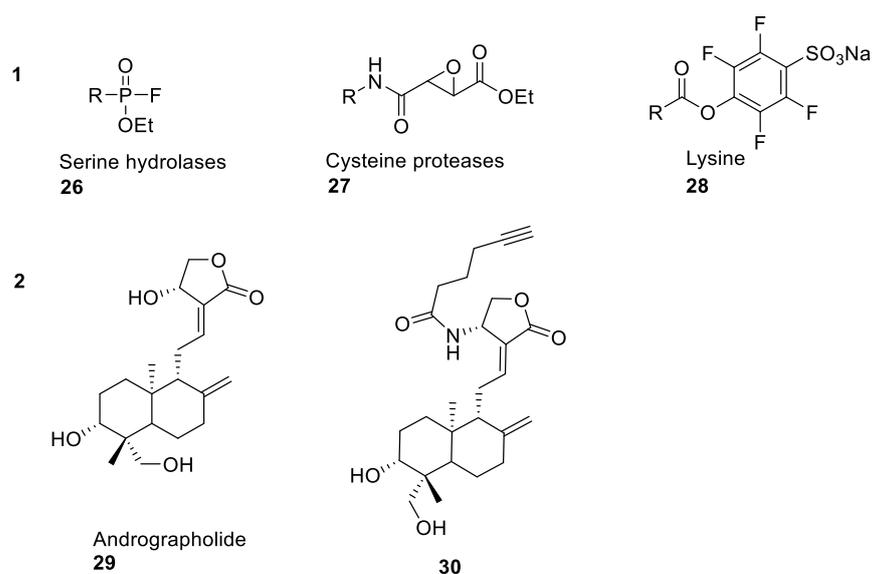


Figure 13 A selection of literature reported warheads and the target residues <sup>101</sup>  
*R*: spacer and binding group. Line 2 shows the parent drug molecule and the derived warhead.

The warhead, also known as a reactive group, selectively and covalently reacts with an active component of a functioning enzyme. <sup>97</sup> The warhead is a challenging component to design, with desired characteristics including specificity within a complex mixture and minimal indiscriminate reactivity. <sup>97,102</sup> This also requires discrimination between the active form of the target enzymes and the inactive, proenzyme forms. <sup>103</sup>

So far, many of the reported probes are electrophilic molecules, reacting with nucleophilic amino acid residues. Notable examples include the labelling of serine with fluorophosphonates (**26**), cysteine with epoxides (**27**) and lysine with esters (**28**) (Figure 13, 1).<sup>100</sup> These probes are considered broad spectrum, enabling the global profiling of almost entire enzyme classes.<sup>103</sup> As well as being used for broad enzyme characterisation, this approach can also be used to identify the targets of inhibitors.<sup>104</sup> This is achieved using competitive inhibition of the ABP with the small molecule inhibitor, which can be a robust way of identifying an inhibitor of an enzyme in a complex environment.<sup>105</sup> This class of ABP is discussed further below as it is directly relevant to this thesis.

In addition to this approach, ABPP can also offer tailor made probes, which are highly selective reacting to a small selection of enzymes or even a single target.<sup>103</sup> In this approach, the probe is usually a derivative of a parent molecule with an unknown mechanism of action.<sup>106</sup> Desired characteristics for the design of tailor-made probes include a similar pharmacological activity to the parent molecule, and functional handles to allow for the enrichment of bound targets.<sup>107</sup> An example of this type of probe is **30**, derived from the parent molecule andrographolide (**29**), a drug with anti-cancer properties (Figure 13, 2). The drug (**29**) was hypothesised to form a covalent interaction with Cys, thus, adding bioorthogonal groups formed **30** and implementing this probe to the lysate of a colorectal cancer cell line (HCT116) revealed that the drug had two targets; NF- $\kappa$ B and actin.<sup>108</sup>

Once a warhead has been verified for specific activity, this can be applied to a wide range of biologically active samples from many different organisms.<sup>97</sup>

## 1.4.2 Spacer and binding group

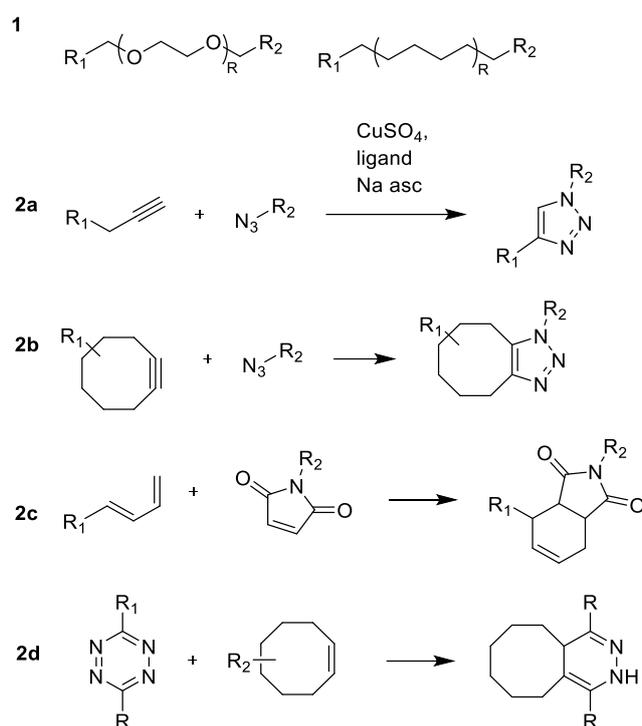


Figure 14 Selection of linkers and available bioorthogonal groups <sup>109</sup>  
*R<sub>1</sub>: warhead; R<sub>2</sub>: reporter tag*

Between the warhead and the reporter functional groups is a flexible linker, which is typically attached to the reporter group. <sup>97</sup> The most basic linkers are long alkyl or polyethylene glycol (PEG) chain, which aims to reduce the steric hindrance of the reporter component, thereby increasing cross-reactivity (Figure 14, 1). <sup>100</sup> These linkers have different properties for example PEG linkers are more hydrophilic, more polar and increased flexibility than the alkyl counterpart, which tend to be more hydrophobic. <sup>110</sup>

To selectively react the linker and the warhead, latent bioorthogonal functional groups are used to limit cross reactivity with active biological elements (Figure 14, 2). <sup>96</sup> These reactions need to occur in physiological conditions without reacting with intrinsic biological functionalities, and cross react with low reagent concentrations. <sup>109</sup> In some probes, the linker intrinsically forms part of the warhead, negating this step (discussed in section 1.4.3).

95

The alkyne-azide functional groups are commonly chosen since they are small residues and thus, have minimal steric impact. <sup>101</sup> However, this system requires copper for the

formation of the 1,2,3-triazole, which causes cell death, therefore this option is undesirable for use in living systems (Figure 14, 2a).<sup>100</sup> To overcome this, a ring strained cyclooctyne was developed to mitigate the use of copper, which can be used for the imaging of live cells (Figure 14, 2b).<sup>100,111</sup> In addition, a Diels-Alder ligation with an electron poor dienophile and an electron rich diene can also be used to crosslink the components without the need of a catalyst. (Figure 14, 2c).<sup>99,109</sup> Whilst cyclooctyne and diene readily react with their counterparts, they are also reactive to thiols, increasing non-specific interactions in a complex mixture, however, using iodoacetamide before implementation can mitigate some of this effect.<sup>109</sup> Another alternative bioorthogonal chemistry is a tetrazine ligation, using 1,2,4,5,-tetrazine, which proceeds through a Carboni-Lindsey reaction (Figure 14, 2d).<sup>112</sup> This reaction occurs rapidly with high specificity, selectivity and yield whilst also allowing for live cell imaging.<sup>112</sup>

#### 1.4.3 Reporter tag

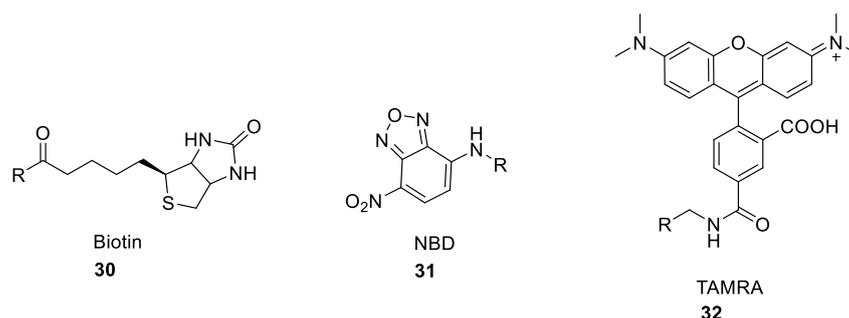


Figure 15 A selection of reporter tags used for ABPP analysis  
R: linker and warhead

The reporter tag is used for the detection or purification of the labelled proteins (Figure 15), and is typically added after the incubation of the warhead.<sup>113</sup> With a biotinylated reporter tag (**30**), streptavidin enrichment and pulldown of the warhead modified enzymes can be performed, resulting in the selective isolation of the targets.<sup>97</sup> One challenge with this is that the conditions required for the elution of biotin from streptavidin involves harsh conditions, using detergents like SDS and reducing agents like urea with the application of heat, typically at 65 °C to denature the streptavidin.<sup>94</sup> This can lead to endogenous biotinylated proteins, non-specifically bound proteins and also streptavidin contaminating the true probe modified eluted proteins.<sup>114</sup> To overcome this issue, an additional functional group can be incorporated into the linker to cleave under milder conditions.<sup>115</sup>

This cleavable linker needs to be stable to biological conditions but still cleave under mild conditions. Common examples are the disulphide linkage (**33**) which can be cleaved on treatment with reducing agents like  $\beta$ -mercaptoethanol (**34**), and levulinoyl esters (**35**) which are cleaved with hydrazine (**36**, Figure 16).<sup>116</sup>

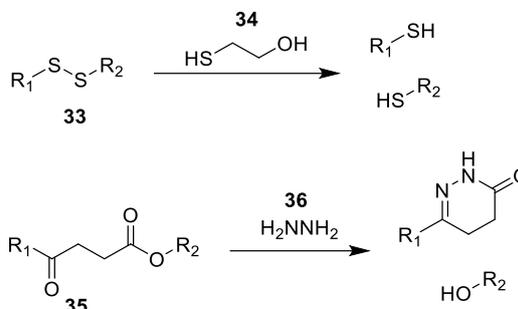


Figure 16 Cleavable linkages

In order to aid detection, a fluorescent reporter tag can be used for sensitive visualisation of the bound products *in vitro* or *in vivo*.<sup>117</sup> Cell permeable fluorophores include BODIPY and NBD (**31**), which can be used for whole cell imaging.<sup>104</sup> Other commonly used fluorophores include TAMRA (**32**), which requires lysis before addition due to limited cell uptake and permeability.<sup>118</sup> Following proteome treatment, identification of labelled proteins can happen by separating the proteins using SDS-PAGE and fluorescent analysis. Alternatively, the protein sample can undergo enrichment then SDS-PAGE analysis followed by band excision, digestion with proteolytic enzymes and identification using mass spectrometry.<sup>95</sup> Using SDS-PAGE separation and fluorescent analysis, the fluorescent signal can be used to rapidly determine competitive inhibition.<sup>100</sup>

Some ABPs, like the fluorophosphonates, commonly have the reporter tag attached to the reactive warhead, allowing for visualisation or enrichment after probe incubation.<sup>95</sup> This one step method is easier to use, negating the need for the following click chemistry reaction, however this is at the cost of reduced permeability and increased steric hindrance (Figure 17).<sup>103</sup>

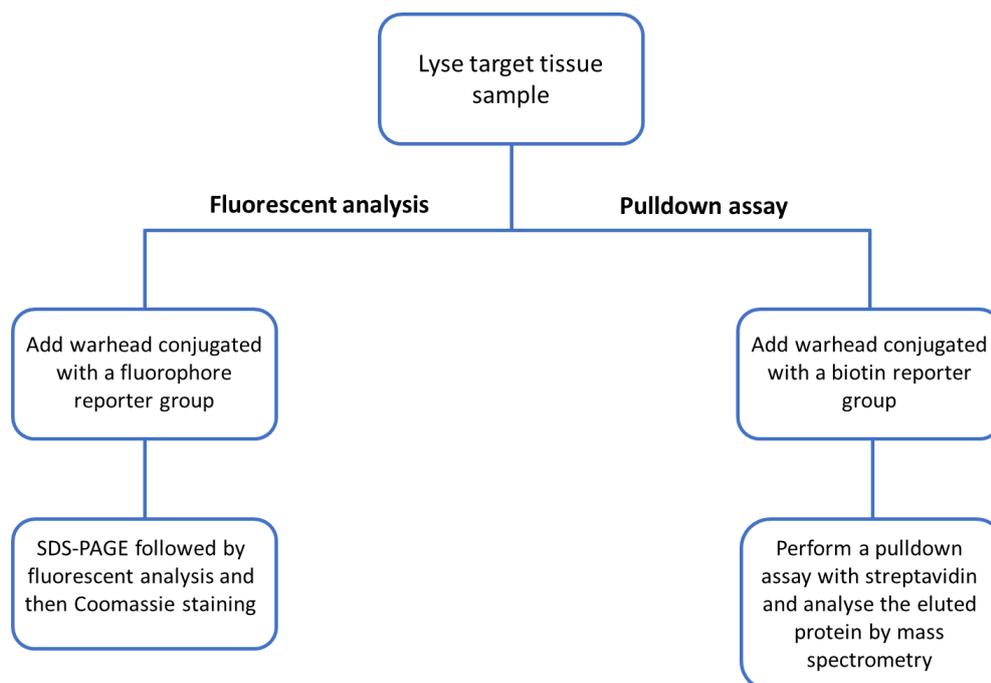


Figure 17 Flow chart demonstrating the process of the two-step ABPP method. Requires the warhead to have the linker and reporter functional group attached.

Alternatively, the warhead can be added first, followed by click chemistry with the linker and reporter group in a two-step method, which is discussed in length below (section 1.4.4).<sup>119</sup> Whilst this is more steps adding complexity, this method has the added flexibility of applying the relevant reporter group for the experiment.<sup>103</sup>

#### 1.4.4 Fishing for targets

In a typical two-step ABPP experiment, the warhead is initially incubated with the desired biological tissue, which is followed by click chemistry with the linker and reporter group allowing for protein target identification through enrichment and fluorescent analysis.<sup>107</sup> Initial warhead labelling can take place *in vitro*, *in situ* and *in vivo* (Figure 18, 1).<sup>97</sup> However, many ABPs have limited membrane permeability, requiring initial homogenisation of the desired sample prior to the addition of the warhead, which risks altering enzyme function.<sup>94</sup> Following probe incubation and tissue homogenisation, if required, the desired reporter tag is then added (Figure 18, 2a and 2b), CuAAC performed, and subsequent complementary analysis executed such as SDS-PAGE or enrichment (Figure 18, 3).<sup>100</sup>

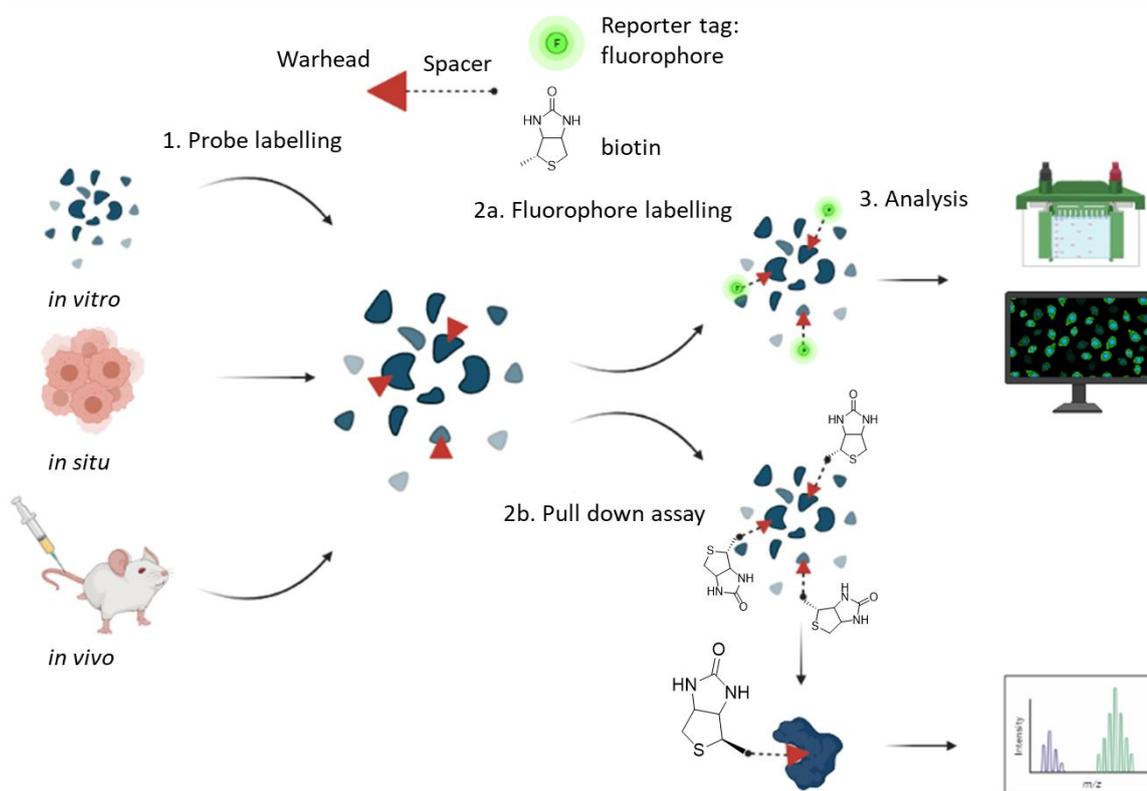


Figure 18 Overview of ABPP process

Red triangle: warhead; F: fluorophore. The warhead is incubated with active protein(s). Using bioorthogonal functional groups, this can be utilised for fluorescent analysis or a pull down assay for protein isolation and subsequent mass spectrometry analysis. Created with BioRender.com

A limitation of activity-based protein profiling is that the proteins analysed need to have a nucleophilic active site residue, meaning that many proteins, like receptors, can be overlooked.<sup>99</sup> When the interaction between the probe and the target proteins are reversible, affinity based protein profiling can be used which involves the incorporation of an additional photoaffinity moiety.<sup>120</sup> When the affinity based probe is exposed to a specific wavelength of light, the formed reactive intermediate will form a covalent bond with the nearby protein, allowing further isolation and identification of bound proteins that may not have an active component.<sup>121</sup>

### 1.5 Application of ABPP

Many of the designed ABPs are electrophilic, therefore this approach targets primarily nucleophilic active site residues like Ser, Cys, Lys, Tyr, Glu, and Asp with recent developments for Arg.<sup>104,122</sup> Additionally, these residues are the most frequently reported residues found in enzyme active sites together with His, which was found to be the most

common residue.<sup>123</sup> Although, there are no ABPs targeting His, progress have been made with the other residues, discussed below.<sup>104,123</sup>

### 1.5.1 Serine hydrolases

Serine hydrolases are one of the largest enzyme classes known, and these employ a conserved nucleophilic serine which serves to hydrolyse amide, ester and thioester bonds.

<sup>124</sup> This class of enzyme have several ABPs targeting them, most notably the fluorophosphonates (FPs, Figure 19).<sup>125,126</sup>

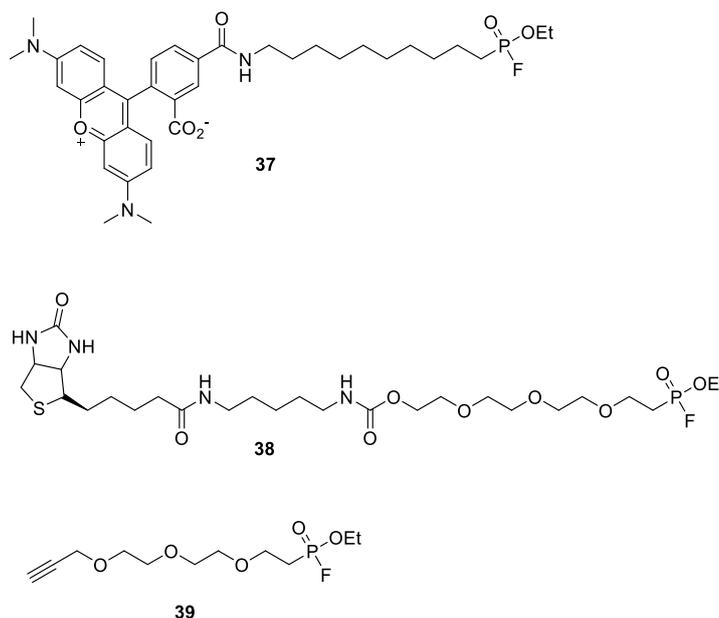


Figure 19 A selection of reported fluorophosphonate probes  
 FP-TAMRA (**33**) is commercially available; FP-Biotin (**34**) is reported by Kidd et al.<sup>127</sup>; FP-alkyne (**35**) is reported by Gillet et al.<sup>126</sup>

First described in 1999, the FPs have been shown to target 80 % of the predicted serine hydrolases present in across multiple mouse and human tissues.<sup>125,128</sup> Whilst the first reported FP-biotin had an alkyl linker instead of a PEG (**38**), this was adapted to increase aqueous solubility.<sup>125,127</sup> To reduce non-specific hydrophobic interactions, increase active site accessibility and achieve membrane permeability, **39** was designed. This provided a different serine hydrolase profile however, it suffered with increased hydrolysis degradation.<sup>126</sup> ABPP for serine residues have aided in the discovery of dysregulated serine hydrolases in diseases and identified several enzymes with therapeutic potential, which demonstrates the value of ABPPs.<sup>129</sup>

## 1.5.2 Cysteine proteases

Cysteine is the least abundant amino acid, however it is also the most intrinsically nucleophilic residue due to the electron rich sulphur present.<sup>130</sup> Along with acting as a nucleophile, this residue also possesses redox activity and can act as a metal ligand<sup>131</sup> Cysteine was the second reported residue to have an ABP designed, with the initial study being published in 2000 (**40**, **41**) (Figure 20).<sup>132,133</sup>

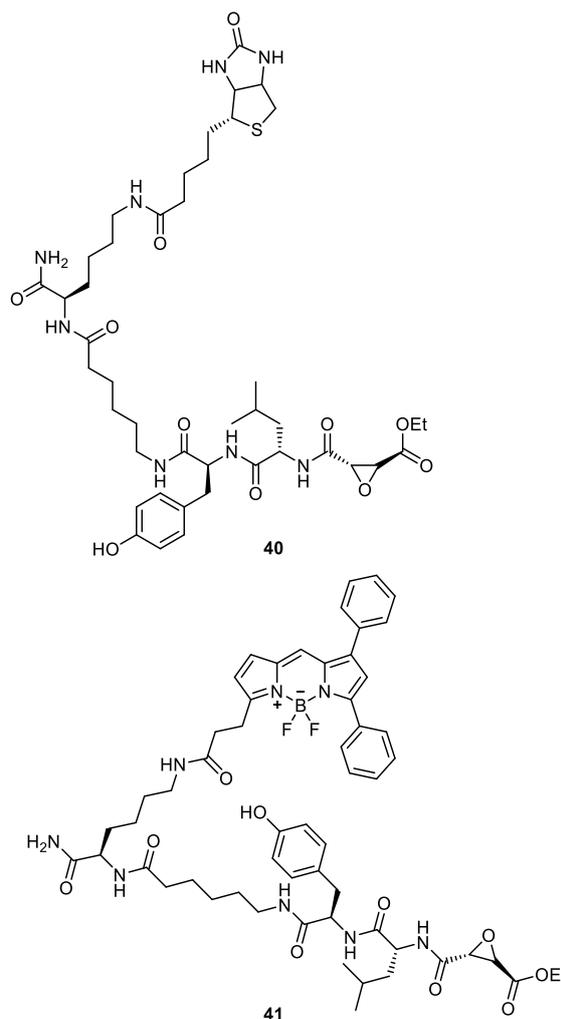


Figure 20 Examples of cysteine reactive probes.<sup>133</sup>  
 Top probe (**40**) was used for pulldown assays whilst bottom probe (**41**) was used for fluorescent analysis on 10 cell lines to study the papain family profile of skin cell carcinomas

**40** and **41** are ABPs reported by Greenbaum *et al.*, which are specific to the papain family of cysteine proteases due to the amino acids mimicking the substrate. Following this a covalent modification of the target occurs via sulphur mediated epoxide ring opening.<sup>133</sup> These probes were implemented to pancreatic tumours to study the importance of

cathepsin cysteine proteases, which found upregulation of several cathepsins during tumorigenesis.<sup>134</sup> In addition, using these probes for competitive inhibition identified small molecules which were specific to enzymes in this family.<sup>135</sup>

### 1.5.3 Lysine

Lysine residues compose 5.9 % of the human proteome.<sup>136</sup> This nucleophilic amino acid has a primary amine residue, which is found in the active sites of proteins such as PI3Ks and lipid kinases.<sup>137</sup>

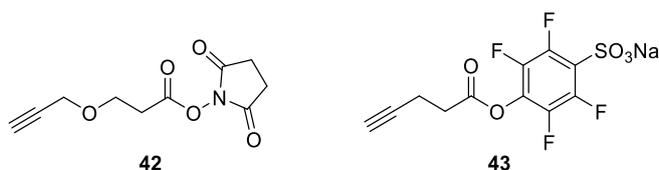


Figure 21 Lysine modifying ABPs

NHS-esters (**42**) are used for global proteomic profiling (Lys, Ser, Thr), with almost half of the labelled residues being lysine (Figure 21).<sup>138</sup> Adjusting this ester further conferred selectivity with lysine residues (**43**, Figure 21). In 2017, a study was performed using a sulfotetrafluorophenyl ester probe (**43**) to profile the human lysine residue proteome. Using **43**, 9,000 lysine containing proteins were identified and quantified, expanding druggable targets of the proteome in humans.<sup>139</sup>

Whilst ABPP has been used to expand knowledge on certain enzyme profiles as discussed above, enzymes are highly diverse and many are not currently targeted using ABPs.<sup>117</sup> One target are enzymes containing an active site histidine residue.

## 1.6 Histidine activity in enzymes

As discussed above, histidine is frequently reported in the active sites of a variety of enzymes making it one of the most common catalytic residues, which can be attributed to the reactivity of the imidazole group present.<sup>123,140</sup> Moreover, whilst the frequency of this amino acid is around 2 % in proteins, it is found in around 20 % of enzyme active sites, making it the most common catalytic residue.<sup>123</sup> This unique imidazole group has higher nucleophilic reactivity than primary and secondary amines, is hydrogen bond donor and acceptor and can coordinate to metals.<sup>141</sup> Despite these factors, there is a niche to study

enzymes with this catalytic residue further, as there have been few attempts to develop histidine modifying reagents, reviewed in section 1.7.<sup>104</sup>

### 1.6.1 Non-specific acid phosphatase

Active site histidine residues can act as a nucleophile, attacking the substrate with the lone pair of electrons found on N-3 of the imidazole.<sup>142</sup> This is an essential mechanism which can be found in a range of enzymes for example, the enzyme families histidine kinases and histidine phosphatases, with these working on phosphorylated substrates (Figure 22).

143,144

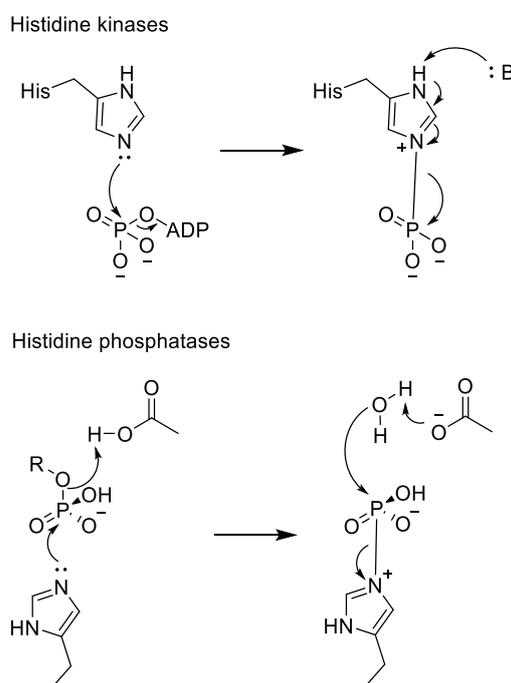


Figure 22 Basic mechanism for histidine kinases and histidine phosphatases<sup>143,145</sup>

Histidine kinases are membrane bound enzymes that are part of the two-component system (TCS).<sup>145</sup> The TCS has been identified in many organisms for example bacteria, archaea, amoeba, plants, fungi and viruses.<sup>146</sup> TCSs are used for environmental sensing, and this pathway begins with a conserved histidine in the histidine kinase (HK) becoming phosphorylated with the  $\gamma$ -phosphoryl group from ATP.<sup>147</sup> Following this, an aspartate on the response regulator (RR) then receives the phosphoryl group, which causes transcriptional changes in response.<sup>148</sup> Whilst most of the histidine kinases detect

extracellular signals for environmental sensing, some are soluble cytosolic enzymes that couple to transmembrane receptors, such as those linked to chemotaxis.<sup>145</sup>

Histidine phosphatases is a superfamily of enzymes that encompass diverse enzymatic functions and are characterised by the conserved nucleophilic histidine in the active site.<sup>143</sup> A large number of histidine phosphatases are located in the nucleus, however they can also be found in the ER, cell surface or secreted.<sup>143,149</sup> Within this family are the acid phosphatases which hydrolyse phosphate esters, a reaction essential for a variety of different metabolic processes.<sup>150</sup> In particular, PhoN is a non-specific acid phosphatase which was used as a model enzyme in Chapter 3.

Non-specific acid phosphatases (NSAPs) dephosphorylate a wide range of unrelated phosphoesters, exhibiting optimal activity in acidic to neutral pH.<sup>151</sup> The first NSAP to be discovered and sequenced was PhoC from *Zymomonas mobilis* in 1989.<sup>152</sup> Since this discovery, many more NSAPs have been identified in multiple species. NSAPs are divided into three distinct classes based on sequence identity, known as Class A, Class B and Class C.<sup>151</sup> Class A are the most phylogenetically distinct NSAPs, range from 25-27 kDa, can be secreted and can be monomeric to oligomeric in structure.<sup>153</sup> Class B are 25 kDa, membrane bound secreted metalloproteins and Class C are the largest NSAPs at 30 kDa and are secreted lipoproteins characterised by the catalytic residues which consist of 4 invariant aspartate residues.<sup>151</sup>

PhoN from *Salmonella typhumirium* is an example of a Class A NSAP which has a determined crystal structure with a resolution of 2.5 Å published in 2003 by Makde *et al* (Figure 23).<sup>154</sup>

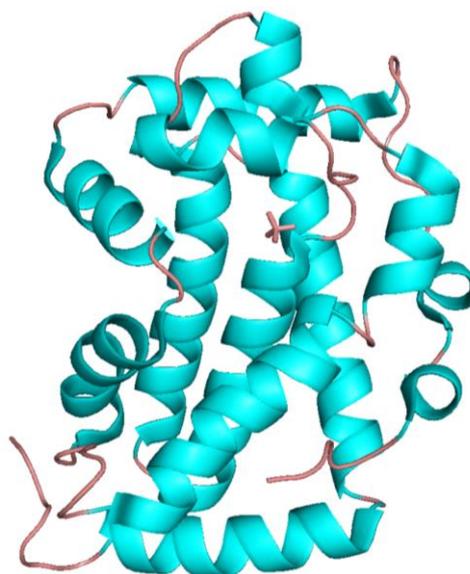


Figure 23 Crystal structure of a monomer of PhoN (PDB: 2A96)<sup>155</sup>  
Crystal structure determined using X-ray diffraction, with a resolution of 2.50 Å

The PhoN active site is composed of a HHD catalytic triad typical of Class A NSAPs, in which His-197 undergoes a nucleophilic attack with the electron-deficient organophosphate substrate, forming a phosphohistidine intermediate. The acid base proton from His-158 stabilises the formed phosphohistidine intermediate before acting as a base and forming the anion from water which then attacks the phosphohistidine releasing the inorganic phosphate (Figure 24).<sup>156</sup>

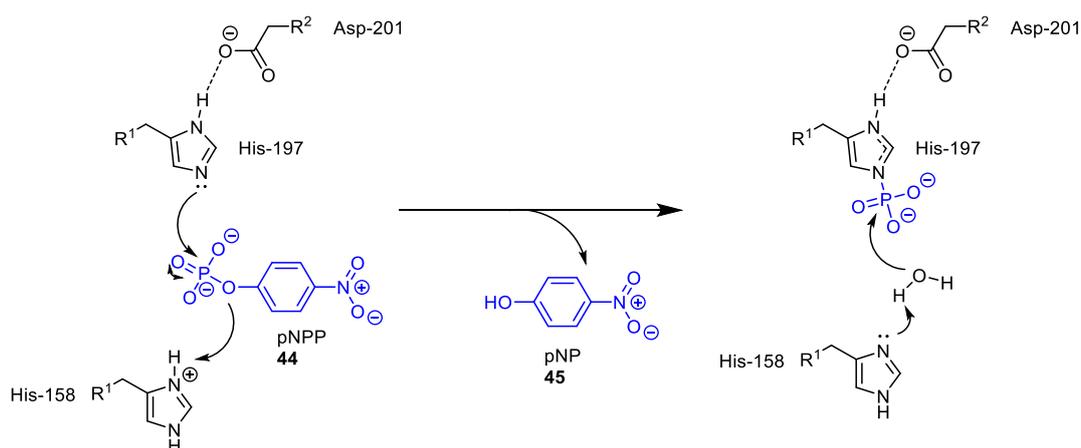


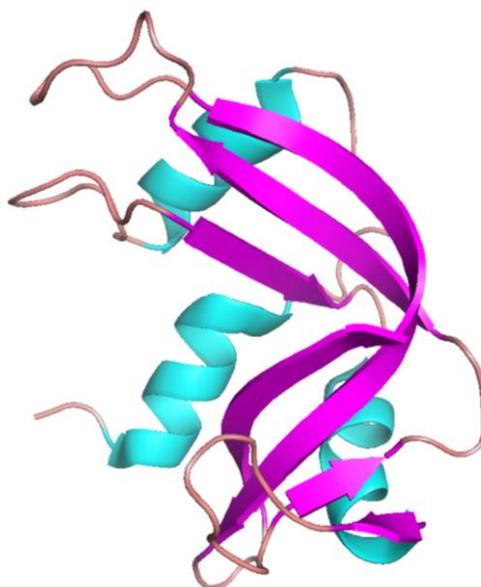
Figure 24 Active site and mechanism of PhoN *S. typhumirium* with para-nitrophenyl phosphate (pNPP) as the substrate<sup>155</sup> Blue: enzyme substrates; black: active site amino acids of PhoN.

Also demonstrated in PhoN, histidine can also act as an acid base proton shuttle. This type of histidine residue is discussed further below.

### 1.6.2 Ribonuclease A

A second function of histidine residues is that it can act as an acid and as a base in the active site of enzymes. This is due to the pKa of N-3 being around 7 therefore, at physiological pH, the amino acid can function as an acid and as a base.<sup>142</sup> This dual functionality is exploited in the active sites of enzymes with an acid base shuttle mechanism.<sup>157</sup> This type of role for histidine residues is the most commonly described role, present in a diverse range of enzyme active sites.<sup>123</sup>

Ribonuclease (RNase) has different isoforms, however RNase A from bovine pancreas was isolated and identified first and therefore, is the most studied (Figure 25).<sup>158</sup>



*Figure 25 Crystal structure of RNase A from bovine pancreas<sup>159</sup>  
Crystal structure determined using X-ray diffraction, with a resolution of 1.40 Å*

This enzyme cleaves the 3',5'-phosphodiester bond, consequently breaking down RNA.<sup>160</sup> The catalytic mechanism of RNase A was first proposed in 1961, and remains the most popular mechanism to date.<sup>161</sup> This describes two steps for the mechanism; an initial transphosphorylation followed by hydrolysis to cleave the phosphodiester bond.<sup>162</sup>

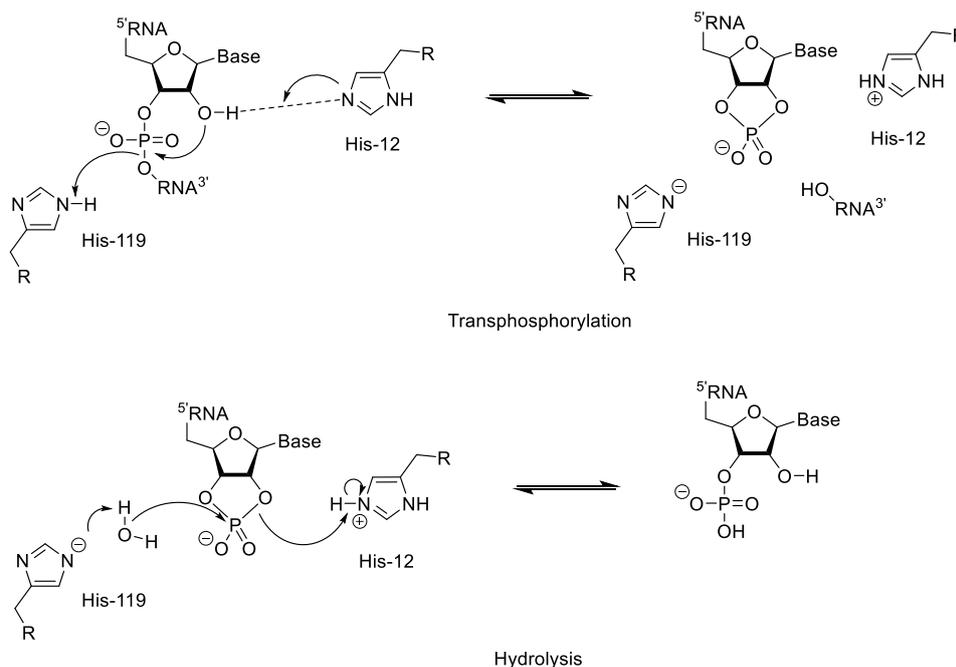


Figure 26 Proposed catalytic mechanism for bovine pancreatic RNase A <sup>163</sup>

Initial transphosphorylation begins with His-12 extracting a proton from 2'-OH, facilitating the nucleophilic attack of the oxygen to the phosphate, ending with deprotonation of His-119. For the second part of this mechanism, His-119 now acts as a base, abstracting a proton from water, leading to an attack from the hydroxyl on the phosphate. This leads to the reformation of the alcohol on the 2' carbon, with His-12 now acting as an acid (Figure 26). <sup>162</sup> Though this is the general method for catalysis, other amino acids do play a role in this process, for example, a mutation at Lys-41 lost catalytic activity by 2%. This loss of activity was thought to be due to the Lys-41 amine stabilising the intermediate negative charge. <sup>164</sup>

### 1.6.3 $\alpha$ -amino- $\beta$ -Carboxymuconate- $\epsilon$ -Semiaidehyde Decarboxylase

Histidine has a high affinity for metal ions, forming strong interactions with the ions and hence, can be found in metalloproteins. <sup>165</sup> This affinity is taken advantage of for the use of a polyhistidine tag in recombinant proteins for protein purification due to the adherence on a transition metal ion matrix. <sup>166</sup> His is considered to be a tridentate ligand, with the potential to coordinate using the amino nitrogen, carboxylate oxygen and more commonly, with N-3 on the imidazole. <sup>165</sup> This amino acid has been found coordinating to a range of

metal ions in proteins including but not limited to iron in heme, copper in tyrosinase and zinc, discussed below.<sup>167,168</sup>

Zinc is a cofactor which is critical for the function of more than 300 enzymes.<sup>169</sup> There are three different forms of zinc binding: structural, catalytic and cocatalytic.<sup>170</sup>

$\alpha$ -amino- $\beta$ -carboxymuconate- $\epsilon$ -semialdehyde decarboxylase (ACMSD) is involved in the metabolic degradation of the amino acid L-tryptophan by removing the carboxylate group from the substrate  $\alpha$ -amino- $\beta$ -carboxymuconate- $\epsilon$ -semialdehyde (ACMS, **46**).<sup>171</sup> In human ACMSD, the conserved residues His-6, His-8, His-174 and Asp-291 coordinate to the  $Zn^{2+}$  ion.<sup>172</sup> The metal ion in human ACMSD was not removed using EDTA, with the enzyme activity being directly related to zinc content, suggesting that catalytic activity is dependent on the zinc.<sup>173</sup>

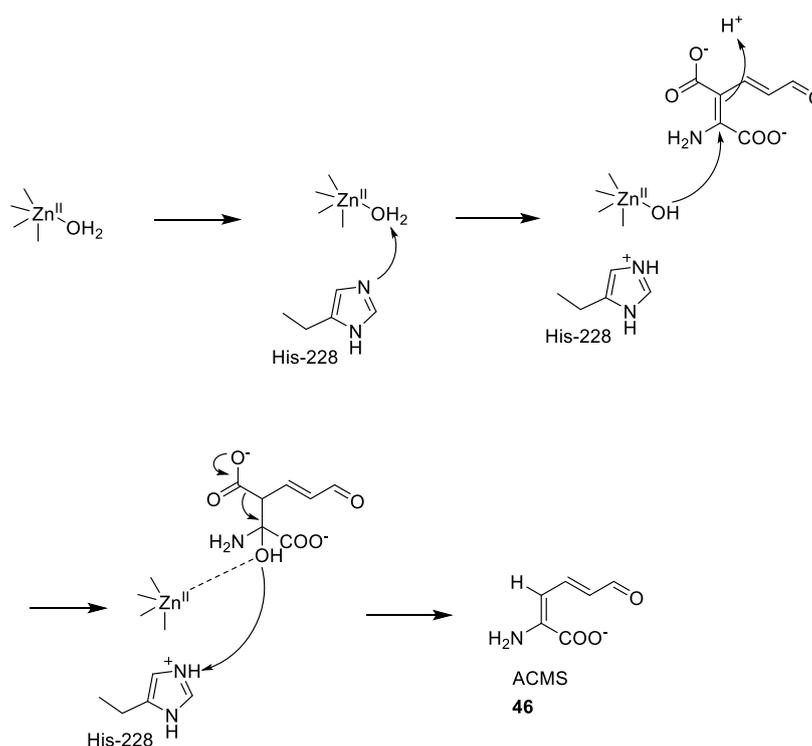


Figure 27 Active site mechanism of ACMSD<sup>174</sup>

In addition to the 3 histidine residues coordinating the zinc ion, His-228 lies opposite to the binding pocket and maintains the hydrolytic water ligand by acting as an acid base (Figure 27).<sup>174</sup> Mutation of His-228 causes a loss of enzyme function due to a lack of metal ion selectivity, incorporating Fe instead.<sup>174</sup>

## 1.7 Histidine modifying reagents

Currently, there is one warhead that reacts specifically with active site histidine residues in enzymes (**47**).<sup>175</sup> Additionally, there is a small selection of compounds which react with histidine residues (Figure 28).

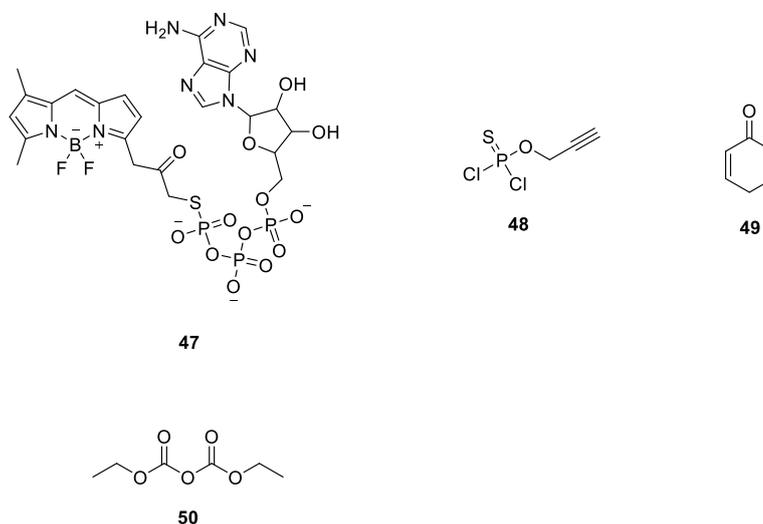


Figure 28 Histidine modifying small molecules

The ATP derivative **47** is a probe first described in 2012 which was used to study two component systems (TCS).<sup>175</sup> As discussed in 1.6.1, TCSs are used for environmental sensing with the histidine kinase (HK) dephosphorylating ATP and transferring this to a response regulator which causes transcriptional changes as a response.<sup>148</sup> To study TCSs, **47** was adapted to include sulphur on the  $\gamma$ -phosphate group, forming a thiophosphate intermediate.<sup>175</sup> This intermediate is more stable than the phosphohistidine intermediate, slowing down the hydrolysis due to being a poor leaving group in comparison.<sup>176</sup> This probe could then be transferred to the RRs in the same way as the substrate, so **47** was found to successfully allow for the visualisation of HKs and RRs.<sup>175</sup>

In a similar method to **47**, the thiophosphate compound (**48**) also used sulphur to stabilise the phosphohistidine intermediate. This probe (**48**) can be used for the labelling of solvent exposed histidine residues, shown with the labelling of His-48 on RNase A as opposed to the active site residues His-12 and His-119.<sup>163,177</sup> This probe had the most effective labelling at pH 8.5 (60%), however at this pH, **48** lost chemoselectivity, labelling some lysine residues

as well. Therefore to compensate for low reactivity at a neutral pH, higher equivalents were used.<sup>177</sup>

2-cyclohexenone (**49**) modifies histidine residues via Michael addition. This probe (**49**) monolabelled 8 enzymes, including RNase A with a modification identified on His-105.<sup>178</sup> This histidine is a solvent exposed histidine, which left the modified RNase A functional.<sup>163,178</sup> The labelled proteins could be fluorescently detected using retro Michael addition with a primary amine containing reporter tag on the formed cyclic ketone.<sup>178</sup>

From the described histidine residue probes, none target active site histidine residues, except DEPC (**50**). DEPC (**50**) reportedly selectively acylates active site histidine residues, which has been demonstrated in a range of enzymes like ribonuclease, glutamate dehydrogenase and phospholipase A<sub>2</sub>.<sup>179</sup>

#### 1.7.1 Diethyl pyrocarbonate (DEPC)

The use of diethyl pyrocarbonate (DEPC, **50**) for the modification of enzymes was first reported by Hullan *et al.*, in 1965 to inactivate trypsin whilst investigating why autolysis was inhibited in DEPC (**50**) treated organisms.<sup>180</sup> Initial fluorescence experimentation using bovine serum albumin (BSA) led to the hypothesis that DEPC (**50**) interacted with tryptophan residues, irreversibly inhibiting enzyme function<sup>181</sup> However, this theory was contested due to the lack of explanation for the strong inhibition of pancreatic ribonuclease, which possess no tryptophan residues.<sup>182</sup> Subsequently Wolf *et al.* (1970) proposed the carbethoxylation mechanism for amines, such as that present on Lys, however there was no mention of the modification on the imidazole of histidine.<sup>182</sup> Two years after this report, Morris and McKinlet-McKee (1972) detailed the inactivation of liver alcohol dehydrogenase using DEPC (**50**). Using fluorescent spectroscopy and varying concentrations of DEPC (**50**), this was shown to corresponded to the quantity of accessible histidine residues in the enzyme.<sup>183</sup> Following this report, there has been a consensus in the literature that DEPC (**50**) modifies N-3 on the imidazole of active site histidine residues.<sup>179</sup> As such, DEPC (**50**) has found an established role for genomic extraction due to the ability to inactivate ribonuclease and deoxyribonuclease, replacing the previously used phenol.<sup>184</sup>

With the established mechanism elucidated, DEPC (**50**) is reported to react with all of the discussed active site histidine residues like a zinc ion coordinator, such as in liver alcohol dehydrogenase (LADH), an acid base shuttle system like in RNase A and a nucleophilic histidine like in glucose-6-phosphatase.<sup>183,185,186</sup> Whilst DEPC (**50**) reacts with active site histidine residues, there are some reports of cross reactivity with other nucleophilic residues such as Cys and Lys.<sup>179</sup>

## 1.8 Project aims

The overall focus of this project was to design and validate a novel ABP based on the dicarbonate structure of DEPC (**50**, Figure 29) that will selectively bind to active site histidine residues. Due to the current lack of histidine binding ABPs, adapting the histidine acylating agent DEPC (**50**) to include a bioorthogonal warhead was hypothesised to allow for ABPP research.

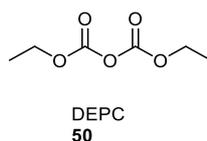


Figure 29 Diethyl pyrocarbonate

Ultimately, the initial focus was on the development of DEPC (**50**) analogues which was then validated in enzyme model systems. Once the scope of the ABP activity was defined, this was then taken into *Leishmania* for proteome exploration.

## 2 Design of a histidine specific probe

### 2.1.1 Synopsis of chapter

As discussed in the previous chapter, the initial aim of the project was to prepare an alkyne containing dicarbonate and validate this as an effective labelling agent for reactive histidine residues. Given this, the first approach was to establish a suitable synthetic method, followed by verification of reactivity with imidazole groups.

### 2.2 Initial probe design

Diethyl pyrocarbonate (DEPC, **50**) is a probe that is reported to selectively acylate active site histidine residues.<sup>179</sup> Therefore, for the initial probe design, it was thought that keeping the structure similar to DEPC (**50**) would be crucial to retaining the histidine labelling selectivity. The initial probe design was a derivative of DEPC (**50**) with only the addition of dialkyl groups. A review of the literature revealed 4 methodologies to synthesise dicarbonate compounds. One study by Plusquellec described a method of synthesising 20 symmetrical carboxylic and carbonic anhydrides from the corresponding chloroformates, including DEPC (**50**, Figure 30).<sup>187</sup> This methodology had the highest yield from the studies evaluated and used reagents that were immediately available.

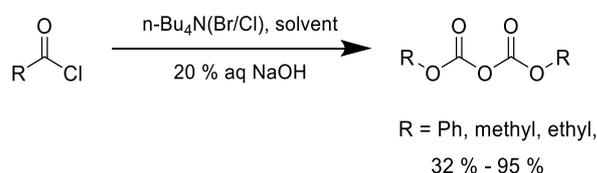


Figure 30 Methodology for the synthesis of dicarbonates<sup>187</sup>

Following this precedent, dipropargyl pyrocarbonate (DPPC, **52**) was generated, as a light yellow oil, from commercially available propargyl chloroformate (**51**, Figure 31).

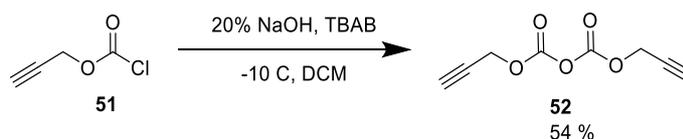


Figure 31 Synthesis of DPPC

Consistent with this structure, DPPC (**52**) provided two peaks in the  $^1\text{H}$  NMR spectrum at  $\delta$  4.76, (2H, d), and 2.54 (4H, t). Additionally, the  $^{13}\text{C}$  NMR spectrum displayed four carbon peaks, notably a carbonyl signal at  $\delta$  154.4, concurring with the recorded data for this molecule obtained by Huber and Diaz.<sup>188</sup> With DPPC (**52**) successfully synthesised, the reactivity was assessed using 4-methylimidazole (**53**) and benzimidazole (**55**) as a primitive model for histidine (Figure 32). Surprisingly, it was found that DPPC (**52**) was stable to both reagents in either MeCN or DCM at room temperature (RT).

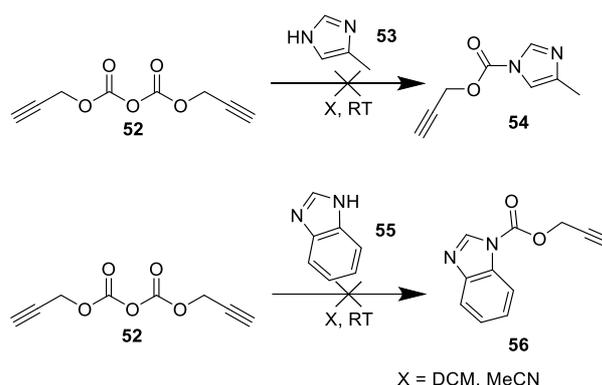


Figure 32 Conditions for the acylation of imidazole

The apparent stability of DPPC (**52**) allowed successful purification by flash column chromatography with the nucleophilic MeOH (DCM: MeOH, 0-30 %). This stability was unexpected, so it was postulated that dipropargyl carbonate (DPC, **57**) had been synthesised as opposed to DPPC (**52**, Figure 33).

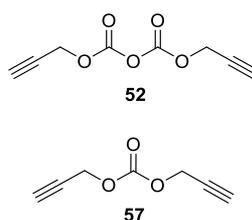


Figure 33 The hypothesised undesired synthesised carbonate

In order to preclude this possibility, the same procedure depicted in Figure 30 was used to synthesise DEPC (**50**). This can be directly compared to an authentic commercially available sample of DEPC (**50**) and diethyl carbonate (DEC, **58**, Figure 34).

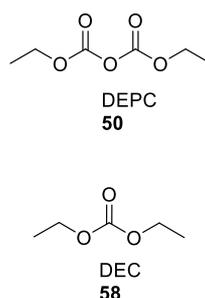


Figure 34 Diethyl pyrocarbonate and diethyl carbonate

Following the same procedure, but starting with ethyl chloroformate (**59**), gave a colourless oil in 49 % yield (Figure 35).

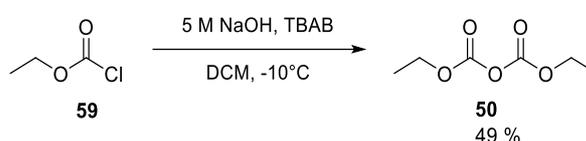


Figure 35 Synthesis of DEPC (**50**) using the methodology from Plusquellec<sup>187</sup>

As seen in Figure 36, <sup>1</sup>H NMR spectroscopy revealed two signals at  $\delta$  4.33 (4H, q,  $J=7.1$ ) and  $\delta$  1.37 (6H, t,  $J=7.1$ ) for synthesised DEPC (**50**) which differed from those obtained for the commercial sample of DEC (**58**). Additionally, <sup>13</sup>C NMR spectroscopy also showed the carbonyl carbon signals for purchased DEPC (**50**) and synthesised DEPC (**50**) were comparable ( $\delta$  148.7 and  $\delta$  148.9 respectively), whilst that for DEC (**58**) was  $\delta$  155.1 (Figure 36). This showed that DEPC (**50**) was successfully synthesised as opposed to DEC (**58**).

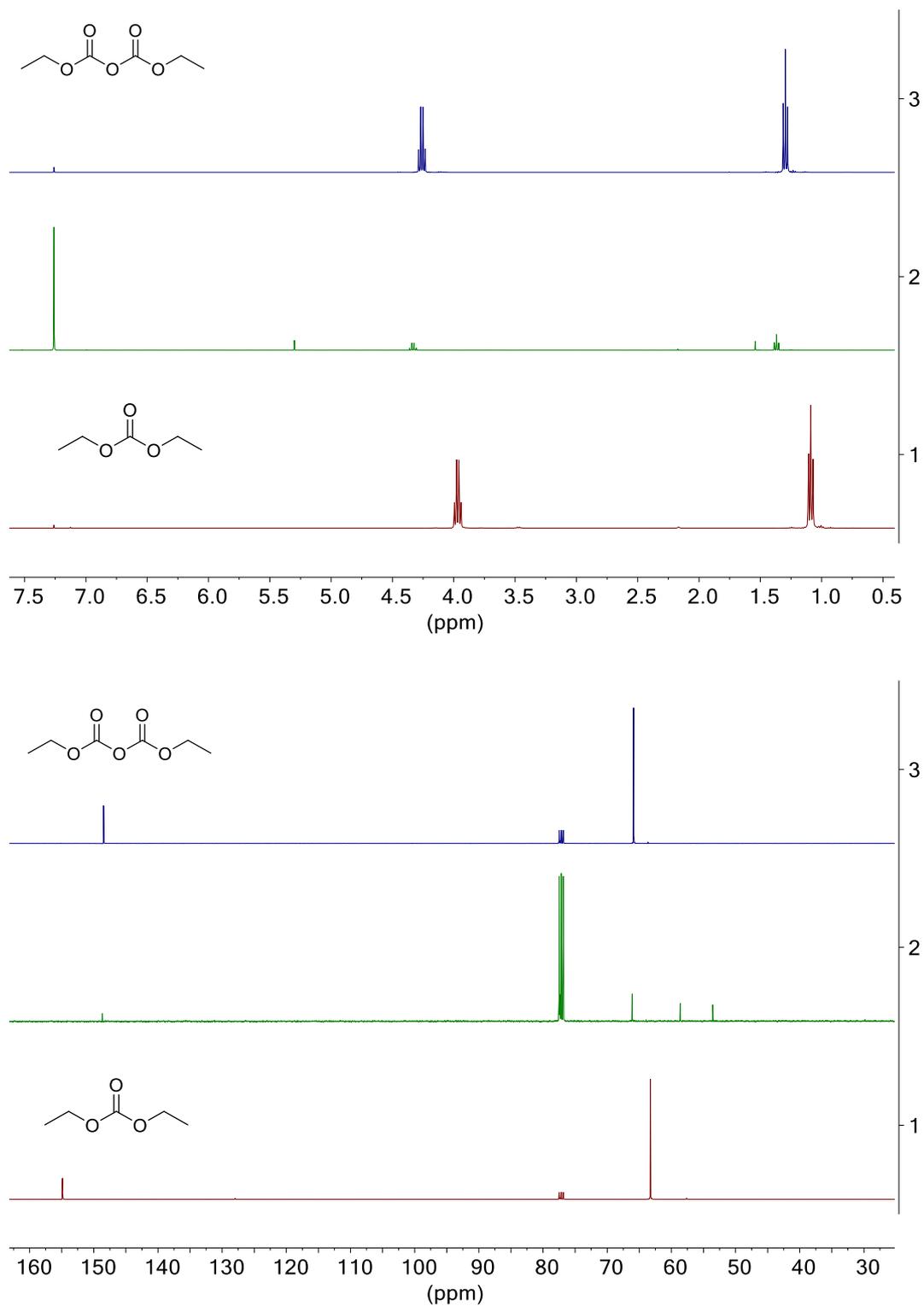


Figure 36 Alignment of the synthesised DEPC with the commercially available DEPC (50) and DEC (58)  
 Blue (top): commercially available DEPC (50); Green (middle): synthesised DEPC; Red (bottom): commercially available DEC (58).

Further support to this conclusion came from MS analysis. Both synthesised and commercial DEPC (**50**) produced a signal at  $m/z$  47 (EtOH+H)<sup>+</sup>, likely due to hydrolyses on the mass spectrometry column. Moreover, a GC-MS chromatogram was produced for DEC (**58**), gave a signal at  $m/z$  139 [M+H]<sup>+</sup>. The collected DEC (**58**) data was supported by literature reports, which affirms <sup>13</sup>C NMR  $\delta$  155.5 for the carbonate carbon and the same mass of  $m/z$  139 [M+H]<sup>+</sup>.<sup>189</sup> With the dicarbonate synthesis methodology (Figure 30) validated with the successful synthesis of DEPC (**50**), this supported the hypothesis that DPPC (**52**) was synthesised as predicted.

The next hypothesis for the lack of DPPC (**52**) reactivity to test was that benzimidazole (**55**) was not a good primitive model for screening histidine residue activity. To ensure that using an imidazole containing compound is a viable model, the reaction between benzimidazole (**55**) and DEPC (**50**) was tested (Figure 37).

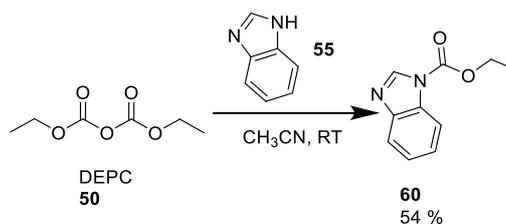


Figure 37 Acylation of benzimidazole with DEPC (**50**)

This reaction resulted in the addition of two aliphatic peaks, which can be seen in the <sup>1</sup>H NMR as a quartet at 4.56 (2H, q,  $J$  = 7.1 Hz) and a triplet at 1.51 (3H, t,  $J$  = 7.1 Hz) arising from the ethyl group. In the <sup>13</sup>C NMR spectroscopy, the carbonyl carbon is present at  $\delta$  149.6. This, in addition to an LC-MS producing a mass at  $m/z$  191 [M+H]<sup>+</sup> suggests that the carbamate (**60**) was synthesised and concluded that benzimidazole can be used to screen histidine reactivity.

4.56 (2H, q,  $J$  = 7.1 Hz, CH<sub>2</sub>), 1.51 (3H, t,  $J$  = 7.1 Hz, CH<sub>3</sub>)

Collectively, these experiments indicated that DPPC (**52**) was surprisingly stable with respect to nucleophilic acylation. Consistent with these observations, a survey of the literature revealed that the reaction of DPPC (**52**), prepared in the same manner (Figure 30), with dihydrazine (**61**) required refluxing at around 80 °C and a considerable amount of time (Figure 38).<sup>188</sup> With the increased nucleophilicity of the primary amine on the

hydrazine (**61**), and harsh conditions required to react with DPPC (**52**) eludes to the inherent stability.<sup>190</sup> Implementing the described conditions for the acylation of DPPC (**52**) would not be ideal in biological conditions, therefore it was determined that this molecule was not a suitable probe.

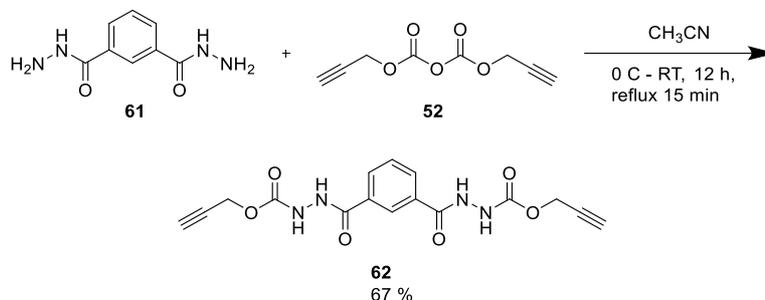


Figure 38 Addition elimination of DPPC with hydrazine from Huber and Diaz<sup>188</sup>

Although DPPC (**52**) was successfully synthesised, the reasons for the observed stability are not obvious.

### 2.3 A new probe design

Given the high reactivity of dicarbonates, it was suggested that the propargylic structure of DPPC (**52**) accounted for the stability. To overcome this idea, the probe was adapted to have 6 carbons either side of the dicarbonate group (Figure 39). These extra carbons distance the alkyne from the dicarbonate functional group, preventing any potential interaction with the dialkyl group. An additional benefit from having a longer probe was that once labelled, the alkyne group would be more likely to protrude from the active site, enhancing the chance of conjugation with a biotinylated azide via Cu(I)-catalyzed azide–alkyne cycloaddition (CuAAC), albeit at the risk of decreasing the solubility of the molecule in H<sub>2</sub>O due to the longer alkyl chain.

To test this hypothesis, access to hexargyl chloroformate (**65**) was required. As the chloroformate (**65**) was not commercially available, it was prepared following a procedure established by Pasquato. In this, triphosgene (**64**) and 5-hexyn-1-ol (**63**) dissolved in DCM were stirred for 3 h at room temperature (Figure 39).<sup>191</sup> The resulting chloroformate (**65**) afforded the derived dicarbonate (**66**) using the same conditions previously established.

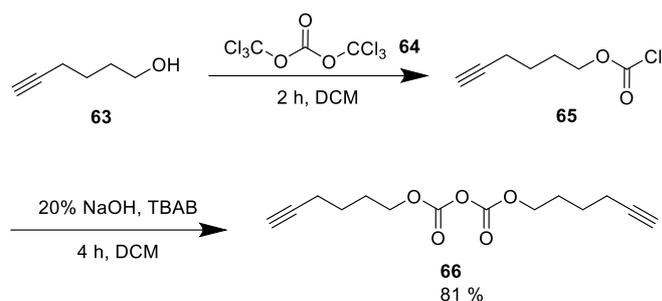


Figure 39 Synthesis of bis(5-hexyn-1-ol) dicarbonate (**66**)

The data for bis(5-hexyn-1-ol) dicarbonate (BHDC, **66**) showed the characteristic O-CH<sub>2</sub> protons being present in <sup>1</sup>H NMR spectroscopy at  $\delta$  4.34 (4H, dt,  $J$  = 8.0, 6.4 Hz) with the dicarbonate carbon present at <sup>13</sup>C NMR  $\delta$  148.0.

### 2.3.1 BHDC imidazole reactivity

The newly designed probe was reacted with benzimidazole (**55**) under the same conditions established for DEPC (**50**) (Figure 40).

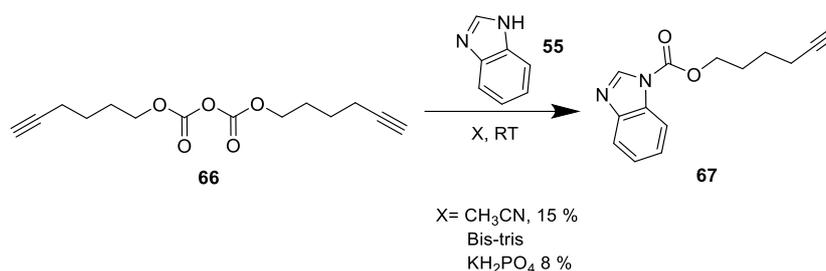


Figure 40 Nucleophilic acylation of BHDC with benzimidazole

Pleasingly, the longer carbon chain increased the reactivity of the dicarbonate **66** to benzimidazole (**55**) leading to the successful isolation of **67** as a white solid, albeit in low yield. After purification by flash chromatography (hexane and EtOAc, 0-100 %), analysis of <sup>1</sup>H NMR spectroscopy of **67** showed the appearance of 4 aromatic peaks, one most notably at  $\delta$  8.05 corresponding to the proton on carbon 2, with the addition of a distinct aliphatic peak at  $\delta$  4.56 (2H, t,  $J$  = 6.5 Hz) corresponding to O-CH<sub>2</sub>. In addition, the expected mass of  $m/z$  243 [M+H]<sup>+</sup> was identified.

To ensure that the probe will still react under biological aqueous conditions, the experiment was repeated with H<sub>2</sub>O as the solvent to ascertain that the competing reaction

by the solvent would not interfere with any protein labelling. This could prevent the tagging of histidine through hydrolysis to 5-hexyn-1-ol (**63**). Following this reaction, the spectroscopy data identified  $m/z$  243  $[M+H]^+$ . However, the reaction was inhibited by the presence of tris (20 mM), which is also reported to decrease DEPC (**50**) half-life.<sup>179</sup> These series of experiments provide evidence that the desired reaction is likely to occur under the conditions in a biological system, without tris.

Following the synthesis of **67**, further evidence was desired to ensure that BHDC (**66**) will bind to histidine itself. It was proposed that a dipeptide model would be a suitable model to achieve this. The dipeptide was synthesised from a racemic mixture of Fmoc His(trt)-OH (**68**) and the simplest amino acid glycine methyl ester.HCl (**69**). As chirality is not predicted to hinder the use of the dipeptide as a histidine model, **68** was kept as a stereoisomer.

Due to the wide variety of coupling reagents in the literature, a few conditions were tested to optimise this reaction. The initial condition used to form the amide bond were the standard coupling reagents EDC.HCl, HOBt and DIPEA achieving 76 % after purification by column chromatography. With further exploration, it was found that the optimal conditions for this condensation was using DMAP and EDC.HCl. These produced the pure product as a white solid after workup, achieving 96 % yield (Figure 41).

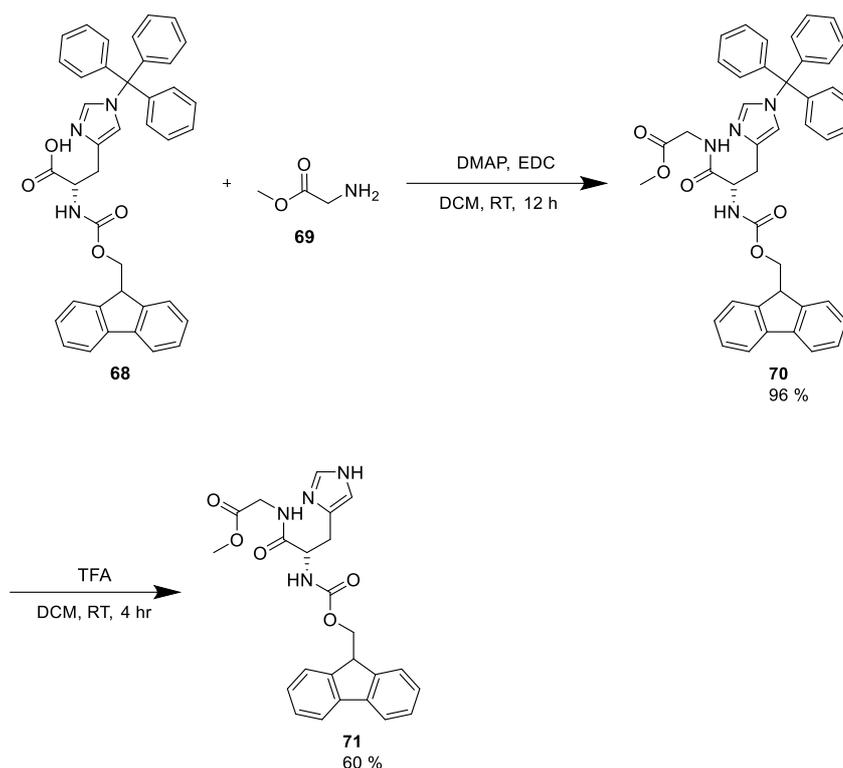


Figure 41 Synthesis of histidine dipeptide to assess BHDC (**66**) binding

The appearance of the methoxy CH<sub>3</sub> signal in the <sup>1</sup>H NMR spectrum  $\delta$  3.71 (3H, s) along with a peak at *m/z* (LCMS ES<sup>+</sup>) 691 [M+H]<sup>+</sup> confirmed the synthesis of **70**. The deprotection of the trityl group with TFA then afforded **71** as a white solid following trituration with Et<sub>2</sub>O (Figure 41). Following the synthesis of **71**, this was then used to test BHDC (**66**) reactivity against histidine (Figure 42).

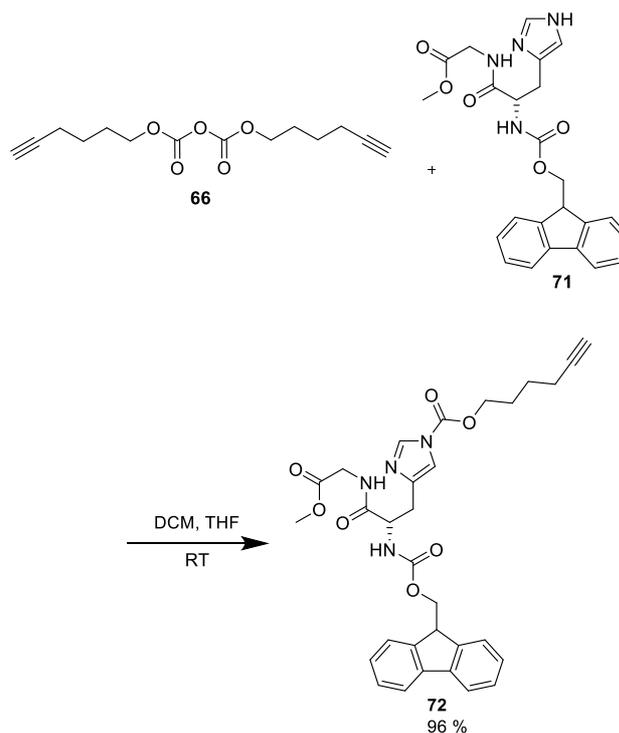


Figure 42 Acylation of imidazole on histidine with bis(5-hexyn-1-ol) dicarbonate

This led to the successful synthesis of the derived adduct **72** as ascertained from the  $^1\text{H}$  NMR spectrum which showed the imidazole proton  $2^{\text{III}}\text{-H}$  at  $\delta$  8.17 and the  $\text{CH}$  at  $\delta$  1.98 (1H, t,  $J = 2.1$  Hz). In addition, there was a peak in the LC-MS trace with  $m/z$  ( $\text{ES}^+$ ) 573  $[\text{M}+\text{H}]^+$ . Collectively, this was promising evidence that **66** will react with histidine in a protein.

To ensure that this was an accurate representation of histidine binding, this process was repeated under the same conditions using DEPC (**50**), instead of BHDC (**66**, Figure 43).

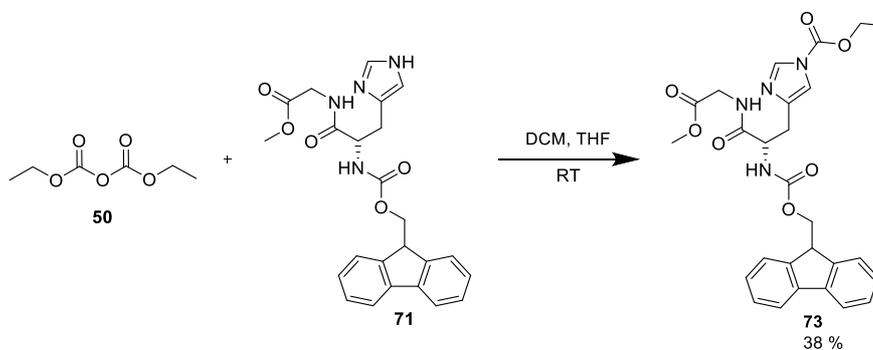


Figure 43 Acylation of histidine imidazole with known histidine inhibitor DEPC (**50**)

This reaction successfully produced **73**, albeit in a low yield. In accordance with this, the  $^1\text{H}$  NMR spectrum showed a peak for the ethoxy  $\text{CH}_2\text{CH}_3$  group present at  $\delta$  4.45 (2H, q,  $J = 7.1$  Hz) and  $\delta$  1.41 (3H, t,  $J = 7.1$  Hz) respectively, with the newly formed carbonate group at  $\delta$  148.2 in the  $^{13}\text{C}$  NMR spectrum.

After establishing the leading probe candidate, BHDC (**66**) by identifying histidine reactivity through a simple dipeptide model, the need to validate the formed carbamate stability remained. This led to the development of an azide containing linker with an NMR signal for analysis of the CuAAC reaction.

#### 2.4 Stability of the carbamate

As discussed in Chapter 1.4, the basic components for activity-based protein profiling (ABPP) are the warhead, linker, and the reporter group.<sup>104</sup> With a leading warhead designed and tested in a model, the next challenge was to ensure all the components are functioning in the predicted manner. Therefore, the click chemistry component was also tested under more controlled, chemical environment. For this, a soluble PEG linker was chosen to prevent solubility issues, with a fluorinated benzyl group to serve as an NMR signal.

To begin, a simple polyethylene glycol (PEG, **74**) linker was synthesised (Figure 44). The methodology initially followed that published by Brauch *et al.* (2012).<sup>192</sup> This involved initial mesylation of the alcohol groups followed by a  $\text{S}_{\text{N}}2$  substitution with  $\text{NaN}_3$  under basic, aqueous conditions resulting in the diazido PEG (**75**). A Staudinger reaction then afforded compound **76** as a colourless oil, with 77 % yield (Figure 44).

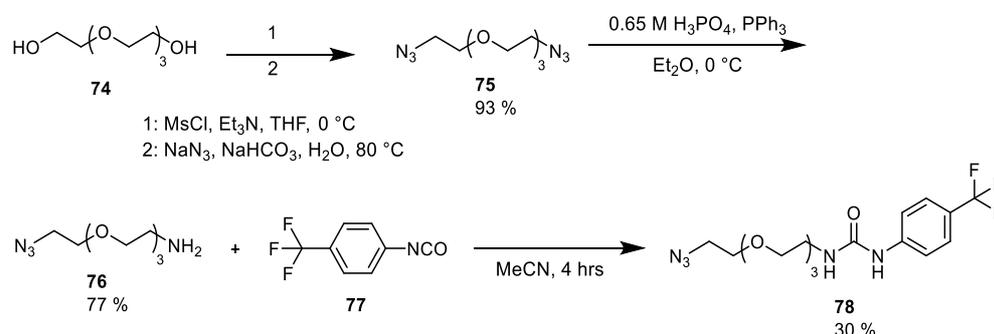


Figure 44 Synthesis of a PEG linker

**76** was then stirred with 4-(trifluoromethyl)benzyl isocyanate (**77**) to produce a fluorinated urea (**78**, Figure 44) with 30 % yield. This was confirmed through the LC-MS showing the expected mass of  $m/z$  ( $ES^+$ ) 406  $[M+H]^+$ . In addition, the  $^1H$  NMR with a peak at  $\delta$  4.34 – 4.29 (2H, m, HN-CH<sub>2</sub>) and 3.36 (2H, t,  $J = 5.1$  Hz, CH<sub>2</sub>-N<sub>3</sub>), the urea carbon present at  $^{13}C$   $\delta$  153.3. Finally, the  $^{19}F$  NMR showing a single peak at  $\delta$  -61.98. CF<sub>3</sub>, which was ideal for use as an NMR signal for future analysis. After synthesis was confirmed, the urea (**78**) was then used to test the Cu(I)-catalyzed azide–alkyne cycloaddition (CuAAC) with the acylated histidine dipeptide (**72**).

#### 2.4.1 Copper-catalyzed azide-alkyne cycloaddition (CuAAC)

To ensure that the click chemistry will successfully form the 1,2,3-triazole necessary for biological analysis, CuAAC was tested. The fluorinated linker (**78**) and probe acylated histidine (**72**) was subsequently added to an aqueous solution containing CuSO<sub>4</sub>·5H<sub>2</sub>O, and Na ascorbate (Figure 45).

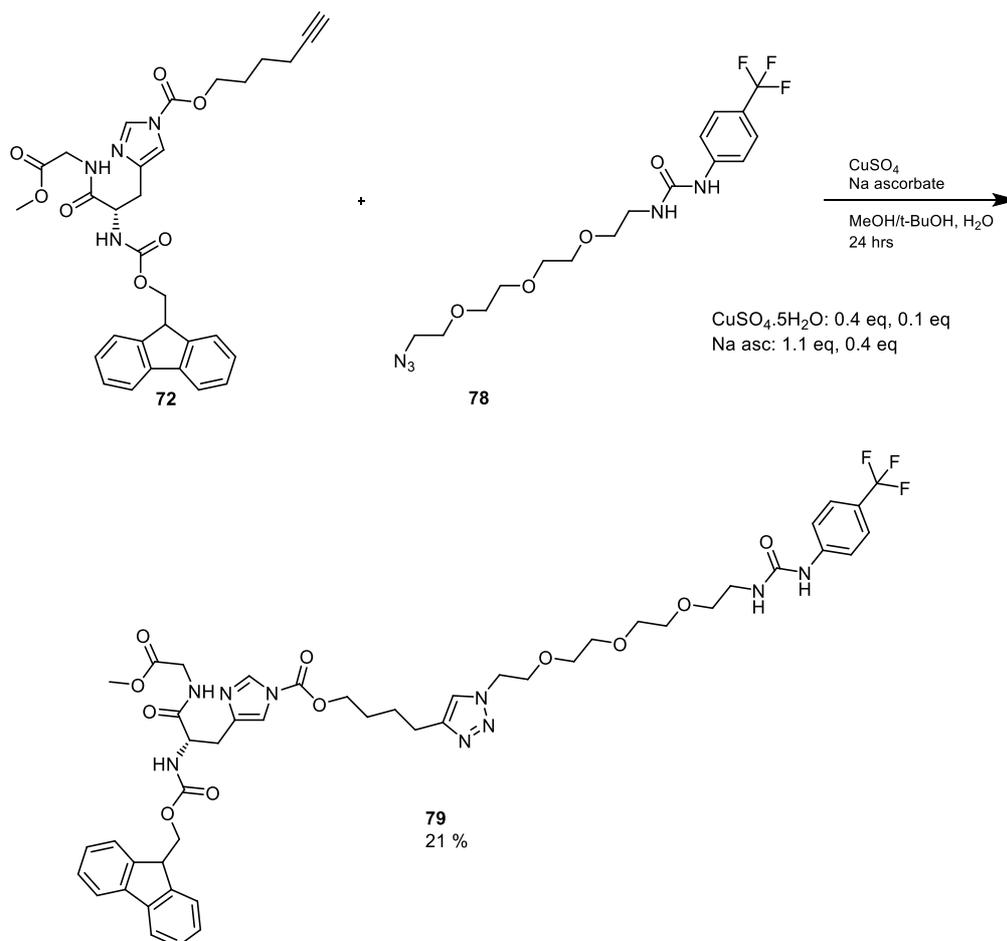


Figure 45 Synthesis of 1,2,3-triazole using CuAAC conditions

Formation of the 1,2,3-triazole using click chemistry initially appeared successful using LC-MS, with a small amount of product at  $m/z$  978  $[\text{M}+\text{H}]^+$  after 12 h. Along with this mass, an additional mass could be seen at  $m/z$  562, which was an unknown product.

Due to the poor conversion of starting material, seen from LC-MS analysis, the reaction was heated to 30 °C and stirred for an additional 12 h. After the addition of heat, the peak at  $m/z$  562 grew in intensity suggesting that the formation of this by-product was favoured in these adjusted conditions. From this reaction mixture, methyl ester **80** was isolated (Figure 46).

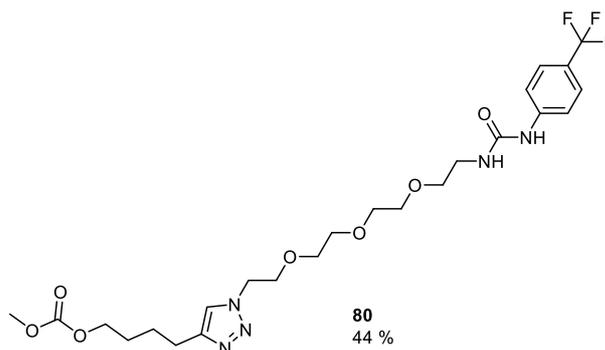


Figure 46 Product formed after methanolysis on the carbamate after CuAAC

Evidence for this structure was obtained from the  $^1\text{H}$  NMR spectrum, which contained a characteristic triazole proton at  $\delta$  7.47 (1H, s), along with a methyl group at  $\delta$  3.76 (3H, s) and the PEG chain at  $\delta$  3.64–3.55 (8H, m) in addition to the  $^{13}\text{C}$  NMR spectrum having a carbonate signal at  $\delta$  155.8. These signals in conjunction with no evidence of the imidazole or Fmoc groups led to the conclusion the product was formed was that in Figure 46. This by-product matches the described mass of  $m/z$  562 observed when monitoring the reaction progression. The hypothesised formation of **80** was due to methanolysis of the carbamate.

Following the failure of the CuAAC, the conditions were adjusted to remove methanol from the reaction, instead replacing it with the less nucleophilic *t*-BuOH. These new conditions led to the successful production of the target compound **79**. This was supported though the  $^1\text{H}$  NMR producing signals for the triazole at  $\delta$  7.48 (1H, s), along with the Fmoc signals at 7.75 (2H, d,  $J = 7.6$  Hz) and 7.38 (2H, t,  $J = 7.6$  Hz) in addition to the glycine signals at 3.69 (3H, s) and 4.00 (2H, q,  $J = 18.2$  Hz). Corroborating with the  $^1\text{H}$  NMR, the carbamate signal at  $^{13}\text{C}$  NMR  $\delta$  148.6 in the spectrum suggested that the desired molecule was synthesised, with further confirmation found in mass spectrometry, which produced a mass which concurs with the hypothesised mass  $m/z$  978  $[\text{M}+\text{H}]^+$  and the  $^{19}\text{F}$  NMR produced one peak at  $\delta$  -61.93. This data strongly suggests that the desired molecule was synthesised.

## 2.5 Alternative dipeptide synthesis

Due to the aforementioned difficulties, it was decided that replacing the Fmoc group would ease analysis of the product. Therefore, the dipeptide model was adjusted, replacing the Fmoc group with 4-fluorobenzoyl chloride (**82**). This fluorinated heterocycle is advantageous over the Fmoc for easier and faster NMR analysis after the CuAAC reaction.

Although the synthetic route initially appeared facile, there were difficulties with the 4-fluorobenzyl chloride (**82**) which repeatedly generated low yields due to the complex mixture formed (Figure 47).

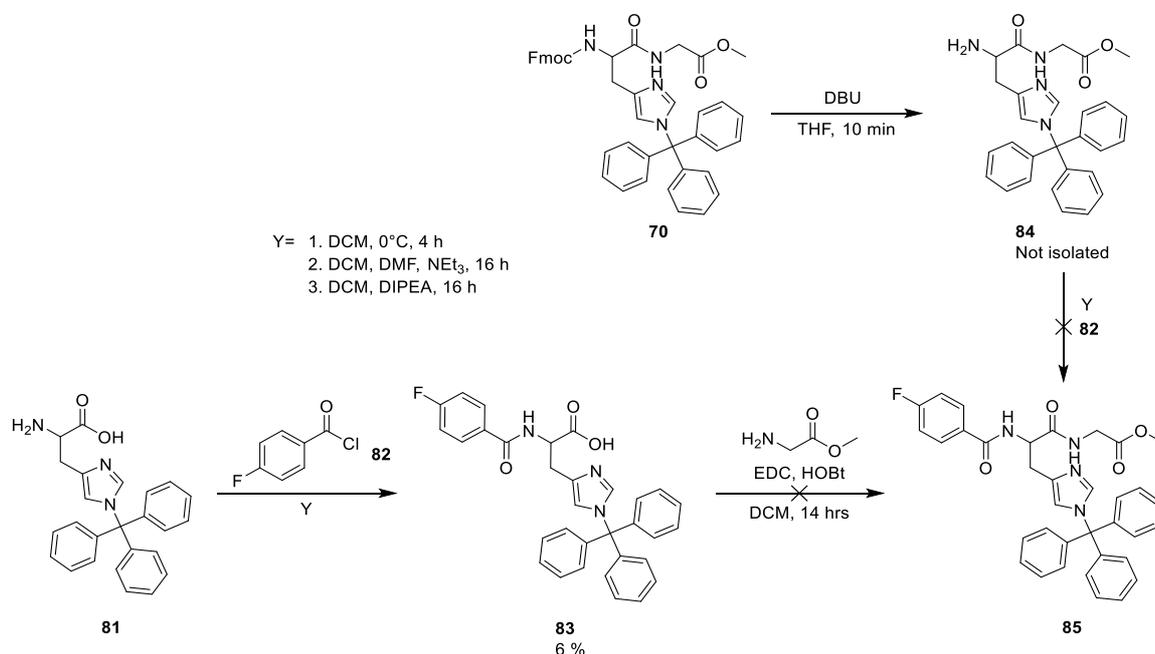


Figure 47 Initial synthetic route for the fluorinated dipeptide

The predominant issue with the conditions for the addition elimination with 4-fluorobenzoyl chloride (**82**) was the generation of 7-14 by-products when reacted with His(trt)OH (**81**). After screening conditions for the production of **83**, the highest isolated yield was 6%. Following the coupling with glycine ME (**69**), no isolatable product (**85**) could be found, which was attributed to the low quantity of starting material **83**. Due to this, a different route was attempted starting with an Fmoc deprotection of the previously synthesised dipeptide **70** (Figure 47). A one pot DBU deprotection followed by the addition of 4-fluorobenzoyl chloride (**82**) resulted in a mixture of compounds from which, the product could not be isolated (**85**).<sup>193</sup> Whilst there is literature precedent for alternative methods to synthesise the aromatic amide bond, such as using TPP as a condensing agent with 4-DMAP, it was decided a different method should be attempted altogether.<sup>194</sup>

To replace the problematic addition elimination, 4-fluorobenzoic acid (**86**) was used. This starts with the previously described one pot method for the DBU deprotection of Fmoc on **70**. After the deprotection was confirmed by TLC, the solvent was removed *en vacuo*, and

the resulting mixture underwent coupling conditions leading to **87** with 61 % crude yield. This crude material then underwent the deprotection of trityl, in which **88** was successfully purified with 40 % yield (Figure 48).

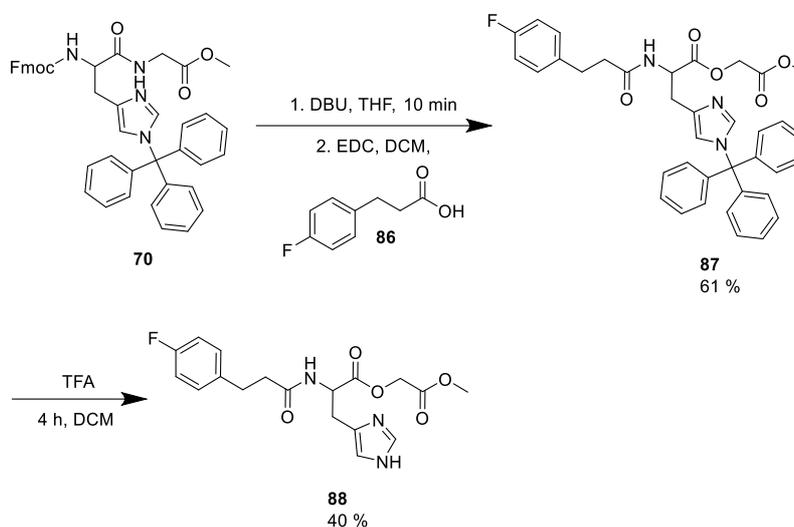


Figure 48 Production of fluorinated dipeptide

This was evident through the lack of trityl group present in the  $^1\text{H}$  NMR, and the corresponding mass  $m/z$  377  $[\text{M}+\text{H}]^+$ . With **88** synthesised, this was acylated with BHDC (**66**) and then the CuAAC reaction was again attempted (Figure 49).

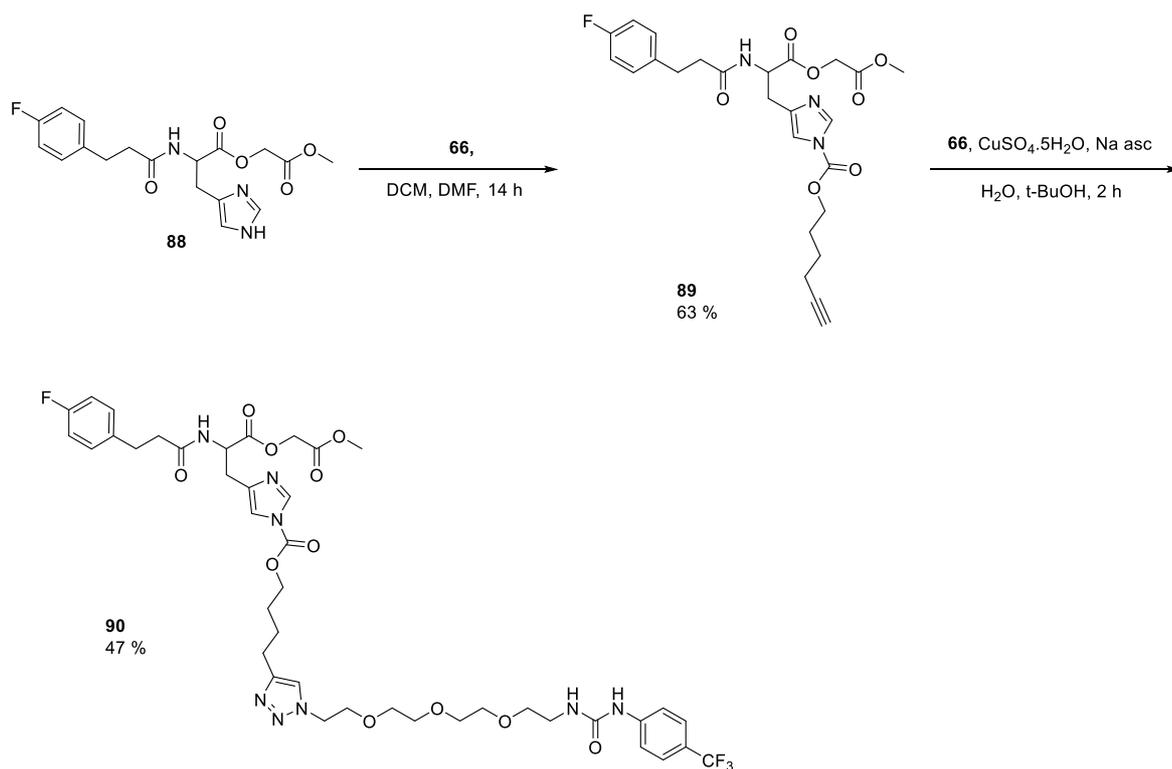


Figure 49 Validation of CuAAC suitability with a difluorinated compound

Using these conditions, the click chemistry reaction successfully produced **90** as a colourless oil within 2 h. The fluorine NMR produced 2 signals with the expected ratio of 3:1 at  $\delta$  -61.95 ( $\text{CF}_3$ ) and  $\delta$  -116.99 (CF) respectively. Analysis via mass spectrometry obtained  $m/z$  907  $[\text{M}+\text{H}]^+$ , corresponding to **90**. Further confirmation from  $^1\text{H}$  NMR displayed the triazole proton at  $\delta$  7.50 (1H, s), imidazole protons at  $\delta$  8.12 (1H, s, 2-*H*) and  $\delta$  7.22 (1H, s, 5-*H*) with the carbamate carbon also visible at  $^{13}\text{C}$   $\delta$  148.3.

These series of experiments led to the confidence that the carbamate group will form in the presence of active site histidine residue and remain attached to the protein during the click chemistry reaction. Important details were elucidated from this series of experiments, such as confirmation of the instability with other nucleophilic elements, such as methanol.

## 2.6 Fluorescent marker and binding group

Before biological testing of BHDC (**66**) could begin, a series of chemical tools were synthesised to perform fluorescent analysis and protein isolation using a pulldown. These form the linker and reporter group and are common techniques utilised in ABPP, as discussed in Chapter 1.4.3. Initial synthetic focus was on the development of a simple azido

containing rhodamine, which would allow for the fluorescent analysis of proteins reactive with BHDC (**66**) in a proteome following CuAAC. Another important component of ABPP is the ability to purify the bound proteins using enrichment with streptavidin. For this series of experiments, a tetra functional biotin tag was developed, which has several advantages over the commercially available biotin linkers, discussed below.

### 2.6.1 Simple rhodamine fluorophore

To explore the inhibition of enzymes, a simple fluorophore was synthesised. This was performed as a one pot reaction following the precedent described by Navarro and Bergstrom (2014).<sup>195</sup> This involved rhodamine B (**91**) activation by NHS with EDC. The solvent was removed and to this crude material, the previously synthesised PEG (**76**, Figure 44) was added. A basic workup followed the completion of this reaction to ensure the rhodamine B remains in the fluorescent state (Figure 50).

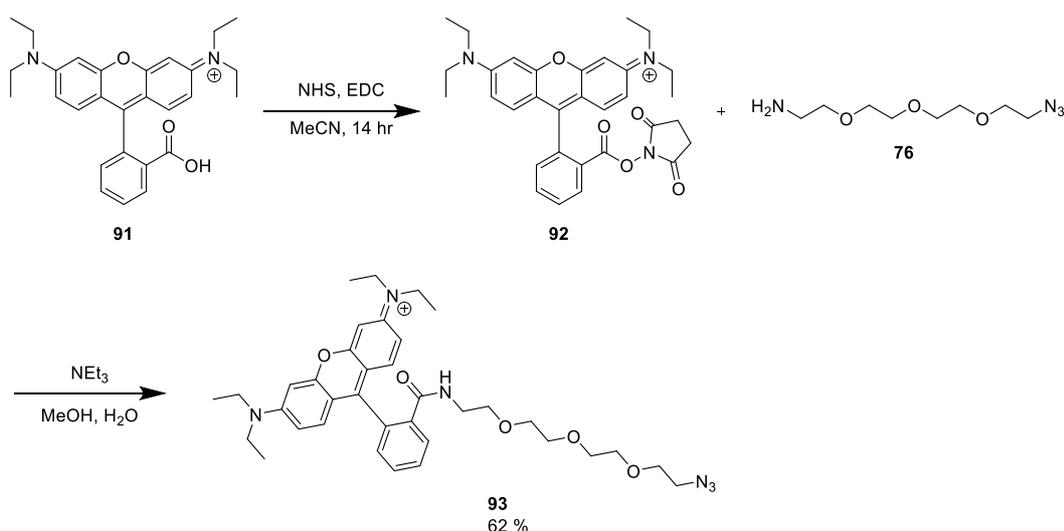


Figure 50 Synthesis of a rhodamine linker for biological analysis

The synthesis of the reporter group **93** was successful, with the fluorescence successfully being retained. The PEG chain O-CH<sub>2</sub> was visible at  $\delta$  3.67 – 3.62 (12H, m) and CH<sub>3</sub> at 1.33 (12H, t,  $J = 7.1$  Hz) with mass spectrometry producing a peak of  $m/z$  644 [M+H]<sup>+</sup>. This was implemented for the fluorescent analysis of the proteome throughout Chapter 3 and Chapter 4.

## 2.6.2 Tetrafunctional reactive linker

For the purification and elution of BHDC (**66**) modified proteins, cross coupling with a biotinylated linker can be used with streptavidin coated magnetic beads. The current commercially available biotinylated molecules range from a simple biotin linker (**94**) to ones that include a cleavable component (Figure 51).

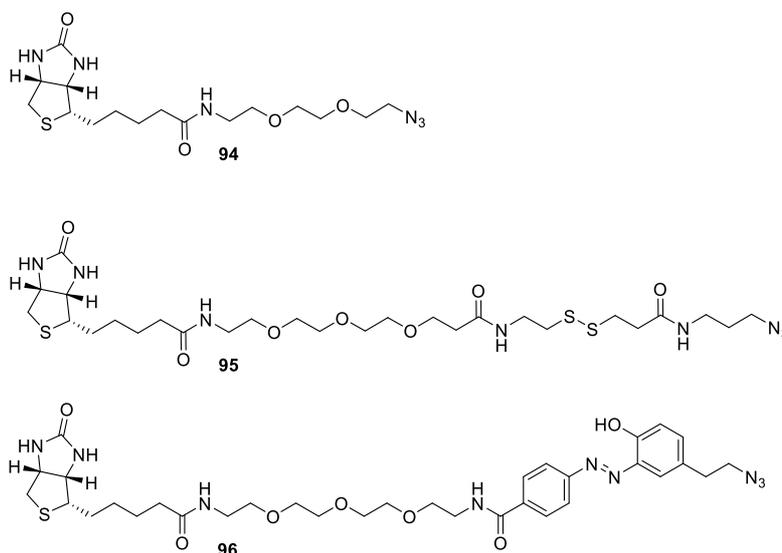


Figure 51 Market available biotinylated azides

**95** and **96** are cleavable using reducing agents like DTT and dithionite, leaving thiol and aminophenol moieties respectively.<sup>196,197</sup> However, the eluted proteins are not fluorescent thereby relying on protein staining for visualisation (Figure 51). One option for fluorescent analysis of the eluted proteins is to perform a Western blot using anti-streptavidin antibodies.<sup>100</sup> However, an alternative option is to design a unique reactive linker, encompassing all the desired characteristics; binds to streptavidin, fluorescent, bioorthogonal reactive group, and cleavable group for easy elution (Figure 52).<sup>198</sup>

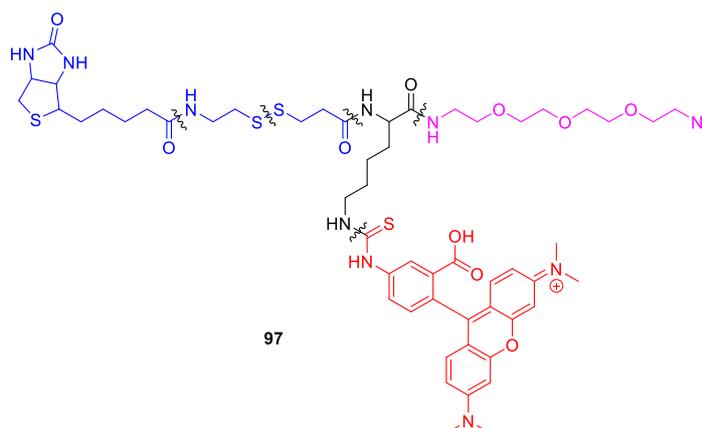


Figure 52 Structure for the trifunctional cleavable linker

The colour refers to the different main sections of the compound: blue: biotinylated component, black: lysine joining the components together, pink: PEG linker with bioorthogonal group, red: TAMRA fluorophore.

Compound **97** can be divided into three sections: azide PEG linker, lysine, and acid cleavable biotin. A lysine joins all the highlighted components in Figure 52 together. To develop this core, Fmoc-Lys(Boc)-OH (**98**) was coupled with the previously described PEG (**76**, Figure 44) using EDC and NHS (Figure 53).

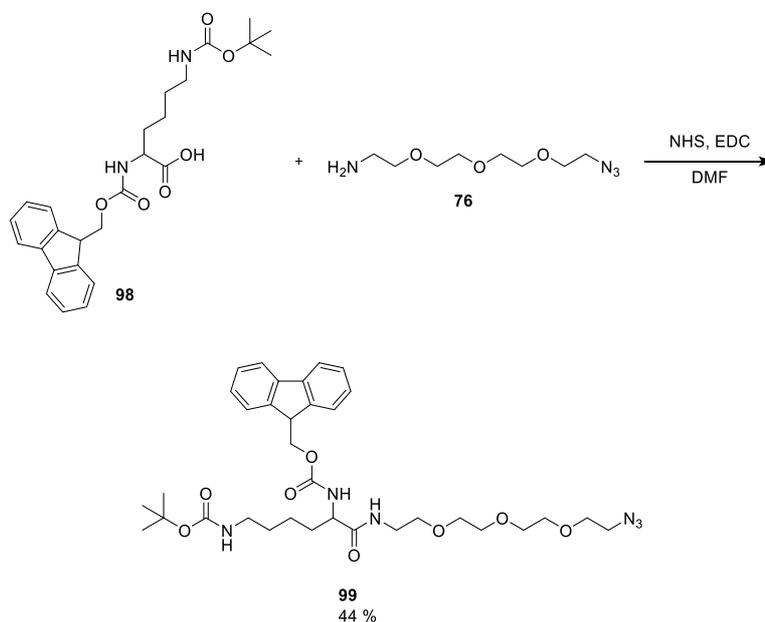


Figure 53 Coupling of lysine with azide linker

This formed (**99**) as a light-yellow oil with 43 % yield. Validation of the product formation occurred using mass spectrometry, producing a mass of  $m/z$  691  $[M+H]^+$ . Additionally,  $^1H$

NMR identified PEG CH<sub>2</sub> at  $\delta$  3.68 – 3.58 (10H, m) and CH<sub>2</sub>-N<sub>3</sub> at  $\delta$  3.36 (2H, t,  $J$  = 5.1 Hz), whilst retaining Boc (1.43, 9H, s).

Following the successful synthesis of **99**, other members of the PGS group then coupled **99** with the disulphide-biotin component (Figure 52, blue) before finally deprotecting Boc and reacting the anime with TAMRA isothiocyanate, forming the desired product **97**.

With these chemical tools synthesised, research progressed to the validation of chemoselectivity of BHDC (**66**) in biology.

### 3 Biological analysis of ABP in a model system

#### 3.1.1 Synopsis of chapter

The initial challenge was to validate BHDC (**66**) acylation of active site histidine residues in biological conditions. To address this, the probe was applied to a biochemical assay testing for inhibition of *Leishmania major* inositol phosphorylceramide synthase (*Lmj*IPCS). Due to difficulties that a membrane bound protein has for mass spectrometric analysis, a model protein with a similar HHD catalytic triad was identified to use as a model. This was used along with a histidine acid base residue model to explore the scope of BHDC (**66**) reactivity. Finally, the probe was used in a streptavidin enrichment pulldown, isolating an overexpressed model protein and intrinsic enzymes from *E. coli*.

#### 3.2 *Lmj*IPCS inhibition assay

Following the synthesis and chemical validation of a leading candidate, bis(5-hexyn-1-ol) dicarbonate (BHDC, **66**) in a dipeptide model, the activity in biological conditions and selectivity needed exploration. To achieve this, the reactivity of **66** with active histidyl residue in protein models was assessed.

As previously described in Chapter 1.2.2, *Leishmania major* IPCS (*Lmj*IPCS). is an enzyme of interest, with evidence suggesting the enzyme is essential and thus, a good drug target.<sup>73</sup> This active site is hypothesised to contain a HHD catalytic triad, with nucleophilic His-264 catalysing the transfer of the phosphate from the PI (**21**) to the ceramide (**17**) to form IPC (**23**). As such, it was proposed that the treatment of *Lmj*IPCS with BHDC (**66**) could lead to the inhibition of this enzyme through covalent labelling of the active site His-264 (Figure 54).

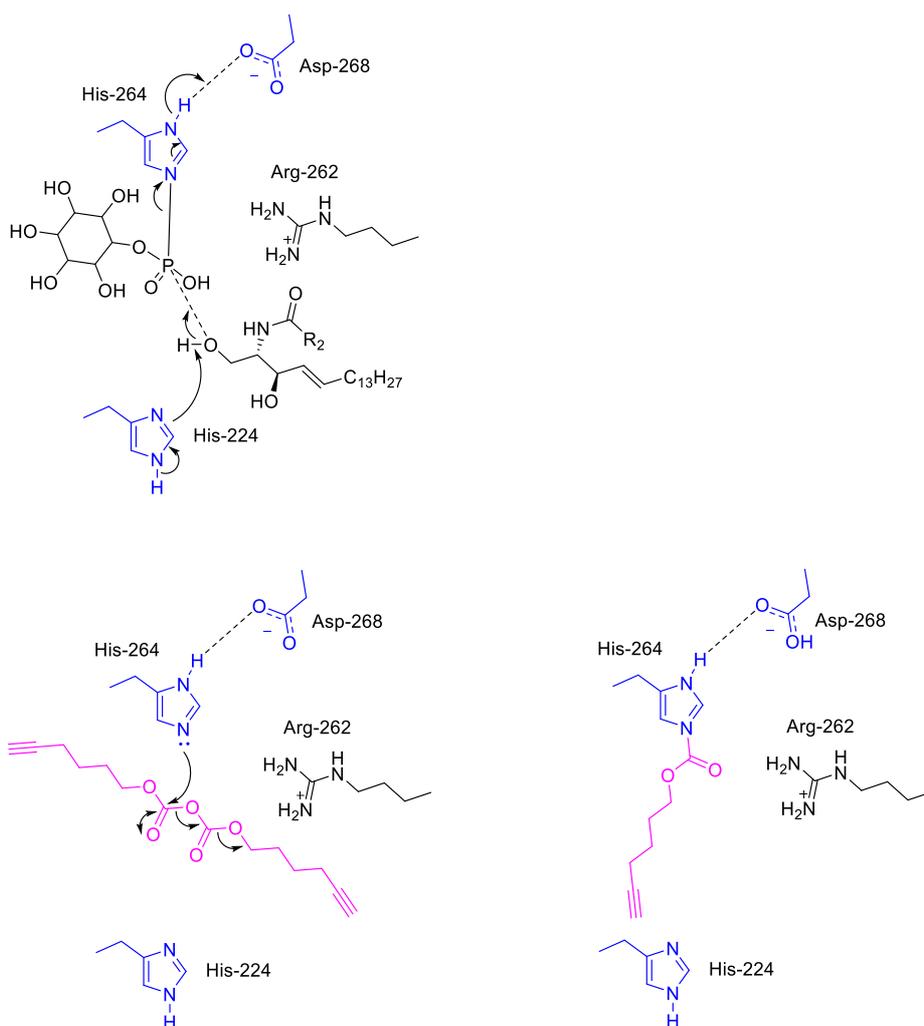


Figure 54 Hypothesised inhibition of *LmjIPCS* active site with BHDC (**66**)  
 Blue: active site catalytic triad residues (HHD); Black: *LmjIPCS* substrates and Arg-262; Pink: BHDC (**66**)

Evaluating inhibition of *LmjIPCS* was possible by reconstituting the established biochemical assay discussed in Chapter 1.2.2. For this, enriched microsomes prepared by Dr John Mina was utilised, which were isolated from *S. cerevisiae* Golgi apparatus expressing *LmjIPCS*.<sup>92</sup> This assay monitors the conversion of NBD-C<sub>6</sub>-ceramide (**100**) to NBD-C<sub>6</sub>-IPC (**101**), exploiting the different retention times with HPTLC to identify conversion (Figure 55). If the probe interacts with the enzyme, then the production of NBD-C<sub>6</sub>-IPC (**101**) will be inhibited.

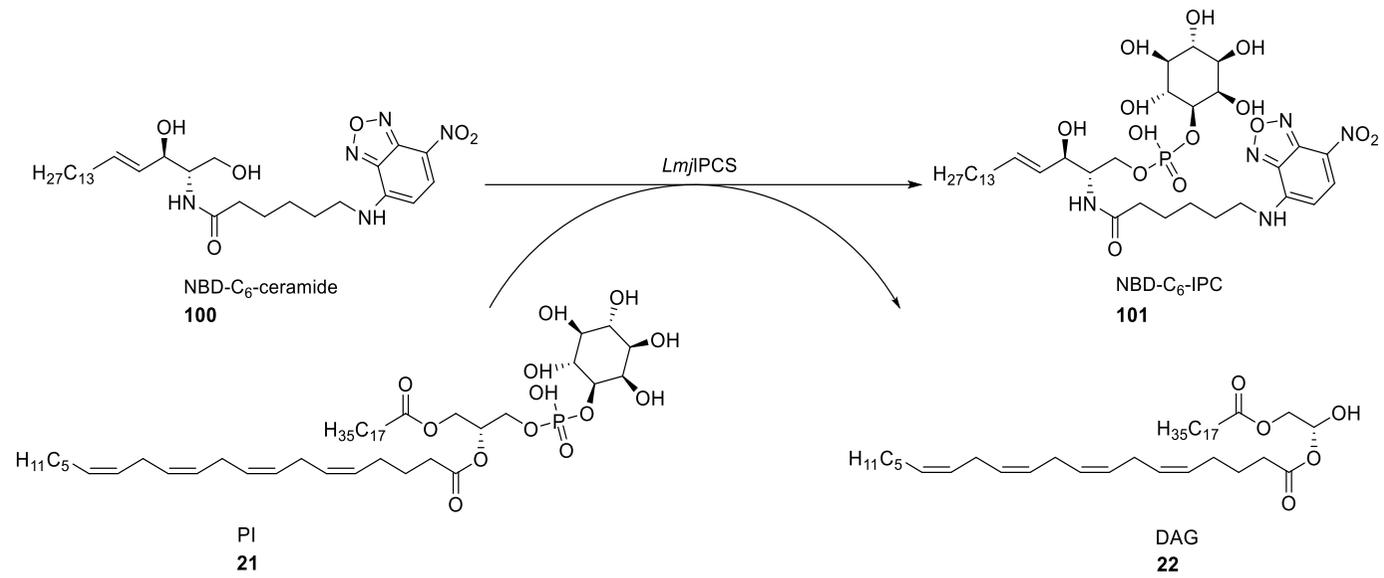


Figure 55 NBD-C<sub>6</sub>-ceramide (**100**) conversion to NBD-C<sub>6</sub>-IPC (**101**)

## 3.2.1 Initial test for inhibition

To begin, the *Lmj*IPCS enriched microsomes were incubated with the desired compounds for 1 h before the addition of the substrates PI (**21**) and NBD-C<sub>6</sub>-ceramide (**100**). The initial biochemical assay tested BHDC (**66**) at 100 μM and the controls 5-hexyn-1-ol (**63**, 100 μM) and DEPC (**50**, 100 μM).

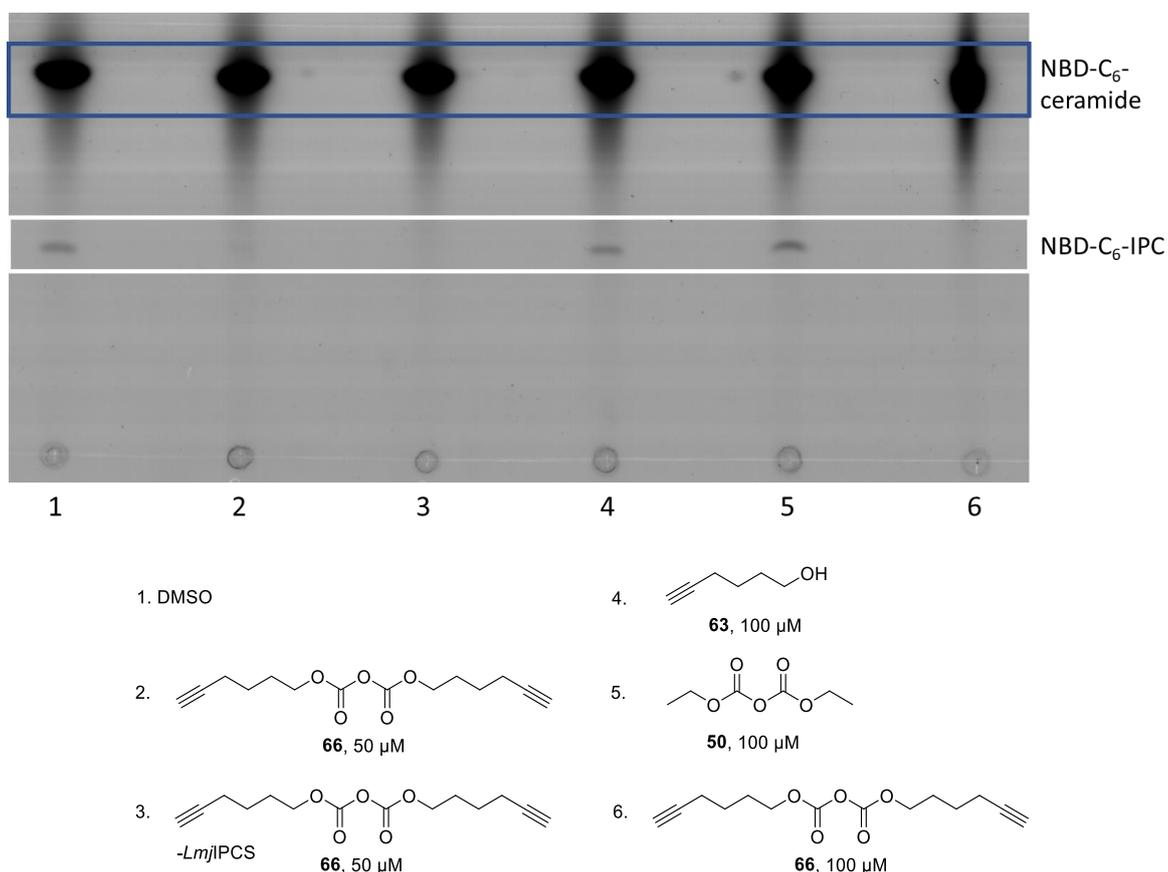


Figure 56 *Lmj*IPCS assay 1

High performance thin layer chromatography (HPTLC) separation of NBD-C<sub>6</sub>-ceramide (top) from NBD-C<sub>6</sub>-IPC (bottom) using the solvent system 55: 45: 10 MeOH, CHCl<sub>3</sub> and (0.25 % KCl) H<sub>2</sub>O. The HPTLC plate was visualised on Fuji FLA-3000. 1: 1 μl DMSO; 2: BHDC (**66**, 50 μM); 3: BHDC (**66**, 50 μM) without *Lmj*IPCS; 4: 5-hexyn-1-ol (**63**, 100 μM); 5: DEPC (**50**, 100 μM); 6: BHDC (**66**, 100 μM). All the tested compounds were dissolved in DMSO. Representative HPTLC from 3 consistent replicates.

As hypothesised, BHDC (**66**) inhibited NBD-C<sub>6</sub>-IPC production (Figure 56, lanes 2), without itself visibly interacting with the substrate (Figure 56, lane 3). By-product 5-hexyn-1-ol (**63**) did not appear to inhibit the enzyme (Figure 56, lane 4), which does suggest that the effect seen with BHDC (**66**) is due to the dicarbonate. A surprising result from this assay was that DEPC (**50**) did not appear to inhibit *Lmj*IPCS (Figure 56, lane 5). The lack of activity from

DEPC (**50**) was surprising as it is reported to bind to active site histidine residues (Chapter 1.7.1), therefore, DEPC (**50**) was invalidated as a positive control.

The initial hypothesis was that due to the instability of DEPC (**50**), the stock compound had hydrolysed and was no longer functional as a histidine acylating reagent. After ruling out decomposition using new, unopened DEPC (**50**) and NMR analysis, it was concluded that DEPC (**50**) inherently does not inhibit *Lmj*IPCS. Another explanation could be due to a difference in lipophilicity, with the more polar DEPC (**50**) not being able to penetrate the lipophilic microsome. To support this, the LogP for each reagent was calculated using molinspiration (<https://www.molinspiration.com>). BHDC (**66**) was predicted to have a LogP of 2.41, ideal for crossing lipid membranes, whereas DEPC (**50**) has a LogP of 0.75. With a predicted LogP of DEPC (**50**) being close to 0, this does suggest that **50** has a reduced capacity to cross the lipid membrane.<sup>199</sup> A literature search of DEPC (**50**) modification of membrane bound proteins revealed that other research groups report that DEPC (**50**) is not able to penetrate the lipid membrane.<sup>200</sup>

To verify that lipophilicity was a key issue and develop an alternative control, a dihexyl analogue was proposed. This new inhibitor has a predicted LogP of 4.89, confirming the increased lipophilicity.

### 3.2.2 Assessing the role of lipophilicity for inhibition

The new inhibitor was successfully synthesised using the same methodology described for BHDC (**66**), however using 1-hexanol (**102**) as the starting material (Figure 57).

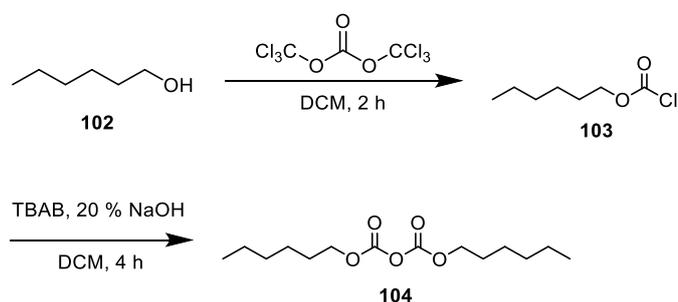
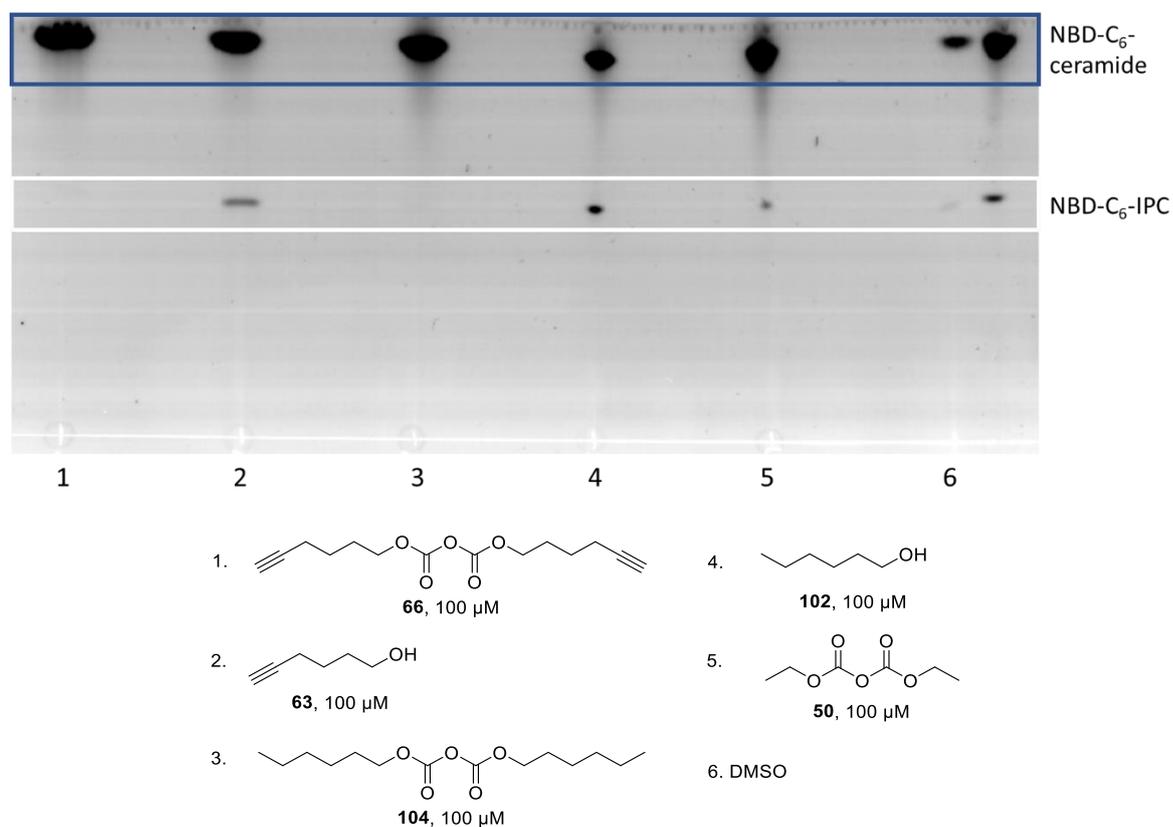


Figure 57 Synthesis of a new DEPC (50) derivative

Following the successful synthesis of **104**, this was then tested in the same biochemical assay described above, monitoring for NBD-C<sub>6</sub>-IPC (**101**) production from NBD-C<sub>6</sub>-ceramide (**100**, Figure 58).



*Figure 58 Validation of *Lmj*IPCS inhibition in a biochemical assay*  
 High performance thin layer chromatography (HPTLC) separation of NBD-C<sub>6</sub>-ceramide (**100**, top, blue rectangle) from NBD-C<sub>6</sub>-IPC (**101**, bottom, white rectangle) using the solvent system 55: 45: 10 MeOH, CHCl<sub>3</sub> and (0.25 % KCl) H<sub>2</sub>O. The HPTLC plate was visualised on Fuji FLA-3000. 1: 100 μM BHDC (**66**); 2: 100 μM 5-hexyn-1-ol (**63**); 3: 100 μM dihexyl dicarbonate (**104**); 4: 100 μM 1-hexanol (**102**); 5: 100 μM DEPC (**50**); 6: DMSO. Representative HPTLC plate over 3 repeats.

Pleasingly, as predicted and in contrast to DEPC (**50**), dihexyl dicarbonate (**104**) showed inhibition of *Lmj*IPCS (Figure 58, lane 3), supporting the hypothesis that lipophilicity plays a role in the inhibition of *Lmj*IPCS. This could indicate that **66** will be able to bind to membrane bound proteins, an advantage over DEPC (**50**).

Having demonstrated that BHDC (**66**) can inhibit *Lmj*IPCS, it was then necessary to validate the mode of action as labelling His-264. The membrane associated with this enzyme would challenge this, and consequently a soluble model enzyme with a similar HHD catalytic triad was then sought to undertake this experiment.

### 3.3 Non-specific acid phosphatase inhibition

To further assess the selectivity of BHDC (**66**) with a nucleophilic histidine, a model enzyme was chosen to avoid the additional difficulties that working with a membrane bound protein like *Lmj*IPCS (Chapter 1.2.2). A model enzyme would therefore have several advantages over the microsomal *Lmj*IPCS that was previously used to assess binding. The critical features for this model enzyme were: small, soluble, HHD catalytic triad, and a known crystal structure.

One protein that fits these narrow criteria was a non-specific acid phosphatase (NSAP) known as PhoN from *Salmonella typhumirium* (UniProt A0A0F6BA10). PhoN is a soluble cytosolic protein, with a known crystal structure (PDB: 2A96).<sup>154,156</sup> Most importantly, PhoN has an active site that proceeds though a similar mechanism as IPCS, with the additional advantage of being a soluble protein (Figure 59).<sup>155</sup>

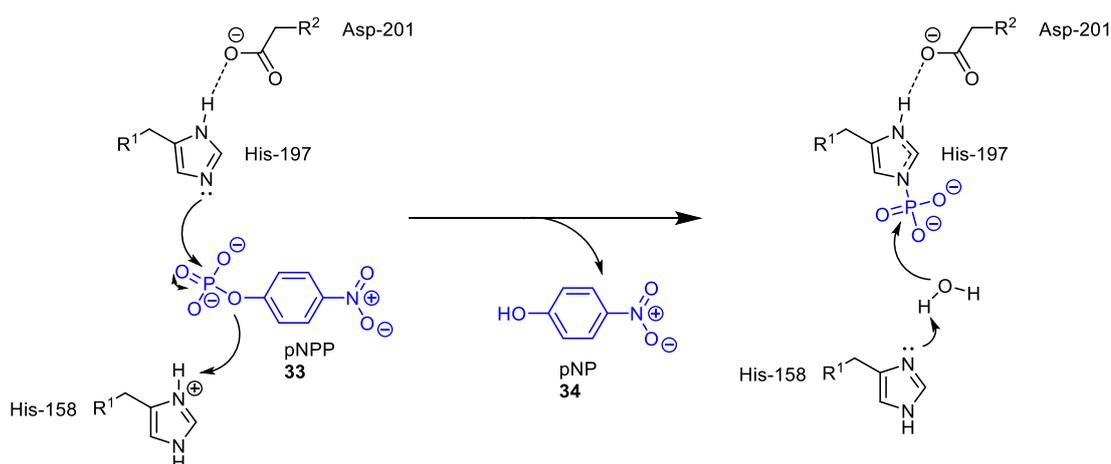


Figure 59 PhoN dephosphorylation catalytic mechanism

The substrate is *para*-nitrophenyl phosphate (pNPP), a known substrate for non-specific acid phosphatases, which releases *para*-nitrophenol (pNP) as a fluorescent product. Blue: catalytic triad residues HHD; Black: active site amino acid residues.

As discussed in Chapter 1.6.1, the electron deficient phosphorylated substrate is attacked by the nucleophilic His-197 in the active site of PhoN, forming a phosphohistidine intermediate. The acid base His-158 then facilitates the release of the product (Figure 59).<sup>156</sup> The hypothesised mechanism for *Lmj*IPCS also results in the formation of the phosphohistidine intermediate with His-264, and the acid base His-224 deprotonates the primary alcohol on the substrate ceramide. The ceramide anion then attacks the phosphate, breaking the phosphohistidine bond (Figure 60).<sup>84</sup>

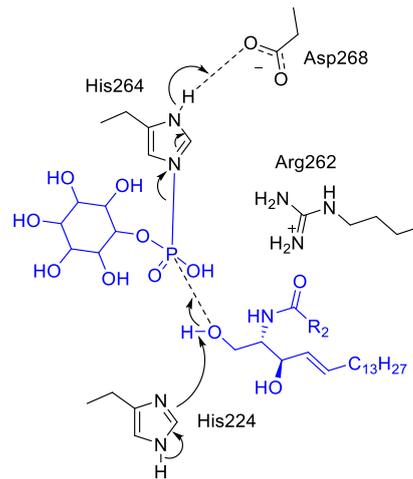


Figure 60 Hypothesised mechanism of LmjIPCS  
 Blue: LmjIPCS substrates PI and ceramide; black: active site amino acid residues

With the model protein chosen, the next challenge was to obtain the enzyme. PhoN is not a commercially available enzyme, therefore recombinant PhoN was produced using competent *E. coli*.

### 3.3.1.1 Cloning

Once the enzyme was identified, the gene was obtained from UniProt (A0A0F6BA10, *Salmonella typhimurium* strain 14028s/SGSC 2262) and ordered through GenScript in a pUC57 cloning plasmid. This was transformed into DH5 $\alpha$  *E. coli* for plasmid propagation. The PhoN gene was then amplified for the transformation into an expression plasmid, with an expected size of 753 bp (Figure 61).

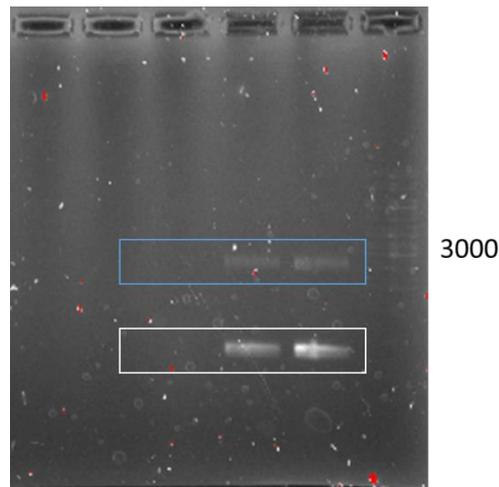
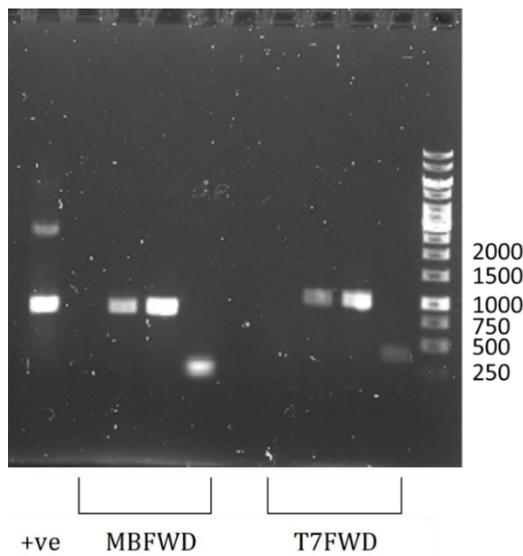


Figure 61 PCR of *PhoN* gene in *pUC57* plasmid from transformed *DH5 $\alpha$*  *E. coli*. Smaller band (bottom; white) expected to be *PhoN* gene; Larger band (top; blue) suspected to be an off-target amplification. The DNA was amplified using the primers designed for *PhoN*.

The designed primers amplified an additional portion of the genome (Figure 61, top band at around 3000 bp), so the desired gene (Figure 61, bottom band) was cut from the agarose gel and purified (Qiagen). Following InFusion cloning into *pOPINF*, this was then transformed into the *E. coli* cell strains BL21 DE3 and BL21 (DE3) *pLysS* and grown, following manufacturers procedures. Successful transformation was checked by colony PCR using the T7 forward primer and the designed forward primer for the expected gene size of 940 and 753 respectively (Figure 62).



**Figure 62 Colony PCR of 3 bacterial colonies following transformation**  
 One bacterial colony was chosen from each transformed species of *E. coli* with a control negative present. Two different primers were used to amplify the colonies. T7FWD: commercial primer for the amplification of the T7 promoter region, which adds 300 b.p onto the gene size; MBFWD: designed primer for the amplification of the beginning of the PhoN gene.

Once the transformed colonies were identified, stock solutions were made with 40 % glycol (v/v) and LB broth at -80 °C for both BL21 DE3 and BL21 (DE3) pLysS. These colonies were sent for Sanger sequencing using T7, MBFWD and MBREV primers (Figure 63).

```

PhoNseq      -----MKSRYLVFFLPLIVAKYTSAETVQPFHSPPEESVNSQFYLP
BL21 (DE3)   MAHHHHHSSGLEVLVFGQPMKSRYLTVFFLPLIVAKYTSAETVQPFHSPPEESVNSQFYLP
BL21pLysS    MAHHHHHSSGLEVLVFGQPMKSRYLTVFFLPLIVAKYTSAETVQPFHSPPEESVNSQFYLP
                *****

PhoNseq      PPGNDDPAYRYDKEAYFKGYAIKGSRWKQAAEDADVSVENIARIFSPVVGAKINPKDTP
BL21 (DE3)   PPGNDDPAYRYDKEAYFKGYAIKGSRWKQAAEDADVSVENIARIFSPVVGAKINPKDTP
BL21pLysS    PPGNDDPAYRYDKEAYFKGYAIKGSRWKQAAEDADVSVENIARIFSPVVGAKINPKDTP
                *****

PhoNseq      ETWNMLKNLLTMGGYYATASAKKYYMRTRPFVLFNHSTCRPEDENTLRKNGSYPSGHTAY
BL21 (DE3)   ETWNMLKNLLTMGGYYATASAKKYYMRTRPFVLFNHSTCRPEDENTLRKNGSYPSGHTAY
BL21pLysS    ETWNMLKNLLTMGGYYATASAKKYYMRTRPFVLFNHSTCRPEDENTLRKNGSYPSGHTAY
                *****

PhoNseq      GTLLALVLSEARPERAQELARRGWEFGQSRVICGAHWQSDVDAGRYVGAVEFARLQTIPA
BL21 (DE3)   GTLLALVLSEARPERAQELARRGWEFGQSRVICGAHWQSDVDAGRYVGAVEFARLQTIPA
BL21pLysS    GTLLALVLSEARPERAQELARRGWEFGQSRVICGAHWQSDVDAGRYVGAVEFARLQTIPA
                *****

PhoNseq      FQKSLAKVREELNDKNNLLSKEDHPKLN
BL21 (DE3)   FQKSLAKVREELNDKNNLLSKEDHPKLN
BL21pLysS    FQKSLAKVREELNDKNNLLSKEDHPKLN
                *****
    
```

**Figure 63 Alignment of reported PhoN sequence with the isolated plasmids**  
 PhoNseq: reported peptide sequence obtained from UniProt (P26976); BL21(DE3): Sanger sequence obtained using the commercial T7 forward primer for the transformed *E. coli* BL21 DE3; BL21pLysS: Sanger sequence obtained using the commercial T7 forward primer for the transformed *E. coli* BL21(DE3) pLysS. Alignment performed using EMBL-EBI (MUSCLE) alignment tool, with the output format using ClustalW

Sanger sequencing determined that no mutations were introduced, and confirmed that the His<sub>6</sub>-tag (Figure 63, 1-19) was positioned on the N-terminus in both conditions, which matches the reported sequence of 'MAHHHHHHSSGLEVLFGGP'.<sup>201</sup>

### 3.3.1.2 Expression of PhoN

To ascertain the ideal conditions for protein expression, 4 different conditions were tested (Table 2) for each of the transformed *E. coli* species.

Table 2 Temperatures and times for the expression of PhoN

Temperature / °C	Time / h
37	4
30	18
25	24
20	24

The transformed colonies were induced at OD 0.6-0.8 using a final concentration of 0.5 mM IPTG. After induction, the conditions shown in Table 2 were followed before lysis using sonication (Figure 64). Alongside the induction conditions, another condition was prepared which was the non-induced condition. This was used as a comparison to the induction condition to aid in the identification of the overexpressed PhoN protein band.

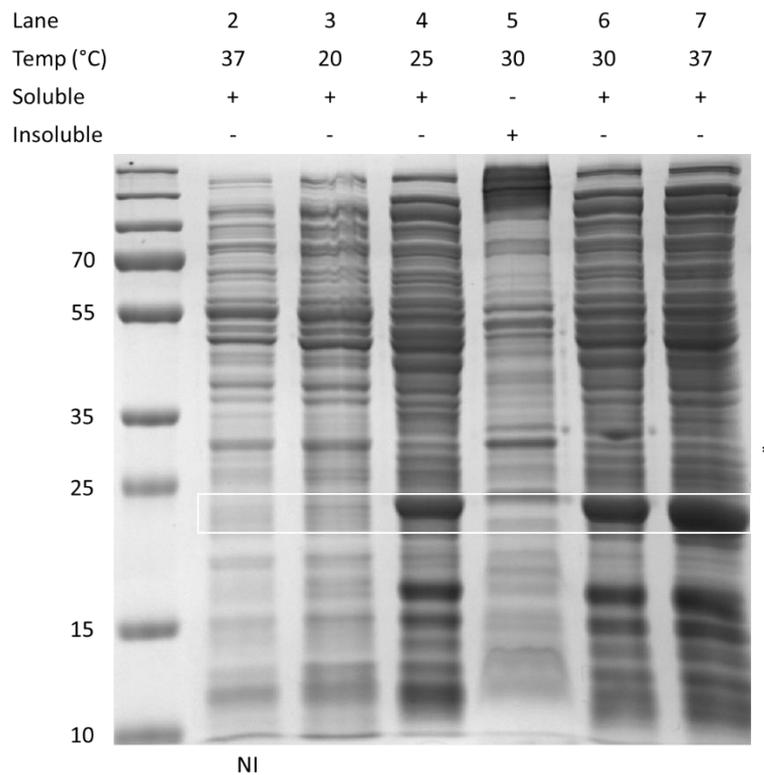
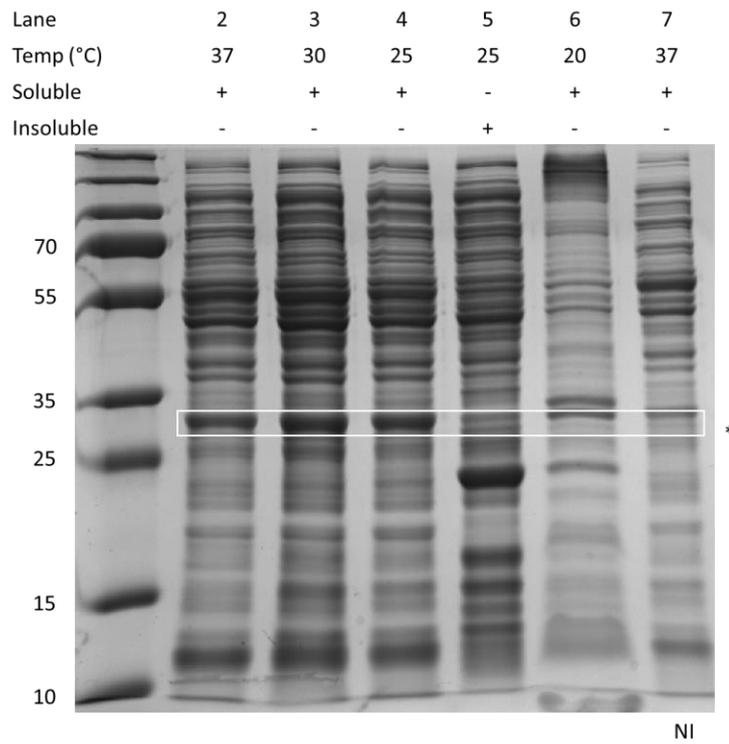


Figure 64 SDS-PAGE analysis of induced lysate of *E. coli* BL21 DE3 (top) and BL21 (DE3) pLysS (bottom) Only some of the insoluble fractions are shown as some of the insoluble material was difficult to load onto the gel. NI: non-induced bacteria culture grown without the addition of IPTG. Star refers to the expected location of the overexpressed PhoN and the box highlights the overexpressed protein. SDS-PAGE gels stained overnight using quick Coomassie protein stain and destained with MilliQ H<sub>2</sub>O.

The results demonstrate that the ideal expression conditions in *E. coli* strain BL21 DE3 was 30 °C overnight (Figure 64; top, lane 3) and at 37 °C for 4 h in BL21 (DE3) pLysS (Figure 64; bottom; lane 6). Whilst a higher quantity of PhoN was achieved in BL21 (DE3) pLysS, it appeared that the recombinant PhoN was 2-4 kDa smaller than the expected mass of 28 kDa (Figure 64, bottom). This was deduced based on the comparison of the non-induced *E. coli* BL21 (DE3) pLysS (Figure 64, bottom, lane 2) and the induction conditions (Figure 64). Therefore, purification initially occurred in the *E. coli* BL21 DE3 expression system.

Following this precedent, recombinant PhoN was overexpressed in 2 L of *E. coli* BL21 DE3 in the optimal condition of 30 °C for 18 h. Following this, the *E. coli* were lysed by sonication and the lysate loaded onto a HisTrap column using a peristaltic pump. The HisTrap was then washed with a wash buffer (20 mM Tris pH 7.4, 500 mM NaCl, 20 mM imidazole), and this loaded onto the protein purification system AKTA to elute the protein using a gradient with 1 M imidazole in the elution buffer. The protein fractions from the AKTA were analysed by SDS-PAGE and Coomassie protein staining along with the *E. coli* BL21 DE3 lysate and proteins eluted during the column wash (Figure 65).

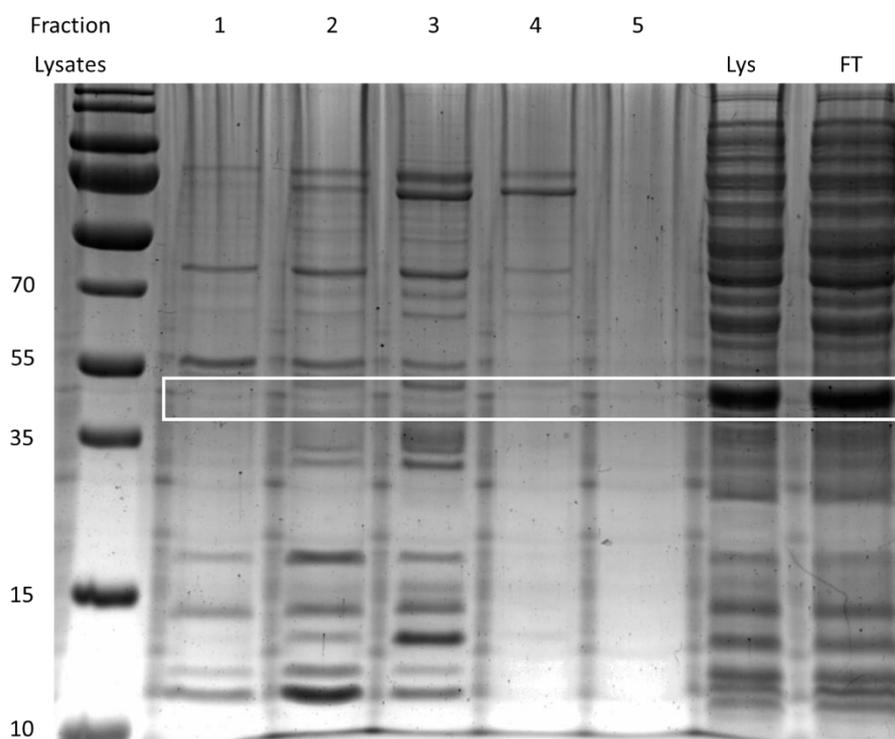


Figure 65 SDS-PAGE analysis of PhoN purification in the expression system *E. coli* BL21 DE3. Lanes 1-5 represent the fractions of eluted protein from the AKTA. Lys: *E. coli* BL21 (DE3) lysate; Wash; proteins collected after washing the HisTrap column. SDS-PAGE gels stained overnight using quick Coomassie protein stain and destained with MilliQ H<sub>2</sub>O. Highlighted is the band where PhoN is expected.

The SDS-PAGE (Figure 65) showed that the purification of recombinant PhoN was unsuccessful in the expression system *E. coli* BL21 DE3. Most of the recombinant PhoN can be seen in the *E. coli* BL21 DE3 wash step, which could be causing the premature elution of the protein. Therefore, the imidazole concentration in the wash buffer was decreased from the standard 20 mM to 10 mM however, this also resulted in no adherence of the protein to the column. It was concluded that the protein coordinates to the Ni<sup>2+</sup> on the column weakly for unknown reasons.

Due to the difficulties with *E. coli* BL21 DE3, the overexpression and purification was then performed using *E. coli* BL21 (DE3) pLysS, and enzyme activity was determined following isolation (Figure 66).

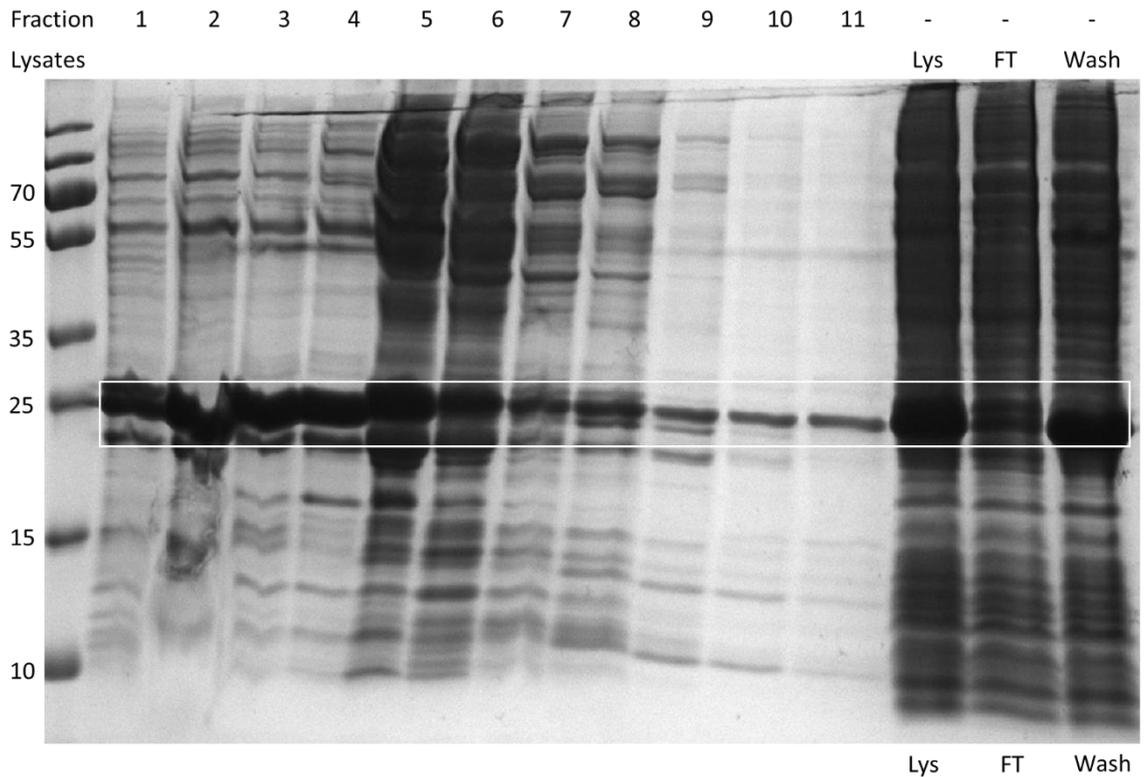
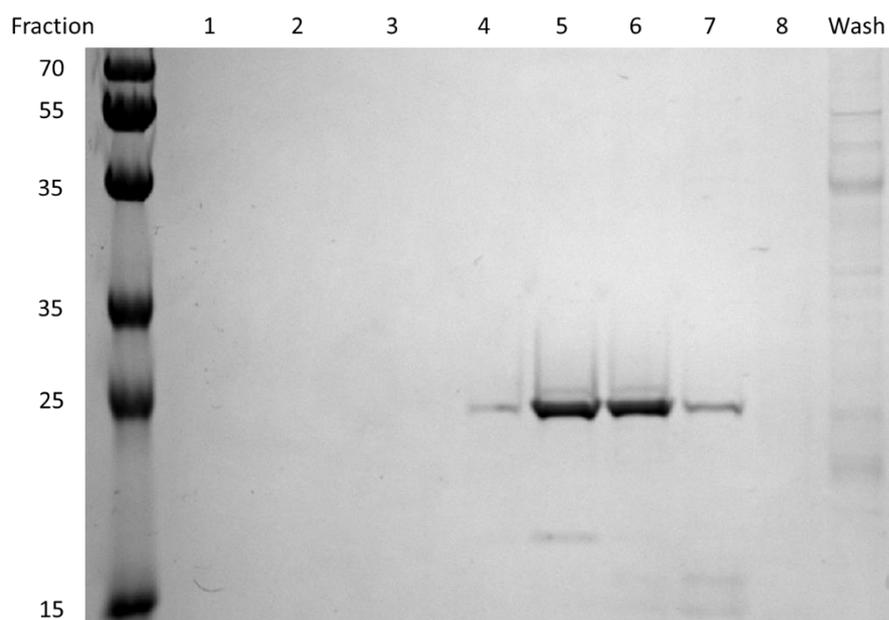


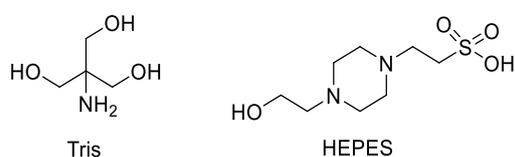
Figure 66 SDS-PAGE analysis of PhoN purification using the expression system *E. coli* BL21 DE3 pLysS. Lanes 1-11 represent the fractions of eluted protein from the AKTA. Lys: *E. coli* BL21 (DE3) pLysS lysate; FT: flowthrough collected after loading the HisTrap column with *E. coli* BL21 (DE3) pLysS lysate using a peristaltic pump; Wash; proteins collected after washing the HisTrap column. Highlighted is the band where PhoN is expected. SDS-PAGE gels stained overnight using quick Coomassie protein stain and destained with MilliQ H<sub>2</sub>O.

Due to the known weak binding of PhoN, no imidazole was present in the wash buffer, and the protein was eluted on a gradient with a maximum concentration of 500 mM imidazole. This HisTrap column purification successfully eluted the desired enzyme however, PhoN eluted nearly instantaneously, which again suggests weak binding with the Ni<sup>2+</sup> (Figure 66). In addition, a large quantity of PhoN came off in the washing stage, with no imidazole present. The eluted fractions A1-A4 were combined, and then these fractions were further purified using MonoQ anion exchange (Figure 67). This type of purification separates proteins using a high salt gradient based on the differences in charge.<sup>202</sup>



*Figure 67 SDS-PAGE analysis of MonoQ anion purification*  
 Fractions 1-8 represent the eluted fractions from the AKTA following MonoQ anion exchange. Wash: proteins eluted from washing the Ion Exchange Column. SDS-PAGE gels stained overnight using quick Coomassie protein stain and destained with MilliQ H<sub>2</sub>O

This additional purification successful generated a total of 4 mg of PhoN. Tris was the buffer used for the purification however it is a known nucleophile and is reported to decrease the half-life of DEPC (**50**) to 1.25 mins at RT.<sup>179</sup> In addition, it was previously shown in Chapter 2.4.1 that the aryl carbamate can be attacked by nucleophilic agents therefore, it is preferable to avoid the use of tris for BHDC (**66**) testing and exploration. Thus, the collective protein was dialysed into 20 mM HEPES (pH 7.4) for all the following described experiments (Figure 68).

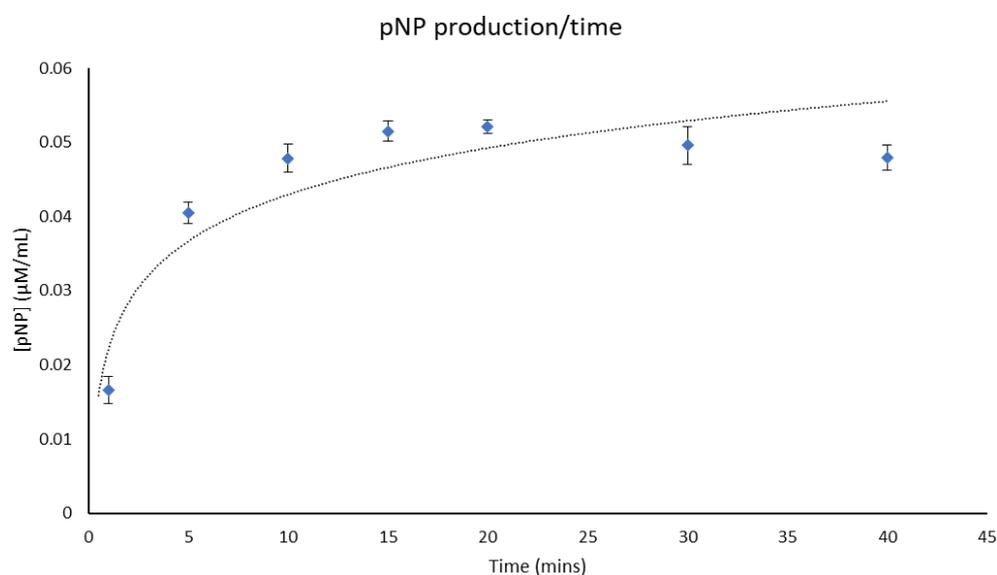


*Figure 68 The structures of Tris and HEPES*

Following purification and dialysis, PhoN enzyme activity was then determined by reconstituting the reported assay for non-specific phosphatases, which was important due to the concerning truncation of the protein.

### 3.3.2 Non-specific acid phosphatase activity

To ensure that the isolated enzyme retained activity, an assay first described by Roche (1950) to measure non-specific acid phosphatase activity was utilised. This assay uses the substrate, para-nitrophenyl phosphate (pNPP, **44**), and monitors the production of the chromogenic product para-nitrophenol (pNP, **45**), with an absorbance of 405 nm.<sup>203</sup> The assay was performed at room temperature with the concentration of 2  $\mu\text{g}/\text{mL}$  PhoN. The assay was quenched with the total reaction volume of 0.2 M NaOH after 5, 10, 15, 20, 30 and 40 mins from the reaction start. With the reported extinction coefficient of pNP (**45**) being  $18,000 \text{ M}^{-1}\text{cm}^{-1}$ , the product concentration was calculated using the Beer-Lambert law from the fluorescent output (Figure 69).<sup>204</sup>

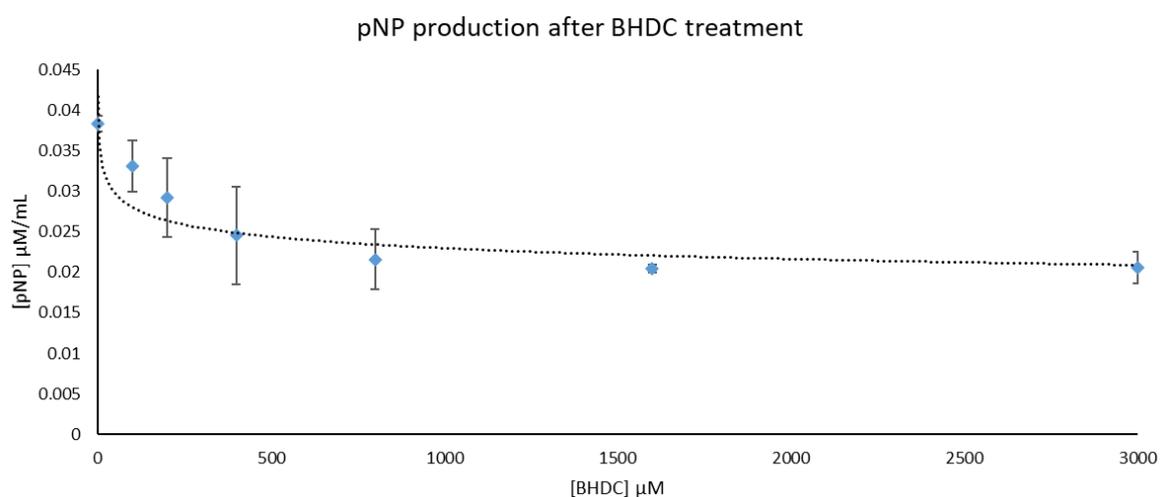


*Figure 69 Non-specific acid phosphatase enzyme activity*  
5 mM pNPP was incubated with the purified recombinant PhoN. The reaction was quenched before the measurement of fluorescence using Cary 5000 UV-Vis-NIR at 405 nm. Data is representative of 4 repeats, with the standard deviation calculated from the repeats and are shown as error bars. The trendline is shown as a logarithmic curve.

With the Beer-Lambert calculation, the data generated at 15 mins was used to calculate the speed of [pNP] production. This was due to the enzyme appearing to reach saturation at this time. Thus, by dividing the time by the [pNP] production, this experiment shows that the activity of the isolated PhoN was 11 nM [pNP]/min/ $\mu\text{g}$  protein, before activity was saturated. After successfully validating the activity of PhoN, the next goal was to validate BHDC (**66**) activity with this enzyme.

### 3.3.3 BHDC inhibition of pNP production

A quick method to determine whether BHDC (**66**) is on target was to perform the NSAP assay following incubation with **66** and measure activity. A decrease in pNP (**45**) production could suggest that the probe is inhibiting the active site. For this experiment, 1 h incubation with varying concentrations of BHDC (**66**) was followed by the addition of pNPP (**44**, 5 mM). The final incubation time with pNPP (**44**) was chosen to be 15 mins, as this was determined to be when the enzyme reached saturation, as described and seen in Figure 69.



*Figure 70 Non-specific acid phosphatase inhibition with BHDC (**66**)*

*BHDC (**66**) was incubated with PhoN before the addition of the substrate pNPP (**44**, 5 mM). The fluorescence was measured using Cary 5000 UV-Vis-NIR at 405 nM. Data is representative of 4 repeats, with standard deviation shown as error bars. The trendline is shown as a logarithmic curve.*

This experiment showed a general decrease in the production of pNP (**45**) after the addition of BHDC (**66**), indicating that the probe does inhibit the enzyme (Figure 70). This was promising as you would expect the quantity of pNP to decrease if BHDC (**66**) does inhibit the enzyme active site, however, this does not eliminate the possibility of multiple acylation events. To preclude this possibility, mass spectrometry was utilised to determine specificity and identify the source behind the truncated protein.

### 3.4 PhoN mass spectrometry analysis

Following the successful PhoN isolation and the NSAP activity identified, the initial experiment focused on validating and identifying the mass of the isolated protein. This was important because the estimated mass of the isolated enzyme from the gels (~ 25 kDa, Figure 67) appeared smaller than the mass of PhoN, which is 28382 Da without the addition

of the His tag. With the addition of the His tag sequence which adds 1.73 kDa, this should bring the total mass of the enzyme to ~ 30112 Da.

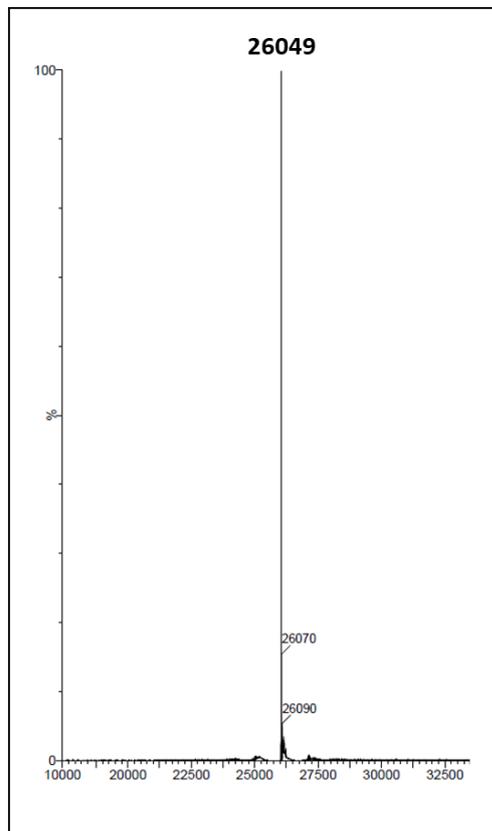


Figure 71 PhoN [ESI]<sup>+</sup> mass spectrometry  
PhoN concentration 0.25 mg/mL

The initial [ESI<sup>+</sup>] mass spectrometry showed that the isolated PhoN was 26049 Da, which is 4 kDa smaller than the predicted mass with the addition of the poly histidine tag (Figure 71). Proteolytic digestion using trypsin and chymotrypsin revealed the isolated enzyme matched PhoN with 85 % coverage. Further analysis of the identified peptide fragments revealed that the first 20 amino acids of the N terminus was not identified. Investigation using a SignalP (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0>) predicted with a probability of 0.89 that a signal peptide exists at the N terminus which would result in a cleavage between amino acids 20 and 21 (Figure 72).<sup>205</sup>

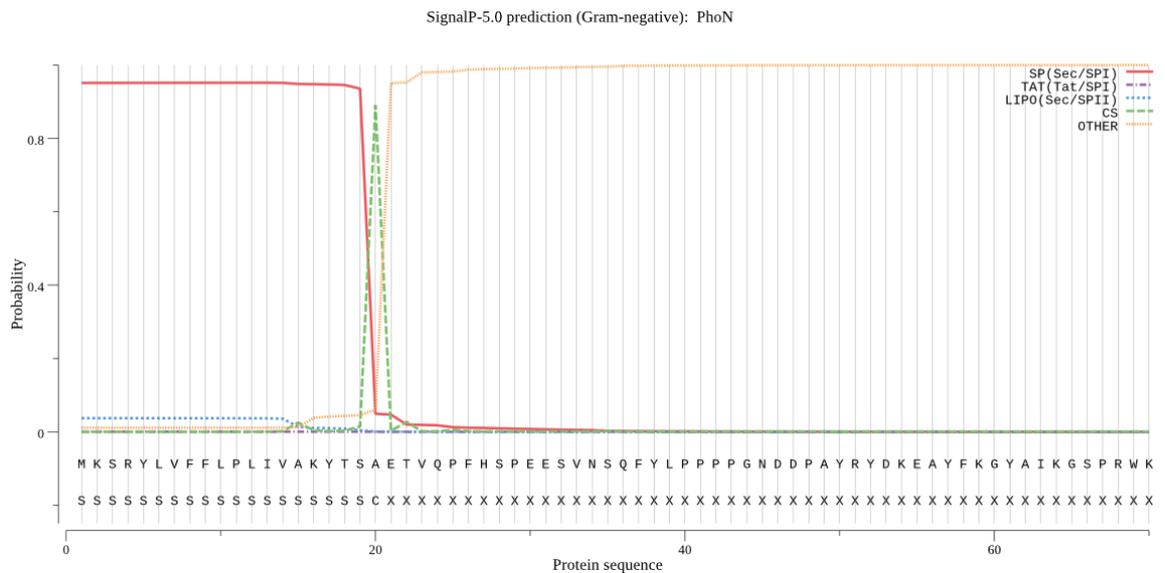


Figure 72 SignalP-5.0 prediction of a cleavage site in PhoN<sup>205</sup>

SP (Sec/SPI): transport by the Sec translocation and cleaved by Signal Peptidase I (*Lep*); TAT (TAT/SPI): transport by the Tat translocation and cleaved by Signal Peptidase II (*Lsp*);

A cleavage in the specified area matches the missing fragment, leading to the hypothesis that this region was cleaved by intrinsic Signal Peptidase I in *E. coli*. With this discovery, the predicted mass without this region is 26.06 kDa ([https://www.bioinformatics.org/sms/prot\\_mw.html](https://www.bioinformatics.org/sms/prot_mw.html)), which is only 10 Da larger than the 26049 Da mass found from mass spectrometry and increases the fragment coverage from the proteolytic digestion to 98 %. This hypothesis would also explain the previous difficulties with the isolation of PhoN using a HisTrap column.

Following the successful isolation and identification of an active enzyme and preliminary BHDC (**66**) inhibition validation, the NSAP PhoN was then utilised in a series of mass spectrometry experiments. For this, increasing concentrations of BHDC (**66**) was incubated with PhoN for 1 h before analysis by (ESI<sup>+</sup>) mass spectrometry with the inclusion of 5-hexyn-1-ol (**63**) as a control (Figure 73).

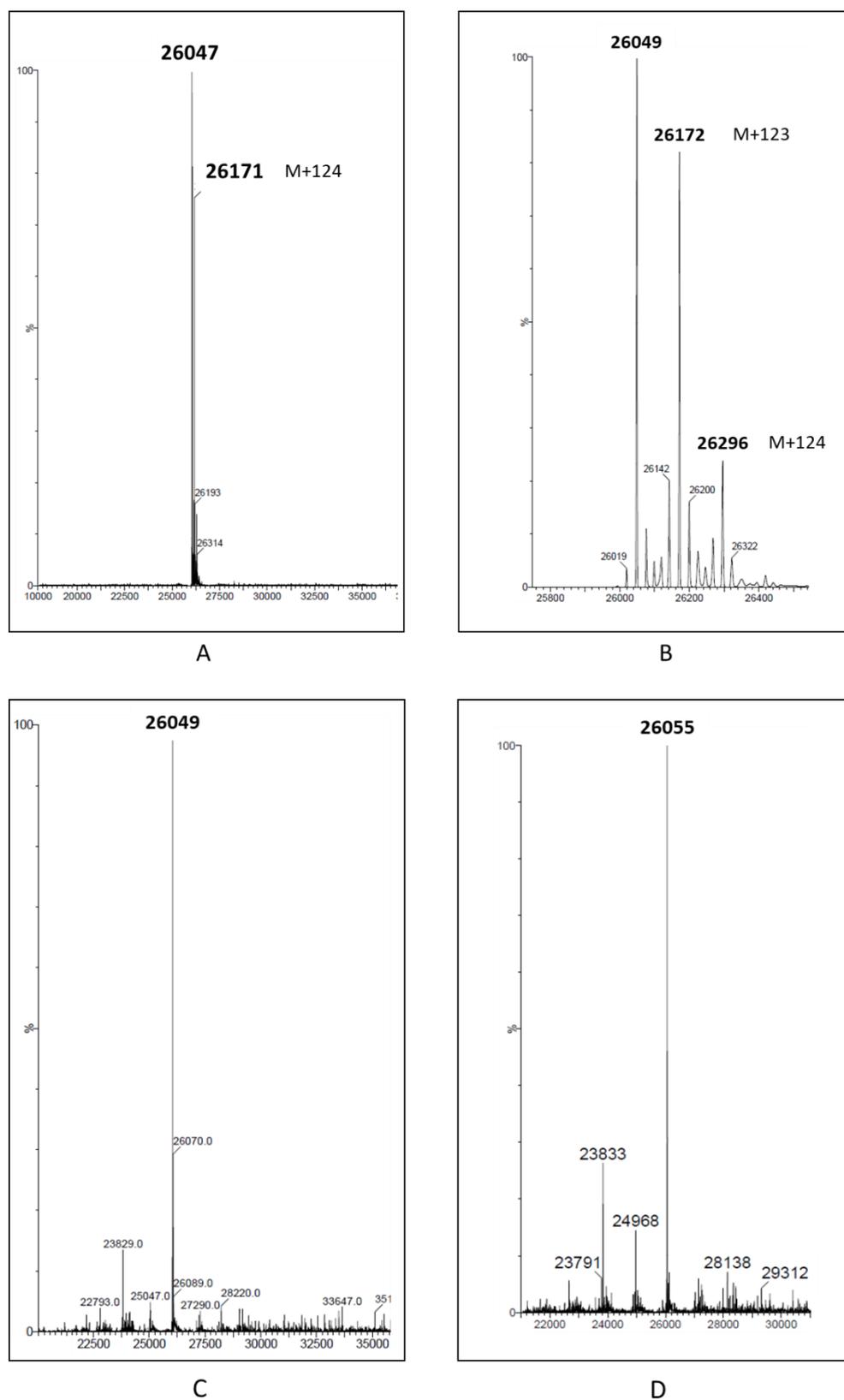
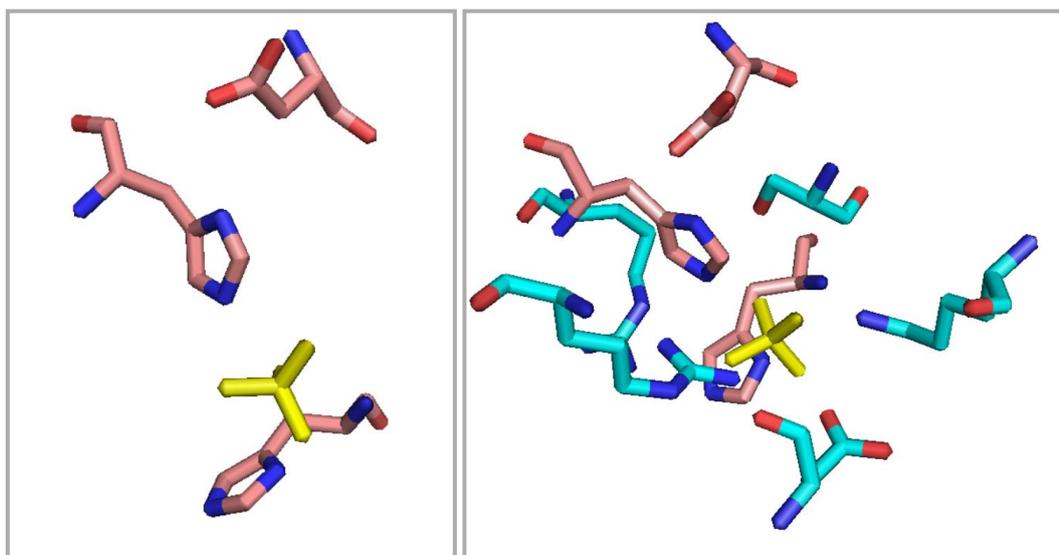


Figure 73 Mass spectrometry of PhoN with BHDC

PhoN was dissolved to 0.2 mg/mL in H<sub>2</sub>O. A: incubation with 100 μM BHDC (66), 1 hr; B: incubation with 200 μM BHDC (66), 1 hr; C: incubation with 100 μM 5-hexyn-1-ol (63), 1 hr; D: incubation with 200 μM 2-cyclohexanone (49), 24 h. [M+124]<sup>+</sup> corresponds with a modification with BHDC (66). Data is representative of 3 repeats.

The expected mass of the BHDC (**66**) modification is +125.05. When 100  $\mu\text{M}$  of BHDC (**66**, 10 eq) was incubated with PhoN for 1 h before analysis, there was an addition of  $[\text{M}+124]^+$ , which corresponds to the modification by **66** (Figure 73, A). When the concentration of BHDC (**66**) was increased to 200  $\mu\text{M}$  (20 eq), there appeared to be multiple acylation events (Figure 73, B). This suggests a loss of chemoselectivity at this concentration. Concurrent with what was found for the *Lmj*/PCS assay, it appears that 5-hexyn-1-ol (**63**), which is produced as a by-product of BHDC (**66**) acylation, had no effect on PhoN (Figure 73, C). Interestingly, when incubated with the reported known histidine inhibitor 2-cyclohexenone (**49**), described in Chapter 1.7, no modification could be seen corresponding to the addition of 2-cyclohexanone (Figure 73, D).<sup>178</sup> This was an important finding showing that BHDC (**66**) has novel activity from other histidine modifying reagents.

A proteolytic digestion of PhoN following treatment with 100  $\mu\text{M}$  BHDC (**66**) by reducing the disulphide bonds with TCEP and then the cysteine residues were modified with iodoacetamide. Then, using trypsin and chymotrypsin showed a modification on the fragment IC(+57.02)GAH(+125.06)WQSDVDAGR. The  $[\text{M}+125]^+$  corresponding to His-197, the known nucleophilic residue in the active site of PhoN along with  $[\text{M}+57]^+$  corresponding to the carbamidomethylation on Cys-194.<sup>206</sup> In addition, the second modification seen at 200  $\mu\text{M}$  BHDC (**66**) was on fragment VLFNH(+125.06)STC(+57.02)RPEDENTLRK, which is His-137, a solvent exposed histidine. Whilst this result was undesired, it determined a limit of <10 equivalents to retain chemoselectivity for a nucleophilic histidine residue. This off-target phenomenon is not unique to BHDC (**66**), with many reported ABPs having non-specific activity if used in excess.<sup>125,178</sup> With the presence of other amino acids that are reported to act as a nucleophile (Figure 72) including 2 cysteine residues, selectivity of the target residue was a promising achievement.<sup>155</sup>



Residue	Active site
His	2
Arg	2
Lys	1
Ser	1
Cys	0

Figure 74 Active site of PhoN *S. typhimurium* (PDB: 2A96)<sup>155</sup>

Phosphate ion present in the active site shown in yellow. Active site residues HHD shown in pink. Residues forming the active site pocket that do not participate in direct enzyme activity shown in cyan.

This series of experiments succinctly shows that within 10 equivalents, the leading probe BHDC (**66**) is selective for the active site nucleophilic histidine in PhoN, in a pure enzyme model context. This was a great achievement however, whether the designed probe will retain activity in a mixture of protein remains to be explored, along with the stability of the carbamate for the bioorthogonal click chemistry.

### 3.5 Fluorescent labelling of PhoN

The experiments described above were indicative that BHDC (**66**) was functioning as a selective probe for the active site nucleophilic His-197 residue in PhoN. However, the stability of the electron deficient aryl carbamate formed after histidine acylation was still unexplored in a biological context. As this had previously shown instability to nucleophiles (described in Chapter 2.4.1) it was important to verify selectivity and stability under the

conditions required for fluorescent labelling and enrichment. The stability of the carbamate is an important component for the full functionality of the ABP for the subsequent fluorescent and enrichment studies.

### 3.5.1 Pure protein labelling

To assess whether BHDC (**66**) is stable enough to allow for the conjugation of fluorophores and biotinylated compounds, initial experiments using purified PhoN was performed to derive suitable conditions. These initial experiments used the simple rhodamine linker (**93**) described in Chapter 2.6.1 (Figure 75).

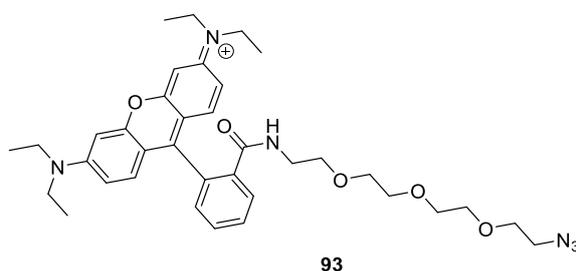


Figure 75 Synthesised simple rhodamine linker used to derive click chemistry conditions

To begin, the protein samples were incubated with 100  $\mu\text{M}$  BHDC (**66**) for 1 h at RT. Following this was the addition of 200  $\mu\text{M}$  of the simple rhodamine linker (Rho-N<sub>3</sub>, **93**) and the other click chemistry components following the procedure established in Speers and Cravatt (2009).<sup>94</sup> It was important to add double the equivalents of the fluorophore (**93**) as the probe (**66**) has two alkyne functional groups per molecule, which could react with the fluorophore. This was initially unsuccessful at producing fluorescently labelled proteins (Figure 76).

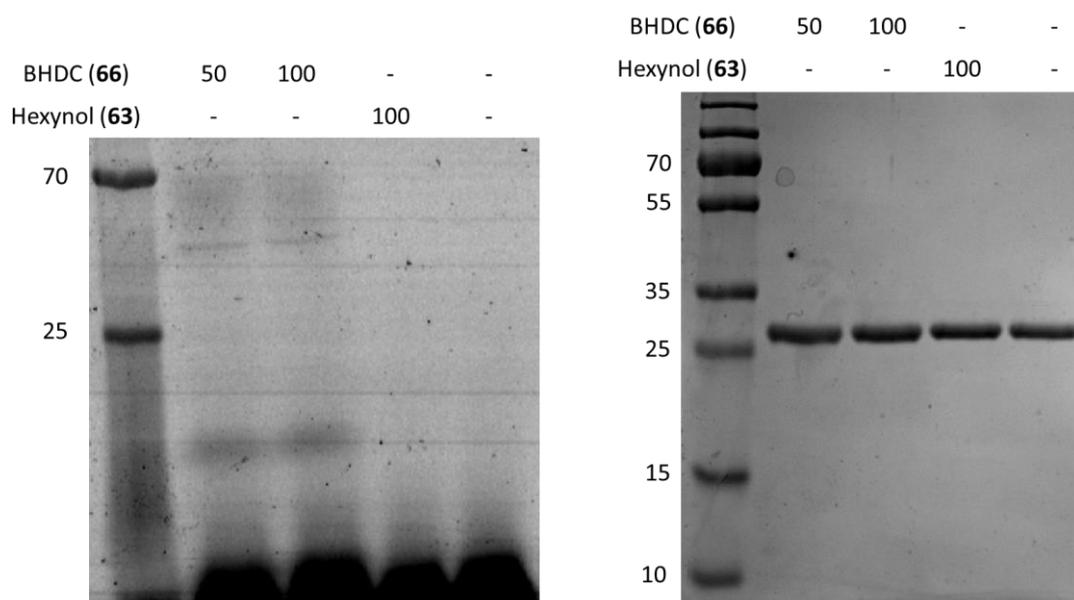
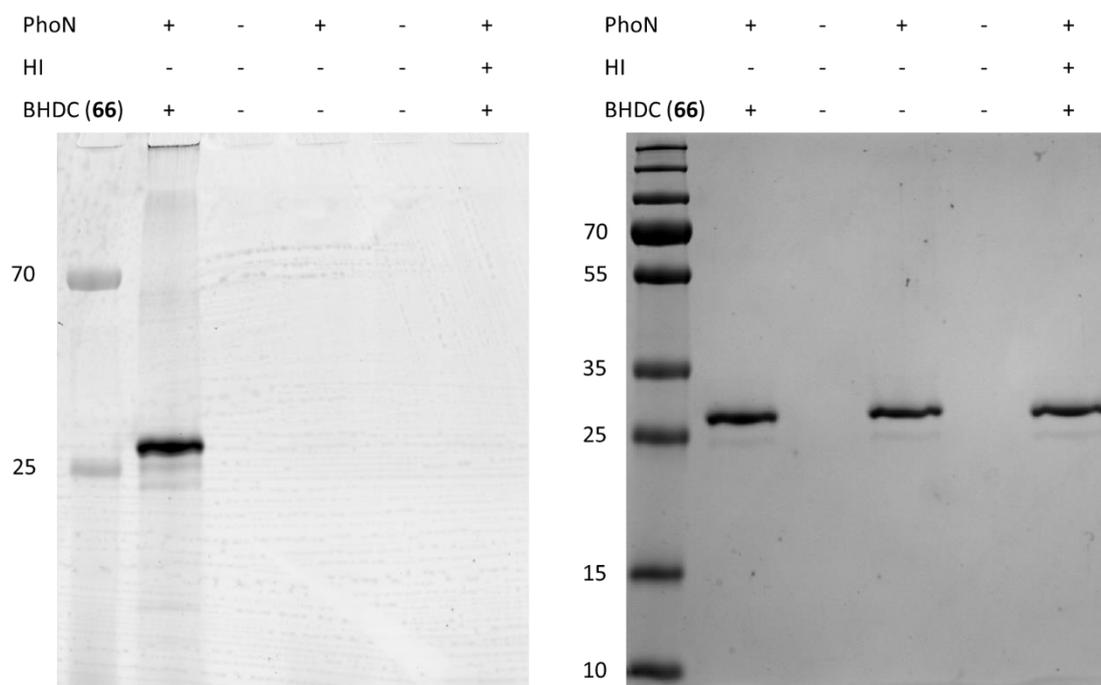


Figure 76 Initial BHDC (66) binding with PhoN  
 Left: fluorescent image taken on Fuji FLA3000 ( $\lambda_{ex}$  473/ $\lambda_{em}$  520 nm); right: corresponding quick Coomassie protein stain. Representative SDS-PAGE from 2 consistent replicates.

Two potential reasons for this could be put forward. Either the carbamate was too unstable for the CuAAC reaction, or the traditional protein loading for SDS-PAGE which recommends heating the sample to 95 °C for 10 min led to decarboxylation. The latter hypothesis has literature support, with Dyer *et al.*, reporting the thermal instability of carbamates resulting in the release of CO<sub>2</sub> and the corresponding alcohol.<sup>207</sup> To test the latter hypothesis, the protein samples were not heated before the addition to the gel (Figure 77).



*Figure 77 Labelling and fluorescent analysis of PhoN with BHDC (66)*  
*PhoN was used to a concentration of 0.5 mg/mL, BHDC (66) was used to a final concentration of 50  $\mu$ M. HI: heat inactivated protein. Left: fluorescent image taken on Typhoon 9400 ( $\lambda_{ex}$  532 nm/ $\lambda_{em}$  580 nm); Right: quick Coomassie protein stain. Representative SDS-PAGE from 3 consistent replicates.*

The adjusted conditions successfully led to the labelling of active PhoN. Therefore, as inferred from Dyer's report, this substantiated that the carbamate is unstable to heat and eliminated from the histidine.<sup>207</sup>

### 3.5.2 Lysate protein labelling

With fluorescent labelling now functioning in the hypothesised manner, the next challenge was to ascertain the chemoselectivity of BHDC (66) labelling in a mixture of proteins. To achieve this, labelling experiments using WT lysate and overexpressing PhoN. To ensure that there was no leaky expression of PhoN in the *E. coli* lysate, *E. coli* BL21 (DE3) pLysS was transformed with the manufacturers supplied control plasmid, which encoded the gene for TNF $\alpha$ . This was tested alongside PhoN overexpressed lysate (Figure 78).

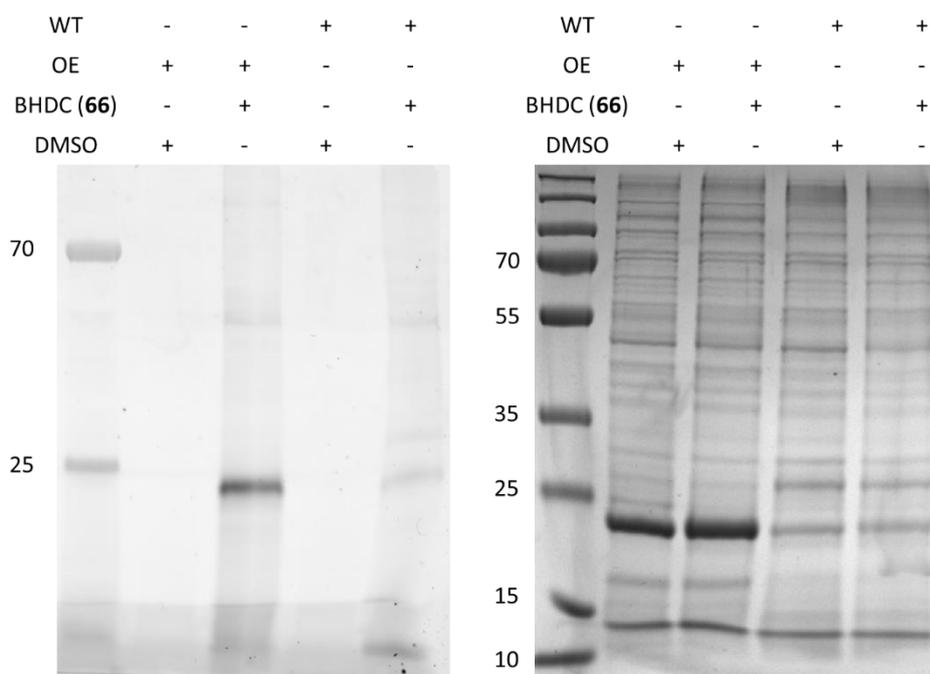


Figure 78 BHDC (66) incubation with overexpressed PhoN lysate and control lysate  
 Left image is a fluorescent image obtained using Typhoon 9400 ( $\lambda_{ex}$  532 nm/  $\lambda_{em}$  580 nm); Right is the corresponding quick Coomassie protein stain. WT: lysate generated using *E. coli* BL21 (DE3) pLysS transformed with the control plasmid; OE: lysate generated from *E. coli* BL21 (DE3) pLysS overexpressing PhoN; BHDC (66): 50  $\mu$ M. Representative SDS-PAGE from 3 consistent replicates.

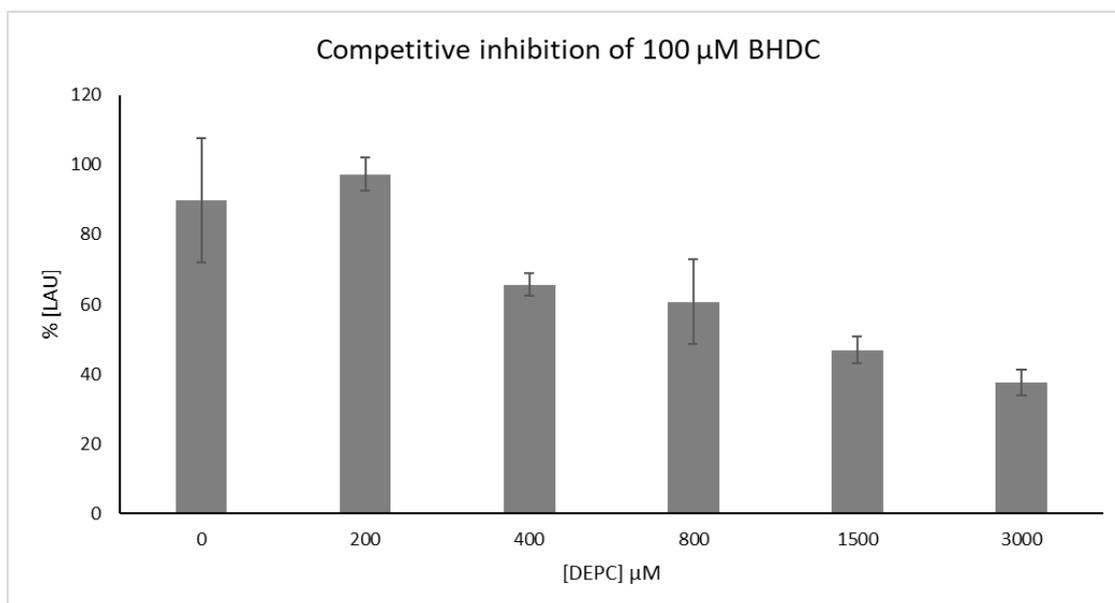
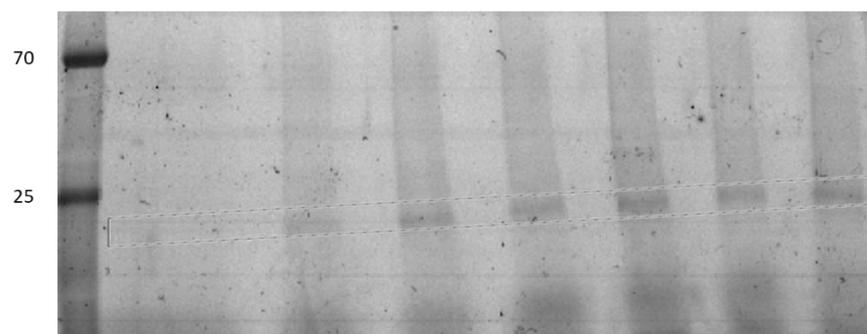
This experiment appeared to successfully label PhoN in the overexpressed lysate, which is suggestive that chemoselectivity was retained in the lysate. Interestingly, a protein of a similar mass to PhoN was labelled in the control BL21 (DE3) pLysS (Figure 78). This is not explained by leaky enzyme expression as the *E. coli* BL21 (DE3) pLysS does not have the plasmid to produce PhoN and accidentally pipetting the wrong sample was ruled out after repeats. Therefore, this is likely indicative of the labelling of an intrinsic *E. coli* protein which has a similar mass to PhoN. This was subsequently discovered to be 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase as determined by mass spectrometry (section 3.7). In conclusion, BHDC (66) bound to PhoN in overexpressed lysate however this was complicated by the presence of a wildtype protein of a similar mass.

### 3.5.3 Competitive inhibition

With the confirmation of BHDC (66) fluorescent labelling achieved in pure enzyme and lysate conditions, the project progressed to a competitive inhibition experiment with DEPC (50). This was to evaluate whether BHDC (66) retained the same histidine residue labelling

activity as that reported for DEPC (**50**) in the proteome. In brief, the PhoN overexpressed *E. coli* lysate was preincubated with DEPC (**50**) at the concentrations of 0  $\mu\text{M}$ , 200  $\mu\text{M}$ , 400  $\mu\text{M}$ , 800  $\mu\text{M}$ , 1.5 mM, 3 mM for 1 h prior to the incubation with 100  $\mu\text{M}$  BHDC (**66**, Figure 79).

BHDC ( $\mu\text{M}$ )	0	100	100	100	100	100	100
DEPC ( $\mu\text{M}$ )	3000	3000	1500	800	400	200	0



*Figure 79 Competitive inhibition of BHDC with DEPC*  
*PhoN concentration 0.25 mg/mL. DEPC (**50**) was preincubated with the samples for 1 h before the addition of BHDC (**66**). The samples were analysed by SDS-PAGE and then Fuji FLA-3000 ( $\lambda_{\text{ex}}$  473/ $\lambda_{\text{em}}$  520 nm) with the fluorescence quantified using AIDA image analyser. Data representative of 3 consistent replicates with standard deviation is shown as error bars from the repeats.*

As seen in Figure 79, increasing concentrations of DEPC (**50**) resulted in a general decrease in measured fluorescence. This supported the hypothesis that BHDC (**66**) retained similar functions to DEPC (**50**).

Thus far, BHDC (**66**) has functioned in the hypothesised manner. The probe (**66**) successfully and selectively labelled PhoN at 10 equivalents, appeared to label PhoN in an overexpressed lysate system in addition to labelling intrinsic *E. coli* proteins from a wildtype lysate. It was presumed that these labelled proteins from the wildtype lysate also have an active site histidine residue. Moreover, a competitive inhibition of PhoN by DEPC (**50**) showed that BHDC (**66**) retained similar activity to **50**, the histidine acylating agent that initiated this project. Whilst this was promising, histidine residues are utilised in many ways in enzyme active sites, as discussed in Chapter 1.6. Given the broad spectrum activity of DEPC (**50**), an acid base histidine residue active site was chosen.<sup>179,183</sup> Significantly, DEPC (**50**) is routinely utilised for the inactivation of ribonuclease A for the extraction of RNA.<sup>208</sup> To explore whether BHDC (**66**) similarly can acylate acid base histidine residues, it was implemented in another model protein ribonuclease A. This was to explore whether selectivity could be retained in this type of enzyme.

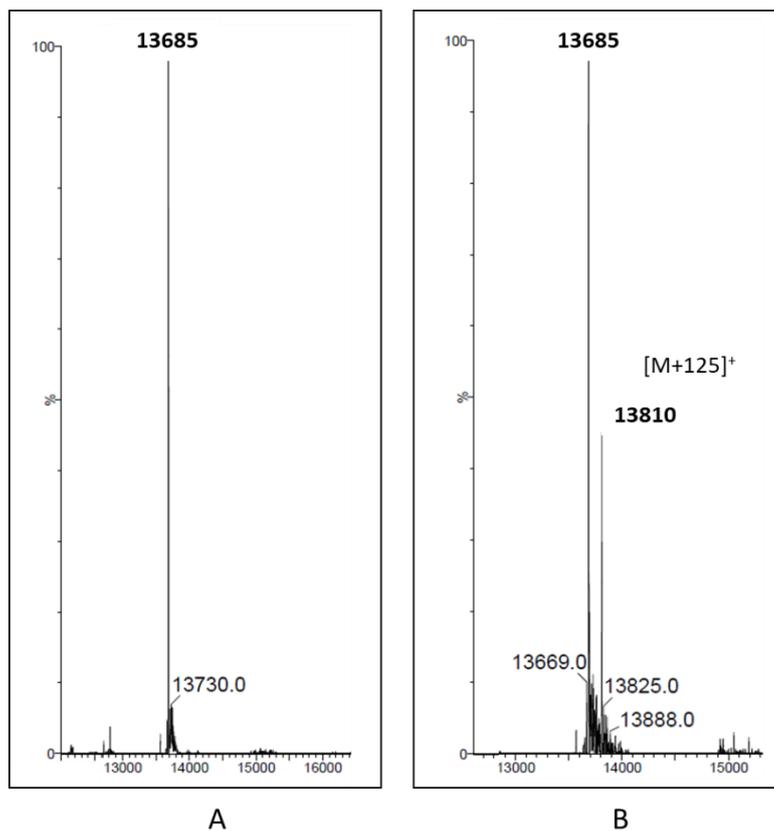
### 3.6 Reactivity in an acid-base active site

Thus far, the experiments were conducted on enzymes (*LmjI*PCS and PhoN) which possess a nucleophilic histidine residue, specifically with the HHD catalytic triad. Whilst there has been promising progress with these models, what has not been explored is a histidine residue with a different function in an enzyme. As discussed in Chapter 1.6, histidine residues have many roles in the active site including functioning as a nucleophile, acid base and a metal ligand. Therefore, to explore the breadth of labelling, another enzyme model was chosen with an acid base histidine residue.

One such enzyme is ribonuclease A (RNase A), which is a small and commercially available enzyme. This enzyme breaks down RNA by cleaving the 3',5'-phosphodiester bond.<sup>160</sup> Interestingly, a common use of DEPC (**50**) is as an additive to buffers (0.1 %) required for RNA purification, due to the ability of DEPC (**50**) to inactivate RNase, and thereby increasing RNA yield.<sup>208</sup> With BHDC (**66**) being a derivative of DEPC (**50**), it was hypothesised that the **66** would also modify this protein.

For initial mass spectrometry analysis, RNase A (Sigma-Aldrich, bovine pancreas, 9001-99-4) was dissolved at 0.2 mg/mL in either H<sub>2</sub>O at 0.2 mg/mL or in PBS (pH 7.4) with no reduction and modification of the cysteine residues. As detailed by the manufacturer

Merck, the purchased RNase A is reported to be 13.6 kDa and requires salt for optimal function. As such, it was hypothesised that RNase A would function better in PBS, a salt-based buffer as opposed to H<sub>2</sub>O, in which activity will be hindered. Therefore, observing BHDC (**66**) reactivity when active in PBS in comparison to the inactivated condition H<sub>2</sub>O can determine the selectivity in functional and non-functional RNase A (Figure 80).



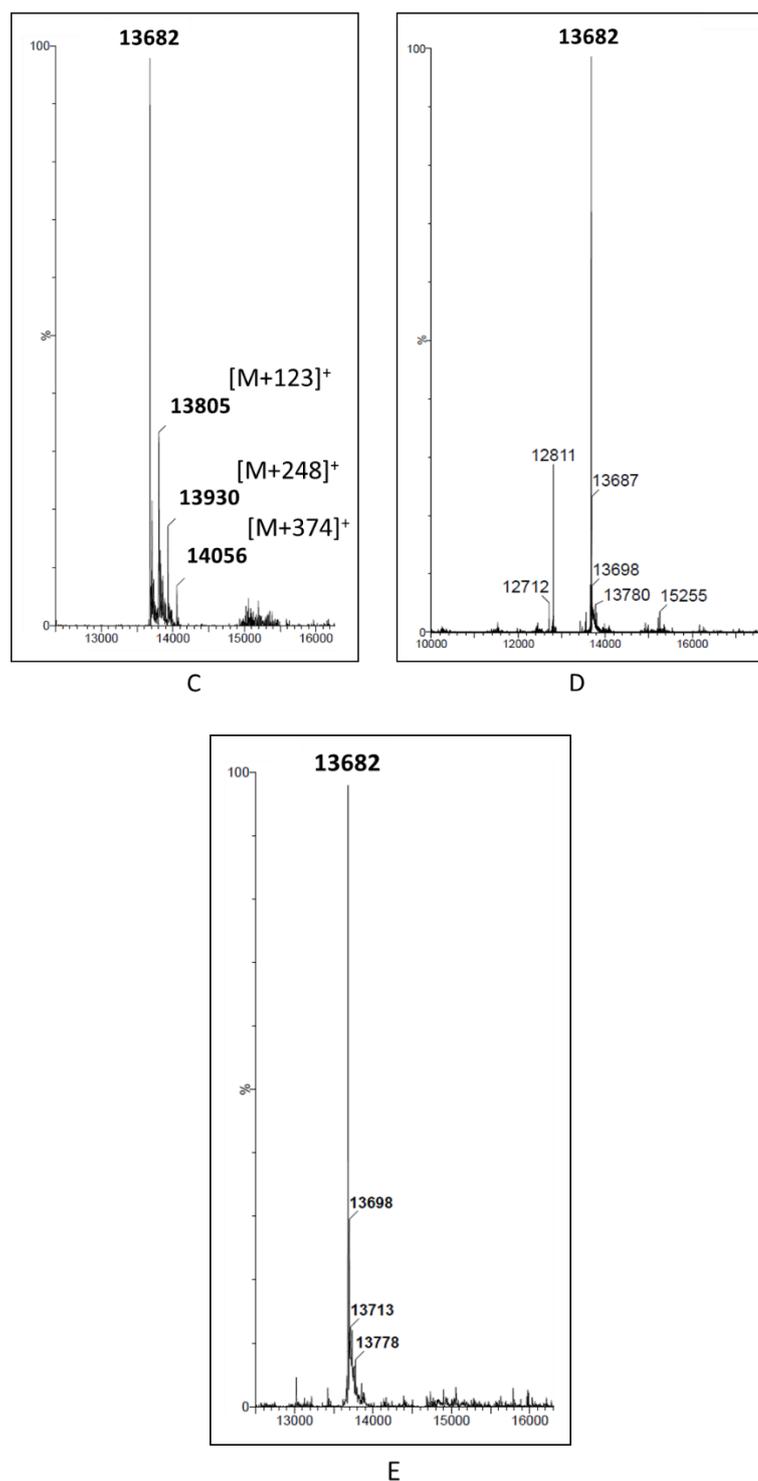
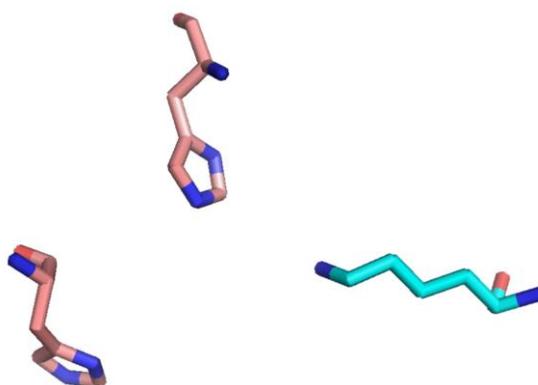


Figure 80 Assess BHDC (**66**) reactivity with RNase A using ESI+ MS  
 A: RNase A in H<sub>2</sub>O; B: RNase A in PBS with 50 μM BHDC (**66**); C: RNase A in PBS with 100 μM BHDC (**66**); D: RNase A in H<sub>2</sub>O with 50 μM BHDC (**66**); E: RNase A in PBS with 100 μM 5-hexyn-1-ol (**63**). Representative from 2 consistent replicates.

Using ESI<sup>+</sup>, it was discovered that when RNase A was incubated with 50 μM (3.3 eq) BHDC (**66**) in PBS, there was a mass [M+124]<sup>+</sup> analogous to a single acylation (Figure 80, B). When

the acylation was performed in H<sub>2</sub>O at the same concentration, no additional peak could be seen (Figure 80, D). This suggests chemoselectivity for the active RNase A and that the labelling of BHDC (**66**) is likely to be at the active site as the non-functional enzyme was not labelled. However, when BHDC (**66**) was added to a final concentration of 100 μM (6.7 equivalents) in PBS, this resulted in 3 additional peaks at  $m/z$  13805 [M+123]<sup>+</sup>, 13930 [M+248]<sup>+</sup> and 14056 [M+374]<sup>+</sup> (Figure 80, C). These masses are consistent with the modification of RNase A with BHDC (**66**) at one, two and three different residues respectively. These results show that chemoselectivity is retained in an acid base active site at 3 eq. Notably, this is a lower threshold to retain specificity than that found in PhoN (10 eq), which could be due to the decreased reactivity of the acid base histidine residue relative to the solvent exposed residues. Consistent with the PhoN analysis, 5-hexyn-1-ol (**63**) did not react with RNase A. In comparison to the previously used PhoN, RNase A is rich in cysteine residues which could increase the likelihood of an off-target reactivity. However, there are 4 disulphide bridges, so all 8 present cysteine residues are oxidised and therefore, are unlikely to react with the probe (**66**).



Residue	Active site
His	2
Arg	0
Lys	1
Ser	0
Cys	0

Figure 81 Structure and active site of RNase A from bovine pancreas (PDB: 1FS3)<sup>159</sup>  
Active site amino acids are in pink whilst the amino acid which forms part of the active site pocket is shown in cyan.

In addition, there are two active site histidine residues which participate to catalyse the degradation of RNA (Figure 82).<sup>163</sup>

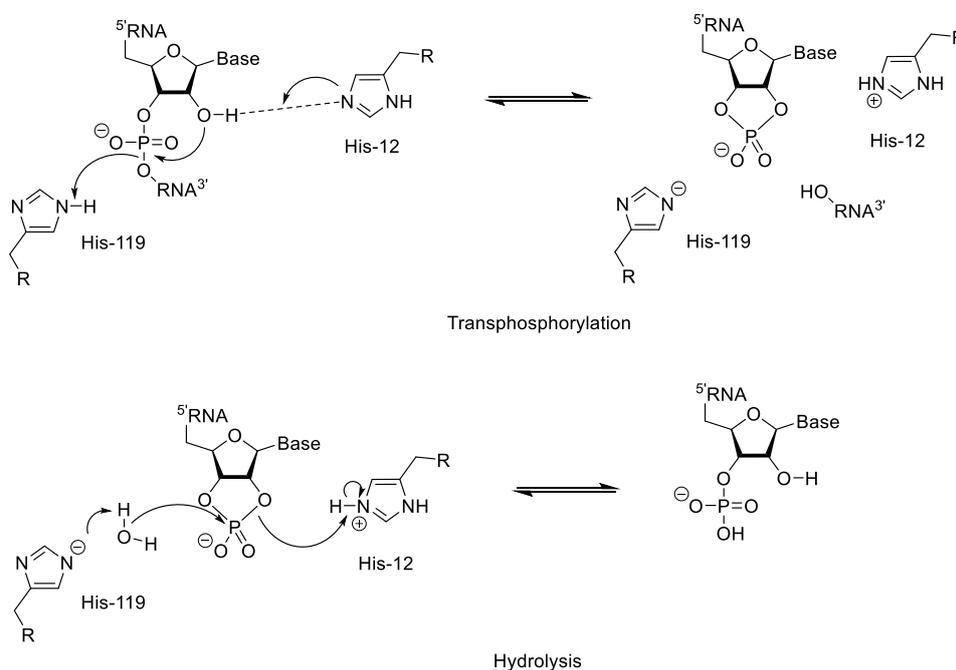
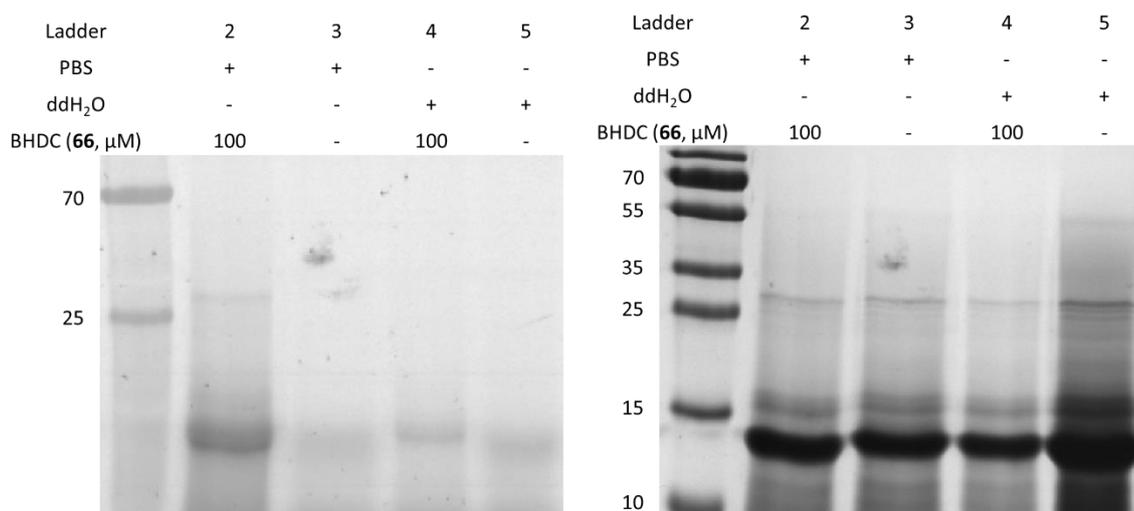


Figure 82 Proposed mechanism of bovine RNase A<sup>163</sup>

As of the time of writing, two samples of RNase A have been submitted for proteolytic digestion and mass spectrometry analysis in order to identify the modified residue in this enzyme. This final bit of data would ensure that the probe (**66**) works as hypothesised in an acid base histidine residue.

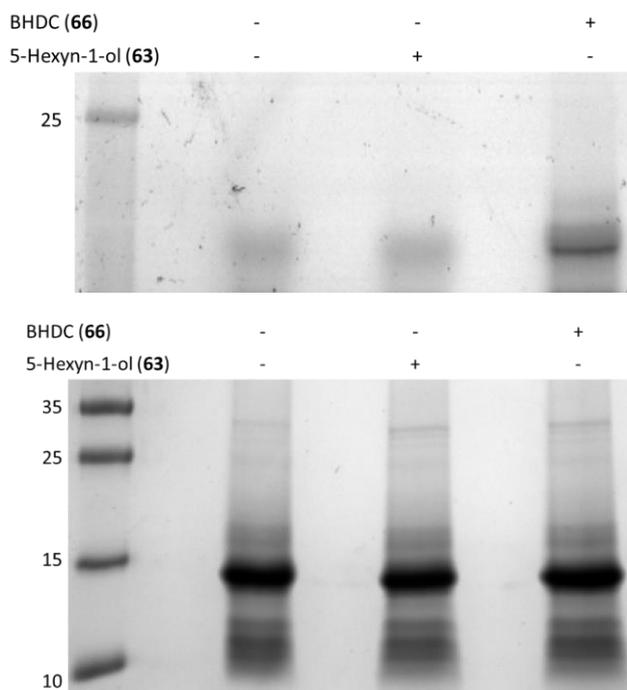
Following mass spectrometry analysis, fluorescent analysis was then performed. Initially, incubation of 1 mg/mL RNase A with 100  $\mu$ M BHDC (**66**) followed by a standard click chemistry procedure with the synthesised simple Rho-N<sub>3</sub> (**93**) previously used was performed (Figure 83).<sup>209</sup>



*Figure 83 Analysis of BHDC (66) reaction with RNase A across different solvents and concentrations. RNase A was prepared to 1 mg/mL across all the conditions. BHDC (66) was incubated for 1 h before the addition of the click chemistry reagents. Top: fluorescent analysis on FLA-3000 (Fujifilm;  $\lambda_{ex}$  473/  $\lambda_{em}$  520 nm); Bottom: overnight quick Coomassie protein stain. Representative SDS-PAGE from 3 consistent replicates.*

The results show that BHDC (66) labelling with RNase A occurred best in PBS at 100  $\mu$ M (Figure 83, lane 2), with reduced labelling seen in the corresponding ddH<sub>2</sub>O (Figure 83, lane 4). This could be due to the reduced enzyme activity without the presence of salt in the PBS buffer.

RNase A in ddH<sub>2</sub>O and PBS was again tested however with the inclusion of 5-hexyn-1-ol (63, 100  $\mu$ M). This result showed that only BHDC (66) interacted with RNase A (Figure 84).



*Figure 84 BHDC (66) and 5-hexyn-1-ol (63) incubation with RNase A*  
 Performed in PBS buffer at 1 mg/mL RNase A. BHDC (66) and 5-hexyn-1-ol (63) incubated to a final concentration of 50  $\mu$ M. Fluorescent image (top) taken on FLA-3000 (Fujifilm;  $\lambda_{ex}$  473/  $\lambda_{em}$  520 nm); overnight quick Coomassie protein stain shown bottom. Representative SDS-PAGE from 3 consistent replicates.

The collective evidence suggests that monoacylation using BHDC (66) is possible with an active site nucleophilic residue and acid base histidine residue. With the boundaries of the probe explored, implementation for enzyme isolation by streptavidin enrichment was the next aim. The carbamate was stable enough to perform fluorescent analyses, however this remained to be tested with the conditions needed for the elution from the streptavidin.

### 3.7 Pulldown purification

As described in Chapter 1.4, another technique that can be utilised in activity-based protein profiling (ABPP) is protein purification by streptavidin enrichment and elution. For this, the previously synthesised and discussed tetrafunctional probe (97, Chapter 2.6.1, Figure 85) was initially utilised to isolate PhoN and RNase A.

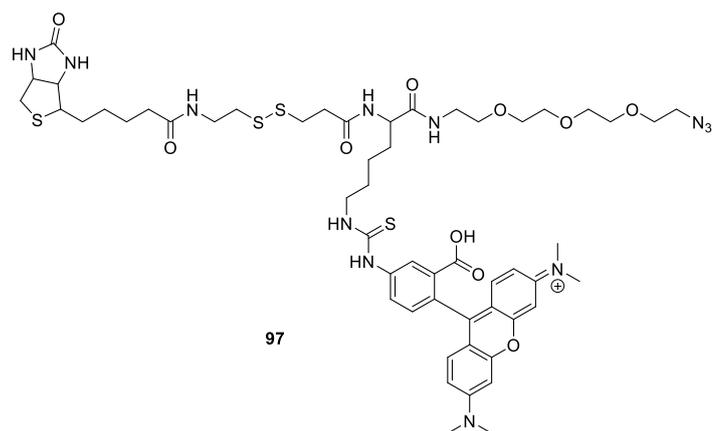
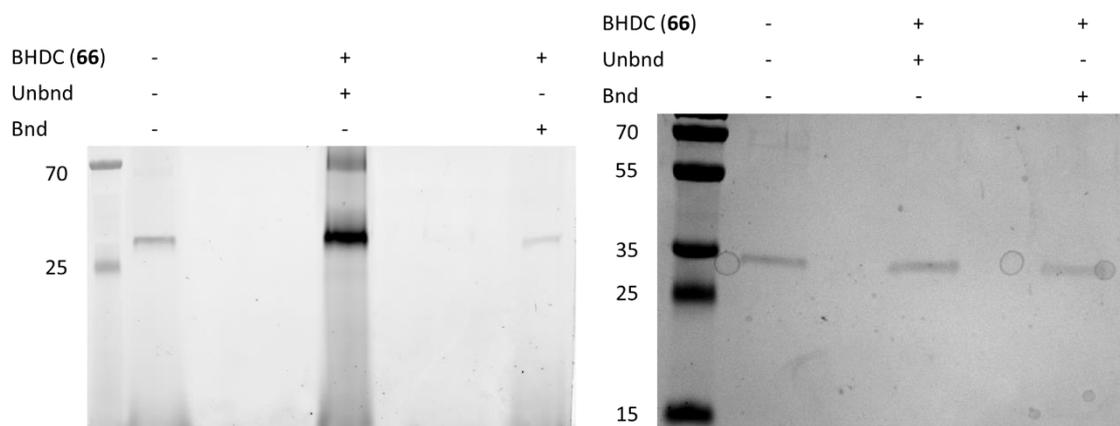
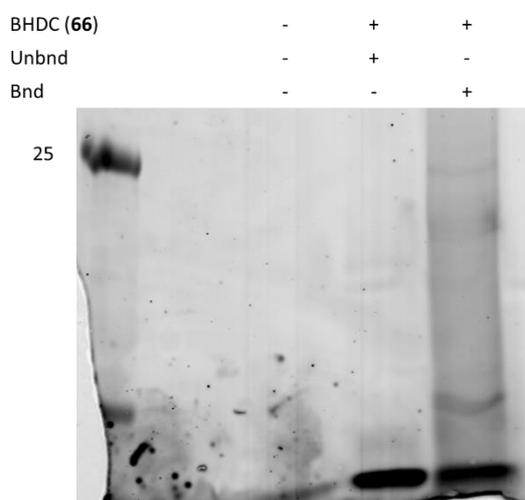


Figure 85 Synthesised tetrafunctional tag (**97**) for biotin enrichment

Briefly, the enzymes were incubated with 2.5  $\mu\text{M}$  BHDC (**66**) before the addition of 5  $\mu\text{M}$  tetrafunctional tag (**97**) and the click chemistry reagents. These concentrations were chosen due to the reported binding capacity of the used magnetic beads. Following a further 1 h incubation, the enzymes were added to the magnetic beads, which were prepared before the addition of the lysate as recommended by the manufacturer. Following 1 h incubation time with the magnetic beads, the unbound proteins were removed and analysed alongside the eluted proteins. To elute the proteins from the magnetic beads, 0.5 M  $\beta\text{ME}$  was added for 0.5 h before analysis by gel electrophoresis (Figure 86 and Figure 87)



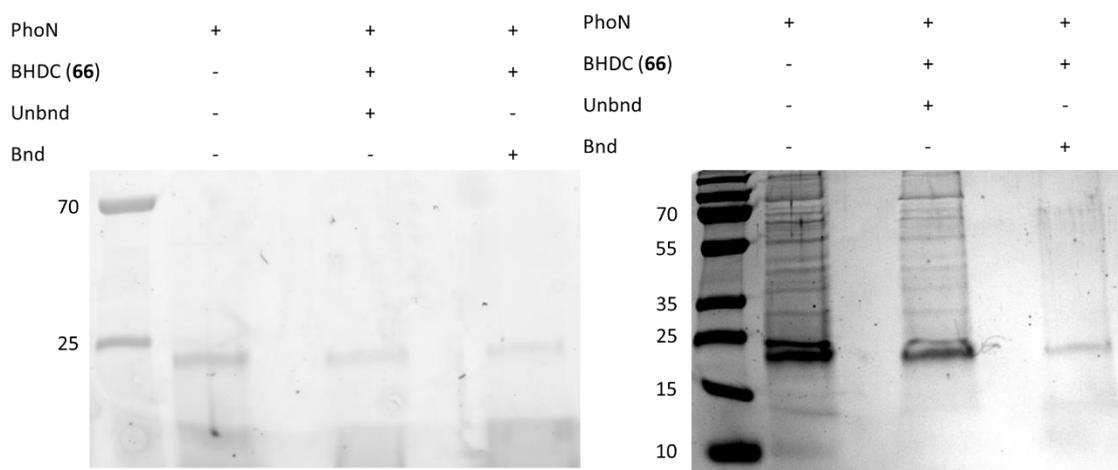
*Figure 86 Pull-down performed with PhoN*  
 2.5  $\mu$ M BHDC (66) was used, with 5  $\mu$ M tetrafunctional probe (97). Cleavage occurred using 0.5 M  $\beta$ ME. Left: fluorescent image obtained using Typhoon 9400 ( $\lambda_{ex}$  532 nm/  $\lambda_{em}$  580 nm); Right: corresponding quick Coomassie protein stain. Representative SDS-PAGE from 3 consistent replicates.



*Figure 87 Pull-down performed on RNase A*  
 2.5  $\mu$ M BHDC (66) was used, with 5  $\mu$ M tetrafunctional probe (97). Cleavage occurred using 0.5 M  $\beta$ ME. A fluorescent image obtained using Typhoon 9400 ( $\lambda_{ex}$  532 nm/  $\lambda_{em}$  580 nm).

Unexpectedly, the biotin pull-down using the tetrafunctional probe (97) repeatedly resulted in the control negative also being fluorescent using the pure PhoN (Figure 86) and pure RNase A (Figure 87). Due to the difficulties that working with a small enzyme like RNase A (13686 kDa) provides, this phenomenon was investigated further using PhoN only.

The tetrafunctional tag (97) was then implemented in overexpressed PhoN lysate to observe the range of this phenomenon (Figure 88).



*Figure 88 Pull-down experiment in PhoN overexpressed lysate*  
 2.5  $\mu$ M BHDC (66) was used, with 5  $\mu$ M tetrafunctional probe (97). Cleavage occurred using 0.5 M  $\beta$ ME. Left: fluorescent image obtained using Typhoon 9400 ( $\lambda_{ex}$  532 nm/ $\lambda_{em}$  580 nm); Right: corresponding quick Coomassie protein stain. Representative SDS-PAGE from 3 consistent replicates.

This repeatedly resulted in a band around the expected size of PhoN being fluorescent. This was not observed in previous experiments when the simple rhodamine azide (93) was implemented, and the initial reason for this was not obvious. Therefore, this was explored further by incubating PhoN with different combinations of reagents to try and identify the cause of this labelling (Figure 89).

BHDC ( <b>66</b> )	-	-	+	+	-	+	-
Tetra-tag ( <b>97</b> )	+	+	+	+	+	-	-
CuSO <sub>4</sub> ·5H <sub>2</sub> O	-	+	-	+	+	-	-
THPTA	-	-	-	+	+	-	-
Na asc	-	-	-	+	+	-	-

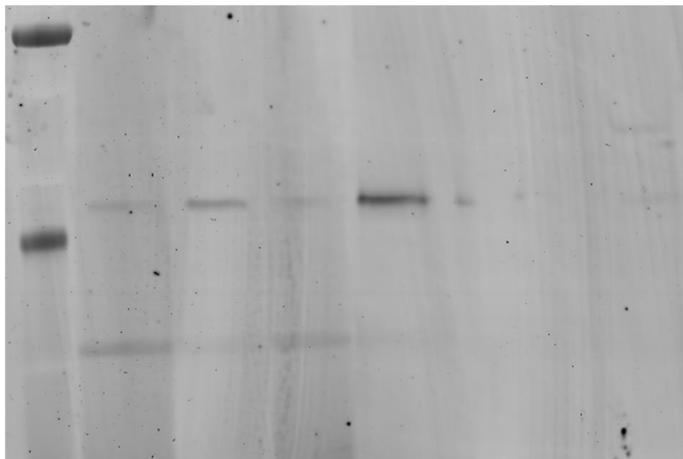


Figure 89 PhoN incubation with click chemistry components  
PhoN (0.25 mg/mL) was incubated with the compounds routinely used to perform click chemistry to the standard concentrations. Representative SDS-PAGE from 2 consistent replicates with different PhoN aliquots.

As seen in Figure 89, when PhoN was incubated with only the tetrafunctional tag (**97**), the enzyme was fluorescent. This was reproducible with a different PhoN aliquot and fresh click chemistry reagents, ruling out the possibility of the probe (**66**) contaminating the stock solutions. This resulted in the hypothesis that the intrinsic disulphide bonds present in PhoN and RNase A were undergoing a disulphide bridge exchange with that in the tetrafunctional linker (**97**). Due to this, the tetrafunctional probe (**97**) is not suitable for analysis with PhoN. Therefore, a commercially available biotin azide (**94**) was purchased (Figure 90).

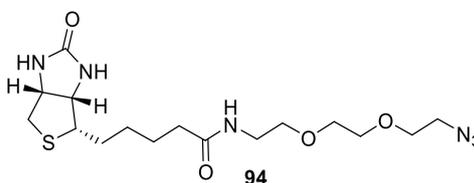


Figure 90 Commercially available biotin azide (**94**)

The subsequent enrichment and pulldown were performed with this compound following similar concentrations and conditions previously described for the tetrafunctional tag.

However, elution occurred by adding 50  $\mu$ L X4 LDS (5 %  $\beta$ ME) to the samples and boiling the magnetic beads for 10 min at 95  $^{\circ}$ C (Figure 91).

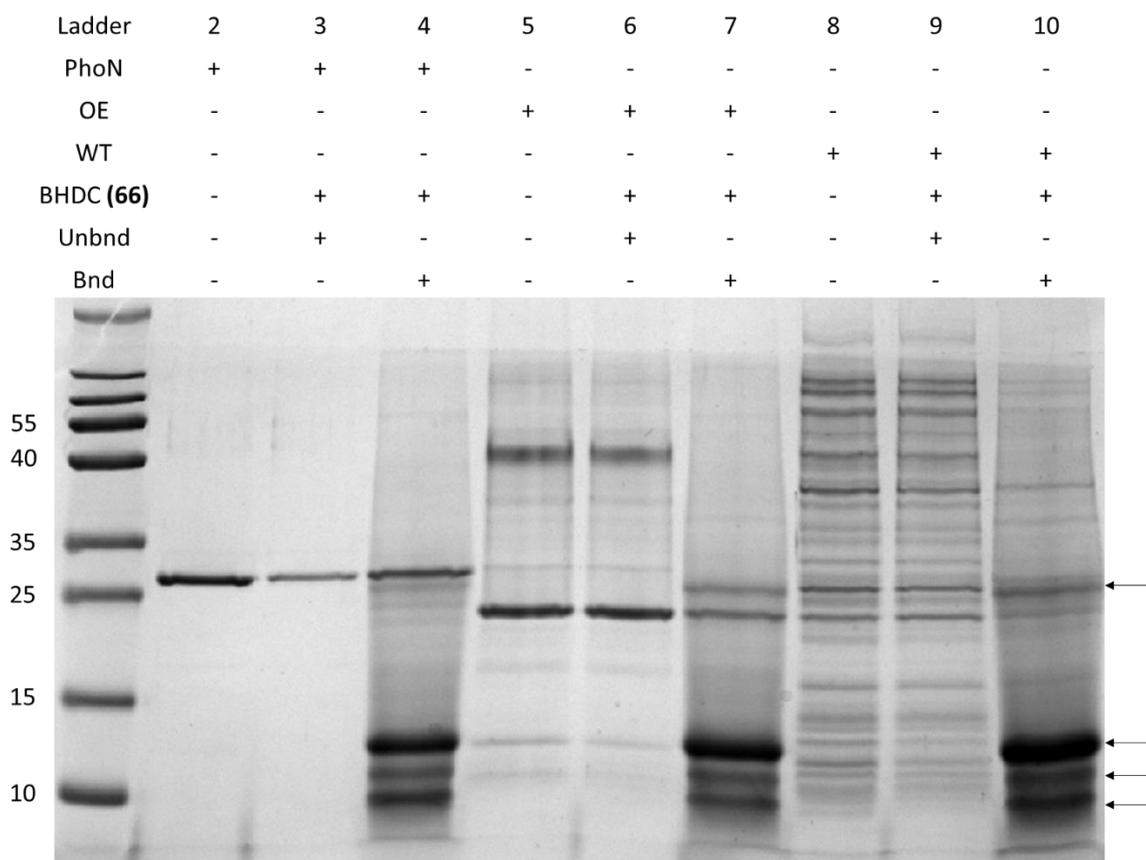


Figure 91 Coomassie stain of pull-down using a commercially available biotin azide. Arrows indicate streptavidin contamination from the elution procedure. OE: overexpressed lysate; WT: *E. coli* BL21 (DE3) pLysS lysate with the control plasmid; Unbnd: proteins removed in the wash step; Bnd: proteins eluted from the magnetic beads. Loading order: protein, washed proteins, eluted proteins. Representative SDS-PAGE from 3 consistent replicates with all the samples having 10  $\mu$ L loaded onto the gel.

Repeating this analysis but with the use of this new commercially available biotin azide (94) resulted in the successful pull-down of PhoN from the purified conditions (Figure 91, lanes 2-4). In addition, a band of a similar mass was isolated in the PhoN overexpressed lysate, which was later confirmed by mass spectrometry to be PhoN (Figure 91, lanes 5-7). In the case of the *E. coli* BL21 (DE3) pLysS with the control plasmid (TNF $\alpha$ ), several proteins were successfully eluted (Figure 91, lanes 8-10).

Following the successful pull-down of the unidentified proteins, the samples were then submitted for mass spectrometry to evaluate the probes (66) selectivity in a complete lysate. For this, the proteins were separated using SDS-PAGE and this was stained overnight

with Coomassie protein stain, and subsequently destained with Milli-Q H<sub>2</sub>O. Then, the protein bands were excised from the gel, eluted and submitted for mass spectrometry (Figure 92).

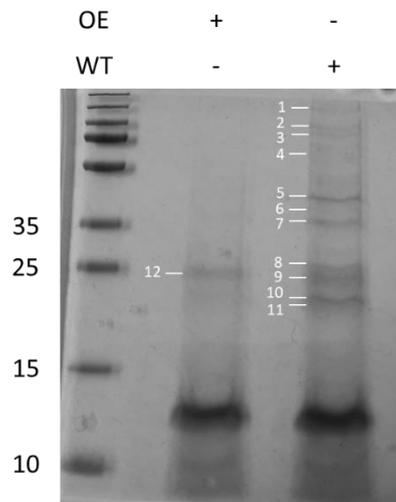


Figure 92 Bands eluted for mass spectrometry analysis  
 OE: lysate in which PhoN was overexpressed; Lys: lysate generated from the *E. coli* with the control plasmid. Gel stained overnight with quick Coomassie protein stain.

The eluted proteins were digested using trypsin and chymotrypsin, then the cysteine residues were reduced with TCEP, and finally the peptide fragments were identified with UniProt KB (see footnote<sup>1</sup>). To begin the analysis of this dataset, the reported protein contaminants resulting from the use of the pulldown technique were removed from the analysis, based on the report from Trinkle-Mulcahy *et al.*<sup>210</sup> This led to the removal of 18 proteins, leaving 298 proteins identified from the pulldown (Figure 93).

<sup>1</sup> This analysis was performed by Dr Adrian Brown (Durham University, Department of Biosciences) who identified the peptide fragments in the dataset

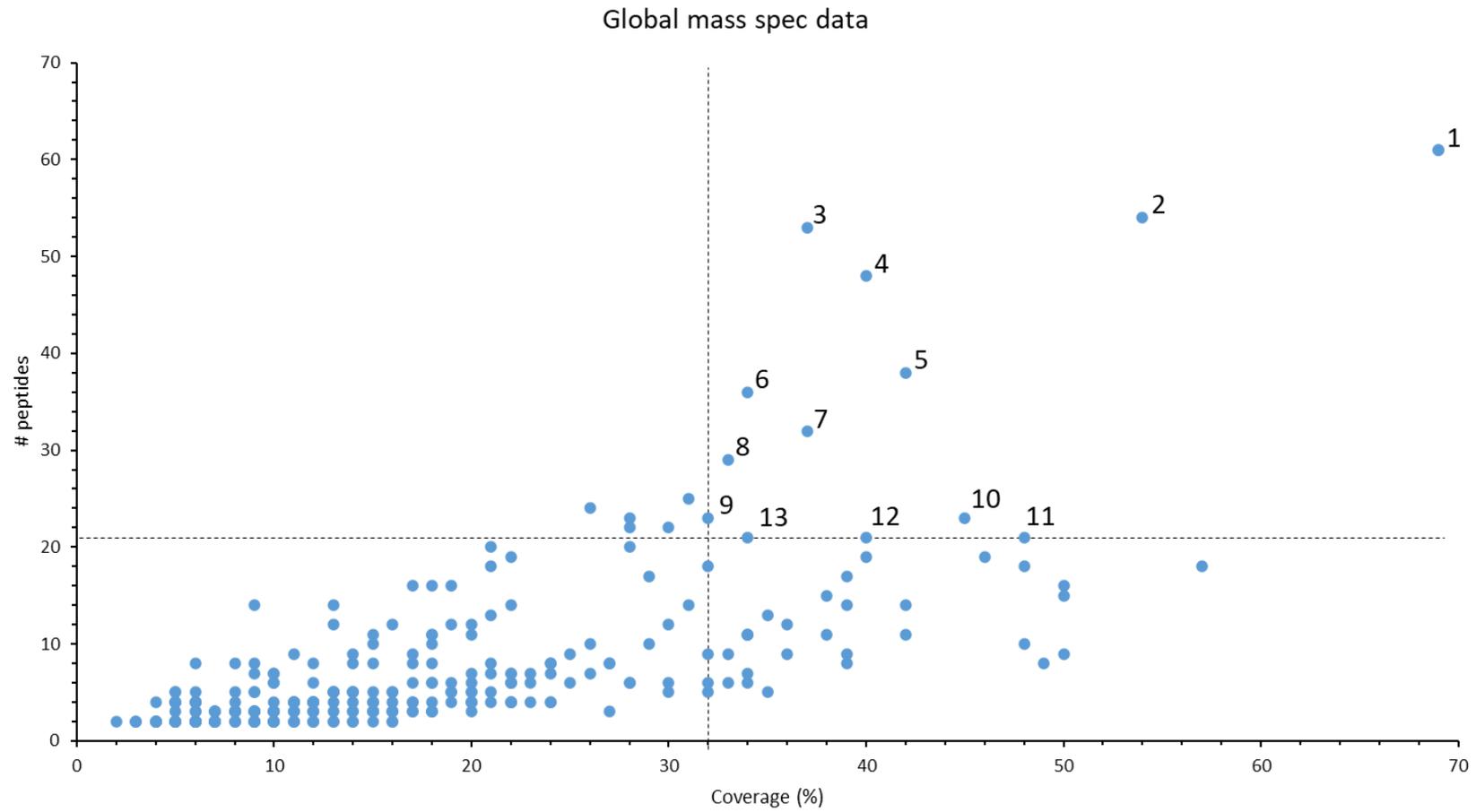


Figure 93 Global mass spectrometry analysis of the isolated proteins

Reported contaminants were removed from this analysis.<sup>210</sup> Lines represent the cut-off point, with the enzymes scoring above these markers being involved in the subsequent analysis.

The proteins were then sorted by the quantity of peptide fragments found of each enzyme (#peptides) and the percentage of the isolated sequence that aligns with the known sequence of the enzymes (% coverage) (Figure 93). Since the presented data is preliminary, with one repeat only, the top scoring proteins in these categories were chosen for in-depth analysis. This arbitrary cut off was chosen at #peptides:  $\geq 21$  and coverage:  $\geq 32\%$ , which divides Figure 93 into 4 quadrants. This cut off resulted in the average enzyme coverage being 44 %, which has some basis in the literature with Hacker *et al.* reporting an average coverage of 40 % when discussing a lysine activity based probe.<sup>139</sup> The majority of identified proteins were found in the bottom left quadrant, scoring low in both categories. Due to the low scores in these categories, these proteins could be due to non-selective binding, intrinsic biotinylation or contamination. The cut off points can be adjusted in future analysis based on pull down repeats and after a control negative performed to identify non-specific binding.

The cut-off point for this analysis may exclude enzymes that are true targets of BHDC (**66**) but have low intrinsic expression which would be reflected with low # peptide numbers. However, these would be hard to identify without repeats, and an internal control negative to identify contaminants. Therefore, the proteins found in the top right quadrant (Figure 93) were chosen for further analysis, which narrowed the dataset down to 12 enzymes. These enzymes were then grouped together based on enzymatic function (Table 3).

Table 3 Summary of isolated enzymes based on catalytic function and features

Uncharacterised: enzyme with an unknown active site function(s); Nucleophilic His: nucleophilic histidine residue in the active site; acid base histidine: a histidine residue which acts as an acid and as a base in the active site; off-target: non-histidine residue reactivity of BHDC (66).

Category	Total Identified	Protein number (n)
Nucleophilic histidine residue	2	10, 11
Acid base histidine residue	6	1, 2, 6, 8, 9, 13
Off target activity	2	5, 12
Uncharacterised active site	2	3, 4, 7

n	Enzyme name	Coenzyme / Cofactor	Coverage (%)	# peptides	Av mass (kDa)
1	Elongation factor Tu 1	GTP	69	61	43284
1	Elongation factor Tu 2	GTP	69	61	43314
2	Elongation factor G	GTP	54	54	77581
3	DNA-directed RNA polymerase subunit beta'	Mg <sup>2+</sup> , Zn <sup>2+</sup>	37	53	155160
4	DNA-directed RNA polymerase subunit beta	Mg <sup>2+</sup> , Zn <sup>2+</sup>	40	48	150632
5	Aconitate hydratase B	[4Fe-4s] cluster	42	38	93498
6	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase	(R)-lipoate, acetyl CoA	34	36	66096
7	Alanine-tRNA ligase	ATP, Zn <sup>2+</sup>	37	32	96032
8	Pyruvate dehydrogenase E1 component	Mg <sup>2+</sup> , thiamine diphosphate	33	29	99669
9	Threonine-tRNA ligase	ATP, Zn <sup>2+</sup>	32	23	74014
10	Succinate-CoA ligase	ATP, Mg <sup>2+</sup>	45	23	41393
11	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	-	48	21	28556
12	Phosphoglycerate kinase	ATP	40	21	41118
13	Polynucleotide phosphorylase	Mg <sup>2+</sup>	34	21	77101

Compared to the bands seen in the initial Coomassie protein stain (Figure 92), the quantity of identified enzymes is relatively low. This was due to some bands (9, 10 and 11) being primarily composed of streptavidin.

### 3.7.1 Pulldown protein composition

In general, the isolated enzymes can be categorised into 4 different groups according to the active site residue: nucleophilic histidine residue, an acid base histidine residue, off-target residue reactivity, and an uncharacterised active site(s). This was based on the function of the active site residues and how they function in the active site.

### 3.7.1.1 Nucleophilic histidine residue

Surprisingly, there was only a limited number of enzymes with a nucleophilic histidine residue identified. One such enzyme was succinate CoA ligase (#peptides: 21; 45 % coverage, n: 10), an enzyme that forms part of the citric acid cycle. In this enzyme, His-246 (*E. coli*) acts as a nucleophile, catalysing the production of succinate (**106**) from succinyl-CoA (**105**, Figure 94).<sup>211</sup>

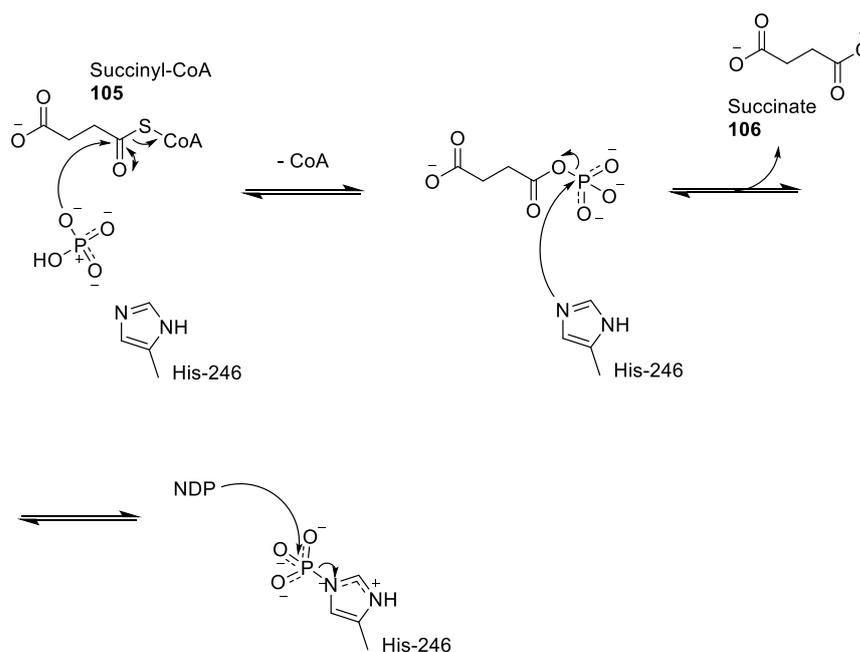


Figure 94 Catalytic mechanism for succinate CoA ligase  
NDP: nucleotide diphosphate which can be either ADP or GDP

This type of active site has been validated to bind to BHDC (**66**), with PhoN used as an example.

The other enzyme identified with a nucleophilic histidine was 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (BPG-dependent PGAM; #peptides: 21; 48 % coverage, n: 11). This enzyme is part of the glycolysis metabolic pathway, forming part of a two-step process to form the product 3-phosphoglycerate (3-PG, **109**) (Figure 95).<sup>212</sup> Along with this enzyme, phosphoglycerate kinase (n: 12) was also identified in the pulldown, which is discussed later.

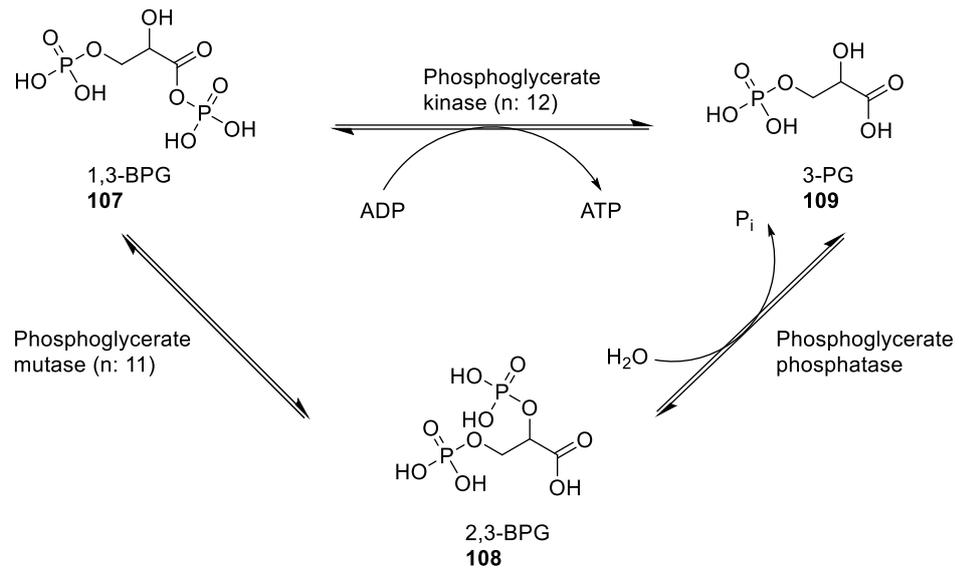


Figure 95 Part of the glycolysis pathway to produce 3-phosphoglycerate <sup>212</sup>

BPG-dependent PGAM has a conserved nucleophilic histidine identified as His-10 in *E. coli* catalysing the formation of 2,3-bisphosphoglycerate (2,3-BPG, **108**) from 1,3-bisphosphoglycerate (1,3-BPG, **107**, Figure 96). <sup>213</sup>

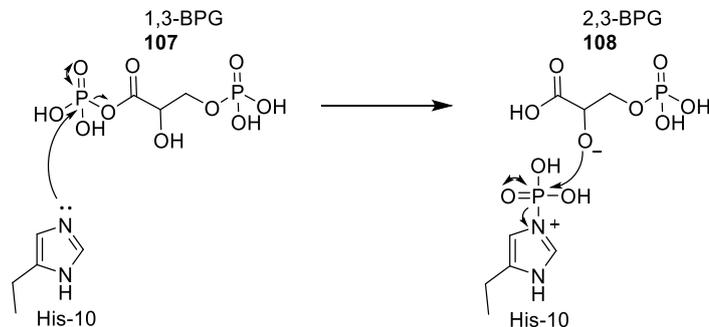


Figure 96 Hypothesised mechanism of 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase <sup>214</sup>

Along with this, it was hypothesised by Bond *et al.*, (2001) that His-183 could act as an acid base residue. <sup>215</sup> Other residues found in the active site region include Glu-88, which could act as a proton source for His-183 and also Arg-61, Arg-9 and Gly-184 which coordinate to the substrate. <sup>215</sup> This enzyme has a mass of 28.5 kDa and was isolated from band 9 (Figure 92). This is around the same area as PhoN is expressed and therefore it was concluded that this was likely the intrinsic *E. coli* enzyme that appears around same mass of PhoN, discussed previously (section 3.5.2).

### 3.7.1.2 Acid base histidine residue

The most common type of active site histidine residue identified from this experiment were acid base histidine residues. This could be due to the higher prevalence of histidine being found in this role, with Ribeiro *et al.* reporting histidine residues are 29 times more likely to be found in an acid base role as opposed to a nucleophilic one.<sup>123</sup> Interestingly, one such enzyme includes an intrinsic ribonuclease known as polynucleotide phosphorylase (#peptides: 21; 34 % coverage, n: 13). In this enzymes active site, His-403 (*E. coli*) is located in a highly conserved region, and mutation of this residue resulted in the loss of catalytic activity.<sup>216</sup> This histidine is positioned near the Mn<sup>2+</sup> metal cofactor and phosphate binding site, where acid base catalysis occurs.<sup>217</sup> Whilst the complete activity of the active site is yet to be elucidated, it is clear that there is a histidine residue that influences enzyme activity.

Other enzymes characterised with an acid base histidine residue included elongation factor Tu 1 (EF-Tu; #peptides: 61; 69 % coverage, n: 1a) and EF-Tu 2 (#peptides: 61; 69 % coverage, n: 1b) which have GTPase activity. These enzymes have the same function with 99.7 % sequence identity, differing only by the last amino acid on the C terminus, with this being Gly-394 in EF-Tu 1 and Ser-394 in Ef-Tu 2.<sup>218</sup> These are abundant enzymes, which have several functions but are primarily described to transport aminoacylated tRNAs to the ribosome during protein translation.<sup>218</sup> This requires forming the ternary complex EF-Tu:aminoacylated-tRNA, which transports to the ribonuclease and releases the product upon GTP hydrolysis when the codon is recognised.<sup>219</sup> To do this, a conserved acid base His-84 (*E. coli*) residue in the switch II region abstracts a proton from water, which then attacks the phosphate on GTP (Figure 98).<sup>220</sup>

Another identified GTPase with an acid base histidine residue in the active site was elongation factor G (#peptides: 54; 54 %, n: 2), which is involved in protein translation. This enzyme has a sequence identity of 16.5 % with EF-Tu, and shares a conserved acid base histidine residue.<sup>221</sup> The conserved His-91 in EF-G (*E. coli*) is an acid base histidine found near a hydrophobic region which has been associated with P<sub>i</sub> release.<sup>222</sup> An alignment between EF-Tu and EF-G showed that His-91 aligns with His-84 from EF-Tu which are found in the aforementioned switch II region (Figure 97).<sup>222</sup>

```

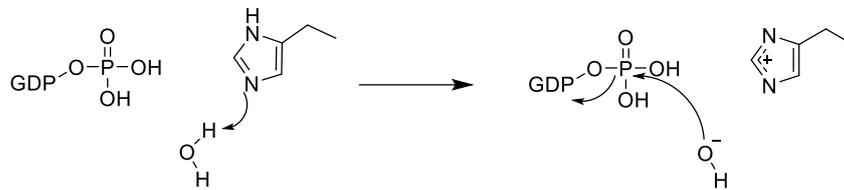
EF-Tu  52  N A P E E K A R G I T I N T S H V E -- Y D T P T R H Y A H ----- V D C P G H A D Y V K N M I T G A A Q M D G A I L 104
EF-G   52  W M E Q E Q E R G I T I T S A A T T A F W S G M A K Q Y E P H R I N I I D T P G H V D F T T I E V E R S M R V L D G A V M 111

```

Figure 97 Alignment of EF-Tu and EF-G

Yellow: conserved histidine residue in the active site; Black: other conserved residues

EF-G maintains the open reading frame during translation by guiding the A-site tRNA and translocating the deacylated tRNA, hydrolysing GTP in the process (Figure 98).<sup>223,224</sup> Given the related structure and function of these enzymes, it is unsurprising that both were identified in this study.

Figure 98 General mechanism for GTP hydrolysis<sup>225</sup>

In *E. coli*, pyruvate dehydrogenase is a complex of three enzyme subunits: pyruvate dehydrogenase (E1p), acetyltransferase (E2p) and lipoamide dehydrogenase (E3).<sup>226</sup> This complex catalyses the production of acetyl-CoA from pyruvate by oxidative decarboxylation, transferring the product of each region to the next active site by lipo-lysine residues.<sup>227,228</sup> Of this complex, the first two components, E1p and E2p, were identified in this pull-down assay, both with acid base histidine residues.

Pyruvate dehydrogenase (#peptides: 29; 33 % coverage, n: 8) E1p subunit catalyses the aerobic decarboxylation of pyruvate to form acetyl-CoA and CO<sub>2</sub>.<sup>226</sup> In *E. coli*, this enzyme subunit has a couple of notable histidine residues; His-142 binds to the cofactor, His-407 and His-640 interact with oxygen atoms, and His-106 binds to a water molecule.<sup>229</sup> Whilst the full catalytic mechanism is yet to be completely elucidated, the highly conserved His-407 residue is hypothesised to function as an acid base, which was likely the target of BHDC (66).<sup>230</sup>

The other subunit isolated from the pyruvate dehydrogenase complex was the E2p subunit, dihydrolipoyllysine-residue acetyltransferase (#peptides: 34; 36 % coverage, n: 6). This enzyme catalyses the transfer of the acetyl group from acetyl-CoA to dihydrolipoamide.<sup>231</sup>

A mutation of His-602 (*E. coli*) resulted in a loss of functionality, leading to the hypothesis that this residue forms part of the active site. <sup>232</sup> In support of this, a recent report by Škerlová *et al.* (2021) identified this histidine residue as an acid base histidine by cryo-EM analysis (Figure 99). <sup>233</sup>

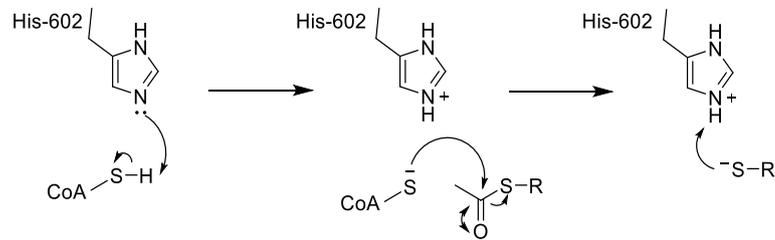


Figure 99 Proposed mechanism for E2 acetyltransferase subunit <sup>231</sup>

In the active site, Asp-607 stabilises His-602, and Ser-551 stabilises the transition state, resulting in the transfer of the acetyl from dihydropyrimidinone to CoA. <sup>233</sup>

Another enzyme identified was threonine tRNA ligase (<sup>Thr</sup>tRNA ligase, #peptides: 23; 32 % coverage, n: 9). This is part of the aminoacyl-tRNA synthetases which link the amino acid to the respective tRNA for protein translation. <sup>234</sup> A mutational analysis on His-309 from intrinsic *E. coli* <sup>Thr</sup>tRNA ligase showed a loss of enzyme function. His-309 was identified as an acid base histidine residue, interacting with the 2'-OH on RNA directly or indirectly through a water molecule (Figure 100). <sup>235</sup>

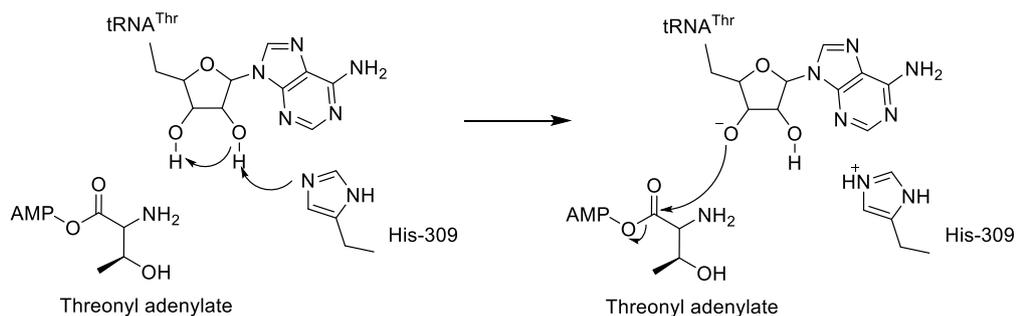


Figure 100 Hypothesised mechanism of <sup>Thr</sup>tRNA ligase <sup>235</sup>

Other notable residues in close proximity and highly conserved are Tyr-462 and Gln-484, both forming hydrogen bonds with the substrate. <sup>235</sup>

### 3.7.1.3 Off target residues

One of the enzymes isolated from the pulldown that had unknown residue reactivity was aconitate hydratase B (#peptides: 38; 42 % coverage, n: 5). In *E. coli*, there are two aconitate hydratase enzymes A and B (sequence identity: 11 %).<sup>236</sup> Aconitate hydratase B is highly expressed in exponential growth and catalyses the reversible isomerisation of citrate and isocitrate.<sup>237</sup> The HEAT-like domain is unique to aconitate hydratase B (*E. coli*) and is a cysteine rich iron-sulphur cluster.<sup>238</sup> However, due to the presence of the nucleophilic cysteine residues, BHDC (**66**) reactivity with this domain is plausible. Whilst a histidine residue has been described as an acid base in both the human and bovine orthologues of this enzyme, no such activity has been described for aconitate hydratase B for *E. coli* to date.<sup>239,240</sup>

The other enzyme isolated that does not have an active site histidine residue was phosphoglycerate kinase (#peptides: 21; 40 % coverage, n: 12). As alluded to previously, this enzyme is part of the glycolysis pathway, catalysing the production of 3-phosphoglycerate (**109**) from 1,3-bisphosphoglycerate (**107**), generating ATP in the process (Figure 95).<sup>241</sup> The substrate 1,3-bisphosphoglycerate (**107**) is bound by 3 arginine and 1 histidine, but more significantly, there are 5 highly conserved Cys residues; two of which are close to the active site and 2 are located near the C terminus which are involved with redox regulation (*Homo sapiens*).<sup>242,243</sup> However, in *H. sapiens*, Lys, Arg and Asp were identified as catalytically active residues, with Lys hypothesised to interact directly with the substrate, ADP.<sup>243</sup> An alignment showed that the catalytic residues were also conserved in the lesser characterised *E. coli* phosphoglycerate kinase, and with a sequence identity of 27.3 %, the catalytic activity are likely similar (Figure 101).

```

HsPGK 1 MSLSNKLTLDKLDVKGKRVMRVDENVPMKNNQITNNQRITKAAVPSIKFCLDNGAKSVVL 60
EcPGK 1 MSV---IKMTDLDLAGKRVFIRADENVVVKDGKVTSDARIRASLPTIELALKQGAKVMV 56

HsPGK 61 MSHLGRP-----DGV----- 70
EcPGK 57 TSHLGRPTEGEYNEEFSLLPVVNYLKDKLSNPVRLVKDYL DGV DVAEGELVVLENVRFNK 116

HsPGK -----
EcPGK 117 GEKKDDETL SKKYAALCDVFVMDAFGTAHRAQASTHGIGKFADVACAGPLLAELDALGK 176

HsPGK 71 -----PFLAILGGAKVADKIQLINMLDKVNEMITGGMAFTFLKVLNNMEIGTSLFD 123
EcPGK 177 ALKEPARPMVAIVGGSKVS TKLTVLDSL SKIADQLIVGGGIANTFIAAQGH-DVGKSLYE 235

HsPGK 124 EEGAKIVKDLMSKAENGVK TITLPVDFVTADKFDENAKTGQTTVASGIPAGWMGLDCGPE 183
EcPGK 236 ADLVDEAKRLLTTCN-----IPVPSDVRVATEFSETAPATLKSVDNDVKADEQILDIGDA 289

HsPGK 184 SSKKYAEAVTRAKQIVWNGPVGVFWEAFARGTKALMDEVVKATSRGCITII GGGDTATC 243
EcPGK 290 SAQELAEILKNAKTI LWNGPVGVFEPNFRKGT EIVANAIADSEA---FSIAGGGDTLAA 346

HsPGK 244 CAKWNTEDKVS HVSTGGGASLELLEGGVLPVGDALSNI--- 281
EcPGK 347 IDLFGIADKISYI STGGGAFLEFVVEGKVLPAVAMLEERAKK 387

```

Figure 101 Alignment of *H. sapiens* and *E. coli* phosphoglycerate kinase

HsPGK: *Homo sapiens* phosphoglycerate kinase; EcPGK: *E. coli* phosphoglycerate kinase. Yellow: the aligned active site residues; black: conserved residues.

The exact target residue of BHDC (66) with phosphoglycerate kinase is unknown however, the likely targets are either one of the conserved cysteine residues or the active site lysine residue.

#### 3.7.1.4 Uncharacterised active site residues

Some of the proteins isolated from the pulldown had unconfirmed active site residues. One such example was an enzyme known as DNA-directed RNA polymerase (#peptides: 53; 37 % coverage, n: 3 and 4). This protein is one of the largest proteins to be found in a bacterial cell with a size of 151 kDa and possessing multiple active sites with separate functions.<sup>244</sup> Due to this, it is unclear whether there is an active histidine residue in this protein. There are reports of point mutations in His-526 (His526Tyr; His526Asp; His526Asn; His526Tyr) being associated with increased resistance to the antibiotic rifampicin, but the reasons for this are unknown.<sup>244</sup>

Another uncharacterised enzyme is alanine tRNA ligase (#peptides: 32; 37 % coverage, n: 7). This enzyme has a domain called a Cys-His box, in which mutation of the histidine inactivates the enzyme.<sup>245</sup> As previously discussed, BHDC (66) did isolate an enzyme with a reactive cysteine residue and due to this, it is challenging to guarantee that BHDC (66) reacted with a histidine residue rather than a cysteine residue.

### 3.7.2 Summary

With the highest scoring the enzymes from the pulldown characterised, there does not appear to be many trends outside of the active site histidine residue. Apart from histidine, 2 enzymes were isolated with off target activity with at least one enzyme, aconitate hydrolase B, in which BHDC (**66**) could be reacting with an active site cysteine residue. Whilst this was undesired, cysteine is the most nucleophilic active site residue, and many histidine probes are reported to cross react with Cys, including DEPC (**50**).<sup>179</sup> However, without an internal control, the analysis is speculative. With appropriate DMSO controls, the off target reactivity can be clearly defined.

### 3.8 Chapter summary

The described experiments thus far show that chemoselectivity of BHDC (**66**) can be retained within <10.0 eq and <3.3 eq in nucleophilic and acid base histidine residues respectively. In addition, the electron deficient aryl carbamate formed is stable to CuAAC for fluorescent analysis and enrichment pull down studies. This supports the hypothesis that BHDC (**66**) can be implemented for the study of enzymes with active site histidine residues in whole cell proteomes.

With these conditions derived, research then progressed to the study of the proteome from *Leishmania*. As discussed in Chapter 1, this is a neglected tropical disease with limited ideal chemotherapy.<sup>2</sup> Increasing the characterisation of the leishmanial proteome can contribute to the advancement of chemotherapies for this disease.<sup>95</sup>

## 4 Biological testing of ABP in *Leishmania*

### 4.1.1 Synopsis of chapter

As discussed throughout Chapter 3, the data for BHDC (**66**) thus far appears to suggest that the probe is functioning in the view of the ultimate project goal. A key objective of this project was to apply this tool for the exploration of the *Leishmania* proteome, which is discussed in this chapter.

### 4.2 Probe application to *L. mexicana* lysate

Whilst BHDC (**66**) appeared to function in the hypothesised manner in *E. coli*, application to a new organism such as *Leishmania* required additional calibration. To begin, lysis conditions were explored to obtain functioning proteins from *Leishmania mexicana*. Firstly, sonication was explored as a lysis method using two different buffers (Figure 102).

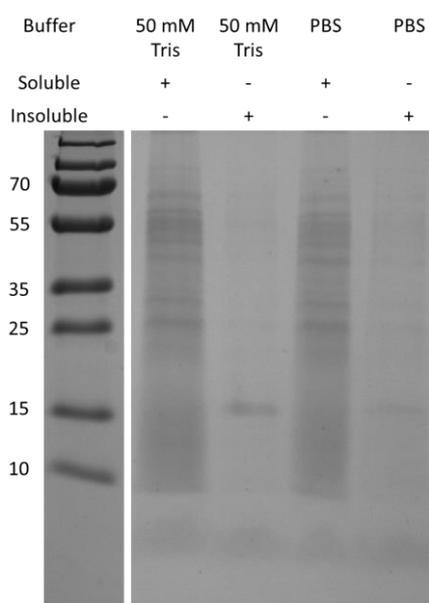
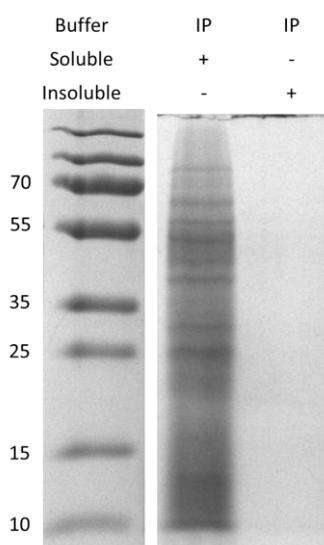


Figure 102 SDS PAGE for the comparison of lysis conditions  
Tris: tris buffer, 50 mM; pH 7.2; PBS: pH 7.4. Both buffers are supplemented with the recommended quantity of protease inhibitors as detailed by the manufacturer (Merck). Representative SDS-PAGE from 3 consistent replicates

Whilst the different buffers produced no discernible difference, the protein yield was poor. Therefore, it was concluded that sonication, the method used for *E. coli* lysis was not an effective lysis method for *Leishmania*. After a brief literature review, it became apparent that detergents were commonly used to produce *Leishmania* lysate.<sup>246–248</sup> To test this, a

modified detergent-based IP lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl and 5 % glycerol) was used following manufacturers procedures (Figure 103).



*Figure 103 Lysis of L. mexicana using IP lysis buffer*  
 IP lysis buffer: 25 mM Tris-HCl pH 7.4, 150 mM NaCl and 5 % glycerol. Lysis was performed following manufacturer procedures. SDS-PAGE stained with Coomassie overnight.

IP lysis buffer successfully produced higher yields of soluble lysate; however limited insoluble protein was detected. Once suitable lysis conditions were determined, the lysate was dialysed into HEPES buffer (20 mM, pH 7.4). This was to remove tris, which acts as a nucleophile on the probe leading to rapid degradation, and EDTA which would chelate to the copper necessary for CuAAC. The lysates were then treated with 10  $\mu$ M BHDC (**66**) for 1 h before SDS-PAGE analysis (Figure 104).

BHDC (66)	+	+	+	+	-	-	-	-
Time (h)	1	1	24	24	1	1	24	24
Soluble	+	-	+	-	+	-	+	-
Insoluble	-	+	-	+	-	+	-	+

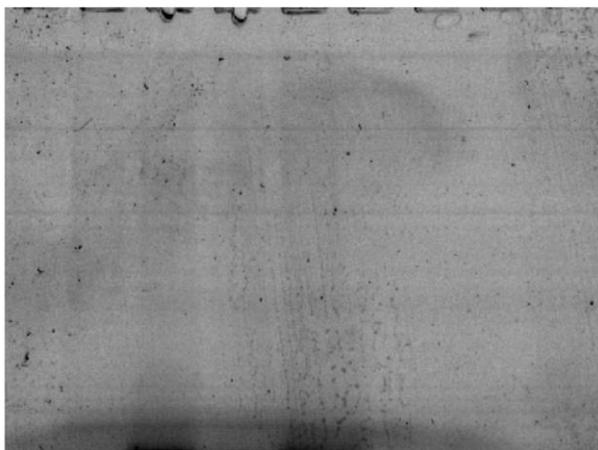


Figure 104 Fluorescent analysis of 1 h and 24 h following incubation with 10  $\mu$ M BHDC (66) BHDC (66) incubated to a final concentration of 10  $\mu$ M. Fluorescent analysis conducted using Y520 filter, 473 nm wavelength on Fuji FLA-3000. Representative result from 3 consistent repeats.

Initial fluorescent analysis repeatedly generated no bands. The Coomassie protein stain was indicative of some protein degradation, with fewer distinct protein bands observed (Figure 105). This was hypothesised to have happened during the dialysis, as the proteins stained soon after lysis appeared in good condition.

BHDC (66)	+	+	+	+	-	-	-	-
Time (h)	1	1	24	24	1	1	24	24
Soluble	+	-	+	-	+	-	+	-
Insoluble	-	+	-	+	-	+	-	+

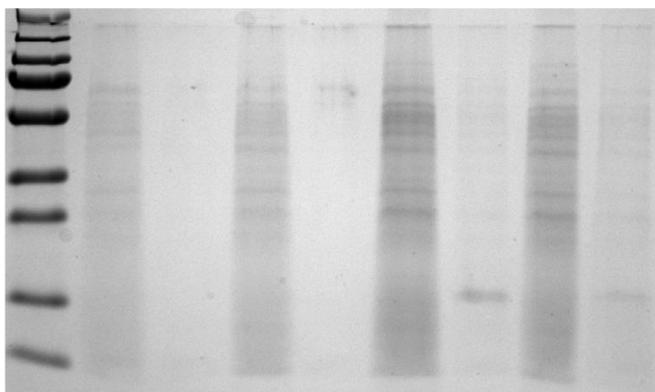


Figure 105 Coomassie stain of *L. mexicana* lysate after BHDC (66) incubation The CuAAC reaction was quenched at the specified time points and the lysate of the parasites then generated

The lack of obvious protein bands from these experiments indicated that the conditions previously used to work with *E. coli* lysate are unsuitable for *Leishmania*. This could be due to variations in protease activity between the organisms, which would reduce the potency of the broad-spectrum protease inhibitor used during the generation of these lysates. This has some literature support, with papers describing *Leishmania* proteases as virulence factors and reporting that cysteine and serine proteases are particularly abundant.<sup>249–251</sup> Therefore, whole cell incubation was thought to be ideal to analyse the proteome of *Leishmania* to ensure the proteins are functional when BHDC (**66**) is used. However, it was uncertain whether the probe will pass through the membrane of *Leishmania* to label the proteome.

### 4.3 Promastigote inhibition assay

Whilst there is evidence that BHDC (**66**) is unhindered by the Golgi apparatus membrane, as discussed in Chapter 3.2, there was still uncertainty whether it will pass through the parasite's plasma membrane. To evaluate the permeability of BHDC (**66**), a promastigote inhibition assay was used. It was hypothesised that if the probe can pass through the plasma membrane, this would lead to cell death due to broad-spectrum inhibition of enzymes.

*Leishmania mexicana* (M379,  $0.5 \times 10^6$  cells mL<sup>-1</sup>) were incubated with a 3 fold dilution series of BHDC (**66**) for 48 h. Cell viability was assessed using the standard cell viability agent alamar blue. Initially, BHDC (**66**) appeared to show no antileishmanial activity, with no cell death detected at 100  $\mu$ M BHDC (**66**). This could be due to a lack of permeability, however both the previously predicted LogP of 2.41 and inhibition of IPC production in a biochemical assay (Chapter 3.2) suggests some ability to navigate the membrane. Therefore, it was hypothesised that BHDC (**66**) may be hydrolysed in the Schneider's medium during the assay preparation before the application to *L. mexicana*.

To overcome this, the conditions applied by Splittstoesser and Wilkinson (1973) for the study of DEPC (**50**) was utilised for this assay.<sup>252</sup> For these adjusted conditions, instead of diluting the probe in Schneider's medium before application to the parasites, the probe was diluted in DMSO to 10 mM and applied to the parasites ( $0.5 \times 10^6$  cells mL<sup>-1</sup>). Subsequently, a serial dilution was performed to achieve a 3-fold dilution. Along with the

probe (**66**), 5-hexyn-1-ol (**63**), and DEPC (**50**) were also tested, with amphotericin B (AmpB) and cycloheximide as positive controls. The viability was quantified by applying alamar blue after 48 h and measuring the fluorescence. After this resulted in the viability of the parasites being  $<100 \mu\text{M}$ , the assay time was reduced to 24 h. This could be due to the fast degradation of the probe allowing surviving parasites to continue replicating. This observation corresponded to what was described by Spittstoesser and Wilkinson reported when using DEPC (**50**).<sup>252</sup>

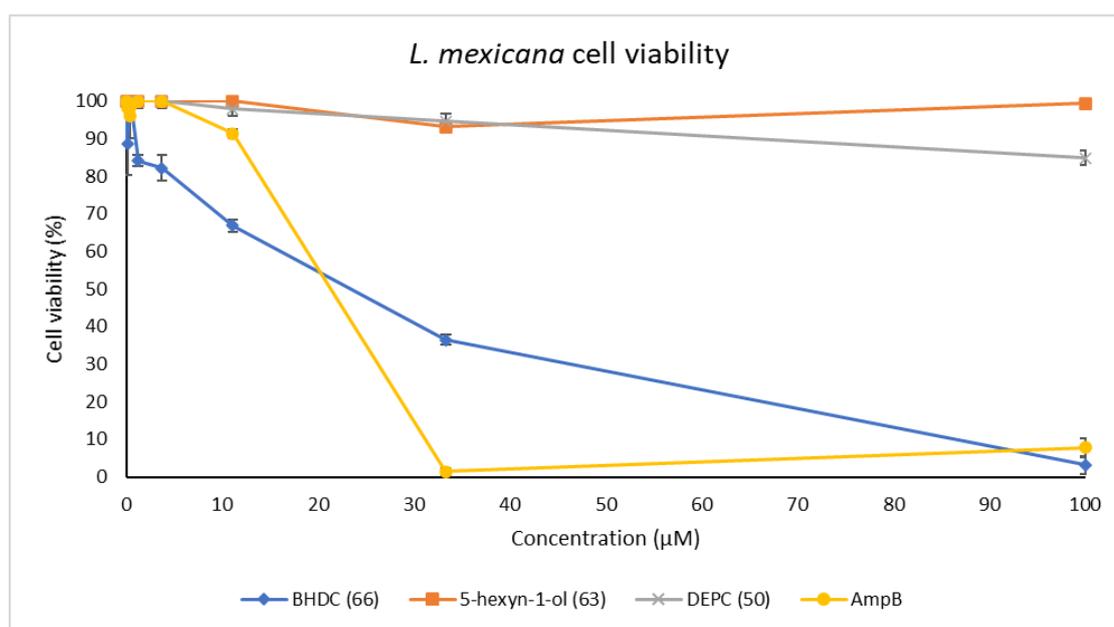


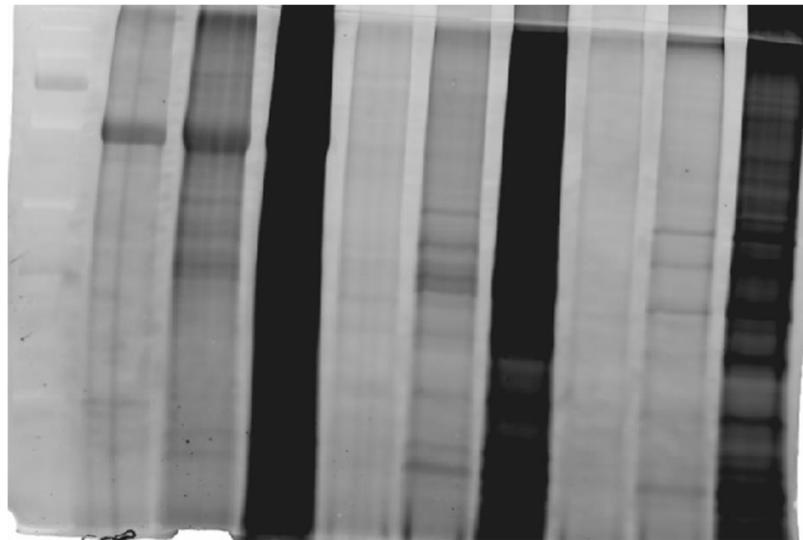
Figure 106 *Leishmania mexicana* cell viability  
Cell viability measured using rezasurin dye. AmpB: amphotericin B which was used as a control.

At a concentration of  $100 \mu\text{M}$ , BHDC (**66**) killed 97 % of *L. mexicana* in comparison to 5-hexyn-1-ol (**63**), which had no detectable effect on the parasites (Figure 106). This shows that the probe is likely passing through the membrane, with the by-product alcohol having no effect on the population. With the validation of whole cell parasite incubation, the project moved on to the application of the probe **66** on whole cell *L. mexicana*.

#### 4.4 Exploration of the *Leishmania* proteome

Following validation of BHDC (**66**) membrane permeability, focus shifted to developing a whole cell methodology for probe application. To begin, the optimal concentration of BHDC (**66**) and incubation time was explored (Figure 107).

Probe ( <b>66</b> , $\mu\text{M}$ )	1	10	100	1	10	100	1	10	100
<i>L. amaz</i>	+	+	+	-	-	-	-	-	-
<i>L. mex</i>	-	-	-	+	+	+	-	-	-
<i>L. maj</i>	-	-	-	-	-	-	+	+	+



Time (h)    0.5    1    2    4    6    24

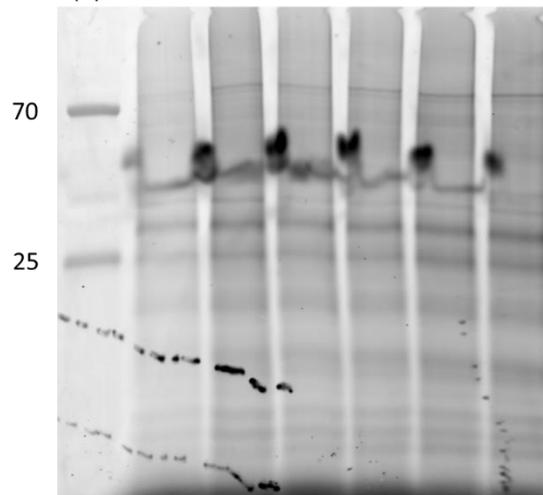


Figure 107 Calibration experiments in *Leishmania*

Fluorescent images taken on the Typhoon 9400 ( $\lambda_{\text{ex}}$  532 nm/  $\lambda_{\text{em}}$  580 nm) and analysed in the programme ImageQuant 5.2. Top: image of several *Leishmania* species with varying concentration of BHDC (**66**); Bottom: *Leishmania mexicana* lysate with 10  $\mu\text{M}$  BHDC (**66**) at varying incubation times. Lysates were generated using IP lysis buffer from which the lysates were stored in for analysis.

As shown in Figure 107, the optimal probe (**66**) concentration in a 15 mL *Leishmania* culture was 10  $\mu\text{M}$ . Following this, the optimal time was analysed using only the species *L. mexicana* due to a functioning CRISPR-Cas9 system, which could be used for the eventual validation of protein targets. From this, 1 h was chosen for future experiments because there was no benefit to longer incubation times. These simple experiments showed that

BHDC (**66**) generated the same profile across multiple time points alluding to specificity of binding in *L. mexicana* whole cell conditions. With the fluorescent analysis now functioning in *L. mexicana*, the isolation of the targeted proteins was then tested.

#### 4.4.1 Pulldown experiment

As previously discussed in Chapter 3.5, the tetrafunctional tag (**97**) was found to be unsuitable for testing with PhoN due to a hypothesised disulphide bridge exchange with PhoN. However, due to the advantages of using **97** for a pulldown such as ease of tagged protein removal, and sensitive fluorescent analysis of the eluted proteins, this tetrafunctional tag (**97**) was implemented for the study of the *L. mexicana* proteome. For this, 10  $\mu$ M BHDC (**66**) was applied to whole cell *L. mexicana* and incubated for 1 h. Then the cells were lysed using IP lysis buffer and 5  $\mu$ M tetrafunctional tag was applied to the lysate with the other click chemistry reagents. Using streptavidin coated magnetic beads, the tagged proteins were isolated, eluted using 0.5 M  $\beta$ ME and analysed by SDS-PAGE (Figure 108).

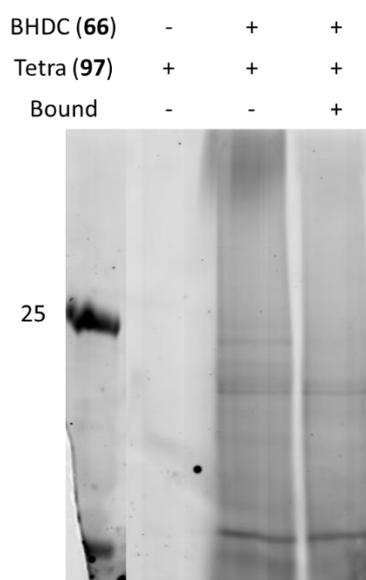


Figure 108 Pulldown performed on *L. mexicana*  
Fluorescent image taken on Typhoon 9400 scanner ( $\lambda_{ex}$  532 nm/  $\lambda_{em}$  580 nm). Bound refers to the proteins eluted from the pulldown experiment

As seen in Figure 108, the pulldown assay led to the isolation of labelled proteins. Unlike in *E. coli*, there did not appear to be any disulphide bridge exchange between the tetrafunctional tag (**97**) and the proteins in *Leishmania*. Whilst this was a promising start,

the identification of the isolated proteins is unknown. The size of *L. mexicana* IPCS is currently unknown however it could be around 37.8 kDa – 42.6 kDa based on *L. major* and *L. donovani* respectively. From the isolated proteins, there is not a prominent band around this size. To validate the inhibition of intrinsic IPCS using fluorescent analysis, differential centrifugation could be utilised to isolate the Golgi apparatus and thus, potentially visualise IPCS by fluorescent analysis. In addition, mass spectrometry analysis can be used, as this was a powerful tool for the identification of intrinsic *E. coli* enzymes, as described in Chapter 3.7.

#### 4.5 Chapter summary

No enzyme labelling was observed when BHDC (**66**) was incubated with the lysate of promastigote *Leishmania*. This was hypothesised to be due to rapid protein decomposition resulting in the inactivation of the enzymes. Upon the validation that BHDC (**66**) can pass through the plasma membrane using a promastigote inhibition assay, fluorescent labelling and a pulldown assay was then achieved. Following this, the proteins were submitted to mass spectrometry for further analysis.

## 5 Conclusions and Future work

### 5.1 Initial hypothesis

The initial hypothesis of this project was to develop an activity-based probe (ABP) based on the structure of diethyl pyrocarbonate (DEPC, **50**), a known acylating agent for active site histidine residues.<sup>179</sup> Following the identification of a lead compound bis(hex-5-yn-1-yl) dicarbonate (**66**), the function and specificity of the probe was explored in enzyme models (Figure 109).

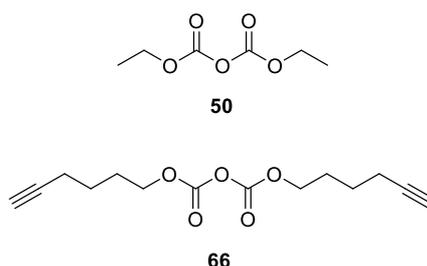


Figure 109 Known histidine labelling agent DEPC (**50**) and developed histidine labelling agent BHDC (**66**)

### 5.2 Conclusion

The scope of BHDC (**66**) labelling was tested on two enzyme models: PhoN with a nucleophilic histidine residue and ribonuclease A with an acid base histidine residue. This resulted in the monoacylation of the two enzyme models, with confirmed labelling on the nucleophilic His-197 and selective labelling of active RNase A. The probe was then implemented into the lysate of *E. coli* BL21 (DE3) pLysS to isolate and identify intrinsic enzymes of similar reactivity leading to 8 out of the 12 isolated proteins having a reported active site histidine residue (Table 4).

Table 4 Profile of the enzymes isolated from *E. coli* BL21 (DE3) pLysS

Category	Total Identified	Protein number (n)
Nucleophilic histidine residue	2	10, 11
Acid base histidine residue	6	1, 2, 6, 8, 9, 13
Off target activity	2	5, 12
Uncharacterised active site	2	3, 4, 7

Whilst there was the undesired off-target reactivity, the majority of the isolated enzymes had an active site histidine residue. This enzyme was then calibrated to function in

*Leishmania*, which produced a protein fingerprint in *L. mexicana*. Due to time constraints, no proteins were identified in this organism.

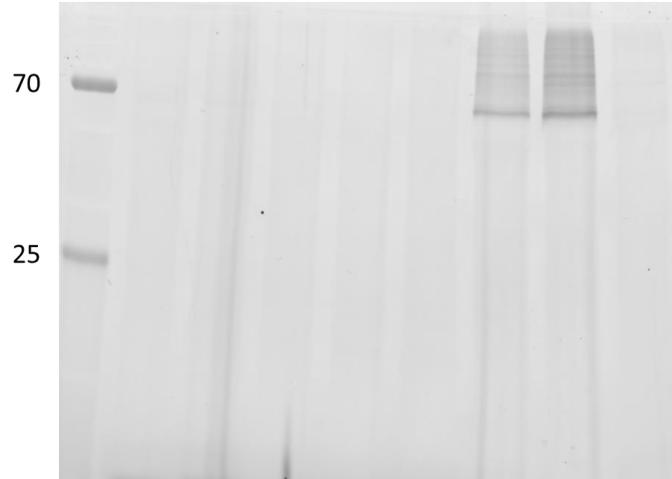
### 5.3 Future work

As described in Chapter 3.6, BHDC (**66**) successfully monoacylated ribonuclease A (RNase A) within 3.3 eq. In addition, this labelling appeared to be selective for the active enzyme, which could imply that the labelling is in the active site. However, without a proteolytic digestion and mass spectrometry analysis, this remains uncertain. To complicate things further with this analysis, there are two active site histidine residues in addition to the cysteine rich sequence. To date, two samples of RNase A have been submitted for restriction digestion analysis for the identification of the residue labelled with BHDC (**66**).

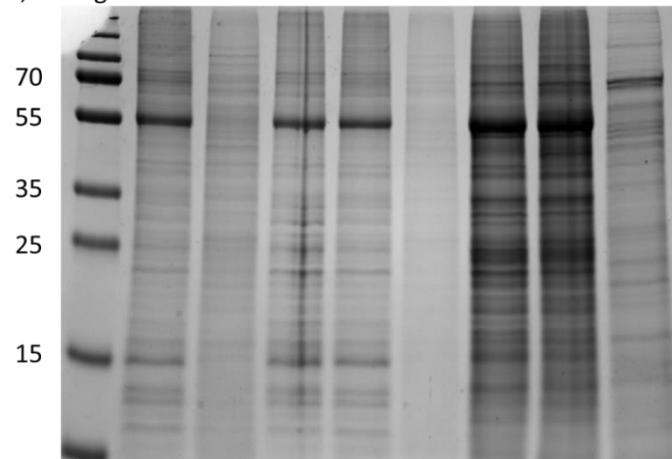
As described in Chapter 3.7.1, BHDC (**66**) was able to isolate a variety of different enzymes from the lysate of *E. coli* BL21 (DE3) pLysS. The presented data was preliminary, with only one repeat and no control negative. Therefore, to reinforce this dataset, two more repeats are currently undergoing analysis, with the addition of a control negative. This will allow for the identification of contaminations, allow for the identification of lower expressed modified enzymes, and increase the reliability of the reported enzymes.

Additionally, to build on the established *Leishmania* work described in Chapter 4, the research on leishmaniasis can be expanded. As previously described in Chapter 1.2.2, IPCS is an enzyme of interest as this has been described to be a good drug target and possesses a HHD catalytic triad.<sup>73,84</sup> As previously discussed in Chapter 3.2, a preliminary assay showed that NBD-C<sub>6</sub>-IPC (**101**) production was inhibited when *Lmj*IPCS was incubated with BHDC (**66**) indicating that the probe potentially modifies this enzyme. To explore whether the probe (**66**) bound to IPCS *in situ*, preliminary work commenced which used differential centrifugation on *L. mexicana* lysate treated with BHDC (**66**). The initial conditions followed those described by Mina *et al.*, for the isolation of the previously used *Lmj*IPCS enriched microsomes.<sup>92</sup> This involved removing large membrane fragments at 10,000 x g, followed by the isolation of mitochondria and larger organelles at 23,000 x g and lastly, the Golgi apparatus at 150,000 x g. To isolate enough material for this experiment, the procedure was performed on 50 mL *L. mexicana* culture in stationary phase (Figure 110).

Ladder	2	3	4	5	6	7	8	9
BHDC (66)	-	-	+	+	+	+	+	+
10,000 x g	S	+	S	+	+	-	-	-
23,000 x g	-	-	-	-	-	+	-	-
150,000 x g	-	-	-	-	-	-	+	S



Ladder	2	3	4	5	6	7	8	9
BHDC (66)	-	-	+	+	+	+	+	+
10,000 x g	S	+	S	+	+	-	-	-
23,000 x g	-	-	-	-	-	+	-	-
150,000 x g	-	-	-	-	-	-	+	S



*Figure 110 Isolation of Golgi apparatus*  
 Numbers correspond to centrifuge speed, with Golgi apparatus reportedly isolated at 150,000 x g; S: remaining supernatant following differential centrifugation. Top: Fluorescent image taken on the Typhoon 9400 ( $\lambda_{ex}$  532 nm/  $\lambda_{em}$  580 nm) and analysed in the programme ImageQuant 5.2; Bottom: corresponding gel stained with quick Coomassie overnight.

After an overnight Coomassie protein stain, the speeds 23,000 x g (Figure 110, lane 7) and 150,000 x g (Figure 110, lane 8) fractions did not appear to have different protein compositions, suggesting that isolation of the Golgi apparatus was unsuccessful. In

addition, the fluorescent image showed no difference in labelling between these fractions and did not separate properly on the SDS-PAGE. Subsequently, to ensure separation between the two fractions, 23,000 x g centrifugation was performed twice to ensure complete removal of these proteins. The proteins were all quantified with a BCA assay and diluted to 1 mg/mL but resolubilizing the Golgi apparatus was difficult, which is reflected in the poor running of the 150,000 x g fraction (Figure 111).

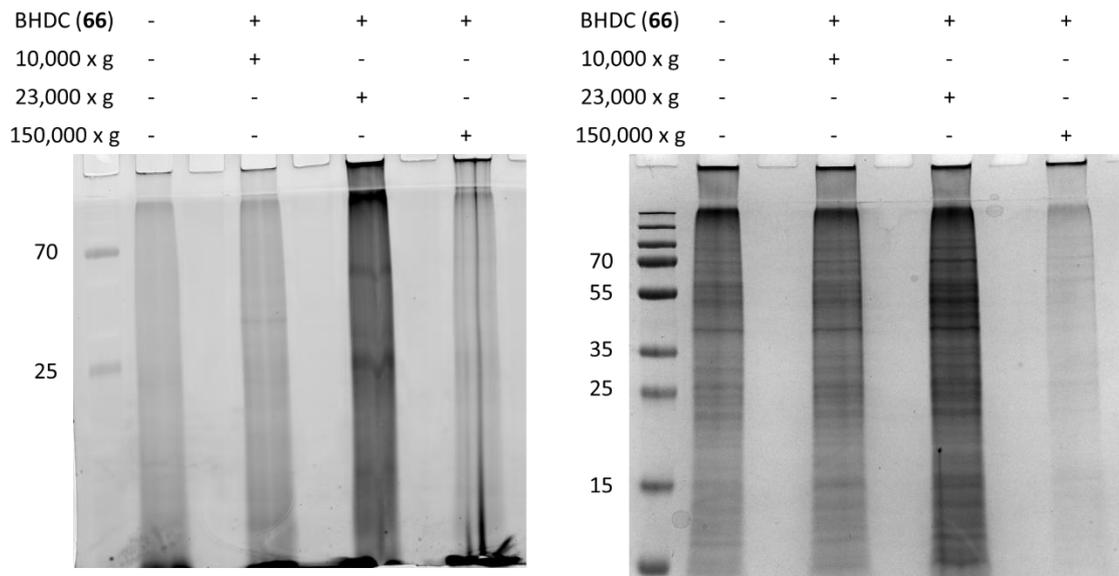


Figure 111 Differential centrifugation on BHDC (66) treated *L. mexicana*. Numbers correspond to centrifuge speed, with Golgi apparatus reportedly isolated at 150,000 x g. Fluorescent Left: image taken on the Typhoon 9400 ( $\lambda_{ex}$  532 nm/ $\lambda_{em}$  580 nm) and analysed in the programme ImageQuant 5.2; Right: the corresponding gel stained with quick Coomassie overnight.

The additional centrifugation at 23,000 x g resulted in the formation of a small pellet upon the second repeat, which could explain the previous result. Additionally, the fluorescent image between 23,000 x g and 150,000 x g conditions now appears different, suggesting successful removal of the mitochondrial fraction from the Golgi apparatus. However, loading and running the Golgi apparatus was now a challenge due to the difficulty resolubilizing the pellet.

Whilst the initial differential centrifugation attempts were unsuccessful at identifying IPCS isolation, the alternative method is the identification of the bound proteins using mass spectrometry, as described for *E. coli* proteins (Chapter 3.7.1). Following this, a fingerprint of the proteome changes in the different lifecycle stages of *Leishmania* can be explored, which would generate an expression protein profile. Validation of the protein targets could

occur using CRISPR-Cas9 to generate a knock-out of the suspected target, which would provide a qualitative profile on the identified proteins. From this pool of enzymes, evaluation for a drug target could lead to a target-based drug discovery project.

Another proposed probe could be developed to visualise labelling in a live cell for fluorescent microscopy. This would require the core dicarbonate reactive centre and a fluorophore that can pass through a plasma membrane, like NBD (**109**, Figure 112)

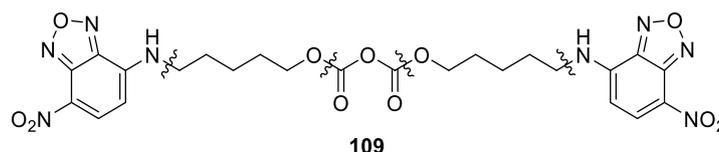


Figure 112 Hypothesised structure of a potential probe for whole cell fluorescent microscopy

This has the advantage of whole cell fluorescence microscopy, visualising which organelles are labelled providing additional evidence for hypothesised enzyme targets. However, there could be issues with steric hindrance thus, preventing active site labelling.

#### 5.4 Broader application

As a new activity-based probe, this can be used as a tool in research outside of the applications described in this thesis. For example, this probe could be used to identify the target of an anti-leishmanial drug. This can be done by preincubating *Leishmania* with the drug and measure the fluorescent decrease in the target.

As well as implementing BHDC (**66**) for the study of the proteome in *Leishmania*, it could be utilised to explore the proteome of vast variety of organisms and cancers. Currently, there are no other reported ABPs that can isolate the enzymes targeted by BHDC (**66**), making this a unique addition to ABPP.

## 6 Biology methodology

### 6.1 General equipment

Centrifugation used CAPP Rondo mini centrifuge CR-1512 (CAPP), Rotina 380 R (Hettich), Beckman Coulter Avanti JXN-26 or the Beckman Coulter ultracentrifuge (Optima XL). The orbital incubator SI500 (Stuart) was used for bacterial incubation. Sonication was performed at the individually specified conditions with Sonopuls (Bandelin). SDS-PAGE gels were made to 1 mm thickness, 12.5 % acrylamide and ran using Bio-rad mini protein tetra system electrophoresis equipment. To visualise DNA gels or Coomassie stained SDS-PAGE, Bio-rad Molecular Imager Gel Doc XR+ was used with the programme Image Lab (Version 6.0.1). Fluorescent images were taken using FLA-3000 (Fujifilm;  $\lambda_{\text{ex}}$  473/  $\lambda_{\text{em}}$  520 nm) and visualised with the programme AIDA image analyser (version 3.52) or Typhoon 9400 (Amersham Biosciences;  $\lambda_{\text{ex}}$  532 nm/  $\lambda_{\text{em}}$  580 nm) with the programme ImageQuant (version 5.2). The Sciex TripleTOF 6600 was used for protein mass spectrometry restriction digestion. Optical density and DNA concentration was quantified using DeNovix DS-11+ Spectrophotometer.

50 mL falcon tubes (Starlab), 15 mL centrifuge tubes (Starlab) and 1.5 mL microcentrifuge tubes (Starlab) were used throughout. Pipette tips (10  $\mu\text{L}$ , 100  $\mu\text{L}$  and 1000  $\mu\text{L}$ ) purchased from Starlab were autoclaved at 121 °C for 15 mins before use. Lysate protein content was determined using Pierce Rapid Gold BCA Protein Assay Kit (Merck) following manufacturers procedures and pure protein content was estimated using DeNovix DS-11+ Spectrophotometer.

### 6.2 Buffers and Medias

Details of the buffers, media components and concentrations used throughout this research are detailed in Table 5.

*Table 5 Buffer compositions used throughout.*

Buffer	Stock solution	Quantity	Final conc.
<b>STE (50 mL)</b>	Tris HCl (1 M)	1.25 mL	25 mM
	Sucrose (1 M)	12.5 mL	250 mM
	EDTA (0.5 M)	1 mL	1 mM
	Protease inhibitor tablet*	1 tablet	

<b>TAE X10 (1 L)</b>	Tris base	48.5 g	40 mM
	Glacial acetic acid	11.4 mL	20 mM
	EDTA (500 mM, pH 8)	20 mL	
<b>Storage (50 mL)</b>	Tris HCl (1 M)	2.5 mL	50 mM
	Glycerol (10 mL)	6.25 mL	20 % v/v
	MgCl <sub>2</sub> .6H <sub>2</sub> O (1 M)	0.25 mL	5 mM
	Protease inhibitor tablet*	1 tablet	
<b>PBS (500 mL)</b>	NaCl	4 g	137 mM
	KCl	100 mg	2.7 mM
	Na <sub>2</sub> HPO <sub>4</sub>	720 mg	10 mM
	KH <sub>2</sub> PO <sub>4</sub>	120 mg	1.8 mM
<b>HEPES (500 mL)</b>	HEPES	2.38 g	20 mM
	1 M HCl (pH adjustment)		
<b>Wash buffer (500 mL)</b>	Tris base (pH 7.4)	1.21 g	20 mM
	NaCl	14.61 g	500 mM
	1 M HCl (pH adjustment)		
	1 M NaOH (pH adjustment)		
	Protease inhibitor tablet*		
<b>Elution buffer (500 mL)</b>	Tris base (pH 7.4)	1.21 g	20 mM
	NaCl	14.61 g	500 mM
	Imidazole	17.02 g	500 mM
	1 M HCl (pH adjustment)		
	1 M NaOH (pH adjustment)		
<b>Salt elution buffer (500 mL)</b>	Tris base (pH 8.5)	1.21 g	20 mM
	NaCl	29.22 g	1 M
<b>SDS electrophoresis buffer (X10, 1L, pH 8.3)</b>	Tris base	30.2 g	25 mM
	Glycine	141.0 g	250 mM
	SDS	10 g	0.1 %
<b>IP lysis buffer (50 mL)</b>	Tris base	0.15 g	25 mM
	NaCl	0.44 g	150 mM
	Triton™ X-100	0.50 mL	1 % (v/v)

	Glycerol	2.50 mL	5 % (v/v)
	Protease inhibitor tablet*	1 tablet	

For buffers with protease inhibitors, cOmplete™ EDTA-free Protease Inhibitor Cocktail (Merck) was used following manufacturers directions.

*Table 6 Media components*

<b>Media</b>	<b>Stock solution</b>	<b>Quantity</b>
<b>LB Broth (1 L)</b>	Tryptone	10 g
	Yeast extract	5 g
	NaCl	10 g
<b>LB Agar (500 mL)</b>	LB Broth powder mixture	17.5 g
	Agar	7.5 g
<b>Schneider's Insect Medium (1 L)</b>	Schneider's Insect Medium	24.5 g
	NaHCO <sub>3</sub>	0.4 g
	CaCl <sub>2</sub>	0.6 g
	1 M HCl (pH adjustment)	
	1 M NaOH (pH adjustment)	

All biological reagents were dissolved in biological grade DMSO (Fisher Bioreagents) before application, unless otherwise specified. Biological grade buffers and medias were prepared using biological grade water from a Milli-Q purification system. The ethanol, sodium dodecyl sulfate (SDS) and PBS tablets were purchased from Fisher Scientific. The LB broth was purchased from Melford and prepared following manufacturer's instructions before being autoclaved at 121 °C for 15 min. Imidazole and glycine were from Acros organics. Trizma® base and Schneider's Insect Medium were from Merck. Agarose was purchased from Bio-Rad. (+)-Sodium L-ascorbate and HEPES were purchased from Fluorochem.

The leading probe BHDC (**66**) synthesis is described in Chapter 7. Before use, the stock solution was made fresh every time. The probe was diluted using DMSO unless specified otherwise.

### 6.3 Microsomal assay materials

NBD-C<sub>6</sub>-ceramide (Invitrogen) made to a concentration of 5 mM using DMSO. PI (Bovine liver, predominant species: 18:0/20:4, 1-*sn*-stearoyl-2-*sn*-arachidonoyl sodium salt, Merck) and PI (Bovine liver, 1,2-Diacyl-*sn*-glycero-3-phospho-(1-D-*myo*-inositol) ammonium salt, Merck) were made to a final solution of 1 mM in CHCl<sub>3</sub>. CHAPS (Melford) was made to 3 mM in Milli-Q grade H<sub>2</sub>O. Eppendorf LoBind Microcentrifuge Tubes (Merck, 1.5 mL) were used throughout this assay for handling microsomal *Lmj*IPCS, Eppendorf® vacuum concentrator 5301 was used when specified. Thinwall Polypropylene tubes (Beckman Coulter) were used for differential centrifugation with the Beckman Coulter ultracentrifuge (Optima XL). TLC plates (F254 silica gel aluminium sheets pre-coated with silica gel, Merck) were used for the separation of NBD-C<sub>6</sub>-ceramide (Merck) and NBD-C<sub>6</sub>-IPC, imaged with the FLA-3000 (Fujifilm) plate reader and analysed with AIDA image analyser software.

#### 6.3.1 *Lmj*IPCS inhibition assay methodology

Stock solution 1 (Table 7) was prepared for each reaction. To this, the desired compounds were added, and incubated for 25 mins at 30 °C, with agitation (150-180 rpm). For stock solution 2, the total volume of needed PI (1 mM) was dried in the bottom of a LoBind Eppendorf tube using the Eppendorf concentrator 5301 for each reaction. To this, the quantities for stock solution 2 (Table 7) were added, with a brief vortex to ensure the components were thoroughly mixed. Following incubation with stock solution 1, 23.75 µL of stock solution 2 was added to each reaction and this was further incubated for 30 mins using the same conditions.

Table 7 The components for each stock solution used for the *Lmj*IPCS inhibition assay

<b>Buffers and reagents</b>	<b>Stock 1 (µL/reaction)</b>	<b>Stock 2 (µL/reaction)</b>	<b>[Final]</b>
Phosphate buffer (71.9 mM, pH 7.0)	17.5	17.5	50 mM
CHAPS (3 mM)	5	5	
Washed IPCS microsome (1 U µL <sup>-1</sup> )	0.25		
Storage buffer	1		
NBD-C <sub>6</sub> -ceramide ( <b>100</b> , 200 µM)		1.25	5 µM
PI ( <b>21</b> , 1 mM)		5	

The samples were quenched with 150  $\mu\text{L}$  of 10  $\text{CHCl}_3$ :10 MeOH:3  $\text{H}_2\text{O}$  and vortexed. The organic layer was separated by centrifugation (8 mins, 13,300 rpm) and the organic layer was transferred to a new LoBind Eppendorf and dried at room temperature. The samples were resuspended in 20  $\mu\text{L}$  of the 10  $\text{CHCl}_3$ :10 MeOH:3  $\text{H}_2\text{O}$  and 3 x 3  $\mu\text{L}$  was spotted on the TLC plate. The products were separated using the solvent system 55:45:10 MeOH:  $\text{CHCl}_3$ : 0.25 %  $\text{KCl}_{\text{aq}}$  and the result was visualised using FLA-3000 (Fujifilm) with the program AIDA image analyser.

#### 6.4 Promastigote inhibition assay materials

*Leishmania mexicana* (strain 0157), *L. major* FV1, and *L. amazonensis* (MHOM/Br/75/JOSEFA) were grown in Nunc EasyFlask 25  $\text{cm}^2$  Nunclon Delta Surface (ThermoScientific) at 30 °C. The promastigotes were grown in Schneider's insect medium (Merck, Table 6) supplemented with 15 % heat inactivated fetal bovine serum (ThermoFisher Scientific, South American origin), and 1 % Gibco™ Penicillin-Streptomycin (PenStrep, 10,000 U/mL, ThermoFisher Scientific), and filter sterilised using Sterile filter flasks (500 mL, 0.22 microns, Merck) and the filtrate stored at 4 °C. The promastigote inhibition assay was conducted in 96 well sterile assay plates (Fisher Scientific), with StarTub Reagent Reservoir (Starlabs) for multichannel pipetting. Resazurin solution (500  $\mu\text{M}$ , 5 mg in 40 mL PBS, Acros organics) was added, and fluorescence measured with fluorescence plate reader (555 – 585 nM, BioTek Synergy HT). All tested controls were dissolved in biological grade EtOH unless otherwise specified and amphotericin B (Apollo Scientific) and cycloheximide (Merck) were used for the control positives.

##### 6.4.1 Promastigote inhibition assay method

All steps were performed under sterile conditions, with a passage number <15. The frozen *Leishmania* stocks were stored in Cryovials (Starlabs) and kept at -150 °C for long term storage. The frozen *Leishmania* aliquot were rapidly defrosted before use and then the *Leishmania* promastigotes were grown in culture flasks with 15  $\text{cm}^3$  supplemented Schneider's Insect Media at 26 °C. The cell count was adjusted to  $0.5 \times 10^6$  cells  $\text{mL}^{-1}$  for the assay. 100  $\mu\text{L}$  of the adjusted  $0.5 \times 10^6$  cells  $\text{mL}^{-1}$  *L. mexicana* in Schneider's Insect Media was pipetted in a 96 well plate for rows B-G and 148.5  $\mu\text{L}$  for row A. The tested compounds

were dissolved in EtOH to a stock concentration of 10 mM, with BHDC (**66**) stock solution being made fresh for every use. 1.5 µl of these stocks was added to row A in triplicate, and then 50 µl of row A was added to row B. This was repeated down the plate to row G, discarding the remaining 50 µl. For the controls cycloheximide and EtOH, 1 µl of the compound [10 mM] was added to 99 µl of the  $0.5 \times 10^6$  cells mL<sup>-1</sup> *L. mexicana*. The plate was then incubated for 20 hrs at 26 °C before the addition of 10 µl resazurin and a further incubation of 4 hrs. The fluorescence was measured using a fluorescence plate reader (555 – 585 nM, BioTek Synergy HT), and the results analysed using the programme GraphPad Prism.

### 6.5 Recombinant *E. coli* materials

The gene PhoN from *Salmonella typhumirium* was obtained from [UniProt \(A0A0F6BAI0\)](#) and was purchased from GenScript (20 µg) in a pUC57 plasmid. Primers were designed using the assistance of OPPF-UK and ordered from Merck. The primer sequences can be found in Table 8.

Table 8 The primer sequence used for pOPINF-PhoN, purchased from Merck

Primer name	Sequence
MB_FWD	aagttctgtttcagggccccgATGAAAAGCCGTTAC
MB_Rev	atggctagaaagctttaTTAATAGTTCAGCTTCG

2x phusion high fidelity enzyme (ThermoFisher) was used for PCR following manufacturers conditions. The Biometra Tone Gradient 96 Thermo Cycler (Analytik Jena) PCR machine was used following the settings in Table 9.

Table 9 Settings for PCR

Step	Temp/ °C	Time/ sec
Initial denaturation	98	30
Denaturation	98	10
Annealing	51	30
Extension	72	30
Final extension	72	300

\*Repeated for 32 cycles

DNA gel electrophoresis was conducted with 0.8 % agarose (Bio-Rad) gel supplemented with SYBR safe DNA gel stain (6X, Invitrogen, 1  $\mu$ L in 30 mL 0.8 % agarose), which was added when the agarose was 40-50 °C. To load the DNA, 1  $\mu$ L DNA Gel Loading Dye (6X, ThermoFisher Scientific) was added to 2.5  $\mu$ L H<sub>2</sub>O and 2.5  $\mu$ L DNA samples with GeneRuler 1 kb ladder (ThermoFisher Scientific) used as a reference. The gels were performed in 1 x TAE buffer at 100 V for 35 mins. DNA purification from the gel used QIAquick Gel Extraction kit (Qiagen), with DNA quantification occurring using DeNovix DS-11+ Spectrophotometer nanodrop function. The expression plasmid pOPINF (Addgene) was linearised using HindIII and Kpml purchased from ThermoFisher Scientific. The PhoN gene was ligated into pOPINF using In-Fusion cloning kit (Takara) then QIAprep Spin Miniprep Kit (Qiagen) for the purification of the plasmid using manufacturer's instructions.

The *E. coli* strains DH5 $\alpha$ , BL21 DE3 and BL21 (DE3) pLysS competent cells were purchased from ThermoFisher Scientific. After transformation, all *E. coli* strains (ThermoFisher Scientific) were grown in the supplied sterile SOC media before being grown on agar plates (Table 6) supplemented with 100  $\mu$ g/mL ampicillin and 50  $\mu$ g/mL chloramphenicol for BL21 (DE3) pLysS only. The transformed *E. coli* were grown in sterile LB broth (Melford) supplemented with 100  $\mu$ g/mL ampicillin (Merck) or with the addition of 50  $\mu$ g/mL chloramphenicol (Acros organics), when specified. Bacterial growth was determined by using 500  $\mu$ l bacterial culture in Fisherbrand curvettes (1.5 mL, ThermoFisher Scientific) and the OD was calculated using DeNovix (DS-11+ Spectrophotometer). Once between OD 0.6-0.8, filter sterilised IPTG (500 mM, Melford) was added to a final concentration of 0.5 mM. Lysis by sonication occurred in wash buffer (Table 5) supplemented with cComplete™ EDTA-free Protease Inhibitor Cocktail (Merck). Proteins were analysed using 12.5 % SDS-PAGE using the recipe in Table 10.

Table 10 Composition of the gels used for SDS-PAGE

	2 x 1 mm		4 x 1 mm	
	Resolution/ $\mu$ L	Stack/ $\mu$ L	Resolution/ $\mu$ L	Stack/ $\mu$ L
<b>Buffer</b>	4800	2400	9600	4800
<b>40 % BisAcrylamide</b>	3750	1425	7500	2850
<b>Milli-Q H<sub>2</sub>O</b>	3150	2974	6300	5948
<b>10 % APS</b>	124.0	62.6	248.0	125.2
<b>TEMED</b>	12.5	6.3	25.0	12.5

The resolving buffer (X4) and stacking buffer (X4) were purchased from Alfa Aesar. Acrylamide:Bis-acrylamide 29:1, 40 % solution was purchased from Fisher Bioreagents. Electrophoresis grade TEMED was purchased from MP Biomedicals and ammonium persulfate was purchased from Merck and made to a 10 % w/v solution with Milli-Q H<sub>2</sub>O. To the protein samples, NuPAGE 4X LDS buffer (ThermoFisher) supplemented with 5 %  $\beta$ -mercaptoethanol (Fluorochem) was added to the final concentration of 1X LDS. The gel was ran in X1 SDS running buffer (Table 5) at 200 V for 50 mins, using PageRuler Plus Prestained Protein Ladder (10-250 kDa, ThermoFisher Scientific) as a reference. The gel was stained using quick Coomassie stain (Generon) overnight before destaining with Milli-Q grade H<sub>2</sub>O. Where necessary, dialysis was achieved using Mini GeBaFlex tube (MWCO: 8 kDa, Generon) or Maxi GeBaFlex tube (MWCO: 8 kDa Generon).

## 6.6 Production of protein from *E. coli*

### 6.6.1 Transformation

The purchased PhoN-pUC57 plasmid was suspended in 20  $\mu$ l Milli-Q H<sub>2</sub>O, and 1  $\mu$ l of this solution was used to transform 50  $\mu$ l chemically competent *E. coli* DH5 $\alpha$  (Clontech), following the manufactures procedure for heat shock. These cells were then added to 500  $\mu$ l pre-warmed SOC media (37 °C), then grown in an orbital incubator for 1 h at 37 °C. The recovered cells were then pelleted by centrifugation for 5 mins, 3000 rpm. 400  $\mu$ l of the SOC media was disposed of and the bacterial pellet was then resuspended in the remaining 100  $\mu$ l LB. These cells were grown on-ampicillin (100  $\mu$ g/mL) impregnated agar plates overnight at 37 °C. A colony was picked for growth followed by plasmid purified using QIAprep Spin Miniprep Kit High-Yield Protocol (Qiagen). Using PCR and gel electrophoresis, GeneJET gel extraction kit (ThermoFisher Scientific) was used to purify the PhoN gene from the gel, ready for ligation into an expression plasmid.

### 6.6.2 Ligation into pOPINF

The expression plasmid pOPINF was linearized using HindIII and KpmI in Tamgo buffer. This was ligated with purified PhoN gene using In-Fusion<sup>®</sup> HD cloning kit (Clontech), following the manufacturers procedure. The ligated pOPINF plasmid was inserted into *E. coli* BL21 DE3 and grown on agar with 100  $\mu$ g/mL ampicillin and for *E. coli* BL21 (DE3) pLysS grown

on agar with 100 µg/mL ampicillin and 50 µg/mL chloramphenicol. Verification of the insertion of the desired gene occurred using colony PCR and electrophoresis.

#### 6.6.3 Sanger sequence

The isolation and purification of plasmids occurred using Qiaprep Spin Miniprep kit (Qiagen). For PCR amplicon purification, GeneJET gel extraction kit (ThermoFisher Scientific) from the excised agarose gel containing the amplicon was used. Sanger sequencing were performed through DBS Genomics (Department of Biosciences, Durham University). The Sanger sequences were analysed on the program 4peaks. The alignments were performed using the programme MUSCLE (EMBL-EBI, ClustalW) and sequence identities were calculated using the programme EMBOSS Needle (EMBL-EBI).

#### 6.6.4 Overexpression and lysis

To overexpress the desired protein in *E. coli* BL21 DE3, the bacteria was grown in LB agar supplemented with 100 µg/mL ampicillin at 37 °C, 150 rpm. To overexpress the desired protein in *E. coli* BL21 (DE3) pLysS, the bacteria was grown in LB agar supplemented with 100 µg/mL ampicillin and 50 µg/mL chloramphenicol at 30 °C, 150 rpm. When the cultures reached 0.6 - 0.8 OD<sub>600</sub>, protein production was induced using a final concentration of 0.5 mM IPTG and incubated at 30 °C overnight at 150 rpm. The cultures were then pelleted at 4000 rpm for 15 mins. The media was decanted from the pellet, and this was resuspended in wash buffer (Table 5).

For lysis, the suspended *E. coli* was sonicated on ice at 30 % power, 1 sec interval for 30 secs (X3) with 1 min rest on ice. This solution was then centrifuged at 13000 rpm at 4 °C for 30 mins to pellet insoluble material.

The proteins were analysed on 12.5 % SDS-PAGE gel by adding LDS to a final concentration of X1 and then analysed by SDS-PAGE at 200 V for 50 mins. The proteins were then stained with quick Coomassie (Generon) and imaged.

#### 6.6.5 Protein purification

For purification, the HisTrap™ Fast Flow 1 column (GE healthcare) and minipuls 3 (Gibson) peristaltic pump were pre-equilibrated with the wash buffer (Table 5) before the addition of the lysate. Following this, the bacterial lysate was loaded onto the HisTrap™ column with the peristaltic pump the following manufacturer's instructions. The column was then

loaded onto the ÄKTA (Cytiva) and purification occurred using a gradient of wash buffer and elution buffer (0-100 %, Table 5). The eluted fractions were analysed by SDS-PAGE, and the desired fractions were combined and further purified using MonoQ anion exchange on the ÄKTA (Cytiva) with 20 mM Tris buffer (pH 8.5) and eluted with salt elution buffer (Table 5) on a gradient (0-100 %). The protein was then dialysed into HEPES (20 mM, pH 7.4) overnight at 4 °C before protein quantification with DeNovix DS-11+ Spectrophotometer. The protein was then supplemented with 20 % glycerol and flash frozen in Cryovials (Starlabs) before storage at -80 °C for long term or stored in a centrifuge tube (1.5 mL, Starlabs) at -20 °C for use.

### 6.7 Non-specific acid phosphatase activity *materials*

PhoN was diluted to 2 µg/mL in HEPES (20 mM, pH 7.4). Para-nitrophenol phosphate (Fluorochem) was prepared in the enzyme buffer (500 mM) before every assay. NaOH (ThermoFisher Scientific) 0.2 M solution in H<sub>2</sub>O. BHDC (X) probe was prepared in DMSO before every assay. Fluorescence was measured using Cary 5000 UV-Vis-NIR at 405 nM and the results analysed on the program Cary WinUV (4.20).

#### 6.7.1 NSAP assay method

To measure the activity of the purified enzyme PhoN, 5 mM para-nitrophenyl phosphate (pNPP) was incubated with 2 µg/mL enzyme in 20 mM HEPES. The reaction was quenched at the specified times by adding 500 µL of 0.2 M NaOH. The resulting fluorescence from the production of para-nitrophenol (pNP) was measured at 405 nM.

To test for enzyme inhibition, BHDC (**66**) was incubated with 2 µg/mL PhoN on a range of 100 µM – 3 mM for 1 h. After this, the substrate pNPP (5 mM) was added and the enzyme was further incubated for an additional 15 min before the reaction was quenched (500 µL of 0.2 M NaOH) and the reading taken at 405 nM.

### 6.8 Fluorescent and pull-down analysis

#### 6.8.1 Materials

Ribonuclease A (bovine pancreas, Merck) stock solution was dissolved in Milli-Q H<sub>2</sub>O at 10 mg/mL. For the working concentration, RNase A was diluted to 1 mg/mL in either Milli-Q H<sub>2</sub>O or PBS buffer. PhoN was diluted to 0.25 mg/mL in 20 mM HEPES buffer (pH 7.4). The lysate from *E. coli* BL21 (DE3) pLysS overexpressing PhoN generated following section 6.6.4.

Following lysis (section 6.6.4), the samples were dialysed using Mini GeBaFlex tube (MWCO: 8 kDa, Generon) or Maxi GeBaFlex tube (MWCO: 8 kDa Generon). in 20 mM HEPES (pH 7.4) overnight at 4 °C. The lysate protein concentration was quantified using Pierce Rapid Gold BCA Protein Assay Kit (Merck) following manufacturers procedure and diluted to 1 mg/mL with 20 mM HEPES (pH 7.4). The rhodamine azide (**93**) was synthesised following Chemical Experimental (Chapter 7). The click chemistry reagents were all prepared in Milli-Q H<sub>2</sub>O and consist of sodium ascorbate (Fluorochem), CuSO<sub>4</sub>.5H<sub>2</sub>O (Merck), THPTA (Merck) and TBTA (Merck). For the pulldown, biotinylated PEG azide (**94**, Merck) was dissolved in DMSO. Streptavidin MagneSphere Paramagnetic Particles (Promega) and DynaMag -2 (Invitrogen) magnet were used for protein purification and 0.5 M β-mercaptoethanol (Fluorochem) for elution.

### 6.8.2 Fluorescent analysis methods

The probe BHDC (**66**) was added to a final concentration of 50 μM for lysates and PhoN, and 25 μM for RNase A, unless otherwise specified. This was incubated at room temperature for 1 h before the addition of the click chemistry reagents. Firstly, the rhodamine azide (**94**) was added and following this was a premixed solution of CuSO<sub>4</sub>.5H<sub>2</sub>O and THPTA/TBTA and finally was the addition of sodium ascorbate in the quantities displayed in Table 11. After briefly vortexing the samples, this mixture was incubated for 1 h at room temperature before quenching with 16.6 μL X4 LDS (5 % β-mercaptoethanol). The samples were loaded to a 12.5 % SDS-PAGE and ran until the solvent front has been removed from the gel. The results were visualised using Fuji FLA3000 with the program AIDA image analyser or Typhoon 9400 (Amersham Biosciences) with the programme ImageQuant (version 5.2) when specified.

Table 11 Click chemistry quantities and components for fluorescent analysis in different lysates <sup>209</sup>

	<b>RNase A</b>	<b>PhoN</b>	<b><i>E. coli</i> lysate</b>	<b><i>Leishmania</i> lysate</b>
BHDC ( <b>66</b> )	25 μM	50 μM	50 μM	50 μM
CuSO <sub>4</sub> .5H <sub>2</sub> O	50 μM	100 μM	100 μM	500 μM
THPTA	250 μM	500 μM	500 μM	-
TBTA	-	-	-	500 μM
Na asc	2.5 mM	5 mM	5 mM	1 mM

Rho-N <sub>3</sub> ( <b>94</b> )	50 $\mu$ M	100 $\mu$ M	100 $\mu$ M	100 $\mu$ M
----------------------------------	------------	-------------	-------------	-------------

### 6.8.3 Pulldown assay

The target protein was incubated with 2.5  $\mu$ M BHDC (**66**) for 1 h before the addition of the click reagents. The first reagent added was the tetrafunctional tag (**97**) or the commercially available biotinylated PEG azide (**94**, Merck) which were both incubated to a final concentration of 5  $\mu$ M. Following this was the addition of a premixed solution of CuSO<sub>4</sub>.5H<sub>2</sub>O and THPTA/TBTA before the final addition of sodium ascorbate to the concentrations documented in Table 12. Following a brief vortex, the protein mixtures were incubated for 1 h at RT. The magnetic beads were prepared following manufacturers instructions, washing the beads with PBS buffer, as advised. The BHDC (**66**)-protein mixtures were added to the beads, and this was incubated at room temperature for 1 h. The unbound protein lysate was removed by inserting the magnetic bead tube into the magnetic rack, waiting for a pellet to form on the side of the tube and carefully removing the supernatant. The beads were then washed X3 with cold PBS (500  $\mu$ L) to remove unbound proteins. For the tetrafunctional tag (**97**), elution occurred by adding 50  $\mu$ L of 0.5 M  $\beta$ -mercaptoethanol and incubating at room temperature for 30 min. Elution of the commercially available biotinylated PEG azide (**94**) occurred by adding 50  $\mu$ L X4 LDS (5 %  $\beta$ -mercaptoethanol) and boiling the samples at 95  $^{\circ}$ C for 10 min. The eluted proteins were analysed by SDS-PAGE and Coomassie protein staining.

Table 12 Click chemistry quantities and components for pull-down assays in different lysates

	<b>RNase A</b>	<b>PhoN</b>	<b><i>E. coli</i> lysate</b>	<b><i>Leishmania</i> lysate</b>
BHDC ( <b>66</b> )	2.5 $\mu$ M	2.5 $\mu$ M	2.5 $\mu$ M	2.5 $\mu$ M
CuSO <sub>4</sub> .5H <sub>2</sub> O	10 $\mu$ M	10 $\mu$ M	10 $\mu$ M	10 $\mu$ M
THPTA	50 $\mu$ M	50 $\mu$ M	50 $\mu$ M	-
TBTA	-	-	-	10 $\mu$ M
Na asc	500 $\mu$ M	500 $\mu$ M	500 $\mu$ M	100 $\mu$ M
Azide ( <b>97</b> or <b>94</b> )	5 $\mu$ M	5 $\mu$ M	5 $\mu$ M	5 $\mu$ M

## 6.9 Proteolytic digestion

The samples submitted to ESI<sup>+</sup> mass spectrometry for the restriction digestion of pure enzymes samples were to the concentrations of 1 mg/mL of protein. After the incubation of the sample with BHDC (**66**), the samples were dialysed into freshly prepared ammonium bicarbonate (50 mM, pH 8.5). These samples were incubated with TCEP (5 mM) for 5 min at RT to reduce disulphide bridges. Following, the protein samples were incubated with iodoacetamide (15 mM) for the modification of cysteine residues, and then the samples were incubated with trypsin (1.25 µL of 1 mg/mL) and separately chymotrypsin (2.5 µL of 1 mg/mL). Further analysis was performed by Dr Adrian Brown, Department of Biosciences at Durham University. The mass spectra obtained from the fragments were queried against *E. coli* sequences from UniProt KB, with a false discovery rate set at 1 % for peptide fragment identification.

## 7 Chemical Experimental

### 7.1 Equipment

IUPAC names were generated using the programme Marvin (ChemAxon). All solvents and reagents were purchased from commercial suppliers. Sensitive reactions were performed with dry solvents in a N<sub>2</sub> enriched environment. NMR spectra were recorded on two instruments: Varian VNMRS-600 instrument with operating frequencies of 600.130 MHz for <sup>1</sup>H and 150.903 MHz for <sup>13</sup>C NMR and Varian VNMRS-700 with operating frequencies of 700.130 MHz for <sup>1</sup>H and 176.048 MHz for <sup>13</sup>C NMR. Spectra were referenced relative to CDCl<sub>3</sub> (dH 7.26 ppm, dC 77.16 ppm) or CD<sub>3</sub>OD (dH 4.87 ppm, dC 49.00 ppm). Chemical shifts are reported in parts per million (ppm), coupling constants (*J*) in hertz (Hz) and multiplicity as singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m) or a combination thereof. All spectral assignments were made with the aid of <sup>1</sup>H-<sup>1</sup>H COSY, <sup>1</sup>H-<sup>13</sup>C HSQC and <sup>1</sup>H-<sup>13</sup>C HMBC NMR experiments. Infra-red spectra were recorded on a Perkin Elmer Paragon 1000 FT-IR spectrometer or a Perkin Elmer RX FT-IR spectrometer with Golden Gate Diamond ATR apparatus and IR assignments are reported in wavenumbers (cm<sup>-1</sup>). Melting points were recorded on Thermo Scientific Electrothermal IA9100 Digital Melting Point apparatus. Thin layer chromatography was performed using Merck F254 silica gel 60 aluminium sheets pre-coated with silica gel. Purification of products used flash chromatography systems CombiFlash Rf 200 (Teledyne ISCO) or CombiFlash NextGen 100 (Teledyne ISCO). High resolution mass spectrometry (HRMS) were recorded on a Waters TQD mass spectrometer ESI-LC water (0.1 % formic acid): MeOH, flow rate 0.6 mL min<sup>-1</sup> with a UPLC BEH C18 1.7 μm (2.1 mm x 50 mm) column.

### 7.2 General procedure A: Formation of Dicarbonates

A 20 % aqueous NaOH (5 M) solution was added to a solution of tetrabutylammonium bromide (1.0 eq) in DCM (1 mL.mmol<sup>-1</sup>). This was cooled to -10 °C, and the chloroformate (10.0 eq) then added and the reaction mixture then stirred at -10 °C for 1 h before warming to and stirring at RT for an additional 3 h. The reaction mixture was then washed with saturated NaHCO<sub>3</sub> (2 x 10.0 ml), H<sub>2</sub>O (1 x 10.0 ml), dried (MgSO<sub>4</sub>) and concentrated to afford the dicarbonate which was used directly with no further purification.

### 7.3 General procedure B: Carbamate Preparation

Pyrocarbonate (1.1 eq) was added to a solution of the benzimidazole (1.0 eq) in MeCN (0.5 M) or His-Gly-OH (1.0 eq) (1:1 DCM and DMF). The resulting solution was stirred at RT for the individually specified time. The organic solvent was removed *in vacuo* and the carbamates were then redissolved in EtOAc (10 ml), which was then washed with NH<sub>4</sub>Cl (1 x 10.0 ml) and H<sub>2</sub>O (2 x 10.0 ml), dried (MgSO<sub>4</sub>) and concentrated *in vacuo*. Purification was achieved using column chromatography.

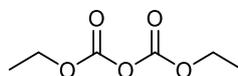
### 7.4 General procedure C: Formation of 1,2,3-triazole

The alkyne (1.0 eq) and azide (1.1 eq) compounds were dissolved in H<sub>2</sub>O and t-BuOH (1:1). To this, CuSO<sub>4</sub>·5H<sub>2</sub>O (0.1 eq) was added followed by sodium ascorbate (0.4 eq) and the resultant solution stirred for 2 h at RT. On completion of the reaction, verified using mass spectrometry, the mixture was extracted with EtOAc (10.0 ml x 4), dried (MgSO<sub>4</sub>) and concentrated *in vacuo*. The product was purified using column chromatography (SiO<sub>2</sub>) using the individually specified solvent system.

### 7.5 General procedure D: Chalcone Preparation

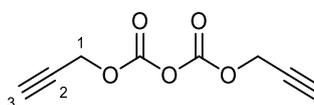
2', 4', 6'-trimethylacetophenone (1.0 eq), Ba(OH)<sub>2</sub>·8H<sub>2</sub>O (0.25 eq) and the aldehyde (1.0 eq) are dissolved in MeOH (0.17 M) and heated to 40 °C overnight (14 h). The reaction was filtered, the solvent removed *in vacuo*, and then the mixture was redissolved in EtOAc (10 mL) and washed with 1 M HCl (1 x 15 mL) and H<sub>2</sub>O (2 x 15 mL), dried (MgSO<sub>4</sub>) and concentrated. Purification occurred by recrystallisation using Et<sub>2</sub>O and hexane.

#### 1. Diethyl pyrocarbonate (DEPC) (50) <sup>187</sup>



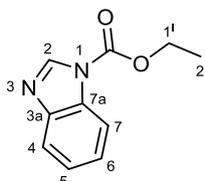
Using general procedure A, ethyl chloroformate (0.05 ml, 0.5 mmol) was used to form DEPC as a clear oil (40.5 mg, 50 %).  $\delta_{\text{H}}$  (700 MHz, CDCl<sub>3</sub>) 4.33 (4H, q,  $J=7.1$ , CH<sub>2</sub>), 1.37 (6H, t,  $J=7.1$ , CH<sub>3</sub>).

#### 2. Dipropargyl pyrocarbonate (52) <sup>188</sup>



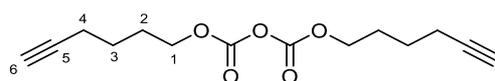
Following general procedure A, use of propargyl chloroformate (2.5 mL, 26 mmol) afforded the title dicarbonate as a clear oil (2.5 g, 54 %).  $\nu_{\max}$  (ATR) 3292, 2955, 2130 (C≡C), 1750 (C=O), 1438, 1386  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (700 MHz,  $\text{CDCl}_3$ ) 4.76 (4H, d,  $J = 2.5$  Hz,  $\text{CH}_2$ ), 2.54 (2H, t,  $J = 2.5$  Hz, 3- $H$ );  $^{13}\text{C}$  NMR (176 MHz,  $\text{CDCl}_3$ )  $\delta$  154.4 (C=O), 77.0 (C-2), 76.3 (C-3), 56.1 ( $\text{CH}_2$ ).

### 3. Ethyl 1H-1,3-benzodiazole-1<sup>l</sup>-carboxylate (60)



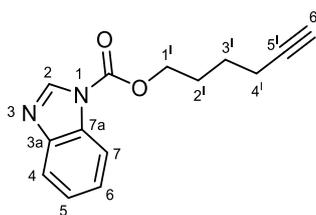
Following general procedure A, DEPC (**50**) (0.1 mL; 0.25 mmol; 1.0 eq) was added to a solution of benzimidazole (59.0 mg; 0.25 mmol; 1.0 eq). Purification by column chromatography ( $\text{SiO}_2$ , 0-100 % hexane: EtOAc) to afford the title carbamate as a white solid (51.0 mg; 54 %).  $\nu_{\max}$  (ATR) 3151- 3078 (C-H), 2983 (C-H), 1750 (C=O), 1610, 1505, 1477, 1448, 1399, 1378  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (700 MHz,  $\text{CDCl}_3$ ) 8.48 (1H, s, 2- $H$ ), 8.03 (1H, d,  $J = 7.8$  Hz, 4- $H$ ), 7.80 (1H, d,  $J = 7.8$  Hz, 7- $H$ ), 7.41 (1H, t,  $J = 7.8$  Hz, 5- $H$ ), 7.39 – 7.35 (1H, m, 6- $H$ ), 4.56 (2H, q,  $J = 7.1$  Hz,  $\text{CH}_2$ ), 1.51 (3H, t,  $J = 7.1$  Hz,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (176 MHz,  $\text{CDCl}_3$ )  $\delta$  149.61 (N-COO), 144.1 (C-Ar), 141.8 (C-2), 131.4 (C-Ar), 125.6 (C-3), 124.7 (C-6), 120.8 (C-Ar), 114.5 (C-Ar), 64.4 ( $\text{CH}_2$ ), 14.4 ( $\text{CH}_3$ );  $m/z$  (LC-MS ES<sup>+</sup>) 191.1894 [M+H]<sup>+</sup>  $\text{C}_{10}\text{H}_{10}\text{N}_2\text{O}_2$ .

### 4. Bis(hex-5-yn-1-yl) dicarbonate (66)<sup>187</sup>



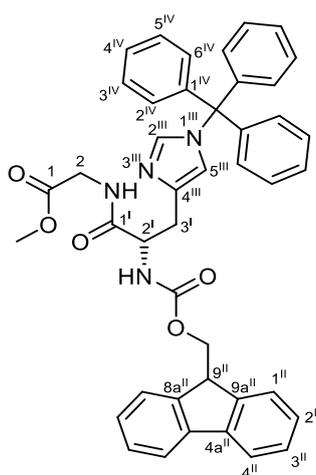
5-Hexyn-1-ol (0.3 mL; 2.5 mmol, 1 eq) was added to a solution of triphosgene (371.0 mg; 1.3 mmol; 0.5 eq) in anhydrous DCM (7.5 ml). The reaction mixture was then stirred at RT for 3 h to generate the chloroformate intermediate. The reaction mixture was concentrated *in vacuo* and the crude chloroformate used immediately for the synthesis of the title dicarbonate (**38**) using general procedure A. Following the final concentration, the product was observed as a light yellow oil (617.0 mg; 92 %).  $\nu_{\max}$  (ATR) 3299, 3004-2888, 2124 (C≡C), 1789 (C=O), 1463, 1387, 1232,  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (700 MHz,  $\text{CDCl}_3$ ) 4.34 (4H, dt,  $J = 8.0$ , 6.4 Hz, 1- $H$ ), 2.26 – 2.23 (4H, m, 4- $H$ ), 1.97 (2H, t,  $J = 2.6$  Hz, 6- $H$ ), 1.91 – 1.83 (4H, m, 2- $H$ ), 1.66 – 1.59 (4H, m, 3- $H$ );  $^{13}\text{C}$  NMR (176 MHz,  $\text{CDCl}_3$ )  $\delta$  148.0 (C=O), 83.5 (C-5), 69.8 (C-1), 69.2 (C-6), 27.4 (C-2), 24.6 (C-3), 18.1 (C-4).

### 5. Hex-5<sup>l</sup>-yn-1<sup>l</sup>-yl 1H-1,3-benzodiazole-1-carboxylate (67)



Following general procedure B, a solution of bis(hex-5-yn-1-yl) dicarbonate (**66**, 266.0 mg; 1 mmol; 1.0 eq) in CH<sub>3</sub>CN (1.0 mL) was added to a solution of benzimidazole (112.0 mg; 1.0 mmol; 1 eq) in CH<sub>3</sub>CN (2.0 mL) and stirred at RT for 2 h. The residue was purified using column chromatography (SiO<sub>2</sub>, hexane: EtOAc, 0-100 %) to afford the final product as a white solid (35 mg; 14 %).  $\nu_{\max}$  (ATR) 2953, 2141 (C≡C), 1793 (C=O), 1749, 1681, 1531, 1448, 1406, 1326 (C-N), 1244, 1212 (C≡C) cm<sup>-1</sup>;  $\delta_{\text{H}}$  (599 MHz, CDCl<sub>3</sub>)  $\delta$  8.55 (1H, s, 2-*H*), 8.05 (1H, d, *J* = 8.0 Hz, Ar-*H*), 7.86 - 7.81 (1H, d, *J* = 8.0, Ar-*H*), 7.47 - 7.37 (2H, m, 5-*H*, 6-*H*), 4.56 (2H, t, *J* = 6.5 Hz, 1<sup>l</sup>-*H*), 2.32 (2H, td, *J* = 6.5, 2.7 Hz, 4<sup>l</sup>-*H*), 2.04 - 2.01 (2H, m, 2<sup>l</sup>-*H*), 2.00 (1H, t, *J* = 2.7 Hz, 6<sup>l</sup>-*H*), 1.76 - 1.71 (2H, m, 3<sup>l</sup>-*H*); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)  $\delta$  149.5 (N-COO), 143.2 (C-Ar), 141.6 (C-2), 131.2 (C-Ar), 126.0 (C-Ar), 125.0 (C-Ar), 120.6 (C-3), 114.6 (C-6), 83.5 (C-13), 69.4 (C-14), 68.1 (C-9), 27.7 (C-10), 24.9 (C-11), 18.2 (C-12); HRMS (ES<sup>+</sup>) 243.1128 [M+H]<sup>+</sup> C<sub>14</sub>H<sub>14</sub>N<sub>2</sub>O<sub>2</sub>.

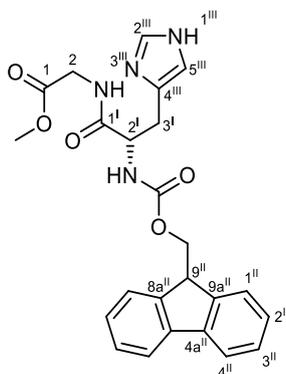
### 6. Methyl 2-[2<sup>l</sup>-{[(9H-fluoren-9<sup>ll</sup>-yl)methoxy]carbonyl}amino]-3<sup>l</sup>-[1-(triphenylmethyl)-1H-imidazol-4<sup>lll</sup>-yl]propanamido]acetate (70)



Glycine methyl ester.HCl (111.0 mg; 0.9 mmol; 1.1 eq), N-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide HCl (EDC) (193.0 mg; 1.0 mmol; 1.3 eq) and 4-dimethylaminopyridine (DMAP) (148.0 mg; 1.2 mmol; 1.5 eq) were dissolved in anhydrous DCM (10.0 mL) under

N<sub>2</sub>. The resulting mixture was stirred for 10 mins at RT when Fmoc His(trt)-OH (500.0 mg; 0.8 mmol; 1.0 eq) was added. The reaction was then stirred for 14 h at RT when it was washed with NaHCO<sub>3</sub> (3 x 15.0 ml), dried (MgSO<sub>4</sub>) and concentrated *in vacuo*. Column chromatography (SiO<sub>2</sub>, 20-100 % hexane: EtOAc) with the desired product eluting at 100 % EtOAc afforded the title dipeptide as a white solid (0.5 g; 96 %).  $\nu_{\max}$  (ATR) 1753, 1720, 1670, 1577-1466, 1445, 1276-1223, 1205, 1054-1015 cm<sup>-1</sup>;  $\delta_{\text{H}}$  (700 MHz, CDCl<sub>3</sub>, rotamer) 7.74 (2H, d,  $J = 7.6$  Hz, 4<sup>II</sup>-H), 7.61 (2H, d,  $J = 7.6$ , 1<sup>II</sup>-H), 7.42 (1H, s, 2<sup>III</sup>-H), 7.37 (2H, t,  $J = 7.6$  Hz, Ar-H), 7.33 – 7.22 (10H, m, Ar-H), 7.12 – 7.10 (6H, m, Ar-H), 6.82 (1H, s, NH), 6.69 (1H, s, 5<sup>III</sup>-H), 4.57 (1H, m, 2<sup>I</sup>-H), 4.34 (2H, d,  $J = 7.5$  Hz, 9<sup>II</sup>-CH<sub>2</sub>), 4.21 (1H, t,  $J = 7.5$  Hz, 9<sup>II</sup>-H), 3.99 (2H, d,  $J = 5.3$  Hz, 2-H), 3.68 (3H, s, CH<sub>3</sub>), 3.13 – 3.10 (1H, m, 3<sup>I</sup>-H), 3.06 – 3.00 (1H, m, 3<sup>I</sup>-H); <sup>13</sup>C NMR (176 MHz, CDCl<sub>3</sub>)  $\delta$  171.8 (C-1<sup>I</sup>), 170.0 (C-1), 159.4 (COON-Fmoc) 144.0 (C-9a<sup>II</sup>), 142.3 (C-Bn<sub>3</sub>), 141.3 (C-4a<sup>II</sup>), 138.6 (C-2<sup>III</sup>), 136.8 (C-4<sup>III</sup>), 129.8 (C-Ar), 128.1 (C-Ar), 128.1 (C-Ar), 127.7 (C-Ar), 127.1 (C-Ar), 127.1 (C-Ar), 125.3 (C-1<sup>III</sup>), 120.0 (C-4<sup>II</sup>), 119.7 (C-5<sup>III</sup>), 67.2 (9<sup>II</sup>-CH<sub>2</sub>), 55.2 (C-2<sup>I</sup>), 52.3 (CH<sub>3</sub>), 47.2 (C-9<sup>II</sup>), 41.3 (C-2), 30.7 (C-3<sup>I</sup>); HRMS (ES<sup>+</sup>) 691.2937 [M+H]<sup>+</sup> C<sub>43</sub>H<sub>38</sub>N<sub>4</sub>O<sub>5</sub>.

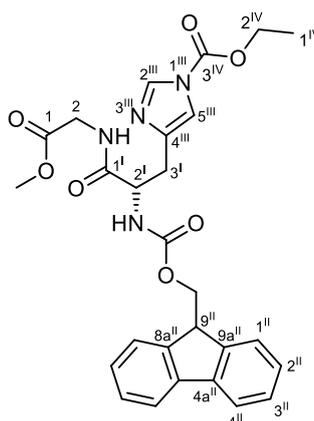
**7. Methyl 2-[2<sup>I</sup>-{[(9H-fluoren-9<sup>II</sup>-yl)methoxy]carbonyl}amino]-3<sup>I</sup>-[1H-imidazol-4<sup>III</sup>-yl]propanamido]acetate (71)**



TFA (2.5 ml) was added dropwise to a solution of methyl 2-[2<sup>I</sup>-{[(9H-fluoren-9<sup>II</sup>-yl)methoxy]carbonyl}amino]-3<sup>I</sup>-[1-(triphenylmethyl)-1H-imidazol-4<sup>III</sup>-yl]propanamido]acetate (**70**, 344.0 mg; 0.5 mmol) in DCM (3.0 mL) and stirred at RT under N<sub>2</sub> for 4 h, with a colour change occurring over the duration of the reaction from yellow to brown. The reaction mixture was then concentrated, the residue dissolved in EtOAc (5.0 ml) and washed with NaHCO<sub>3</sub> (2 x 5.0 ml) and H<sub>2</sub>O (2 x 5.0 ml). The organic layer was dried (MgSO<sub>4</sub>), concentrated. The product was the triturated in cold Et<sub>2</sub>O (10.0 mL), affording the

product as a white solid (89.0 mg; 40 %).  $\nu_{\max}$  (ATR) 3355-3265 (NH), 2953, 2516-2369, 1738 (C=O), 1678, 1645, 1584-1534, 1441 (C-H), 1365, 1348 (C-N)  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$  (700 MHz,  $\text{CD}_3\text{OD}$ , rotamer) 8.04 (1H, s, 2<sup>III</sup>-H), 7.78 (2H, d,  $J = 7.6$  Hz, 4<sup>II</sup>-H), 7.59 (2H, d,  $J = 7.6$  Hz, 1<sup>II</sup>-H), 7.37 (2H, t,  $J = 7.6$ , 3<sup>II</sup>-H), 7.32 – 7.26 (2H, m, 2<sup>II</sup>-H), 7.04 (1H, s, 5<sup>III</sup>-H), 4.46 (1H, t,  $J = 7.5$  Hz, 2<sup>I</sup>-H), 4.38 (1H, dd,  $J = 10.7, 7.0$  Hz, 9<sup>II</sup>-CH<sub>2</sub>), 4.30 (1H, dd,  $J = 10.7, 7.0$  Hz, 9<sup>II</sup>-CH<sub>2</sub>), 4.18 – 4.17 (1H, m, 9<sup>II</sup>-H), 4.00 – 3.87 (2H, m, 2-H), 3.70 (3H, s, CH<sub>3</sub>), 3.16 (1H, dd,  $J = 15.0, 5.4$  Hz, 3<sup>I</sup>-H), 2.95 (1H, dd,  $J = 15.0, 5.4$  Hz, 3<sup>I</sup>-H); <sup>13</sup>C NMR (176 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  172.4 (C-1<sup>I</sup>), 170.2 (C-1), 156.8 (N-COO), 143.8 (C-9a<sup>II</sup>), 141.1 (C-4a<sup>II</sup>), 134.4 (C-2<sup>III</sup>), 127.4 (C-3<sup>II</sup>), 126.7 (C-2<sup>II</sup>), 124.7 (C-1<sup>II</sup>), 119.5 (C-4<sup>II</sup>), 117.1 (C-5<sup>III</sup>), 69.8 (C-2<sup>I</sup>), 66.6 (9<sup>II</sup>-CH<sub>2</sub>), 54.6 (C-2), 51.2 (CH<sub>3</sub>), 48.0 (C-9<sup>II</sup>), 40.5 (C-2), 28.5 (C-3<sup>I</sup>); HRMS (ES<sup>+</sup>) 449.1805 [M+H]<sup>+</sup> C<sub>24</sub>H<sub>24</sub>N<sub>4</sub>O<sub>5</sub>; MT 154.0-155.4 °C.

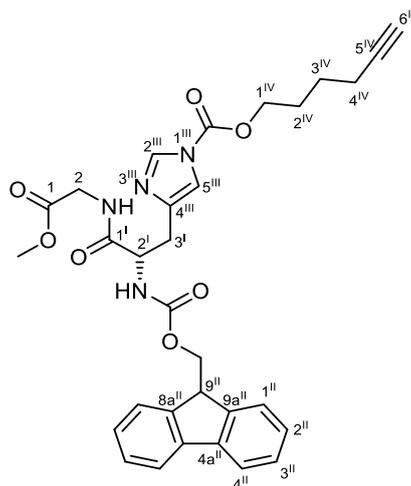
### 8. Methyl 2-[2<sup>I</sup>-({[(9H-fluoren-9<sup>II</sup>-yl)methoxy]carbonyl}amino)-3<sup>I</sup>-[(1-ethoxycarboxy)-1H-imidazol-4<sup>III</sup>-yl]]propanamido]acetate (**73**)



Following general procedure B, DEPC (**50**, 33.0  $\mu\text{L}$ ; 0.2 mmol; 2.2 eq) was added to a solution of methyl 2-[2<sup>I</sup>-({[(9H-fluoren-9<sup>II</sup>-yl)methoxy]carbonyl}amino)-3<sup>I</sup>-(1H-imidazol-4<sup>III</sup>-yl)propanamido]acetate (**71**, 46.0 mg; 0.1 mmol; 1.0 eq) in dry DCM (5.0 mL) and dry DMF (1.0 mL) under N<sub>2</sub> and stirred overnight. The reaction mixture was washed with H<sub>2</sub>O (4 x 10.0 ml), then the residue was purified by column chromatography (SiO<sub>2</sub>, 0-100 % hexane: acetone) to produce the carbamate as a white solid (22.0 mg, 42 %).  $\delta_{\text{H}}$  NMR (700 MHz,  $\text{CDCl}_3$ , rotamer).  $\nu_{\max}$  (ATR) 3422-3187 (N-H), 2982, 1753 (C=O), 1712, 1538, 1448 (CH<sub>3</sub>), 1407, 1249, 1215  $\text{cm}^{-1}$ ; <sup>1</sup>H NMR (599 MHz,  $\text{CDCl}_3$ )  $\delta$  8.20 (1H, s, 2<sup>III</sup>-H), 7.76 (2H, d,  $J = 7.6$  Hz, 4<sup>II</sup>-H), 7.59 (2H, d,  $J = 7.6$  Hz, 1<sup>II</sup>-H), 7.39 (2H, t,  $J = 7.4$  Hz, 3<sup>II</sup>-H), 7.31 (2H, t,  $J = 7.4$  Hz, 2<sup>II</sup>-H), 6.49 (1H, s, 5<sup>III</sup>-H), 4.61 (1H, m, 2<sup>I</sup>-H), 4.45 (2H, q,  $J = 7.1$  Hz, 2<sup>IV</sup>-H), 4.37 (2H, d,  $J = 7.3$  Hz, 9<sup>II</sup>-CH<sub>2</sub>), 4.22 (1H, t,  $J = 7.3$  Hz, 9<sup>II</sup>-H), 4.01 (2H, s, 2-H), 3.71 (3H, s, O-CH<sub>3</sub>), 3.16 – 3.08

(2H, m, 3<sup>I</sup>-H), 1.41 (3H, t,  $J = 7.1$  Hz, 1<sup>IV</sup>-H); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)  $\delta$  171.3 (C-1<sup>I</sup>), 170.0 (C-1), 156.5 (COON), 148.2 (C-3<sup>IV</sup>), 143.9 (C-9a<sup>II</sup>), 141.4 (C-4a<sup>II</sup>), 136.7 (C-2<sup>III</sup>), 127.9 (C-3<sup>II</sup>), 127.2 (C-2<sup>II</sup>), 125.3 (C-1<sup>II</sup>), 120.1 (C-4<sup>II</sup>), 67.4 (9<sup>II</sup>-CH<sub>2</sub>), 65.0 (C-2<sup>IV</sup>), 54.6 (C-2<sup>I</sup>), 52.4 (O-CH<sub>3</sub>), 47.3 (C-9<sup>II</sup>), 41.4 (C-2), 30.2 (C-3<sup>I</sup>), 14.3 (C-1<sup>IV</sup>); HRMS (ES<sup>+</sup>) 521.2038 [M+H]<sup>+</sup> C<sub>27</sub>H<sub>28</sub>N<sub>4</sub>O<sub>7</sub>.

**9. Methyl 2-[2<sup>I</sup>-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-3<sup>I</sup>-{(1-hex-5<sup>IV</sup>-ynoxycarboxy)-1H-imidazol-4<sup>III</sup>-yl)]propanamido]acetate (72)**

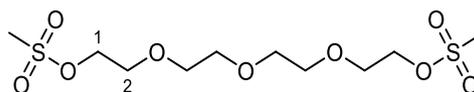


**Procedure 1:** Following general procedure B, bis(hex-5-yn-1-yl) dicarbonate (**66**, 24.3 mg; 0.1 mmol; 1.0 eq) dissolved in 3.2 mL DCM was added to methyl 2-[2<sup>I</sup>-({[(9H-fluoren-9<sup>II</sup>-yl)methoxy]carbonyl}amino)-3<sup>I</sup>-(1H-imidazol-4<sup>III</sup>-yl)]propanamido]acetate (**71**, 22.7 mg; 0.1 mmol; 1.0 eq) dissolved in 3.2 ml DMF and stirred at RT for 6 h. To purify the product, column chromatography (SiO<sub>2</sub>, 0-100 % hexane: acetone) was performed, with the pure product eluting at 40 % acetone. The named product forms a white solid when dry (24.0 mg; 32 %).

**Procedure 2:** Bis(hex-5-yn-1-yl) dicarbonate (**66**, 29.0 mg; 0.1 mmol; 1.1 eq) was added to methyl 2-[2<sup>I</sup>-({[(9H-fluoren-9<sup>II</sup>-yl)methoxy]carbonyl}amino)-3<sup>I</sup>-(1H-imidazol-4<sup>III</sup>-yl)]propanamido]acetate (**71**, 45.0 mg; 0.1 mmol; 1.0 eq) in 8.0 ml tris-bis buffer (0.1 M) and stirred overnight.  $\nu_{\max}$  (ATR) 3194- 3509 (N-H), 2980-2885, 2215 (C $\equiv$ C), 1793, 1747, 1716, 1680, 1526, 1242 cm<sup>-1</sup>; <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>, rotamer)  $\delta$  8.17 (1H, s, 2<sup>III</sup>-H), 7.75 (2H, d,  $J = 7.5$  Hz, 4<sup>II</sup>-H), 7.59 (2H, d,  $J = 7.5$ , 1<sup>I</sup>-H), 7.39 (2H, t,  $J = 7.5$  Hz, 3<sup>I</sup>-H), 7.30 (2H, t,  $J = 7.5$ , 2<sup>II</sup>-H) 6.52 (1H, s, 5<sup>III</sup>-H), 4.61 (1H, m, 2<sup>I</sup>-H), 4.47 – 4.30 (4H, m, 1<sup>IV</sup>-H, 9<sup>II</sup>-CH<sub>2</sub>), 4.22 (1H, t,  $J = 7.1$  Hz, 9<sup>II</sup>-H), 4.00 (2H, s, 2-H), 3.70 (3H, s, CH<sub>3</sub>), 3.10 (2H, d,  $J = 14.4$ , 3<sup>I</sup>-H), 2.30 – 2.22 (2H, m, 4<sup>IV</sup>-H), 1.98 (1H, t,  $J = 2.1$  Hz, 6<sup>IV</sup>-H), 1.90 (2H, p,  $J = 6.7$  Hz, 2<sup>IV</sup>-H), 1.63 (2H, p,  $J =$

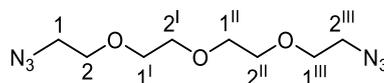
6.7 Hz, 3<sup>IV</sup>-H); <sup>13</sup>C NMR (176 MHz, CDCl<sub>3</sub>) δ 171.4 (C-1<sup>I</sup>), 170.1 (C-1), 156.4 (C=O), 148.3 (C=O), 143.9 (C-9a<sup>II</sup>), 141.4 (C-4a<sup>II</sup>), 139.0 (C-4<sup>III</sup>), 136.9 (C-2<sup>III</sup>), 127.8 (C-3<sup>II</sup>), 127.2 (C-2<sup>II</sup>), 125.3 (C-1<sup>II</sup>), 120.1 (C-4<sup>II</sup>), 115.1 (C-5<sup>III</sup>), 83.5 (C-5<sup>IV</sup>), 69.4 (C-6<sup>IV</sup>), 68.2 (C-1<sup>IV</sup>), 67.4 (9<sup>II</sup>-CH<sub>2</sub>), 54.6 (C-2<sup>I</sup>), 52.4 (CH<sub>3</sub>), 47.2 (C-9<sup>II</sup>), 41.3 (C-2), 30.2 (C-3<sup>I</sup>), 27.5 (C-2<sup>IV</sup>), 24.7 (C-3<sup>IV</sup>), 18.1 (C-4<sup>IV</sup>); HRMS (ES<sup>+</sup>) 573.2361 [M+H]<sup>+</sup> C<sub>31</sub>H<sub>32</sub>N<sub>4</sub>O<sub>7</sub>.

#### 10. 2-(2-{2-[2-(methanesulfonyloxy)ethoxy]ethoxy}ethoxy)ethyl methanesulfonate <sup>192</sup>



NEt<sub>3</sub> (10.0 mL, 74.1 mmol, 2.2 eq) in dry THF (10.0 ml) was added dropwise at 0 °C to a solution of tetraethylene glycol (6.5 g, 33.7 mmol, 1.0 eq) and methanesulfonyl chloride (5.7 mL, 74.1 mmol, 2.2 eq) in dry THF (20.0 ml) under N<sub>2</sub>. The reaction was stirred for 1 h at 0 °C, and then a further 3 h at RT. The THF was then removed *in vacuo* and the residue was re-dissolved in EtOAc (25.0 mL) and washed with H<sub>2</sub>O (30.0 mL x 3), dried (MgSO<sub>4</sub>) and concentrated. Purification by column chromatography (SiO<sub>2</sub>, hexane: EtOAc 0-100 %) gave the title ester as a colourless oil (9.3 g, 78 %).  $\nu_{\max}$  (ATR) 3034, 3005-2741 (CH<sub>3</sub>), 1460, 1347, 1252, 1175 cm<sup>-1</sup>; <sup>1</sup>H NMR (599 MHz, CDCl<sub>3</sub>) δ 4.35 – 4.34 (4H, m, 1-H<sub>2</sub>), 3.75 – 3.73 (4H, m, 2-H<sub>2</sub>), 3.65 – 3.60 (8H, m, O-CH<sub>2</sub>CH<sub>2</sub>-O), 3.05 (6H, s, CH<sub>3</sub>); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 70.7 (O-CH<sub>2</sub>CH<sub>2</sub>-O), 70.5 (O-CH<sub>2</sub>CH<sub>2</sub>-O), 69.4 (C-2), 69.1 (C-3), 37.7 (CH<sub>3</sub>); HRMS (ES<sup>+</sup>) 351.0791 [M+H]<sup>+</sup> C<sub>10</sub>H<sub>22</sub>O<sub>9</sub>S<sub>2</sub>.

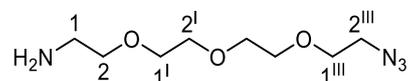
#### 11. 1-azido-2-{2<sup>I</sup>-[2<sup>II</sup>-(2<sup>III</sup>-azidoethoxy)ethoxy]ethoxy}ethane (75) <sup>192</sup>



NaHCO<sub>3</sub> (1.2 g, 13.8 mmol, 0.5 eq) and NaN<sub>3</sub> (3.7 g, 56.2 mmol, 2.2 eq) were dissolved in 10 ml H<sub>2</sub>O, then this mixture was added to 2-(2-{2-[2-(methanesulfonyloxy)ethoxy]ethoxy}ethoxy)ethyl methanesulfonate (8.9 g, 25.5 mmol, 1.0 eq). The resulting solution was heated under reflux at 80 °C for 12 h. The product was then extracted with DCM (20.0 ml x 5), dried (MgSO<sub>4</sub>) and concentrated affording the product (**75**) as a colourless oil (4.7 g, 75 %).  $\nu_{\max}$  (ATR) 2103 (N<sub>3</sub>), 1450, 1351, 1288, 1121 cm<sup>-1</sup>; <sup>1</sup>H NMR (599 MHz, CDCl<sub>3</sub>) δ 3.70 – 3.63 (12H, m, O-CH<sub>2</sub>), 3.38 (4H, t, *J* = 5.1 Hz, N<sub>3</sub>-CH<sub>2</sub>); <sup>13</sup>C

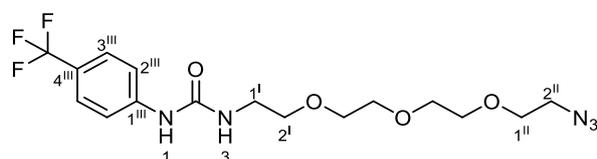
NMR (151 MHz, CDCl<sub>3</sub>) 70.8 (O-CH<sub>2</sub>), 70.1 (CH<sub>2</sub>CH<sub>2</sub>-N<sub>3</sub>), 50.8 (CH<sub>2</sub>-N<sub>3</sub>); HRMS (ES<sup>+</sup>) 267.1185 [M+Na]<sup>+</sup> C<sub>8</sub>H<sub>16</sub>N<sub>6</sub>O<sub>3</sub>.

## 12. 1<sup>I</sup>-[2-(2-aminoethoxy)ethoxy]-2<sup>I</sup>-(2<sup>II</sup>-azidoethoxy)ethane (**76**)<sup>192</sup>



Triphenylphosphine (1.8 g, 6.9 mmol, 0.7 eq) dissolved in dry Et<sub>2</sub>O (16.0 ml) was added dropwise to a solution of diazido PEG (**75**, 2.0 g, 8.1 mmol, 1 eq) dissolved in 0.65 M aqueous phosphoric acid (16.0 mL) cooled to 0 °C. The reaction was stirred for 1 h at 0 °C, then at RT for 18 h. The aqueous layer was washed with Et<sub>2</sub>O (20.0 ml x 3) and then adjusted to pH 10 by the addition of KOH pellets and extracted with DCM (15.0 ml x 4). The combined DCM extracts were concentrated *in vacuo*, dried (MgSO<sub>4</sub>) and the resultant oil triturated in Et<sub>2</sub>O. Following trituration to remove the triphenylphosphine oxide, the filtrate was concentrated, and the residue purified using column chromatography (SiO<sub>2</sub>, 0-30 % CHCl<sub>3</sub>: MeOH), to afford the titled product eluting at 30 % MeOH as a colourless oil (1.2 g, 62 %).  $\nu_{\max}$  (ATR) 3725-3086 (NH<sub>2</sub>), 3017-2745, 2109 (N<sub>3</sub>), 1460, 1291, 1126 (C-O) cm<sup>-1</sup>; <sup>1</sup>H NMR (599 MHz, CDCl<sub>3</sub>)  $\delta$  3.72 (2H, m, CH<sub>2</sub>-NH<sub>2</sub>), 3.69 – 3.65 (10H, m, O-CH<sub>2</sub>), 3.61 (2H, t, *J* = 4.7, 2-H), 3.39 (2H, t, *J* = 4.7 Hz, CH<sub>2</sub>-N<sub>3</sub>); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)  $\delta$  72.6 (C-2), 70.9 (O-CH<sub>2</sub>), 70.8 (O-CH<sub>2</sub>), 70.8 (O-CH<sub>2</sub>), 70.5 (O-CH<sub>2</sub>), 70.2 (O-CH<sub>2</sub>), 61.9 (CH<sub>2</sub>-NH<sub>2</sub>), 50.8 (CH<sub>2</sub>-N<sub>3</sub>); HRMS (ES<sup>+</sup>) 245.1377 [M+H]<sup>+</sup> C<sub>8</sub>H<sub>18</sub>N<sub>4</sub>O<sub>3</sub>.

## 13. 3-(2<sup>I</sup>-{2-[2-(2<sup>II</sup>-azidoethoxy)ethoxy]ethoxy}ethyl)-1-[4<sup>III</sup>-(trifluoromethyl)phenyl]urea (**78**)

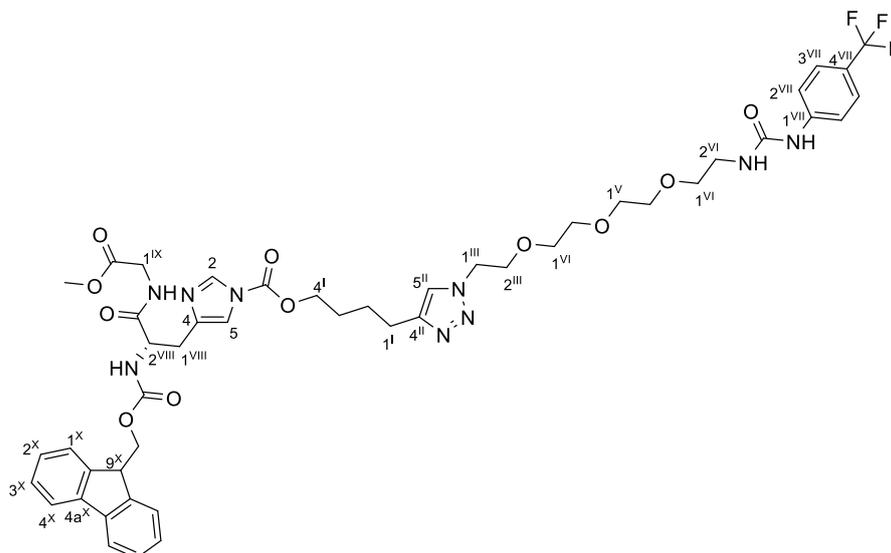


4-(trifluoromethyl)phenyl isocyanate (0.3 mL, 1.7 mmol, 1.0 eq) was added dropwise to a solution of 1<sup>I</sup>-[2-(2-aminoethoxy)ethoxy]-2<sup>I</sup>-(2<sup>II</sup>-azidoethoxy)ethane (**76**, 380.0 mg, 1.7 mmol, 1.0 eq) dissolved in dry MeCN (10.0 ml) and stirred the resultant mixture at RT for 12 h. The solvent was removed *in vacuo* and the product dissolved in EtOAc (10.0 ml) and washed with H<sub>2</sub>O (10.0 ml x 3), dried (MgSO<sub>4</sub>) and concentrated. Purification by column chromatography was used (20-100 % PET Et<sub>2</sub>O and EtOAc) gave the title named urea as a



(2H, m, 7<sup>II</sup>-H), 3.64 – 3.55 (8H, m, O-CH<sub>2</sub>), 2.75 (2H, t, *J* = 7.2 Hz, 4-H), 1.81 – 1.70 (4H, m, 2-H and 3-H); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>) δ 155.8 (O=CO=O), 153.2 (N-CO-N), 147.4 (C-4<sup>I</sup>), 141.4 (C-CF<sub>3</sub>), 126.2 (C-Ar, *q*, *J* = 3.7 Hz), 125.9 (CF<sub>3</sub>, *q*, *J* = 237.7 Hz), 125.1 (C-Ar, *d*, *J* = 34.9 Hz), 122.0 (C-5<sup>I</sup>), 118.0 (C-Ar), 70.6 (O-CH<sub>2</sub>), 70.6 (O-CH<sub>2</sub>), 70.6 (O-CH<sub>2</sub>), 70.5 (O-CH<sub>2</sub>), 69.6 (C-2<sup>II</sup>), 69.4 (C-7<sup>II</sup>), 67.8 (C-1), 64.5 (HN-CH<sub>2</sub>), 54.7 (CH<sub>3</sub>), 50.2 (C-1<sup>II</sup>), 28.2 (CH<sub>2</sub>), 25.6 (CH<sub>2</sub>), 25.2 (C-4); <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -62.02; HRMS (ES<sup>+</sup>) 563.2338 [M+2H]<sup>+</sup> C<sub>24</sub>H<sub>34</sub>N<sub>4</sub>O<sub>8</sub>F<sub>3</sub>

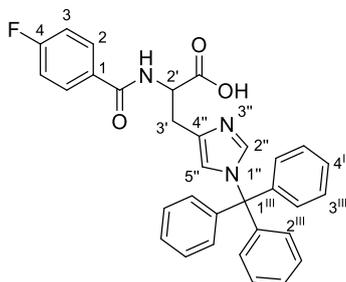
**15. 4<sup>II</sup>-{1<sup>II</sup>-[2<sup>III</sup>-(2<sup>IV</sup>-{2<sup>V</sup>-[2<sup>VI</sup>-{[4<sup>VII</sup>-(trifluoromethyl)phenyl]carbamoyl}amino)ethoxy]ethoxy}ethoxy)ethyl]-1H-1<sup>II</sup>,2<sup>II</sup>,3<sup>II</sup>-triazol-4<sup>II</sup>-yl}butyl 4-[2<sup>VIII</sup>-{[(9H-fluoren-9<sup>X</sup>-yl)methoxy]carbonyl}amino]-2<sup>IX</sup>-[(2-methoxy-2-oxoethyl)carbamoyl]ethyl]-1H-imidazole-1-carboxylate (79)**



Methyl 2-[2<sup>I</sup>-{[(9H-fluoren-9-yl)methoxy]carbonyl}amino]-3<sup>I</sup>-[(1-hex-5<sup>IV</sup>-ynoxycarboxy)-1H-imidazol-4<sup>III</sup>-yl]]propanamido]acetate (**71**, 85.0 mg, 0.2 mmol, 1.0 eq) and the urea (**78**, 66.0 mg, 0.2 mmol, 1.1 eq) were combined following general procedure C. The crude reaction mixture was dissolved in, EtOAc and the resulting solution washed with H<sub>2</sub>O (4.0 ml x 2) and brine (4.0 ml x 1), dried (MgSO<sub>4</sub>) and concentrated *in vacuo*. Purification by column chromatography (SiO<sub>2</sub>, 0-100 % hexane: acetone) produced the title triazole compound as a colourless oil (30.0 mg, 21 %).  $\nu_{\max}$  (ATR) 3451-3179 (NH), 2998-2809, 2250, 1729, 1612 (C=O), 1545, 1415 (CF), 1327, 1224 cm<sup>-1</sup>; <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>) δ 8.08 (1H, s, 2), 7.92 (1H, s, NH), 7.75 (2H, d, *J* = 7.6 Hz, 4<sup>X</sup>-H), 7.58 – 7.53 (6H, m, Ar-H), 7.48 (1H, s, 5<sup>II</sup>), 7.38 (2H, t, *J* = 7.6 Hz, 3<sup>X</sup>-H), 7.29 (2H, t, *J* = 7.6 Hz, 2<sup>X</sup>-H), 7.24 (1H, s, 5), 6.50 (1H, s,

NH), 4.57 (1H, s, CH), 4.50 (2H, t,  $J = 5.1$  Hz,  $1^{\text{III}}\text{-H}$ ), 4.40-4.38 (4H, m,  $4^{\text{I}}\text{-H}$ ,  $\text{CH}_2$ ), 4.30 (2H, t,  $J = 4.6$  Hz,  $2^{\text{VI}}\text{-H}$ ), 4.22 (1H, t,  $J = 7.3$  Hz, CH), 4.00 (2H, q,  $J = 18.2$  Hz,  $1^{\text{IX}}\text{-H}$ ), 3.86 (2H, t,  $J = 5.1$  Hz,  $2^{\text{III}}\text{-H}$ ), 3.70 – 3.69 (5H, m,  $\text{CH}_3$  and  $1^{\text{VI}}$ ), 3.60- 3.56 (8H, m, PEG O- $\text{CH}_2$ ), 3.14 (1H, d,  $J = 15.0$  Hz,  $1^{\text{VIII}}\text{-H}$ ), 3.02 (1H, d,  $J = 15.0$  Hz,  $1^{\text{VIII}}\text{-H}$ ), 2.77 (2H, t,  $J = 7.1$  Hz,  $1^{\text{I}}\text{-H}$ ), 1.84 – 1.79 (4H, m,  $2^{\text{I}}$  and  $3^{\text{I}}$ );  $^{13}\text{C}$  NMR (176 MHz,  $\text{CDCl}_3$ )  $\delta$  170.0 (C- $2^{\text{IX}}$ ), 156.5 (C=O), 153.4 (N-CO-N), 148.6 (Imid-COO), 147.2 (C=O), 143.9 (C-Ar), 141.6 (C-Ar), 141.4 (C-Ar), 137.0 (C-2), 127.9 (C- $3^{\text{X}}$ ), 127.2 (C- $2^{\text{X}}$ ), 126.3 (C- $2^{\text{VII}}$ /C- $3^{\text{VII}}$ , q,  $J = 3.7$  Hz), 125.2 (C- $2^{\text{VII}}$ /C- $3^{\text{VII}}$ , q,  $J = 18.75$ ), 124.3 ( $\text{CF}_3$ , q,  $J = 274.1$ ), 122.2 (C- $5^{\text{II}}$ ), 120.1 (C- $4^{\text{X}}$ ), 118.2 (C-Ar), 115.0 (C-5), 70.7 (O- $\text{CH}_2$ ), 70.7 (O- $\text{CH}_2$ ), 70.6 (O- $\text{CH}_2$ ), 70.6 (O- $\text{CH}_2$ ), 69.7 (O- $\text{CH}_2$ ), 69.6 (O- $\text{CH}_2$ ), 69.5 (O- $\text{CH}_2$ ), 68.2 (C- $4^{\text{I}}$ / $\text{CH}_2$ ), 67.4 (C- $4^{\text{I}}$ / $\text{CH}_2$ ), 64.5 (C- $2^{\text{VI}}$ ), 54.7 (CH), 52.4 ( $\text{CH}_3$ ), 50.4 (C- $1^{\text{III}}$ ), 47.3 (CH), 41.4 (C- $1^{\text{IX}}$ ), 30.2 (C- $1^{\text{VIII}}$ ), 28.1 ( $2^{\text{I}}$ / $3^{\text{I}}$ ), 25.8 ( $2^{\text{I}}$ / $3^{\text{I}}$ ), 25.2 (C- $1^{\text{I}}$ );  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -61.93; HRMS ( $\text{ES}^+$ ) 978.3973 [ $\text{M}+\text{H}$ ] $^+$   $\text{C}_{47}\text{H}_{54}\text{F}_3\text{N}_9\text{O}_{11}$ .

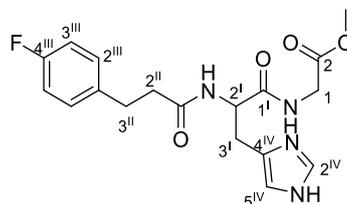
**16. 2'-[(4-fluorophenyl)formamido]-3''-[1'-(triphenylmethyl)-1H-imidazol-4'-yl]propanoic acid (83)**



4-fluorobenzoyl chloride (70  $\mu\text{L}$ , 0.6 mmol, 1.2 eq) was combined to a solution of 2-amino-3-[1-(triphenylmethyl)-1H-imidazol-4-yl]propanoic acid (200.0 mg, 0.5 mmol, 1.0 eq,) in DCM (5.0 mL), DMF (1.0 mL) and  $\text{NEt}_3$  (0.2 mL, 1.3 mmol, 2.5 eq) cooled on ice. After 1 h, the reaction mixture was washed with  $\text{H}_2\text{O}$  (10.0 mL x 4), dried ( $\text{MgSO}_4$ ) and concentrated *in vacuo*. The product was purified using column chromatography ( $\text{SiO}_2$ , 0-30 % DCM: MeOH) producing the titled acid as a light yellow oil (20.0 mg, 8 %).  $\nu_{\text{max}}$  (ATR) 3331-3035 (OH), 2700-2293, 2279-2171, 2076, 1677, 1516, 1433 (CF), 1207  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (700 MHz,  $\text{CDCl}_3$ )  $\delta$  7.69 (1H, s,  $2^{\text{II}}\text{-H}$ ), 7.64 (2H, dd,  $J = 8.5, 5.3$  Hz, 2-H), 7.31 (3H, t,  $J = 7.4$  Hz,  $4^{\text{III}}\text{-H}$ ), 7.27 – 7.25 (6H, m, Ar-H), 7.04 – 7.01 (8H, m, Ar-H), 6.62 (1H, s,  $5^{\text{II}}\text{-H}$ ), 4.71 (1H, s,  $2^{\text{I}}\text{-H}$ ), 3.46 – 3.44 (1H, m,  $3^{\text{I}}\text{-H}$ ), 3.36 – 3.33 (1H, m,  $3^{\text{I}}\text{-H}$ );  $^{13}\text{C}$  NMR (176 MHz,  $\text{CDCl}_3$ )  $\delta$  174.0 (COOH), 165.4 (CON), 164.9 (C-4, d,  $J = 252.0$  Hz), 141.6 (C-Ar), 137.4 (C- $1^{\text{III}}$ ), 134.3 (C- $4^{\text{II}}$ ), 130.4 (C-1, d,  $J = 2.9$  Hz), 129.7 (C- $2^{\text{II}}$ ), 129.4 (C-2, d,  $J = 8.8$  Hz), 128.6 (C- $4^{\text{III}}$ ), 128.4 (C- $2^{\text{III}}$

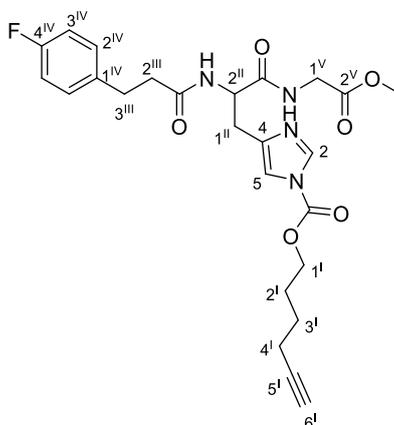
and 3<sup>III</sup>), 120.8 (C-5<sup>II</sup>), 115.6 (C-3, d,  $J = 21.9$  Hz), 53.2 (C-2<sup>I</sup>), 28.9 (C-3<sup>I</sup>); <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>)  $\delta$  -108.27; HRMS (ES<sup>+</sup>) 520.2040 [M+H]<sup>+</sup> C<sub>32</sub>H<sub>26</sub>FN<sub>3</sub>O<sub>3</sub>.

**17. Methyl 2-{2<sup>I</sup>-[3<sup>II</sup>-(4<sup>III</sup>-fluorophenyl)propanimido]-3<sup>I</sup>-(1H-imidazol-4<sup>IV</sup>-yl)propanamido}acetate (88)**



Under N<sub>2</sub>, DBU (0.12 mL, 0.7 mmol, 2.0 eq) was added to a solution of methyl 2-[2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino]-3-[1-(triphenylmethyl)-1H-imidazol-4-yl]propanamido}acetate (**70**, 0.3 g, 0.4 mmol, 1.0 eq,) in dry THF (4.0 mL) which was stirred for 10 mins, monitoring the reaction by TLC until the complete removal of Fmoc. The solvent was then removed *en vacuo* and redissolved in dry DCM (4.0 mL). To this, EDC (16.0 mg, 0.9 mmol, 1.1 eq), HOBt (115.0 mg, 0.9 mmol, 1.1 eq) and NEt<sub>3</sub> (500.0  $\mu$ L, 1.2 mmol, 1.5 eq) were combined to a solution of methyl 2-{2-amino-3-[1-(triphenylmethyl)-1H-imidazol-4-yl]propanamido}acetate) (362.0 mg, 0.8 mmol, 1.0 eq) dissolved in dry DCM (6.0 mL) under N<sub>2</sub> and the reaction stirred for 14 h. The organic layer was washed with NaHCO<sub>3</sub> (10.0 ml x 2) and H<sub>2</sub>O (10.0 ml x 2), dried (MgSO<sub>4</sub>) and concentrated *en vacuo*. The resulting mixture was dissolved in DCM (7.5 mL) and TFA (4.0 ml) was added dropwise and stirred at RT for 4 h. The solvent was removed *in vacuo* and the product triturated using cold Et<sub>2</sub>O producing the deprotected product as a white solid (120.3 g, 40 %).  $\nu_{\max}$  (ATR) 3498-3249 (NH), 2659-2322, 2235, 2073, 1676 (C=O), 1458 (CH<sub>3</sub>), 1363 (CF), 1232 cm<sup>-1</sup>; <sup>1</sup>H NMR (700 MHz, MeOD, rotamer)  $\delta$  7.58 (1H, s, 2<sup>IV</sup>-H), 7.03 (2H, td,  $J = 8.7, 5.4$  Hz, 2<sup>III</sup>-H), 6.83 (2H, td,  $J = 8.5, 2.3$  Hz, 3<sup>III</sup>-H), 6.77 (1H, s, 5<sup>IV</sup>-H), 4.54 (1H, dd,  $J = 8.7, 5.4$  Hz, 2<sup>I</sup>-H), 3.86 – 3.78 (2H, m, 1-H), 3.59 (3H, s, CH<sub>3</sub>), 2.99 (1H, dd,  $J = 15.0, 5.4$  Hz, 3<sup>I</sup>-H), 2.80 (1H, dd,  $J = 15.0, 5.4$  Hz, 3<sup>I</sup>-H), 2.71 (2H, t,  $J = 7.6$  Hz, 3<sup>II</sup>-H), 2.38 (2H, td,  $J = 7.6$  Hz, 2<sup>II</sup>-H); <sup>13</sup>C NMR (176 MHz, MeOD)  $\delta$  173.7 (C-1<sup>II</sup>), 172.7 (C-1<sup>I</sup>), 170.2 (C-2), 161.4 (d,  $J = 242.3$  Hz, CF), 136.6 (d,  $J = 3.2$ , C-1<sup>III</sup>), 134.9 (C-2<sup>IV</sup>), 129.6 (d,  $J = 8.0$  Hz, C-2<sup>III</sup>), 126.6 (C-4<sup>IV</sup>), 117.5 (C-5<sup>IV</sup>), 114.6 (d,  $J = 21.3$  Hz, C-3<sup>III</sup>), 53.1 (C-2<sup>I</sup>), 51.3 (CH<sub>3</sub>), 40.5 (C-1), 37.2 (C-2<sup>II</sup>), 30.3 (C-3<sup>II</sup>), 28.6 (C-3<sup>I</sup>), 13.0; <sup>19</sup>F NMR (376 MHz, MeOD)  $\delta$  -119.20; HRMS (ES<sup>+</sup>) 377.1624 [M+H]<sup>+</sup> C<sub>18</sub>H<sub>21</sub>FN<sub>4</sub>O<sub>4</sub>.

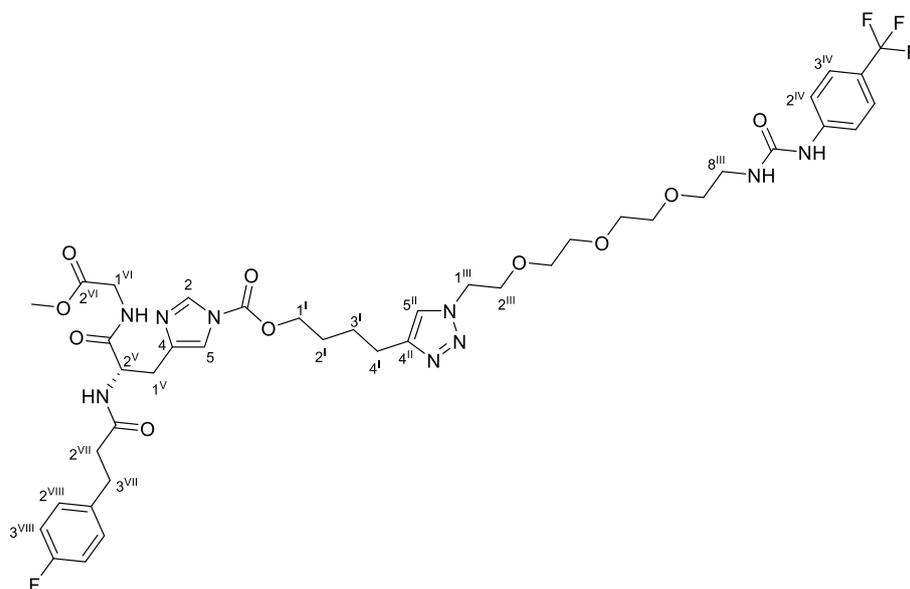
**18. Hex-5<sup>I</sup>-yn-1<sup>I</sup>-yl 4-{2<sup>II</sup>-[3<sup>III</sup>-(4<sup>IV</sup>-fluorophenyl)propanamido]-2<sup>II</sup>-[(2<sup>V</sup>-methoxy-2<sup>V</sup>-oxoethyl)carbamoyl]ethyl}-1H-imidazole-1<sup>V</sup>-carboxylate (**89**)**



Bis(hex-5-yn-1-yl) dicarbonate (**66**, 10.0 mg, 0.04 mmol, 1.2 eq,) dissolved in dry DCM (1.0 mL) was added dropwise to methyl 2-{2<sup>I</sup>-[3<sup>II</sup>-(4<sup>III</sup>-fluorophenyl)propanimido]-3<sup>I</sup>-(1H-imidazol-4<sup>IV</sup>-yl)propanamido}acetate (**88**, 12.0 mg, 0.03 mmol, 1 eq) in DMF (1.0 mL) and the reaction was stirred at RT overnight. The organic layer was washed with NaHCO<sub>3</sub> (1 x 5.0 mL) followed by H<sub>2</sub>O (3 x 5.0 mL), dried (MgSO<sub>4</sub>) and concentrated. To purify, column chromatography was performed (SiO<sub>2</sub>, 50-100 % hexane: acetone) producing a clear oil (10.0 mg, 63 %).  $\nu_{\max}$  (ATR) 3434-3164, 2960, 2872, 2240 (C≡C), 1760, 1651, 1539, 1514, 1439, 1411 (CF), 1384, 1249, 1219, 1187 cm<sup>-1</sup>; <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>)  $\delta$  8.08 (1H, s, 2-H), 7.22 (1H, s, 5-H), 7.16 (2H, t, *J* = 8.6, Hz, 2<sup>IV</sup>-H), 6.93 (2H, t, *J* = 8.6 Hz, 3<sup>IV</sup>-H), 4.71 (1H, q, *J* = 5.7 Hz, 2<sup>II</sup>-H), 4.43 (2H, td, *J* = 6.5, 2.2 Hz, 1<sup>I</sup>-H), 3.97 (1H, dd, *J* = 18.1, 5.5 Hz, 1<sup>V</sup>-H), 3.88 (1H, dd, *J* = 18.1, 5.5 Hz, 1<sup>V</sup>-H), 3.70 (3H, s, CH<sub>3</sub>), 3.09 (1H, dd, *J* = 15.0, 5.5 Hz, 1<sup>II</sup>-H), 2.95 (2H, t, *J* = 7.6 Hz, 3<sup>III</sup>-H), 2.89 (1H, dd, *J* = 15.0, 5.5 Hz, 1<sup>II</sup>-H), 2.56 (2H, t, *J* = 7.6 Hz, 2<sup>III</sup>-H), 2.28 (2H, td, *J* = 6.9, 2.6 Hz, 4<sup>I</sup>-H), 1.98 (1H, t, *J* = 2.6 Hz, 6<sup>I</sup>-H), 1.92 (2H, dt, *J* = 14.4, 7.0 Hz, 2<sup>I</sup>-H), 1.66 (2H, dt, *J* = 14.4, 7.0 Hz, 3<sup>I</sup>-H); <sup>13</sup>C NMR (176 MHz, CDCl<sub>3</sub>)  $\delta$  172.4 (C-1<sup>III</sup>), 171.2 (C-3<sup>II</sup>), 170.0 (C-2<sup>V</sup>), 161.6 (d, *J* = 244.2 Hz, C-4<sup>IV</sup>), 148.4 (Imid-COO), 139.3 (C-4), 136.7 (C-2), 136.4 (d, *J* = 3.3 Hz, C-1<sup>IV</sup>), 129.9 (d, *J* = 7.8 Hz, C-2<sup>IV</sup>), 115.3 (d, *J* = 21.1 Hz, C-3<sup>IV</sup>), 115.1 (C-5), 83.4 (C-5<sup>I</sup>), 69.4 (C-6<sup>I</sup>), 68.2 (C-1<sup>I</sup>), 52.8 (C-2<sup>II</sup>), 52.4 (CH<sub>3</sub>), 41.2 (C-1<sup>V</sup>), 38.3 (C-2<sup>III</sup>), 30.7 (C-3<sup>III</sup>), 29.4 (C-1<sup>II</sup>), 27.6 (C-2<sup>I</sup>), 24.7 (C-3<sup>I</sup>), 18.1 (C-4<sup>I</sup>); <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>)  $\delta$  -116.98; HRMS (ES<sup>+</sup>) 501.2153 [M+H]<sup>+</sup> C<sub>25</sub>H<sub>29</sub>FN<sub>4</sub>O<sub>6</sub>.

**19. 4-{1-[2-(2-[2-2-({4<sup>IV</sup>-(trifluoromethyl)phenyl]carbamoyl)amino)ethoxy]ethoxy]ethoxy}ethyl]-1H-1<sup>II</sup>,2<sup>II</sup>,3<sup>II</sup>-**

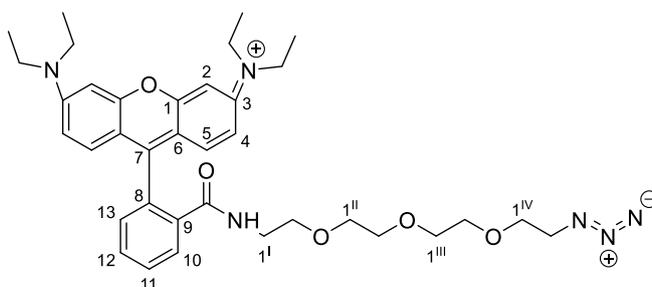
triazol-4<sup>l</sup>-yl)butyl 4-{2<sup>v</sup>-[3<sup>vii</sup>-(4<sup>viii</sup>-fluorophenyl)propanamido]-2-[(2-methoxy-2-oxoethyl)carbamoyl]ethyl}-1H-imidazole-1<sup>vi</sup>-carboxylate (**90**)



A solution of hex-5-yn-1-yl 4-{2-[3-(4-fluorophenyl)propanamido]-2-[(2-methoxy-2-oxoethyl)carbamoyl]ethyl}-1H-imidazole-1-carboxylate (**89**, 43.0 mg, 0.1 mmol, 1.0. eq) and 3-(2-{2-[2-(2-azidoethoxy)ethoxy]ethoxy}ethyl)-1-[4-(trifluoromethyl)phenyl]urea (**78**, 38.0 mg, 0.1 mmol, 1.1 eq) formed the desired product following the procedure detailed in general procedure C. The reaction mixture was purified using column chromatography (SiO<sub>2</sub>, 0-100 % acetone: MeCN) with the product appearing as a colourless oil (36.0 mg, 47 %).  $\nu_{\max}$  (ATR) 3377-3182 (NH), 2988-2821, 2251, 1660 (C=O), 1614, 1546, 1411 (CF), 1324, 1225, 1114, 1071 cm<sup>-1</sup>; <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>)  $\delta$  8.12 (1H, s, 2-H), 7.95 (1H, s, N-H), 7.60 – 7.51 (4H, m, 2<sup>iv</sup>-H, 3<sup>iv</sup>-H), 7.50 (1H, s, 5<sup>ii</sup>-H), 7.22 (1H, s, 5-H), 7.15 (2H, dd, *J* = 8.5, 5.5 Hz, 2<sup>viii</sup>-H), 6.95 – 6.90 (2H, dd, *J* = 8.5, 5.5 Hz, 3<sup>viii</sup>-H), 4.73 (1H, q, *J* = 5.8 Hz, 2<sup>v</sup>-H), 4.51 (2H, t, *J* = 5.1 Hz, 1<sup>iii</sup>-H), 4.41 (2H, td, *J* = 6.1, 2.1 Hz, 1<sup>l</sup>-H), 4.31 (2H, t, *J* = 4.5 Hz, 8<sup>iii</sup>-H), 3.96 (1H, dd, *J* = 18.1, 5.7 Hz, 1<sup>vi</sup>-H), 3.90 (1H, d, *J* = 5.7 Hz, 1<sup>vi</sup>-H), 3.87 (2H, t, *J* = 5.1 Hz, 2<sup>iii</sup>-H), 3.70 (2H, t, *J* = 4.5 Hz, 7<sup>iii</sup>-H), 3.69 (3H, s, CH<sub>3</sub>), 3.63 – 3.56 (8H, m, O-CH<sub>2</sub>), 3.08 (1H, dd, *J* = 15.1, 5.8 Hz, 1<sup>v</sup>-H), 2.93 (3H, m, 1<sup>v</sup>-H, 3<sup>vii</sup>-H), 2.78 (2H, t, *J* = 7.1 Hz, 4<sup>l</sup>-H), 2.56 (2H, t, *J* = 7.6 Hz, 2<sup>vii</sup>-H), 1.89 – 1.78 (4H, m, 3<sup>l</sup>-H, 2<sup>l</sup>-H); <sup>13</sup>C NMR (176 MHz, CDCl<sub>3</sub>)  $\delta$  172.6 (C-1<sup>vii</sup>), 171.3 (C-3<sup>v</sup>), 170.0 (C-2<sup>vi</sup>), 161.6 (d, *J* = 244.1 Hz, CF), 153.4 (N-CO-N), 148.3 (Imid-COO), 147.2 (C-4<sup>ii</sup>), 141.6 (C-Ar), 139.0 (C-4), 136.6 (C-2), 136.4 (d, *J* = 2.7 Hz, C-1<sup>viii</sup>), 130.0 (d, *J* = 7.9 Hz, C-2<sup>viii</sup>), 126.4 (d, *J* = 3.8 Hz, C-Ar), 124.3 (q, *J* = 255.0 Hz, CF<sub>3</sub>), 122.3 (C-5<sup>ii</sup>), 118.2 (C-Ar), 115.3 (d, *J* = 21.1 Hz, C-3<sup>viii</sup>), 115.2 (C-5), 70.7 (O-CH<sub>2</sub>), 70.7 (O-CH<sub>2</sub>), 70.6 (O-CH<sub>2</sub>), 69.7 (C-2<sup>iii</sup>),

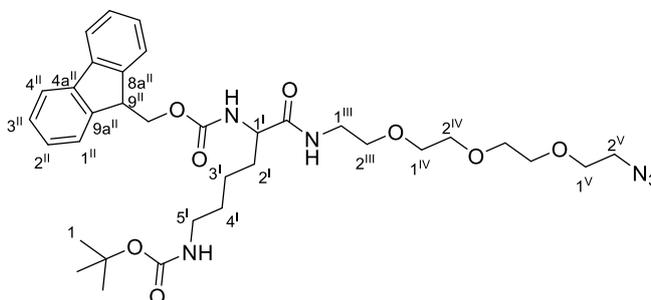
69.5 (O-CH<sub>2</sub>), 68.5 (C-1<sup>I</sup>), 64.5 (C-8<sup>III</sup>), 52.8 (C-2<sup>V</sup>), 52.4 (CH<sub>3</sub>), 50.4 (C-1<sup>III</sup>), 41.3 (C-1<sup>VI</sup>), 38.3 (C-2<sup>VII</sup>), 30.7 (C-3<sup>VII</sup>), 29.4 (C-1<sup>V</sup>), 28.1 (C-3<sup>I</sup>), 25.7 (C-2<sup>I</sup>), 25.2 (C-4<sup>I</sup>); <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -61.94, -116.99; HRMS (ES<sup>+</sup>) 907.3647 [M+2H]<sup>+</sup> C<sub>41</sub>H<sub>51</sub>F<sub>4</sub>N<sub>9</sub>O<sub>10</sub>.

**20. 9-{2<sup>I</sup>-[(2<sup>II</sup>-{2<sup>III</sup>-[2<sup>IV</sup>-(2<sup>IV</sup>-azidoethoxy)ethoxy]ethoxy)ethoxy]ethyl}carbonyl]phenyl}-6<sup>V</sup>-(diethylamino)-N,N-diethyl-3H-xanthen-3-iminium (93)**



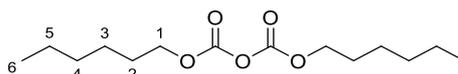
EDC (96.0 mg, 0.5 mmol, 1.0 eq) was added to the solution of rhodamine B (200.0 mg, 0.5 mmol, 1.0 eq) and NHS (58.0 mg, 0.5 mmol, 1.0 eq) dissolved in MeCN (5.0 mL) heated to 45 °C and stirred overnight. The solvent was removed *in vacuo* and then the rhodamine-NHS was dissolved in DCM (2.0 mL) and NEt<sub>3</sub> (0.2 mL, 1.5 mmol, 3.0 eq) under N<sub>2</sub> in which a solution of 1-[2-(2-aminoethoxy)ethoxy]-2-(2-azidoethoxy)ethane (**76**, 109.0 mg, 0.5 mmol, 1.0 eq,) dissolved in DCM (2.0 mL) was added dropwise. The reaction was stirred overnight (18 h) with the reaction wrapped in foil to limit light exposure. The solvent was removed *in vacuo* and this dark oil was purified by column chromatography (SiO<sub>2</sub>, 0-30 % DCM: MeOH), with the product eluting at 30 % MeOH, appearing as a dark brown oil (198.0 mg, 62 %).  $\nu_{\max}$  (ATR) 2981, 2939, 2189 (N<sub>3</sub>), 1753, 1640 (C=O), 1594, 1524, 1471, 1415, 1340, 1276, 1252, 1185 cm<sup>-1</sup>; <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>) δ 8.35 – 8.31 (1H, m, 10-H), 7.81 (1H, td, *J* = 7.8, 1.4 Hz, 12-H), 7.74 (1H, td, *J* = 7.8, 1.4 Hz, 11-H), 7.32 – 7.29 (1H, m, 13-H), 7.07 (2H, d, *J* = 9.5 Hz, 5-H), 6.91 (2H, dd, *J* = 9.5, 2.4 Hz, 4-H), 6.83 (2H, d, *J* = 2.4 Hz, 2-H), 4.20 – 4.15 (2H, m, 1<sup>I</sup>-H), 3.68 – 3.58 (17H, m, CH<sub>2</sub>, NH), 3.56 (4H, tt, *J* = 6.6, 3.8 Hz, CH<sub>2</sub>), 3.36 (2H, t, *J* = 5.0 Hz, CH<sub>2</sub>-N<sub>3</sub>), 1.33 (12H, t, *J* = 7.1 Hz, CH<sub>3</sub>); <sup>13</sup>C NMR (176 MHz, CDCl<sub>3</sub>) δ 165.1 (C=ON), 159.0 (C-7), 157.9 (C-1), 155.7 (C-3), 133.9 (C-Ar), 133.3 (C-12), 131.7 (C-10), 131.5 (C-Ar), 130.5 (C-Ar), 130.4 (C-Ar), 129.9 (C-Ar), 114.4 (C-4), 133.7 (C-6), 96.5 (C-2), 70.8 (CH<sub>2</sub>), 70.7 (CH<sub>2</sub>), 70.7 (CH<sub>2</sub>), 70.6 (CH<sub>2</sub>), 70.6 (CH<sub>2</sub>), 70.2 (CH<sub>2</sub>), 68.8 (CH<sub>2</sub>), 64.8 (C-1<sup>I</sup>), 50.8 (CH<sub>2</sub>-N<sub>3</sub>), 46.3 (CH<sub>2</sub>), 12.8 (CH<sub>3</sub>); HRMS (ES<sup>+</sup>) 644.3467 [M+H]<sup>+</sup> C<sub>36</sub>H<sub>47</sub>N<sub>6</sub>O<sub>5</sub>.

**21. (9H-fluoren-9-yl)methyl N-{1-[(2-{2-[2-(2-azidoethoxy)ethoxy]ethoxy}ethyl)carbamoyl]-5-[[tert-butoxy]carbonyl]amino}pentyl}carbamate (99)**



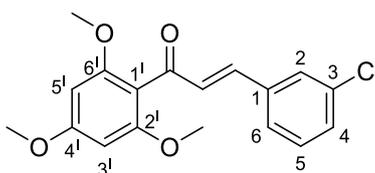
NHS (109.0 mg, 0.9 mmol, 2.0 eq) and EDC (181.0 mg, 0.9 mmol, 2.0 eq) was added to a solution of (2S)-6-[[tert-butoxy]carbonyl]amino-2-[[9H-fluoren-9-yl)methoxy]carbonyl]amino hexanoic acid (221.0 mg, 0.5 mmol, 1.0 eq) dissolved in DMF (8.0 ml) and stirred for 15 min at RT. 1-[2-(2-aminoethoxy)ethoxy]-2-(2-azidoethoxy)ethane (**76**, 100.0 mg, 0.5 mmol, 1.0 eq) was combined and the reaction stirred for 28 h. 15 % citric acid was added to the reaction mixture and the product was extracted using EtOAc (4 x 10.0 ml). The organic layer was then washed with brine (1 x 15.0 mL), dried (MgSO<sub>4</sub>) and concentrated. Purification occurred using column chromatography (C18-bonded SiO<sub>2</sub>, 0-100 % H<sub>2</sub>O:MeOH), which gives the product as a light yellow oil (136.0 mg, 44 %).  $\nu_{\max}$  (ATR) 2992-2814, 2107 (N<sub>3</sub>), 1706 (C=O), 1665, 1522, 1451, 1249, 1169 cm<sup>-1</sup>; <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>)  $\delta$  7.76 (2H, d,  $J = 7.5$  Hz, 4<sup>II</sup>-H), 7.60 (2H, d,  $J = 7.5$  Hz, 1<sup>I</sup>-H), 7.40 (2H, t,  $J = 7.5$  Hz, 3<sup>II</sup>-H), 7.32 (2H, t,  $J = 7.5$  Hz, 2<sup>II</sup>-H), 6.54 (1H, s, NH), 5.55 (1H, s, NH), 4.65 (1H, s, NH), 4.40 (2H, t,  $J = 6.6$  Hz, 9<sup>II</sup>-CH<sub>2</sub>), 4.21 (1H, t,  $J = 6.6$  Hz, 9<sup>II</sup>-H), 4.17 – 4.11 (1H, m, 1<sup>I</sup>-H), 3.68 – 3.58 (10H, m, 1<sup>IV</sup>-1<sup>V</sup>-H), 3.56 (2H, s, 2<sup>III</sup>-H), 3.47 (2H, s, 1<sup>III</sup>-H), 3.36 (2H, t,  $J = 5.1$  Hz, CH<sub>2</sub>-N<sub>3</sub>), 3.10 (2H, m, 5<sup>I</sup>-H), 1.84 (2H, m, 2<sup>I</sup>-H), 1.48 (2H, m, 4<sup>I</sup>-H), 1.43 (9H, s, CH<sub>3</sub>), 1.36 (2H, t,  $J = 8.0$  Hz, 3<sup>I</sup>-H); <sup>13</sup>C NMR (176 MHz, CDCl<sub>3</sub>)  $\delta$  171.7 (CON-PEG), 156.3 (C-3), 144.0 (Fmoc-COON), 144.0 (C-9a<sup>II</sup>), 141.5 (C-4a<sup>II</sup>), 127.9 (C-3<sup>II</sup>), 127.2 (C-2<sup>II</sup>), 125.2 (C-1<sup>II</sup>), 120.1 (C-4<sup>II</sup>), 70.8 (C-1<sup>IV</sup> to 1<sup>V</sup>), 70.8 (C-1<sup>IV</sup> to 1<sup>V</sup>), 70.7 (C-1<sup>IV</sup> to 1<sup>V</sup>), 70.4 (C-2<sup>II</sup> to 1<sup>V</sup>), 70.1 (C-1<sup>IV</sup> to 1<sup>V</sup>), 69.7 (C-2<sup>III</sup>), 67.1 (Fmoc-CH<sub>2</sub>), 55.0 (C-1<sup>I</sup>), 50.8 (CH<sub>2</sub>-N<sub>3</sub>), 47.3 (C-9<sup>II</sup>), 42.9 (C-5<sup>I</sup>), 39.5 (C-1<sup>III</sup>), 32.7 (C-2<sup>I</sup>), 29.8 (C-4<sup>I</sup>), 28.6 (CH<sub>3</sub>), 22.6 (C-3<sup>I</sup>); HRMS (ES<sup>+</sup>) 691.3420 [M+H]<sup>+</sup> C<sub>34</sub>H<sub>48</sub>N<sub>6</sub>O<sub>8</sub>.

**22. Dihexyl dicarbonate (104)**



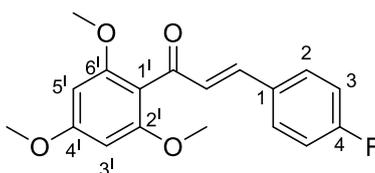
1-hexanol (0.13 mL; 2.0 mmol, 1.0 eq) was added to a solution of triphosgene (1484.0 mg; 1.0 mmol; 0.5 eq) in anhydrous DCM (2 ml). The reaction mixture was then stirred at RT for 3 h to generate the chloroformate intermediate. The reaction mixture was concentrated *in vacuo* and the crude chloroformate used immediately for the synthesis of the title dicarbonate using general procedure A. Following the final concentration, the product was observed as a light yellow oil (228.0 mg; 42 %).  $\nu_{\max}$  (ATR) 2980, 1984, 1797 (C=O), 1462, 1385  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.33 – 4.26 (4H, m, 1-H), 1.77 – 1.73 (4H, m, 2-H), 1.43 – 1.27 (12H, m,  $\text{CH}_2$ ), 0.98 – 0.82 (6H, m,  $\text{CH}_3$ );  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  148.1 (C=O), 72.6 (C-1), 31.4 (C-2), 28.4 ( $\text{CH}_2$ ), 25.6 ( $\text{CH}_2$ ), 22.6 ( $\text{CH}_2$ ), 14.1 ( $\text{CH}_3$ ).

### 23. (2E)-3-(3-chlorophenyl)-1'-(2', 4', 6'-trimethoxyphenyl)prop-2-en-1-one



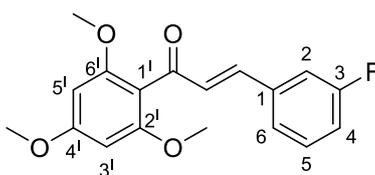
Following procedure D, the acetophenone (120.0 mg, 0.6 mmol, 1.0 eq),  $\text{Ba}(\text{OH})_2 \cdot \text{H}_2\text{O}$  (46.7 mg, 0.2 mmol, 0.25 eq) and 3-chlorobenzaldehyde (91.0  $\mu\text{L}$ , 0.8 mmol, 1.4 eq) were mixed together in dry MeOH (5.0 mL) overnight at 40 °C. After workup and recrystallisation, the pure product was obtained as a yellow solid (99.0 mg, 50 %).  $\nu_{\max}$  (ATR) 2980, 2889, 1649 (C=C), 1605, 1463 ( $\text{CH}_3$ ), 1384, 1207  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (599 MHz,  $\text{CDCl}_3$ )  $\delta$  7.49 (1H, s, 2-H), 7.40 (1H, d,  $J = 7.4$ , 4-H), 7.34 – 7.27 (2H, m, 5-H, 6-H), 7.30 (1H,  $J = 16.0$  Hz,  $\beta\text{H}$ ), 6.94 (1H, d,  $J = 16.0$  Hz,  $\alpha\text{H}$ ), 6.16 (2H, s, 3'-H, 5'-H), 3.86 (3H, s, 4'- $\text{CH}_3$ ), 3.78 (6H, s, 2'- $\text{CH}_3$ , 6'- $\text{CH}_3$ );  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  193.8 (C=O), 162.8 (C-4'), 159.1 (C-2', C-6'), 142.1 (C $\beta$ ), 137.1 (C-1), 135.0 (C-3), 130.3 (C $\alpha$ ), 130.2 (C-Ar), 130.1 (C-Ar), 128.3 (C-2), 126.6 (C-4), 111.8 (C-1'), 90.9 (C-3', C-5'), 56.1 (2'- $\text{CH}_3$ , 6'- $\text{CH}_3$ ), 55.6 (4'- $\text{CH}_3$ ); HRMS ( $\text{ES}^+$ ) 333.0905 [ $\text{M}+\text{H}^+$ ]  $\text{C}_{18}\text{H}_{17}^{35}\text{ClO}_4$ .

### 24. (2E)-3-(4-fluorophenyl)-1'-(2', 4', 6'-trimethoxyphenyl)prop-2-en-1-one



Following procedure D, the acetophenone (80.0 mg, 0.4 mmol, 1.0 eq), Ba(OH)<sub>2</sub>·H<sub>2</sub>O (30.0 mg, 0.1 mmol, 0.3 eq) and 4-fluorobenzaldehyde (47.2 μL, 0.4 mmol, 1.0 eq) were mixed together in dry MeOH (4.0 mL) overnight at 40 °C. Upon confirmation through TLC that the reaction had finished, the reaction was filtered then the solvent removed *in vacuo*. The resulting yellow oil was then dissolved in EtOAc (10.0 mL) then quenched with 1 M HCl (5.0 mL), then washed with H<sub>2</sub>O (10.0 ml x 2) and brine (10.0 ml x 1), dried (MgSO<sub>4</sub>) and concentrated. The pure product was obtained as a yellow solid by recrystallising the mixture with Et<sub>2</sub>O and hexane (115 mg, 90 %).  $\nu_{\max}$  (ATR) 2981, 1651 (C=C), 1602, 1458 (CH<sub>3</sub>), 1227, 1157, 1129 cm<sup>-1</sup>; <sup>1</sup>H NMR (599 MHz, CDCl<sub>3</sub>)  $\delta$  7.52 – 7.49 (2H, m, 2-*H*), 7.33 (1H, d, *J* = 16.0 Hz,  $\beta$ *H*), 7.05 (2H, t, *J* = 8.6 Hz, 3-*H*), 6.88 (1H, d, *J* = 16.0 Hz,  $\alpha$ *H*), 6.16 (2H, s, 3'*-H*, 5'*-H*), 3.86 (3H, s, 4'*-CH*<sub>3</sub>), 3.77 (6H, s, 2'*-CH*<sub>3</sub>, 6'*-CH*<sub>3</sub>); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)  $\delta$  194.1 (C=O), 164.0 (d, *J* = 251.2 Hz, CF), 162.6 (C-4'), 159.0 (C-2', 6'), 142.8 (C $\beta$ ), 131.4 (d, *J* = 3.3 Hz, C-1), 130.4 (d, *J* = 8.5 Hz, C-2), 128.9 (C $\alpha$ ), 116.1 (d, *J* = 21.8 Hz, C-3), 111.9 (C-1'), 90.9 (C-3', 5'), 56.1 (2', 6'*-CH*<sub>3</sub>), 55.6 (4'*-CH*<sub>3</sub>); <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>)  $\delta$  -61.94; HRMS (ES<sup>+</sup>) 317.1214 [M+H<sup>+</sup>] C<sub>18</sub>H<sub>17</sub>FO<sub>4</sub>.

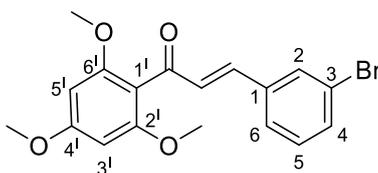
#### 25. (2E)-3-(3-fluorophenyl)-1'-(2', 4', 6'-trimethoxyphenyl)prop-2-en-1-one



Following procedure D, the acetophenone (80.0 mg, 0.4 mmol, 1.0 eq), Ba(OH)<sub>2</sub>·H<sub>2</sub>O (30.0 mg, 0.1 mmol, 0.3 eq) and 3-fluorobenzaldehyde (40.0 μL, 0.4 mmol, 1.0 eq) were mixed together in dry MeOH (4.0 mL) overnight at 40 °C. Upon confirmation through TLC that the reaction had finished, the reaction was filtered then the solvent removed *in vacuo*. The resulting yellow oil was then dissolved in EtOAc (10.0 mL) then quenched with 1 M HCl (5.0 mL), then washed with H<sub>2</sub>O (10.0 ml x 2) and brine (10.0 ml x 1), dried (MgSO<sub>4</sub>) and concentrated. The pure product was obtained as a yellow solid by recrystallising the mixture with Et<sub>2</sub>O and hexane (101.3 mg, 84 %).  $\nu_{\max}$  (ATR) 2972, 1655 (C=C), 1605, 1456 (CH<sub>3</sub>),

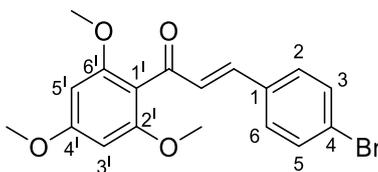
1227, 1157, 1128  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.36 – 7.25 (3H, m, Ar-H,  $\beta\text{H}$ ) 7.22 – 7.20 (1H, m, 2-H), 7.09 – 7.02 (1H, m, 4-H), 6.94 (1H, d,  $J = 16.0$  Hz,  $\alpha\text{H}$ ), 6.16 (2H, s, 3<sup>I</sup>-H, 5<sup>I</sup>-H), 3.86 (3H, s, 4<sup>I</sup>-CH<sub>3</sub>), 3.78 (6H, s, 2<sup>I</sup>-CH<sub>3</sub>, 6<sup>I</sup>-CH<sub>3</sub>);  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  193.8 (C=O), 163.1 (d,  $J = 246.6$  Hz, CF), 162.8 (C-4<sup>I</sup>), 159.1 (C-2<sup>I</sup>, 6<sup>I</sup>), 142.3 (C $\beta$ ), 137.6 (d,  $J = 7.7$  Hz, C-1), 130.4 (d,  $J = 8.2$  Hz, C-5), 124.5 (d,  $J = 2.9$  Hz, C-6), 117.1 (d,  $J = 21.7$  Hz, C-4), 114.7 (d,  $J = 21.7$  Hz, C-2), 111.8 (C-1<sup>I</sup>), 90.9 (C-3<sup>I</sup>, 5<sup>I</sup>), 56.1 (2<sup>I</sup>, 6<sup>I</sup>-CH<sub>3</sub>), 55.6 (4<sup>I</sup>-CH<sub>3</sub>);  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -112.8; HRMS (ES<sup>+</sup>) 317.1187 [M+H<sup>+</sup>] C<sub>18</sub>H<sub>17</sub>FO<sub>4</sub>.

## 26. (2E)-3-(3-bromophenyl)-1<sup>I</sup>-(2<sup>I</sup>, 4<sup>I</sup>, 6<sup>I</sup>-trimethoxyphenyl)prop-2-en-1-one



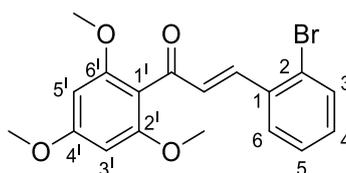
The acetophenone (200.0 mg, 1.0 mmol, 1.0 eq), Ba(OH)<sub>2</sub>·H<sub>2</sub>O (75.0 mg, 0.2 mmol, 0.3 eq) and 3-bromobenzaldehyde (110.0  $\mu\text{L}$ , 1.0 mmol, 1.0 eq) were mixed together in dry MeOH (5.0 mL) overnight at 40 °C. Upon confirmation through TLC that the reaction had finished, the reaction was filtered then the solvent removed *in vacuo*. The resulting yellow oil was then dissolved in EtOAc (10.0 mL) then quenched with 1 M HCl (10.0 mL), then washed with H<sub>2</sub>O (15.0 mL x 2) and brine (15.0 mL x 1), dried (MgSO<sub>4</sub>) and concentrated. The pure product was obtained as a yellow solid by recrystallising the mixture with Et<sub>2</sub>O and hexane (182.0 mg, 51 %).  $\nu_{\text{max}}$  (ATR) 2981, 1649 (C=C), 1605, 1463 (CH<sub>3</sub>), 1393, 1207, 1157, 1129  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (599 MHz,  $\text{CDCl}_3$ )  $\delta$  7.65 (1H, t,  $J = 1.8$  Hz, 2-H), 7.52 – 7.43 (2H, m, Ar-H), 7.29 (1H, d,  $J = 16.0$  Hz,  $\beta\text{H}$ ), 7.24 (1H, t,  $J = 7.9$  Hz, Ar-H), 6.94 (1H, d,  $J = 16.0$  Hz,  $\alpha\text{H}$ ), 6.16 (2H, s, 5<sup>II</sup>-H, 3<sup>II</sup>-H), 3.86 (3H, s, 4<sup>II</sup>-CH<sub>3</sub>), 3.78 (6H, s, 2<sup>II</sup>-H, 6<sup>II</sup>-CH<sub>3</sub>);  $^{13}\text{C}$  NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  193.8 (C=O), 162.8 (C-4<sup>I</sup>), 159.1 (C-2<sup>I</sup>, C-6<sup>I</sup>), 142.0 (C $\beta$ ), 137.4 (C-1), 133.0 (Ar-H), 131.2 (C-2), 130.4 (C-Ar), 130.3 (C $\alpha$ ), 127.0 (Ar-H), 123.1 (C-Br), 111.8 (C-1), 90.9 (C-3<sup>I</sup>, 5<sup>I</sup>), 56.1 (2<sup>II</sup>, 6<sup>II</sup>-CH<sub>3</sub>), 55.6 (4<sup>I</sup>-CH<sub>3</sub>); HRMS (ES<sup>+</sup>) 377.0378 [M+H<sup>+</sup>] C<sub>18</sub>H<sub>17</sub><sup>79</sup>BrO<sub>4</sub>.

## 27. (2E)-3-(4-bromophenyl)-1<sup>I</sup>-(2<sup>I</sup>, 4<sup>I</sup>, 6<sup>I</sup>-trimethoxyphenyl)prop-2-en-1-one



Following procedure D, the acetophenone (140.0 mg, 0.7 mmol, 1.0 eq), Ba(OH)<sub>2</sub>·H<sub>2</sub>O (52.5 mg, 0.2 mmol, 0.3 eq) and 4-bromobenzaldehyde (123.2 mg, 0.7 mmol, 1.0 eq) were dissolved in dry MeOH (5.0 mL) and stirred overnight at 40 °C. After an initial filtration to remove a white precipitate, the pure product recrystallised as the solution cooled as a light yellow solid (91.0 mg, 41 %).  $\nu_{\max}$  (ATR) 2979, 1648 (C=C), 1605, 1456 (CH<sub>3</sub>), 1206, 1157, 1130 cm<sup>-1</sup>; <sup>1</sup>H NMR (599 MHz, CDCl<sub>3</sub>)  $\delta$  7.51 – 7.48 (2H, m, 3-*H*, 5-*H*), 7.39 – 7.37 (2H, m, 2-*H*, 6-*H*), 7.31 (1H, d, *J* = 16.0 Hz,  $\beta$ *H*), 6.94 (1H, d, *J* = 16.0 Hz,  $\alpha$ *H*), 6.16 (2H, s, 3'*-H*, 5'*-H*), 3.86 (3H, s, 4'*-OCH*<sub>3</sub>), 3.77 (6H, s, 2'*-OCH*<sub>3</sub>, 6'*-OCH*<sub>3</sub>); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)  $\delta$  193.9 (C=O), 162.7 (C-4'), 159.1 (C-4', C-6'), 142.5 (C $\beta$ ), 134.2 (C-1), 132.2 (C-3, C-5), 129.9 (C-2, C-4), 129.7 (C $\alpha$ ), 124.5 (C-Br), 111.9 (C-1'), 91.0 (C-3', C-5'), 56.1 (2'*-OCH*<sub>3</sub>, 6'*-OCH*<sub>3</sub>), 55.6 (4'*-OCH*<sub>3</sub>); HRMS (ES<sup>+</sup>) 377.0399 [M+H<sup>+</sup>] C<sub>18</sub>H<sub>17</sub><sup>79</sup>BrO<sub>4</sub>.

### 28. (2E)-3-(2-bromophenyl)-1'-(2', 4', 6'-trimethoxyphenyl)prop-2-en-1-one



Following procedure D, the acetophenone (140.0 mg, 0.7 mmol, 1.0 eq), Ba(OH)<sub>2</sub>·H<sub>2</sub>O (53.0 mg, 0.2 mmol, 0.3 eq) and 2-bromobenzaldehyde (78.0  $\mu$ L, 0.7 mmol, 1.4 eq) dissolved in dry MeOH (4.0 mL) and stirred overnight at 40 °C. The pure product was obtained as a yellow solid by recrystallising the mixture with Et<sub>2</sub>O and hexane (149.0 mg, 59 %).  $\nu_{\max}$  (ATR) 2981, 1653 (C=C), 1605, 1465 (CH<sub>3</sub>), 1207, 1157, 1128 cm<sup>-1</sup>; <sup>1</sup>H NMR (599 MHz, CDCl<sub>3</sub>)  $\delta$  7.72 (1H, d, *J* = 16.0 Hz, *H* $\beta$ ), 7.65 (1H, dd, *J* = 7.9, 1.7 Hz, 6-*H*), 7.58 (1H, dd, *J* = 7.9, 1.7 Hz, 3-*H*), 7.32 (1H, m, 5-*H*), 7.20 (1H, td, *J* = 7.9, 1.7 Hz, 4-*H*), 6.86 (1H, d, *J* = 16.0 Hz, *H* $\alpha$ ), 6.16 (2H, s, 3'*-H*, 5'*-H*), 3.86 (3H, s, 4'*-CH*<sub>3</sub>), 3.78 (6H, s, 2'*-CH*<sub>3</sub>, 6'*-CH*<sub>3</sub>); <sup>13</sup>C NMR (151 MHz, CDCl<sub>3</sub>)  $\delta$  194.3 (C=O), 162.7 (C-4'), 159.1 (C-2', C-6'), 142.9 (C $\beta$ ), 135.3 (C-1), 133.5 (C-3), 131.6 (C $\alpha$ ), 131.2 (C-4), 128.1 (C-6), 127.8 (C-5), 125.7 (C-Br), 111.5 (C-1'), 90.9 (C-3', C-5'), 56.1 (2'*-OCH*<sub>3</sub>, 6'*-OCH*<sub>3</sub>), 55.6 (4'*-OCH*<sub>3</sub>); HRMS (ES<sup>+</sup>) 377.0392 [M+H<sup>+</sup>] C<sub>18</sub>H<sub>17</sub><sup>79</sup>BrO<sub>4</sub>.

## 8 Appendix 1

Due to the coronavirus lockdown, the biological aspect of this work was unable to proceed because of issues with laboratory access. Therefore, during this period, a brief SAR project was undertaken synthesising drugs with a chalcone scaffold.

The first reported chalcone was from Fu *et al.* (2004) which was isolated from licorice root (*Glycyrrhiza glabra*). This initial structure (**110**) was found to be effective at causing cell arrest in prostate cancer cells (Figure 113).<sup>253</sup> The basic scaffold of chalcones is an  $\alpha,\beta$ -unsaturated carbonyl bridge in the middle of two aromatic regions.

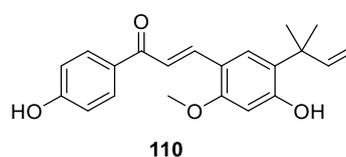


Figure 113 Licochalcone first isolated from licorice root

Following the initial discovery, this structure was explored for a variety of diseases including leishmaniasis. Initial exploration of chalcones as an antileishmanial resulted in cytotoxicity with human cell lines, with a  $CC_{50}$  range of 13 - >55  $\mu\text{M}$ , with one scoring over >100  $\mu\text{M}$ .<sup>254</sup> A brief SAR campaign from Boeck *et al.* (2006) explored 2',6'-dihydroxy-4'-methoxychalcones to decrease human cell line cytotoxicity whilst retaining the anti-leishmanial properties. The reported chalcones produced  $IC_{50}$  values of 0.4 - >100  $\mu\text{M}$  for promastigotes and 3.4 - >100  $\mu\text{M}$  for amastigotes with a  $CC_{50}$  range of 7.6 - >100  $\mu\text{M}$ .<sup>255</sup> Similarly to this reported scaffold, 2',4',6'-trimethoxyacetophenone (**111**) was used in this project for the left-hand ring. The chalcone series was synthesised by base catalysed Claisen–Schmidt condensation (Figure 114).<sup>256</sup>

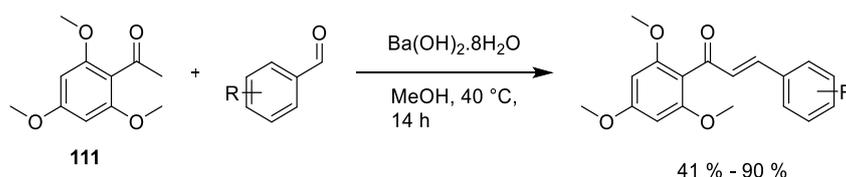


Figure 114 Chalcone synthesis

Using the established synthetic route, a series of chalcone derivatives were synthesised to explore the optimal halogen substituent and position on the right-hand ring (Figure 115).

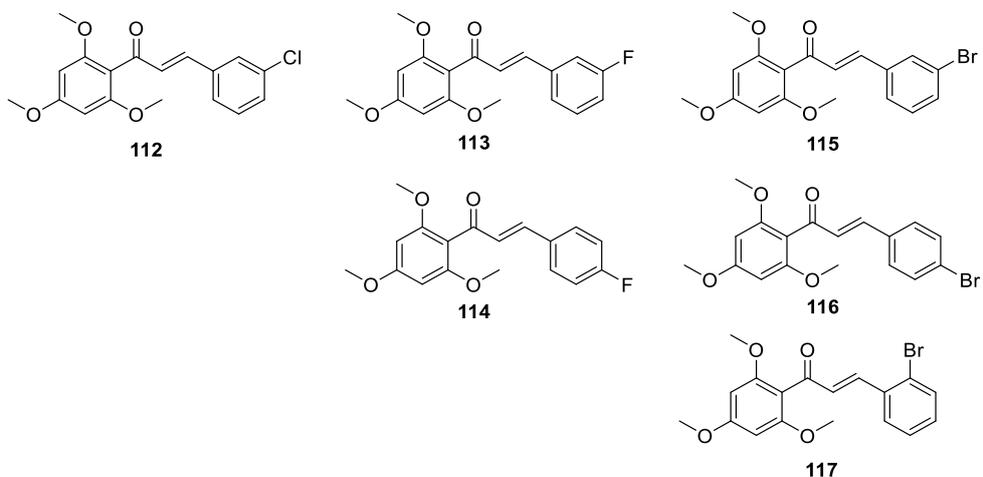


Figure 115 Series of chalcone analogues

The synthesised chalcones will be tested in a promastigote inhibition assay to assess the  $IC_{50}$ , and thus determine the optimal halogen substituent and position for antileishmanial activity.

## 9 Appendix 2

Table 13 Reported enzymes discussed in Chapter 3.7

n	Protein Group	Protein ID	Protein Accession	-10lgP	Coverage (%)	Coverage (%) Sample 2	#Peptides	#Unique	#Spec Sample 2	PTM	Avg. Mass	Cellular location
1	1	3332	P0CE47 EFTU1_ECOLI	449.98	69	69	61	61	97	Oxidation (M)	43284	Cytoplasm
1	1	393	P0CE48 EFTU2_ECOLI	449.98	69	69	61	61	97	Oxidation (M)	43314	Cytoplasm
2	3	9965	P0A6M8 EFG_ECOLI	446.26	54	54	54	54	62	Oxidation (M)	77581	Cytoplasm
3	6	8642	P0A8T7 RPOC_ECOLI	425.29	37	37	53	53	54	Oxidation (M)	155160	Nucleoid
4	9	9527	P0A8V2 RPOB_ECOLI	460.25	40	40	48	48	49	Oxidation (M)	150632	Nucleoid
5	9	9971	P36683 ACNB_ECOLI	419.55	42	42	38	38	43	Oxidation (M)	93498	Cytosol
6	34	8269	P06959 ODP2_ECOLI	339.77	34	34	36	36	19	Oxidation (M)	66096	Cytoplasm
7	18	10021	P00957 SYA_ECOLI	339.69	37	37	32	32	32	Oxidation (M)	96032	Cytoplasm
8	15	8121	P0AFG8 ODP1_ECOLI	338.69	33	33	29	29	31	Oxidation (M)	99669	Mitochondria
9	18	11629	P0A8M3 SYT_ECOLI	312.79	32	32	23	23	24	Oxidation (M)	74014	Cytoplasm
10	19	7866	P0A836 SUCC_ECOLI	303.58	45	45	23	23	23	Oxidation (M)	41393	Mitochondria
11	30	280	P62707 GPMA_ECOLI	263.42	48	48	21	21	13		28556	Cytoplasm
12	24	11635	P0A799 PGK_ECOLI	320.39	40	40	21	21	17	Oxidation (M)	41118	Cytoplasm
13	17	8347	P05055 PNP_ECOLI	322.27	34	34	21	21	22	Oxidation (M)	77101	Cytoplasm

Table 14 Unreported proteins with a high coverage (%) but low peptide fragment numbers (#peptide)

Protein Group	Protein ID	Protein Accession	-10lgP	Coverage		#Peptides	#Unique	#Spec Sample 2	PTM	Avg. Mass
				Coverage (%)	Coverage (%) Sample 2					
19	9342	P0A870 TALB_ECOLI	302.29	50	50	16	16	18	Oxidation (M)	35219
22	14320	P08312 SYFA_ECOLI	274.27	50	50	15	15	15	Oxidation (M)	36832
38	3285	P69441 KAD_ECOLI	237.28	50	50	9	9	9	Oxidation (M)	23586
41	10662	P0A8F0 UPP_ECOLI	288.14	49	49	8	8	8	Oxidation (M)	22533
14	9528	P08200 IDH_ECOLI	337.74	48	48	18	18	20	Oxidation (M)	45757
33	10085	P0ACJ8 CRP_ECOLI	230.55	48	48	10	10	11	Oxidation (M)	23640
23	12538	P00509 AAT_ECOLI	289.88	46	46	19	19	19	Oxidation (M)	43573
17	12548	P0A6B7 ISCS_ECOLI	297.99	42	42	14	14	14	Oxidation (M)	45090
23	10663	P0AE08 AHPC_ECOLI	263.97	42	42	11	11	13	Oxidation (M)	20761
16	7849	P0A9B2 G3P1_ECOLI	297.27	40	40	19	19	22	Oxidation (M)	35533
31	14321	P0A796 PFKA_ECOLI	233.27	39	39	17	17	11	Oxidation (M)	34842
37	8319	P62620 ISPG_ECOLI	271.36	39	39	14	14	14	Oxidation (M)	40684
38	7786	P0A7V8 RS4_ECOLI	232.31	39	39	9	9	9	Oxidation (M)	23469
42	3304	P0A9V1 LPTB_ECOLI	222.97	39	39	8	8	8	Oxidation (M)	26801
19	9041	P76373 UDG_ECOLI	257.83	38	38	15	15	15	Oxidation (M)	43657
32	10169	P67910 HLDD_ECOLI	225.93	38	38	11	11	11	Oxidation (M)	34893
39	1012	P76034 YCIT_ECOLI	213.33	36	36	9	9	9	Oxidation (M)	27603
44	11883	P0AB77 KBL_ECOLI	269.25	36	36	12	12	12	Oxidation (M)	43117
41	12540	P0AEB2 DACA_ECOLI	255.66	35	35	13	12	13	Oxidation (M)	44444
55	10667	P45799 NUDE_ECOLI	199.51	35	35	5	5	5	Oxidation (M)	21153
28	9432	P0A817 METK_ECOLI	286.7	34	34	11	11	11	Oxidation (M)	41952
40	12541	P77690 ARNB_ECOLI	296.05	34	34	11	11	11	Oxidation (M)	42238
41	1535	P07014 SDHB_ECOLI	200.24	34	34	7	7	7	Oxidation (M)	26770
46	10664	P0A7D1 PTH_ECOLI	187.08	34	34	6	6	6	Oxidation (M)	21082

48	12542	P08506 DACC_ECOLI	244.43	34	34	11	10	11	Oxidation (M)	43609
24	7778	P0A7V3 RS3_ECOLI	210.88	33	33	6	6	6		25983
59	10890	P0A7G6 RECA_ECOLI	244.59	33	33	9	9	9	Oxidation (M)	37973
19	8399	P33602 NUOG_ECOLI	357.33	32	32	28	28	29	Oxidation (M)	100299
28	9040	P77398 ARNA_ECOLI	273.07	32	32	18	18	18		74289
31	229	P23836 PHOP_ECOLI	182.2	32	32	6	6	6	Oxidation (M)	25535
48	10668	P69503 APT_ECOLI	183.34	32	32	5	5	5		19859
55	12545	P23721 SERC_ECOLI	234.62	32	32	9	9	9	Oxidation (M)	39783
23	10773	P0AFG3 ODO1_ECOLI	339.61	31	31	25	25	25	Oxidation (M)	105062

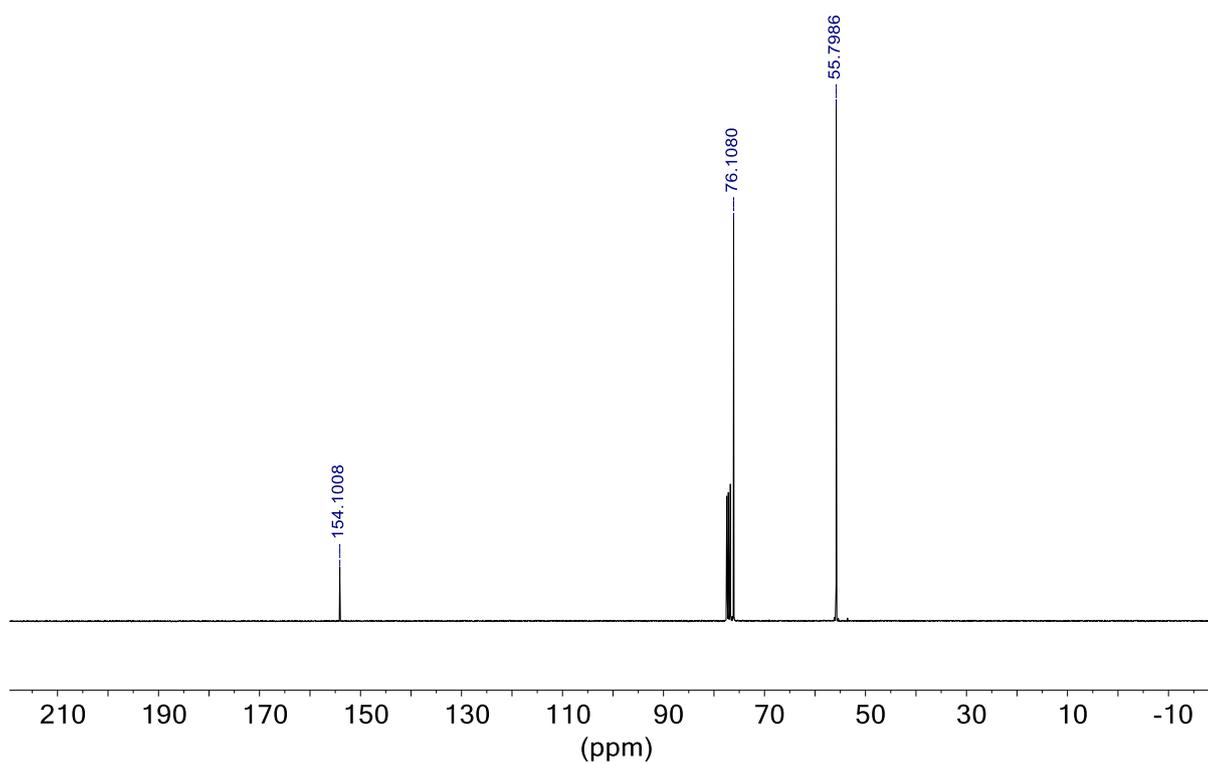
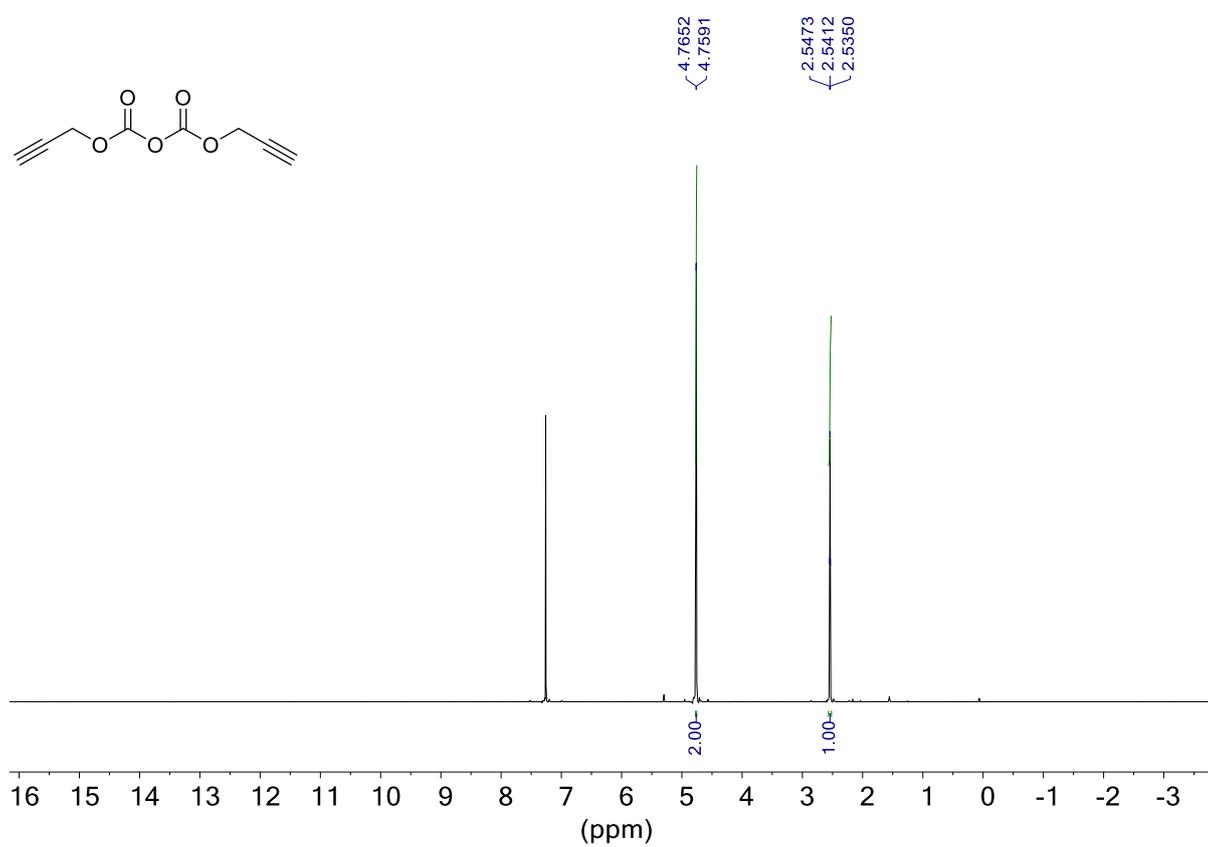
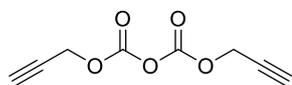
Table 15 Unreported proteins with a high coverage (%) but low peptide fragment numbers (#peptide) description in chronological order from Table 14

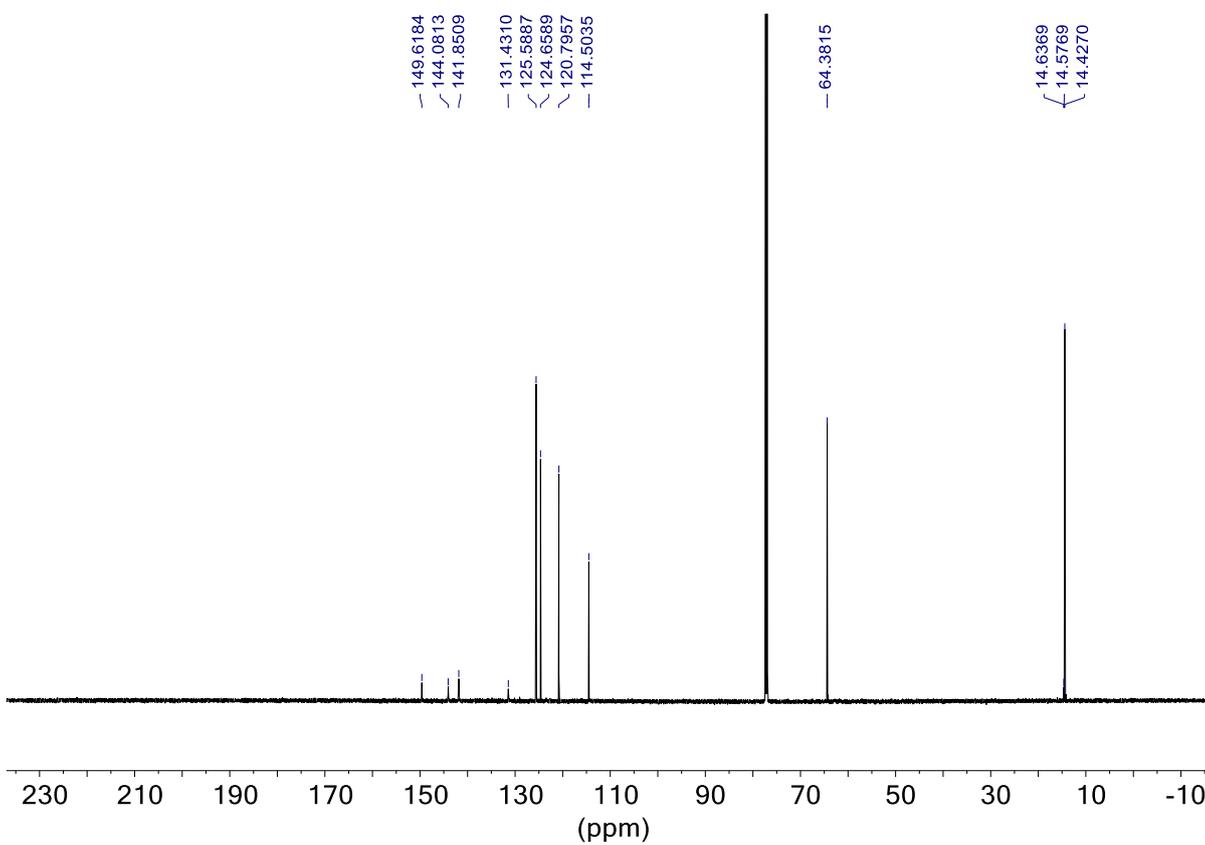
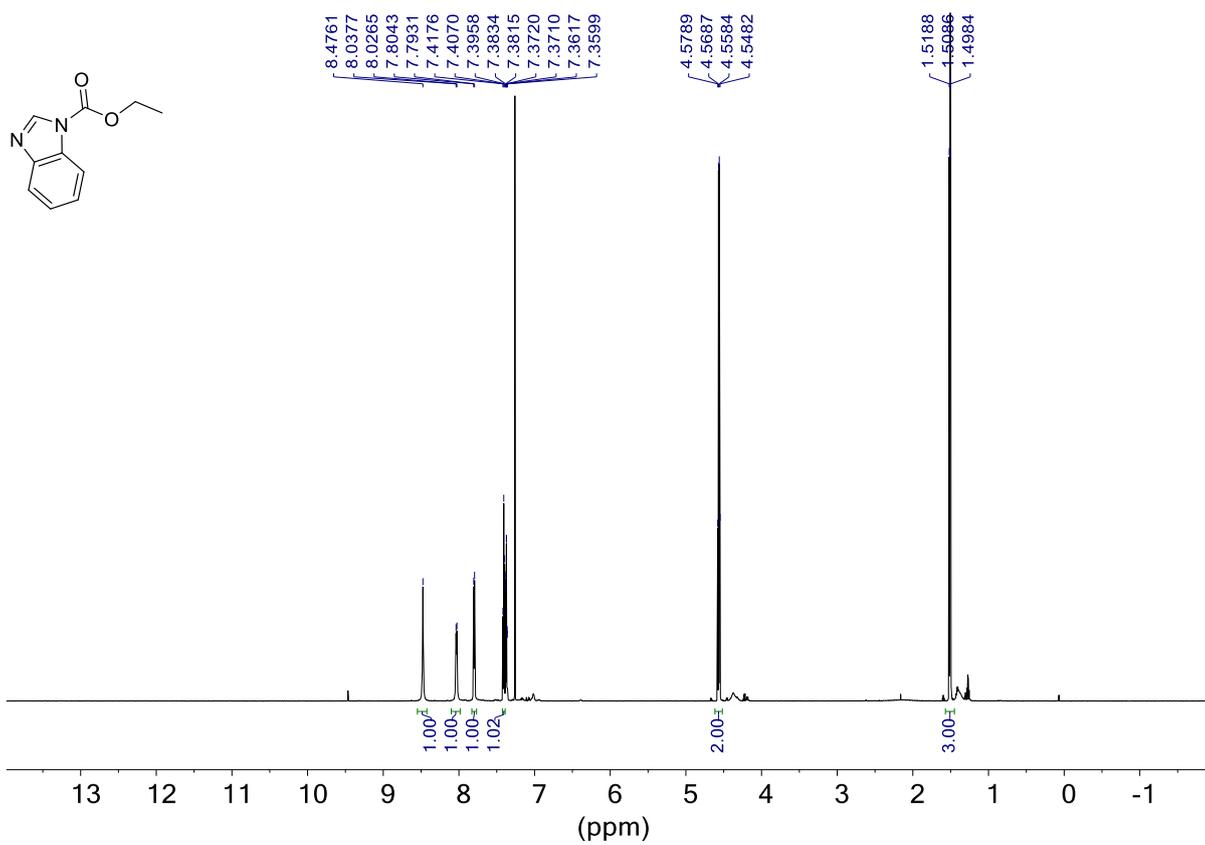
Protein Group	Protein ID	Description
19	9342	Transaldolase B OS=Escherichia coli (strain K12) GN=talB PE=1 SV=2
22	14320	Phenylalanine--tRNA ligase alpha subunit OS=Escherichia coli (strain K12) GN=pheS PE=1 SV=2
38	3285	Adenylate kinase OS=Escherichia coli (strain K12) OX=833333 GN=adk PE=1 SV=1
41	10662	Uracil phosphoribosyltransferase OS=Escherichia coli (strain K12) GN=upp PE=1 SV=1
14	9528	Isocitrate dehydrogenase [NADP] OS=Escherichia coli (strain K12) GN=icd PE=1 SV=1
33	10085	cAMP-activated global transcriptional regulator CRP OS=Escherichia coli (strain K12) GN=crp PE=1 SV=1
23	12538	Aspartate aminotransferase OS=Escherichia coli (strain K12) GN=aspC PE=1 SV=1
17	12548	Cysteine desulfurase IscS OS=Escherichia coli (strain K12) GN=iscS PE=1 SV=1
23	10663	Alkyl hydroperoxide reductase subunit C OS=Escherichia coli (strain K12) GN=ahpC PE=1 SV=2
16	7849	Glyceraldehyde-3-phosphate dehydrogenase A OS=Escherichia coli (strain K12) GN=gapA PE=1 SV=2
31	14321	ATP-dependent 6-phosphofructokinase isozyme 1 OS=Escherichia coli (strain K12) GN=pfkA PE=1 SV=1
37	8319	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin) OS=Escherichia coli (strain K12) GN=ispG PE=1 SV=1
38	7786	30S ribosomal protein S4 OS=Escherichia coli (strain K12) GN=rpsD PE=1 SV=2
42	3304	Lipopolysaccharide export system ATP-binding protein LptB OS=Escherichia coli (strain K12) OX=833333 GN=lptB PE=1 SV=2

19 9041 UDP-glucose 6-dehydrogenase OS=Escherichia coli (strain K12) GN=ugd PE=1 SV=1  
32 10169 ADP-L-glycero-D-manno-heptose-6-epimerase OS=Escherichia coli (strain K12) GN=hldD PE=1 SV=1  
39 1012 Uncharacterized HTH-type transcriptional regulator YciT OS=Escherichia coli (strain K12) OX=83333 GN=yciT PE=4 SV=1  
44 11883 2-amino-3-ketobutyrate coenzyme A ligase OS=Escherichia coli (strain K12) GN=kbl PE=1 SV=1  
41 12540 D-alanyl-D-alanine carboxypeptidase DacA OS=Escherichia coli (strain K12) GN=dacA PE=1 SV=1  
55 10667 ADP compounds hydrolase NudE OS=Escherichia coli (strain K12) GN=nudE PE=1 SV=1  
28 9432 S-adenosylmethionine synthase OS=Escherichia coli (strain K12) GN=metK PE=1 SV=2  
40 12541 UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase OS=Escherichia coli (strain K12) GN=arnB PE=1 SV=3  
41 1535 Succinate dehydrogenase iron-sulfur subunit OS=Escherichia coli (strain K12) OX=83333 GN=sdhB PE=1 SV=1  
46 10664 Peptidyl-tRNA hydrolase OS=Escherichia coli (strain K12) GN=pth PE=1 SV=1  
48 12542 D-alanyl-D-alanine carboxypeptidase DacC OS=Escherichia coli (strain K12) GN=dacC PE=1 SV=2  
24 7778 30S ribosomal protein S3 OS=Escherichia coli (strain K12) GN=rpsC PE=1 SV=2  
59 10890 Protein RecA OS=Escherichia coli (strain K12) GN=recA PE=1 SV=2  
19 8399 NADH-quinone oxidoreductase subunit G OS=Escherichia coli (strain K12) GN=nuoG PE=1 SV=4  
28 9040 Bifunctional polymyxin resistance protein ArnA OS=Escherichia coli (strain K12) GN=arnA PE=1 SV=1  
31 229 Transcriptional regulatory protein PhoP OS=Escherichia coli (strain K12) OX=83333 GN=phoP PE=1 SV=1  
48 10668 Adenine phosphoribosyltransferase OS=Escherichia coli (strain K12) GN=apt PE=1 SV=1  
55 12545 Phosphoserine aminotransferase OS=Escherichia coli (strain K12) GN=serC PE=1 SV=4  
23 10773 2-oxoglutarate dehydrogenase E1 component OS=Escherichia coli (strain K12) GN=sucA PE=1 SV=1

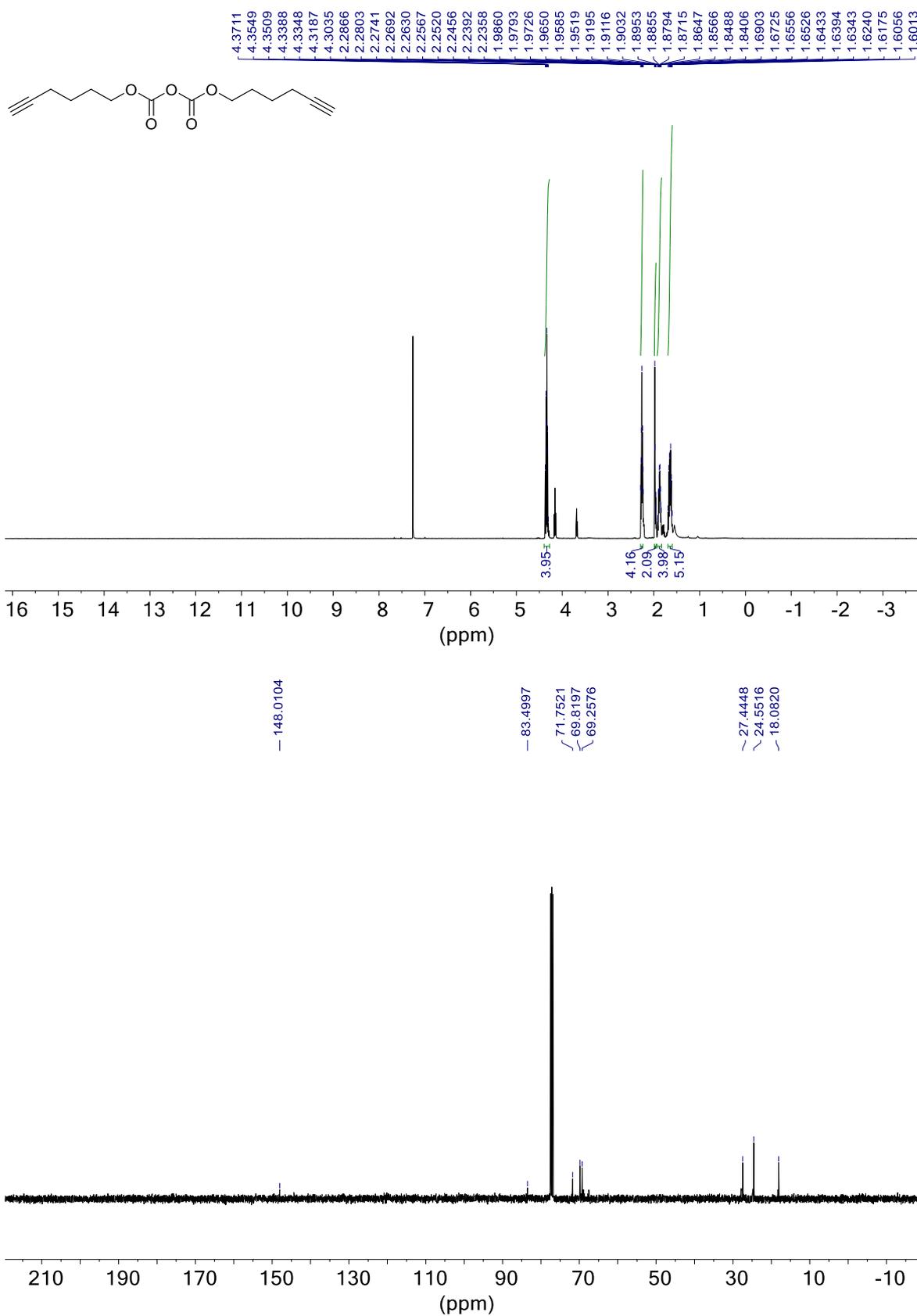
## 10 Appendix 3

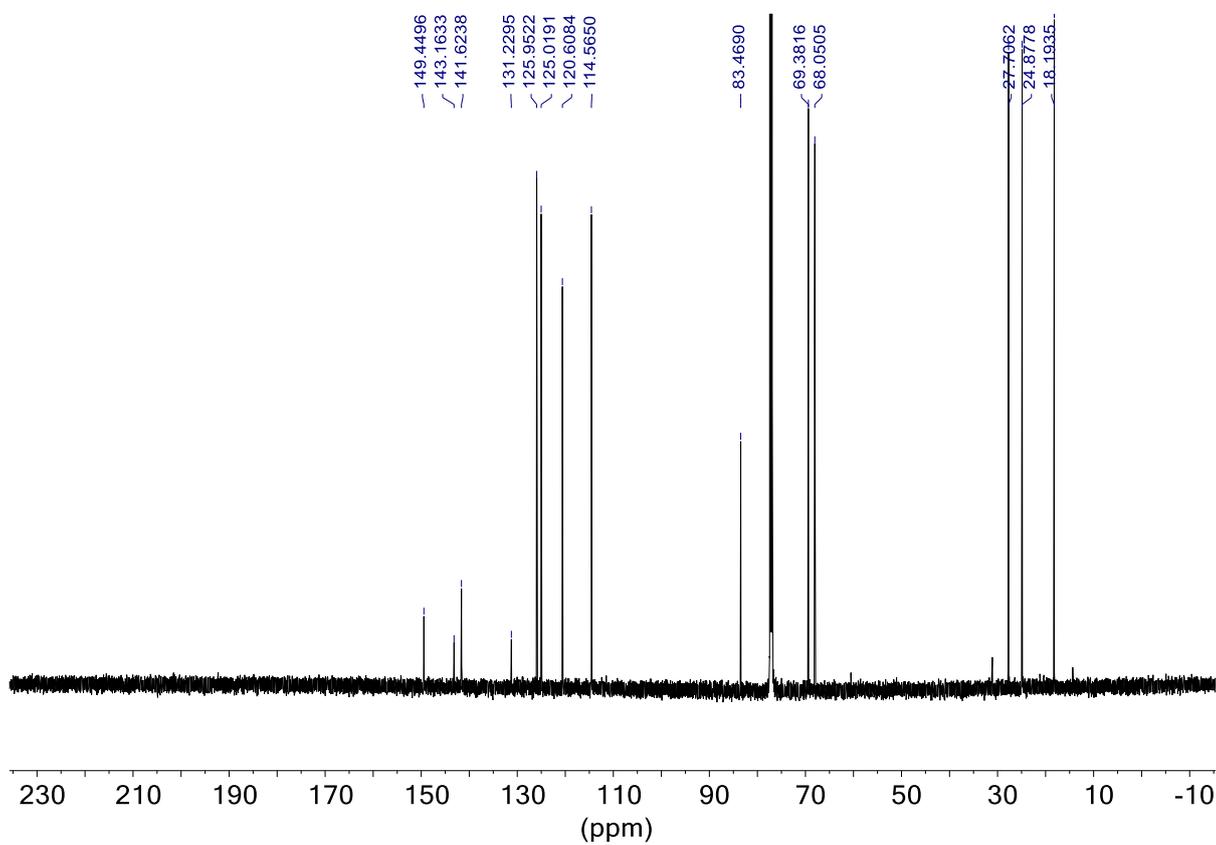
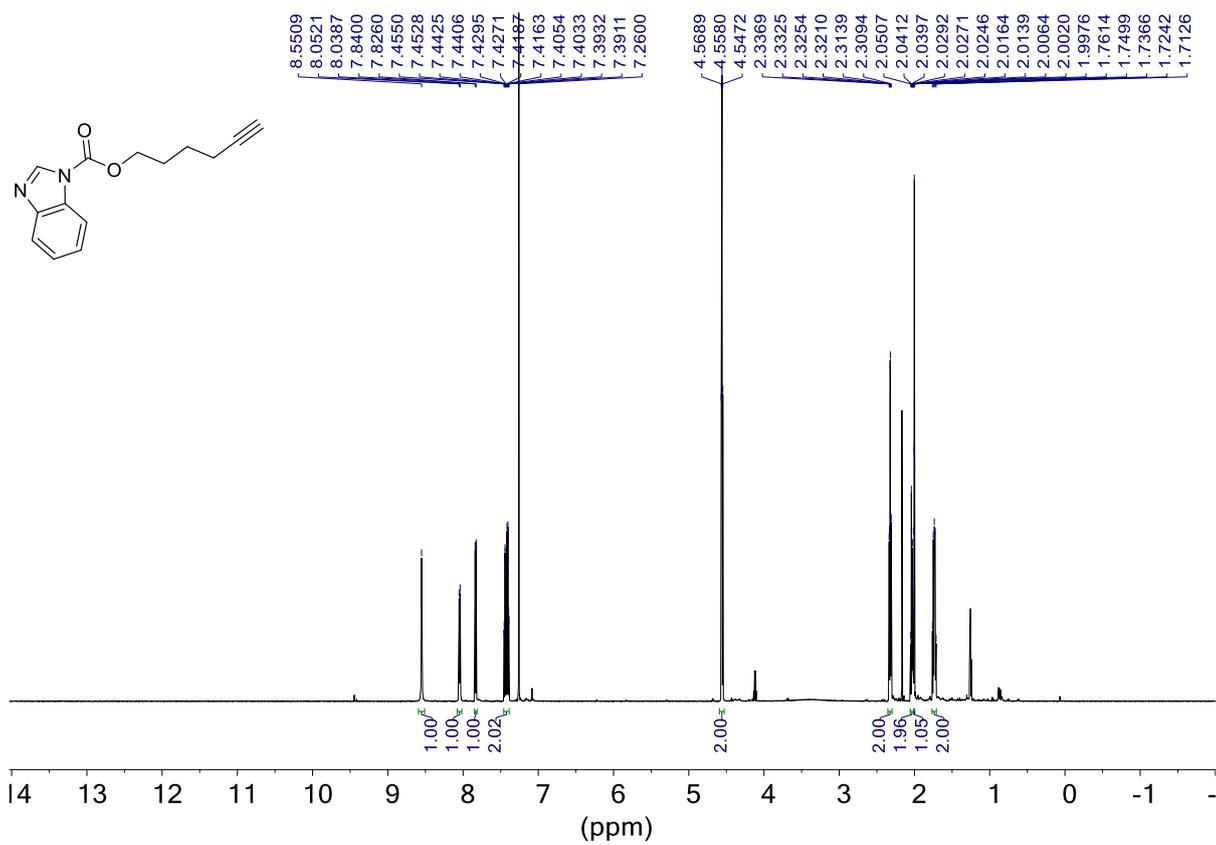
## Dipropargyl pyrocarbonate (52)



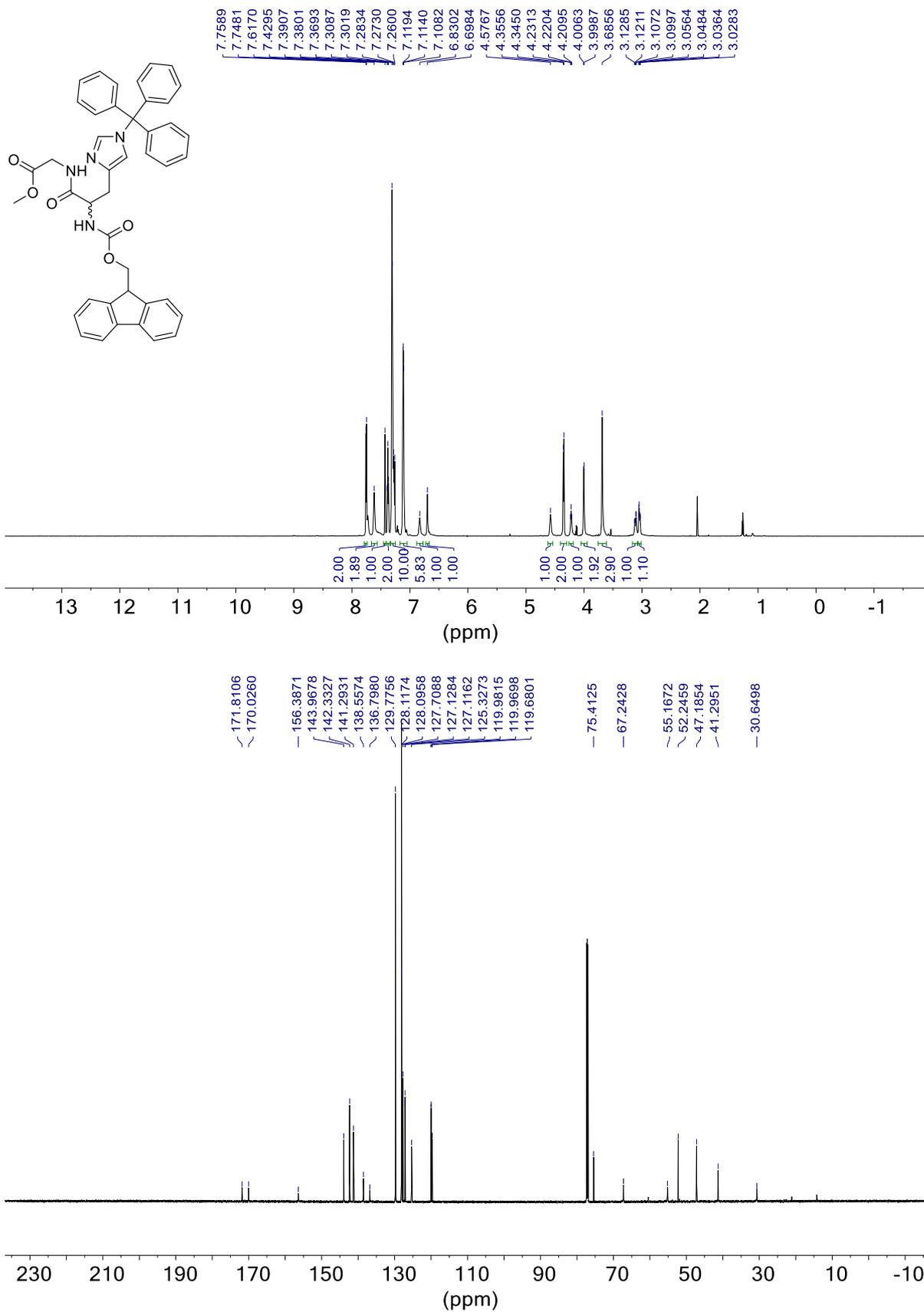
Ethyl 1H-1,3-benzodiazole-1<sup>l</sup>-carboxylate (60)

## Bis(hex-5-yn-1-yl) dicarbonate (66)

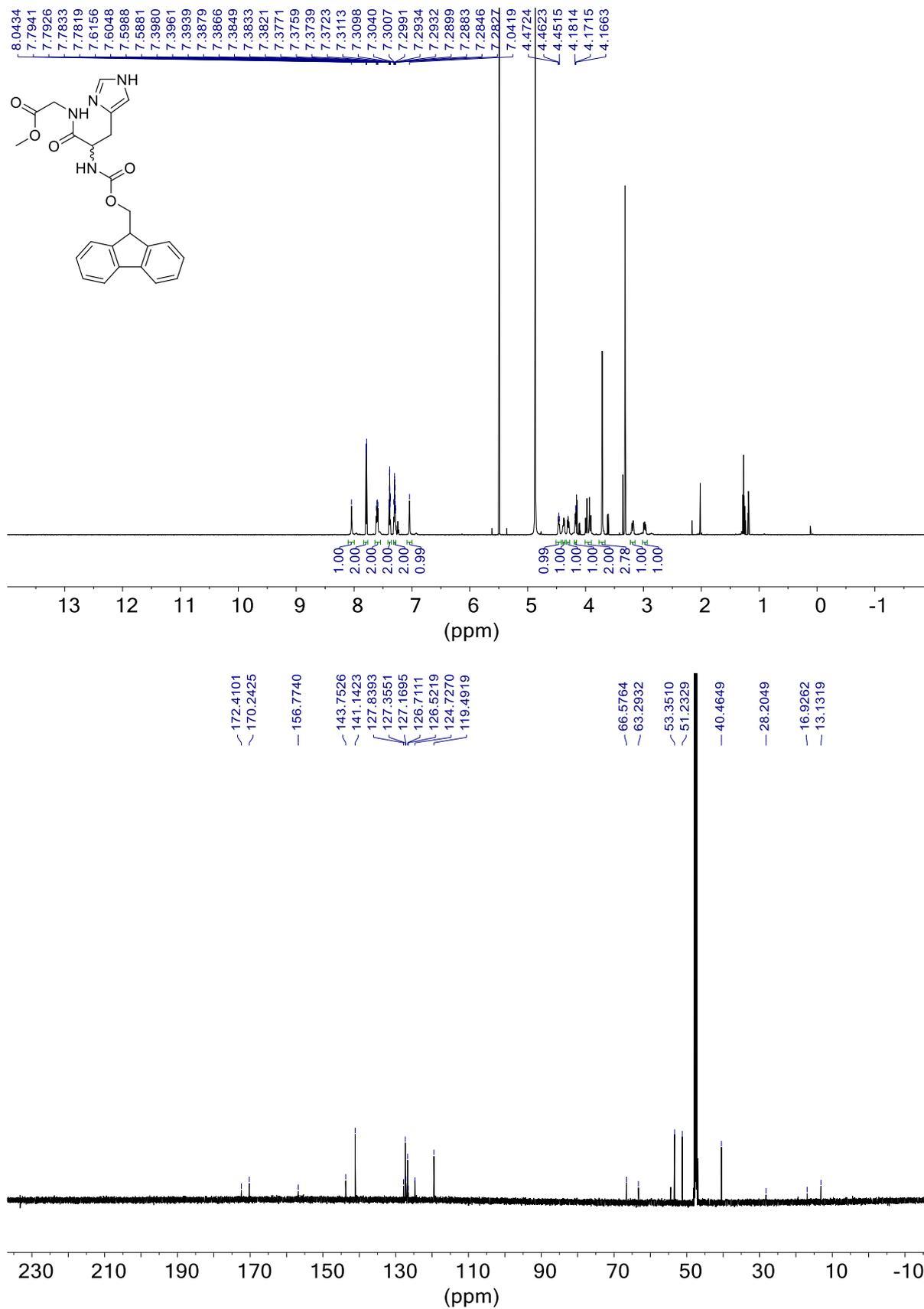


Hex-5<sup>l</sup>-yn-1<sup>l</sup>-yl 1H-1,3-benzodiazole-1-carboxylate (67)

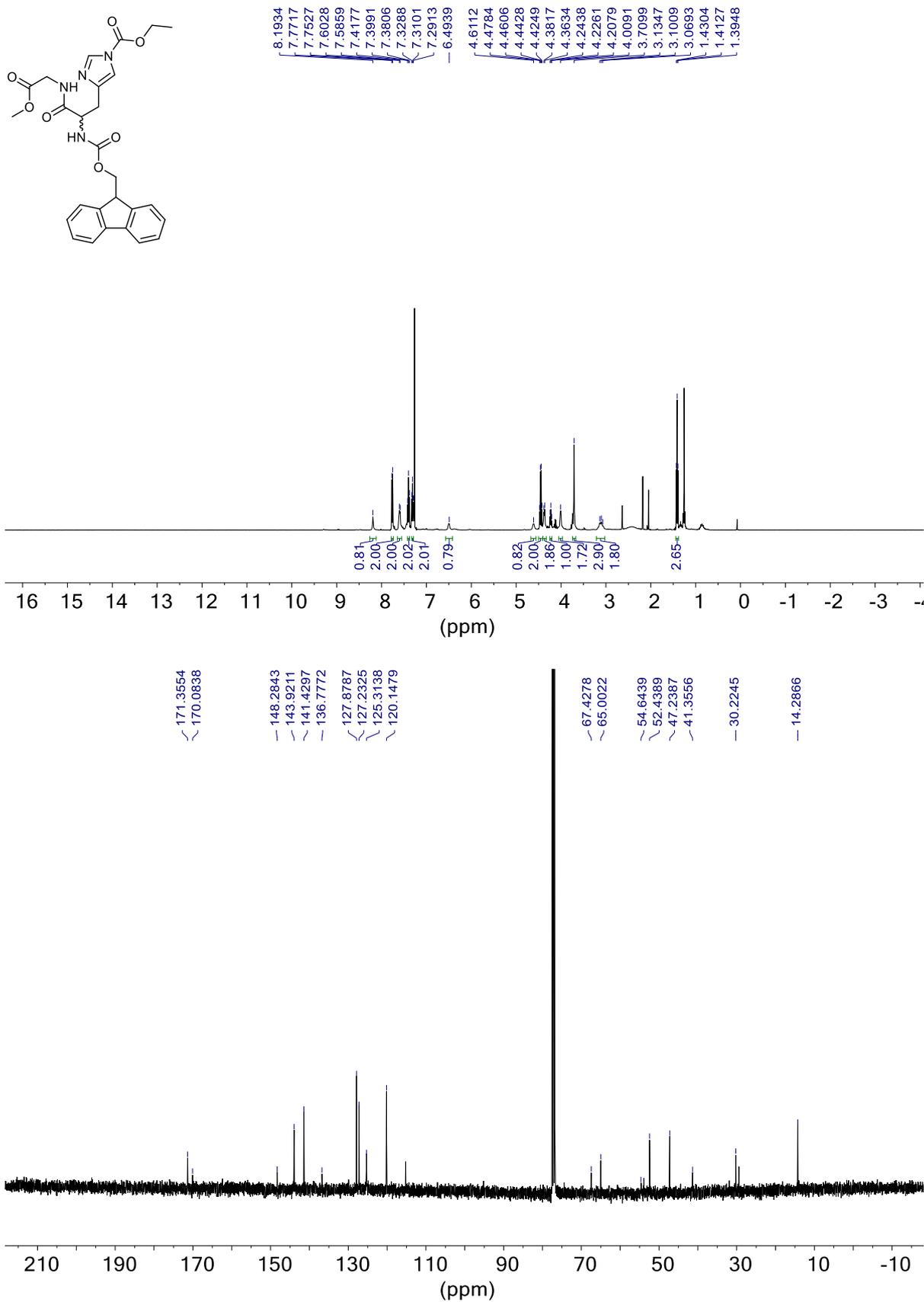
**Methyl 2-[2'-{[(9H-fluoren-9-yl)methoxy]carbonyl}amino]-3'-[1-(triphenylmethyl)-1H-imidazol-4-yl]propanamido]acetate (70)**



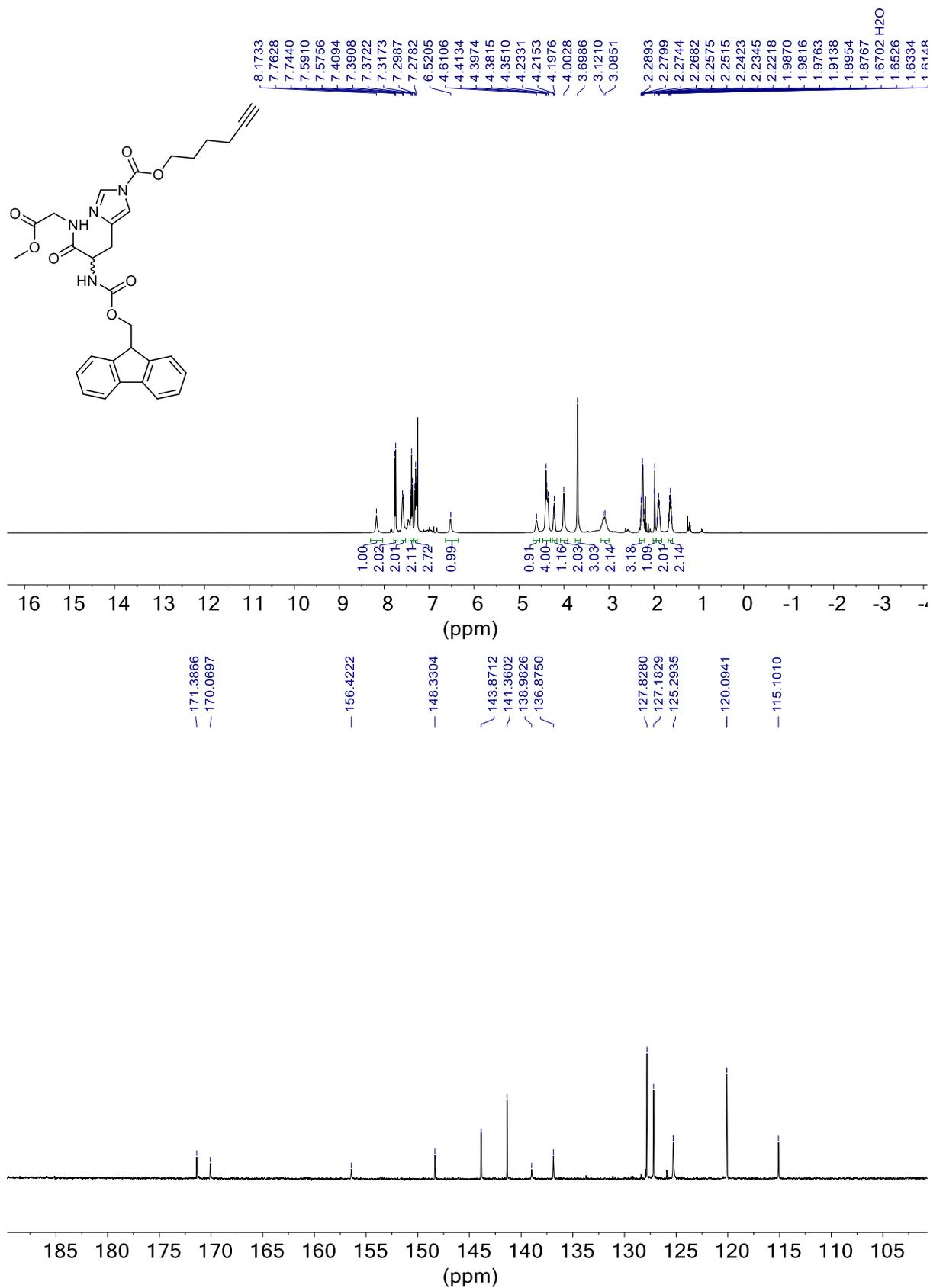
Methyl 2-[2'-{[(9H-fluoren-9-yl)methoxy]carbonyl}amino]-3'-(1H-imidazol-4-yl)propanamido]acetate (71)



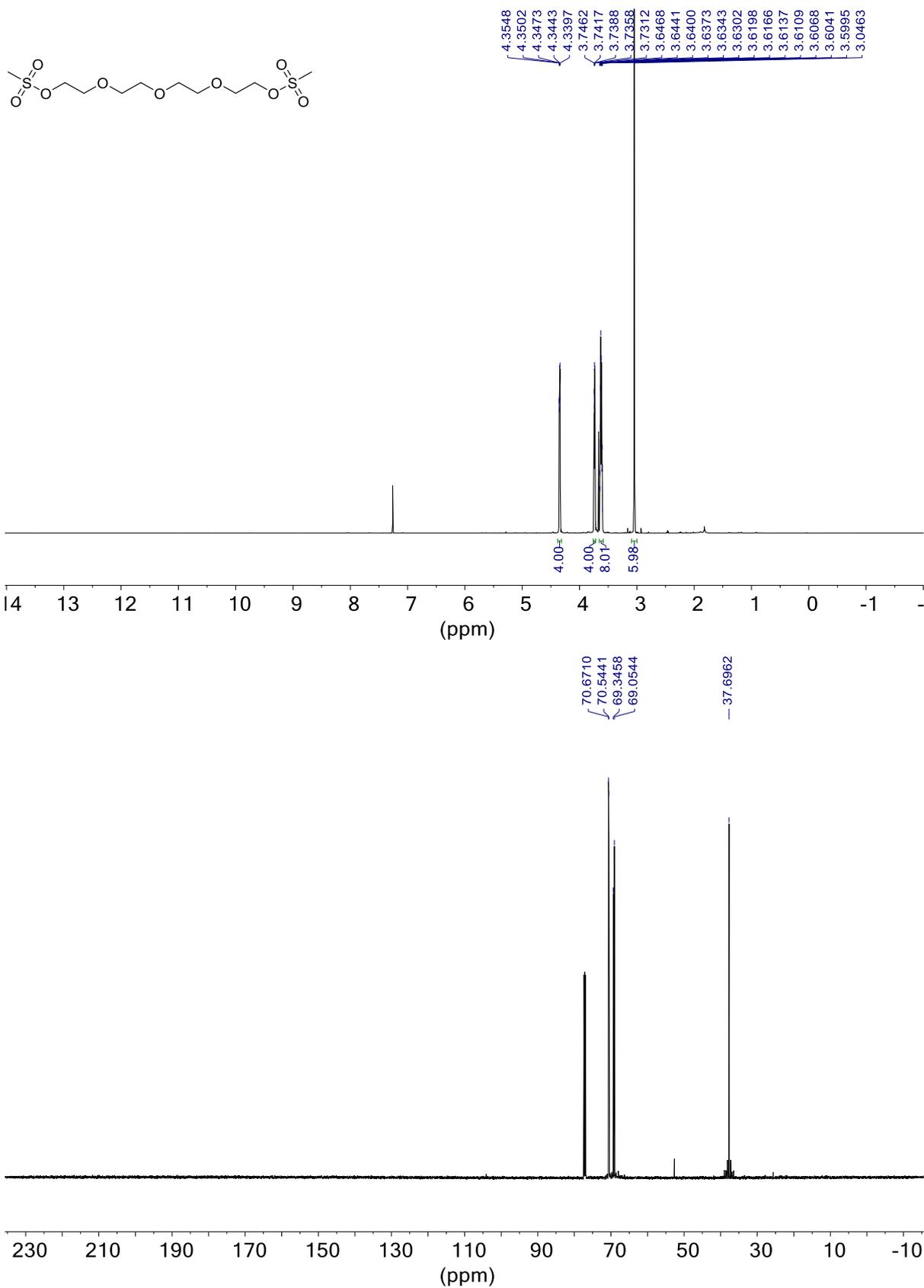
**Methyl 2-[2'-((9H-fluoren-9-yl)methoxy)carbonyl]amino-3'-[(1-ethoxycarbony)-1H-imidazol-4-yl]]propanamido]acetate (73)**

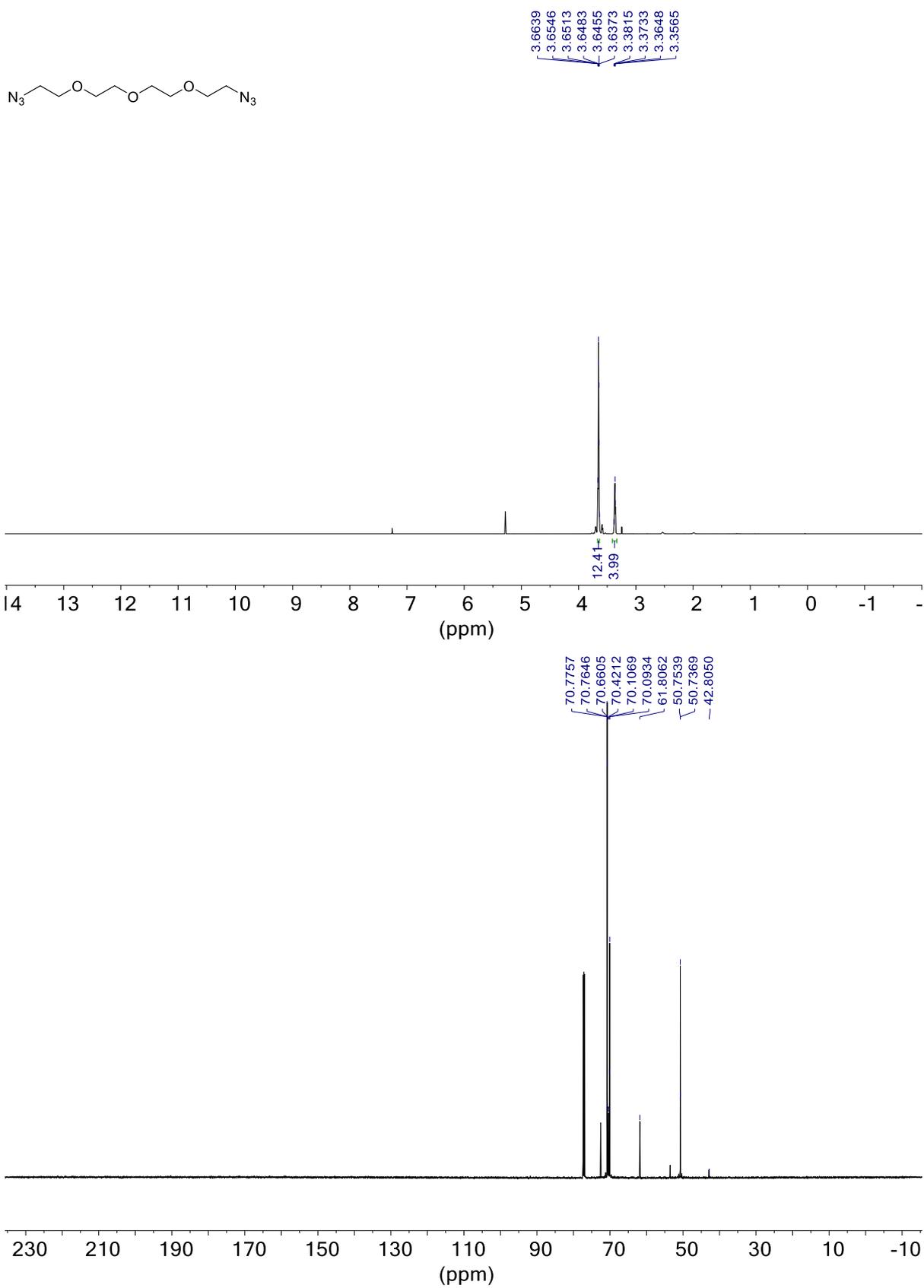


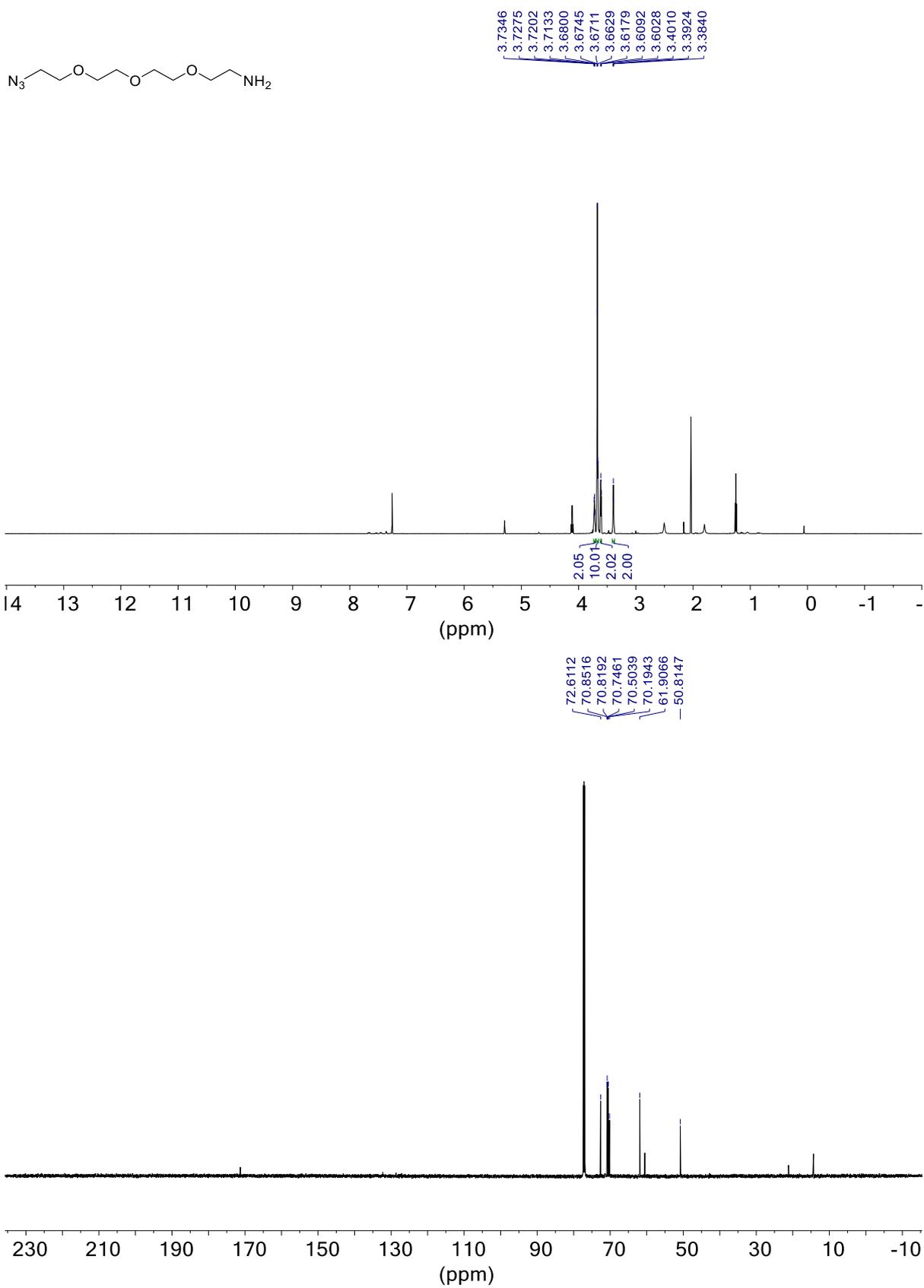
**Methyl 2-[2'-{[(9H-fluoren-9-yl)methoxy]carbonyl}amino]-3'-[(1-hex-5<sup>IV</sup>-ynoxycarboxy)-1H-imidazol-4<sup>III</sup>-yl]]propanamido]acetate (72)**



## 2-(2-{2-[2-(methanesulfonyloxy)ethoxy]ethoxy}ethoxy)ethyl methanesulfonate

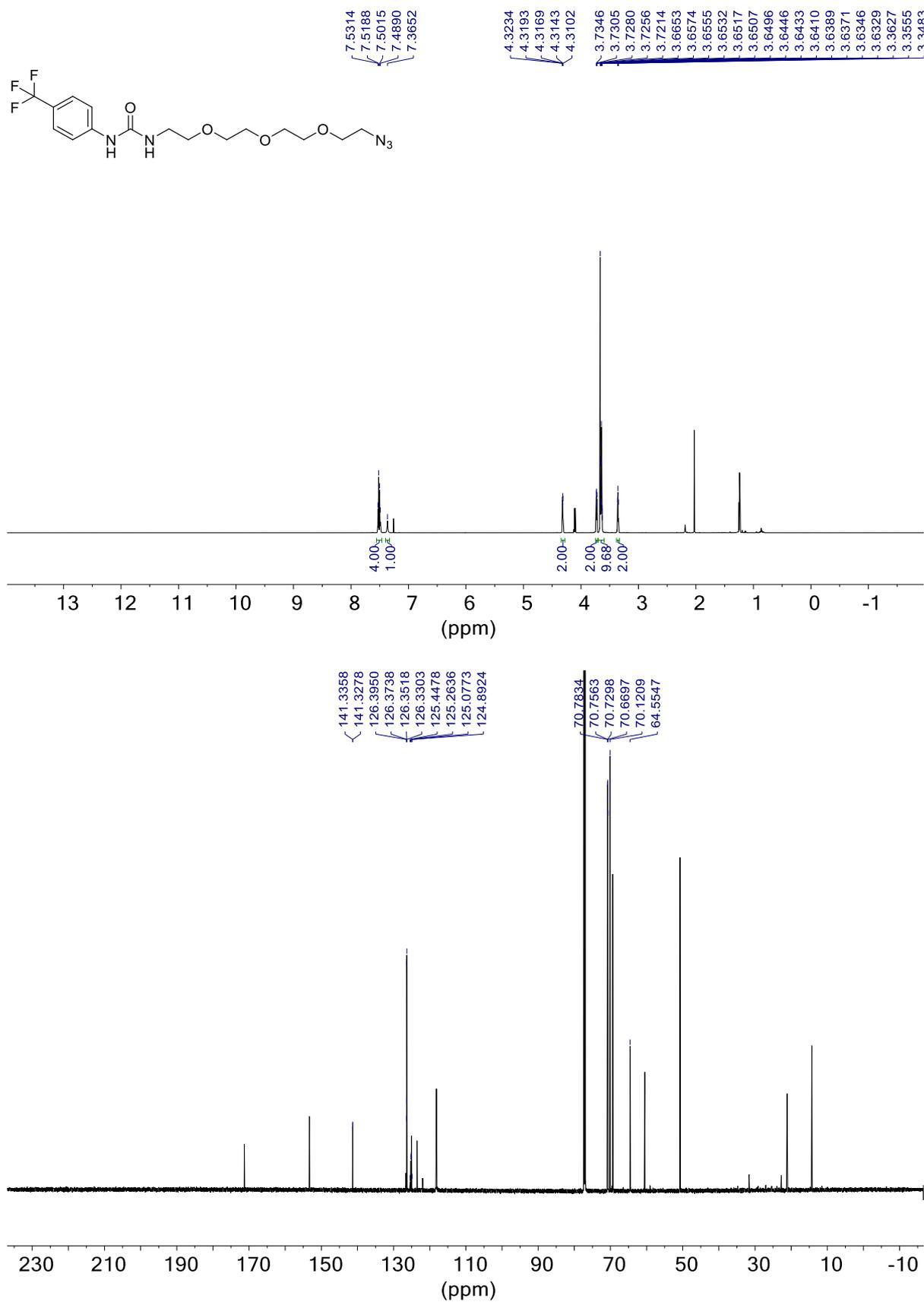


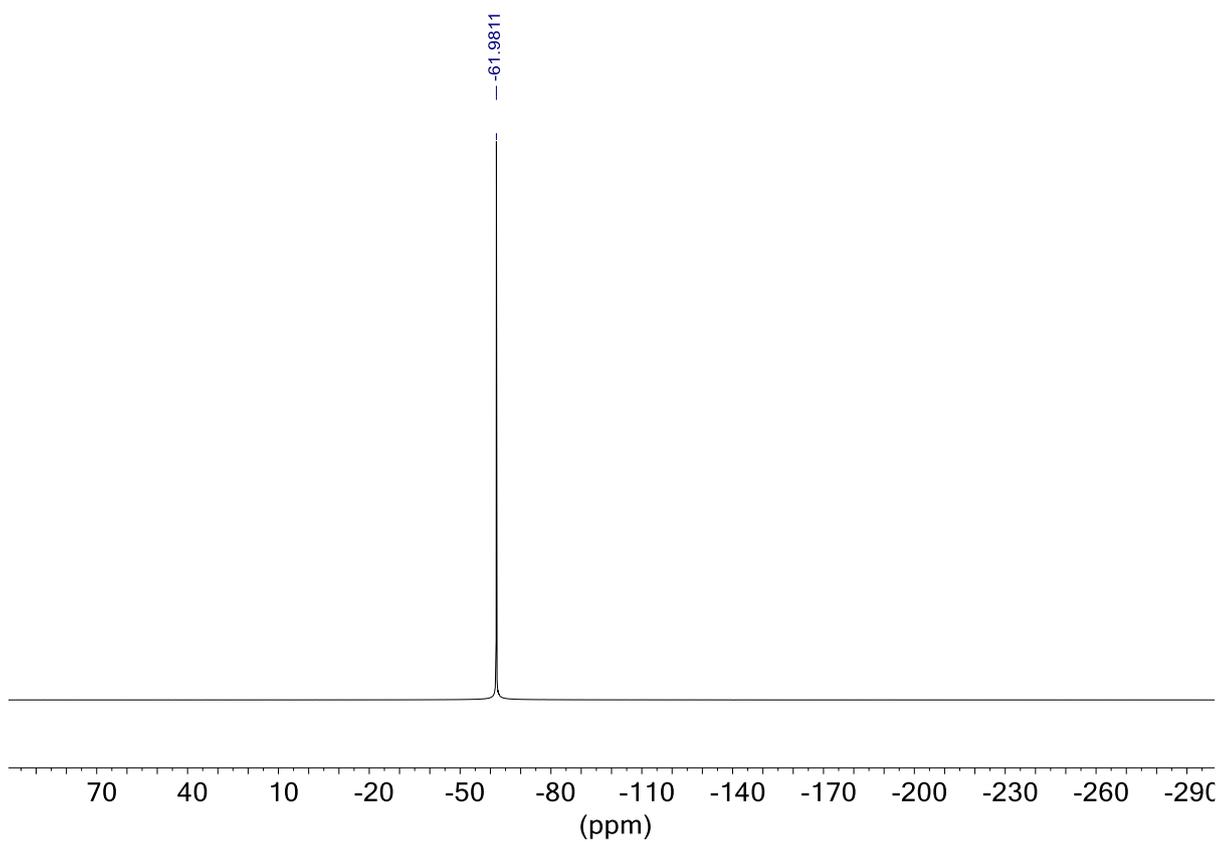
**1-azido-2-{2<sup>I</sup>-[2<sup>II</sup>-[2<sup>III</sup>-azidoethoxy]ethoxy]ethoxy}ethane (75)**

**1<sup>I</sup>-[2-(2-aminoethoxy)ethoxy]-2<sup>I</sup>-(2<sup>II</sup>-azidoethoxy)ethane (76)**

## 3-(2'-{2-[2-(2''-azidoethoxy)ethoxy]ethoxy}ethyl)-1-[4'''-(trifluoromethyl)phenyl]urea

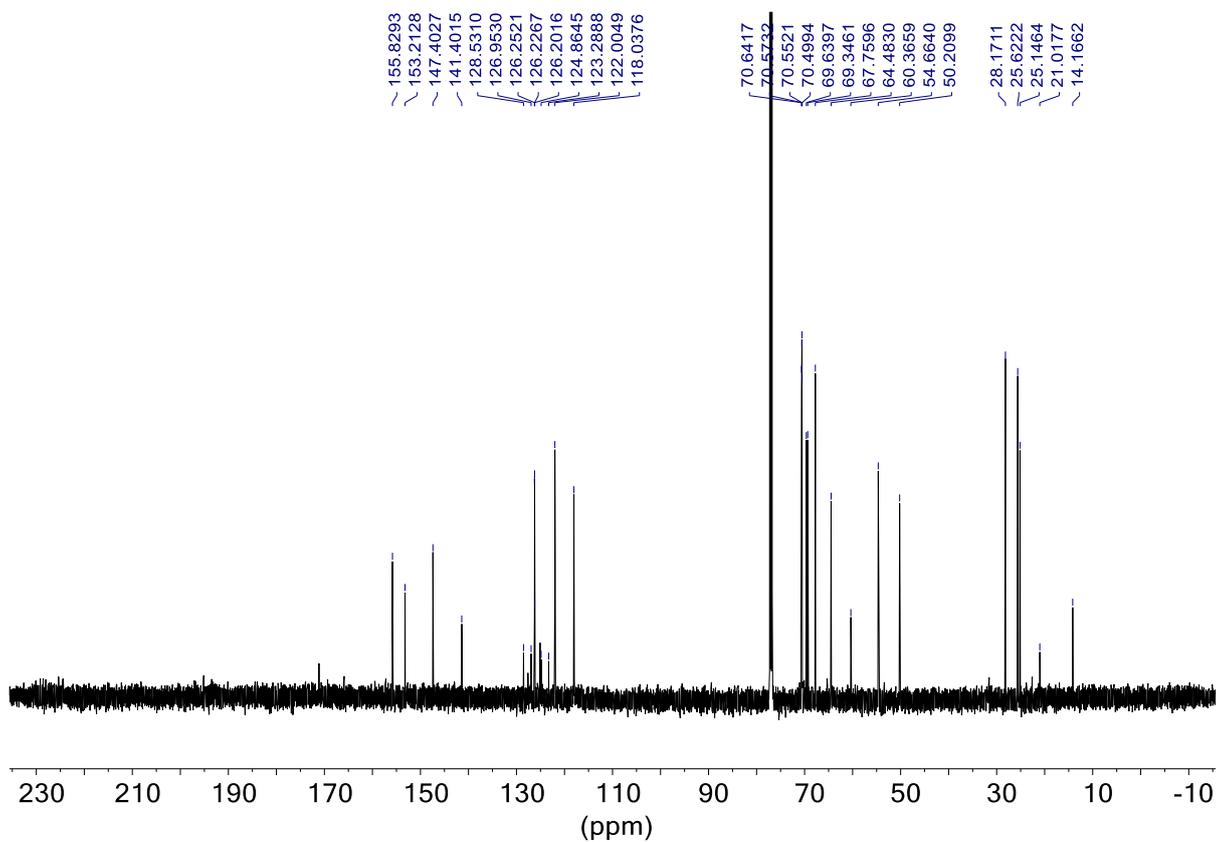
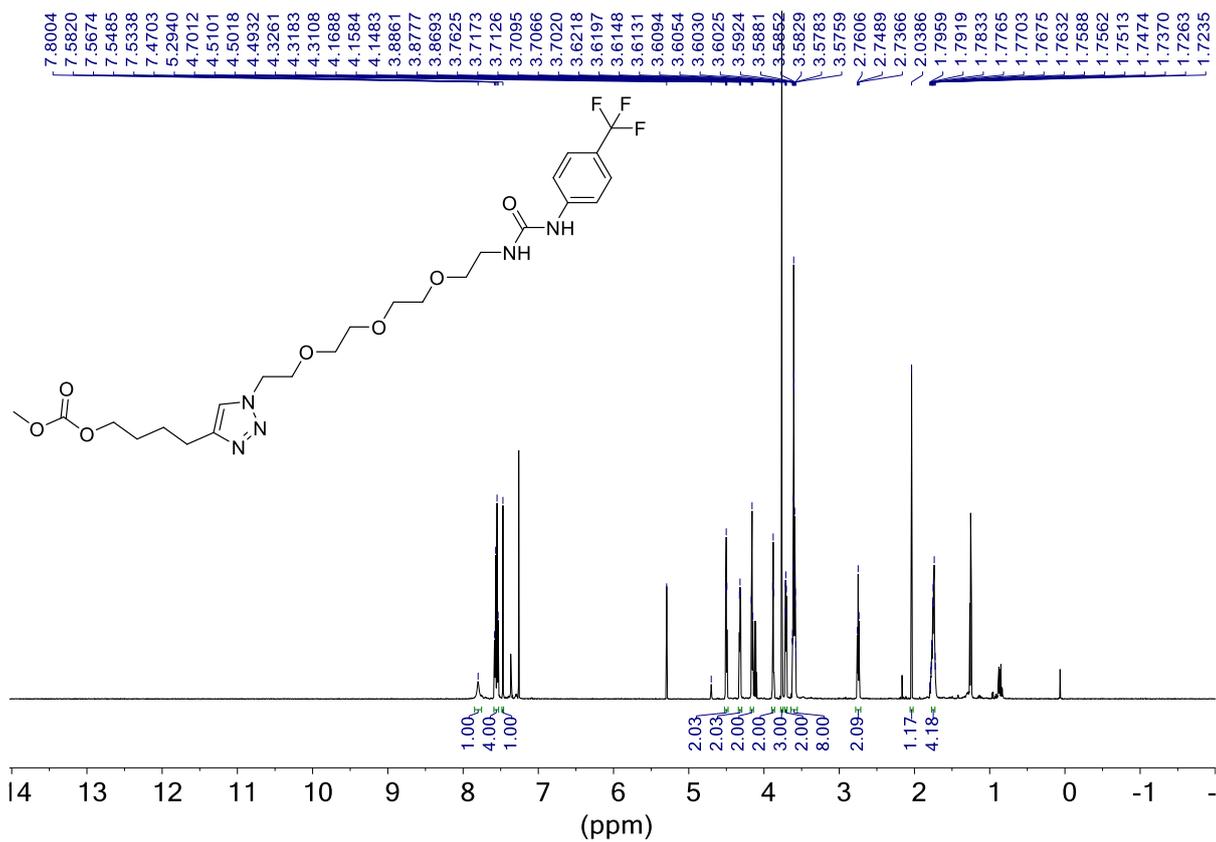
(78)



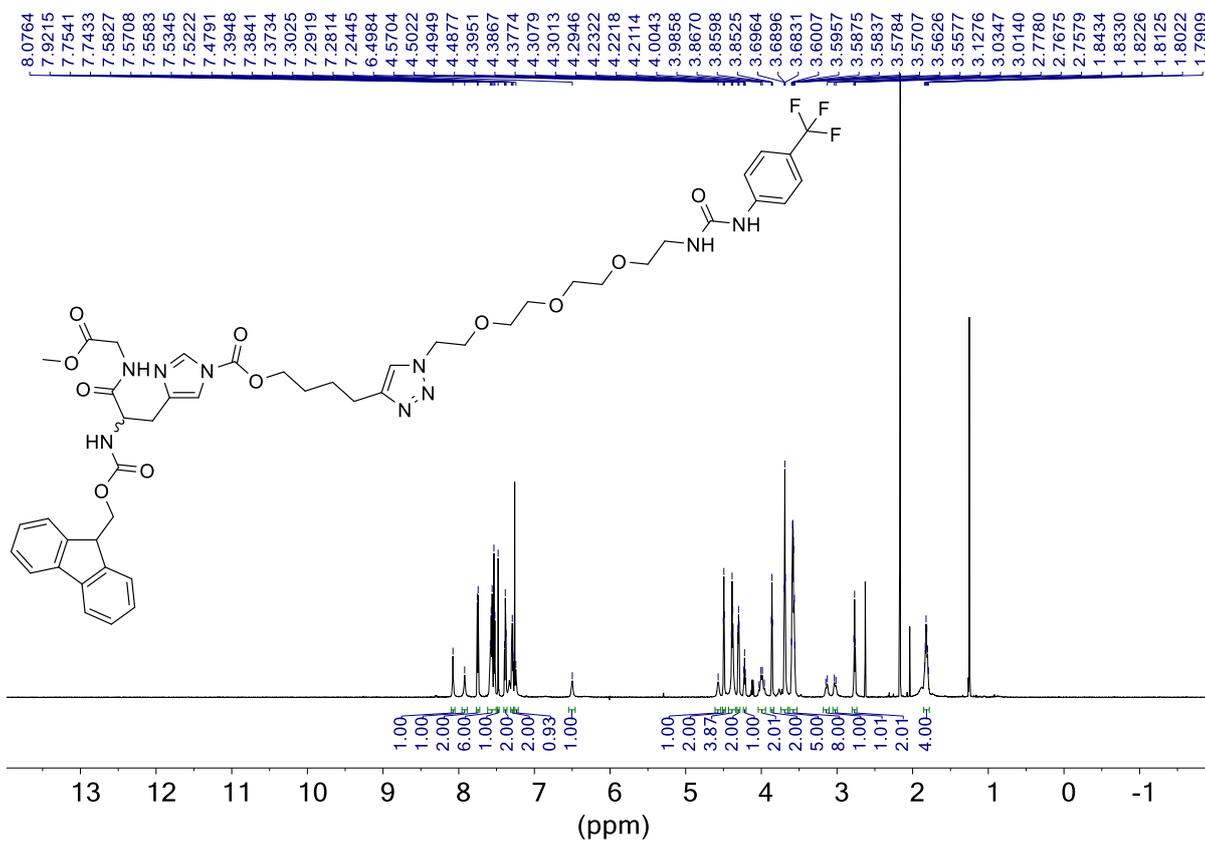


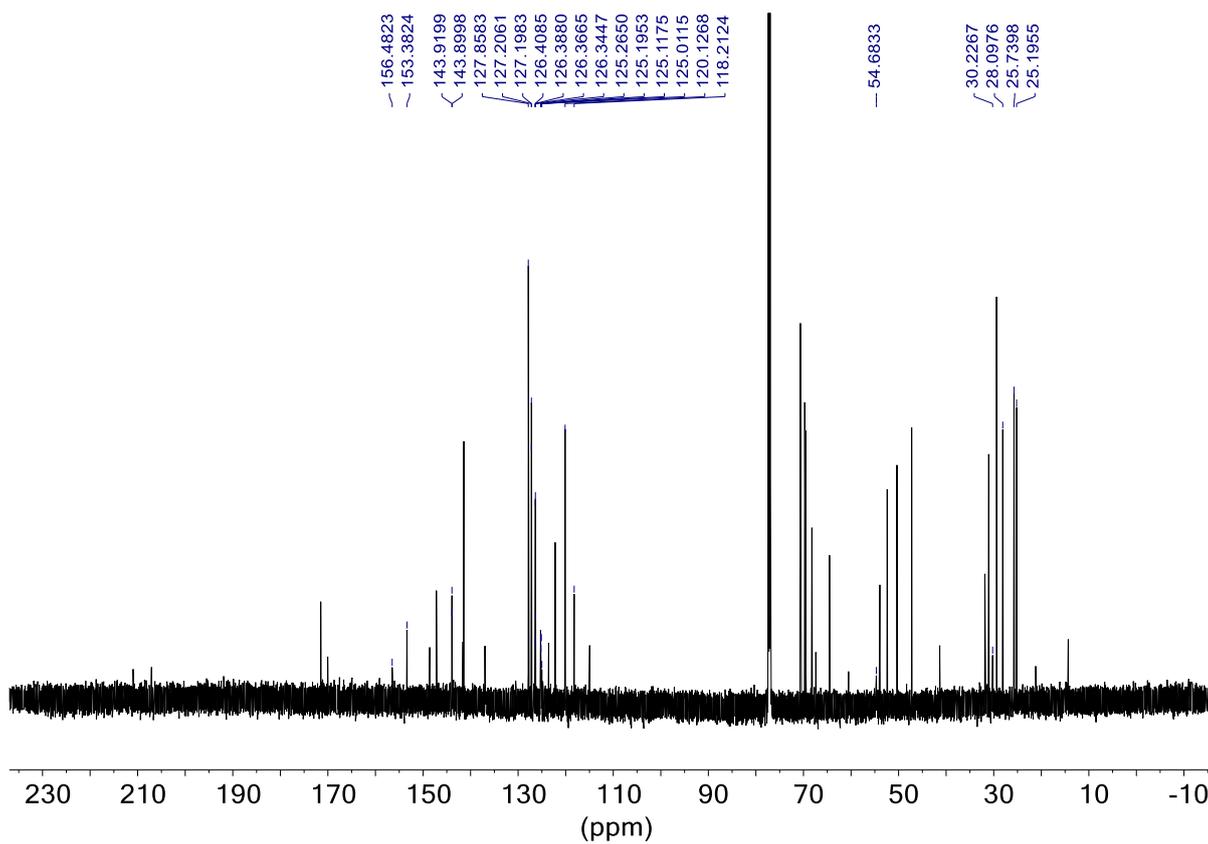
**Methyl**

**4<sup>III</sup>-{1-[2-(2-{2-[2-({[4-(trifluoromethyl)phenyl]carbonyl]amino)ethoxy]ethoxy}ethoxy)ethyl]-1H-1,2,3-triazol-4'-yl]}butyl carbonate (80)**

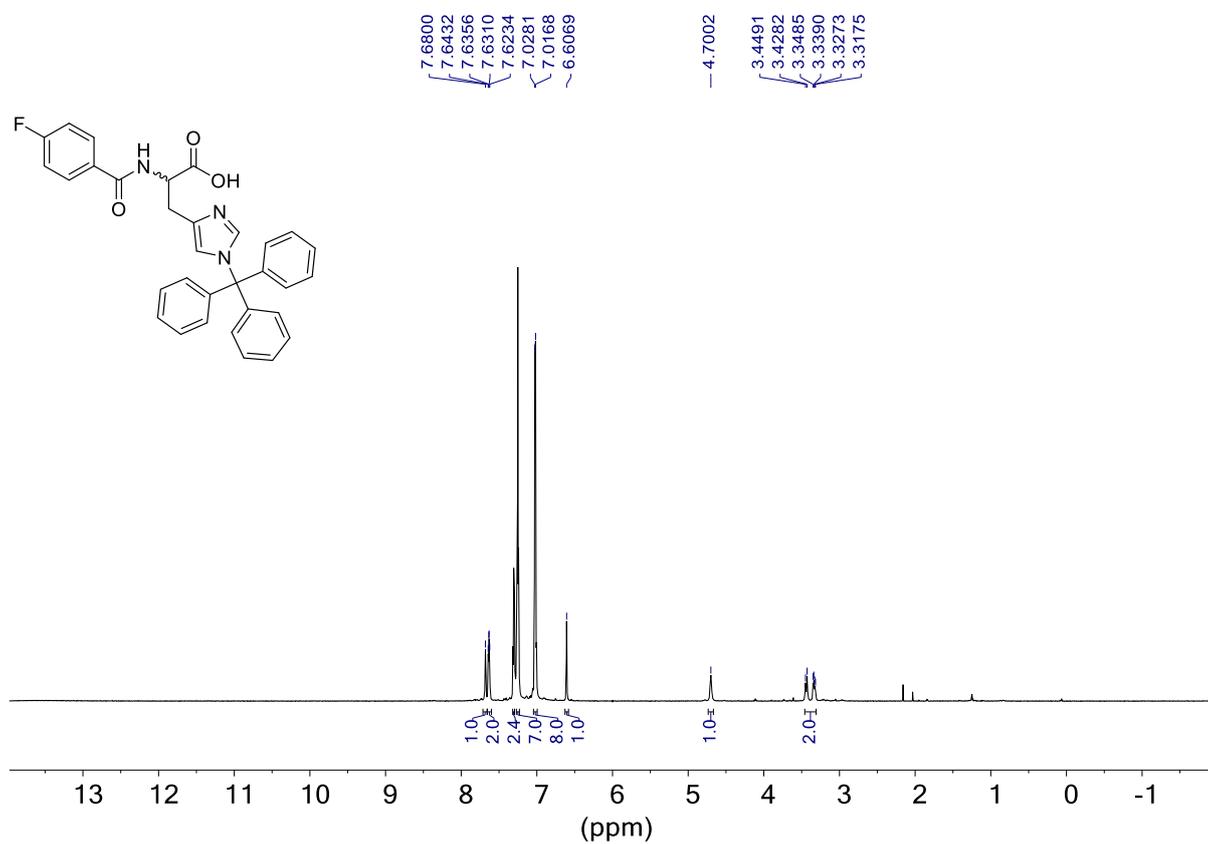


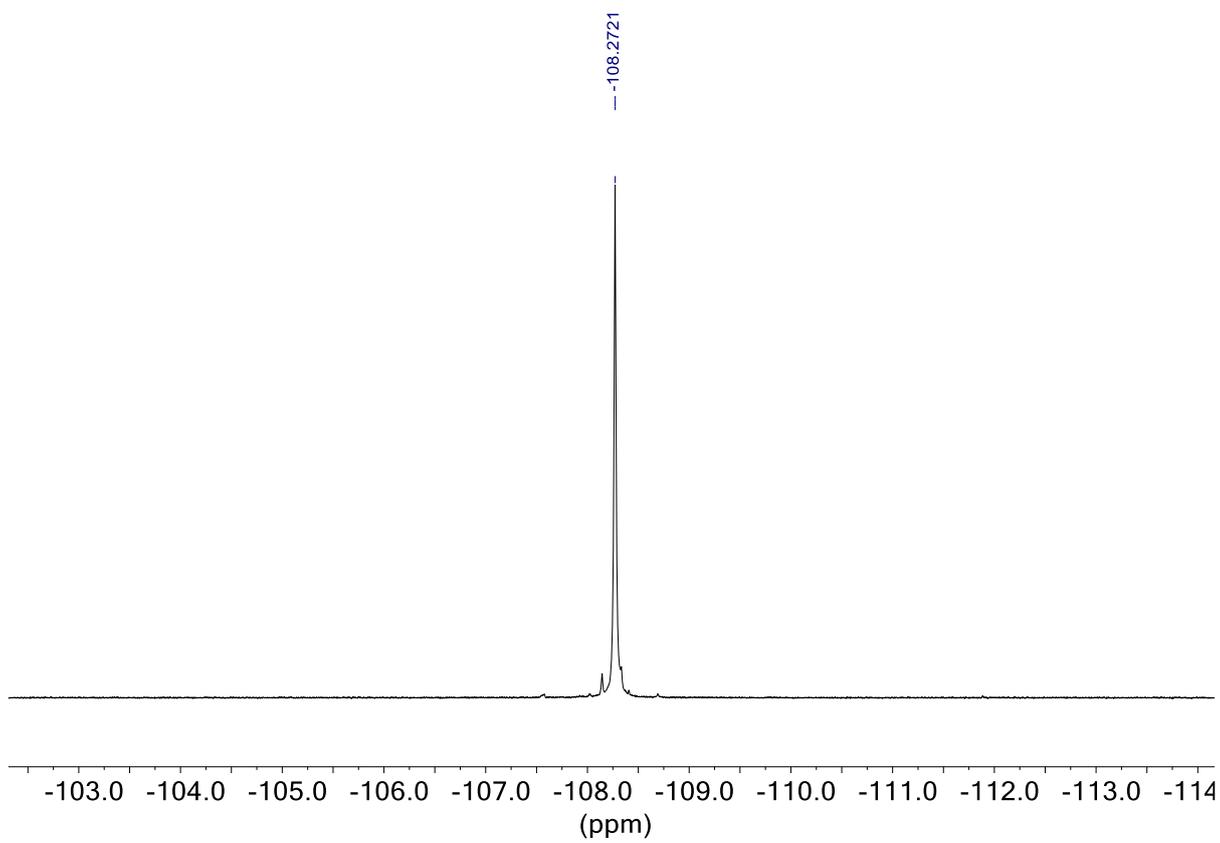
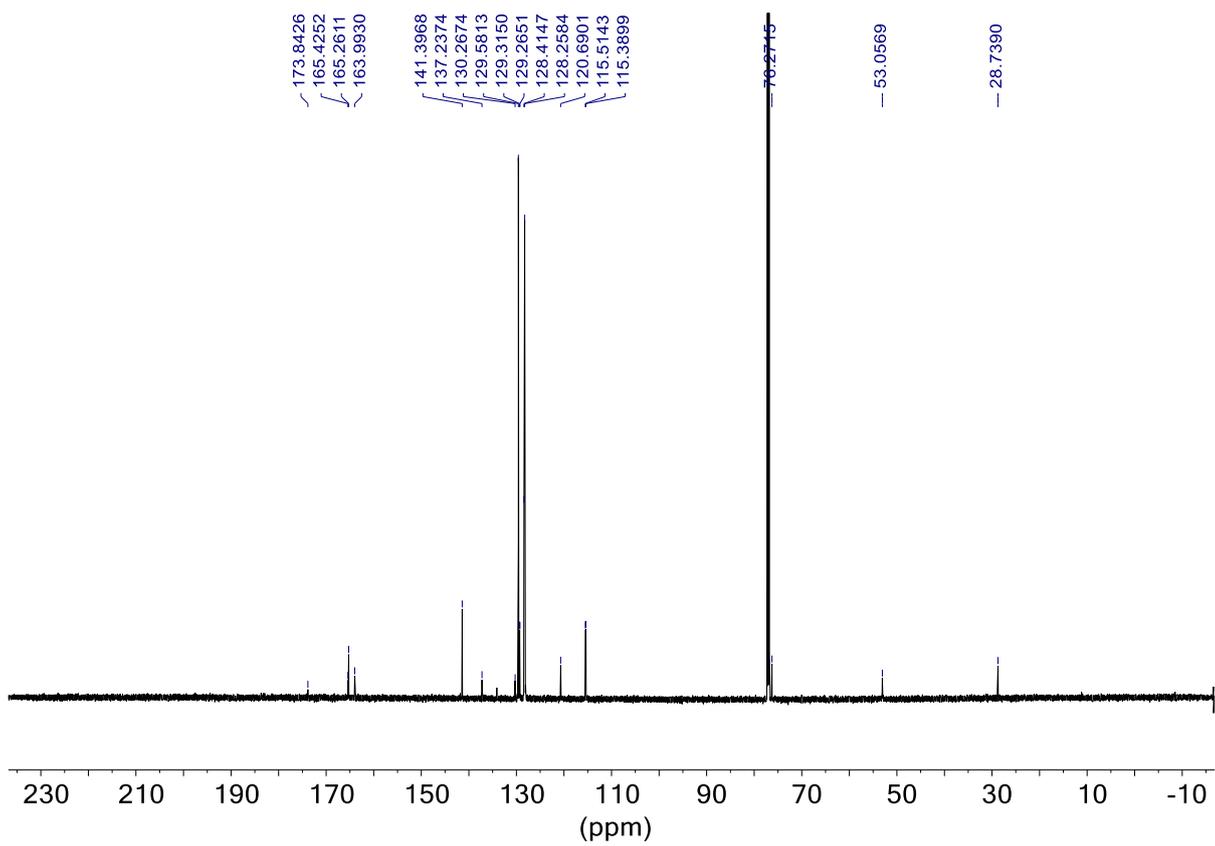
**4<sup>II</sup>-{1<sup>II</sup>-[2<sup>III</sup>-(2<sup>IV</sup>-{2<sup>V</sup>-[2<sup>VI</sup>-{[4<sup>VII</sup>-(trifluoromethyl)phenyl]carbamoyl}amino)ethoxy]ethoxy}ethoxy)ethyl]-1H-1<sup>II</sup>,2<sup>II</sup>,3<sup>II</sup>-triazol-4<sup>II</sup>-yl]butyl 4-[2<sup>VIII</sup>-{[(9H-fluoren-9<sup>X</sup>-yl)methoxy]carbonyl}amino]-2<sup>IX</sup>-[(2-methoxy-2-oxoethyl)carbamoyl]ethyl]-1H-imidazole-1-carboxylate (79)**



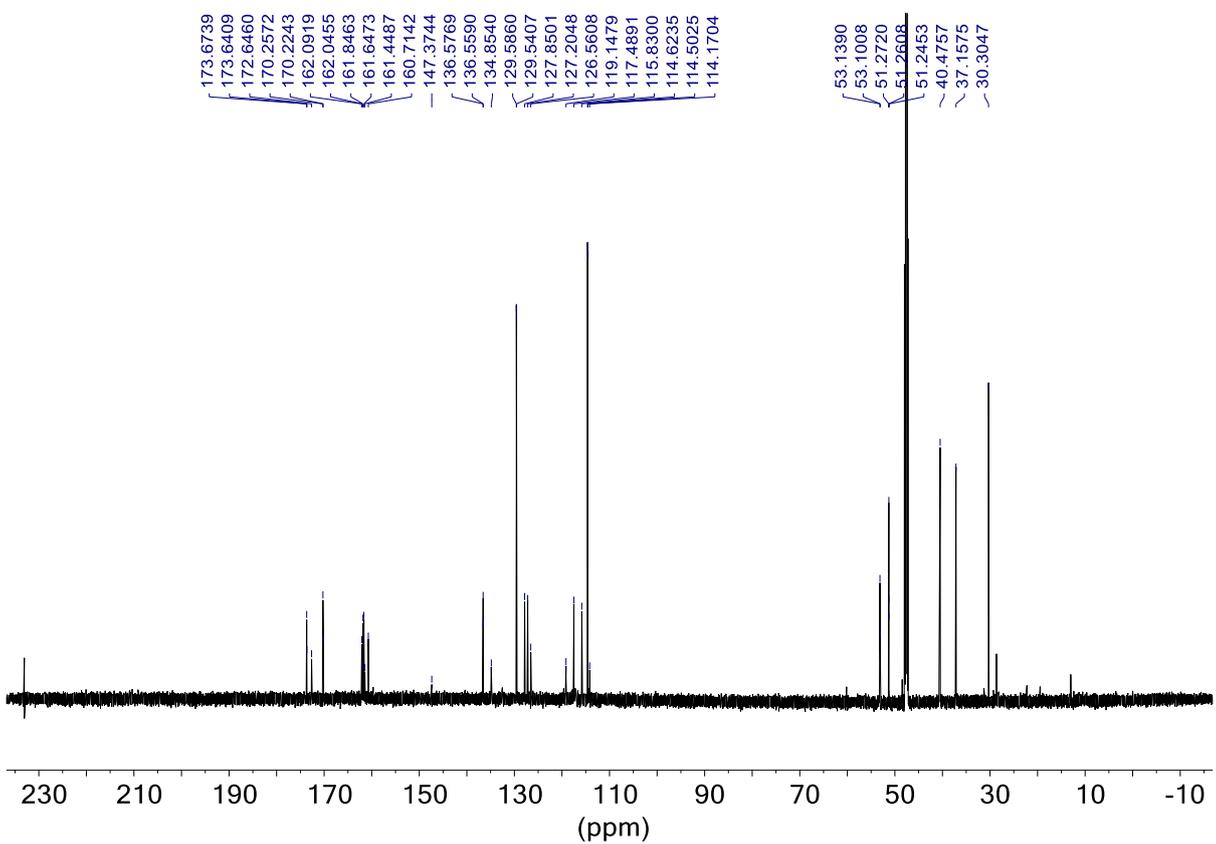
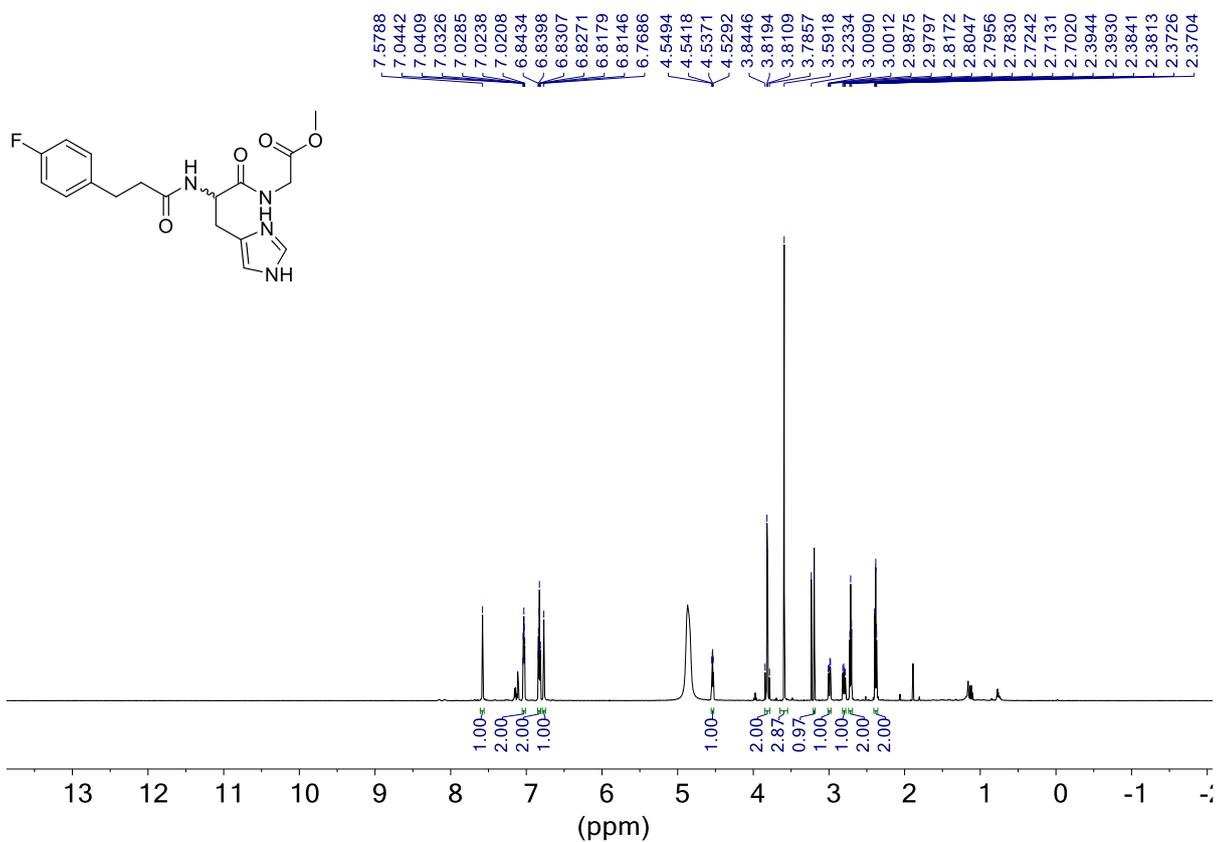


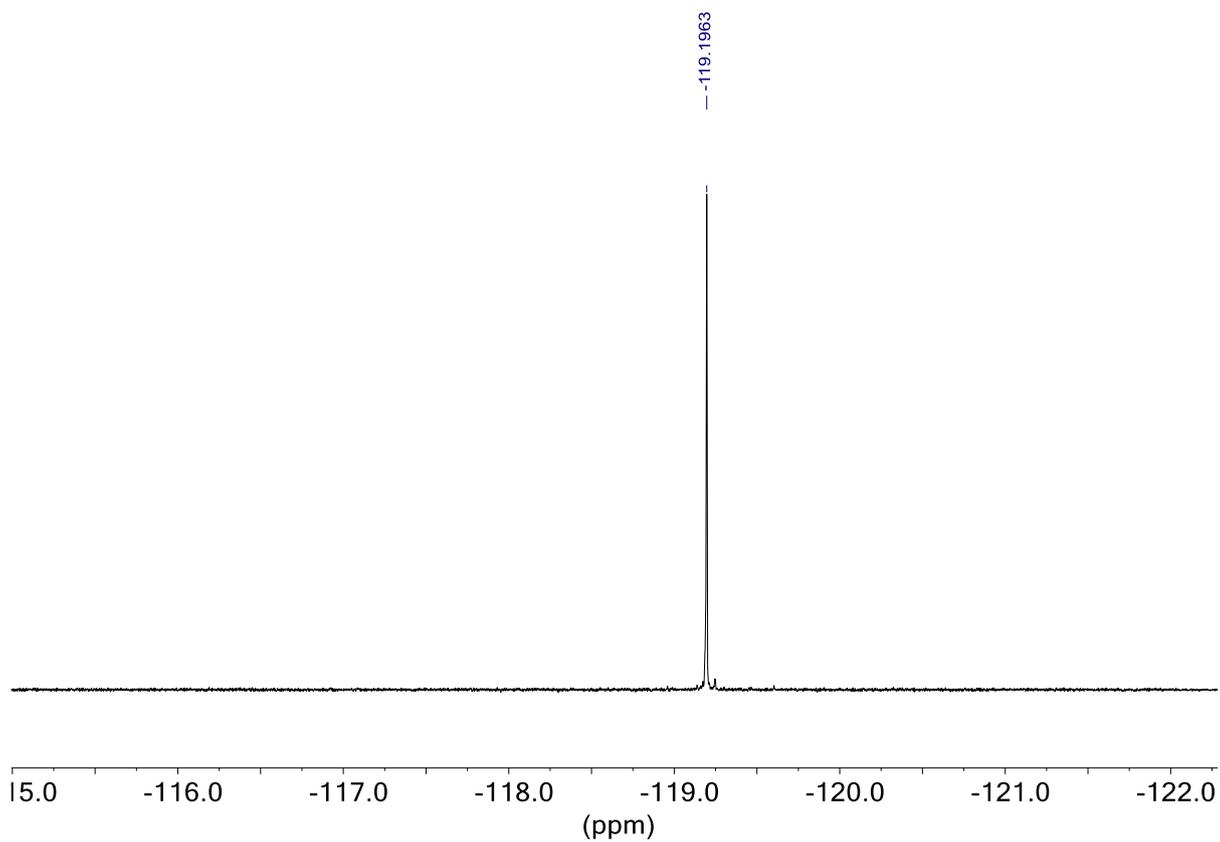
**2<sup>1</sup>-[(4-fluorophenyl)formamido]-3<sup>11</sup>-[1<sup>1</sup>-(triphenylmethyl)-1H-imidazol-4<sup>1</sup>-yl]propanoic acid (83)**



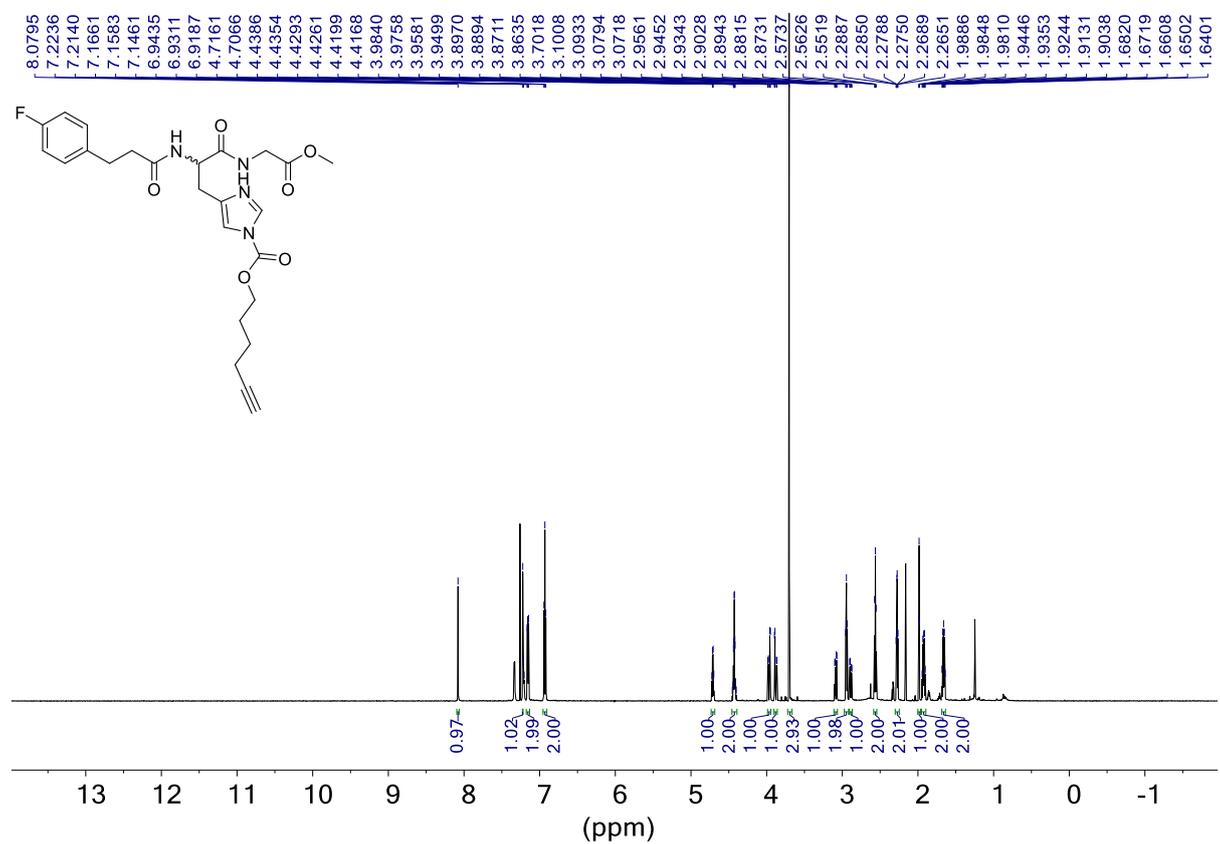


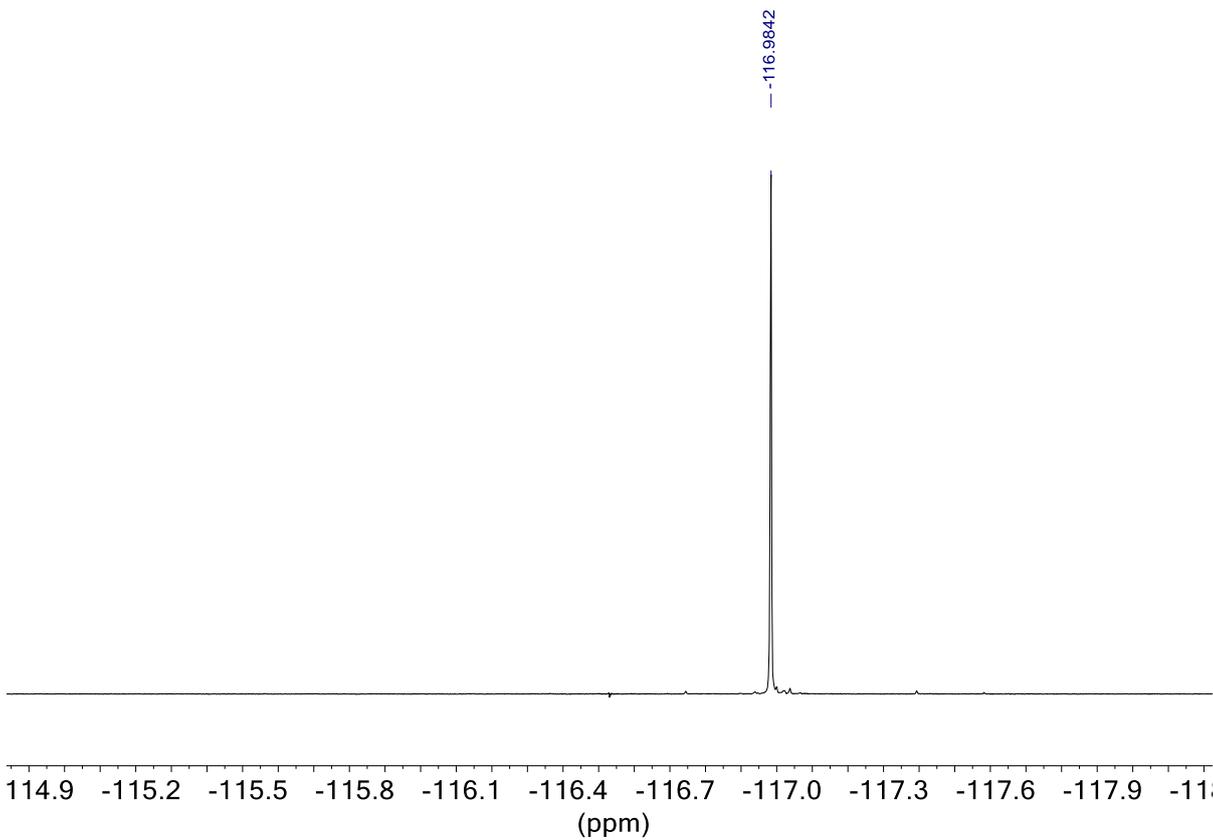
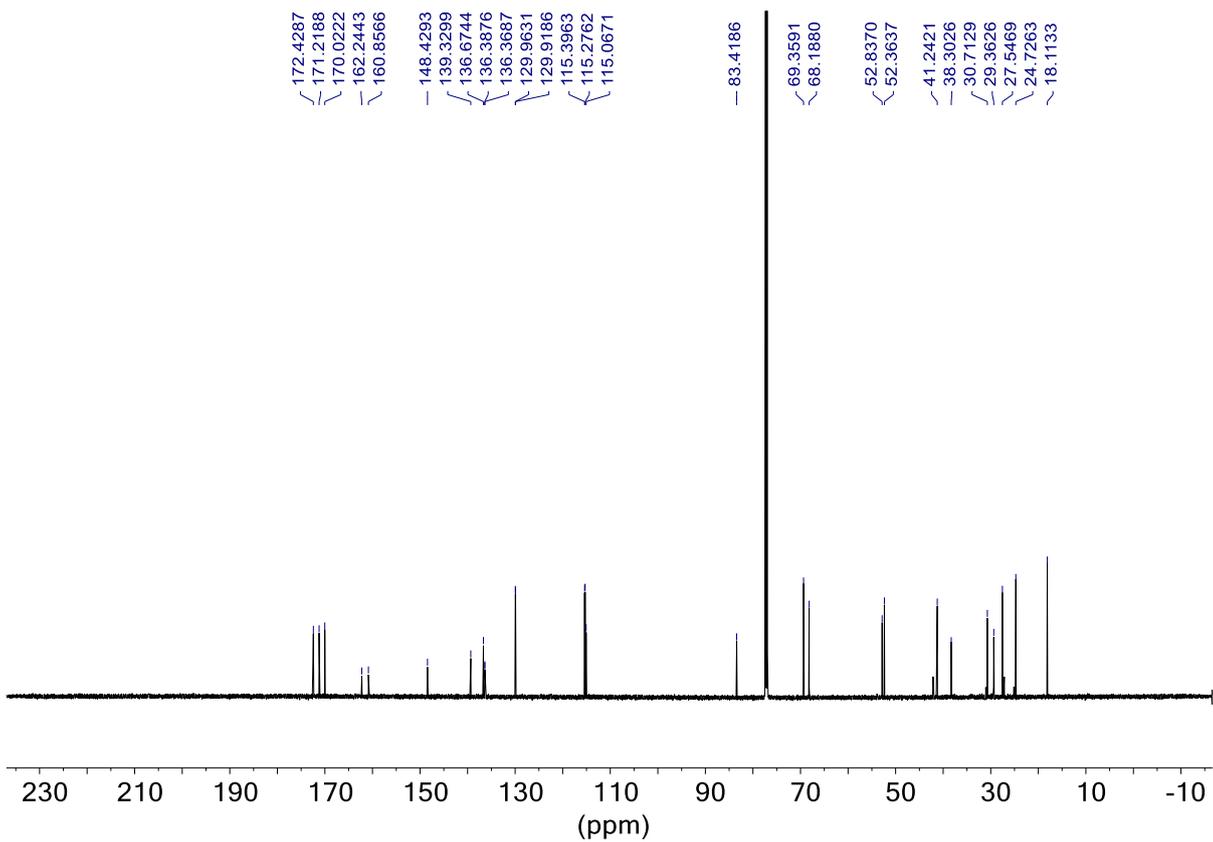
Methyl 2-{2'-[3''-(4'''-fluorophenyl)propanimido]-3'-[(1H-imidazol-4-yl)propanamido]acetate (88)



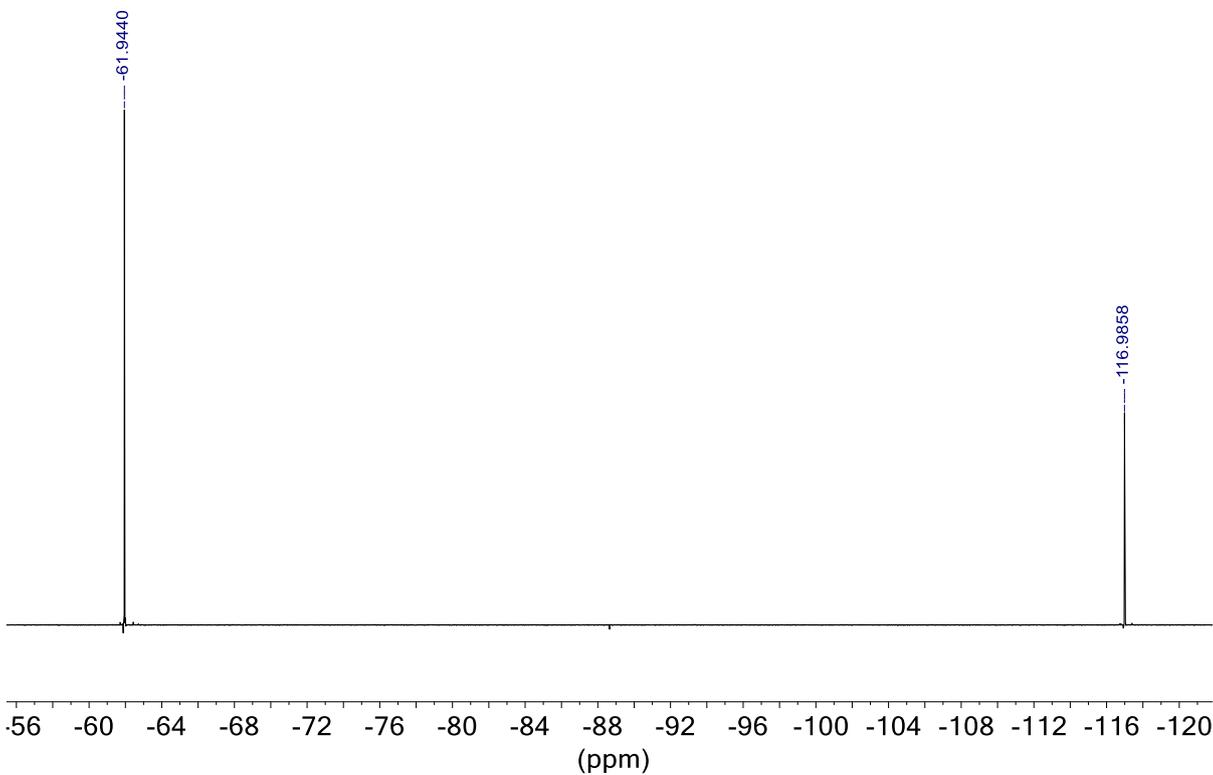
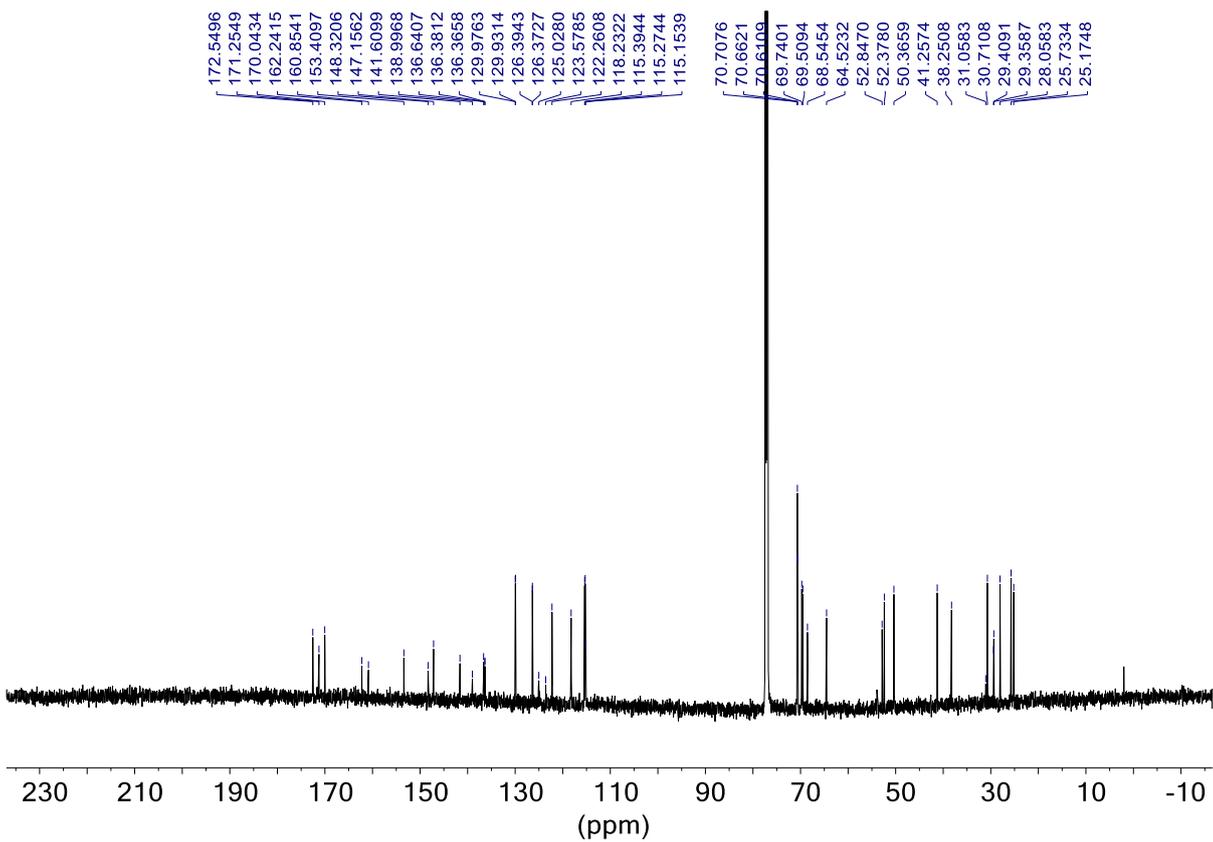


**Hex-5<sup>l</sup>-yn-1<sup>l</sup>-yl 4-{2<sup>ll</sup>-[3<sup>lll</sup>-(4<sup>IV</sup>-fluorophenyl)propanamido]-2<sup>ll</sup>-[(2<sup>V</sup>-methoxy-2<sup>V</sup>-oxoethyl)carbamoyl]ethyl}-1H-imidazole-1-carboxylate (89)**



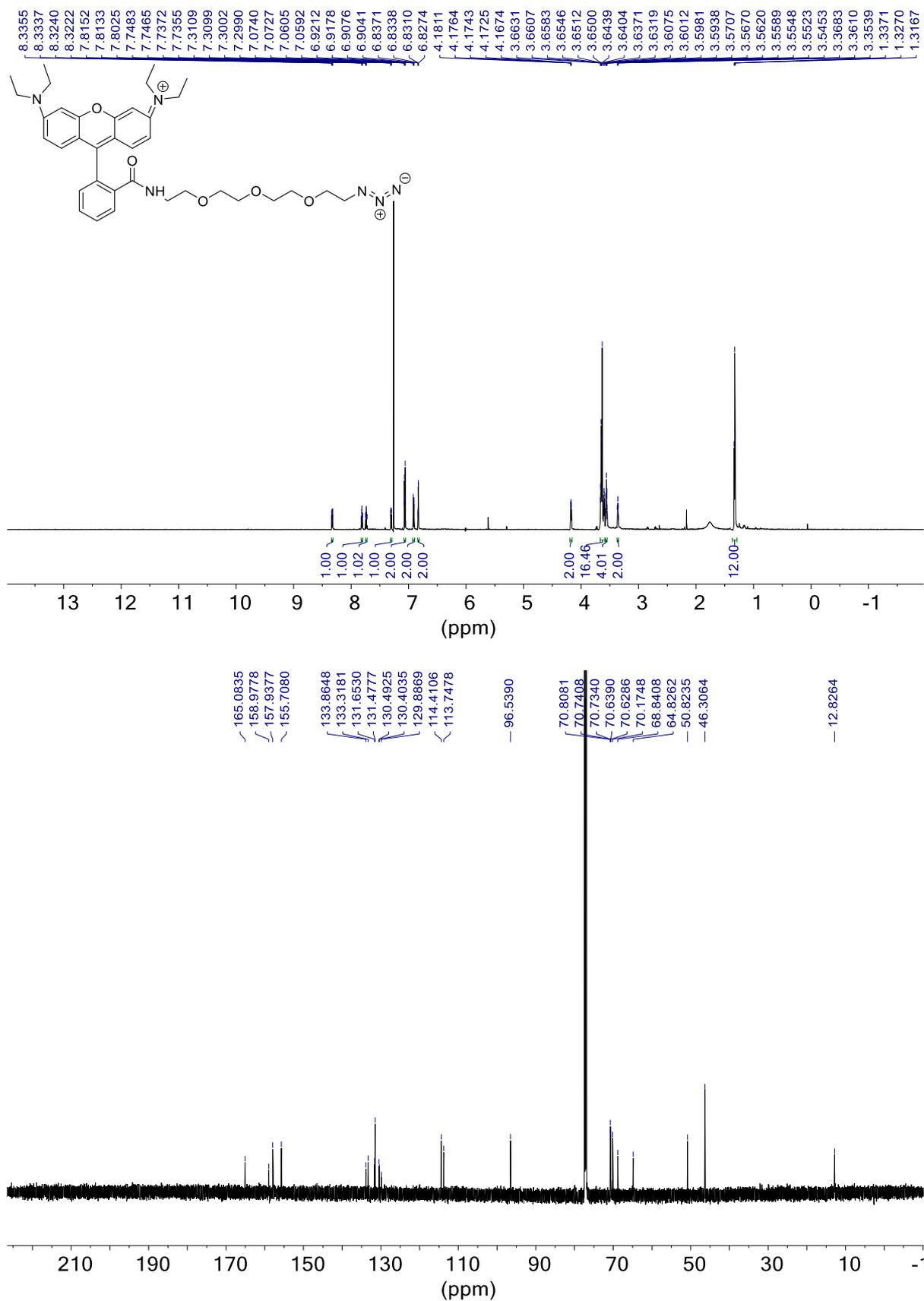




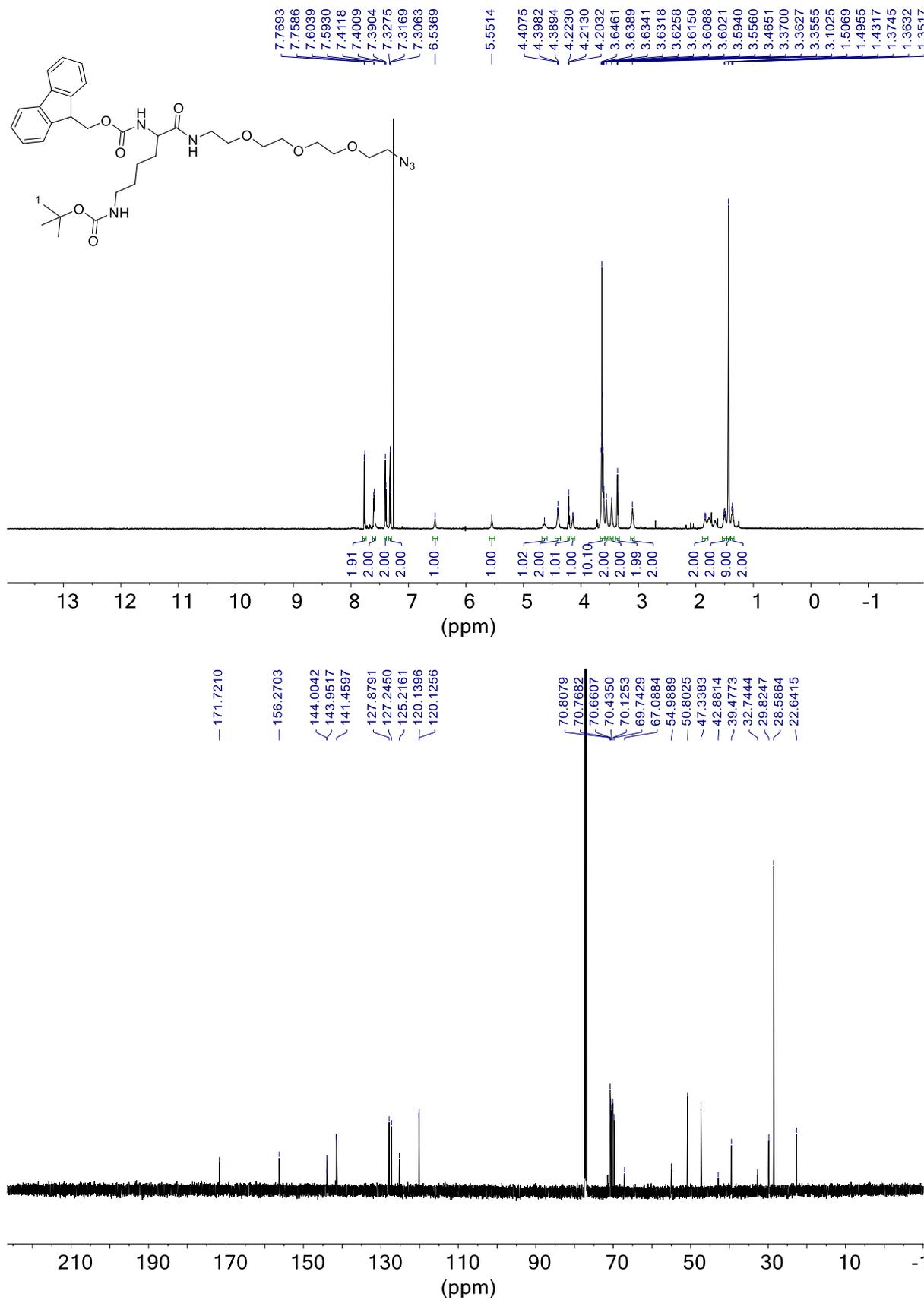


9-{2<sup>I</sup>-[(2<sup>II</sup>-{2<sup>III</sup>-[2<sup>IV</sup>-(2<sup>IV</sup>-azidoethoxy)ethoxy]ethoxy)ethyl]carbamoyl]phenyl}-6<sup>V</sup>-

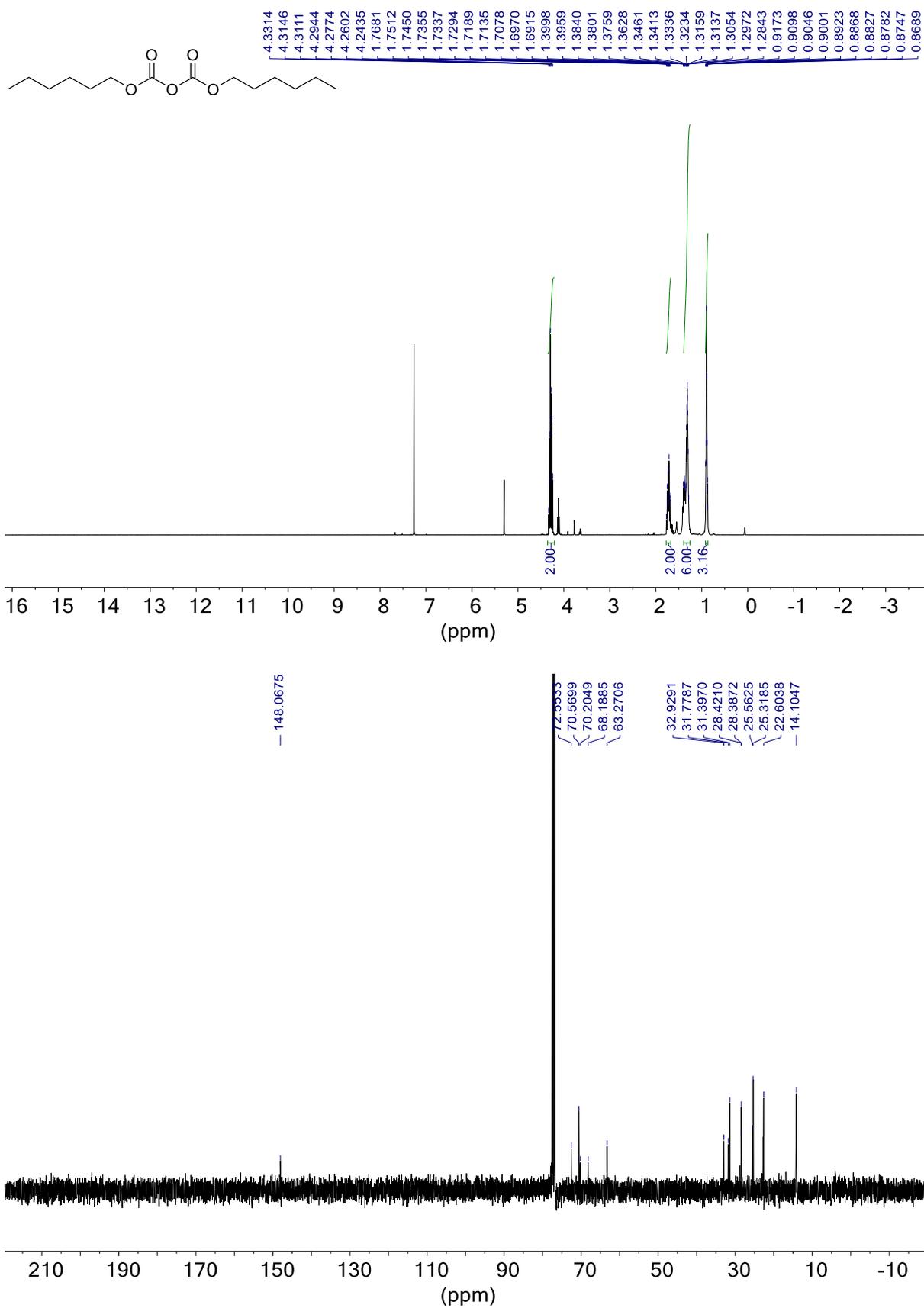
## (diethylamino)-N,N-diethyl-3H-xanthen-3-iminium (93)

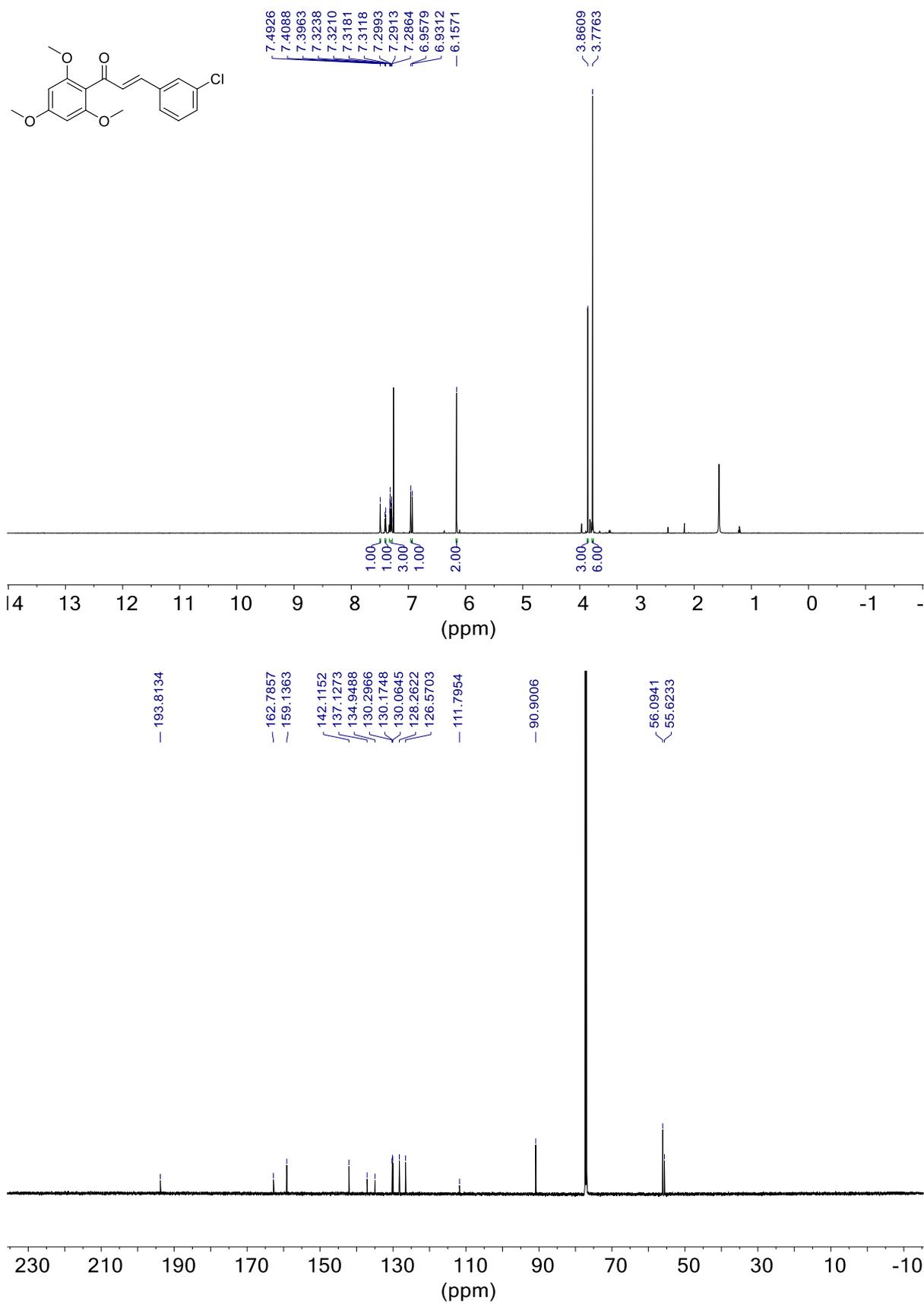


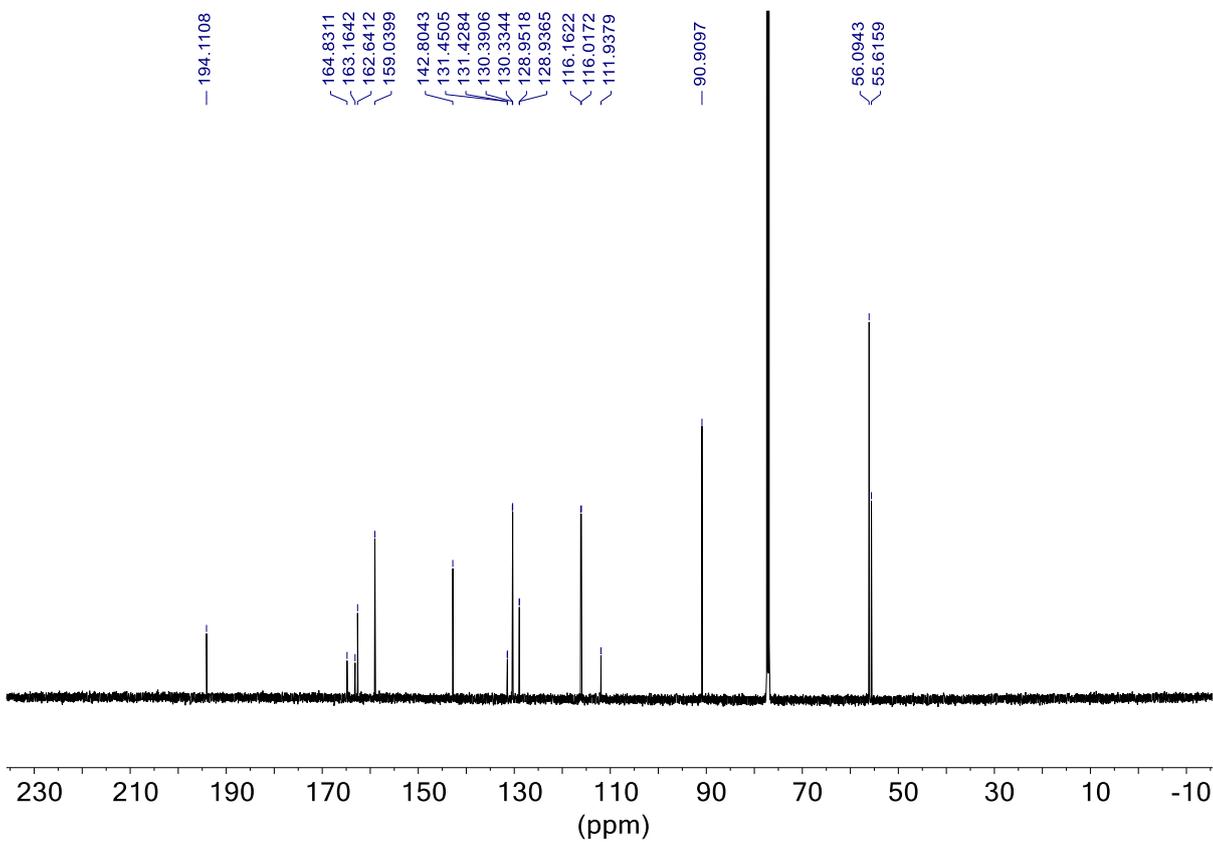
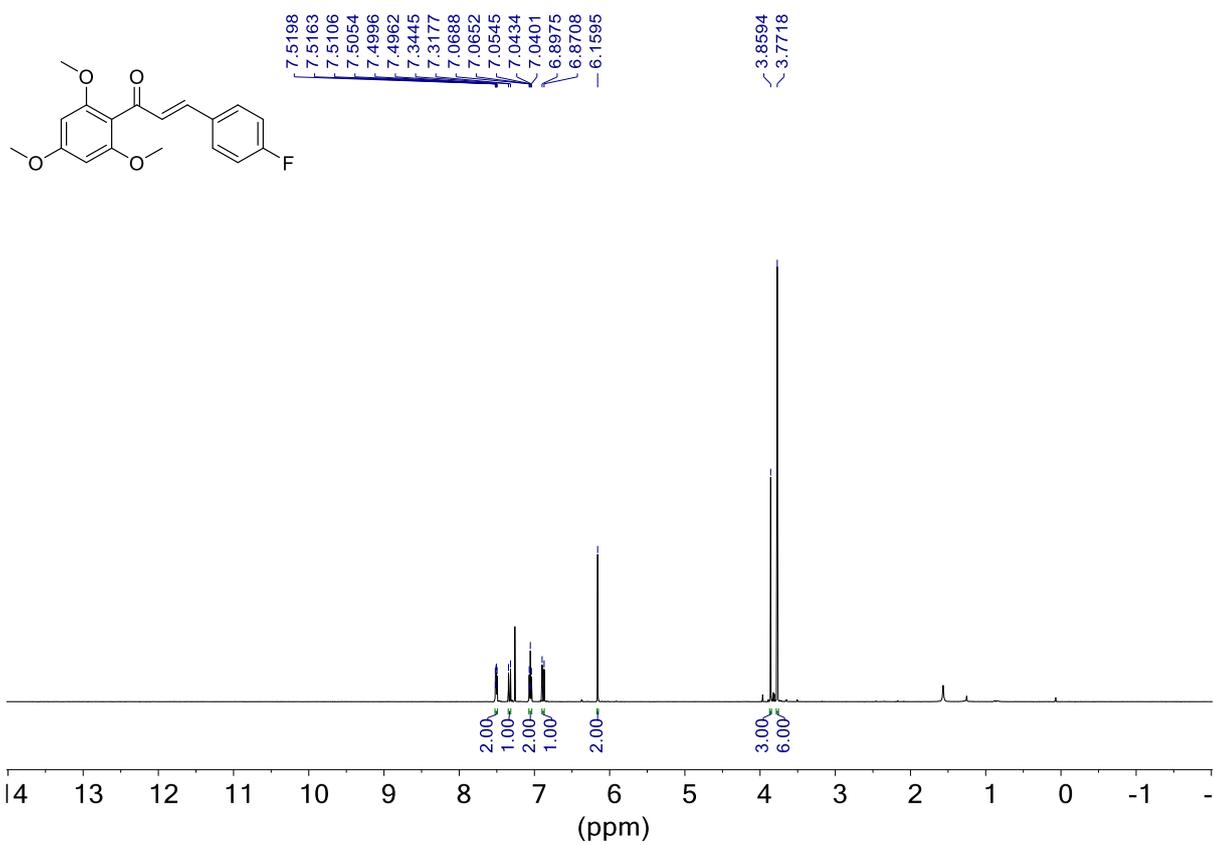
(9H-fluoren-9-yl)methyl N-{1-[(2-{2-[2-(2-azidoethoxy)ethoxy]ethoxy}ethoxy)ethyl]carbamoyl]-5-[[tert-butoxy]carbonyl]amino}pentyl}carbamate (99)

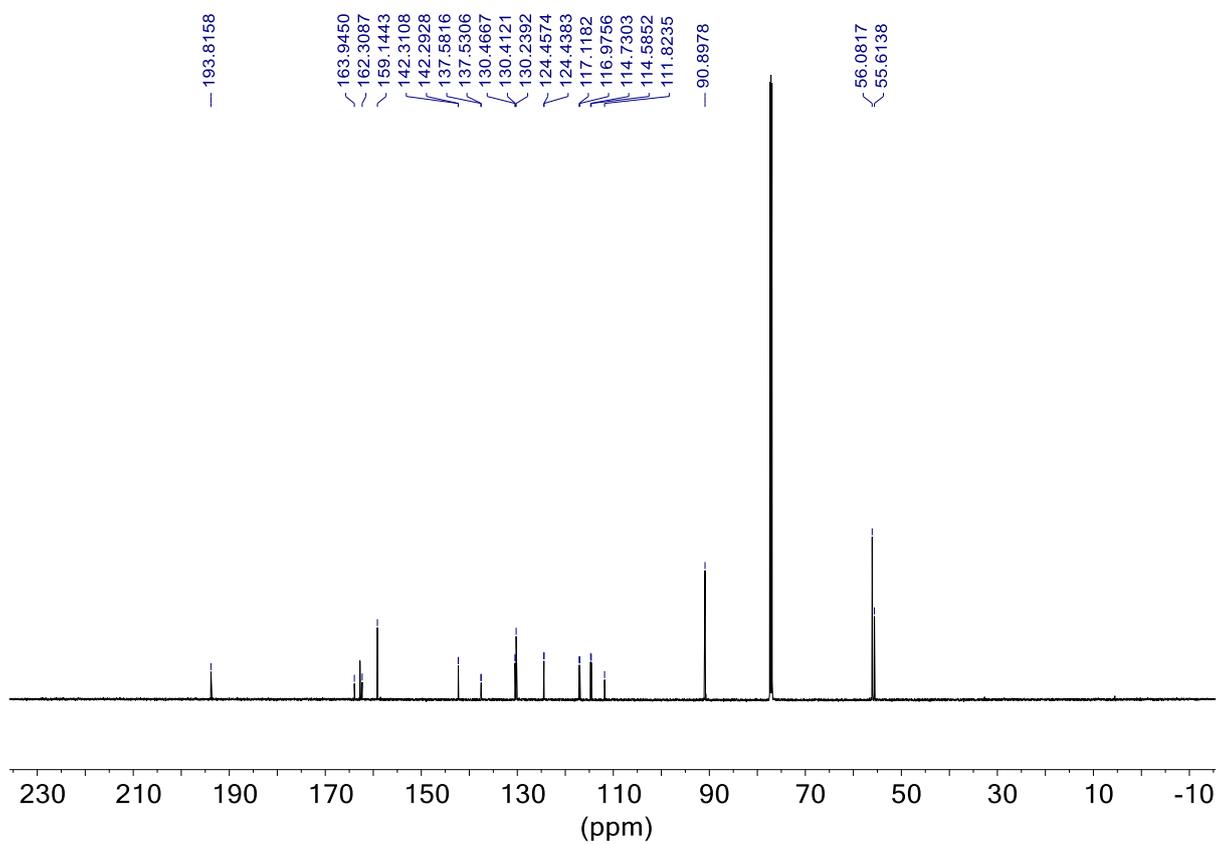
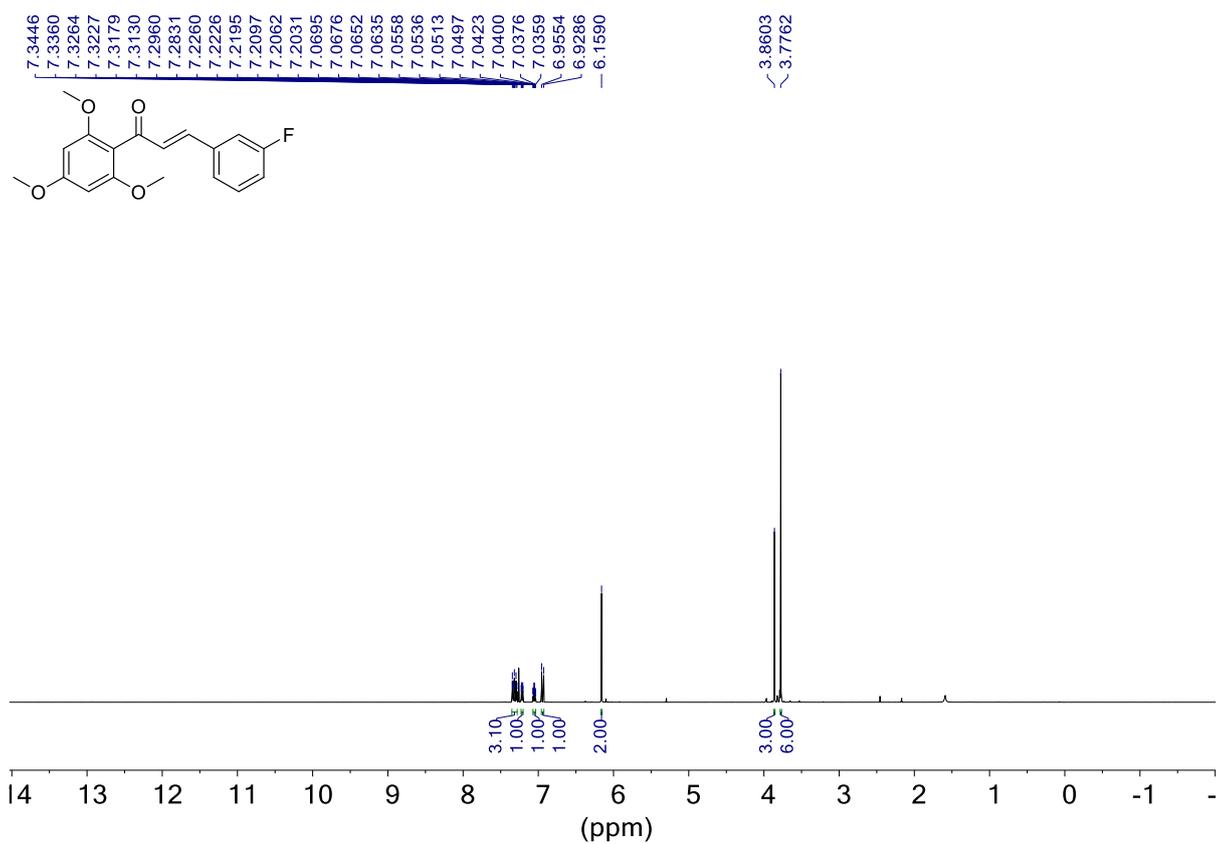


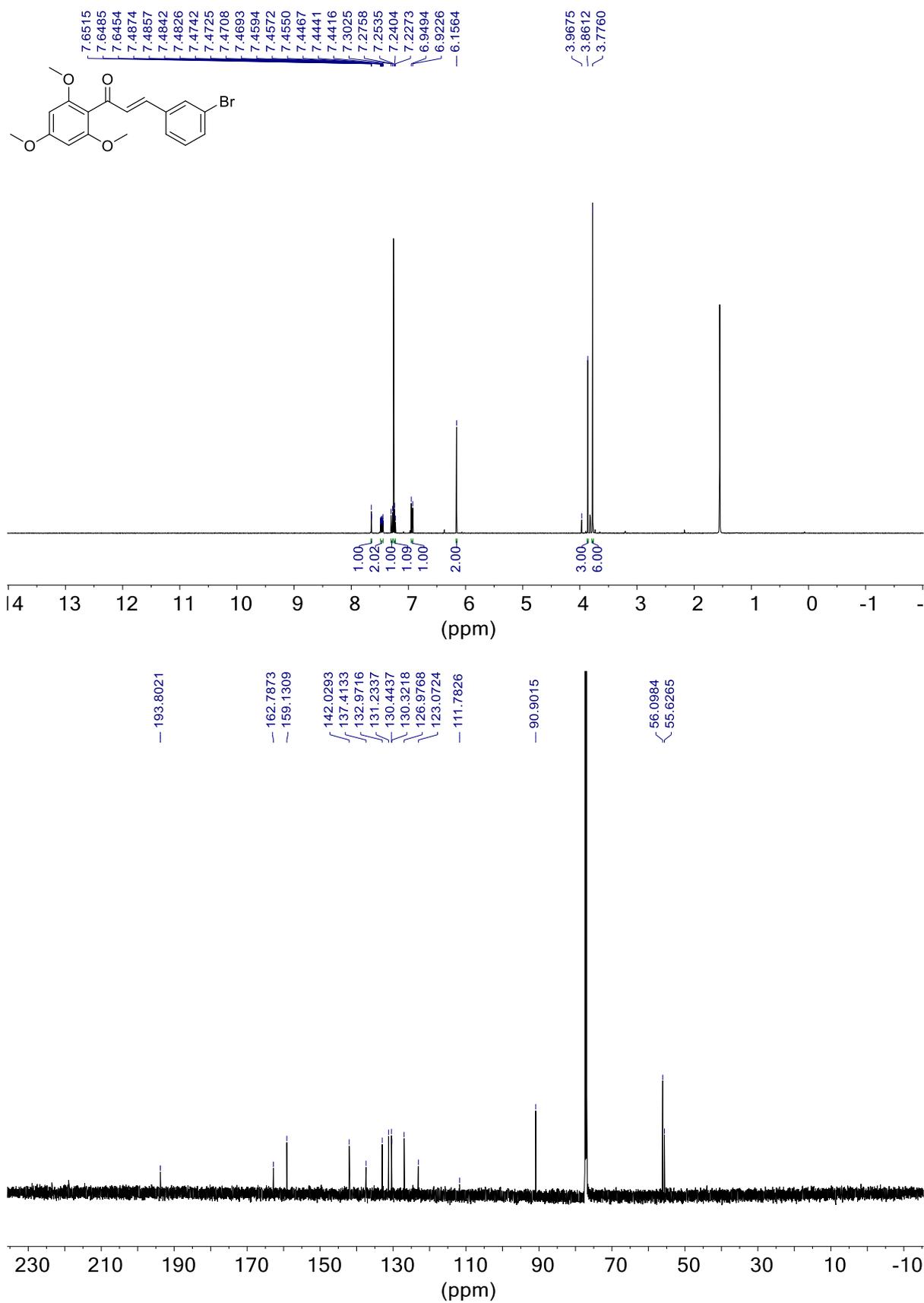
## Dihexyl dicarbonate (104)

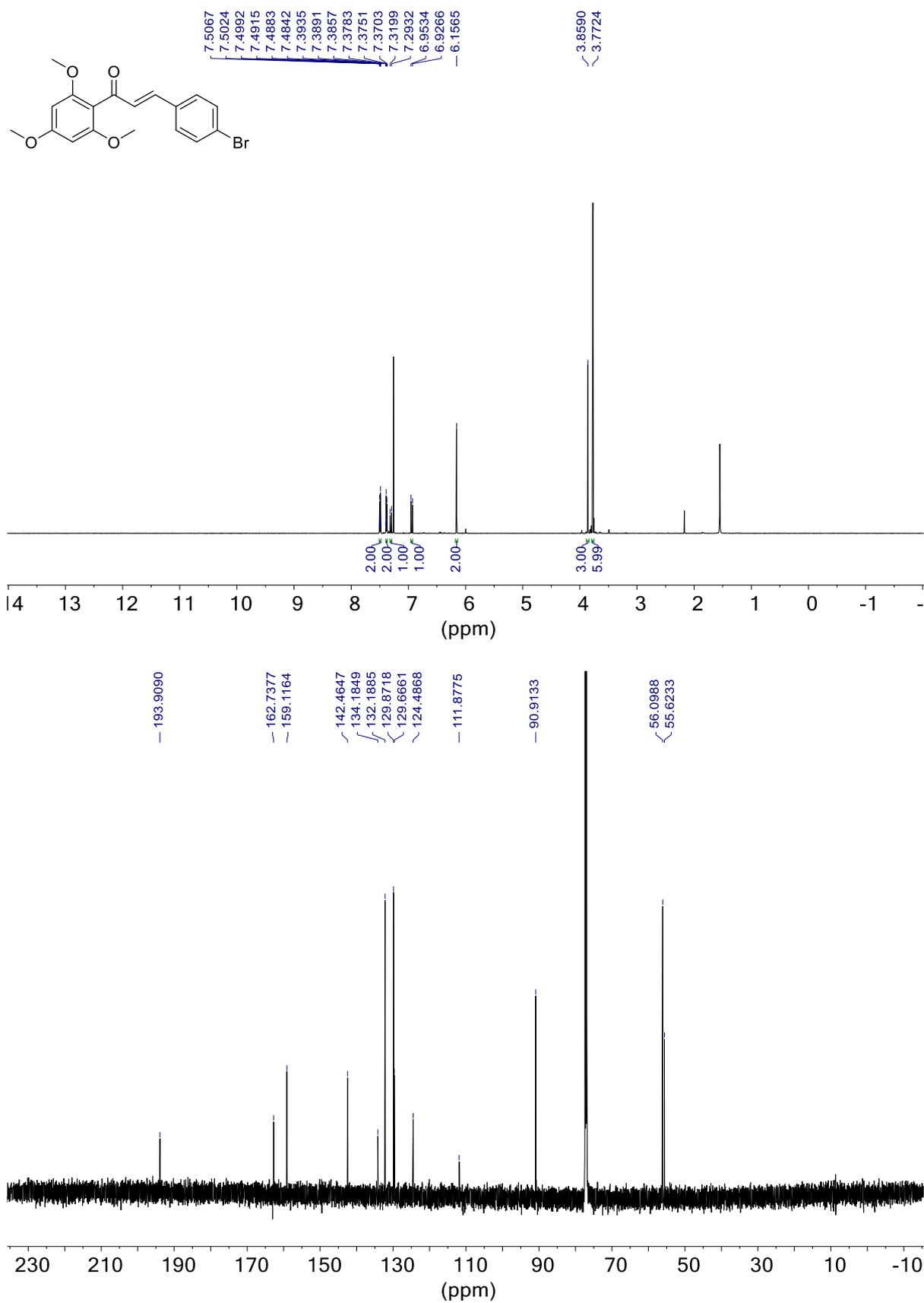


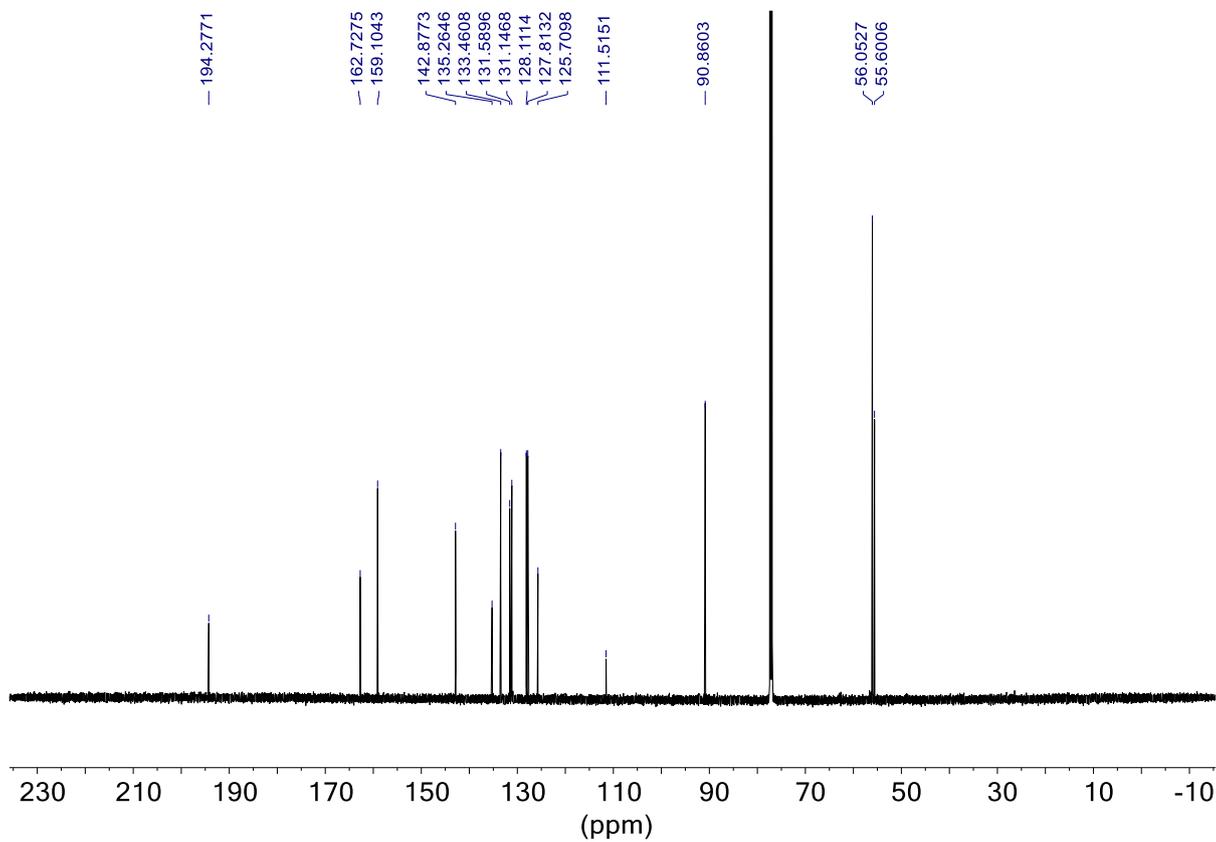
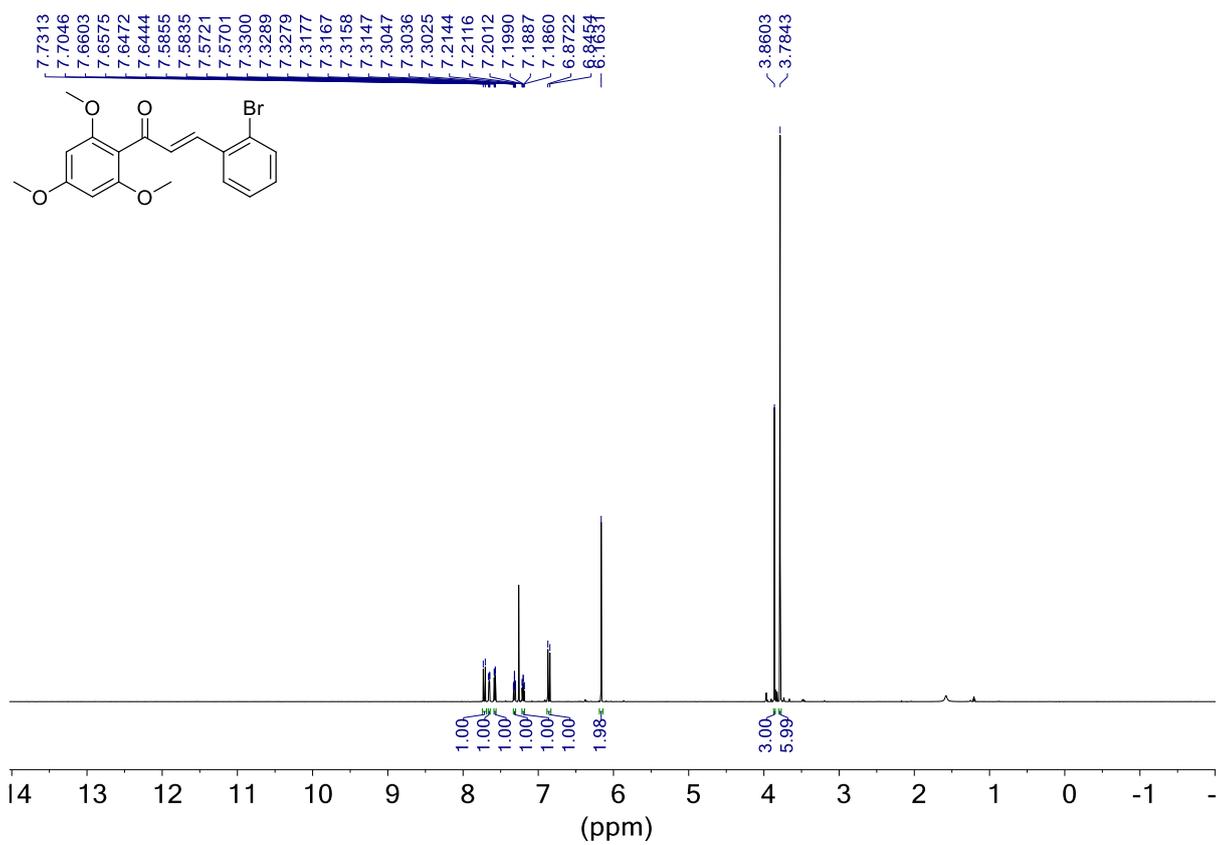
**(2E)-3-(3-chlorophenyl)-1'-(2', 4', 6'-trimethoxyphenyl)prop-2-en-1-one (112)**

**(2E)-3-(4-fluorophenyl)-1-(2,4,6-trimethoxyphenyl)prop-2-en-1-one (114)**

**(2E)-3-(3-fluorophenyl)-1'-(2', 4', 6'-trimethoxyphenyl)prop-2-en-1-one (113)**

**(2E)-3-(3-bromophenyl)-1<sup>1</sup>-(2<sup>1</sup>, 4<sup>1</sup>, 6<sup>1</sup>-trimethoxyphenyl)prop-2-en-1-one (115)**

**(2E)-3-(4-bromophenyl)-1-(2,4,6-trimethoxyphenyl)prop-2-en-1-one (116)**

**(2E)-3-(2-bromophenyl)-1-(2,4,6-trimethoxyphenyl)prop-2-en-1-one (117)**

## 11 References

- (1) Stuart, K.; Brun, R.; Croft, S.; Fairlamb, A.; Gürtler, R. E.; McKerrow, J.; Reed, S.; Tarleton, R. Kinetoplastids: Related Protozoan Pathogens, Different Diseases. *J. Clin. Invest.* **2008**, *118* (4), 1301–1310. <https://doi.org/10.1172/JCI33945>.
- (2) Mann, S.; Frasca, K.; Scherrer, S.; Henao-Martínez, A. F.; Newman, S.; Ramanan, P.; Suarez, J. A. A Review of Leishmaniasis: Current Knowledge and Future Directions. *Curr. Trop. Med. Reports* **2021**, *8* (2), 1–12. <https://doi.org/10.1007/s40475-021-00232-7>.
- (3) Georgiadou, S. P.; Makaritsis, K. P.; Dalekos, G. N. Leishmaniasis Revisited: Current Aspects on Epidemiology, Diagnosis and Treatment. *J. Transl. Intern. Med.* **2015**, *3* (2), 43–45. <https://doi.org/10.1515/jtim-2015-0002>.
- (4) Jain, K.; Jain, N. K. Vaccines for Visceral Leishmaniasis: A Review. *J. Immunol. Methods* **2015**, *422*, 1–12. <https://doi.org/10.1016/j.jim.2015.03.017>.
- (5) CDC. Leishmaniasis <http://www.cdc.gov/parasites/leishmaniasis/>.
- (6) Lindoso, J. A. L.; Cunha, M. A.; Queiroz, I. T.; Moreira, C. H. V. Leishmaniasis–HIV Coinfection: Current Challenges. *HIV/AIDS - Res. Palliat. Care* **2016**, *8*, 147–156. <https://doi.org/10.2147/HIV.S93789>.
- (7) de Oliveira Guerra, J. A.; Prestes, S. R.; Silveira, H.; Coelho, L. I. de A. R. C.; Gama, P.; Moura, A.; Amato, V.; Barbosa, M. das G. V.; de Lima Ferreira, L. C. Mucosal Leishmaniasis Caused by *Leishmania (Viannia) Braziliensis* and *Leishmania (Viannia) Guyanensis* in the Brazilian Amazon. *PLoS Negl. Trop. Dis.* **2011**, *5* (3), e980. <https://doi.org/10.1371/journal.pntd.0000980>.
- (8) Alvar, J.; Vélez, I. D.; Bern, C.; Herrero, M.; Desjeux, P.; Cano, J.; Jannin, J.; de Boer, M. Leishmaniasis Worldwide and Global Estimates of Its Incidence. *PLoS One* **2012**, *7* (5), e35671. <https://doi.org/10.1371/journal.pone.0035671>.
- (9) Bailey, M. S.; Lockwood, D. N. J. Cutaneous Leishmaniasis. *Clin. Dermatol.* **2007**, *25* (2), 203–211. <https://doi.org/10.1016/j.clindermatol.2006.05.008>.
- (10) Dowlati, Y. Cutaneous Leishmaniasis: Clinical Aspect. *Clin. Dermatol.* **1996**, *14* (5), 425–431. [https://doi.org/10.1016/0738-081X\(96\)00058-2](https://doi.org/10.1016/0738-081X(96)00058-2).
- (11) Palumbo, E. Treatment Strategies for Mucocutaneous Leishmaniasis. *J. Glob. Infect. Dis.* **2010**, *2* (2), 147. <https://doi.org/10.4103/0974-777X.62879>.
- (12) Walton, B. C.; Velasco, O. The Distribution and Aetiology of Diffuse Cutaneous Leishmaniasis in the New World. In *Leishmaniasis*; Springer: Boston, MA, 1989; pp 149–157. [https://doi.org/10.1007/978-1-4613-1575-9\\_20](https://doi.org/10.1007/978-1-4613-1575-9_20).
- (13) Scorza, B. M.; Carvalho, E. M.; Wilson, M. E. Cutaneous Manifestations of Human and Murine Leishmaniasis. *Int. J. Mol. Sci.* **2017**, *18* (6), 1296. <https://doi.org/10.3390/ijms18061296>.

- (14) Machado, G. U.; Prates, F. V.; Machado, P. R. L. Disseminated Leishmaniasis: Clinical, Pathogenic, and Therapeutic Aspects. *An. Bras. Dermatol.* **2019**, *94*, 9–16. <https://doi.org/10.1590/abd1806-4841.20198775>.
- (15) Ahluwalia, S. Mucocutaneous Leishmaniasis: An Imported Infection among Travellers to Central and South America. *BMJ* **2004**, *329* (7470), 842–844. <https://doi.org/10.1136/bmj.329.7470.842>.
- (16) Yanik, M.; Gurel, M. S.; Simsek, Z.; Kati, M. The Psychological Impact of Cutaneous Leishmaniasis. *Clin. Exp. Dermatol.* **2004**, *29* (5). <https://doi.org/10.1111/j.1365-2230.2004.01605.x>.
- (17) Lindoso, J. A.; Cota, G. F.; da Cruz, A. M.; Goto, H.; Maia-Elkhoury, A. N. S.; Romero, G. A. S.; de Sousa-Gomes, M. L.; Santos-Oliveira, J. R.; Rabello, A. Visceral Leishmaniasis and HIV Coinfection in Latin America. *PLoS Negl. Trop. Dis.* **2014**, *8* (9), e3136. <https://doi.org/10.1371/journal.pntd.0003136>.
- (18) Karimi, A.; Alborzi, A.; Amanati, A. Visceral Leishmaniasis: An Update and Literature Review. *Arch. Pediatr. Infect. Dis.* **2016**, *4* (3). <https://doi.org/10.5812/pedinfect.31612>.
- (19) Zijlstra, E. E.; Musa, A. M.; Khalil, E. A. G.; El Hassan, I. M.; El-Hassan, A. M. Post-Kala-Azar Dermal Leishmaniasis. *Lancet Infect. Dis.* **2003**, *3* (2), 87–98. [https://doi.org/10.1016/S1473-3099\(03\)00517-6](https://doi.org/10.1016/S1473-3099(03)00517-6).
- (20) Mukhopadhyay, D.; Dalton, J. E.; Kaye, P. M.; Chatterjee, M. Post Kala-Azar Dermal Leishmaniasis: An Unresolved Mystery. *Trends Parasitol.* **2014**, *30* (2), 65–74. <https://doi.org/10.1016/j.pt.2013.12.004>.
- (21) Ganguly, S.; Das, N. K.; Barbhuiya, J. N.; Chatterjee, M. Post-Kala-Azar Dermal Leishmaniasis - an Overview. *Int. J. Dermatol.* **2010**, *49* (8), 921–931. <https://doi.org/10.1111/j.1365-4632.2010.04558.x>.
- (22) Burza, S.; Sinha, P. K.; Mahajan, R.; Sanz, M. G.; Lima, M. A.; Mitra, G.; Verma, N.; Das, P. Post Kala-Azar Dermal Leishmaniasis Following Treatment with 20 Mg/Kg Liposomal Amphotericin B (Ambisome) for Primary Visceral Leishmaniasis in Bihar, India. *PLoS Negl. Trop. Dis.* **2014**, *8* (1), e2611. <https://doi.org/10.1371/journal.pntd.0002611>.
- (23) Rahman, K. M.; Islam, S.; Rahman, M. W.; Kenah, E.; Galive, C. M.; Zahid, M. M.; Maguire, J.; Rahman, M.; Haque, R.; Luby, S. P.; et al. Increasing Incidence of Post-Kala-Azar Dermal Leishmaniasis in a Population-Based Study in Bangladesh. *Clin. Infect. Dis.* **2010**, *50* (1), 73–76. <https://doi.org/10.1086/648727>.
- (24) Ramírez, J. D.; Hernández, C.; León, C. M.; Ayala, M. S.; Flórez, C.; González, C. Taxonomy, Diversity, Temporal and Geographical Distribution of Cutaneous Leishmaniasis in Colombia: A Retrospective Study. *Sci. Rep.* **2016**, *6* (1), 1–10. <https://doi.org/10.1038/srep28266>.
- (25) Oryan, A.; Akbari, M. Worldwide Risk Factors in Leishmaniasis. *Asian Pac. J. Trop. Med.* **2016**, *9* (10), 925–932. <https://doi.org/10.1016/j.apjtm.2016.06.021>.
- (26) Alvar, J.; Yactayo, S.; Bern, C. Leishmaniasis and Poverty. *Trends Parasitol.* **2006**, *22*,

- 552–557. <https://doi.org/10.1016/j.pt.2006.09.004>.
- (27) De Pablos, L. M.; Ferreira, T. R.; Walrad, P. B. Developmental Differentiation in Leishmania Lifecycle Progression: Post-Transcriptional Control Conducts the Orchestra. *Curr. Opin. Microbiol.* **2016**, *34*, 82–89. <https://doi.org/10.1016/j.mib.2016.08.004>.
- (28) Bates, P. A. Transmission of Leishmania Metacyclic Promastigotes by Phlebotomine Sand Flies. *Int. J. Parasitol.* **2007**, *37* (10), 1097–1106. <https://doi.org/10.1016/j.ijpara.2007.04.003>.
- (29) Rogers, M. E.; Bates, P. A. Leishmania Manipulation of Sand Fly Feeding Behavior Results in Enhanced Transmission. *PLoS Pathog.* **2007**, *3* (6), 0818–0825. <https://doi.org/10.1371/journal.ppat.0030091>.
- (30) Liu, D.; Uzonna, J. E. The Early Interaction of Leishmania with Macrophages and Dendritic Cells and Its Influence on the Host Immune Response. *Front. Cell. Infect. Microbiol.* **2012**, *2*, 83. <https://doi.org/10.3389/fcimb.2012.00083>.
- (31) Palatnik-de-Sousa, C. B. Vaccines for Canine Leishmaniasis. *Front. Immunol.* **2012**, *3*, 69. <https://doi.org/10.3389/fimmu.2012.00069>.
- (32) de Vries, H. J. C.; Reedijk, S. H.; Schallig, H. D. F. H. Cutaneous Leishmaniasis: Recent Developments in Diagnosis and Management. *Am. J. Clin. Dermatol.* **2015**, *16* (2), 99–109. <https://doi.org/10.1007/s40257-015-0114-z>.
- (33) David, C. V.; Craft, N. Cutaneous and Mucocutaneous Leishmaniasis. **2009**, *22* (6), 491–502. <https://doi.org/10.1111/j.1529-8019.2009.01272.x>.
- (34) van Griensven, J.; Gadisa, E.; Aseffa, A.; Hailu, A.; Beshah, A. M.; Diro, E. Treatment of Cutaneous Leishmaniasis Caused by Leishmania Aethiopica: A Systematic Review. *PLoS Negl. Trop. Dis.* **2016**, *10* (3), 1–20. <https://doi.org/10.1371/journal.pntd.0004495>.
- (35) Zerpa, O.; Ulrich, M.; Blanco, B.; Polegre, M.; Avila, A.; Matos, N.; Mendoza, I.; Pratlong, F.; Ravel, C.; Convit, J. Diffuse Cutaneous Leishmaniasis Responds to Miltefosine but Then Relapses. *Br. J. Dermatol.* **2007**, *156* (6), 1328–1335. <https://doi.org/10.1111/j.1365-2133.2007.07872.x>.
- (36) Kato, K. C.; Morais-Teixeira, E.; Reis, P. G.; Silva-Barcellos, N. M.; Salaün, P.; Campos, P. P.; Dias Corrêa-Junior, J.; Rabello, A.; Demicheli, C.; Frézard, F. Hepatotoxicity of Pentavalent Antimonial Drug: Possible Role of Residual Sb(III) and Protective Effect of Ascorbic Acid. *Antimicrob. Agents Chemother.* **2014**, *58* (1), 481–488. <https://doi.org/10.1128/AAC.01499-13>.
- (37) Ponte-Sucre, A.; Gamarro, F.; Dujardin, J. C.; Barrett, M. P.; López-Vélez, R.; García-Hernández, R.; Pountain, A. W.; Mwenechanya, R.; Papadopoulou, B. Drug Resistance and Treatment Failure in Leishmaniasis: A 21st Century Challenge. *PLoS Negl. Trop. Dis.* **2017**, *11* (12), e0006052. <https://doi.org/10.1371/journal.pntd.0006052>.
- (38) Haldar, A. K.; Sen, P.; Roy, S. Use of Antimony in the Treatment of Leishmaniasis: Current Status and Future Directions. *Mol. Biol. Int.* **2011**, *2011*, 1–23.

- <https://doi.org/10.4061/2011/571242>.
- (39) Maltezou, H. C. Drug Resistance in Visceral Leishmaniasis. *J. Biomed. Biotechnol.* **2009**, *2010*. <https://doi.org/10.1155/2010/617521>.
- (40) Frézard, F.; Demicheli, C.; Ribeiro, R. R. Pentavalent Antimonials: New Perspectives for Old Drugs. *Molecules* **2009**, *14* (7), 2317–2336. <https://doi.org/10.3390/molecules14072317>.
- (41) Sundar, S.; Goyal, N. Molecular Mechanisms of Antimony Resistance in Leishmania. *J. Med. Microbiol.* **2007**, *56* (2), 143–153. <https://doi.org/10.1099/jmm.0.46841-0>.
- (42) Jeddi, F.; Piarroux, R.; Mary, C. Antimony Resistance in Leishmania, Focusing on Experimental Research. *J. Trop. Med.* **2011**, *2011* (695382). <https://doi.org/10.1155/2011/695382>.
- (43) Balasegaram, M.; Ritmeijer, K.; Lima, M. A.; Burza, S.; Ortiz Genovese, G.; Milani, B.; Gaspani, S.; Potet, J.; Chappuis, F. Liposomal Amphotericin B as a Treatment for Human Leishmaniasis. *Expert Opin. Emerg. Drugs* **2012**, *17* (4), 493–510. <https://doi.org/10.1517/14728214.2012.748036>.
- (44) Deray, G. Amphotericin B Nephrotoxicity. *J. Antimicrob. Chemother.* **2002**, *49*, 37–41.
- (45) Neves, L. O.; Talhari, A. C.; Gadelha, E. P. N.; Silva Júnior, R. M. da; Guerra, J. A. de O.; Ferreira, L. C. de L.; Talhari, S. A Randomized Clinical Trial Comparing Meglumine Antimoniate, Pentamidine and Amphotericin B for the Treatment of Cutaneous Leishmaniasis by Leishmania Guyanensis. *An. Bras. Dermatol.* **2011**, *86* (6), 1092–1101. <https://doi.org/10.1590/S0365-05962011000600005>.
- (46) Hamill, R. J. Amphotericin B Formulations: A Comparative Review of Efficacy and Toxicity. *Drugs* **2013**, *73* (9), 919–934. <https://doi.org/10.1007/s40265-013-0069-4>.
- (47) Sundar, S.; Gupta, L. B.; Rastogi, V.; Agrawal, G.; Murray, H. W. Short-Course, Cost-Effective Treatment with Amphotericin B-Fat Emulsion Cures Visceral Leishmaniasis. *Trans. R. Soc. Trop. Med. Hyg.* **2000**, *94* (2), 200–204. [https://doi.org/10.1016/S0035-9203\(00\)90277-3](https://doi.org/10.1016/S0035-9203(00)90277-3).
- (48) Shirzadi, M. R. Liposomal Amphotericin B: A Review of Its Properties, Function, and Use for Treatment of Cutaneous Leishmaniasis. *Res. Rep. Trop. Med.* **2019**, *10*, 11. <https://doi.org/10.2147/rrtm.s200218>.
- (49) Purkait, B.; Kumar, A.; Nandi, N.; Sardar, A. H.; Das, S.; Kumar, S.; Pandey, K.; Ravidas, V.; Kumar, M.; De, T.; et al. Mechanism of Amphotericin B Resistance in Clinical Isolates of Leishmania Donovanii. *Antimicrob. Agents Chemother.* **2012**, *56* (2), 1031–1041. <https://doi.org/10.1128/AAC.00030-11>.
- (50) Hafiz, S.; Kyriakopoulos, C. *Pentamidine*; StatPearls Publishing: Treasure Island (FL), 2021.
- (51) Sundar, S.; Chakravarty, J. An Update on Pharmacotherapy for Leishmaniasis. *Expert Opin. Pharmacother.* **2015**, *16* (2), 237–252. <https://doi.org/10.1517/14656566.2015.973850>.

- (52) Chakravarty, J.; Sundar, S. Drug Resistance in Leishmaniasis. *J. Glob. Infect. Dis.* **2010**, *2* (2), 167–176. <https://doi.org/10.4103/0974-777x.62887>.
- (53) Basselin, M.; Denise, H.; Coombs, G. H.; Barrett, M. P. Resistance to Pentamidine in *Leishmania Mexicana* Involves Exclusion of the Drug from the Mitochondrion. *Antimicrob. Agents Chemother.* **2002**, *46* (12), 3731–3738. <https://doi.org/10.1128/AAC.46.12.3731-3738.2002>.
- (54) Mukherjee, A.; Padmanabhan, P. K.; Sahani, M. H.; Barrett, M. P.; Madhubala, R. Roles for Mitochondria in Pentamidine Susceptibility and Resistance in *Leishmania Donovanii*. *Mol. Biochem. Parasitol.* **2006**, *145* (1), 1–10. <https://doi.org/10.1016/j.molbiopara.2005.08.016>.
- (55) Sun, T.; Zhang, Y. Pentamidine Binds to tRNA through Non-Specific Hydrophobic Interactions and Inhibits Aminoacylation and Translation. *Nucleic Acids Res.* **2008**, *36* (5), 1654–1664. <https://doi.org/10.1093/nar/gkm1180>.
- (56) Wiwanitkit, V. Interest in Paromomycin for the Treatment of Visceral Leishmaniasis (Kala-Azar). *Ther. Clin. Risk Manag.* **2012**, *8*, 323. <https://doi.org/10.2147/TCRM.S30139>.
- (57) De Menezes, J. P. B.; Guedes, C. E. S.; De Oliveira Almeida Petersen, A. L.; Fraga, D. B. M.; Veras, P. S. T. Advances in Development of New Treatment for Leishmaniasis. *Biomed Res. Int.* **2015**, *2015*. <https://doi.org/10.1155/2015/815023>.
- (58) Bhattacharya, A.; Leprohon, P.; Bigot, S.; Padmanabhan, P. K.; Mukherjee, A.; Roy, G.; Gingras, H.; Mestdagh, A.; Papadopoulou, B.; Ouellette, M. Coupling Chemical Mutagenesis to next Generation Sequencing for the Identification of Drug Resistance Mutations in *Leishmania*. *Nat. Commun.* **2019**, *10* (1), 5627. <https://doi.org/10.1038/s41467-019-13344-6>.
- (59) Fernández, M. M.; Malchiodi, E. L.; Algranati, I. D. Differential Effects of Paromomycin on Ribosomes of *Leishmania Mexicana* and Mammalian Cells. *Antimicrob. Agents Chemother.* **2011**, *55* (1), 86–93. <https://doi.org/10.1128/AAC.00506-10>.
- (60) Fourmy, D.; Yoshizawa, S.; Puglisi, J. D. Paromomycin Binding Induces a Local Conformational Change in the A-Site of 16 S rRNA. *J. Mol. Biol.* **1998**, *277* (2), 333–345. <https://doi.org/10.1006/jmbi.1997.1551>.
- (61) Maarouf, M. In Vivo Interference of Paromomycin with Mitochondrial Activity of *Leishmania*. *Exp. Cell Res.* **1997**, *232* (2), 339–348. <https://doi.org/10.1006/excr.1997.3500>.
- (62) Chawla, B.; Jhingran, A.; Panigrahi, A.; Stuart, K. D.; Madhubala, R. Paromomycin Affects Translation and Vesicle-Mediated Trafficking as Revealed by Proteomics of Paromomycin -Susceptible -Resistant *Leishmania Donovanii*. *PLoS One* **2011**, *6* (10), e26660. <https://doi.org/10.1371/journal.pone.0026660>.
- (63) Croft, S. L.; Neal, R. A.; Pendergast, W.; Chan, J. H. The Activity of Alkyl Phosphorylcholines and Related Derivatives against *Leishmania Donovanii*. *Biochem. Pharmacol.* **1987**, *36* (16), 2633–2636. <https://doi.org/10.1016/0006->

- 2952(87)90543-0.
- (64) Singh, N.; Kumar, M.; Singh, R. K. Leishmaniasis: Current Status of Available Drugs and New Potential Drug Targets. *Asian Pac. J. Trop. Med.* **2012**, *5* (6), 485–497. [https://doi.org/10.1016/S1995-7645\(12\)60084-4](https://doi.org/10.1016/S1995-7645(12)60084-4).
- (65) Monge-Maillo, B.; López-Vélez, R.; Saravolatz, L. D. Miltefosine for Visceral and Cutaneous Leishmaniasis: Drug Characteristics and Evidence-Based Treatment Recommendations. *Clin. Infect. Dis.* **2015**, *60* (9), 1398–1404. <https://doi.org/10.1093/cid/civ004>.
- (66) Rakotomanga, M.; Saint-Pierre-Chazalet, M.; Loiseau, P. M. Alteration of Fatty Acid and Sterol Metabolism in Miltefosine-Resistant *Leishmania Donovanii* Promastigotes and Consequences for Drug-Membrane Interactions. *Antimicrob. Agents Chemother.* **2005**, *49* (7), 2677–2686. <https://doi.org/10.1128/AAC.49.7.2677-2686.2005>.
- (67) Dorlo, T. P. C.; Balasegaram, M.; Beijnen, J. H.; de vries, P. J. Miltefosine: A Review of Its Pharmacology and Therapeutic Efficacy in the Treatment of Leishmaniasis. *J. Antimicrob. Chemother.* **2012**, *67* (11), 2576–2597. <https://doi.org/10.1093/jac/dks275>.
- (68) Luque-Ortega, J. R.; Rivas, L. Miltefosine (Hexadecylphosphocholine) Inhibits Cytochrome c Oxidase in *Leishmania Donovanii* Promastigotes. *Antimicrob. Agents Chemother.* **2007**, *51* (4), 1327–1332. <https://doi.org/10.1128/AAC.01415-06>.
- (69) Bush, J. T.; Wasunna, M.; Alves, F.; Alvar, J.; Olliaro, P. L.; Otieno, M.; Sibley, C. H.; Strub Wourgaft, N.; Guerin, P. J. Systematic Review of Clinical Trials Assessing the Therapeutic Efficacy of Visceral Leishmaniasis Treatments: A First Step to Assess the Feasibility of Establishing an Individual Patient Data Sharing Platform. *PLoS Negl. Trop. Dis.* **2017**, *11* (9), e0005781. <https://doi.org/10.1371/journal.pntd.0005781>.
- (70) Hodiamont, C. J.; Kager, P. A.; Bart, A.; de Vries, H. J. C.; van Thiel, P. P. A. M.; Leenstra, T.; de Vries, P. J.; van Vugt, M.; Grobusch, M. P.; van Gool, T. Species-Directed Therapy for Leishmaniasis in Returning Travellers: A Comprehensive Guide. *PLoS Negl. Trop. Dis.* **2014**, *8* (5), e2832. <https://doi.org/10.1371/journal.pntd.0002832>.
- (71) Santos, D. O.; Coutinho, C. E. R.; Madeira, M. F.; Bottino, C. G.; Vieira, R. T.; Nascimento, S. B.; Bernardino, A.; Bourguignon, S. C.; Corte-Real, S.; Pinho, R. T.; et al. Leishmaniasis Treatment - A Challenge That Remains: A Review. *Parasitol. Res.* **2008**, *103* (1), 1–10. <https://doi.org/10.1007/s00436-008-0943-2>.
- (72) den Boer, M.; Argaw, D.; Jannin, J.; Alvar, J. Leishmaniasis Impact and Treatment Access. *Clin. Microbiol. Infect.* **2011**, *17* (10), 1471–1477. <https://doi.org/10.1111/j.1469-0691.2011.03635.x>.
- (73) Denny, P. W.; Shams-Eldin, H.; Price, H. P.; Smith, D. F.; Schwarz, R. T. The Protozoan Inositol Phosphorylceramide Synthase: A Novel Drug Target That Defines a New Class of Sphingolipid Synthase. *J. Biol. Chem.* **2006**, *281* (38), 28200–28209. <https://doi.org/10.1074/jbc.M600796200>.
- (74) Zulfiqar, B.; Shelper, T. B.; Avery, V. M. Leishmaniasis Drug Discovery: Recent

- Progress and Challenges in Assay Development. *Drug Discov. Today* **2017**, *22* (10), 1516–1531. <https://doi.org/10.1016/j.drudis.2017.06.004>.
- (75) Mohapatra, S. Drug Resistance in Leishmaniasis: Newer Developments. *Trop. Parasitol.* **2014**, *4* (1), 4. <https://doi.org/10.4103/2229-5070.129142>.
- (76) Pandey, K.; Pandey, B.; Pun, S. Relapse of Kala-Azar after Use of Multiple Drugs: A Case Report and Brief Review of Literature. *Indian J. Med. Microbiol.* **2012**, *30* (2), 227. <https://doi.org/10.4103/0255-0857.96703>.
- (77) Mandlik, V.; Singh, S. Molecular Docking and Molecular Dynamics Simulation Study of Inositol Phosphorylceramide Synthase – Inhibitor Complex in Leishmaniasis: Insight into the Structure Based Drug Design. *F1000Research* **2016**, *5*, 1610. <https://doi.org/10.12688/f1000research.9151.1>.
- (78) Hannich, J. T.; Umebayashi, K.; Riezman, H. Distribution and Functions of Sterols and Sphingolipids. *Cold Spring Harb. Perspect. Biol.* **2011**, *3* (5), 1–14. <https://doi.org/10.1101/cshperspect.a004762>.
- (79) Futerman, A. H.; Riezman, H. The Ins and Outs of Sphingolipid Synthesis. *Trends Cell Biol.* **2005**, *15* (6), 312–318. <https://doi.org/10.1016/j.tcb.2005.04.006>.
- (80) Hanada, K. Serine Palmitoyltransferase, a Key Enzyme of Sphingolipid Metabolism. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2003**, *1632* (1–3), 16–30. [https://doi.org/10.1016/S1388-1981\(03\)00059-3](https://doi.org/10.1016/S1388-1981(03)00059-3).
- (81) Školová, B.; Kováčik, A.; Tesař, O.; Opálka, L.; Vávrová, K. Phytosphingosine, Sphingosine and Dihydrosphingosine Ceramides in Model Skin Lipid Membranes: Permeability and Biophysics. *Biochim. Biophys. Acta - Biomembr.* **2017**, *1859* (5), 824–834. <https://doi.org/10.1016/j.bbamem.2017.01.019>.
- (82) Merrill, A. H. De Novo Sphingolipid Biosynthesis: A Necessary, but Dangerous, Pathway. *J. Biol. Chem.* **2002**, *277* (29), 25843–25846. <https://doi.org/10.1074/jbc.R200009200>.
- (83) Gault, C. R.; Obeid, L. M.; Hannun, Y. A. An Overview of Sphingolipid Metabolism: From Synthesis to Breakdown. *Adv. Exp. Med. Biol.* **2010**, 1–23. [https://doi.org/10.1007/978-1-4419-6741-1\\_1](https://doi.org/10.1007/978-1-4419-6741-1_1).
- (84) Mina, J. G.; Mosely, J. A.; Ali, H. Z.; Denny, P. W.; Steel, P. G. Exploring Leishmania Major Inositol Phosphorylceramide Synthase (LmjIPCS): Insights into the Ceramide Binding Domain. *Org. Biomol. Chem.* **2011**, *9* (6), 1823–1830. <https://doi.org/10.1039/c0ob00871k>.
- (85) Koval, M.; Pagano, R. E. Intracellular Transport and Metabolism of Sphingomyelin. *Biochim. Biophys. Acta* **1991**, *1082* (2), 113–125. [https://doi.org/10.1016/0005-2760\(91\)90184-J](https://doi.org/10.1016/0005-2760(91)90184-J).
- (86) Zhang, K.; Beverley, S. M. Phospholipid and Sphingolipid Metabolism in Leishmania. *Mol. Biochem. Parasitol.* **2010**, *170* (2), 55–64. <https://doi.org/10.1016/j.molbiopara.2009.12.004>.
- (87) Ding, T.; Li, Z.; Hailemariam, T.; Mukherjee, S.; Maxfield, F. R.; Wu, M. P.; Jiang, X. C.

- SMS Overexpression and Knockdown: Impact on Cellular Sphingomyelin and Diacylglycerol Metabolism, and Cell Apoptosis. *J. Lipid Res.* **2008**, *49* (2), 376–385. <https://doi.org/10.1194/jlr.M700401-JLR200>.
- (88) Norcliffe, J. L.; Mina, J. G.; Alvarez, E.; Cantizani, J.; De Dios-Anton, F.; Colmenarejo, G.; Valle, S. G. Del; Marco, M.; Fiandor, J. M.; Martin, J. J.; et al. Identifying Inhibitors of the Leishmania Inositol Phosphorylceramide Synthase with Antiprotozoal Activity Using a Yeast-Based Assay and Ultra-High Throughput Screening Platform. *Sci. Rep.* **2018**, *8* (1), 1–10. <https://doi.org/10.1038/s41598-018-22063-9>.
- (89) Sutterwala, S. S.; Hsu, F. F.; Sevova, E. S.; Schwartz, K. J.; Zhang, K.; Key, P.; Turk, J.; Beverley, S. M.; Bangs, J. D. Developmentally Regulated Sphingolipid Synthesis in African Trypanosomes. *Mol. Microbiol.* **2008**, *70* (2), 281–296. <https://doi.org/10.1111/j.1365-2958.2008.06393.x>.
- (90) Moraes, I.; Evans, G.; Sanchez-Weatherby, J.; Newstead, S.; Stewart, P. D. S. Membrane Protein Structure Determination - The next Generation. *Biochim. Biophys. Acta - Biomembr.* **2014**, *1838* (1), 78–87. <https://doi.org/10.1016/j.bbamem.2013.07.010>.
- (91) Yin, H.; Flynn, A. D. Drugging Membrane Protein Interactions. *Annu. Rev. Biomed. Eng.* **2016**, *18*, 51–76. <https://doi.org/10.1146/annurev-bioeng-092115-025322>.
- (92) Mina, J. G.; Mosely, J. A.; Ali, H. Z.; Shams-Eldin, H.; Schwarz, R. T.; Steel, P. G.; Denny, P. W. A Plate-Based Assay System for Analyses and Screening of the Leishmania Major Inositol Phosphorylceramide Synthase. *Int. J. Biochem. Cell Biol.* **2010**, *42* (9), 1553–1561. <https://doi.org/10.1016/j.biocel.2010.06.008>.
- (93) Fischl, A. S.; Liu, Y.; Browdy, A.; Cremesti, A. E. Inositolphosphoryl Ceramide Synthase from Yeast. *Methods Enzymol.* **2000**, *311*, 123–130. [https://doi.org/10.1016/S0076-6879\(00\)11073-0](https://doi.org/10.1016/S0076-6879(00)11073-0).
- (94) Speers, A. E.; Cravatt, B. F. Activity-Based Protein Profiling (ABPP) and Click Chemistry (CC)-ABPP by MudPIT Mass Spectrometry. *Curr. Protoc. Chem. Biol.* **2011**, *1* (1), 29–41. <https://doi.org/10.1002/9780470559277.ch090138>.
- (95) Sieber, S. A.; Cravatt, B. F. Analytical Platforms for Activity-Based Protein Profiling - Exploiting the Versatility of Chemistry for Functional Proteomics. *Chem. Commun.* **2006**, *22*, 2311–2319. <https://doi.org/10.1039/b600653c>.
- (96) Barglow, K. T.; Cravatt, B. F. Activity-Based Protein Profiling for the Functional Annotation of Enzymes. *Nat. Methods* **2007**, *4* (10), 822–827. <https://doi.org/10.1038/nmeth1092>.
- (97) Berger, A. B.; Vitorino, P. M.; Bogoyo, M. Activity-Based Protein Profiling: Applications to Biomarker Discovery, in Vivo Imaging and Drug Discovery. *Am. J. Pharmacogenetics* **2004**, *4* (6), 371–381. <https://doi.org/10.2165/00129785-200404060-00004>.
- (98) Yang, P.; Liu, K. Activity-Based Protein Profiling: Recent Advances in Probe Development and Applications. *ChemBioChem* **2015**, *16* (5), 712–724. <https://doi.org/10.1002/cbic.201402582>.

- (99) Fang, H.; Peng, B.; Ong, S. Y.; Wu, Q.; Li, L.; Yao, S. Q. Recent Advances in Activity-Based Probes (ABPs) and Affinity-Based Probes (AfBPs) for Profiling of Enzymes. *Chem. Sci.* **2021**, *12*, 8288–8310. <https://doi.org/10.1039/d1sc01359a>.
- (100) Wang, S.; Tian, Y.; Wang, M.; Wang, M.; Sun, G.; Sun, X. Advanced Activity-Based Protein Profiling Application Strategies for Drug Development. *Front. Pharmacol.* **2018**, *9*, 353. <https://doi.org/10.3389/fphar.2018.00353>.
- (101) Böttcher, T.; Pitscheider, M.; Sieber, S. A. Natural Products and Their Biological Targets: Proteomic and Metabolomic Labeling Strategies. *Angew. Chemie - Int. Ed.* **2010**, *49* (15), 2680–2698. <https://doi.org/10.1002/anie.200905352>.
- (102) Bennis, H. J.; Wincott, C. J.; Tate, E. W.; Child, M. A. Activity- and Reactivity-Based Proteomics: Recent Technological Advances and Applications in Drug Discovery. *Curr. Opin. Chem. Biol.* **2021**, *60*, 20–29. <https://doi.org/10.1016/j.cbpa.2020.06.011>.
- (103) Deng, H.; Lei, Q.; Wu, Y.; He, Y.; Li, W. Activity-Based Protein Profiling: Recent Advances in Medicinal Chemistry. *Eur. J. Med. Chem.* **2020**, *191*, 112151. <https://doi.org/10.1016/j.ejmech.2020.112151>.
- (104) Niphakis, M. J.; Cravatt, B. F. Enzyme Inhibitor Discovery by Activity-Based Protein Profiling. *Annu. Rev. Biochem.* **2014**, *83* (1), 341–377. <https://doi.org/10.1146/annurev-biochem-060713-035708>.
- (105) Patricelli, M. P. Activity-Based Probes for Functional Proteomics. *Briefings Funct. Genomics Proteomics* **2002**, *1* (2), 151–158. <https://doi.org/10.1093/bfgp/1.2.151>.
- (106) Gerry, C. J.; Schreiber, S. L. Chemical Probes and Drug Leads from Advances in Synthetic Planning and Methodology. *Nat. Rev. Drug Discov.* **2018**, *17* (5), 333–352. <https://doi.org/10.1038/nrd.2018.53>.
- (107) Chen, X.; Wang, Y.; Ma, N.; Tian, J.; Shao, Y.; Zhu, B.; Wong, Y. K.; Liang, Z.; Zou, C.; Wang, J. Target Identification of Natural Medicine with Chemical Proteomics Approach: Probe Synthesis, Target Fishing and Protein Identification. *Signal Transduct. Target. Ther.* **2020**, *5* (1), 1–13. <https://doi.org/10.1038/s41392-020-0186-y>.
- (108) Wang, J.; Tan, X. F.; Nguyen, V. S.; Yang, P.; Zhou, J.; Gao, M.; Li, Z.; Lim, T. K.; He, Y.; Ong, C. S.; et al. A Quantitative Chemical Proteomics Approach to Profile the Specific Cellular Targets of Andrographolide, a Promising Anticancer Agent That Suppresses Tumor Metastasis. *Mol. Cell. Proteomics* **2014**, *13* (3), 876–886. <https://doi.org/10.1074/mcp.M113.029793>.
- (109) Verhelst, S. H. L.; Bongers, K. M.; Willems, L. I. Bioorthogonal Reactions in Activity-Based Protein Profiling. *Molecules* **2020**, *25* (24), 5994. <https://doi.org/10.3390/molecules25245994>.
- (110) Meng, B.; Liu, J.; Wang, L. Oligo(Ethylene Glycol) as Side Chains of Conjugated Polymers for Optoelectronic Applications. *Polym. Chem.* **2020**, *11* (7), 1261–1270. <https://doi.org/10.1039/c9py01469a>.
- (111) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo,

- A.; Codelli, J. A.; Bertozzi, C. R. Copper-Free Click Chemistry for Dynamic in Vivo Imaging. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (43), 16793–16797. <https://doi.org/10.1073/pnas.0707090104>.
- (112) Kang, K.; Park, J.; Kim, E. Tetrazine Ligation for Chemical Proteomics. *Proteome Sci.* **2017**, *15* (1), 1–13. <https://doi.org/10.1186/s12953-017-0121-5>.
- (113) Zweerink, S.; Kallnik, V.; Ninck, S.; Nickel, S.; Verheyen, J.; Blum, M.; Wagner, A.; Feldmann, I.; Sickmann, A.; Albers, S. V.; et al. Activity-Based Protein Profiling as a Robust Method for Enzyme Identification and Screening in Extremophilic Archaea. *Nat. Commun.* **2017**, *8* (1), 1–12. <https://doi.org/10.1038/ncomms15352>.
- (114) Yang, Y.; Hahne, H.; Kuster, B.; Verhelst, S. H. L. A Simple and Effective Cleavable Linker for Chemical Proteomics Applications. *Mol. Cell. Proteomics* **2013**, *12* (1), 237–244. <https://doi.org/10.1074/mcp.M112.021014>.
- (115) Willems, L. I.; Van Der Linden, W. A.; Li, N.; Li, K. Y.; Liu, N.; Hoogendoorn, S.; Van Der Marel, G. A.; Florea, B. I.; Overkleeft, H. S. Bioorthogonal Chemistry: Applications in Activity-Based Protein Profiling. *Acc. Chem. Res.* **2011**, *44* (9), 718–729. <https://doi.org/10.1021/ar200125k>.
- (116) Geurink, P. P.; Florea, B. I.; Li, N.; Witte, M. D.; Verasdonck, J.; Kuo, C. L.; Van Der Marel, G. A.; Overkleeft, H. S. A Cleavable Linker Based on the Levulinoyl Ester for Activity-Based Protein Profiling. *Angew. Chemie - Int. Ed.* **2010**, *49* (38), 6802–6805. <https://doi.org/10.1002/anie.201001767>.
- (117) Roberts, A. M.; Ward, C. C.; Nomura, D. K. Activity-Based Protein Profiling for Mapping and Pharmacologically Interrogating Proteome-Wide Ligandable Hotspots. *Curr. Opin. Biotechnol.* **2017**, *43*, 25–33. <https://doi.org/10.1016/j.copbio.2016.08.003>.
- (118) Li, L.; Zhang, Z. Development and Applications of the Copper-Catalyzed Azide-Alkyne Cycloaddition (CuAAC) as a Bioorthogonal Reaction. *Molecules* **2016**, *21* (10), 1393. <https://doi.org/10.3390/molecules21101393>.
- (119) Fonović, M.; Bogoy, M. Activity-Based Probes as a Tool for Functional Proteomic Analysis of Proteases. *Expert Rev. Proteomics* **2008**, *5* (5), 721–730. <https://doi.org/10.1586/14789450.5.5.721>.
- (120) Smith, E.; Collins, I. Photoaffinity Labeling in Target-and Binding-Site Identification. *Future Med. Chem.* **2015**, *7* (2), 159–183. <https://doi.org/10.4155/fmc.14.152>.
- (121) Dubinsky, L.; Krom, B. P.; Meijler, M. M. Diazirine Based Photoaffinity Labeling. *Bioorganic Med. Chem.* **2012**, *20* (2), 554–570. <https://doi.org/10.1016/j.bmc.2011.06.066>.
- (122) Wanigasekara, M. S. K.; Huang, X.; Chakrabarty, J. K.; Bugarin, A.; Chowdhury, S. M. Arginine-Selective Chemical Labeling Approach for Identification and Enrichment of Reactive Arginine Residues in Proteins. *ACS Omega* **2018**, *3* (10), 14229–14235. <https://doi.org/10.1021/acsomega.8b01729>.
- (123) Ribeiro, A. J. M.; Tyzack, J. D.; Borkakoti, N.; Holliday, G. L.; Thornton, J. M. A Global Analysis of Function and Conservation of Catalytic Residues in Enzymes. *J. Biol.*

- Chem.* **2020**, 295 (2), 314–324. <https://doi.org/10.1074/jbc.REV119.006289>.
- (124) Long, J. Z.; Cravatt, B. F. The Metabolic Serine Hydrolases and Their Functions in Mammalian Physiology and Disease. *Chem. Rev.* **2011**, 111 (10), 6022–6063. <https://doi.org/10.1021/cr200075y>.
- (125) Liu, Y.; Patricelli, M. P.; Cravatt, B. F. Activity-Based Protein Profiling: The Serine Hydrolases. *Proc. Natl. Acad. Sci.* **1999**, 96 (26), 14694–14699. <https://doi.org/10.1073/pnas.96.26.14694>.
- (126) Gillet, L. C. J.; Namoto, K.; Ruchti, A.; Hoving, S.; Boesch, D.; Inverardi, B.; Mueller, D.; Coulot, M.; Schindler, P.; Schweigler, P.; et al. In-Cell Selectivity Profiling of Serine Protease Inhibitors by Activity-Based Proteomics. *Mol. Cell. Proteomics* **2008**, 7 (7), 1241–1253. <https://doi.org/10.1074/mcp.M700505-MCP200>.
- (127) Kidd, D.; Liu, Y.; Cravatt, B. F. Profiling Serine Hydrolase Activities in Complex Proteomes. *Biochemistry* **2001**, 40 (13), 4005–4015. <https://doi.org/10.1021/bi002579j>.
- (128) Bachovchin, D. A.; Ji, T.; Li, W.; Simon, G. M.; Blankman, J. L.; Adibekian, A.; Hoover, H.; Niessen, S.; Cravatt, B. F. Superfamily-Wide Portrait of Serine Hydrolase Inhibition Achieved by Library-versus-Library Screening. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107 (49), 20941–20946. <https://doi.org/10.1073/pnas.1011663107>.
- (129) Bachovchin, D. A.; Cravatt, B. F. The Pharmacological Landscape and Therapeutic Potential of Serine Hydrolases. *Nat. Rev. Drug Discov.* **2012**, 11 (1), 52–68. <https://doi.org/10.1038/nrd3620>.
- (130) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative Reactivity Profiling Predicts Functional Cysteines in Proteomes. *Nature* **2010**, 468 (7325), 790–797. <https://doi.org/10.1038/nature09472>.
- (131) Bak, D. W.; Bechtel, T. J.; Falco, J. A.; Weerapana, E. Cysteine Reactivity across the Subcellular Universe. *Curr. Opin. Chem. Biol.* **2019**, 48, 96–105. <https://doi.org/10.1016/j.cbpa.2018.11.002>.
- (132) Shannon, D. A.; Weerapana, E. Covalent Protein Modification: The Current Landscape of Residue-Specific Electrophiles. *Curr. Opin. Chem. Biol.* **2015**, 24, 18–26. <https://doi.org/10.1016/j.cbpa.2014.10.021>.
- (133) Greenbaum, D.; Medzihradszky, K. F.; Burlingame, A.; Bogyo, M. Epoxide Electrophiles as Activity-Dependent Cysteine Protease Profiling and Discovery Tools. *Chem. Biol.* **2000**, 7 (8), 569–581. [https://doi.org/10.1016/S1074-5521\(00\)00014-4](https://doi.org/10.1016/S1074-5521(00)00014-4).
- (134) Joyce, J. A.; Baruch, A.; Chehade, K.; Meyer-Morse, N.; Giraudo, E.; Tsai, F. Y.; Greenbaum, D. C.; Hager, J. H.; Bogyo, M.; Hanahan, D. Cathepsin Cysteine Proteases Are Effectors of Invasive Growth and Angiogenesis during Multistage Tumorigenesis. *Cancer Cell* **2004**, 5 (5), 443–453. [https://doi.org/10.1016/S1535-6108\(04\)00111-4](https://doi.org/10.1016/S1535-6108(04)00111-4).
- (135) Greenbaum, D.; Baruch, A.; Hayrapetian, L.; Darula, Z.; Burlingame, A.; Medzihradszky, K. F.; Bogyo, M. Chemical Approaches for Functionally Probing the Proteome. *Mol. Cell. proteomics* **2002**, 1 (1), 60–68.

<https://doi.org/10.1074/mcp.T100003-MCP200>.

- (136) Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Guerreiro, A.; Cal, P. M. S. D.; Bertoldo, J.; Maneiro, M.; Perkins, E.; Howard, J.; Deery, M. J.; et al. Chemo- and Regioselective Lysine Modification on Native Proteins. *J. Am. Chem. Soc.* **2018**, *140* (11), 4004–4017. <https://doi.org/10.1021/jacs.7b12874>.
- (137) Wymann, M. P.; Bulgarelli-Leva, G.; Zvelebil, M. J.; Pirola, L.; Vanhaesebroeck, B.; Waterfield, M. D.; Panayotou, G. Wortmannin Inactivates Phosphoinositide 3-Kinase by Covalent Modification of Lys-802, a Residue Involved in the Phosphate Transfer Reaction. *Mol. Cell. Biol.* **1996**, *16* (4), 1722–1733. <https://doi.org/10.1128/mcb.16.4.1722>.
- (138) Ward, C. C.; Kleinman, J. I.; Nomura, D. K. NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem. Biol.* **2017**, *12* (6), 1478–1483. <https://doi.org/10.1021/acschembio.7b00125>.
- (139) Hacker, S. M.; Backus, K. M.; Lazear, M. R.; Forli, S.; Correia, B. E.; Cravatt, B. F. Global Profiling of Lysine Reactivity and Ligandability in the Human Proteome. *Nat. Chem.* **2017**, *9* (12), 1181–1190. <https://doi.org/10.1038/nchem.2826>.
- (140) Li, X.; Ma, H.; Dong, S.; Duan, X.; Liang, S. Selective Labeling of Histidine by a Designed Fluorescein-Based Probe. *Talanta* **2004**, *62* (2), 367–371. <https://doi.org/10.1016/j.talanta.2003.08.004>.
- (141) Liao, S. M.; Du, Q. S.; Meng, J. Z.; Pang, Z. W.; Huang, R. B. The Multiple Roles of Histidine in Protein Interactions. *Chem. Cent. J.* **2013**, *7* (1), 1–12. <https://doi.org/10.1186/1752-153X-7-44>.
- (142) Sundberg, R. J.; Martin, B. Interactions of Histidine and Other Imidazole Derivatives with Transition Metal Ions in Chemical and Biological Systems. *Chem. Rev.* **1974**, *74* (4), 471–517. <https://doi.org/10.1021/cr60290a003>.
- (143) Rigden, D. J. The Histidine Phosphatase Superfamily: Structure and Function. *Biochem. J.* **2008**, *409* (2), 333–348. <https://doi.org/10.1042/BJ20071097>.
- (144) Adam, K.; Hunter, T. Histidine Kinases and the Missing Phosphoproteome from Prokaryotes to Eukaryotes. *Lab. Investig.* **2018**, *98* (2), 233–247. <https://doi.org/10.1038/labinvest.2017.118>.
- (145) Wolanin, P. M.; Thomason, P. A.; Stock, J. B. Histidine Protein Kinases: Key Signal Transducers Outside the Animal Kingdom. *Genome Biol.* **2002**, *3* (10), 1–8. <https://doi.org/10.1186/gb-2002-3-10-reviews3013>.
- (146) Kabbara, S.; Herivaux, A.; De Bernonville, T. D.; Courdavault, V.; Clastre, M.; Gastebois, A.; Osman, M.; Hamze, M.; Mark Cock, J.; Schaap, P.; et al. Diversity and Evolution of Sensor Histidine Kinases in Eukaryotes. *Genome Biol. Evol.* **2019**, *11* (1), 86–108. <https://doi.org/10.1093/gbe/evy213>.
- (147) Dago, A. E.; Schug, A.; Procaccini, A.; Hoch, J. A.; Weigt, M.; Szurmant, H. Structural Basis of Histidine Kinase Autophosphorylation Deduced by Integrating Genomics, Molecular Dynamics, and Mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (26), E1733–E1742. <https://doi.org/10.1073/pnas.1201301109>.

- (148) Bhate, M. A. P.; Molnar, K. A. S.; Goulian, M.; Degrado, W. F. Signal Transduction in Histidine Kinases: Insights from New Structures. *Structure* **2015**, *23* (6), 981–994. <https://doi.org/10.1016/j.str.2015.04.002>.
- (149) Coker, O. O.; Warit, S.; Rukseree, K.; Summpunn, P.; Prammananan, T.; Palittapongarnpim, P. Functional Characterization of Two Members of Histidine Phosphatase Superfamily in Mycobacterium Tuberculosis. *BMC Microbiol.* **2013**, *13* (1), 1–12. <https://doi.org/10.1186/1471-2180-13-292>.
- (150) Alshami, A.; Varon, J. *Acid Phosphatase*, 1st ed.; StatPearls Publishing: Treasure Island (FL), 2020.
- (151) Gandhi, N. U.; Chandra, S. B. A Comparative Analysis of Three Classes of Bacterial Non-Specific Acid Phosphatases and Archaeal Phosphoesterases: Evolutionary Perspective. *Acta Inform. Medica* **2012**, *20* (3), 167–173. <https://doi.org/10.5455/aim.2012.20.167-173>.
- (152) Pond, J. L.; Eddy, C. K.; Mackenzie, K. F.; Conway, T.; Borecky, D. J.; Ingram, L. O. Cloning, Sequencing, and Characterization of the Principal Acid Phosphatase, the PhoC+ Product, from *Zymomonas Mobilis*. *J. Bacteriol.* **1989**, *171* (2), 767–774. <https://doi.org/10.1128/jb.171.2.767-774.1989>.
- (153) Rossolini, G. M.; Schippa, S.; Riccio, M. L.; Berlutti, F.; Macaskie, L. E.; Thaller, M. C. Bacterial Nonspecific Acid Phosphohydrolases: Physiology, Evolution and Use as Tools in Microbial Biotechnology. *Cell. Mol. Life Sci.* **1998**, *54* (8), 833–850. <https://doi.org/10.1007/s000180050212>.
- (154) Makde, R. D.; Kumar, V.; Gupta, G. D.; Jasti, J.; Singh, T. P.; Mahajan, S. K. Expression, Purification, Crystallization and Preliminary X-Ray Diffraction Studies of Recombinant Class B Non-Specific Acid Phosphatase of *Salmonella Typhimurium*. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **2003**, *59* (3), 515–518. <https://doi.org/10.1107/S0907444903018006>.
- (155) Makde, R. D.; Mahajan, S. K.; Kumar, V. Structure and Mutational Analysis of the PhoN Protein of *Salmonella Typhimurium* Provide Insight into Mechanistic Details. *Biochemistry* **2007**, *46* (8), 2079–2090. <https://doi.org/10.1021/bi062180g>.
- (156) Makde, R. D.; Dikshit, K.; Kumar, V. Protein Engineering of Class-A Non-Specific Acid Phosphatase (PhoN) of *Salmonella Typhimurium*: Modulation of the PH-Activity Profile. *Biomol. Eng.* **2006**, *23* (5), 247–251. <https://doi.org/10.1016/j.bioeng.2006.06.004>.
- (157) Rebek, J. On the Structure of Histidine and Its Role in Enzyme Active Sites. *Struct. Chem.* **1990**, *1* (1), 129–131. <https://doi.org/10.1007/BF00675792>.
- (158) Rosenberg, H. F. RNase A Ribonucleases and Host Defense: An Evolving Story. *J. Leukoc. Biol.* **2008**, *83* (5), 1079–1087. <https://doi.org/10.1189/jlb.1107725>.
- (159) Chatani, E.; Hayashi, R.; Moriyama, H.; Ueki, T. Conformational Strictness Required for Maximum Activity and Stability of Bovine Pancreatic Ribonuclease A as Revealed by Crystallographic Study of Three Phe120 Mutants at 1.4 Å Resolution. *Protein Sci.* **2002**, *11* (1), 72–81. <https://doi.org/10.1110/ps.ps.31102>.

- (160) Nogués, M. V.; Vilanova, M.; Cuchillo, C. M. Bovine Pancreatic Ribonuclease A as a Model of an Enzyme with Multiple Substrate Binding Sites. *Biochim. Biophys. Acta (BBA)/Protein Struct. Mol.* **1995**, *1253* (1), 16–24. [https://doi.org/10.1016/0167-4838\(95\)00138-K](https://doi.org/10.1016/0167-4838(95)00138-K).
- (161) Findly, D.; Herries, D. G.; Mathias, A. P.; Rabin, B. R.; Ross, C. A. The Active Site and Mechanism of Action of Bovine Pancreatic Ribonuclease. *Nature* **1961**, *190* (4776), 781–784. <https://doi.org/10.1038/190781a0>.
- (162) Raines, R. T. Ribonuclease A. *Chem. Rev.* **1998**, *98* (3), 1045–1065. <https://doi.org/10.1021/cr960427h>.
- (163) Park, C.; Schultz, L. W.; Raines, R. T. Contribution of the Active Site Histidine Residues of Ribonuclease A to Nucleic Acid Binding. *Biochemistry* **2001**, *40* (16), 4949–4956. <https://doi.org/10.1021/bi0100182>.
- (164) Trautwein, K.; Holliger, P.; Stackhouse, J.; Benner, S. A. Site-Directed Mutagenesis of Bovine Pancreatic Ribonuclease: Lysine-41 and Aspartate-121. *FEBS Lett.* **1991**, *281* (1.2), 275–277. [https://doi.org/10.1016/0014-5793\(91\)80410-5](https://doi.org/10.1016/0014-5793(91)80410-5).
- (165) Silva, K. I.; Michael, B. C.; Geib, S. J.; Saxena, S. ESEEM Analysis of Multi-Histidine Cu(II)-Coordination in Model Complexes, Peptides, and Amyloid- $\beta$ . *J. Phys. Chem. B* **2014**, *118* (30), 8935–8944. <https://doi.org/10.1021/jp500767n>.
- (166) Bornhorst, J. A.; Falke, J. J. Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzymol.* **2000**, *326*, 245–254. [https://doi.org/10.1016/s0076-6879\(00\)26058-8](https://doi.org/10.1016/s0076-6879(00)26058-8).
- (167) Herzik, M. A.; Jonnalagadda, R.; Kuriyan, J.; Marletta, M. A. Structural Insights into the Role of Iron-Histidine Bond Cleavage in Nitric Oxide-Induced Activation of H-NOX Gas Sensor Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (40), E4156–E4164. <https://doi.org/10.1073/pnas.1416936111>.
- (168) Noh, H.; Lee, S. J.; Jo, H. J.; Choi, H. W.; Hong, S.; Kong, K. H. Histidine Residues at the Copper-Binding Site in Human Tyrosinase Are Essential for Its Catalytic Activities. *J. Enzyme Inhib. Med. Chem.* **2020**, *35* (1), 726–732. <https://doi.org/10.1080/14756366.2020.1740691>.
- (169) McCall, K. A.; Huang, C. C.; Fierke, C. A. Function and Mechanism of Zinc Metalloenzymes. *J. Nutr.* **2000**, *130* (5), 1437S–1446S. <https://doi.org/10.1093/jn/130.5.1437s>.
- (170) Crichton, R. R. Zinc: Lewis Acid and Gene Regulator. In *Biological Inorganic Chemistry*; Elsevier: Amsterdam, The Netherlands, 2008; pp 197–210. <https://doi.org/10.1016/b978-044452740-0.50012-0>.
- (171) Fukuoka, S. I.; Ishiguro, K.; Yanagihara, K.; Tanabe, A.; Egashira, Y.; Sanada, H.; Shibata, K. Identification and Expression of a cDNA Encoding Human  $\alpha$ -Amino- $\beta$ -Carboxymuconate- $\epsilon$ -Semiaidehyde Decarboxylase (ACMSD): A Key Enzyme for the Tryptophan-Niacine Pathway and “Quinolate Hypothesis.” *J. Biol. Chem.* **2002**, *277* (38), P35162–35167. <https://doi.org/10.1074/jbc.M200819200>.
- (172) Garavaglia, S.; Perozzi, S.; Galeazzi, L.; Raffaelli, N.; Rizzi, M. The Crystal Structure of

- Human  $\alpha$ -Amino- $\beta$ -Carboxymuconate- $\epsilon$ - Semialdehyde Decarboxylase in Complex with 1,3-Dihydroxyacetonephosphate Suggests a Regulatory Link between NAD Synthesis and Glycolysis. *FEBS J.* **2009**, *276* (22), 6615–6623. <https://doi.org/10.1111/j.1742-4658.2009.07372.x>.
- (173) Huo, L.; Liu, F.; Iwaki, H.; Li, T.; Hasegawa, Y.; Liu, A. Human  $\alpha$ -Amino- $\beta$ -Carboxymuconate- $\epsilon$ -Semialdehyde Decarboxylase (ACMSD): A Structural and Mechanistic Unveiling. *Proteins Struct. Funct. Bioinforma.* **2015**, *83* (1), 178–187. <https://doi.org/10.1002/prot.24722>.
- (174) Huo, L.; Fielding, A. J.; Chen, Y.; Li, T.; Iwaki, H.; Hosler, J. P.; Chen, L.; Hasegawa, Y.; Que, L.; Liu, A. Evidence for a Dual Role of an Active Site Histidine in  $\alpha$ -Amino- $\beta$ -Carboxymuconate- $\epsilon$ -Semialdehyde Decarboxylase. *Biochemistry* **2012**, *51* (29), 5811–5821. <https://doi.org/10.1021/bi300635b>.
- (175) Wilke, K. E.; Francis, S.; Carlson, E. E. Activity-Based Probe for Histidine Kinase Signaling. *J. Am. Chem. Soc.* **2012**, *134* (22), 9150–9153. <https://doi.org/10.1021/ja3041702>.
- (176) Chlebowski, J. F.; Coleman, J. E. Mechanisms of Hydrolysis of O Phosphorothioates and Inorganic Thiophosphate by Escherichia Coli Alkaline Phosphatase. *J. Biol. Chem.* **1974**, *249* (22), 7192–7202. [https://doi.org/10.1016/s0021-9258\(19\)42092-9](https://doi.org/10.1016/s0021-9258(19)42092-9).
- (177) Jia, S.; He, D.; Chang, C. J. Bioinspired Thiophosphorodichloridate Reagents for Chemoselective Histidine Bioconjugation. *J. Am. Chem. Soc.* **2019**, *141* (18), 7294–7301. <https://doi.org/10.1021/jacs.8b11912>.
- (178) Joshi, P. N.; Rai, V. Single-Site Labeling of Histidine in Proteins, on-Demand Reversibility, and Traceless Metal-Free Protein Purification. *Chem. Commun.* **2019**, *55* (8), 1100–1103. <https://doi.org/10.1039/c8cc08733d>.
- (179) Miles, E. W. [41] *Modification of Histidyl Residues in Proteins by Diethylpyrocarbonate*; Academic Press, 1977; Vol. 47. [https://doi.org/10.1016/0076-6879\(77\)47043-5](https://doi.org/10.1016/0076-6879(77)47043-5).
- (180) Hullán, L.; Szontagh, T.; Turtóczky, I.; Fedorcsák, I. The Inactivation of Trypsin by Diethyl Pyrocarbonate. *Acta Chem. Scand.* **1965**, *19* (10), 2440–2441. <https://doi.org/10.3891/acta.chem.scand.19-2440>.
- (181) Rosén, C. G.; Fedorcsák, I. Studies on the Action of Diethyl Pyrocarbonate on Proteins. *Biochim. Biophys. Acta (BBA)-General Subj.* **1966**, *130* (2), 401–405. [https://doi.org/10.1016/0304-4165\(66\)90236-4](https://doi.org/10.1016/0304-4165(66)90236-4).
- (182) Wolf, B.; Lesnaw, J. A.; Reichmann, M. E. A Mechanism of the Irreversible Inactivation of Bovine Pancreatic Ribonuclease by Diethylpyrocarbonate: A General Reaction of Diethylpyrocarbonate with Proteins. *Eur. J. Biochem.* **1970**, *13* (3), 519–525. <https://doi.org/10.1111/j.1432-1033.1970.tb00955.x>.
- (183) Morris, D. L.; Mckinley-mckee, J. S. The Histidines in Liver Alcohol Dehydrogenase: Chemical Modification with Diethylpyrocarbonate. *Eur. J. Biochem.* **1972**. <https://doi.org/10.1111/j.1432-1033.1972.tb02016.x>.
- (184) Solymosy, F.; Fedorcsák, I.; Gulyás, A.; Farkas, G. L.; Ehrenberg, L. A New Method

- Based on the Use of Diethyl Pyrocarbonate as a Nuclease Inhibitor for the Extraction of Undegraded Nucleic Acid from Plant Tissues. *Eur. J. Biochem.* **1968**, *5* (4), 520–527. <https://doi.org/10.1111/j.1432-1033.1968.tb00401.x>.
- (185) Safarian, S.; Moosavi-Movahedi, A. A.; Hosseinkhani, S.; Xia, Z. C.; Habibi-Rezaei, M.; Hosseini, G.; Sorenson, C.; Sheibani, N. The Structural and Functional Studies of His119 and His12 in RNase A via Chemical Modification. *J. Protein Chem.* **2003**, *22* (7), 643–654. <https://doi.org/10.1023/B:JOPC.0000008729.20730.59>.
- (186) Arion, W. J.; Burchell, B.; Burchell, A. Specific Inactivation of the Phosphohydrolase Component of the Hepatic Microsomal Glucose-6-Phosphatase System by Diethyl Pyrocarbonate. *Biochem. J.* **1984**, *220* (3), 835–842. <https://doi.org/10.1042/bj2200835>.
- (187) Plusquellec, D.; Roulleau, F.; Lefevre, M.; Brown, E. A New Synthesis of Carboxylic and Carbonic Acid Anhydrides Using Phase Transfer Reactions. *Tetrahedron* **1988**, *44* (9), 2471–2476. [https://doi.org/10.1016/S0040-4020\(01\)81698-7](https://doi.org/10.1016/S0040-4020(01)81698-7).
- (188) Huber, T.; Díaz, D. D. Dipropargyl 2,2'-Isophthaloylbis(Hydrazinecarboxylate). *Molbank* **2010**, *2010* (4), M701. <https://doi.org/10.3390/M701>.
- (189) Ca', N. Della; Gabriele, B.; Ruffolo, G.; Veltri, L.; Zanetta, T.; Costa, M. Effective Guanidine-Catalyzed Synthesis of Carbonate and Carbamate Derivatives from Propargyl Alcohols in Supercritical Carbon Dioxide. *Adv. Synth. Catal.* **2011**, *353* (1), 133–146. <https://doi.org/10.1002/adsc.201000607>.
- (190) Nigst, T. A.; Antipova, A.; Mayr, H. Nucleophilic Reactivities of Hydrazines and Amines: The Futile Search for the  $\alpha$ -Effect in Hydrazine Reactivities. *J. Org. Chem.* **2012**, *77* (18), 8142–8155. <https://doi.org/10.1021/jo301497g>.
- (191) Pasquato, L.; Modena, G.; Cotarca, L.; Delogu, P.; Mantovani, S. Conversion of Bis(Trichloromethyl) Carbonate to Phosgene and Reactivity of Triphosgene, Diphosgene, and Phosgene with Methanol. *J. Org. Chem.* **2000**, *65* (24), 8224–8228. <https://doi.org/10.1021/jo000820u>.
- (192) Brauch, S.; Henze, M.; Osswald, B.; Naumann, K.; Wessjohann, L. A.; van Berkel, S. S.; Westermann, B. Fast and Efficient MCR-Based Synthesis of Clickable Rhodamine Tags for Protein Profiling. *Org. Biomol. Chem.* **2012**, *10* (5), 958–965. <https://doi.org/10.1039/C1OB06581E>.
- (193) Ralhan, K.; KrishnaKumar, V. G.; Gupta, S. Piperazine and DBU: A Safer Alternative for Rapid and Efficient Fmoc Deprotection in Solid Phase Peptide Synthesis. *RSC Adv.* **2015**, *5* (126), 104417–104425. <https://doi.org/10.1039/c5ra23441g>.
- (194) Chiriac, C. I.; Onciu, M.; Tanasa, F. Synthesis of Aromatic Amides at Room Temperature Using Triphenyl Phosphite-4-Dimethylaminopyridine as Reagent. *Des. Monomers Polym.* **2004**, *7* (4), 331–335. <https://doi.org/10.1163/1568555041475284>.
- (195) Navarro, J. R. G.; Bergström, L. Labelling of N-Hydroxysuccinimide-Modified Rhodamine B on Cellulose Nanofibrils by the Amidation Reaction. *RSC Adv.* **2014**, *4* (105), 60757–60761. <https://doi.org/10.1039/c4ra06559j>.

- (196) Howden, A. J. M.; Geoghegan, V.; Katsch, K.; Efstathiou, G.; Bhushan, B.; Boutureira, O.; Thomas, B.; Trudgian, D. C.; Kessler, B. M.; Dieterich, D. C.; et al. QuaNCAT: Quantitating Proteome Dynamics in Primary Cells. *Nat. Methods* **2013**, *10* (4), 343–346. <https://doi.org/10.1038/nmeth.2401>.
- (197) Traut, R. R.; Bollen, A.; Sun, T. T.; Hershey, J. W. B.; Sundberg, J.; Pierce, L. R. Methyl 4-Mercaptobutyrimidate as a Cleavable Cross-Linking Reagent and Its Application to the Escherichia Coli 30S Ribosome. *Biochemistry* **1973**, *12* (17), 3266–3273. <https://doi.org/10.1021/bi00741a019>.
- (198) Miyajima, R.; Sakai, K.; Otani, Y.; Wadatsu, T.; Sakata, Y.; Nishikawa, Y.; Tanaka, M.; Yamashita, Y.; Hayashi, M.; Kondo, K.; et al. Novel Tetrafunctional Probes Identify Target Receptors and Binding Sites of Small-Molecule Drugs from Living Systems. *ACS Chem. Biol.* **2020**, *15* (9), 2364–2373. <https://doi.org/10.1021/acscchembio.0c00335>.
- (199) Hansch, C.; Clayton, J. Lipophilic Character and Biological Activity of Drugs II: The Parabolic Case. *J. Pharm. Sci.* **1973**, *62* (1), 1–21. <https://doi.org/10.15171/PS.2016.01>.
- (200) Groeneweg, S.; Lima De Souza, E. C.; Visser, W. E.; Peeters, R. P.; Visser, T. J. Importance of His192 in the Human Thyroid Hormone Transporter MCT8 for Substrate Recognition. *Endocrinology* **2013**, *154* (7), 2525–2532. <https://doi.org/10.1210/en.2012-2225>.
- (201) Berrow, N. S.; Alderton, D.; Sainsbury, S.; Nettleship, J.; Assenberg, R.; Rahman, N.; Stuart, D. I.; Owens, R. J. A Versatile Ligation-Independent Cloning Method Suitable for High-Throughput Expression Screening Applications. *Nucleic Acids Res.* **2007**, *35* (6), e45. <https://doi.org/10.1093/nar/gkm047>.
- (202) Hager, D. A.; Jin, D. J.; Burgess, R. R. Use of Mono Q High-Resolution Ion-Exchange Chromatography To Obtain Highly Pure and Active Escherichia Coli RNA Polymerase. *Biochemistry* **1990**, *29* (34), 7890–7894. <https://doi.org/10.1021/bi00486a016>.
- (203) ROCHE, J.; THOAI, N. V. Phosphatase Alkaline. *Adv. Enzymol. Relat. Subj. Biochem.* **1950**, *10*. <https://doi.org/10.1002/9780470114735.hawley12780>.
- (204) Biggs, A. I. A Spectrophotometric Determination of the Dissociation Constants of P-Nitrophenol and Papaverine. *Trans. Faraday Soc.* **1954**, *50*, 800–802. <https://doi.org/10.1039/tf9545000800>.
- (205) Almagro Armenteros, J. J.; Tsirigos, K. D.; Sønderby, C. K.; Petersen, T. N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks. *Nat. Biotechnol.* **2019**, *37* (4), 420–423. <https://doi.org/10.1038/s41587-019-0036-z>.
- (206) Gundry, R. L.; White, M.; Murray, C.; Lesley, K.; Qin, F.; Stanley, B.; Van Eyk, J. Facile Preparation of Peptides for Mass Spectrometry Analysis in Bottom-Up Proteomics Workflows. *Curr. Protoc.* **2009**, *1* (90), 10–25. <https://doi.org/10.1002/cpz1.85>.
- (207) Dyer, E.; Newborn, G. E. Thermal Degradation of Carbamates of Methylenebis-(4-Phenyl Isocyanate). *J. Am. Chem. Soc.* **1958**, *80* (20), 5495–5498.

- <https://doi.org/10.1021/ja01553a045>.
- (208) Green, M. R.; Sambrook, J. How to Win the Battle with RNase. *Cold Spring Harb. Protoc.* **2019**, *2019* (2), 95–98. <https://doi.org/10.1101/pdb.top101857>.
- (209) Presolski, I. S.; Vu, P. H.; Finn, M. . Copper-Catalyzed Azide–Alkyne Click Chemistry for Bioconjugation. *Curr. Protoc. Chem. Biol.* **2011**, *3* (4), 153–162.
- (210) Trinkle-Mulcahy, L.; Boulon, S.; Lam, Y. W.; Urcia, R.; Boisvert, F. M.; Vandermoere, F.; Morrice, N. A.; Swift, S.; Rothbauer, U.; Leonhardt, H.; et al. Identifying Specific Protein Interaction Partners Using Quantitative Mass Spectrometry and Bead Proteomes. *J. Cell Biol.* **2008**, *183* (2), 223–239. <https://doi.org/10.1083/jcb.200805092>.
- (211) Fraser, M. E.; James, M. N. G.; Bridger, W. A.; Wolodko, W. T. A Detailed Structural Description of Escherichia Coli Succinyl-CoA Synthetase. *J. Mol. Biol.* **1999**, *285* (4), 1633–1653. <https://doi.org/10.1006/jmbi.1998.2324>.
- (212) Fothergill-Gilmore, L. A.; Watson, H. C. *The Phosphoglycerate Mutases*; Meister, A., Ed.; John Wiley & Sons Inc, 1989; Vol. 62. <https://doi.org/10.1002/9780470123089.ch6>.
- (213) Johnsen, U.; Schönheit, P. Characterization of Cofactor-Dependent and Cofactor-Independent Phosphoglycerate Mutases from Archaea. *Extremophiles* **2007**, *11* (5), 647. <https://doi.org/10.1007/s00792-007-0094-x>.
- (214) Davies, D. R.; Staker, B. L.; Abendroth, J. A.; Edwards, T. E.; Hartley, R.; Leonard, J.; Kim, H.; Rychel, A. L.; Hewitt, S. N.; Myler, P. J.; et al. An Ensemble of Structures of Burkholderia Pseudomallei 2,3- Bisphosphoglycerate-Dependent Phosphoglycerate Mutase. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2011**, *67* (9), 1044–1050. <https://doi.org/10.1107/S1744309111030405>.
- (215) Bond, C. S.; White, M. F.; Hunter, W. N. High Resolution Structure of the Phosphohistidine-Activated Form of Escherichia Coli Cofactor-Dependent Phosphoglycerate Mutase. *J. Biol. Chem.* **2001**, *276* (5), P3247-3253. <https://doi.org/10.1074/jbc.M007318200>.
- (216) Jarrige, A. C.; Bréchemier-Baey, D.; Mathy, N.; Duché, O.; Portier, C. Mutational Analysis of Polynucleotide Phosphorylase from Escherichia Coli. *J. Mol. Biol.* **2002**, *321* (3), 397–409. [https://doi.org/10.1016/S0022-2836\(02\)00645-9](https://doi.org/10.1016/S0022-2836(02)00645-9).
- (217) Nurmohamed, S.; Vaidialingam, B.; Callaghan, A. J.; Luisi, B. F. Crystal Structure of Escherichia Coli Polynucleotide Phosphorylase Core Bound to RNase E, RNA and Manganese: Implications for Catalytic Mechanism and RNA Degradosome Assembly. *J. Mol. Biol.* **2009**, *389* (1), 17–33. <https://doi.org/10.1016/j.jmb.2009.03.051>.
- (218) Harvey, K. L.; Jarocki, V. M.; Charles, I. G.; Djordjevic, S. P. The Diverse Functional Roles of Elongation Factor Tu (Ef-Tu) in Microbial Pathogenesis. *Front. Microbiol.* **2019**, *10*, 2351. <https://doi.org/10.3389/fmicb.2019.02351>.
- (219) Stark, H.; Rodnina, M. V.; Rinke-Appel, J.; Brimacombe, R.; Wintermeyer, W.; Van Heel, M. Visualization of Elongation Factor Tu on the Escherichia Coli Ribosome. *Nature* **1997**, *389* (6649), 403–406. <https://doi.org/10.1038/38770>.

- (220) Aleksandrov, A.; Field, M. Mechanism of Activation of Elongation Factor Tu by Ribosome: Catalytic Histidine Activates GTP by Protonation. *RNA* **2013**, *19* (9), 1218–1225. <https://doi.org/10.1261/rna.040097.113>.
- (221) Savelsbergh, A.; Rodnina, M. V.; Wintermeyer, W. Distinct Functions of Elongation Factor G in Ribosome Recycling and Translocation. *RNA* **2009**, *15* (5), 772–780. <https://doi.org/10.1261/rna.1592509>.
- (222) Koripella, R. K.; Holm, M.; Dourado, D.; Mandava, C. S.; Flores, S.; Sanyal, S. A Conserved Histidine in Switch-II of EF-G Moderates Release of Inorganic Phosphate. *Sci. Rep.* **2015**, *5* (1), 1–10. <https://doi.org/10.1038/srep12970>.
- (223) Peng, B. Z.; Bock, L. V.; Belardinelli, R.; Peske, F.; Grubmüller, H.; Rodnina, M. V. Active Role of Elongation Factor G in Maintaining the mRNA Reading Frame during Translation. *Sci. Adv.* **2019**, *5* (12), eaax8030. <https://doi.org/10.1126/sciadv.aax8030>.
- (224) Cunha, C. E.; Belardinelli, R.; Peske, F.; Holtkamp, W.; Wintermeyer, W.; Rodnina, M. V. Dual Use of GTP Hydrolysis by Elongation Factor G on the Ribosome. *Translation* **2013**, *1* (1), e24315. <https://doi.org/10.4161/trla.24315>.
- (225) Carvalho, A. T. P.; Szeler, K.; Vavitsas, K.; Åqvist, J.; Kamerlin, S. C. L. Modeling the Mechanisms of Biological GTP Hydrolysis. *Arch. Biochem. Biophys.* **2015**, *582*, 80–90. <https://doi.org/10.1016/j.abb.2015.02.027>.
- (226) Cassey, B.; Guest, J. R.; Attwood, M. M. Environmental Control of Pyruvate Dehydrogenase Complex Expression in Escherichia Coli. *FEMS Microbiol. Lett.* **1998**, *159* (2), 325–329. <https://doi.org/10.1111/j.1574-6968.1998.tb12878.x>.
- (227) Miles, J. S.; Guest, J. R.; Radford, S. E.; Perham, R. N. A Mutant Pyruvate Dehydrogenase Complex of Escherichia Coli Deleted in the (Alanine + Proline)-Rich Region of the Acetyltransferase Component. *Biochim. Biophys. Acta (BBA)/Protein Struct. Mol.* **1987**, *913* (2), 117–121. [https://doi.org/10.1016/0167-4838\(87\)90319-0](https://doi.org/10.1016/0167-4838(87)90319-0).
- (228) Jones, D. D.; Stott, K. M.; Howard, M. J.; Perham, R. N. Restricted Motion of the Lipoyl-Lysine Swinging Arm in the Pyruvate Dehydrogenase Complex of Escherichia Coli. *Biochemistry* **2000**, *39* (29), 8448–8459. <https://doi.org/10.1021/bi992978i>.
- (229) Arjunan, P.; Sax, M.; Brunskill, A.; Chandrasekhar, K.; Nemeria, N.; Zhang, S.; Jordan, F.; Furey, W. A Thiamin-Bound, Pre-Decarboxylation Reaction Intermediate Analogue in the Pyruvate Dehydrogenase E1 Subunit Induces Large Scale Disorder-to-Order Transformations in the Enzyme and Reveals Novel Structural Features in the Covalently Bound Adduct. *J. Biol. Chem.* **2006**, *281* (22), 15296–15303. <https://doi.org/10.1074/jbc.M600656200>.
- (230) Nemeria, N.; Arjunan, P.; Brunskill, A.; Sheibani, F.; Wei, W.; Yan, Y.; Zhang, S.; Jordan, F.; Furey, W. Histidine 407, a Phantom Residue in the E1 Subunit of the Escherichia Coli Pyruvate Dehydrogenase Complex, Activates Reductive Acetylation of Lipoamide on the E2 Subunit. An Explanation for Conservation of Active Sites between the E1 Subunit and Transketo. *Biochemistry* **2002**, *41* (52), 15459–15467. <https://doi.org/10.1021/bi0205909>.

- (231) Mattevi, A.; Obmolova, G.; Kalk, K.; Teplyakov, A.; Hol, W. G. J. Crystallographic Analysis of Substrate Binding and Catalysis in Dihydrolipoyl Transacetylase (E2p). *Biochemistry* **1993**, *32* (15), 3887–3901. <https://doi.org/10.1021/bi00066a007>.
- (232) Uchida, K.; Stadtman, E. R. Modification of Histidine Residues in Proteins by Reaction with 4-Hydroxynonenal. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (10), 4544–4548. <https://doi.org/10.1073/pnas.89.10.4544>.
- (233) Škerlová, J.; Berndtsson, J.; Nolte, H.; Ott, M.; Stenmark, P. Structure of the Native Pyruvate Dehydrogenase Complex Reveals the Mechanism of Substrate Insertion. *Nat. Commun.* **2021**, *12* (1), 1–10. <https://doi.org/10.1038/s41467-021-25570-y>.
- (234) Nie, A.; Sun, B.; Fu, Z.; Yu, D. Roles of Aminoacyl-TRNA Synthetases in Immune Regulation and Immune Diseases. *Cell Death Dis.* **2019**, *10* (12), 1–14. <https://doi.org/10.1038/s41419-019-2145-5>.
- (235) Minajigi, A.; Francklyn, C. S. RNA-Assisted Catalysis in a Protein Enzyme: The 2'-Hydroxyl of TRNAThr A76 Promotes Aminoacylation by Threonyl-TRNA Synthetase. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (46), 17748–17753. <https://doi.org/10.1073/pnas.0804247105>.
- (236) Tang, Y.; Quail, M. A.; Artymiuk, P. J.; Guest, J. R.; Green, J. Escherichia Coli Aconitases and Oxidative Stress: Post-Transcriptional Regulation of SodA Expression. *Microbiology* **2002**, *148* (4), 1027–1037. <https://doi.org/10.1099/00221287-148-4-1027>.
- (237) Bradbury, A. J.; Gruer, M. J.; Rudd, K. E.; Guest, J. R. The Second Aconitase (AcnB) of Escherichia Coli. *Microbiology* **1996**, *142* (2), 389–400. <https://doi.org/10.1099/13500872-142-2-389>.
- (238) Williams, C. H.; Stillman, T. J.; Barynin, V. V.; Sedelnikova, S. E.; Tang, Y.; Green, J.; Guest, J. R.; Artymiuk, P. J. E. Coli Aconitase B Structure Reveals a HEAT-like Domain with Implications for Protein-Protein Recognition. *Nat. Struct. Biol.* **2002**, *9* (6), 447–452. <https://doi.org/10.1038/nsb801>.
- (239) Chen, F.; Lukat, P.; Iqbal, A. A.; Saile, K.; Kaefer, V.; van den Heuvel, J.; Blankenfeldt, W.; Büssow, K.; Pessler, F. Crystal Structure of Cis-Aconitate Decarboxylase Reveals the Impact of Naturally Occurring Human Mutations on Itaconate Synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (41), 20644–20654. <https://doi.org/10.1073/pnas.1908770116>.
- (240) Lauble, H.; Kennedy, M. C.; Beinert, H.; Stout, C. D. Crystal Structures of Aconitase with Trans-Aconitate and Nitro-citrate Bound. *J. Mol. Biol.* **1994**, *237* (4), 437–451. <https://doi.org/10.1006/jmbi.1994.1246>.
- (241) Berg, J.; Tymoczko, J.; Stryer, L. *Biochemistry*, 5th Editio.; W.H Freeman and Company: New York, 2002.
- (242) Tsukamoto, Y.; Fukushima, Y.; Hara, S.; Hisabori, T. Redox Control of the Activity of Phosphoglycerate Kinase in Synechocystis Sp. PCC6803. *Plant Cell Physiol.* **2013**, *54* (4), 484–491. <https://doi.org/10.1093/pcp/pct002>.
- (243) Zerrad, L.; Merli, A.; Schröder, G. F.; Varga, A.; Grácz, É.; Pernot, P.; Round, A.; Vas,

- M.; Bowler, M. W. A Spring-Loaded Release Mechanism Regulates Domain Movement and Catalysis in Phosphoglycerate Kinase. *J. Biol. Chem.* **2011**, *286* (16), 14040–14048. <https://doi.org/10.1074/jbc.M110.206813>.
- (244) Goldstein, B. P. Resistance to Rifampicin: A Review. *J. Antibiot. (Tokyo)*. **2014**, *67* (9), 625–630. <https://doi.org/10.1038/ja.2014.107>.
- (245) Miller, W. T.; Schimmel, P.; Hill, K. A. W. Evidence for a “Cysteine-Histidine Box” Metal-Binding Site in an Escherichia Coli Aminoacyl-TRNA Synthetase. *Biochemistry* **1991**, *30* (28), 6970–6976. <https://doi.org/10.1021/bi00242a023>.
- (246) de Souza Lima, B. S.; Esteves, B. B.; Fialho-Júnior, L. C.; de Oliveira Mendes, T. A.; da Fonseca Pires, S.; Chapeourouge, A.; Perales, J.; de Andrade, H. M. Study of the Differentially Abundant Proteins among Leishmania Amazonensis, L. Braziliensis, and L. Infantum. *PLoS One* **2020**, *15* (10), e0240612. <https://doi.org/10.1371/journal.pone.0240612>.
- (247) Sardar, A. H.; Jardim, A.; Ghosh, A. K.; Mandal, A.; Das, S.; Saini, S.; Abhishek, K.; Singh, R.; Verma, S.; Kumar, A.; et al. Genetic Manipulation of Leishmania Donovanii to Explore the Involvement of Argininosuccinate Synthase in Oxidative Stress Management. *PLoS Negl. Trop. Dis.* **2016**, *10* (3), e0004308. <https://doi.org/10.1371/journal.pntd.0004308>.
- (248) Ali, D.; Abbady, A. Q.; Kweider, M.; Soukkarieh, C. Cloning, Expression, Purification and Characterization of Leishmania Tropica PDI-2 Protein. *Open Life Sci.* **2016**, *11* (1), 166–176. <https://doi.org/10.1515/biol-2016-0022>.
- (249) Silva-Almeida, M.; Souza-Silva, F.; Pereira, B. A. S.; Ribeiro-Guimarães, M. L.; Alves, C. R. Overview of the Organization of Protease Genes in the Genome of Leishmania Spp. *Parasites and Vectors* **2014**, *7* (1), 1–7. <https://doi.org/10.1186/1756-3305-7-387>.
- (250) Mahmoudzadeh-Niknam, H.; McKerrow, J. H. Leishmania Tropica: Cysteine Proteases Are Essential for Growth and Pathogenicity. *Exp. Parasitol.* **2004**, *106* (3–4), 158–163. <https://doi.org/10.1016/j.exppara.2004.03.005>.
- (251) Yao, C.; Donelson, J. E.; Wilson, M. E. Internal and Surface-Localized Major Surface Proteases of Leishmania Spp. and Their Differential Release from Promastigotes. *Eukaryot. Cell* **2007**, *6* (10), 1905–1912. <https://doi.org/10.1128/EC.00073-07>.
- (252) Splittstoesser, D. F.; Wilkison, M. Some Factors Affecting the Activity of Diethylpyrocarbonate as a Sterilant. *Appl. Microbiol.* **1973**, *25* (6), 853–857.
- (253) Fu, Y.; Hsieh, T. C.; Guo, J.; Kunicki, J.; Lee, M. Y. W. T.; Darzynkiewicz, Z.; Wu, J. M. Licochalcone-A, a Novel Flavonoid Isolated from Licorice Root (Glycyrrhiza Glabra), Causes G2 and Late-G1 Arrests in Androgen-Independent PC-3 Prostate Cancer Cells. *Biochem. Biophys. Res. Commun.* **2004**, *322* (1), 263–270. <https://doi.org/10.1016/j.bbrc.2004.07.094>.
- (254) de Mello, T. F. P.; Bitencourt, H. R.; Pedroso, R. B.; Aristides, S. M. A.; Lonardoni, M. V. C.; Silveira, T. G. V. Leishmanicidal Activity of Synthetic Chalcones in Leishmania (Viannia) Braziliensis. *Exp. Parasitol.* **2014**, *136*, 27–34.

<https://doi.org/10.1016/j.exppara.2013.11.003>.

- (255) Boeck, P.; Bandeira Falcão, C. A.; Leal, P. C.; Yunes, R. A.; Filho, V. C.; Torres-Santos, E. C.; Rossi-Bergmann, B. Synthesis of Chalcone Analogues with Increased Antileishmanial Activity. *Bioorganic Med. Chem.* **2006**, *14* (5), 1538–1545. <https://doi.org/10.1016/j.bmc.2005.10.005>.
- (256) Zhuang, C.; Zhang, W.; Sheng, C.; Zhang, W.; Xing, C.; Miao, Z. Chalcone: A Privileged Structure in Medicinal Chemistry. *Chem. Rev.* **2017**, *117* (12), 7762–7810. <https://doi.org/10.1021/acs.chemrev.7b00020>.