

AN ACTIVITY-BASED PROBE FOR THE ACYL PROTEIN THIOESTERASES

BY

YIMING CHEN

B. S., UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA, 2009

M. A., BROWN UNIVERSITY, 2012

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN THE  
DEPARTMENT OF CHEMISTRY AT BROWN UNIVERSITY

PROVIDENCE, RHODE ISLAND

MAY 2016

© Copyright 2016 by

Yiming Chen



This dissertation by Yiming Chen is accepted in its present form by the Department of Chemistry at satisfying the dissertation requirement for the degree of

Doctor of Philosophy

Date \_\_\_\_\_

Dr. Christopher T. Seto, Advisor

Recommended to the Graduate Council by

Date \_\_\_\_\_

Dr. Jason K. Sello, Reader

Date \_\_\_\_\_

Dr. Paul G. Williard, Reader

Approved by the Graduate Council

Date \_\_\_\_\_

Dr. Peter Weber  
Dean of the Graduate School

## Curriculum Vitae

Yiming Chen was born on January 16<sup>th</sup>, 1987 in Lanzhou, Gansu Province, China. He started his elementary education at Anxilu Elementary School (安西路小学) in Lanzhou. He then attended The Middle School Attached to Northwestern Normal University (西北师范大学附属中学), now renamed to The High School Attached to Northwestern Normal University), and graduated in 2005. Yiming began his undergraduate studies in 2005, pursuing degrees in the Department of Chemistry at the University of Science and Technology of China (USTC) in Hefei, Anhui Province. Upon completion of his Bachelor's degree in 2009, Yiming began graduate studies toward a Ph. D. degree in the Department of Chemistry at University of Nebraska - Lincoln, and transferred to Brown University in 2010. His graduate research was conducted under the supervision of Dr. Christopher T. Seto, and was concerned with the synthesis and application of an activity-based probe for acyl protein thioesterases (APTs) using in protein profiling of complex biological systems. Yiming has or will have co-authored the following publications:

1. Chen, Y., Zompa M. A., Bisiewicz, R. H., Seto, C. T. "Development of an Activity-Based Probe for Acyl Protein Thioesterases." (In preparation)
2. Chen, Y., Li, H. "Intermolecular Interaction in Water Hexamer", *J. Phys. Chem. A*, **2010**, *114* (43), pp 11719–11724
3. Chen, Y., Shi, J., Rui, L., Guo, Q. "Theoretical study on C<sub>32</sub> fullerenes and their endohedral complexes with noble gas atoms", *J. of Mol. Struct.: THEOCHEM*, **2009**, *907* (1–3), pp 104–108
4. Zompa, M. A., Chen, Y., Bisiewicz, R. H., Seto, C. T. "Activity-Based Labeling of Serine/Threonine Phosphatases via a Spontaneous Cyclization Mechanism." (In preparation)

## Acknowledgments

Though only my name appears on the cover of this dissertation, a great many people have contributed to its production. I owe my gratitude to all those people who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever.

My deepest gratitude is to my advisor, Prof. Christopher Seto. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. Prof. Seto taught me how to question thoughts and express ideas. His patience and support helped me overcome many crisis situations and finish this dissertation.

My thesis committee members, Prof. Paul Williard and Prof. Jason Sello, have been always there to listen and give me great advices. I am deeply grateful to them for their insightful comments and encouragement, but also for the hard question which incentivized me to widen my research from various perspectives. I am also especially grateful to Dr. Tun-Li Shen, who over the years, taught me an incredible amount about the mass spectrometry.

I also would like to thank my laboratory colleagues, not only for introducing me to organic chemistry techniques, but also for contributing very largely to a nice and friendly working atmosphere, especially Dr. Michael Zompa, Xuefeng Guo and Holly Sorenson. I would like also to thank the undergraduate students who worked in the group, especially Rachel Bisiewicz and Rosario Dispenza, for the nice time spend together. Furthermore, I would like to thank all the members of the Cane Group, for their help in both experiments and instrumentations. Especially, I want to thank Dr. Wayne Chou for teaching me a great deal of biological experiment techniques and answering my questions. Without their help, I

cannot finish my research project.

Many friends have helped me stay sane through these difficult years. Their support and care helped me overcome setbacks and stay focused on my graduate study. I would like to thank Marius Dräger, Jie Guang, Jian He, Ji Huang, Helene Kuhn, Philip Lukulay, Liheng Wu, Yan Yang and Rui Zhang. I greatly value their friendship and I deeply appreciate their belief in me.

I also want to give my thankfulness to Jorge Cham, the author of *Piled Higher and Deeper* comics and director of *The PHD Movie*. His works comforted me in Graduate School and helped me killing lots of time when I procrastinated.

Most importantly, none of this would have been possible without the love and patience of my family. My wife, Weili Li, has been a constant source of love, concern, support and strength all these years. I also would like to express my heart-felt gratitude to my parents for their love and support for the years.

## Table of Contents

|               |                                                                                  |     |
|---------------|----------------------------------------------------------------------------------|-----|
| Chapter One   | Background Introduction .....                                                    | 1   |
| 1             | The Background of Cancer .....                                                   | 3   |
| 2             | The Rat Sarcoma (Ras) Protein Family .....                                       | 4   |
| 3             | The Palmitoylation Cycle of Ras .....                                            | 5   |
| 4             | The Acyl Protein Thioesterases (APTs) .....                                      | 8   |
| 5             | Ras Depalmitoylation as a Molecular Target for Potential Antitumor<br>Drugs..... | 9   |
| 6             | Activity-Based Protein Profiling.....                                            | 11  |
| 7             | References.....                                                                  | 23  |
| Chapter Two   | Design and Synthesis of the ABP for APTs .....                                   | 35  |
| 1             | Abstract.....                                                                    | 37  |
| 2             | Design of the ABP .....                                                          | 37  |
| 3             | Validation of ABP's design by Molecular Dynamics Simulation .....                | 39  |
| 4             | Synthesis of the ABP .....                                                       | 42  |
| 5             | Synthesis and Characterization of Compounds .....                                | 49  |
| 6             | NMR Spectra and Mass Spectra .....                                               | 64  |
| 7             | References.....                                                                  | 104 |
| Chapter Three | Validation of the ABP via Kinetic Assay and In-gel Fluorescence...               | 107 |
| 1             | Abstract.....                                                                    | 109 |

|                                                                            |                                                                                                    |     |
|----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|-----|
| 2                                                                          | Measurement of the $K_m$ and $V_{max}$ of APT1 using the <i>p</i> NPO Colorimetric Substrate ..... | 110 |
| 3                                                                          | A Synthetic Fluorogenic Substrate for APT Activity Assays.....                                     | 114 |
| 4                                                                          | APT1 Inhibition Kinetic Assay.....                                                                 | 118 |
| 5                                                                          | In-gel Fluorescence of Fluorescently Labeled APTs.....                                             | 123 |
| 6                                                                          | References.....                                                                                    | 132 |
| Chapter Four Using Mass Spectrometry to Investigate APT-ABP Labeling ..... |                                                                                                    | 135 |
| 1                                                                          | Abstract .....                                                                                     | 137 |
| 2                                                                          | Introduction to the Protein Mass Spectrometry .....                                                | 137 |
| 3                                                                          | Whole Protein ESI-MS Validates the Labeling on APTs.....                                           | 144 |
| 4                                                                          | Using Tandem Mass Spectrometry to Reveal the Labeling Site .....                                   | 154 |
| 5                                                                          | References.....                                                                                    | 159 |
| Chapter Five Using the ABP to Profile Thioesterase Activity in Cells.....  |                                                                                                    | 165 |
| 1                                                                          | Abstract .....                                                                                     | 167 |
| 2                                                                          | Experiment Design.....                                                                             | 167 |
| 3                                                                          | HeLa Cell Culture .....                                                                            | 169 |
| 4                                                                          | Cell Proteome Labeling Experiment.....                                                             | 171 |
| 5                                                                          | References.....                                                                                    | 180 |
| Appendix I: Identified Cellular Proteins Labeled by ABP .....              |                                                                                                    | 183 |

## List of Schemes

|                                                                                                                         |    |
|-------------------------------------------------------------------------------------------------------------------------|----|
| <b>Scheme 1.1</b> Mechanism for the palmitoylation of a cysteine residue catalyzed by<br>PATs .....                     | 6  |
| <b>Scheme 1.2</b> First ABP developed for the APTs .....                                                                | 16 |
| <b>Scheme 1.3</b> Several activity-based probes developed for the serine hydrolases.....                                | 18 |
| <b>Scheme 1.4</b> The mechanism of the fluorophosphonate-based ABP for serine<br>hydrolases .....                       | 19 |
| <b>Scheme 1.5</b> Reaction mechanisms of the reactive warheads utilized in the ABPs for<br>cysteine proteases. ....     | 19 |
| <b>Scheme 1.6</b> Mechanism of the ABP for histone deacetylases based on the inhibitor<br>suberanolhydroxamic acid..... | 20 |
| <b>Scheme 1.7</b> ABP for the matrix metalloproteinase based on the known reversible<br>inhibitor ilomastat.....        | 21 |
| <b>Scheme 1.8</b> Examples of ABPs for phosphatases and their mechanism of actions ....                                 | 23 |
| <b>Scheme 2.1</b> Design of the clickable ABP for APTs.....                                                             | 38 |
| <b>Scheme 2.2</b> Proposed reaction mechanism for the ABP designed for APTs .....                                       | 39 |
| <b>Scheme 2.3</b> Retrosynthetic analysis of the ABP .....                                                              | 42 |
| <b>Scheme 2.4</b> Retrosynthesis of Synthon A .....                                                                     | 43 |
| <b>Scheme 2.5</b> Retrosynthesis of Synthon B.....                                                                      | 43 |
| <b>Scheme 2.6</b> Synthesis of Synthon A.....                                                                           | 44 |
| <b>Scheme 2.7</b> Selective protection of the benzylic hydroxyl group .....                                             | 45 |
| <b>Scheme 2.8</b> Synthesis of Synthon B.....                                                                           | 46 |
| <b>Scheme 2.9</b> The intramolecular acyl-transfer reaction of the <i>O</i> -acyl <i>L</i> -serine.....                 | 46 |

|                                                                                                                                                      |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>Scheme 2.10</b> Linkage of Synthon A with Synthon B.....                                                                                          | 47  |
| <b>Scheme 2.11</b> Acylation, deprotection and deoxyfluorination of the carbamate<br>intermediate .....                                              | 48  |
| <b>Scheme 2.12</b> Syntheses of the biotin-azide tag .....                                                                                           | 49  |
| <b>Scheme 3.1</b> The enzyme-catalyzed hydrolysis of <i>p</i> -nitrophenyl octanoate .....                                                           | 111 |
| <b>Scheme 3.2</b> Synthesis of the 4-methylumbelliferyl octanoate (4-MUO) .....                                                                      | 115 |
| <b>Scheme 3.3</b> Proposed mechanism of the copper (I)-catalyzed click chemistry .....                                                               | 124 |
| <b>Scheme 3.4</b> Synthesis of the tris [(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine ....                                                           | 127 |
| <b>Scheme 4.1</b> The mechanism of <i>N</i> -terminal gluconoylation which introduces the mass<br>shift of +178 Da in some recombinant proteins..... | 148 |
| <b>Scheme 5.1</b> Acyl-Coenzyme A thioesterase 7 (ACOT7) catalyzes the hydrolysis of<br>palmitoyl coenzyme A thioester .....                         | 178 |
| <b>Scheme 5.2</b> The hydrolysis reaction of <i>S</i> -formylglutathione catalyzed by<br><i>S</i> -formylglutathione hydrolases .....                | 179 |



## List of Figures

|                                                                                                                       |     |
|-----------------------------------------------------------------------------------------------------------------------|-----|
| <b>Figure 1.1</b> Comparison of cell division between normal cells and cancer cells .....                             | 3   |
| <b>Figure 1.2</b> Regulation of Ras activity by the GAPs and GEFs .....                                               | 4   |
| <b>Figure 1.3</b> The palmitoylation cycle of the Ras protein.....                                                    | 7   |
| <b>Figure 1.4</b> The crystal structure of APT1 .....                                                                 | 8   |
| <b>Figure 1.5</b> Structures of the reported APT1/APT2 inhibitors.....                                                | 10  |
| <b>Figure 1.6</b> The concepts of an activity-based probe (ABP) and activity-based protein<br>profiling (ABPP). ..... | 12  |
| <b>Figure 1.7</b> Example of a “tag-free” ABP .....                                                                   | 13  |
| <b>Figure 1.8</b> Three analysis platforms for the ABP-labeled proteins.....                                          | 15  |
| <b>Figure 2.1</b> The structure of the energy-minimized complex in the molecular dynamics<br>simulation.....          | 41  |
| <b>Figure 2.2</b> The RMSD and potential energies versus simulation time .....                                        | 42  |
| <b>Figure 3.1</b> The standard curve for <i>p</i> NP .....                                                            | 112 |
| <b>Figure 3.2</b> Data for <i>p</i> -nitrophenyl octanoate hydrolysis by 150 nM APT1 .....                            | 113 |
| <b>Figure 3.3</b> Standard curve of 4-MU fluorescence intensity .....                                                 | 116 |
| <b>Figure 3.4</b> The rate/concentration plots for APT1 and APT2.....                                                 | 117 |
| <b>Figure 3.5</b> Concept of the inhibition kinetic assay.....                                                        | 121 |
| <b>Figure 3.6</b> The inhibition of APT1 by the ABP .....                                                             | 122 |
| <b>Figure 3.7</b> The concept of in-gel fluorescence detection of APT activity .....                                  | 126 |
| <b>Figure 3.8</b> In-gel fluorescent labeling experiment result of APT1 .....                                         | 129 |
| <b>Figure 3.9</b> Concentration-dependent fluorescence labeling of APT1 .....                                         | 130 |

|                                                                                                             |     |
|-------------------------------------------------------------------------------------------------------------|-----|
| <b>Figure 3.10</b> In-gel fluorescence experiment using APT2 .....                                          | 131 |
| <b>Figure 3.11</b> Concentration-dependent fluorescence labeling of APT2.....                               | 132 |
| <b>Figure 4.1</b> Mechanism of electrospray ionization.....                                                 | 140 |
| <b>Figure 4.2</b> The Principle of matrix-assisted laser desorption/ionization .....                        | 141 |
| <b>Figure 4.3</b> The mass spectra of mouse recombinant APT1 protein .....                                  | 147 |
| <b>Figure 4.4</b> The mass spectra of human recombinant APT2 protein.....                                   | 149 |
| <b>Figure 4.5</b> Deconvoluted ESI-MS spectra of APT1 .....                                                 | 151 |
| <b>Figure 4.6</b> Deconvoluted ESI-MS spectra of APT2 .....                                                 | 152 |
| <b>Figure 4.7</b> Co-incubation of the ABP with cytochrome C and APT1 .....                                 | 154 |
| <b>Figure 4.8</b> Principle of tandem mass spectrometry in peptide sequencing.....                          | 155 |
| <b>Figure 4.9</b> Common fragmentation patterns observed for a tetrapeptide .....                           | 156 |
| <b>Figure 4.10</b> The tandem mass spectrum of the ABP-labeled peptide of APT1 .....                        | 158 |
| <b>Figure 4.11</b> The tandem mass spectrum of the <i>N</i> -terminal peptide from His-tagged<br>APT1 ..... | 159 |
| <b>Figure 5.1</b> Experiment design of the ABPP labeling experiments .....                                  | 168 |
| <b>Figure 5.2</b> Microscopic images of HeLa cells at different densities.....                              | 170 |
| <b>Figure 5.3</b> In-gel fluorescence analysis of the ABP-labeled cellular proteome.....                    | 173 |
| <b>Figure 5.4</b> Quaternary structure of ACOT7 .....                                                       | 176 |
| <b>Figure 5.5</b> Binding of the Coenzyme A within the active site of ACOT7.....                            | 177 |

## List of Tables

|                                                                                                                                                      |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>Table 3.1</b> Conditions and results of the APT1 $K_m$ and $V_{max}$ measurement.....                                                             | 114 |
| <b>Table 3.2</b> The $K_m$ and $V_{max}$ of APT1 using <i>p</i> NPPO as the substrate.....                                                           | 114 |
| <b>Table 3.3</b> The $V_{max}$ and $K_m$ of APT1 and APT2 using 4-MUO as the substrate .....                                                         | 117 |
| <b>Table 3.4a</b> The apparent rate constants obtained from the incubations .....                                                                    | 123 |
| <b>Table 3.4b</b> The inhibition constant ( $K_I$ ) and inactivation rate constant ( $k_i$ ) of the<br>ABP for APT1 .....                            | 123 |
| <b>Table 4.1</b> The common matrices for analysis of large biological molecule by<br>MALDI.....                                                      | 142 |
| <b>Table 4.2</b> The sequences and the calculated molecular weights of the AP1 and APT2<br>proteins .....                                            | 145 |
| <b>Table 5.1</b> The enzymes with known thioesterase activity which were identified as being<br>labeled by the ABP by tandem mass spectrometry ..... | 178 |
| <b>Table S1</b> The proteins from Mascot search with at least two unique peptides scoring<br>above 40.0.....                                         | 185 |

Dedicated to the great explorers  
who lost on the journeys to the paradise

*Abstract of An Activity-Based Probe for the Acyl Protein Thioesterases, by Yiming Chen, Ph.D., Brown University, May 2016*

The S-palmitoylated signal-transducing rat sarcoma GTPase (Ras), which plays important roles in cellular growth, division, and differentiation, is frequently mutated in various types of cancer. Dynamic palmitoylation and depalmitoylation of Ras confers proper localization and signaling in cells. So far, only two enzymes have been characterized as being involved in depalmitoylation: including acyl protein thioesterase 1 and acyl protein thioesterase 2. To decipher the physiological functions of the enzymes involved in S-depalmitoylation, we have designed a clickable activity-based probe for acyl protein thioesterases. The probe consists of a palmitoylated serine substrate linked to a trapping moiety that becomes reactive after enzymatic hydrolysis. This activity-based probe is critical for profiling the activities of acyl protein thioesterases in biological systems and for inhibitor screening.



# **CHAPTER ONE**

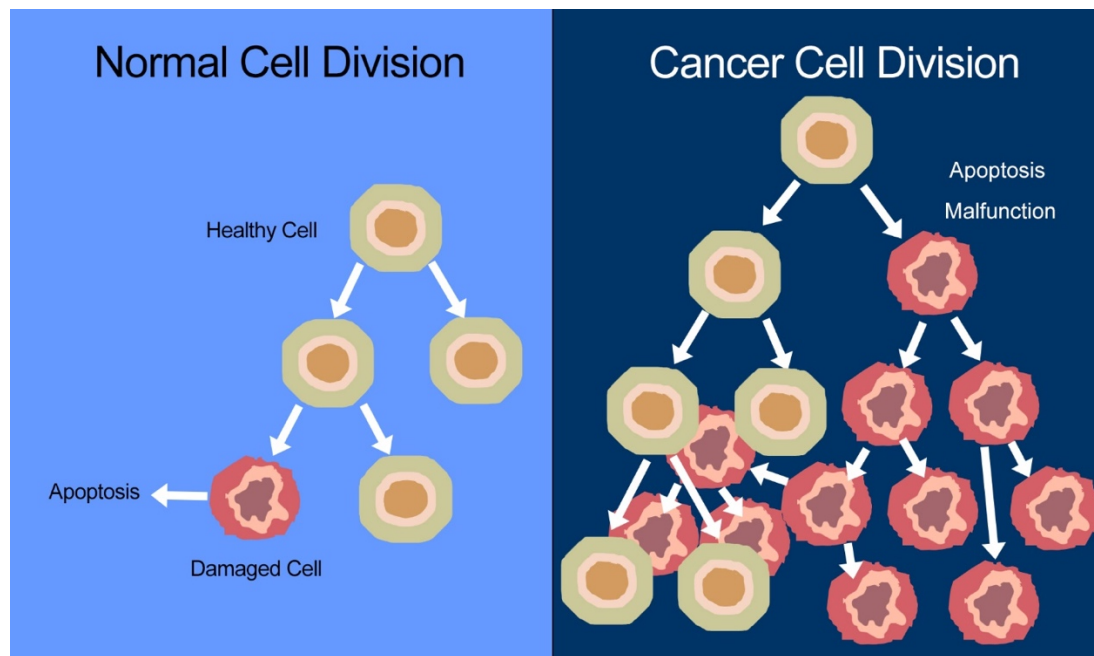
## Background Introduction





## 1. The Background of Cancer

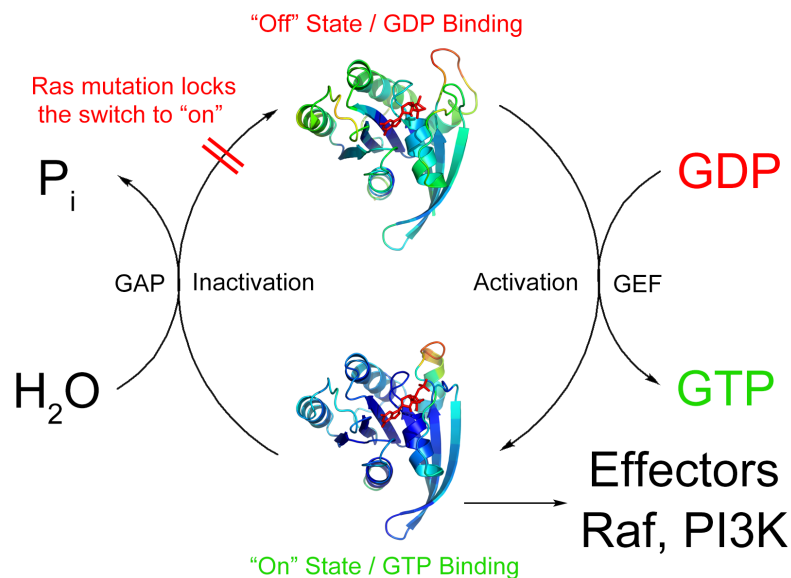
Cancer is a group of diseases involving abnormal cell growth that affects various tissues.<sup>1</sup> Cancer is characterized by an unregulated cell proliferation that invades and spreads to the other parts of the body, and often leads to the cancer-related deaths (Figure 1). A metastatic tumor can spread from its original site through the bloodstream or lymphatic circulation to distant sites. This metastasis is frequently encountered in the late stages of cancers and is associated with significantly reduced prognosis for survival.<sup>2</sup> Nowadays, based on estimates from the American Cancer Society,<sup>3</sup> cancer has become an enormous global health issue accounting for one in every eight deaths worldwide. In 2008, 12 million new cancer cases were diagnosed and 7 million people died from cancer globally.



**Figure 1.1 Comparison of cell division between normal cells and cancer cells.** The programmed apoptosis of a damaged cell occurs in the normal cell division cycles. However, when proto-oncogenes mutate into oncogenes involved in specific cellular signaling pathways, cancer will lead to unregulated cell division.

## 2. The Rat Sarcoma (Ras) Protein Family

Rat sarcoma (Ras) is a family of related proto-oncogene proteins <sup>4</sup> that is ubiquitously expressed in all cell lineages, and organs of mammalian animals. The Ras subfamily comprising the H-Ras, N-Ras, and K-Ras proteins forms a family of GTP-regulated cell signal-transducing molecular switches cycling between the inactive GDP-bound and active GTP-bound state. When Ras is “switched on” by the incoming signals, it subsequently activates other downstream molecular signal transducers, which ultimately enables the genes involved in cell growth, differentiation and survival. The reactivation of Ras requires replacement of bound GDP by GTP, catalyzed by the nucleotide exchange factors (GEF), whereas the inactivation of Ras occurs through its intrinsic GTPase activity, accelerated by the GTPase activating proteins (GAPs).



**Figure 1.2 Regulation of Ras activity by the GAPs and GEFs.** The GTP-binding Ras is switched to “on” status and enables the downstream signal transducers Raf and PI3K. GAP inactivates Ras by promoting the intrinsic hydrolysis of bound GTP to GDP. Then GEF activates the Ras again by stimulating the replacement of bound GDP by GTP. The mutation of Ras inhibits the intrinsic GTPase activity and locks the protein into the “on” state, which leads to unregulated cell proliferation.

Once Ras is activated, Ras proteins signal through several cell-signaling pathways involved in cell proliferation, for example, MAPK or PI3K. This aspect of Ras dynamic regulation can be deregulated in Ras mutations at the binding interface of the GTP-Ras-GAP complex, typically at residues Gly-12, Gly-13 and Gln-61. The mutation prevents the approach of the regulatory protein GAP.<sup>5</sup>

As a result, mutations in Ras genes can lead to the permanent activation of Ras proteins. This can cause overactive signaling, even without the incoming signals. Because these signals result in cell growth and division, permanently-activated Ras can ultimately lead to cancer.<sup>6</sup> The three Ras genes in humans (HRAS, KRAS, and NRAS) are the most common oncogenes in human cancer; mutations that permanently activate Ras are found in 25% of all tumors.<sup>7</sup> For this reason, Ras inhibitors are being investigated as treatments for related cancers.

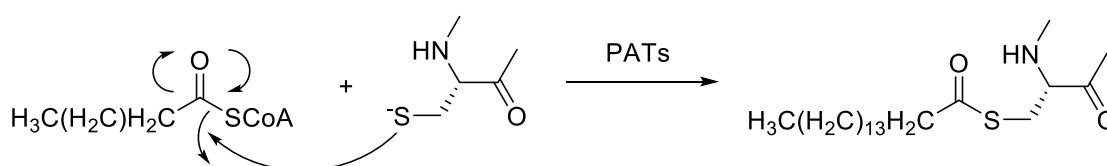
### 3. The Palmitoylation Cycle of Ras

Palmitoylation is a post-translational modification that incorporates a 16C palmitate moiety onto cysteine residues of a protein by a thioester linkage. This modification enhances the protein's overall hydrophobicity, which results in the protein being localized to the hydrophobic plasma membrane.<sup>8</sup> Palmitoylation plays significant roles in intracellular Ras signal propagation. Similar post-translational modifications are prenylation and myristoylation,<sup>9</sup> which localize proteins to membrane surfaces. However, in contrast to these other types of modification, a notable feature of palmitoylation is its reversibility.<sup>10</sup>

Attachment of fatty acid moieties to amino acid residues changes the properties and

localization of proteins in the cellular environment. If palmitoylation occurs on proteins that have also undergone some other lipidation, then localization of the protein to the plasma membrane is a result of the composite effects of both modifications. For example, the Sarcoma family of tyrosine kinases is modified by both *N*-myristoylation and palmitoylation with the addition of both lipids being required to localize the kinases to the plasma membrane.<sup>11</sup> Other examples are the H-Ras and N-Ras proteins. Both *S*-prenylation and palmitoylation are required for the localization of these proteins to the plasma membrane.

Palmitoylation is catalyzed by protein acetyltransferases (PATs), which contain a highly conserved DHHC (Asp-His-His-Cys) cysteine-rich domain. The palmitoylation reaction involves a nucleophilic attack by a thiolate ion ( $\text{S}^-$ ) of the target cysteine residue on the carbonyl carbon of palmitoyl coenzyme A (Scheme 1.1). Coenzyme A leaves as a leaving group, and the palmitate group is transferred to the cysteine residue by formation of a new thioester bond.

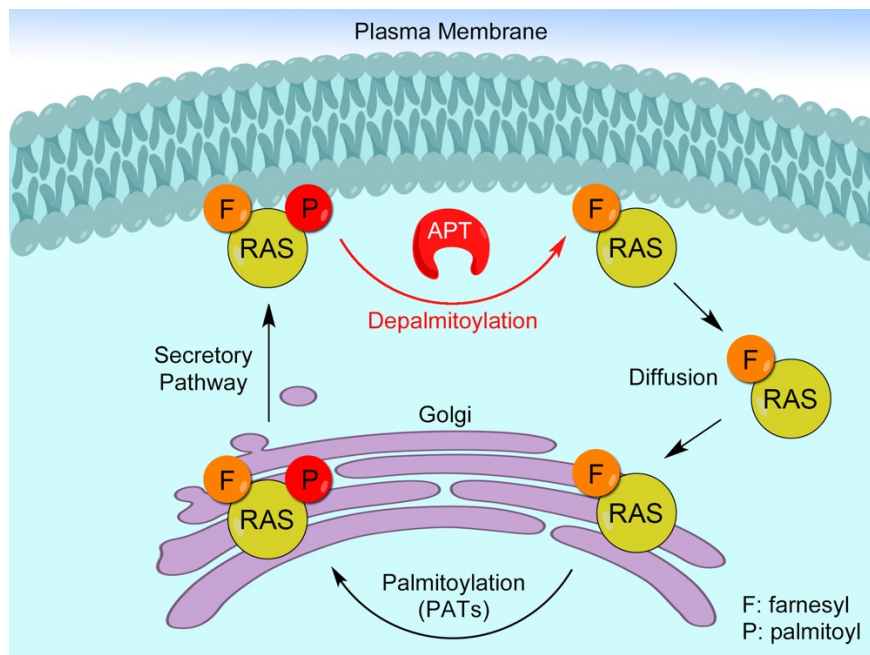


**Scheme 1.1 Mechanism for the palmitoylation of a cysteine residue catalyzed by PATs.**

Depalmitoylation is the reverse process of palmitoylation, and acts to remove the C16 fatty acid chains from Ras, which allows the release of Ras from the cell membrane for return to the Golgi apparatus.<sup>12</sup> This process is catalyzed by acyl protein thioesterases (APT<sub>s</sub>). While several palmitoyl transferases have been discovered, knowledge of the

depalmitoylation enzymes is still limited. Only two enzymes have been identified that catalyze protein depalmitoylation *in vivo*: acyl protein thioesterase 1 (APT1) and acyl protein thioesterase 2 (APT2).<sup>13</sup>

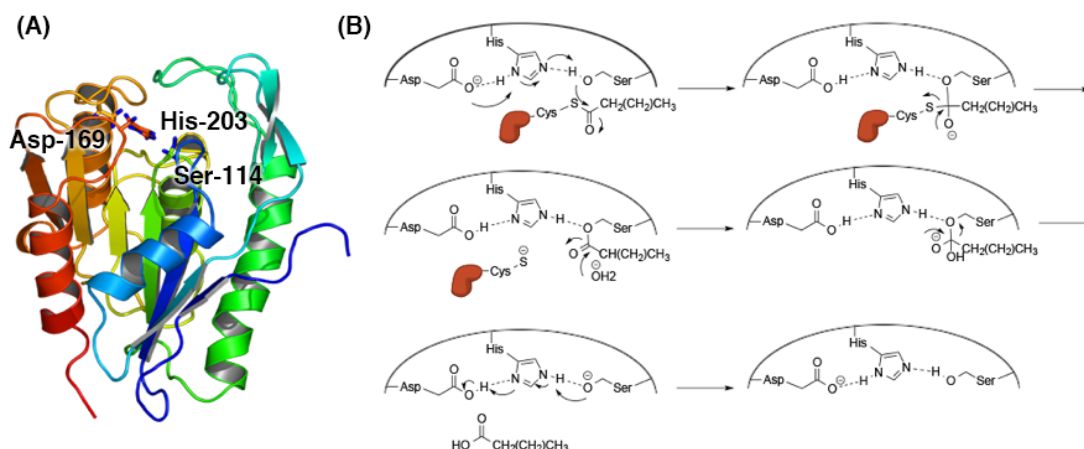
The roles of H-Ras and N-Ras proteins in the Ras proto-oncogene family might be the best examples for illustrating how the localization of proteins is dynamically mediated by palmitoylation (Figure 1.3).<sup>11b, 14</sup> Palmitoylation of H-Ras and N-Ras occurs at the Golgi apparatus and confers association of H/N-Ras with the plasma membrane by vesicular trafficking. There, depalmitoylation decreases the hydrophobicity of H-Ras and N-Ras, thus returning them to the cytoplasm. They then return to the Golgi apparatus for the next round of palmitoylation. Since the H/N-Ras proteins activate different downstream cellular signals, reversible palmitoylation cycles also regulate signaling via protein localization.



**Figure 1.3 The palmitoylation cycle of the Ras protein.**

#### 4. The Acyl Protein Thioesterases (APTs)

Acyl protein thioesterase 1 (APT1), also known as lysophospholipase 1 (LYPA1) is a 25 kDa cytosolic hydrolase first isolated from rat liver.<sup>13</sup> First described as a lysophospholipase, LYPA1 was subsequently renamed as acyl protein thioesterase1 (APT1) due to its stronger depalmitoylation activity. APT1 is the first member of the APT family discovered. The crystal structure of human APT1 was solved showing a catalytic triad consisting in Ser114, His203 and Asp169 (Figure 1.4).<sup>15</sup>



**Figure 1.4 The crystal structure of APT1.** (A) The catalytic triad (Ser-114, Asp-169, His-203) is presented as stick model. (B) The mechanism of the APT1 catalyzed hydrolysis.

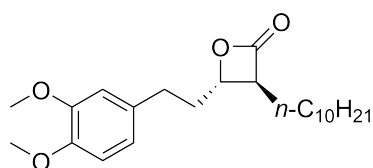
APT1 depalmitoylates *in vitro* a wide variety of substrates such as H-Ras proteins, G-protein subunits,<sup>13</sup> eNOs,<sup>16</sup> RGS4,<sup>16</sup> SNAP-23,<sup>17</sup> and ghrelin.<sup>18</sup> Acyl protein thioesterase 2 (APT2) was first purified from pig stomach.<sup>19</sup> Human APT2 revealed 66% sequence identity with human APT1, in particular with a conserved catalytic triad. In analogy to APT1, its close homologue APT2 has a highly conserved amino acid sequence among

various species. The crystal structure of APT2 is very similar to the structure of APT1 as shown by the superposition of their respective crystal structures. Recently, the depalmitoylation of growth-associated protein 43 (GAP-43) was observed in cells overexpressing APT2, suggesting that APT2 is probably involved in the GAP-43 depalmitoylation.<sup>20</sup> However, experimental evidence of the involvement of APT2 in Ras depalmitoylation *in vivo* is still lacking.

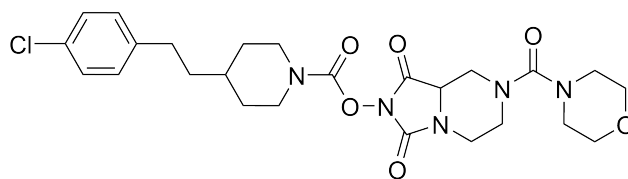
### 5. Ras Depalmitoylation as a Molecular Target for Potential Antitumor Drugs

Given the importance of Ras localization for biological function, a series of inhibitors targeting either Ras post-translational modifications or Ras depalmitoylation have been developed (Figure 1.5). The inhibition of APTs may represent an innovative method to disrupt aberrant oncogenic Ras signaling in tumor cells. To date several APT inhibitors have been reported, including the substrate-mimicking inhibitor Raspalin 3,<sup>21</sup> the  $\beta$ -lactone inhibitor Palmostatin B,<sup>22</sup> triazole urea inhibitor ML211,<sup>23</sup> and APT1/APT2 selective inhibitors ML348 / ML349 (Figure 1.5).<sup>24</sup>

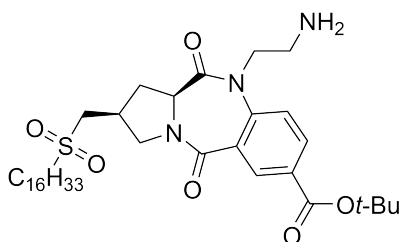
However, there is no method currently available to profile APT activity directly for *in vitro* and *in vivo* experiments, which makes the development of new APT inhibitors a significant challenge. To validate the potencies and specificities of APT inhibitors, a quick enzymatic profiling method is highly desirable.



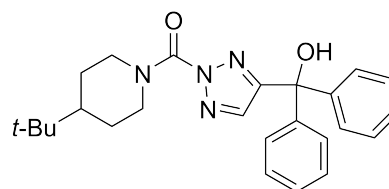
Palmostatin B  
 $IC_{50} = 0.67 \mu M$  (APT1)



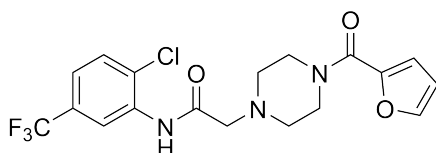
ML378  
 $IC_{50} = 87 \text{ nM}$  (APT1) /  $143 \text{ nM}$  (APT2)



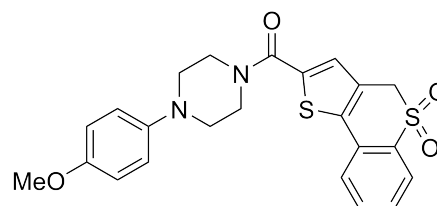
Raspalin 3  
 $IC_{50} = 148 \text{ nM}$  (APT1)



ML211  
 $IC_{50} = 17 \text{ nM}$  (APT1) /  $30 \text{ nM}$  (APT2)



ML348  
 $IC_{50} = 210 \text{ nM}$  (APT1) /  $> 3000 \text{ nM}$  (APT2)



ML349  
 $IC_{50} = > 3000 \text{ nM}$  (APT1) /  $144 \text{ nM}$  (APT2)

**Figure 1.5 Structures of the reported APT1/APT2 inhibitors.**

## 6. Activity-Based Protein Profiling

### 6.1. Introduction to activity-based protein profiling

Activity-based protein profiling (ABPP) is a proteomic technique that employs synthetic molecules to probe mechanistically related enzymes.<sup>25</sup> To assay enzyme activities in biological system, various methods have been developed. For example, genomic DNA screening reveals genes that are associated with particular enzymes, but cannot discern the transcription levels of the genes. In contrast, cDNA screening can reveal the existence of



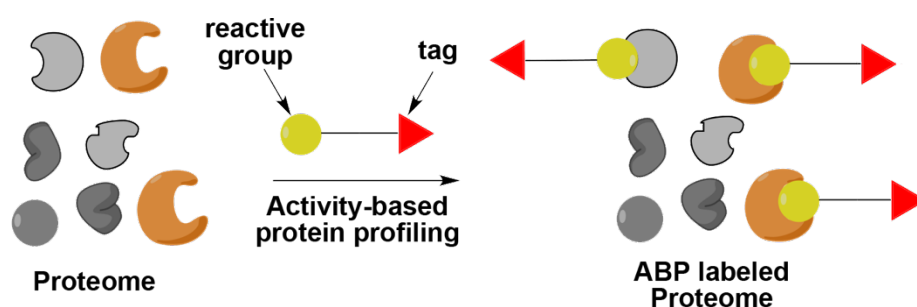
the transcriptome, but is not able to discern whether it has been translated. On the other hand, although protein expression investigations target the enzymes in the proteome, this technique cannot judge whether a protein is correctly folded and/or post-translational modified. For instance, immunoaffinity methods such as Western blotting or proteomic mass spectrometry can confirm the expression of a protein, but it cannot address its enzymatic activity directly. More complicated technologies, like gene knockout, attempt to directly profile an enzyme's role in a particular biological pathway, but the invasive nature of such studies is highly undesirable.

In contrast with these traditional protein profiling methods, activity-based protein profiling (ABPP) can provide a comprehensive view of targeted enzyme's activity in biological system with a minimum interference.<sup>25a, 26</sup>

## 6.2. Activity-based probe

The basic unit of ABPP is the activity-based probe (ABP), a synthetic molecule that typically consists of two essential components: an active-site-directed reactive group that reacts with and labels protein covalently based on its mechanism, and a tag that can be utilized to detect or isolate labeled protein (Figure 1.6). Some probes may also contain an extra binding component, which enhances selectivity by binding to specific site on a protein with high affinity. The reactive group usually includes a highly-reactive electrophile (or the precursor to such an electrophile), which can be covalently linked to a nucleophilic residue in the active site of an active target enzyme. If the enzyme is not active (for example, inactivated by inhibitor or not post-translational modified correctly), it will

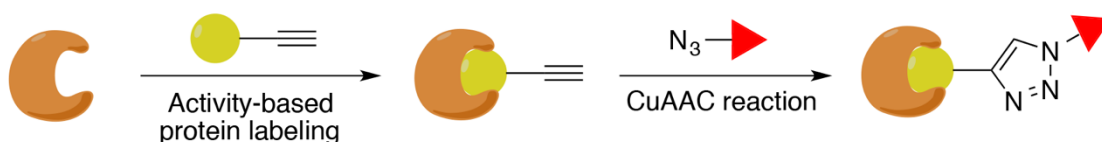
not be labeled by the ABP. As another crucial component of an ABP, the tag may be either a reporter such as a fluorophore, or an affinity label such as biotin.



**Figure 1.6 The concepts of an activity-based probe (ABP) and activity-based protein profiling (ABPP).** The ABP containing a reactive group and a tag reacts with proteins that have a specific catalytic activity to label them covalently.

### 6.3. “Tag-free” activity-based probes

In recent years, “tag-free” ABPs were developed. Instead of integrating the tag into the ABP, the “tag-free” ABP only has a relatively small conjugation site that can be linked to the tag in situ. The advantage of this innovative strategy is that: it avoids the possible problem of the bulky tag component interfering with enzyme labeling the the protein labeling step, and the ABP can be functionalized with greater versatility. According to this design, highly-efficient bioorthogonal conjugation reactions, such as the copper (I)-catalyzed alkyne-azide cycloaddition (CuAAC) reaction,<sup>27</sup> was utilized to attach the synthetic tag to the ABP-labeled proteins (Figure 1.7).



**Figure 1.7 Example of a “tag-free” ABP.** First, the ABP with terminal alkyne reacts with target protein in a mechanism-based manner to label it covalently. Subsequently, the azide tag is conjugated with the ABP-protein complex via the CuAAC reaction.

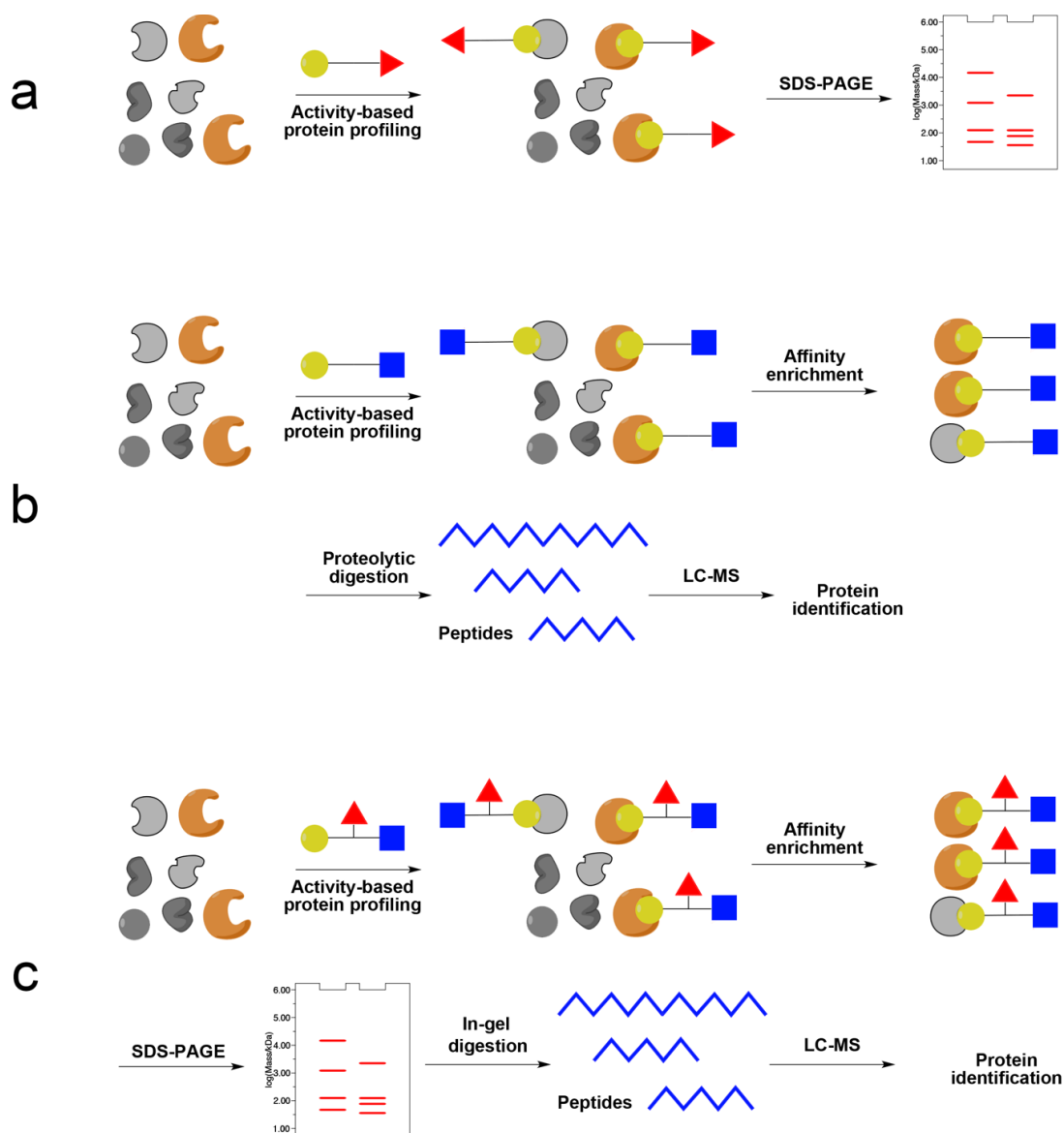
#### 6.4. Analysis of the ABP-labeled proteins

Currently there are two major platforms used to analyze the ABP-labeled proteins: gel electrophoresis and liquid chromatography-mass spectrometry (LC-MS). Initially ABPP was designed as a gel-based approach to profile enzyme activities. In this method, ABP-treated proteomes are subjected to polyacrylamide gel electrophoresis (PAGE) for resolving. The ABP-labeled proteins in the gel were visualized by either in-gel fluorescence or immunoblotting techniques (Figure 1.8a). Although other high-resolution methods have emerged in recent years, this gel-based protocol is still widely applied because of its robustness and efficiency.<sup>25b, 28</sup>

Another widely used analysis platform is liquid chromatography-mass spectrometry (LC-MS) (Figure 1.8b). Because it lacks of high resolution and the ability to identify multiple target proteins, the gel-based approach has limitations for analyzing complicated proteome samples. To address this issue, proteomic mass spectrometry was introduced. There are two steps in this method: first, the ABP is tagged with a biotin tag and immobilized biotin-binding proteins,<sup>25a, 29</sup> like streptavidin or avidin, are utilized to enrich the ABP-labeled proteins on a stationary phase. This affinity purification method is highly desirable since the unlabeled proteins in the proteome can be easily washed away from the target proteins bound to the stationary phase. Second, the enriched ABP-labeled proteins on the stationary

phase are proteolytically digested by proteases, such as trypsin or chymotrypsin. The resulting peptide fragments are analyzed by tandem mass spectrometry to give the identification information.

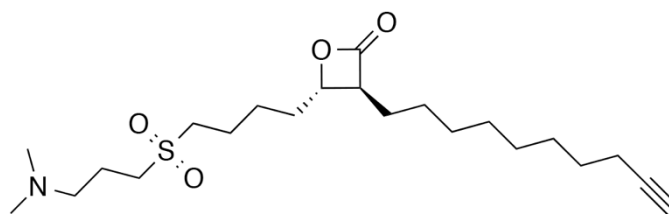
In addition, these two approaches can be combined to afford a new comprehensive analysis platform (Figure 1.8c). In this approach, a dual-functional tag which contains both reporter tag and affinity tag is employed. The ABP-labeled proteins are enriched and purified by an affinity resin or column; the resulting proteins are subjected to polyacrylamide gel electrophoresis. After the proteins in the gel are resolved and visualized, interested gel bands are excised and digested by proteases. The resulting peptides are extracted from the gel and subjected to tandem mass spectrometry analysis.<sup>30</sup> This combined platform makes full use of the advantages of both gel-based and proteomic mass spectrometry methods, and is highly desired in ABPP experiments.



**Figure 1.8 Three analysis platforms for the ABP-labeled proteins.** The yellow circle denotes the reactive group of the ABP; the red triangle denotes the reporter tag for visualization; the blue square denotes the affinity tag for enrichment. (a) Traditional gel-based method for visualizing ABP-labeled proteins. (b) LC-MS based method for identification of the ABP-labeled proteins. (c) A combined method which utilizes the advantages of both.

### 6.5. Activity-based probe for APTs

In 2011, the first ABP for APTs based on the  $\beta$ -lactone (Scheme 1.2) was synthesized.<sup>31</sup> This ABP has been successfully applied to assay APT activity in a biological system. However, there are three issues with this ABP, which need to be addressed. First, the  $\beta$ -lactone is a highly reactive species that can be rapidly hydrolyzed in aqueous solution at neutral pH. Second, the strong reducing reagent glutathione (GSH) exists in animal cells (5 mM) and can react with the  $\beta$ -lactone to form inactive lactathione. Third, although the  $\beta$ -lactone was considered as an irreversible inhibitor, it has been shown that the adduct formed between an enzyme and the  $\beta$ -lactone can be hydrolyzed, resulting in the detachment of the ABP and recovery of enzyme activity. For these reasons, the application of  $\beta$ -lactones as the reactive group in ABPs has been limited.<sup>32</sup>



**Scheme 1.2 First ABP developed for the APTs.** The ABP is based on a  $\beta$ -lactone inhibitor.

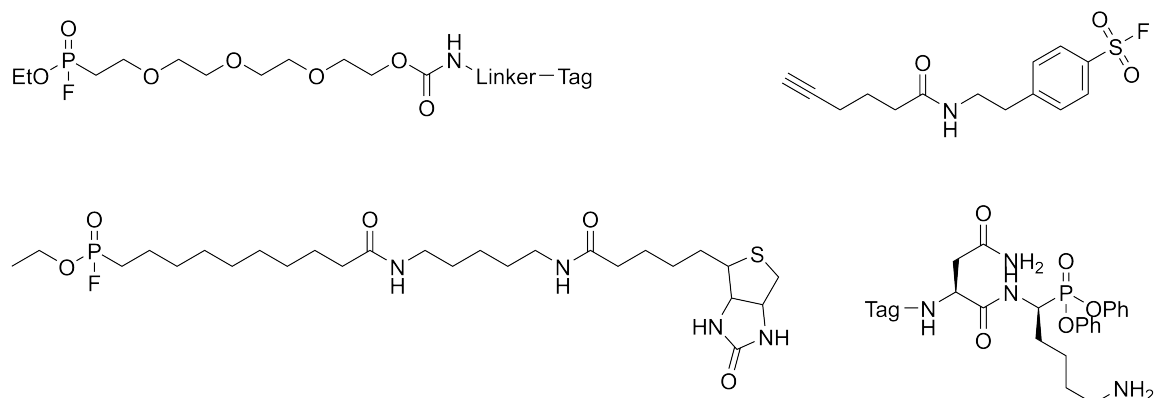
With the aim of overcoming these limitations, we have constructed a quinone methide-based ABP for APTs with a novel mechanism of action. This mechanism of action has allowed us to incorporate a substrate mimic of the natural substrate of the APTs. This ABP was synthesized and then validated using kinetic inactivation assays, gel-based ABPP methods and LC-MS based method. Finally, this ABP was applied to a complex biological system to assay the activity of thioesterases.

## 6.6. Activity-based probes for different classes of enzymes

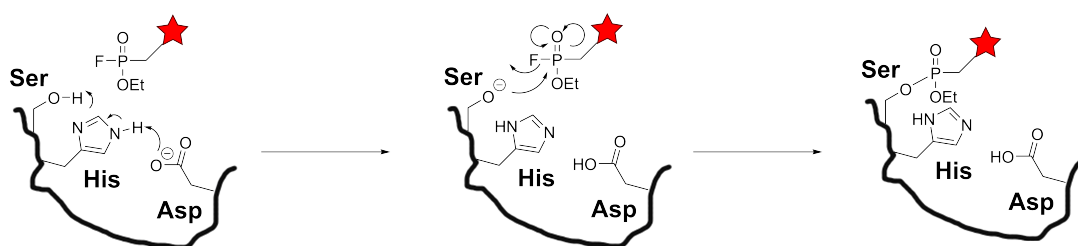
Effort have been made by scientists in recent decades to develop ABPs for various enzyme classes. With the help of the ABP technique, enzyme function in a wide range of biological systems has been investigated. In this section, several ABPs for different classes of enzymes are introduced.

### 6.6.1. ABP for serine hydrolases

Serine hydrolases are a diverse enzyme class containing proteases, lipases, esterases, thioesterases and peptidases.<sup>33</sup> These enzymes constitute about 1% of all the predicted proteins encoded by eukaryotic genomes. Most of the serine hydrolases share a similar catalytic mechanism involving a highly conserved serine amino acid residue for nucleophilic addition to electrophilic substrates, including esters, thioesters, peptides and proteins. For this reason, several ABPs containing an electrophilic reactive group have been developed to label the nucleophilic serine residue of serine hydrolases. The reactive functional groups applied in these ABPs includes fluorophosphonates, sulfonyl fluorides, and carbamates. For instances, fluorophosphonates exhibited reactivity with broad range of serine hydrolases (Scheme 1.3).<sup>25b, 28a, 34</sup> These ABPs have been applied in different biological systems to profile the serine hydrolase activity, including cancer,<sup>35</sup> atherosclerosis,<sup>36</sup> and nervous system signaling.<sup>37</sup> As an example, the reaction mechanism of the fluorophosphonate ABP is demonstrated in Scheme 1.4.



**Scheme 1.3** Several activity-based probes developed for the serine hydrolases. Fluorophosphonate, aryl phosphonate, or sulfonyl fluoride were used as the reactive components of the ABPs.



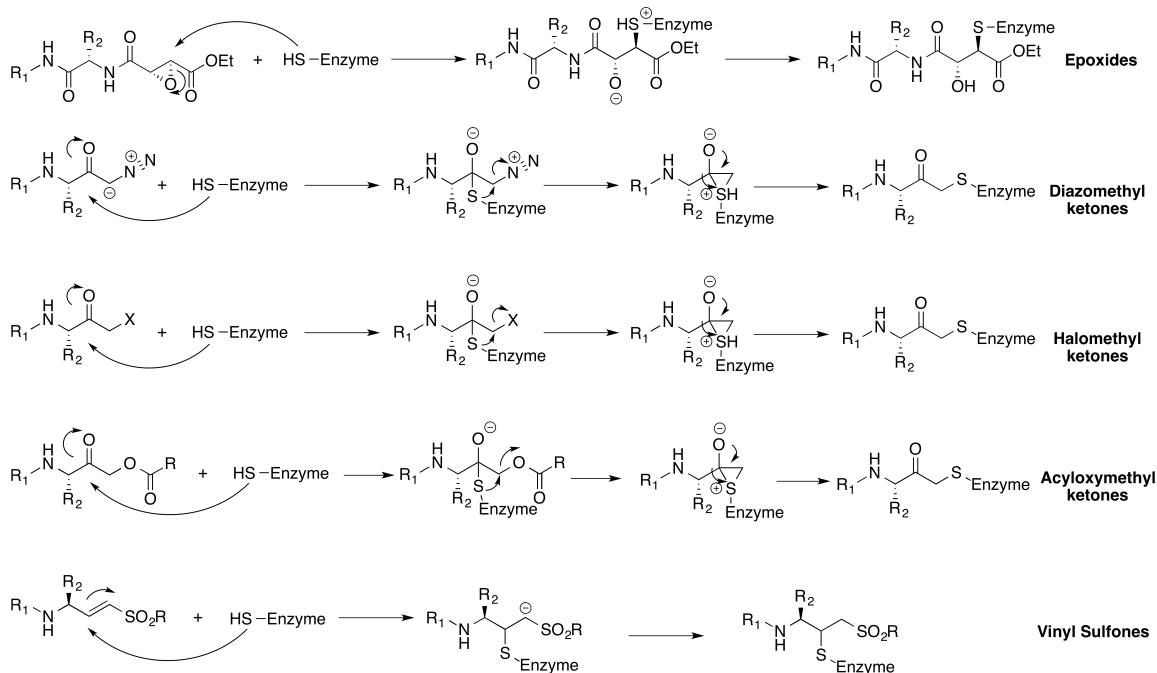
**Scheme 1.4** The mechanism of the fluorophosphonate-based ABP for serine hydrolases. The red star denotes the reporter tag linked to the fluorophosphonate reactive core.

### 6.6.2. ABPs for cysteine proteases

Similar to serine hydrolases, cysteine proteases are also a huge family of enzymes that perform peptide bond hydrolyzing reactions.<sup>38</sup> The mechanism of this type of hydrolysis reaction involves the attack by the nucleophilic cysteine residue in the active site, which enables the possibility to label the catalytic cysteine residue with several electrophiles. Because the nucleophilicity of cysteine is much stronger than that of serine, weaker electrophiles are usually employed in ABPs for cysteine proteases. Epoxides,<sup>39</sup> diazomethyl ketones,<sup>40</sup> halomethylketone,<sup>41</sup> acyloxymethyl ketones,<sup>42</sup> and vinyl sulfones,<sup>43</sup>



have been utilized as the reactive warheads of the ABPs. The mechanisms of few a ABPs for cysteine proteases are illustrated in Scheme 1.5.



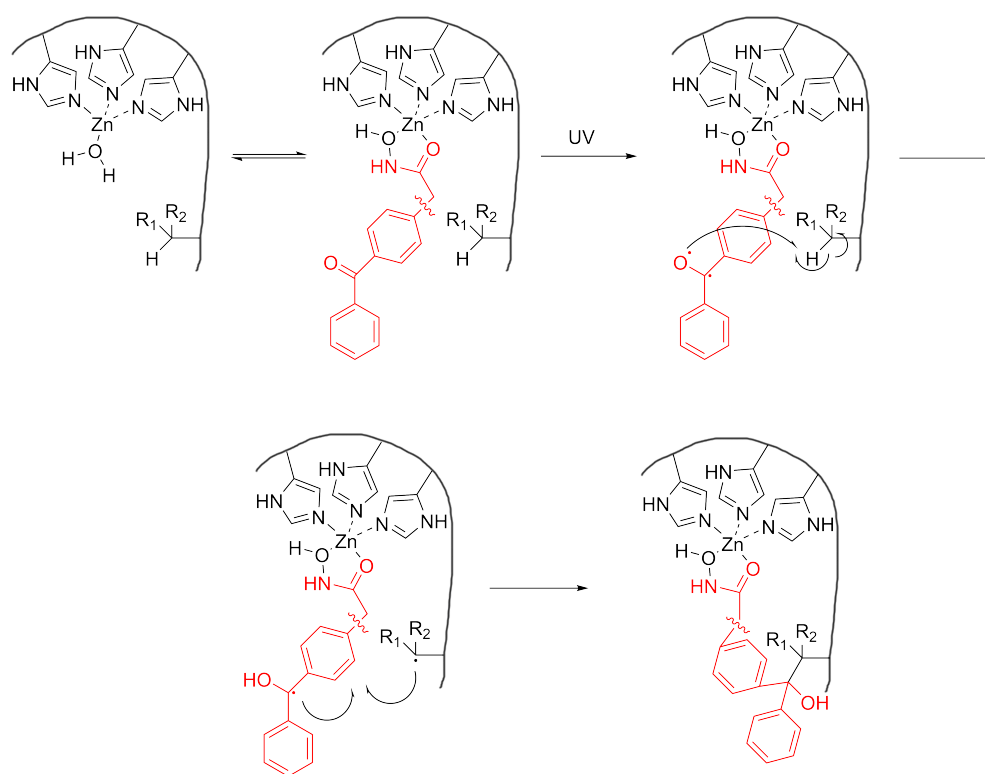
**Scheme 1.5** Reaction mechanisms of the reactive warheads utilized in ABPs for cysteine proteases.

### 6.6.3. ABPs for metallohydrolases

Besides the serine hydrolases family and cysteine proteases family, a third major category of hydrolase are the metallohydrolases. Different from the serine or cysteine hydrolases that use nucleophilic amino acid as catalytic residues, the metallohydrolases adopt a metal ion-activated water molecule in active sites to hydrolyze the substrate.<sup>44</sup> There are two major classes of metallohydrolases: metalloproteases and histone deacetylases. Due to the lack of catalytic nucleophile in the active sites, rather than mechanism-based electrophilic

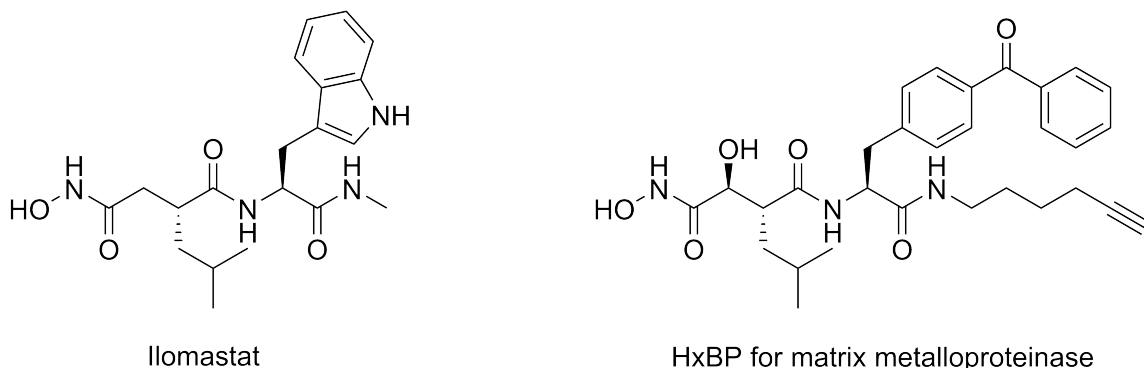
ABPs, several photoreactive ABPs have been developed to trap the enzyme by photo-induced reactions.

Based on the structure of suberanolhydroxamic acid, which is a known histone deacetylases inhibitor and has been approved by the FDA for the treatment of cutaneous T cell lymphoma, a photoreactive ABP for histone deacetylases has been developed.<sup>45</sup> The mechanism of this ABP is presented in Scheme 1.6. First, the hydroxamic acid group of the ABP chelates with the zinc ion in the active site, then the complex is exposed to UV radiation to generate a radical from benzophenone group *in situ* as a trapping device. Finally, the adjacent amino acid residue is linked with the ABP via a covalent bond.



**Scheme 1.6 Mechanism of the ABP for histone deacetylases based on the inhibitor suberanolhydroxamic acid.**

Another ABP for matrix metalloproteinases was also developed based on a reversible inhibitor ilomastat, which is also a hydroxamic acid (Scheme 1.7). This ABP selectively labels only the active form of the matrix metalloproteinases, but not the zymogen or endogenous inhibitor-bound enzymes.<sup>46</sup> The mechanism of this ABP is similar to the probe for the histone deacetylases described above.



**Scheme 1.7 ABP for the matrix metalloproteinase (right) based on the known reversible inhibitor ilomastat (left).**

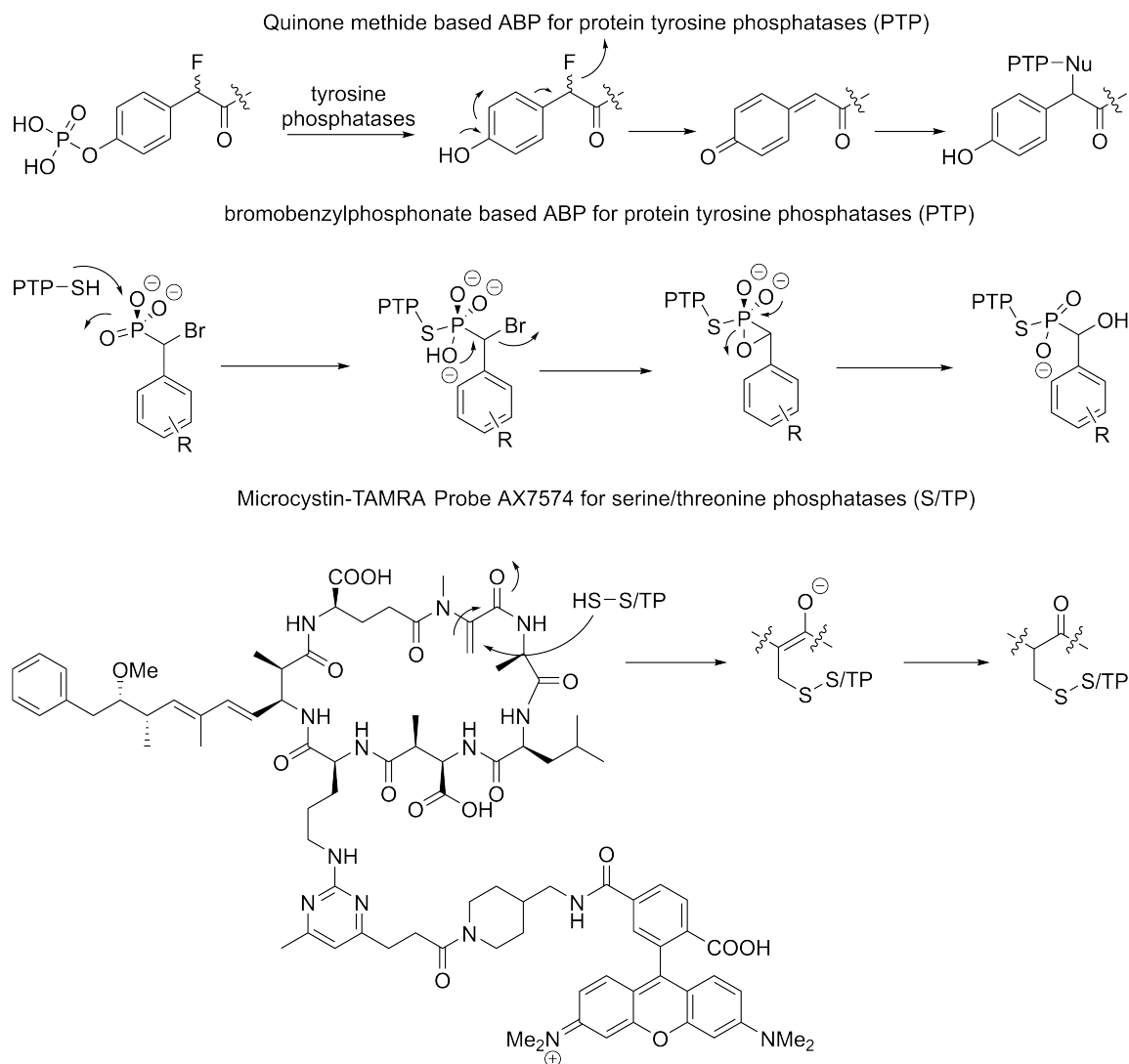
#### 6.6.4. ABPs for glycosidases

Glycosidases hydrolyze glycosidic bonds in glycoproteins and oligosaccharides. These efficient and highly specific catalysts play important roles in biological processes. Thus, a detailed knowledge of glycosidase function is invaluable for understanding and controlling diseases and for industrial applications. To profile the activity of glycosidase, several ABPs have been designed and constructed. Tsai and colleagues synthesized an ABP for glycosidases, which employed a precursor of a quinone methide electrophile to label the nucleophilic residue.<sup>47</sup> Vocadlo and Bertozzi constructed an azide-modified fluorosugar as a probe for  $\beta$ -glycosidases.<sup>48</sup> Hekmat and colleagues designed and synthesized a 1-(2,4-dinitrophenyl)-2-fluorosugar probe with a biotin tag linked via a cleavable disulfide

linker.<sup>30</sup> By reducing the disulfide bond, the proteins enriched on immobilized streptavidin can be released and analyzed.

#### 6.6.5. ABPs for protein phosphatases

Protein phosphorylation is a reversible post-translational modification.<sup>49</sup> Two major categories of phosphatases, the serine/threonine phosphatases and tyrosine phosphatases are responsible for cleaving phosphate from amino acid residues in proteins. Serine/threonine phosphatases constitute a large majority of the protein dephosphorylation activity in eukaryotic cells.<sup>50</sup> In recent years, by utilizing ABPs designed for phosphatases, researchers have identified 107 suspected tyrosine phosphatases and about 40 serine/threonine phosphatases thus far.<sup>26c, 51</sup> Zhang and colleagues have constructed a bromobenzylphosphonate probe that tagged several purified tyrosine phosphatase.<sup>51a</sup> By using a quinone methide generated *in situ* as reactive group, Lo and Zhu have developed two ABPs targeting tyrosine phosphatases, independently.<sup>51b, 52</sup> Shreder and colleagues synthesized a rhodamine-conjugated derivative of the natural inhibitor microcystin and used this agent to profile serine/threonine phosphatase activities in cellular proteomes.<sup>53</sup> Mechanisms of these ABPs are presented in Scheme 1.8 as examples.



**Scheme 1.8 Examples of ABPs for phosphatases and their mechanism of actions.**

## 7. References

1. Anand, P.; Kunnumakara, A. B.; Sundaram, C.; Harikumar, K. B.; Tharakan, S. T.; Lai, O. S.; Sung, B. Y.; Aggarwal, B. B., Cancer is a Preventable Disease that Requires Major Lifestyle Changes. *Pharm. Res-Dord* **2008**, 25 (9), 2097-2116.
2. Klein, C. A., Cancer - The metastasis cascade. *Science* **2008**, 321 (5897), 1785-1787.
3. Global Cancer Facts & Figures (2nd Edition). *American Cancer Society* **2011**.

4. (a) Ellis, R. W.; Defeo, D.; Shih, T. Y.; Gonda, M. A.; Young, H. A.; Tsuchida, N.; Lowy, D. R.; Scolnick, E. M., The P21 Src Genes of Harvey and Kirsten Sarcoma-Viruses Originate from Divergent Members of a Family of Normal Vertebrate Genes. *Nature* **1981**, 292 (5823), 506-511; (b) Chang, E. H.; Gonda, M. A.; Ellis, R. W.; Scolnick, E. M.; Lowy, D. R., Human Genome Contains 4 Genes Homologous to Transforming Genes of Harvey and Kirsten Murine Sarcoma-Viruses. *P. Natl. Acad. Sci. U.S.A.* **1982**, 79 (16), 4848-4852; (c) Parada, L. F.; Tabin, C. J.; Shih, C.; Weinberg, R. A., Human E<sub>j</sub> Bladder-Carcinoma Oncogene Is Homolog of Harvey Sarcoma-Virus Ras Gene. *Nature* **1982**, 297 (5866), 474-478.
5. (a) Scheffzek, K.; Ahmadian, M. R.; Kabsch, W.; Wiesmuller, L.; Lautwein, A.; Schmitz, F.; Wittinghofer, A., The Ras-RasGAP complex: Structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **1997**, 277 (5324), 333-338; (b) Wittinghofer, A.; Waldmann, H., Ras - A molecular switch involved in tumor formation. *Angew. Chem. Int. Ed.* **2000**, 39 (23), 4193-4214.
6. Schubbert, S.; Shannon, K.; Bollag, G., Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer.* **2007**, 7 (4), 295-308.
7. Downward, J., Targeting ras signalling pathways in cancer therapy. *Nat. Rev. Cancer.* **2003**, 3 (1), 11-22.
8. (a) Greaves, J.; Chamberlain, L. H., Palmitoylation-dependent protein sorting. *J. Cell Biol.* **2007**, 176 (3), 249-254; (b) Linder, M. E.; Deschenes, R. J., Palmitoylation: policing protein stability and traffic. *Nat. Rev. Mol. Cell Bio.* **2007**, 8 (1), 74-84.
9. (a) Zhang, F. L.; Casey, P. J., Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* **1996**, 65, 241-269; (b) Farazi, T. A.; Waksman, G.;

Gordon, J. I., The biology and enzymology of protein *N*-myristoylation. *J. Biol. Chem.* **2001**, *276* (43), 39501-39504.

10. Mumby, S. M., Reversible palmitoylation of signaling proteins. *Curr. Opin. Cell. Biol.* **1997**, *9* (2), 148-154.

11. (a) Laude, A. J.; Prior, I. A., Palmitoylation and localisation of RAS isoforms are modulated by the hypervariable linker domain. *J. Cell Sci.* **2008**, *121* (4), 421-427; (b) Conibear, E.; Davis, N. G., Palmitoylation and depalmitoylation dynamics at a glance. *J. Cell Sci.* **2010**, *123* (23), 4007-4010; (c) Vojtek, A. B.; Der, C. J., Increasing complexity of the Ras signaling pathway. *J. Biol. Chem.* **1998**, *273* (32), 19925-19928.

12. Goodwin, J. S.; Drake, K. R.; Rogers, C.; Wright, L.; Lippincott-Schwartz, J.; Philips, M. R.; Kenworthy, A. K., Depalmitoylated Ras traffics to and from the Golgi complex via a nonvesicular pathway. *J. Cell Biol.* **2005**, *170* (2), 261-272.

13. Duncan, J. A.; Gilman, A. G., A cytoplasmic acyl-protein thioesterase that removes palmitate from G protein alpha subunits and p21(RAS). *J. Biol. Chem.* **1998**, *273* (25), 15830-15837.

14. (a) Rocks, O.; Gerauer, M.; Vartak, N.; Koch, S.; Huang, Z. P.; Pechlivanis, M.; Kuhlmann, J.; Brunsveld, L.; Chandra, A.; Ellinger, B.; Waldmann, H.; Bastiaens, P. I. H., The palmitoylation machinery Is a spatially organizing system for peripheral membrane proteins. *Cell* **2010**, *141* (3), 458-471; (b) Rocks, O.; Peyker, A.; Kahms, M.; Verveer, P. J.; Koerner, C.; Lumbierres, M.; Kuhlmann, J.; Waldmann, H.; Wittinghofer, A.; Bastiaens, P. I. H., An acylation cycle regulates localization and activity of palmitoylated Ras isoforms. *Science* **2005**, *307* (5716), 1746-1752.

15. Devedjiev, Y.; Dauter, Z.; Kuznetsov, S. R.; Jones, T. L. Z.; Derewenda, Z. S., Crystal structure of the human acyl protein thioesterase I from a single X-ray data set to 1.5 angstrom. *Structure* **2000**, 8 (11), 1137-1146.
16. Yeh, D. C.; Duncan, J. A.; Yamashita, S.; Michel, T., Depalmitoylation of endothelial nitric-oxide synthase by acyl-protein thioesterase 1 is potentiated by  $\text{Ca}^{2+}$ -calmodulin. *J. Biol. Chem.* **1999**, 274 (46), 33148-33154.
17. Sim, D. S.; Dilks, J. R.; Flaumenhaft, R., Platelets possess and require an active protein palmitoylation pathway for agonist-mediated activation and in vivo thrombus formation. *Arterioscl. Throm. Vas.* **2007**, 27 (6), 1478-1485.
18. Satou, M.; Nishi, Y.; Yoh, J.; Hattori, Y.; Sugimoto, H., Identification and characterization of acyl-protein thioesterase 1/lysophospholipase I as a ghrelin deacylation/lysophospholipid hydrolyzing enzyme in fetal bovine serum and conditioned medium. *Endocrinology* **2010**, 151 (10), 4765-4775.
19. (a) Sunaga, H.; Sugimoto, H.; Nagamachi, Y.; Yamashita, S., Purification and Properties of Lysophospholipase Isoenzymes from Pig Gastric-Mucosa. *Biochem. J.* **1995**, 308, 551-557; (b) Toyoda, T.; Sugimoto, H.; Yamashita, S., Sequence, expression in *Escherichia coli*, and characterization of lysophospholipase II. *Biochim. Biophys. Acta* **1999**, 1437 (2), 182-193.
20. Tomatis, V. M.; Trenchi, A.; Gomez, G. A.; Daniotti, J. L., Acyl-Protein Thioesterase 2 Catalyzes the Deacylation of Peripheral Membrane-Associated GAP-43. *Plos One* **2010**, 5 (11).
21. Deck, P.; Pendzialek, D.; Biel, M.; Wagner, M.; Popkirova, B.; Ludolph, B.; Kragol, G.; Kuhlmann, J.; Giannis, A.; Waldmann, H., Development and biological evaluation of



acyl protein thioesterase 1 (APT1) inhibitors. *Angew. Chem. Int. Ed.* **2005**, *44* (31), 4975-4980.

22. Dekker, F. J.; Rocks, O.; Vartak, N.; Menninger, S.; Hedberg, C.; Balamurugan, R.; Wetzel, S.; Renner, S.; Gerauer, M.; Scholermann, B.; Rusch, M.; Kramer, J. W.; Rauh, D.; Coates, G. W.; Brunsveld, L.; Bastiaens, P. I. H.; Waldmann, H., Small-molecule inhibition of APT1 affects Ras localization and signaling. *Nat. Chem. Biol.* **2010**, *6* (6), 449-456.

23. Adibekian, A.; Martin, B. R.; Speers, A. E.; Brown, S. J.; Spicer, T.; Fernandez-Vega, V.; Ferguson, J.; Cravatt, B. F.; Hodder, P.; Rosen, H., Optimization and characterization of a triazole urea dual inhibitor for lysophospholipase 1 (LYPLA1) and lysophospholipase 2 (LYPLA2). In *Probe Reports from the NIH Molecular Libraries Program*, Bethesda (MD), 2010.

24. (a) Adibekian, A.; Martin, B. R.; Chang, J. W.; Hsu, K. L.; Tsuboi, K.; Bachovchin, D. A.; Speers, A. E.; Brown, S. J.; Spicer, T.; Fernandez-Vega, V.; Ferguson, J.; Cravatt, B. F.; Hodder, P.; Rosen, H., Characterization of a Selective, Reversible Inhibitor of Lysophospholipase 2 (LYPLA2). In *Probe Reports from the NIH Molecular Libraries Program*, Bethesda (MD), 2010; (b) Adibekian, A.; Martin, B. R.; Chang, J. W.; Hsu, K. L.; Tsuboi, K.; Bachovchin, D. A.; Speers, A. E.; Brown, S. J.; Spicer, T.; Fernandez-Vega, V.; Ferguson, J.; Cravatt, B. F.; Hodder, P.; Rosen, H., Characterization of a Selective, Reversible Inhibitor of Lysophospholipase 1 (LYPLA1). In *Probe Reports from the NIH Molecular Libraries Program*, Bethesda (MD), 2010.

25. (a) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W., Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **2008**, *77*, 383-414; (b)

Liu, Y.; Patricelli, M. P.; Cravatt, B. F., Activity-based protein profiling: the serine hydrolases. *P. Natl. Acad. Sci. U.S.A.* **1999**, *96* (26), 14694-14699.

26. (a) Evans, M. J.; Cravatt, B. F., Mechanism-based profiling of enzyme families. *Chem. Rev.* **2006**, *106* (8), 3279-3301; (b) Uttamchandani, M.; Li, J.; Sun, H.; Yao, S. Q., Activity-based protein profiling: new developments and directions in functional proteomics. *Chembiochem* **2008**, *9* (5), 667-675; (c) Krishnamurthy, D.; Barrios, A. M., Profiling protein tyrosine phosphatase activity with mechanistic probes. *Curr. Opin. Chem. Biol.* **2009**, *13* (4), 375-381.

27. (a) Speers, A. E.; Adam, G. C.; Cravatt, B. F., Activity-based protein profiling in vivo using a copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. *J. Am. Chem. Soc.* **2003**, *125* (16), 4686-4687; (b) Meldal, M.; Tornøe, C. W., Cu-Catalyzed Azide-Alkyne Cycloaddition. *Chem. Rev.* **2008**, *108* (8), 2952-3015; (c) Kolb, H. C.; Sharpless, K. B., The growing impact of click chemistry on drug discovery. *Drug Discov. Today* **2003**, *8* (24), 1128-1137; (d) Lutz, J. F., 1,3-dipolar cycloadditions of azides and alkynes: A universal ligation tool in polymer and materials science. *Angew. Chem. Int. Ed.* **2007**, *46* (7), 1018-1025.

28. (a) Kidd, D.; Liu, Y. S.; Cravatt, B. F., Profiling serine hydrolase activities in complex proteomes. *Biochemistry-Us* **2001**, *40* (13), 4005-4015; (b) Patricelli, M. P.; Giang, D. K.; Stamp, L. M.; Burbaum, J. J., Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* **2001**, *1* (9), 1067-1071.

29. Okerberg, E. S.; Wu, J. Y.; Zhang, B. H.; Samii, B.; Blackford, K.; Winn, D. T.; Shreder, K. R.; Burbaum, J. J.; Patricelli, M. P., High-resolution functional proteomics by active-site peptide profiling. *P. Natl. Acad. Sci. U.S.A.* **2005**, *102* (14), 4996-5001.

30. Hekmat, O.; Kim, Y. W.; Williams, S. J.; He, S. M.; Withers, S. G., Active-site peptide "fingerprinting" of glycosidases in complex mixtures by mass spectrometry - Discovery of a novel retaining beta-1,4-glycanase in *Cellulomonas fimi*. *J. Biol. Chem.* **2005**, *280* (42), 35126-35135.
31. Rusch, M.; Zimmermann, T. J.; Burger, M.; Dekker, F. J.; Gormer, K.; Triola, G.; Brockmeyer, A.; Janning, P.; Bottcher, T.; Sieber, S. A.; Vetter, I. R.; Hedberg, C.; Waldmann, H., Identification of Acyl Protein Thioesterases 1 and 2 as the Cellular Targets of the Ras-Signaling Modulators Palmostatin B and M. *Angew. Chem. Int. Ed.* **2011**, *50* (42), 9838-9842.
32. (a) Rusch, M. Design, synthesis and application of small molecule acyl protein thioesterase inhibitors. Ph.D. Dissertaion, Technische Universität Dortmund, Dortmund, 2011; (b) Davda, D.; Martin, B. R., Acyl protein thioesterase inhibitors as probes of dynamic *S*-palmitoylation. *Medchemcomm* **2014**, *5* (3), 268-276.
33. Simon, G. M.; Cravatt, B. F., Activity-based Proteomics of Enzyme Superfamilies: Serine Hydrolases as a Case Study. *J. Biol. Chem.* **2010**, *285* (15), 11051-11055.
34. (a) Pan, Z. Y.; Jeffery, D. A.; Chehade, K.; Beltman, J.; Clark, J. M.; Grothaus, P.; Bogyo, M.; Baruch, A., Development of activity-based probes for trypsin-family serine proteases. *Bioorg. Med. Chem. Lett.* **2006**, *16* (11), 2882-2885; (b) Mahrus, S.; Craik, C. S., Selective chemical functional probes of granzymes A and B reveal granzyme B is a major effector of natural killer cell-mediated lysis of target cells. *Chem. Biol.* **2005**, *12* (5), 567-577; (c) Shannon, D. A.; Gu, C.; McLaughlin, C. J.; Kaiser, M.; van der Hoorn, R. A. L.; Weerapana, E., Sulfonyl Fluoride Analogues as Activity-Based Probes for Serine Proteases. *Chembiochem* **2012**, *13* (16), 2327-2330.

35. (a) Jessani, N.; Niessen, S.; Wei, B. Q. Q.; Nicolau, M.; Humphrey, M.; Ji, Y. R.; Han, W. S.; Noh, D. Y.; Yates, J. R.; Jeffrey, S. S.; Cravatt, B. F., A streamlined platform for high-content functional proteomics of primary human specimens. *Nat. Methods* **2005**, *2* (9), 691-697; (b) Jessani, N.; Liu, Y. S.; Humphrey, M.; Cravatt, B. F., Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. *P. Natl. Acad. Sci. U.S.A.* **2002**, *99* (16), 10335-10340; (c) Jessani, N.; Humphrey, M.; McDonald, W. H.; Niessen, S.; Masuda, K.; Gangadharan, B.; Yates, J. R.; Mueller, B. M.; Cravatt, B. F., Carcinoma and stromal enzyme activity profiles associated with breast tumor growth in vivo. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101* (38), 13756-13761.
36. Milner, J. M.; Kevorkian, L.; Young, D. A.; Jones, D.; Wait, R.; Donell, S. T.; Barksby, E.; Patterson, A. M.; Middleton, J.; Cravatt, B. F.; Clark, I. M.; Rowan, A. D.; Cawston, T. E., Fibroblast activation protein alpha is expressed by chondrocytes following a pro-inflammatory stimulus and is elevated in osteoarthritis. *Arthrit. Res. Ther.* **2006**, *8* (1).
37. Alexander, J. P.; Cravatt, B. F., Mechanism of carbamate inactivation of FAAH: Implications for the design of covalent inhibitors and in vivo functional probes for enzymes. *Chem. Biol.* **2005**, *12* (11), 1179-1187.
38. (a) Sanman, L. E.; Bogoy, M., Activity-Based Profiling of Proteases. *Annu. Rev. Biochem.* **2014**, *83*, 249-273; (b) Fonovic, M.; Bogoy, M., Activity based probes for proteases: Applications to biomarker discovery, molecular imaging and drug screening. *Curr. Pharm. Des.* **2007**, *13* (3), 253-261.
39. Greenbaum, D.; Medzihradszky, K. F.; Burlingame, A.; Bogoy, M., Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem. Biol.* **2000**, *7* (8), 569-581.

40. Mason, R. W.; Wilcox, D.; Wikstrom, P.; Shaw, E. N., The Identification of Active Forms of Cysteine Proteinases in Kirsten-Virus-Transformed Mouse Fibroblasts by Use of a Specific Radiolabeled Inhibitor. *Biochem. J.* **1989**, *257* (1), 125-129.
41. (a) Rhoads, M. L.; Fetterer, R. H., Developmentally-Regulated Secretion of Cathepsin L-Like Cysteine Proteases by *Haemonchus-Contortus*. *J. Parasitol.* **1995**, *81* (4), 505-512; (b) Bedner, E.; Smolewski, P.; Amstad, P.; Darzynkiewicz, Z., Activation of caspases measured in situ by binding of fluorochrome-labeled inhibitors of caspases (FLICA): Correlation with DNA fragmentation. *Exp. Cell Res.* **2000**, *259* (1), 308-313; (c) Dobrota, C.; Fasci, D.; Hadade, N. D.; Roiban, G. D.; Pop, C.; Meier, V. M.; Dumitru, I.; Matache, M.; Salvesen, G. S.; Funeriu, D. P., Glycine Fluoromethylketones as SENP-Specific Activity Based Probes. *Chembiochem* **2012**, *13* (1), 80-84.
42. (a) Bromme, D.; Smith, R. A.; Coles, P. J.; Kirschke, H.; Storer, A. C.; Krantz, A., Potent Inactivation of Cathepsin-S and Cathepsin-L by Peptidyl (Acyloxy)Methyl Ketones. *Biol. Chem. H-S* **1994**, *375* (5), 343-347; (b) Kato, D.; Boatright, K. M.; Berger, A. B.; Nazif, T.; Blum, G.; Ryan, C.; Chehade, K. A. H.; Salvesen, G. S.; Bogoy, M., Activity-based probes that target diverse cysteine protease families. *Nat. Chem. Biol.* **2005**, *1* (1), 33-38; (c) Blum, G.; Mullins, S. R.; Keren, K.; Fonovic, M.; Jedeszko, C.; Rice, M. J.; Sloane, B. F.; Bogoy, M., Dynamic imaging of protease activity with fluorescently quenched activity-based probes. *Nat. Chem. Biol.* **2005**, *1* (4), 203-209.
43. (a) Wang, G.; Mahesh, U.; Chen, G. Y. J.; Yao, S. Q., Solid-phase synthesis of peptide vinyl sulfones as potential inhibitors and activity-based probes of cysteine proteases. *Org. Lett.* **2003**, *5* (5), 737-740; (b) Borodovsky, A.; Ovaa, H.; Meester, W. J. N.; Venanzi, E. S.; Bogoy, M. S.; Hekking, B. G.; Ploegh, H. L.; Kessler, B. M.; Overkleeft, H. S., Small-

molecule inhibitors and probes for ubiquitin- and ubiquitin-like-specific proteases. *Chembiochem* **2005**, *6* (2), 287-291; (c) Bogyo, M.; McMaster, J. S.; Gaczynska, M.; Tortorella, D.; Goldberg, A. L.; Ploegh, H., Covalent modification of the active site threonine of proteasomal beta subunits and the Escherichia coli homolog HslV by a new class of inhibitors. *P. Natl. Acad. Sci. U.S.A.* **1997**, *94* (13), 6629-6634.

44. (a) Mitić, N.; Miraula, M.; Selleck, C.; Hadler, K. S.; Uribe, E.; Pedroso, M. M.; Schenk, G., Chapter Three - Catalytic Mechanisms of Metallohydrolases Containing Two Metal Ions. In *Advances in Protein Chemistry and Structural Biology*, Christo, Z. C., Ed. Academic Press: 2014; Vol. Volume 97, pp 49-81; (b) Wilcox, D. E., Binuclear metallohydrolases. *Chem. Rev.* **1996**, *96* (7), 2435-2458; (c) Mitic, N.; Smith, S. J.; Neves, A.; Guddat, L. W.; Gahan, L. R.; Schenk, G., The catalytic mechanisms of binuclear metallohydrolases. *Chem. Rev.* **2006**, *106* (8), 3338-3363; (d) Schenk, G.; Mitic, N.; Gahan, L. R.; Ollis, D. L.; McGeary, R. P.; Guddat, L. W., Binuclear Metallohydrolases: Complex Mechanistic Strategies for a Simple Chemical Reaction. *Acc. Chem. Res.* **2012**, *45* (9), 1593-1603; (e) McGeary, R. P.; Schenk, G.; Guddat, L. W., The applications of binuclear metallohydrolases in medicine: Recent advances in the design and development of novel drug leads for purple acid phosphatases, metallo-beta-lactamases and arginases. *Eur. J. Med. Chem.* **2014**, *76*, 132-144.

45. (a) Salisbury, C. M.; Cravatt, B. F., Activity-based probes for proteomic profiling of histone deacetylase complexes. *P. Natl. Acad. Sci. U.S.A.* **2007**, *104* (4), 1171-1176; (b) Salisbury, C. M.; Cravatt, B. F., Optimization of activity-based probes for proteomic profiling of histone deacetylase complexes. *J. Am. Chem. Soc.* **2008**, *130* (7), 2184-2194.

46. Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, B. F., Activity-based probes for the proteomic profiling of metalloproteases. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101* (27), 10000-10005.
47. Tsai, C. S.; Li, Y. K.; Lo, L. C., Design and synthesis of activity probes for glycosidases. *Org. Lett.* **2002**, *4* (21), 3607-3610.
48. Vocadlo, D. J.; Bertozzi, C. R., A strategy for functional proteomic analysis of glycosidase activity from cell lysates. *Angew. Chem. Int. Ed.* **2004**, *43* (40), 5338-5342.
49. (a) Johnson, L. N., The regulation of protein phosphorylation. *Biochem. Soc. Trans.* **2009**, *37*, 627-641; (b) Ciesla, J.; Fraczyk, T.; Rode, W., Phosphorylation of basic amino acid residues in proteins: important but easily missed. *Acta Biochim. Pol.* **2011**, *58* (2), 137-147; (c) Burnett, G.; Kennedy, E. P., The Enzymatic Phosphorylation of Proteins. *J. Biol. Chem.* **1954**, *211* (2), 969-980.
50. (a) Mansuy, I. M.; Shenolikar, S., Protein serine/threonine phosphatases in neuronal plasticity and disorders of learning and memory. *Trends Neurosci.* **2006**, *29* (12), 679-686; (b) Hunter, T.; Sefton, B. M., Transforming Gene-Product of Rous-Sarcoma Virus Phosphorylates Tyrosine. *P. Natl. Acad. Sci. U.S.A.* **1980**, *77* (3), 1311-1315; (c) Sim, A. T. R.; Baldwin, M. L.; Rostas, J. A. P.; Holst, J.; Ludowyke, R. I., The role of serine/threonine protein phosphatases in exocytosis. *Biochem. J.* **2003**, *373*, 641-659.
51. (a) Kumar, S.; Zhou, B.; Liang, F. B.; Wang, W. Q.; Huang, Z. H.; Zhang, Z. Y., Activity-based probes for protein tyrosine phosphatases. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101* (21), 7943-7948; (b) Zhu, Q.; Huang, X.; Chen, G. Y. J.; Yao, S. Q., Activity-based fluorescent probes that target phosphatases. *Tetra. Lett.* **2003**, *44* (13), 2669-2672.

52. Lo, L. C.; Chiang, Y. L.; Kuo, C. H.; Liao, H. K.; Chen, Y. J.; Lin, J. J., Study of the preferred modification sites of the quinone methide intermediate resulting from the latent trapping device of the activity probes for hydrolases. *Biochem. Bioph. Res. Co.* **2005**, 326 (1), 30-35.
53. Shreder, K. R.; Liu, Y. S.; Nomanhboy, T.; Fuller, S. R.; Wong, M. S.; Gai, W. Z.; Wu, J. Y.; Leventhal, P. S.; Lill, J. R.; Corral, S., Design and synthesis of AX7574: A microcystin-derived, fluorescent probe for serine/threonine phosphatases. *Bioconjugate Chem.* **2004**, 15 (4), 790-798.



# **CHAPTER TWO**

## **Design and Synthesis of the ABP for APTs**



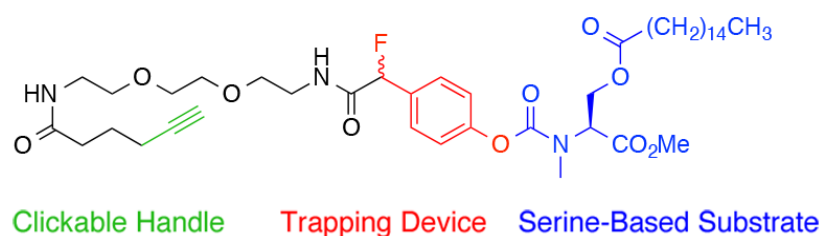
## 1. Abstract

In order to maximize the specificity of the ABP for activity-based protein profiling of acyl protein thioesterases (APTs), we propose a probe that mimics the structure of a palmitoylated cysteine residue. In contrast to other ABPs targeting APTs, the reactive center of our ABP is not directly bound to its recognition unit. Rather, the two are coupled at opposite ends of a central carbamate linkage. Synthesis of the ABP can be accomplished through a convergent route. While multiple synthetic routes were attempted in pursuit of one that conferred formation of a key phenyl carbamate, the final iteration of the ABP's synthesis was a 15-step convergent route that called for use of *N*-methyl serine methyl ester in the carbamate-forming reaction step. All important synthetic products and intermediates are well characterized by NMR spectroscopy and high-resolution mass spectrometry.

## 2. Design of the ABP

Based on the mechanism and substrate specificity of the APTs, an activity-based probe (ABP) was designed to react with only the active forms of APTs. Our ABP for APTs consists of three major components (Scheme 2.1): a recognition substrate, a trapping device, and a clickable reporter handle. In our proposed mechanism of action (Scheme 2.2), the *O*-palmitoylated serine, which mimics the natural substrate of APTs, is hydrolyzed by an APT. Upon enzymatic cleavages of the palmitoyl group on serine, the released serine carbamate intermediate undergoes rapid cyclization, resulting in the formation of *p*-fluoromethylphenol.<sup>1</sup> The *p*-fluoromethylphenol undergoes 1,6-elimination to produce an electrophilic quinone methide (QM). This quinone methide generated *in situ* is able to trap

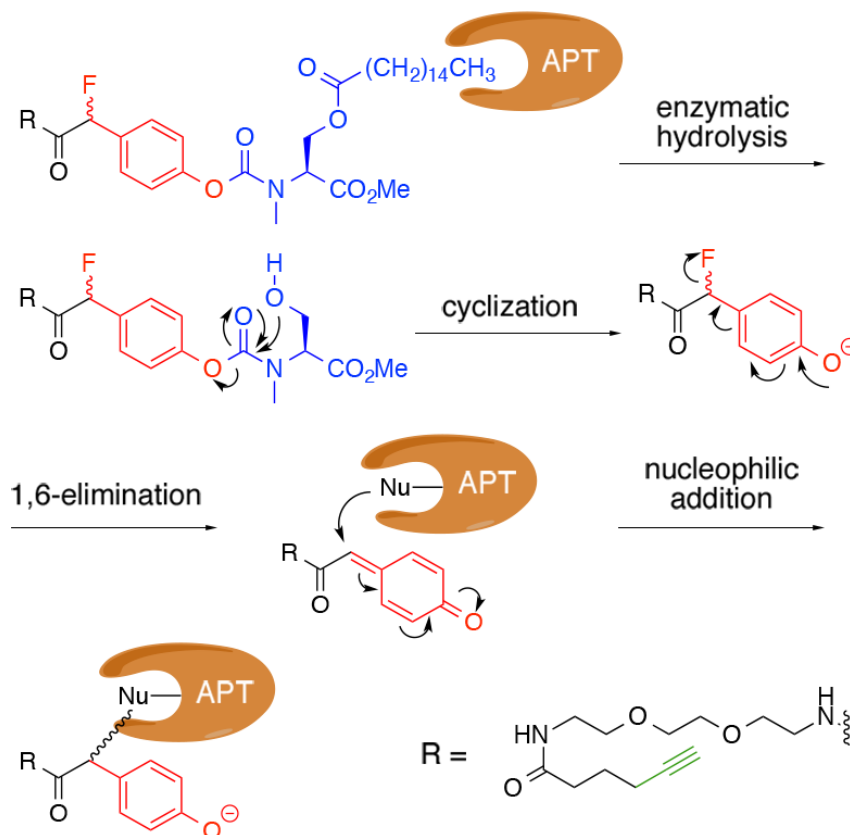
a nearby nucleophilic amino acid residue in the APT's active site, thereby covalently labeling the APT in an activity-dependent manner.<sup>2</sup> Subsequently, a fluorescent- or affinity-conjugated azide tag can be attached to the clickable alkyne handle via click chemistry. It should be noted that the amino group of the serine component is methylated due to the enhanced stability of *N*-methylated carbamate in basic aqueous solution.



**Scheme 2.1 Design of the clickable ABP for APTs.** The ABP has three major components: an *O*-palmitoylated serine-based substrate, a trapping device that functions as a precursor to a strong electrophile, and a versatile clickable tag.

*p*-Fluoromethylphenol derivatives have been successfully applied in ABPP targeting tyrosine phosphatases due to their structural similarity with tyrosine.<sup>2</sup> However, a fluoromethylphenol-based ABP has not yet been applied to studies targeting post-translational modifications at serine, threonine, or cysteine residues. Compared to the traditional use of quinone methides, the introduction of a cyclizable serine carbamate moiety not only enhances the substrate similarity between the probe and its natural substrate, but also expands the opportunity to target other enzyme families that hydrolyze post-translational modifications on serine or cysteine residues. In addition, compared to the  $\beta$ -lactone inhibitors developed for APT1/APT2, the inherent hydrolysis of our APT-probe

complex should not be a problem that interferes with the labeling of the enzyme.<sup>3</sup>



**Scheme 2.2 Proposed reaction mechanism for the ABP designed for APTs.** Enzymatic depalmitoylation initiates the spontaneous cyclization of serine, which activates the trapping device through 1,6-elimination of the liberated fluoromethyl phenoxide. The quinone methide generated *in situ* traps a nearby nucleophilic amino acid residue to covalently label the APT target.

### 3. Validation of ABP's design by Molecular Dynamics Simulation

In an effort to validate the design of the ABP, a molecular dynamics (MD) simulation of the ABP-APT1 reaction was performed *in silico* using the known crystal structure of APT1 (PDB: 1fj2).<sup>4</sup> The quinone methide obtained from ABP hydrolysis/elimination was docked

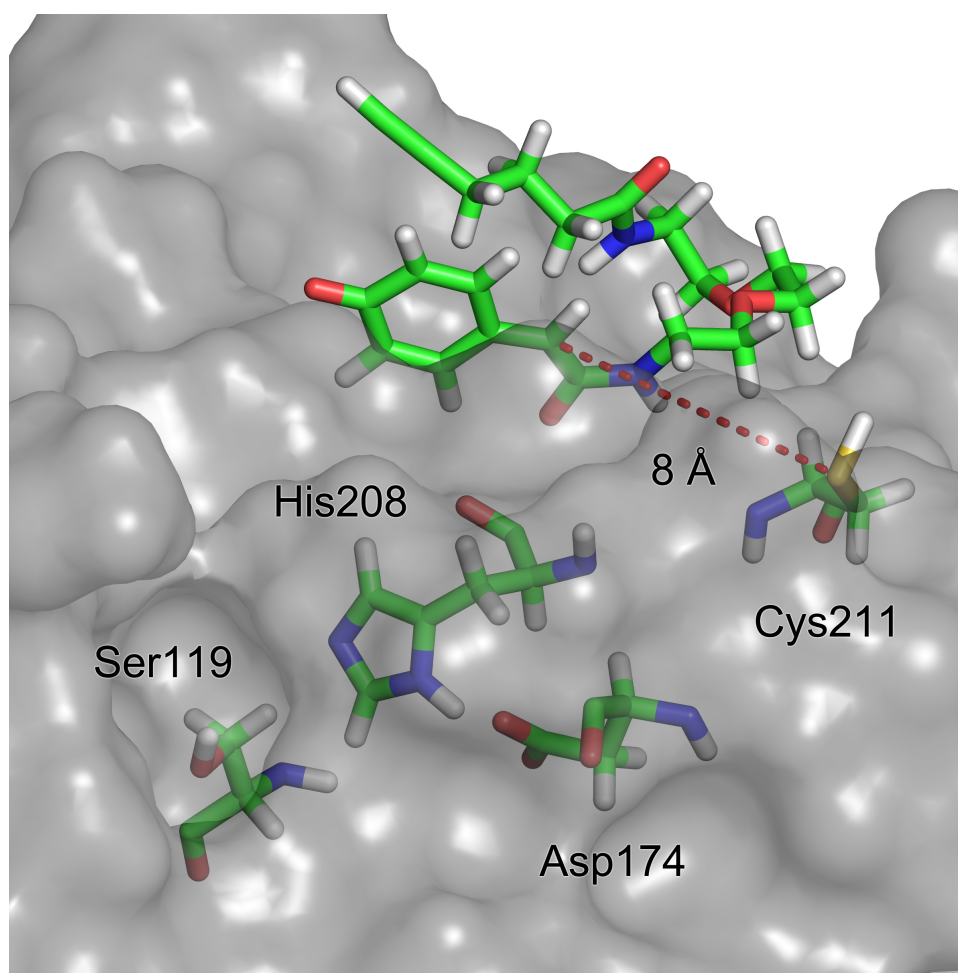
into the active site with the AutoDock Vina software package.<sup>5</sup> The most stable complex structure was utilized as the initial structure for the molecular dynamics simulation. The Gromacs (Groningen Machine for Chemical Simulations) software package was employed to carry out the MD simulation.<sup>6</sup> The simulation result was analyzed and visualized by PyMOL.<sup>7</sup>

To achieve the simulation, the crystal structure PDB file of human APT1 protein was downloaded from the RCSB Protein Data Bank.<sup>4</sup> The water molecules and bromide ions were removed and only the monomer of APT1 (25 kDa) was used. In addition, all the non-polar hydrogen atoms on protein residues were added using the UCSF Chimera. The topologies and parameters of the quinone methide ligand were generated by using ACPYPE and Amber 9.0.<sup>8</sup>

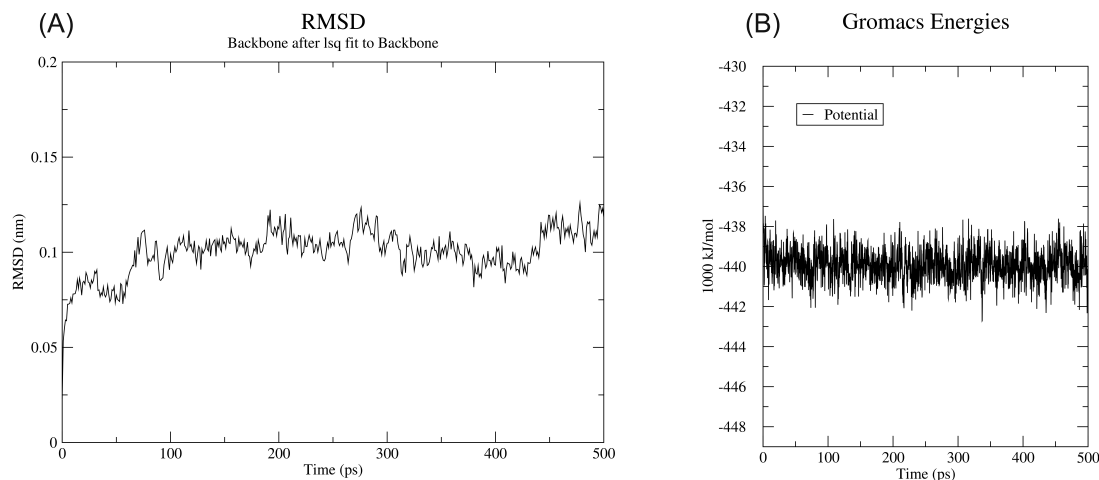
In the simulation, the Amber99SB force field and TIP3P water model were used.<sup>5a</sup> To make the whole system neutral, 0.15 M NaCl was added into the octahedron box ( $d = 1.0$ ). First, the energy of the complex was minimized using a conjugate gradient algorithm in 200 steps. Second, a 100 ps position-restrained dynamics run was performed under NVT ensemble to relax the water molecules. Finally, a full 500 ps dynamics run was performed under NPT ensemble. All the calculations were carried out in the Center for Computation and Visualization of Brown University.

The structure of the complex with minimized energies is presented in Figure 2.1. The quinone methide, the potential nucleophile, and the catalytic residues are labeled and presented as ball-and-stick models. The distance between the benzylic carbon atom and the Cys211 thiol was measured. Results from this simulation suggest that the Cys211 residue is the most proximal nucleophile to the benzylic carbon atom of the ABP, which is

distanced 8 Å from the thiol of the Cys211 (Figure 2.1). The RMSD (root-mean-square deviation of atomic positions) and potential energies are also plotted versus simulation time (Figure 2.2). This molecular dynamics simulation result lends strong support to the design concept that the ABP covalently labels in proximity to the APT1 active site in accordance with its proposed mechanism of action.



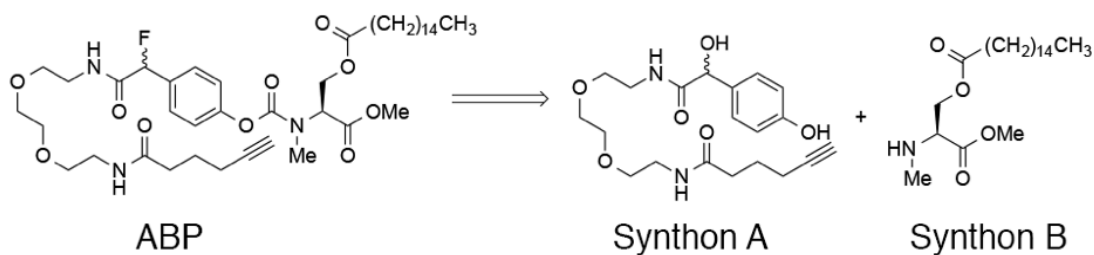
**Figure 2.1 The structure of the energy-minimized complex in the molecular dynamics simulation.** Part of the protein surface is shown. The quinone methide, the potential nucleophile, and the catalytic residues are labeled and presented as ball-and-stick models. The distance between the benzylic carbon atom and the Cys211 thiol was measured.



**Figure 2.2** The RMSD and potential energies versus simulation time.

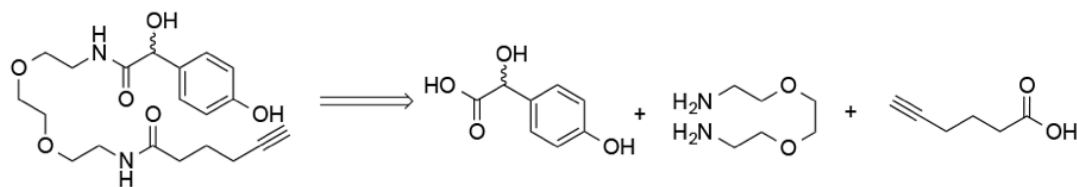
#### 4. Synthesis of the ABP

The retrosynthesis of the ABP is presented in Scheme 2.3. The synthesis of the ABP can be accomplished by reacting the Synthon A and Synthon B to give a carbamate. The retrosyntheses of the both synthons are illustrated in the Scheme 2.4 and Scheme 2.5.

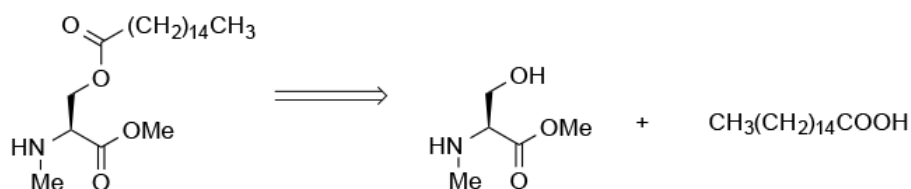


**Scheme 2.3 Retrosynthetic analysis of the ABP.** The synthesis of the ABP can be accomplished by reacting the Synthon A and Synthon B to give a carbamate.





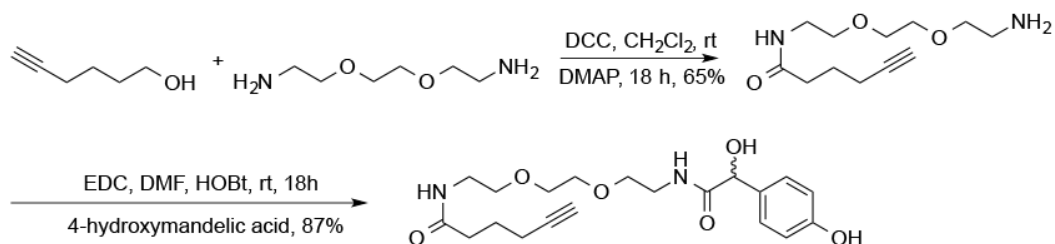
**Scheme 2.4 Retrosynthesis of Synthon A.** The synthesis of the Synthon A can be accomplished by coupling 6-hexynoic acid and 4-hydroxymandelic acid to each amino group of the diamine linker.



**Scheme 2.5 Retrosynthesis of Synthon B.** Synthon B can be synthesized by the *O*-acylation of the *N*-methyl serine methyl ester.

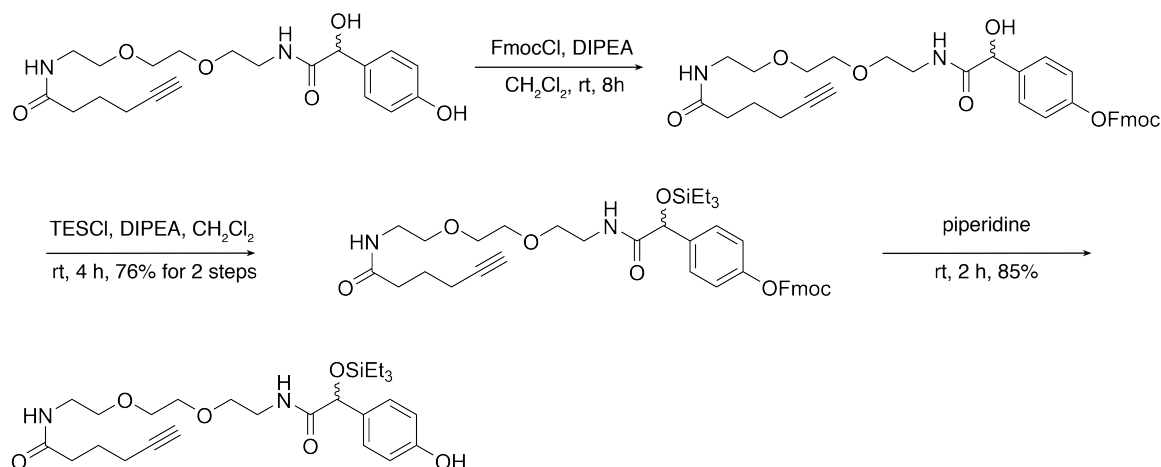
#### 4.1. Synthesis of the Synthon A.

Based on the above retrosynthesis, the synthetic pathway toward Synthon A was designed and carried out (Scheme 2.6). First, 6-hexynoic acid was coupled to one of the two amino groups of the linker 2,2'-(ethane-1,2-diylbis(oxy))bis(ethan-1-amine). To optimize the yield of the targeted product, several different peptide coupling reagents and different number of equivalents of starting materials were tested. After flash chromatography column purification, the product *N*-(2-(2-(2-aminoethoxy)ethoxy)ethyl)hex-5-ynamide was coupled with 4-hydroxymandelic acid to give Synthon A.



**Scheme 2.6 Synthesis of Synthon A.** 5-Hexynoic acid and 4-hydroxymandelic acid were coupled to the diamine linker with DCC and EDC as amide coupling reagents, respectively.

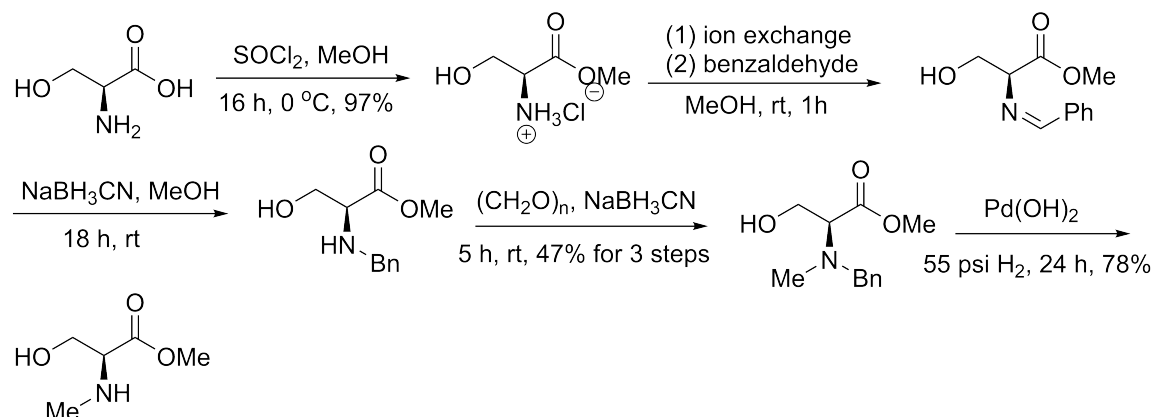
To join Synthons A and B, the phenolic hydroxyl group was reacted with phosgene to give the corresponding chloroformate. This chloroformate was coupled with the amino group of Synthon B to form a carbamate linkage. However, the benzylic hydroxyl group was found to be reactive with the intermediate chloroformate. Thus, the reaction of phosgene with Synthon A produced a carbonate formed between two molecules of Synthon A. To avoid this undesirable side-reaction, the benzylic hydroxyl group was protected prior to reaction with phosgene. This strategy is outlined in Scheme 2.7, and was adopted in order to achieve the desired carbamate. Once the Synthon A was obtained, the phenolic hydroxyl group was protected by fluorenylmethyloxycarbonyl chloride (Fmoc-Cl) selectively due to the acidity difference between the two hydroxyl groups. Subsequently, the benzylic hydroxyl group was orthogonally protected as the triethylsilyl ether (Et<sub>3</sub>SiO-), followed by cleavage of the fluorenylmethyloxycarbonyl group under basic condition.



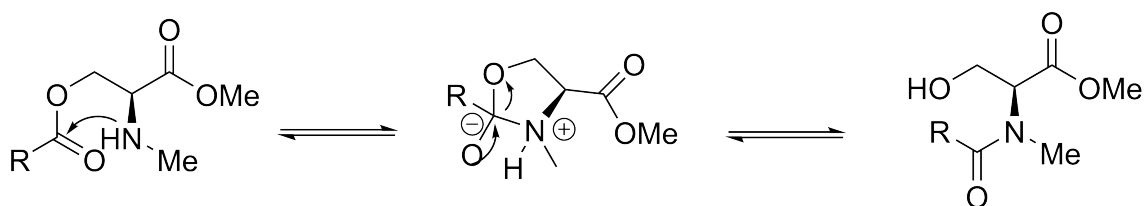
**Scheme 2.7 Selective protection of the benzylic hydroxyl group.** The phenolic hydroxyl group was selectively protected using the fluorenylmethyloxycarbonyl (Fmoc) group because of the difference in acidity between the two hydroxyl groups, followed by protection of the benzylic  $\text{-OH}$  group with  $\text{Et}_3\text{SiCl}$ . Finally, the Fmoc- group was deprotected using an organic base.

#### 4.2. Synthesis of the Synthon B

Based on the retrosynthetic analysis (Scheme 2.5), the synthesis of the Synthon B was accomplished as described in Scheme 2.8.<sup>9</sup> First, the esterification of the L-serine was achieved by reaction with thionyl chloride ( $\text{SOCl}_2$ ) in methanol. The amino group of the L-serine methyl ester was subsequently protected with a benzyl group through a reductive amination reaction, and then methylated. The benzylation of the  $\text{-NH}_2$  group is essential since *N,N*-dimethyl serine is the major product of the methylation with unprotected serine. The cleavage of the benzyl protecting group was accomplished by Parr hydrogenation in Parr hydrogenation apparatus. In addition, to prevent the potential intramolecular acyl-transfer reaction of the palmitoyl group from oxygen to nitrogen (Scheme 2.9), the coupling reaction was done before the *O*-acylation of the serine component.



**Scheme 2.8 Synthesis of Synthon B.** L-serine was esterified using  $\text{SOCl}_2$  in methyl alcohol, and *N*-protected with a benzyl group. The benzylated amine was subsequently methylated by reductive amination. The final cleavage of the benzyl protecting group was accomplished by high pressure catalytic hydrogenation.

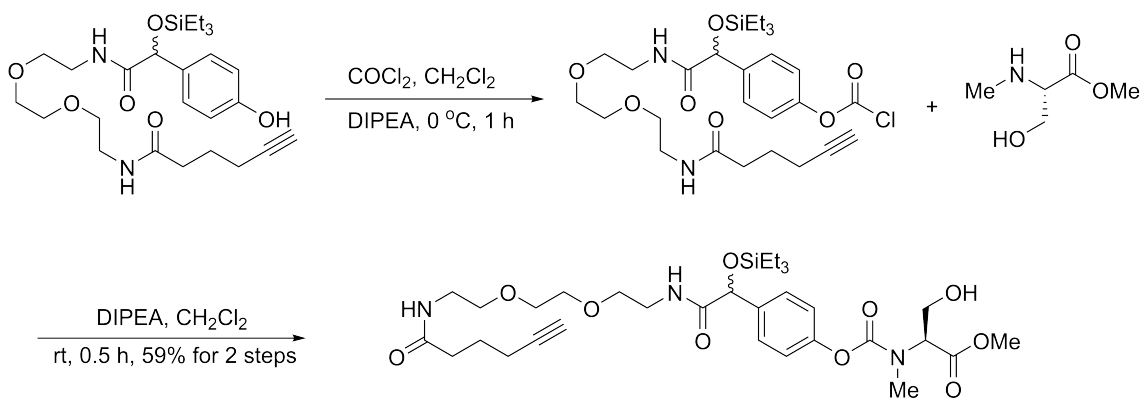


**Scheme 2.9 The intramolecular acyl-transfer reaction of the *O*-acyl *L*-serine.** The secondary amino group can attack the carbonyl of the ester through a five-member ring transition state.

#### 4.3. Linkage of the Synthon A and B

Synthon A successfully underwent reaction with phosgene to give the desired chloroformate (Scheme 2.10). It should be noted that this reaction failed when triphosgene was used instead of phosgene. This finding was surprising, as recent literature reports suggest that triphosgene is equivalent to phosgene in most synthetic reactions. To the solution of the chloroformate was added the Synthon B. These synthons reacted to yield the desired carbamate intermediate. This carbamate intermediate is labile to both strong acid (due to the liability of the triethylsilyl ether), and strong base (because the serine

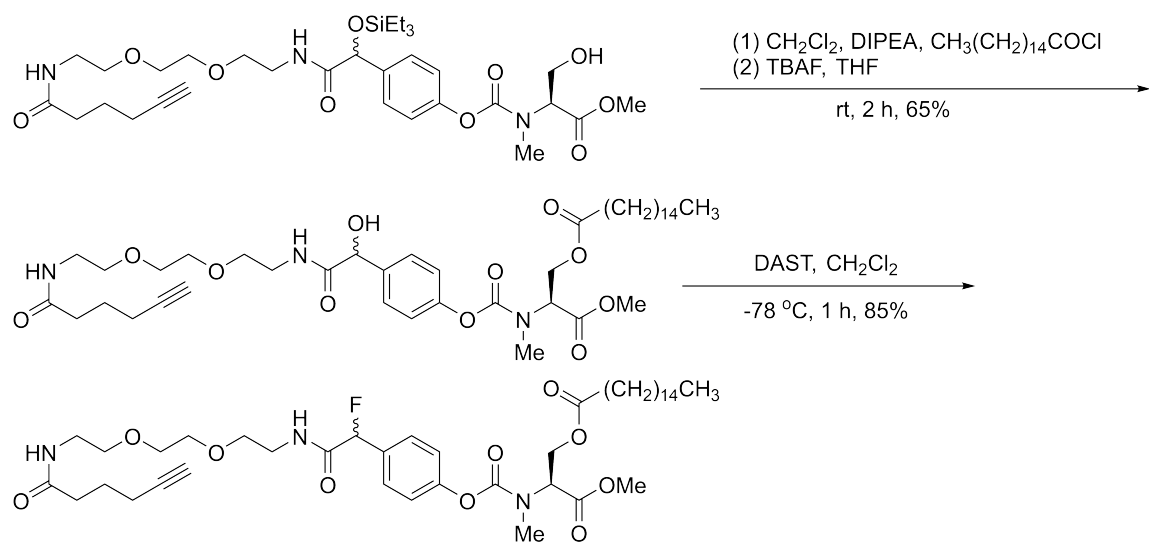
component can rapidly cyclize at pH > 8). To stabilize this intermediate for purification, acetic acid (pKa = 4.76) was added to the reaction before the flash column.



**Scheme 2.10 Linkage of Synthon A with Synthon B.** Synthon A was reacted with phosgene, and the resulting chloroformate was subsequently coupled with Synthon B to afford the desired carbamate intermediate.

#### 4.4. Acylation, deprotection and deoxyfluorination of the carbamate intermediate

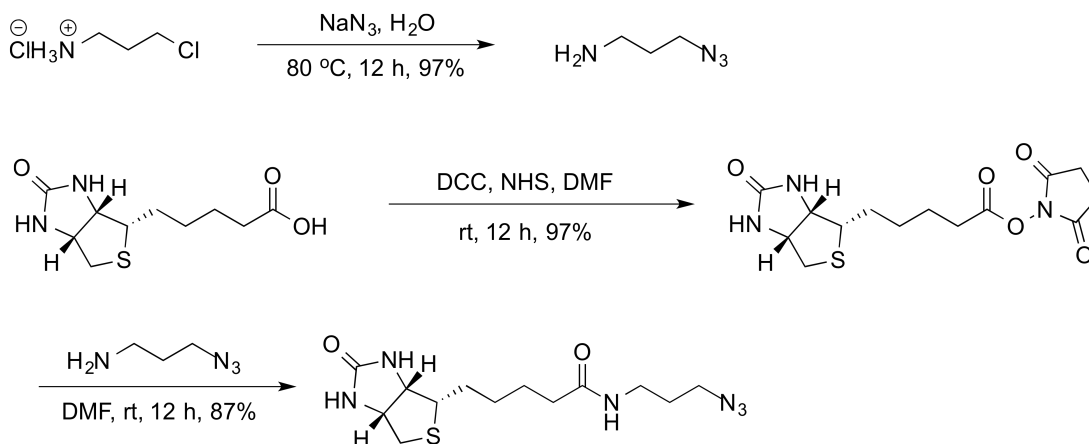
The unstable carbamate intermediate was *O*-palmitoylated with palmitoyl chloride (Scheme 2.11). Subsequently, the benzylic triethylsilyl ether was cleaved using *n*-Bu<sub>4</sub>NF. The final step of the ABP's synthesis was deoxyfluorination of the benzylic hydroxyl group, achieved by reaction with diethylaminosulfur trifluoride (DAST). The final product of the synthesis was characterized by <sup>1</sup>H-, <sup>13</sup>C- and <sup>19</sup>F- NMR as well as high resolution mass spectrometry.



**Scheme 2.11** Acylation, deprotection and deoxyfluorination of the carbamate intermediate.

#### 4.5. Synthesis of the biotin-azide tag

The ABP was designed to be a clickable probe which can be conjugated with an azide tag through copper (I)-catalyzed click chemistry for in-gel visualization and on-bead affinity enrichment. To fulfill this target, biotin-azide tags were synthesized (Scheme 2.12).



**Scheme 2.12 Syntheses of the biotin-azide tag.** The synthesis of the 3-azidopropylamine precursor is also presented.

To accomplish the synthesis of the biotin-azide tag, D-biotin was first coupled with *N*-hydroxysuccinimide (NHS) to afford the reactive biotin-NHS ester, which was subsequently linked to the 3-azido-1-propanamine. The product was purified by flash chromatography column.

## 5. Synthesis and Characterization of Compounds

### 5.1. General experimental procedures

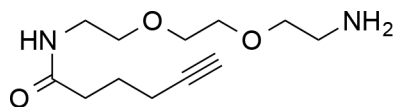
All reactions were carried out in oven-dried glassware under a positive pressure of nitrogen. Thin-layer chromatography (TLC) was performed using silica gel 60 F<sup>254</sup> plates from EMD Chemicals Inc. TLC plates were visualized by exposure to 254 nm ultraviolet (UV) light and/or were stained by phosphomolybdic acid (10% in ethanol) or ninhydrin (0.3% in 3% AcOH/96.7% *n*-butanol). Flash column chromatography was employed to purify products using a positive pressure of nitrogen.

Commercial reagents and solvents were used as received. 5-Hexynoic acid was obtained from GFS Chemicals. 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC), fluorenylmethyloxycarbonyl chloride (FmocCl), and chlorotriethylsilane were obtained from Oakwood Chemical. Diethylaminosulfur trifluoride was obtained from Alfa Aesar. All the other reagents were purchased from Sigma-Aldrich.

Proton nuclear magnetic resonance ( $^1\text{H}$ -NMR) spectra were measured on Bruker Avance III HD Ascend 600 MHz, Bruker Avance III Ultra-Shield 400 MHz, or Bruker Avance 300 MHz spectrometers. Chemical shifts are reported in parts per million (ppm,  $\delta$  scale) and are calibrated using residual non-deuterated solvent as an internal reference ( $\text{CHCl}_3$ :  $\delta$  7.26). Chemical shift data for all  $^1\text{H}$ -NMR spectra are recorded in the following format: chemical shift (multiplicity, coupling constant, integration). The multiplicity of each resonance is reported as: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. Carbon nuclear magnetic resonance ( $^{13}\text{C}$ -NMR) spectra were measured using the same spectrometers as was used for  $^1\text{H}$ -NMR. Carbon chemical shifts are recorded in units of parts per million (ppm,  $\delta$  scale) and are referenced to the carbon resonances of the deuterated solvent ( $\text{CDCl}_3$ :  $\delta$  77.0). Fluorine nuclear magnetic resonance ( $^{19}\text{F}$ -NMR) spectra were recorded using the same spectrometers, and all chemical shift values are reported in ppm on the  $\delta$  scale.

High-resolution fast atom bombardment (FAB) mass spectra were obtained with a JEOL JMS-600H mass spectrometer. High-resolution electrospray ionization (ESI) mass spectra were obtained with an Agilent Technologies 6530 Accurate-Mass Q-TOF mass spectrometer.



5.2. Synthesis of *N*-(2-(2-(2-aminoethoxy)ethoxy)ethyl)hex-5-ynamide (**1**)

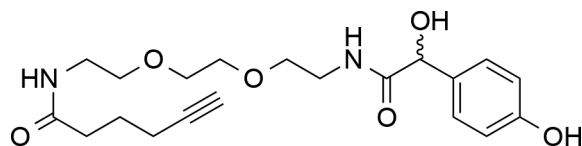
To an cooled solution at 0 °C of 2,2'-(ethylenedioxy)bis(ethylamine) (6.36 mL, 6.46 g, 42.9 mmol) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (40 mL) was added a pre-mixed solution of 5-hexynoic acid (3.95 mL, 4.01 g, 35.8 mmol), *N,N'*-dicyclohexylcarbodiimide (11.1 g, 53.7 mmol), and 4-dimethylaminopyridine (0.309 g, 2.49 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) dropwise via addition funnel. The reaction mixture was allowed to warm to room temperature. After 20 h of stirring, the crude mixture was filtered and the solvent was removed under vacuum. The product was purified by flash column chromatography (5% MeOH in CH<sub>2</sub>Cl<sub>2</sub> with 1% Et<sub>3</sub>N) to give **1** as a colorless oil (6.13 g, 25.4 mmol, 71%).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ 6.87 (br s, 1H), 3.43 (s, 4H), 3.32-3.38 (m, 4H), 3.22-3.26 (m, 2H), 2.68 (t, *J* = 5.2 Hz, 2H), 2.14 (t, *J* = 7.6 Hz, 2H), 2.04-2.08 (m, 2H), 1.86 (t, *J* = 2.4 Hz, 1H), 1.62-1.70 (m, 2H), 1.45 (br s, 2H).

<sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>) δ 172.4, 83.6, 73.0, 70.1, 70.0, 69.8, 69.1, 41.5, 39.0, 34.9, 24.2, 17.8.

HRMS-FAB: calculated for C<sub>12</sub>H<sub>22</sub>N<sub>2</sub>O<sub>3</sub> ([M+ Na]<sup>+</sup>) 265.1528, observed 265.1533.

5.3. Synthesis of *N*-(2-(2-(2-(2-hydroxy-2-(4-hydroxyphenyl)acetamido)ethoxy)ethoxy)ethyl)hex-5-ynamide (**2**)



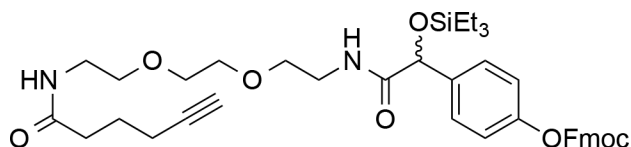
To a solution of 4-hydroxymandelic acid (14.4 mmol, 2.68 g) and **1** (9.59 mmol, 2.32 g) in *N,N*-dimethylformamide (50 mL) was added EDC (14.4 mmol, 2.76 g) and hydroxybenzotriazole (0.96 mmol, 0.130 g). The reaction was heated at 80 °C. After 19 h, the solvent was removed under vacuum. The product was purified by flash chromatography (5% MeOH in EtOAc) to give **2** as a colorless oil (3.27 g, 8.34 mmol, 87%).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ 7.37 (d, *J* = 8.0 Hz, 1H), 7.16 (d, *J* = 12.8 Hz, 2H), 7.08 (t, *J* = 5.6 Hz, 1H), 6.70 (d, *J* = 9.6 Hz, 2H), 6.33 (t, *J* = 9.6 Hz, 1H), 4.96 (s, 1H), 4.50 (s, 1H), 3.57 (m, 4H), 3.47 (m, 4H), 3.39 (m, 4H), 2.28 (t, *J* = 7.2 Hz, 2H), 2.23 (m, 2H), 2.01 (t, *J* = 2.6 Hz, 1H), 1.82 (m, *J* = 7.1 Hz, 2H).

<sup>13</sup>C-NMR (100 MHz, CDCl<sub>3</sub>) δ 173.4, 173.2, 157.0, 130.7, 128.4, 115.8, 83.6, 74.0, 70.4, 70.1, 69.8, 69.5, 69.4, 39.4, 39.1, 35.0, 24.1, 17.8.

HRMS-FAB: calculated for C<sub>20</sub>H<sub>28</sub>N<sub>2</sub>O<sub>6</sub> ([M+ Na]<sup>+</sup>) 415.1845, observed 415.1860.

#### 5.4. Synthesis of (9*H*-fluoren-9-yl)methyl(4-(3,3-diethyl-6,17-dioxo-4,10,13-trioxa-7,16-diaza-3-siladocos-21-yn-5-yl)phenyl) carbonate (**3**)



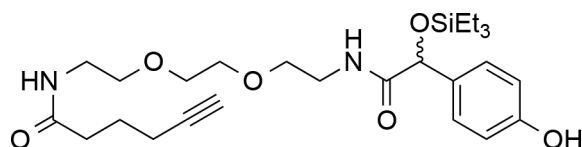
To a solution of **2** (3.20 g, 8.16 mmol) in THF (100 mL) was added diisopropylethylamine (1.71 mL, 1.27 g, 9.80 mmol) and 9-fluorenylmethoxycarbonyl chloride (2.54 g, 9.80 mmol dissolved in 20 mL of CH<sub>2</sub>Cl<sub>2</sub>). The reaction was stirred for 3 h at room temperature and then quenched by addition of 10 mL of MeOH. The solvent was removed under vacuum. The reaction mixture was partitioned between EtOAc (100 mL) and 0.1 M HCl (100 mL). The organic layer was dried over MgSO<sub>4</sub> and concentrated under vacuum. The crude product was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (100 mL), and to this solution was added chlorotriethylsilane (3.00 mL, 2.69 g, 17.9 mmol) and diisopropylethylamine (5.92 mL, 4.40 g, 34.0 mmol). After 12 h, the reaction mixture was partitioned between EtOAc (100 mL) and 0.1 M HCl (100 mL). The organic layer was dried over MgSO<sub>4</sub> and concentrated under vacuum. The product was purified by flash column chromatography (20% to 100% EtOAc in hexanes) to give **3** as a yellow oil (3.4 g, 4.73 mmol, 58%).

<sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>) δ 7.81 (d, *J* = 7.8 Hz, 2H), 7.66 (d, *J* = 7.8 Hz, 2H), 7.53 (d, *J* = 8.4, 2H), 7.45 (d, *J* = 7.2 Hz, 2H), 7.37 (d, *J* = 7.8 Hz, 2H), 7.25 (br s, 1H), 7.18 (d, *J* = 8.4 Hz, 2H) 6.03 (br s, 1H), 5.12 (s, 1H), 4.55 (d, *J* = 7.8 Hz, 2H), 3.54-3.61 (m, 12H), 2.34 (t, *J* = 7.2 Hz, 2H), 2.27 (d, *J* = 2.4 Hz, 2H), 1.99 (s, 1H), 1.88 (t, *J* = 7.2 Hz, 2H), 0.94 (t, *J* = 7.8 Hz, 9H), 0.62-0.66 (m, 6H).

<sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>) δ 172.1, 172.0, 153.5, 150.8, 143.18, 143.16, 141.4, 137.8, 128.0, 127.3, 125.2, 120.8, 120.2, 83.6, 74.8, 70.5, 70.30, 70.28, 70.0, 69.8, 69.1, 60.4, 46.7, 39.2, 38.8, 35.0, 24.2, 17.9, 6.6, 4.6.

HRMS-FAB: calculated for C<sub>41</sub>H<sub>52</sub>N<sub>2</sub>O<sub>8</sub>Si ([M+Na]<sup>+</sup>) 751.3391, observed 751.3375.

5.5. Synthesis of *N*-(3,3-diethyl-5-(4-hydroxyphenyl)-6-oxo-4,10,13-trioxa-7-aza-3-silapentadecan-15-yl)hex-5-ynamide (**4**)



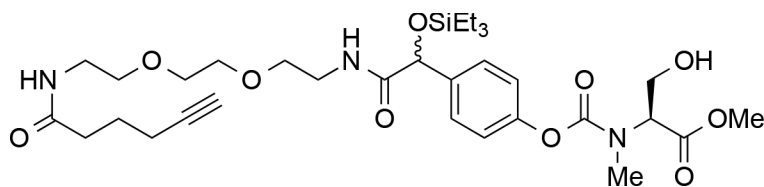
To a solution of **3** (2.60 g, 3.57 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was added piperidine (4.18 mL, 3.60 g, 42.3 mmol). The reaction was stirred at room temperature for 10 h. The solvent was removed under vacuum. The crude product was partitioned between CH<sub>2</sub>Cl<sub>2</sub> (100 mL) and 0.1 M HCl (100 mL). The organic layer was washed with 0.1 M HCl (2 × 100 mL), dried over MgSO<sub>4</sub>, and then concentrated under vacuum. The product was purified by flash column chromatography (0% to 5% MeOH in CH<sub>2</sub>Cl<sub>2</sub>) to give **4** as a colorless oil (1.5 g, 2.89 mmol, 81%).

<sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>) δ 7.37 (br m, 1H), 7.18 (d, *J* = 8.4 Hz, 2H), 6.65 (d, *J* = 8.4 Hz, 2H), 6.13 (br s, 1H), 5.00 (s, 1H), 3.45-3.62 (m, 12H), 2.34 (t, *J* = 7.2 Hz, 2H), 2.26-2.27 (m, 2H), 2.00 (t, *J* = 3.0 Hz, 1H), 1.88 (t, *J* = 7.2 Hz, 2H), 0.90 (t, *J* = 6.4 Hz, 9H), 0.59 (m, 6H).

<sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>) δ 173.4, 172.7, 156.8, 130.8, 127.8, 115.5, 83.5, 75.3, 70.3, 70.2, 69.9, 69.8, 69.2, 39.2, 38.8, 35.0, 24.2, 17.8, 6.6, 4.6.

HRMS-FAB: calculated for  $\text{C}_{26}\text{H}_{42}\text{N}_2\text{O}_6\text{Si}$  ( $[\text{M} + \text{Na}]^+$ ) 529.2710, observed 529.2725.

5.6. Synthesis of methyl *N*-((4-(3,3-diethyl-6,17-dioxo-4,10,13-trioxa-7,16-diaza-3-siladocos-21-yn-5-yl)phenoxy)carbonyl)-*N*-methyl-*L*-serinate (**5**)



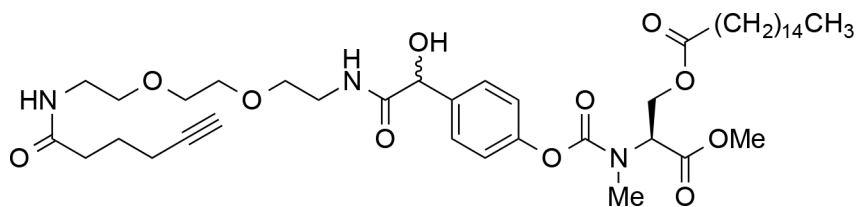
To a solution of phosgene (3.75 mL of a 15% solution in toluene, 5.25 mmol) in  $\text{CH}_2\text{Cl}_2$  (30 mL) cooled to 0 °C in an ice-water bath, a pre-mixed solution of **4** (1.33 g, 2.63 mmol) and diisopropylethylamine (0.548 mL, 0.407 g, 3.15 mmol) in  $\text{CH}_2\text{Cl}_2$  (10 mL) was added dropwise via an addition funnel. The reaction was stirred at 0 °C for 30 min. A pre-mixed solution of methyl *N*-methyl *L*-serinate (0.870 g, 6.57 mmol) and diisopropylethylamine (1.14 mL, 0.847 g, 6.57 mmol) in  $\text{CH}_2\text{Cl}_2$  (10 mL) was added. The reaction was stirred for 10 min at room temperature and was then acidified to pH = 4.0 with glacial acetic acid. The solvent was removed under vacuum and the product was purified by flash column chromatography (1% to 8% MeOH in  $\text{CH}_2\text{Cl}_2$ ) to give **5** as a colorless oil (1.40 g, 2.10 mmol, 80%).

$^1\text{H}$ -NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.32 (d,  $J$  = 9.0 Hz, 2H), 7.16 (br s, H), 6.98 (d,  $J$  = 9.0 Hz, 2H), 6.40 (br s, 1H), 4.95 (s, 1H), 4.64 (br s, 1H), 3.84-3.96 (m, 2H), 3.63 (s, 3H), 3.27-3.45 (m, 12H), 2.92-3.02 (d, 3H, rotamers of carbamate), 2.17 (t,  $J$  = 6.0 Hz, 2H), 2.07-2.13 (m, 2H), 1.90 (s, 1 H), 1.66-1.75 (m, 2H), 0.79 (t,  $J$  = 9.0 Hz, 9H), 0.44-0.52 (m, 6H).

$^{13}\text{C}$ -NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  172.6, 172.1, 170.0, 155.3, 151.0, 136.8, 126.9, 121.3, 83.5, 74.8, 70.1, 69.64, 69.56, 69.2, 60.1, 52.2, 39.0, 38.7, 34.8, 24.2, 17.8, 6.5, 4.4.

HRMS-FAB: calculated for  $\text{C}_{32}\text{H}_{51}\text{N}_3\text{O}_{10}\text{Si}$  ( $[\text{M}^+ \text{Na}]^+$ ) 688.3241, observed 688.3266.

5.7. Synthesis of (2*S*)-2-(((4-(1-hydroxy-2,13-dioxo-6,9-dioxo-3,12-diazaoctadec-17-yn-1-yl) phenoxy)carbonyl)(methyl)amino)-3-methoxy-3-oxopropyl palmitate (**6**)



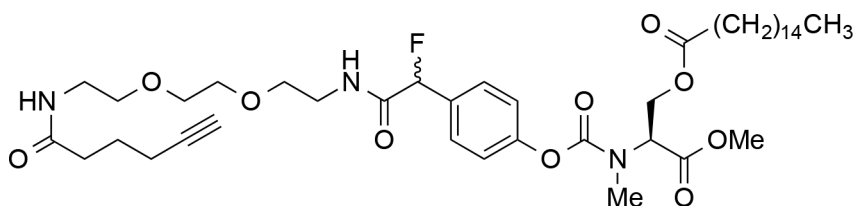
To a cooled solution of **5** (0.430 g, 0.646 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (9 mL) was added a pre-mixed solution of palmitoyl chloride (0.259 mL, 0.235 g, 0.855 mmol) and diisopropylethylamine (0.148 mL, 0.110 g, 0.855 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (9 mL). The reaction mixture was stirred at room temperature for 10 h. After concentration under vacuum, the crude reaction mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> (100 mL) and 0.1 M NaOH (100 mL). The organic layer was dried over MgSO<sub>4</sub> and then concentrated under vacuum. The crude material was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (5 mL) and cooled to -78 °C. Tetra-*n*-butylammonium fluoride (0.33 mL of a 1.0 M solution in THF, 0.33 mmol) was added dropwise. The reaction was stirred at 0 °C for 1 h, and then quenched by addition of water (1.0 mL). The crude reaction mixture was partitioned between water (20 mL) and CH<sub>2</sub>Cl<sub>2</sub> (20 mL). The organic layer was dried over MgSO<sub>4</sub> and concentrated under vacuum. The crude product was purified by flash column chromatography (25% THF in CH<sub>2</sub>Cl<sub>2</sub>) to give **6** as a white solid (0.28 g, 0.35 mmol, 54%).

<sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>) δ 7.47 (d, *J* = 6.6 Hz, 2H), 7.14 (d, *J* = 8.4 Hz, 2H), 6.08 (br s, 1H), 5.03 (s, 1H), 4.51-4.79 (m, 2H), 3.80 (s, 3H), 3.41-3.61 (m, 12H), 3.05-3.14 (d, 3H, rotamers of carbamate), 2.26-2.37 (m, 4H), 2.00 (s, 1H), 1.84-1.88 (m, 2H), 1.62-1.65 (t, *J* = 6.6 Hz, 2H), 1.26 (m, 24H), 0.89 (t, *J* = 7.2 Hz, 3H).

$^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  173.3, 173.2, 172.7, 172.3, 168.8, 151.5, 135.8, 125.5, 121.9, 121.8, 121.7, 83.6, 73.6, 70.5, 70.2, 69.9, 69.6, 69.21, 69.17, 39.3, 39.24, 39.18, 34.2, 31.9, 30.3, 29.7, 29.65, 29.61, 29.5, 29.4, 29.2, 29.1, 24.9, 24.2, 22.7, 17.8, 14.1.

HRMS-FAB: calculated for  $\text{C}_{42}\text{H}_{67}\text{N}_3\text{O}_{11}$  ( $[\text{M} + \text{Na}]^+$ ) 812.4673, observed 812.4648.

5.8. Synthesis of (2*S*)-2-(((4-(1-fluoro-2,13-dioxo-6,9-dioxo-3,12-diazaoctadec-17-yn-1-yl) phenoxy)carbonyl)(methyl)amino)-3-methoxy-3-oxopropyl palmitate (**7**)



To an ice-cooled solution of **6** (100 mg, 0.127 mmol) in  $\text{CH}_2\text{Cl}_2$  (5 mL) was added a solution of diethylaminosulfur trifluoride (16.7  $\mu\text{L}$ , 20.4 mg, 0.127 mmol) in  $\text{CH}_2\text{Cl}_2$  (1.0 mL). The reaction was stirred at 0  $^\circ\text{C}$  for 30 min, and was then quenched by addition of a saturated solution of NaCl (5.0 mL), and allowed to stir for an additional 15 min. The crude reaction mixture was partitioned between  $\text{CH}_2\text{Cl}_2$  (20 mL) and water (20 mL). The organic layer was dried over  $\text{Na}_2\text{SO}_4$  and concentrated under vacuum. The crude product was purified by flash column chromatography (10% THF in  $\text{CH}_2\text{Cl}_2$ ) to yield **7** as a white solid (80 mg, 0.10 mmol, 80%).

$^1\text{H}$ -NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.49 (d,  $J$  = 4.8 Hz, 2H), 7.19 (d,  $J$  = 8.4 Hz, 2H), 7.01 (br s, 1H), 6.08 (br s, 1H), 5.80 (d,  $J$  = 48.6 Hz, 1H), 4.50-4.68 (m, 2H), 3.80 (s, 3H), 3.46-3.65 (m, 12H), 3.05-3.14 (d, 3H, rotamers of carbamate), 2.24-2.36 (m, 4H), 1.99 (t,  $J$  =

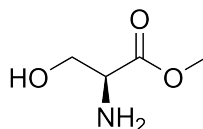
3.0 Hz, 1H), 1.85-1.88 (m, 2H), 1.62-1.65 (t,  $J = 7.2$  Hz, 2H), 1.26 (m, 24H), 0.89 (t,  $J = 7.2$  Hz, 3H).

$^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  173.3, 172.4, 168.8, 168.3, 154.9, 151.5, 135.8, 127.9, 127.85, 127.79, 121.94, 121.90, 121.8, 91.9, 90.7, 83.6, 70.2, 70.0, 69.4, 69.2, 52.7, 39.2, 38.9, 35.0, 31.9, 30.3, 29.71, 29.67, 29.6, 29.5, 29.4, 29.3, 29.13, 29.12, 24.9, 24.2, 22.7, 17.8, 14.2.

$^{19}\text{F}$ - NMR (565 MHz)  $\delta$  -176.95 (d,  $J = 49.0$  Hz).

HRMS-FAB: calculated for  $\text{C}_{42}\text{H}_{66}\text{FN}_3\text{O}_{10}$  ( $[\text{M} + \text{Na}]^+$ ) 814.4630, observed 814.4655.

### 5.9. Synthesis of methyl *L*-serinate (**8**)<sup>9</sup>



To an ice-cold solution of *L*-serine (13.0g, 0.12 mol) in MeOH (300 mL) was added 37 mL  $\text{SOCl}_2$  (60.7 g, 0.51 mol) dropwise. The reaction mixture was stirred at 0 °C for 6 hr, and was allowed to warm to room temperature overnight. The solvent was removed *in vacuo* and 200 mL ice-cold diethyl ether was added to precipitate the product. The product was filtered and washed with 20mL ice-cold diethyl ether twice and **8** (13.6 g, 95%) was used without further purification.

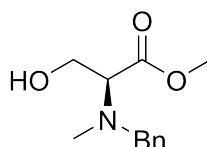
$^1\text{H}$ -NMR (600 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.18 (t,  $J = 4.2$  Hz, 1H), 4.00-4.03 (m, 1H), 3.90-3.92 (m, 1H), 3.78 (s, 3H).



$^{13}\text{C}$ -NMR (151 MHz,  $\text{D}_2\text{O}$ )  $\delta$  169.2, 59.4, 54.7, 53.6.

HRMS-FAB: calculated for  $\text{C}_4\text{H}_9\text{NO}_3$  ( $[\text{M} + \text{H}]^+$ ) 120.0661, observed 120.0665.

#### 5.10. Synthesis of methyl *N*-benzyl-*N*-methyl-*L*-serinate (**9**)



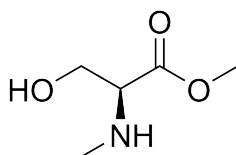
To a solution of **8** (0.5 g, 3.2 mmol) in MeOH (10 mL) was added Amberlite IRA-401 resin (2 g,  $^-\text{OH}$  form), and the mixture was stirred at room temperature for 30 min. The reaction was filtered to remove the resin. The filtrate was concentrated *in vacuo* and dissolved in MeOH (40 mL). To the solution of the neutral ester was added newly distilled benzaldehyde (0.34 mL, 3.37 mmol). The reaction mixture was stirred at room temperature for 1 hr. Subsequently,  $\text{NaBH}_3\text{CN}$  (212 mg, 3.37 mmol) was added to the solution and the mixture was stirred at room temperature for 16 hr. To the reaction mixture was added a 20% formaldehyde solution in MeOH (0.61 mL, 3.21 mmol, prepared by dissolving paraformaldehyde in MeOH at 70  $^\circ\text{C}$  with few piece of KOH pellets), and the mixture was stirred at room temperature for 5 hr before the  $\text{NaBH}_3\text{CN}$  (0.21 g 3.37 mmol) was added. The reaction was allowed to stir at room temperature for 18 hr, and then was filtered to remove the solid. The filtrate was concentrated *in vacuo*. The crude product was purified by flash chromatography column (30% EtOAc in hexanes) to afford **9** as a colorless oil (0.47 g, 67%).

$^1\text{H}$ -NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.29-7.37 (m, 5H), 3.90 (d,  $J = 18$  Hz, 1H), 3.75-3.84 (m, 6H), 3.58-3.61 (m, 1H), 2.38 (s, 3H).

$^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ )  $\delta$  171.0, 138.2, 129.0, 128.5, 127.5, 65.8, 59.2, 58.9, 51.4, 37.4.

HRMS-FAB: calculated for  $\text{C}_{12}\text{H}_{17}\text{NO}_3$  ( $[\text{M} + \text{Na}]^+$ ) 241.1106, observed 246.1102.

#### 5.11. Synthesis of methyl *N*-methyl-*L*-serinate (**10**)

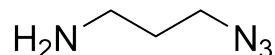


To a solution of **9** (6 g, 26.9 mmol) in MeOH (200 mL) was added  $\text{Pd}(\text{OH})_2$  on carbon (20 wt %, 1 g). The reaction mixture was shaken on a Parr hydrogenation apparatus under an  $\text{H}_2$  atmosphere (50 psi) for 24 hrs. The reaction was filtered to remove the catalyst and the filtrate was concentrated *in vacuo*. The crude product was purified by flash chromatography column (5% to 10% MeOH in  $\text{CH}_2\text{Cl}_2$ ) to afford compound **10** as an off-white powder (2.4 g, 69%).

$^1\text{H}$ -NMR (300 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  3.77 (m, 5H), 3.30 (m, 1H), 2.39 (s, 3H).

$^{13}\text{C}$ -NMR (75 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  173.0, 64.4, 61.9, 50.9, 33.0.

HRMS-FAB: calculated for  $\text{C}_5\text{H}_{11}\text{NO}_3$  ( $[\text{M} + \text{Na}]^+$ ) 156.0637, observed 156.0633.

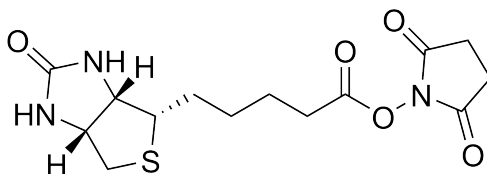
5.12. Synthesis of 3-azidopropylamine (**11**)<sup>10</sup>

To a stirred solution of 3-chloropropylamine hydrochloride (6.0 g, 4.6 mmol) in water (50 mL) was added sodium azide (9.0 g, 138 mmol). The reaction mixture was heated to 80 °C and stirred for 15 hr. The reaction was allowed to cool to room temperature before KOH pellets (4g) were added. The aqueous layer was extracted with three 50 mL aliquots of diethyl ether. The organic layers were combined and dried over anhydrous MgSO<sub>4</sub> before the solvent was carefully removed *in vacuo*. Compound **11** was obtained as a colorless liquid (4.4 g, 97%).

<sup>1</sup>H-NMR (600 MHz, CDCl<sub>3</sub>) δ 3.30 (t, *J* = 6 Hz, 2H), 2.72 (t, *J* = 7 Hz, 2H), 1.55-1.68 (m, 2H), 1.55 (t, *J* = 7 Hz, 2H).

<sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>) δ 49.1, 39.2, 32.3.

HRMS-ESI: calculated for C<sub>3</sub>H<sub>8</sub>N<sub>4</sub> ([M+ H]<sup>+</sup>) 101.0827, observed 101.0824.

5.13. Synthesis of 2,5-dioxopyrrolidin-1-yl 5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno [3,4-*d*] imidazol-4-yl)pentanoate (**12**)<sup>11</sup>

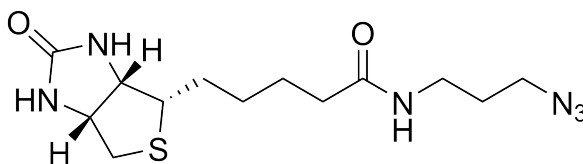
To a solution of D-biotin (4.88 g, 20 mmol) in DMF (200 mL) was added *N*-hydroxysuccinimide (3.44 g, 30 mmol). *N,N*-dicyclohexylcarbodiimide (4.5 g, 22 mmol) was added and the reaction mixture was stirred at room temperature for 24 hr. The DCU was removed by filtration before the filtrate was concentrated *in vacuo*. The product was precipitated by adding 200 mL of ice-cold diethyl ether. Compound **12** was obtained without further purification (6.2 g, 91%).

<sup>1</sup>H-NMR (600 MHz, DMSO) δ 6.41 (d, *J* = 3.6 Hz, 2H), 4.31 (t, *J* = 6.0 Hz, 1H), 4.14-4.16 (m, 1H), 3.10-3.13 (m, 1H), 2.82-2.85 (m, 4H), 2.67-2.69 (m, 2H), 1.62-1.68 (m, 3H), 1.41-1.45 (m, 3H).

<sup>13</sup>C-NMR (151 MHz, DMSO) δ 170.8, 169.4, 163.1, 61.4, 59.6, 55.7, 30.5, 28.3, 28.0, 26.0, 24.8.

HRMS-ESI: calculated for C<sub>14</sub>H<sub>19</sub>N<sub>3</sub>O<sub>5</sub>S ([M+ H]<sup>+</sup>) 342.1124, observed 342.1128.

#### 5.14. Synthesis of *N*-(3-azidopropyl)-5-((3*aS*,4*S*,6*aR*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamide (**13**)<sup>12</sup>



To a solution of **12** (1.0 g, 2.9 mmol) in DMF (10 mL) was added **11** (0.44 g, 4.4 mmol), and the mixture was stirred at room temperature overnight. The mixture was concentrated

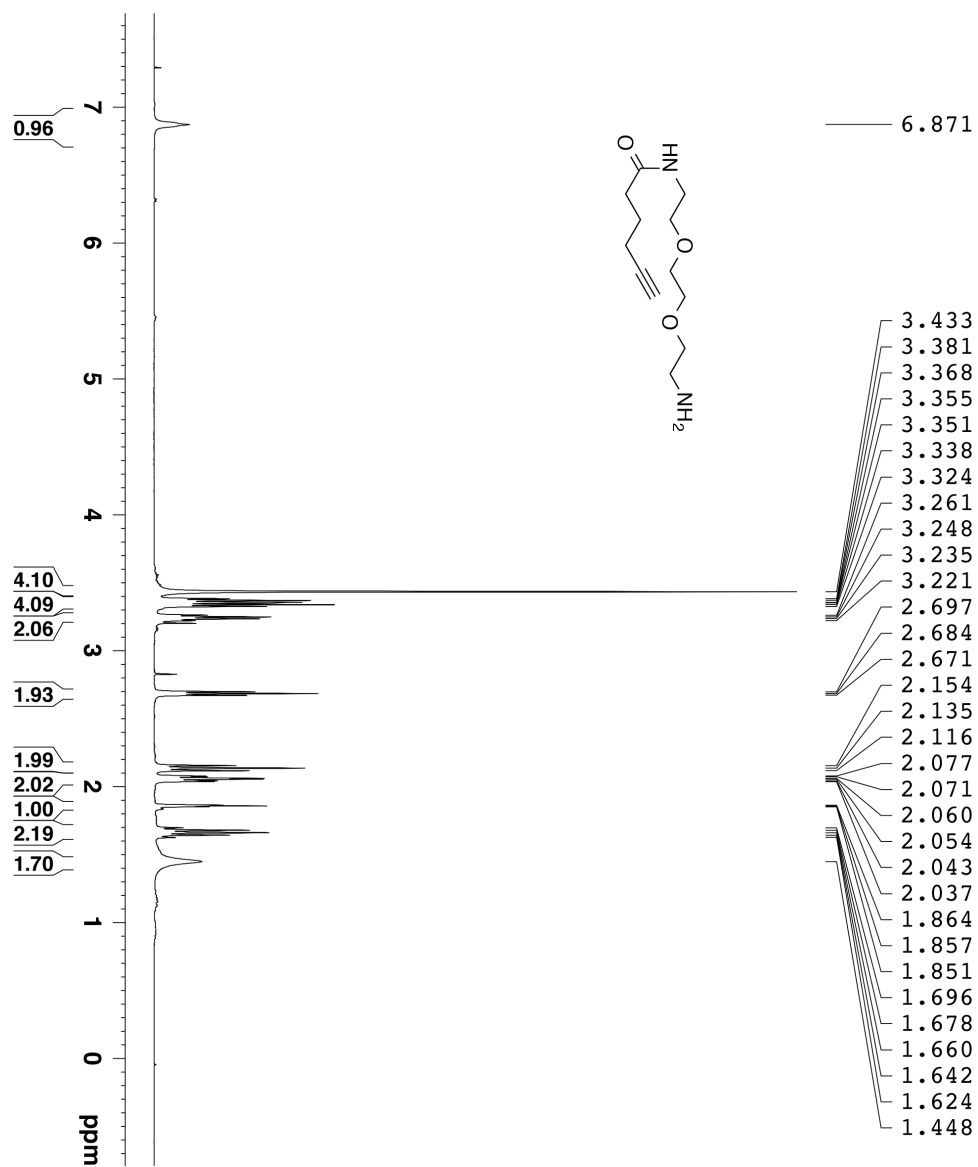
*in vacuo*, and the crude product was purified by flash chromatography column (10% MeOH in EtOAc) to afford the **13** as white solid (0.72 g, 76%).

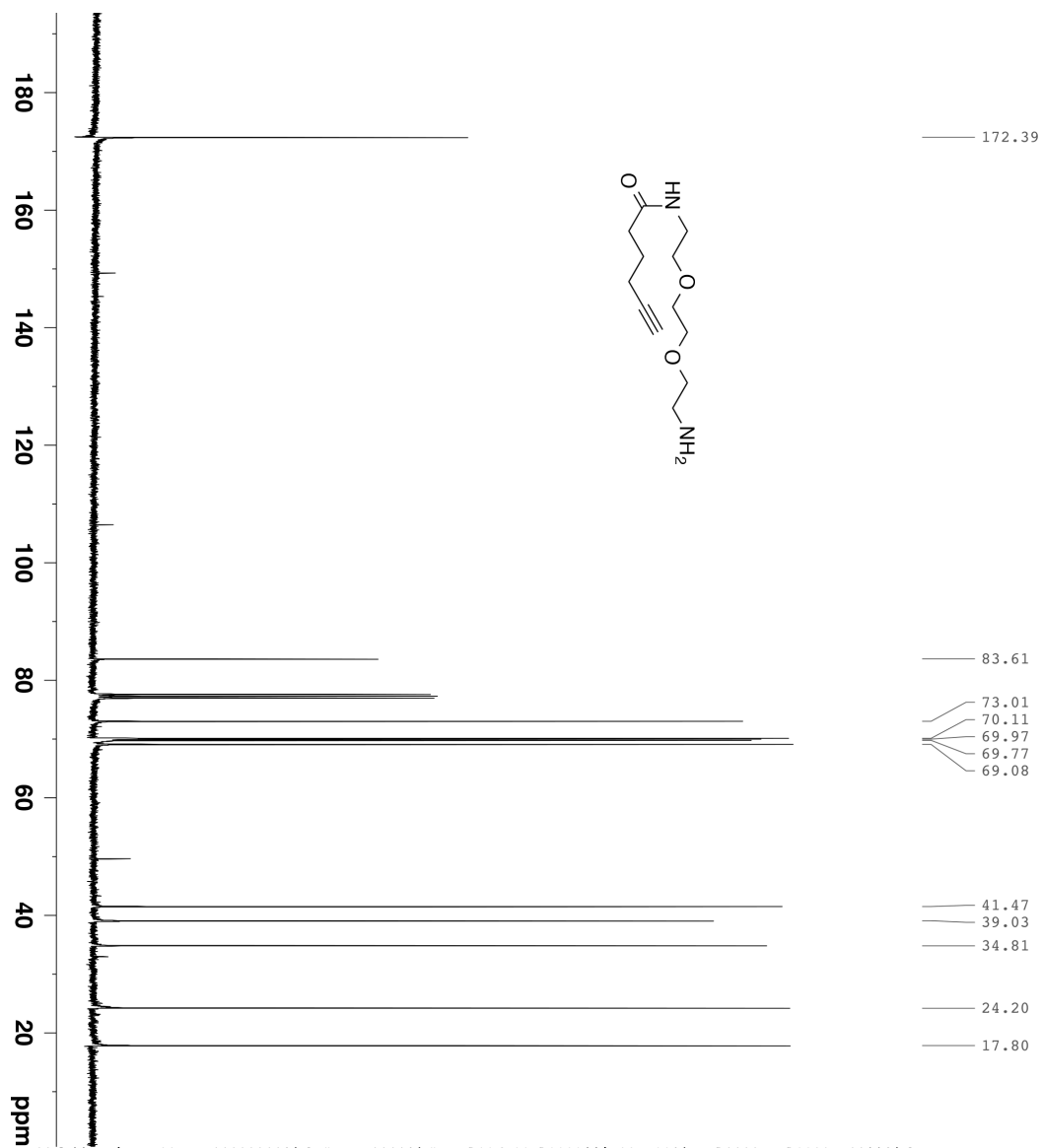
<sup>1</sup>H-NMR (600 MHz, CD<sub>3</sub>OD) δ 4.50-4.52 (m, 1H), 4.31-4.33 (m, 1H), 3.22-3.39 (m, 5H), 2.95 (m, 1H), 2.72 (d, *J* = 11 Hz, 1H), 2.23 (t, *J* = 7.0 Hz, 2H), 1.63-1.80 (m, 6H), 1.45-1.47 (m, 2H).

<sup>13</sup>C-NMR (151 MHz, CD<sub>3</sub>OD) δ 174.8, 164.7, 62.0, 60.2, 55.6, 39.6, 36.3, 35.4, 28.4, 28.3, 28.1, 25.4.

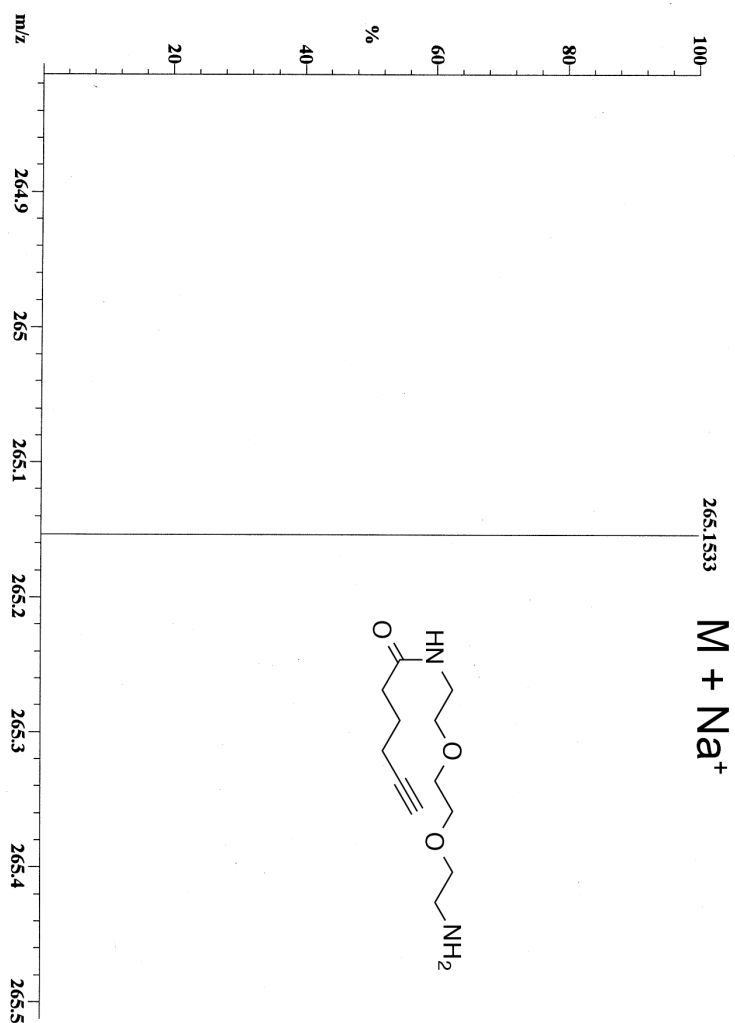
HRMS-ESI: calculated for C<sub>13</sub>H<sub>22</sub>N<sub>6</sub>O<sub>2</sub>S ([M+ H]<sup>+</sup>) 327.1603, observed 327.1596.

## 6. NMR Spectra and Mass Spectra

6.1. Spectra of *N*-(2-(2-(2-aminoethoxy)ethoxy)ethyl)hex-5-ynamide (**1**)<sup>1</sup>H-NMR

$^{13}\text{C}$ -NMR

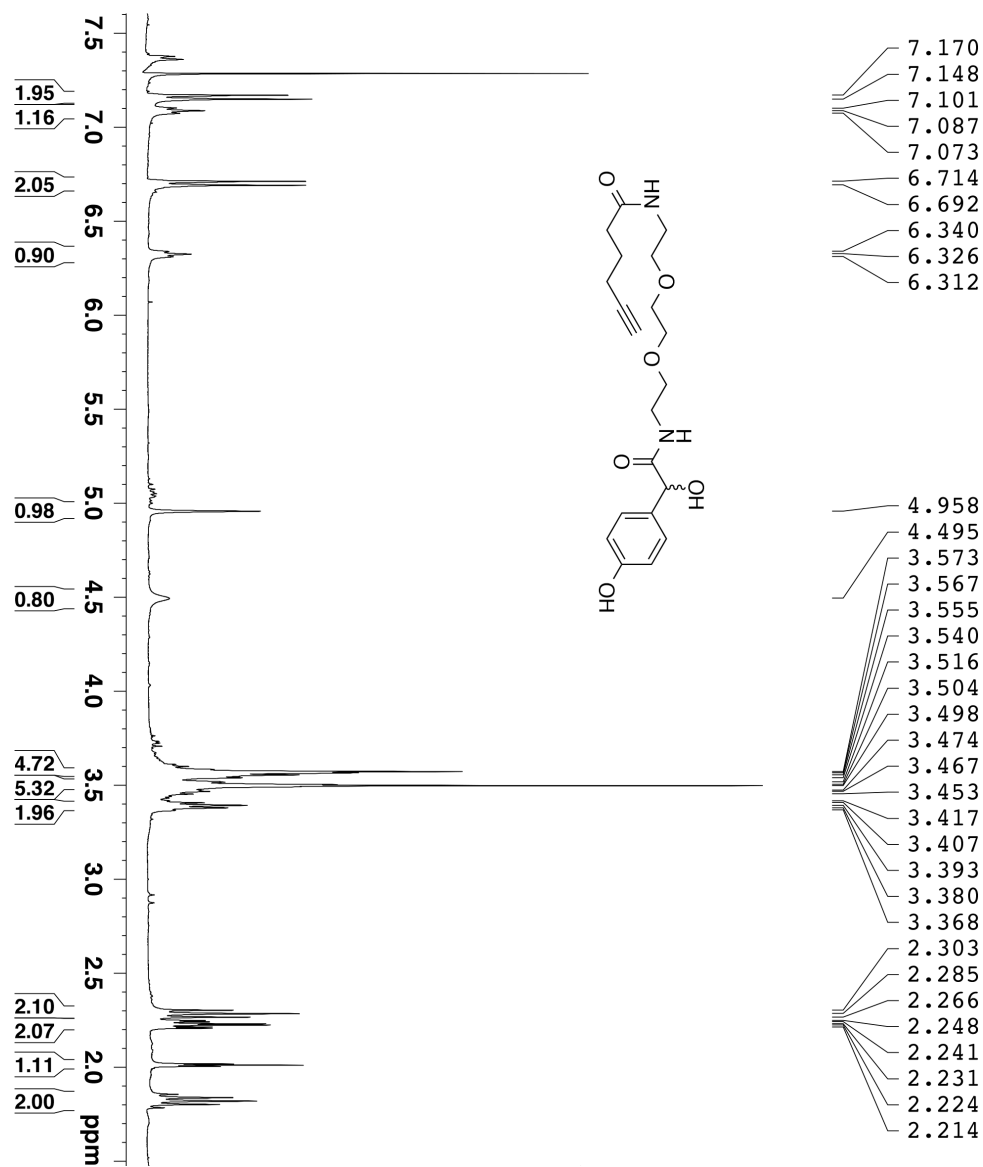
## HRMS

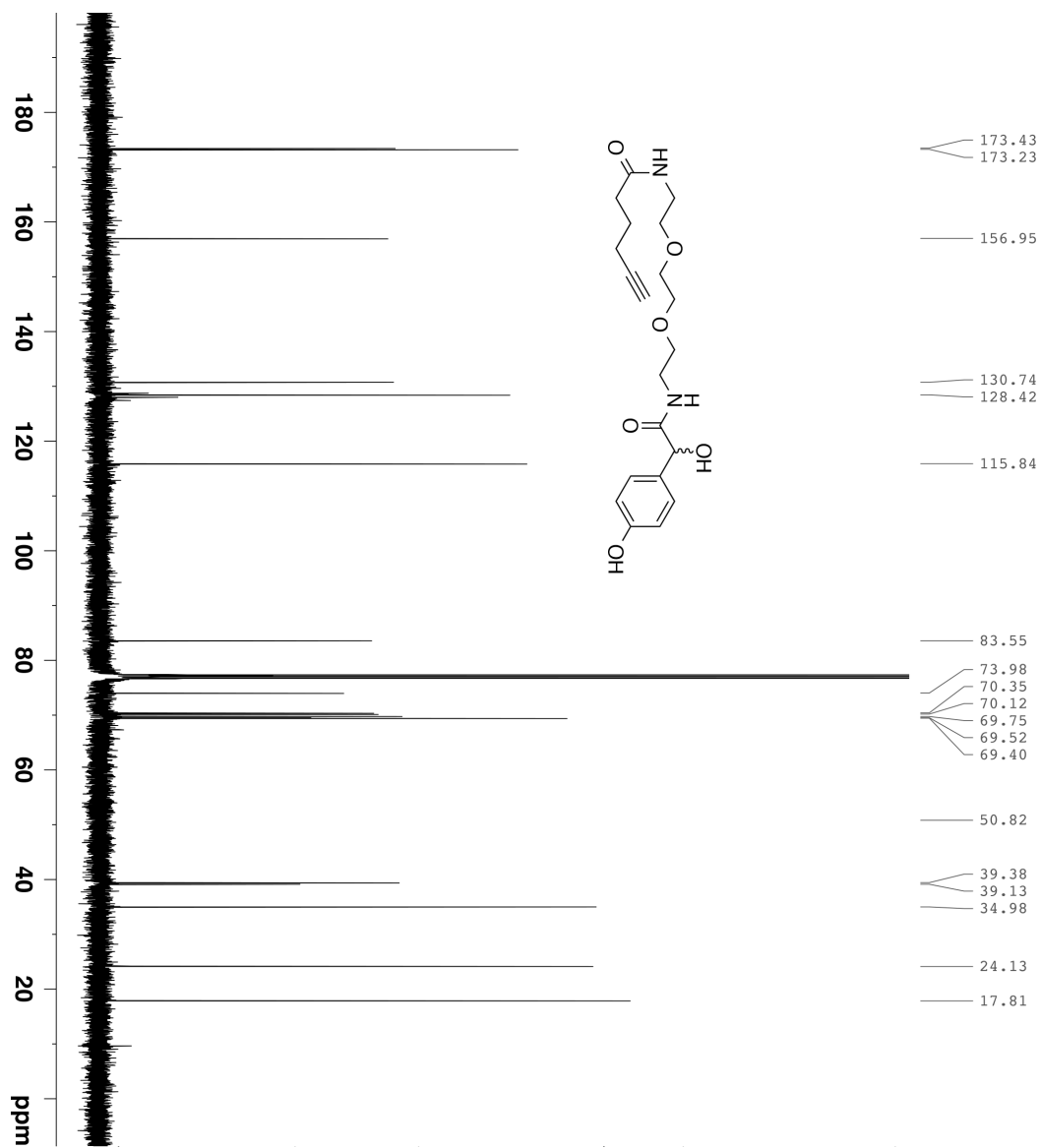




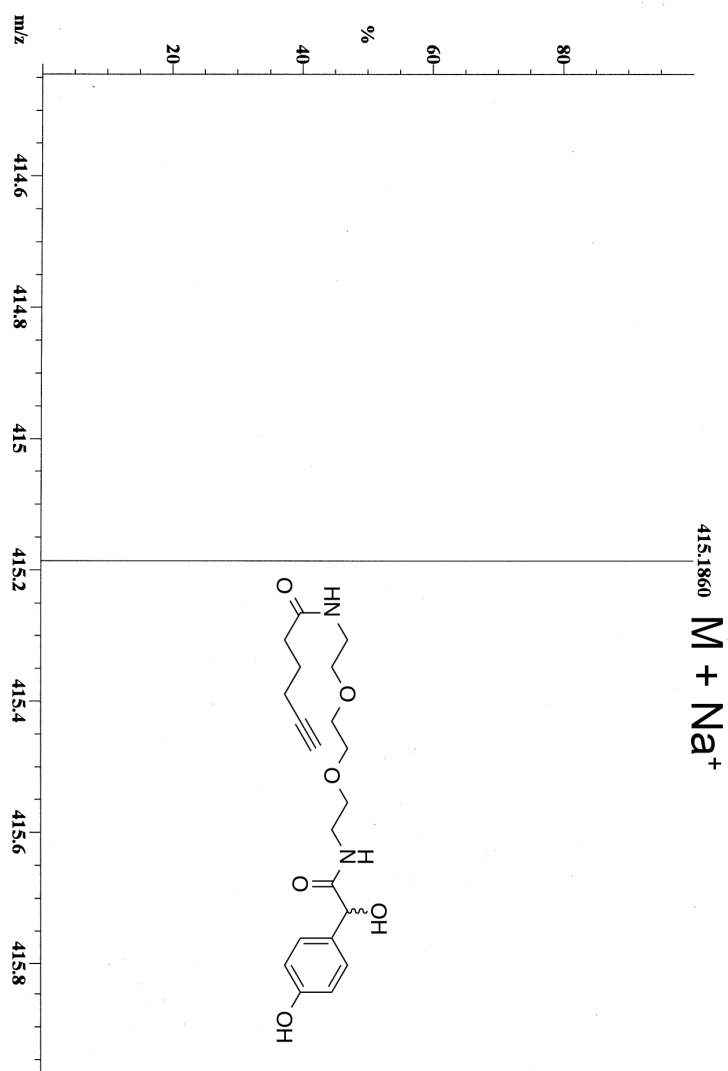
6.2. Spectra of *N*-(2-(2-(2-(2-hydroxy-2-(4-hydroxyphenyl)acetamido)ethoxy)ethoxy)ethyl)hex-5-ynamide (2)

$^1\text{H-NMR}$



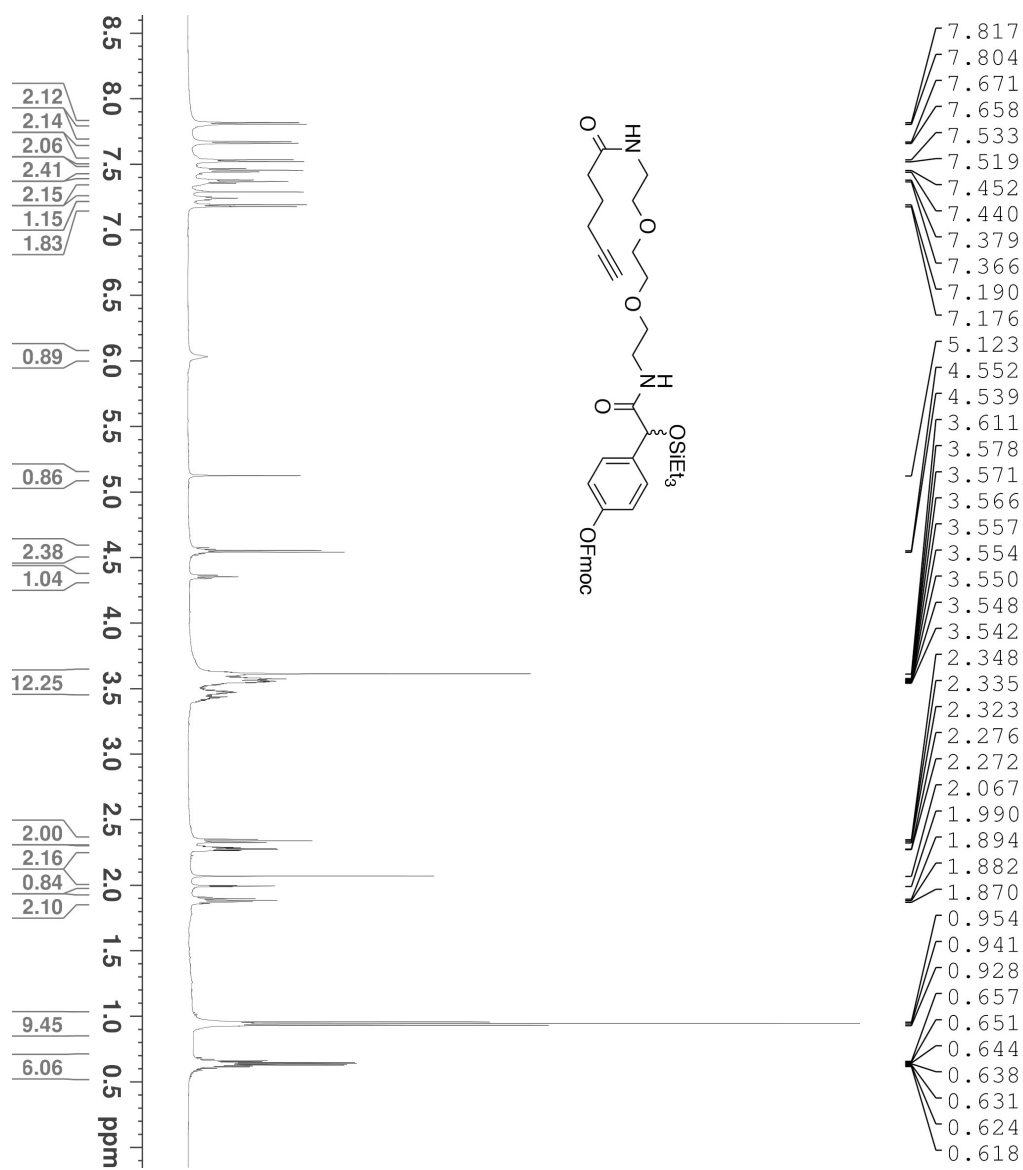
$^{13}\text{C}$ -NMR

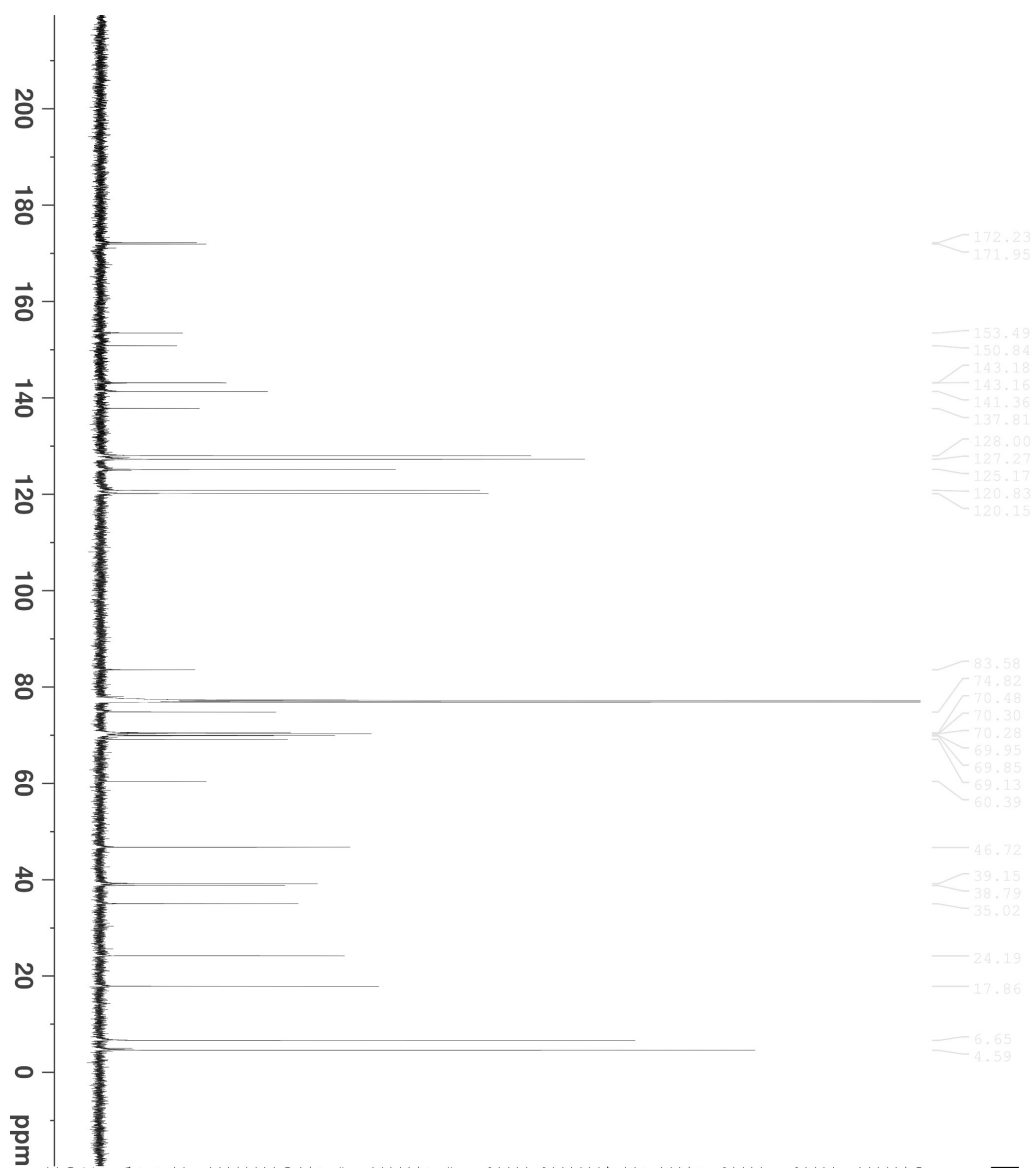
## HRMS



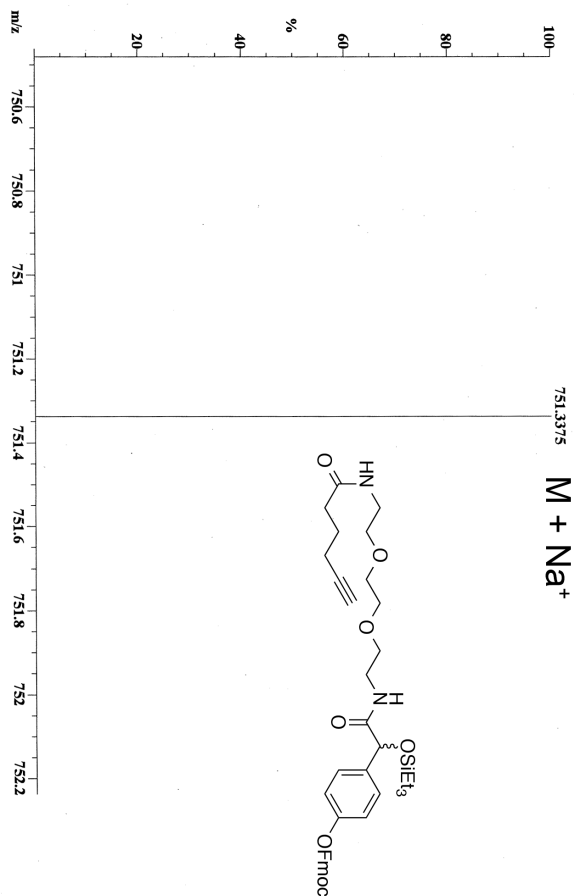
6.3. Spectra of (9*H*-fluoren-9-yl)methyl(4-(3,3-diethyl-6,17-dioxo-4,10,13-trioxa- 7,16-diaza-3-siladocos-21-yn-5-yl)phenyl) carbonate (**3**)

<sup>1</sup>H-NMR

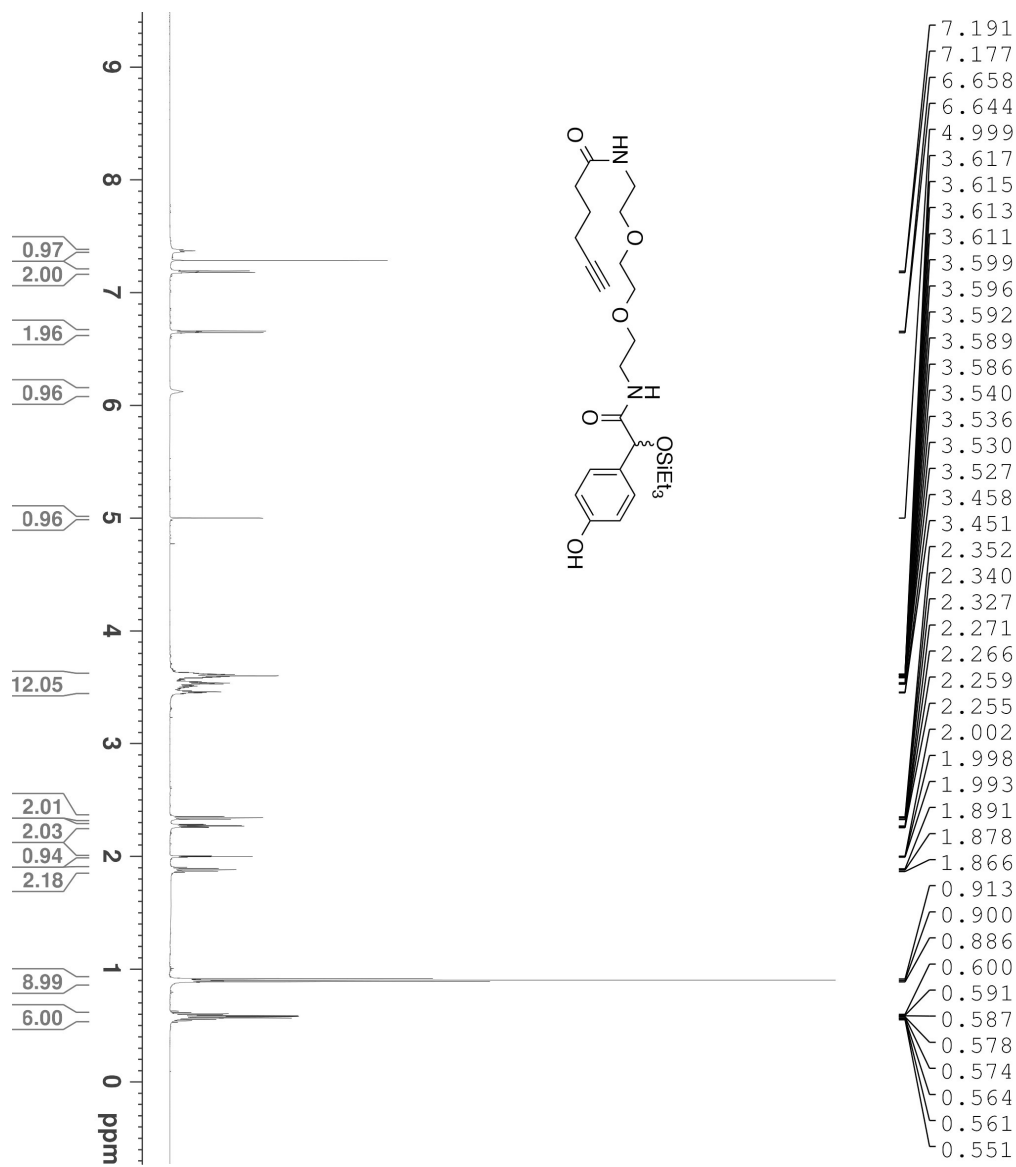


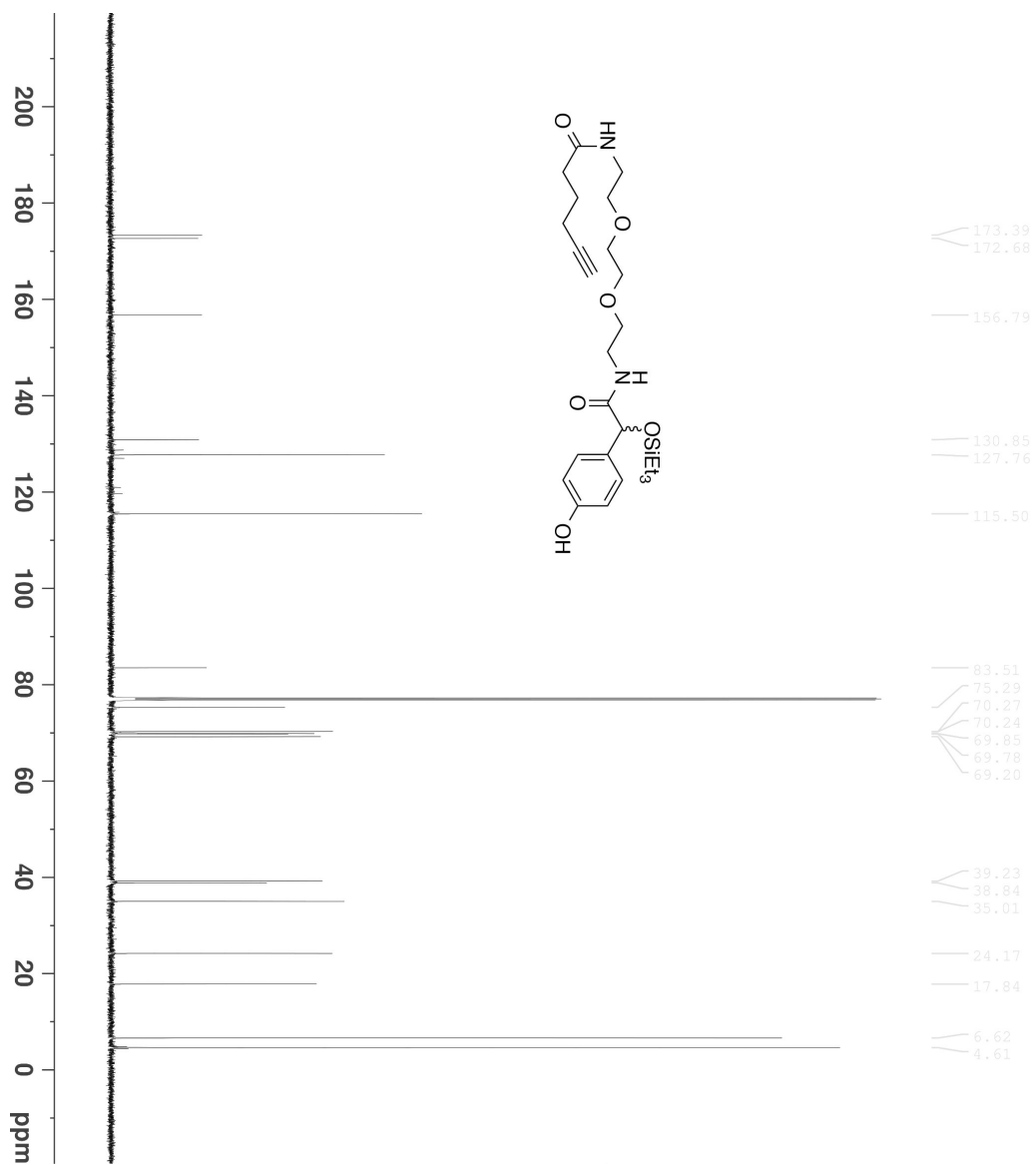
$^{13}\text{C}$ -NMR

## HRMS



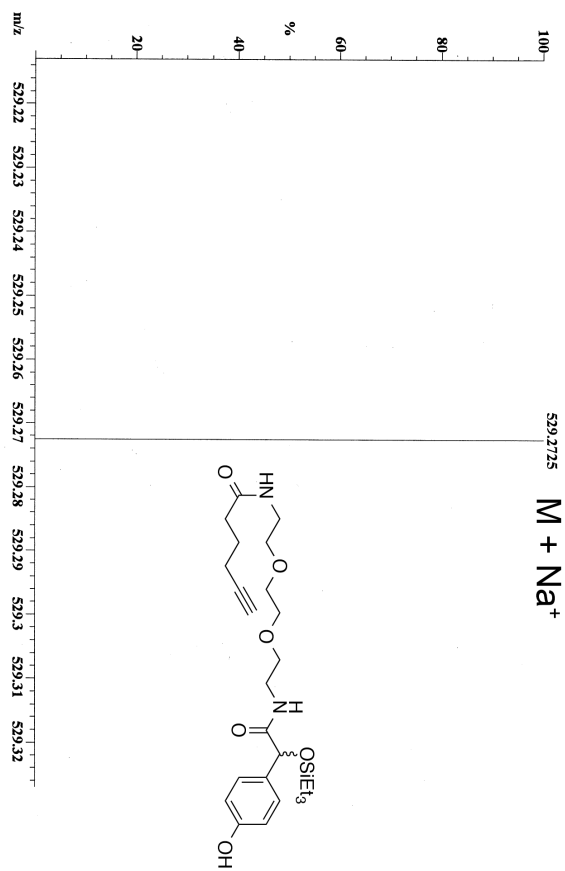
#### 6.4. Spectra of *N*-(3,3-diethyl-5-(4-hydroxyphenyl)-6-oxo-4,10,13-trioxa-7-aza-3-silapentadecan-15-yl)hex-5-ynamide (**4**)

<sup>1</sup>H-NMR

<sup>13</sup>C-NMR

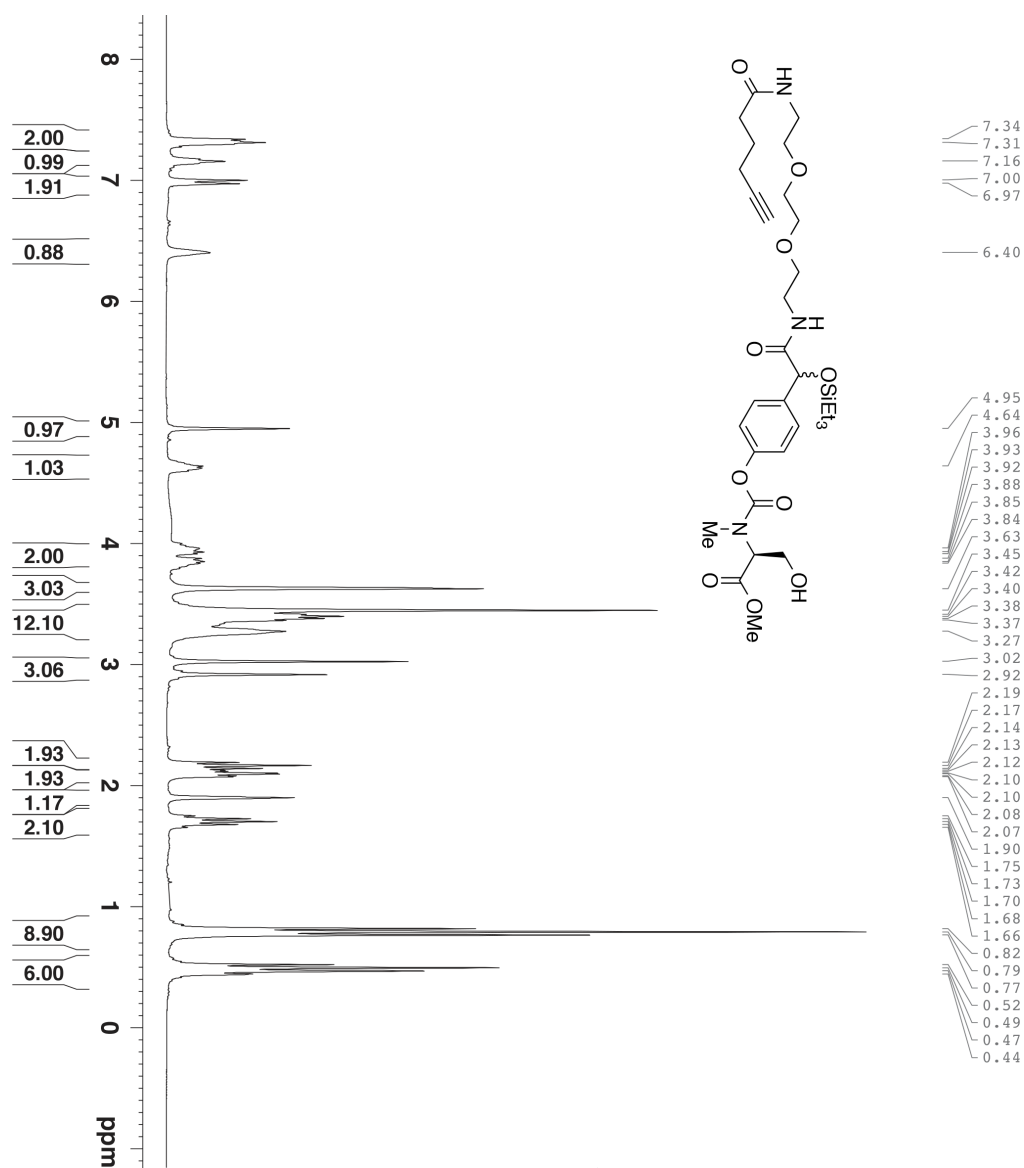


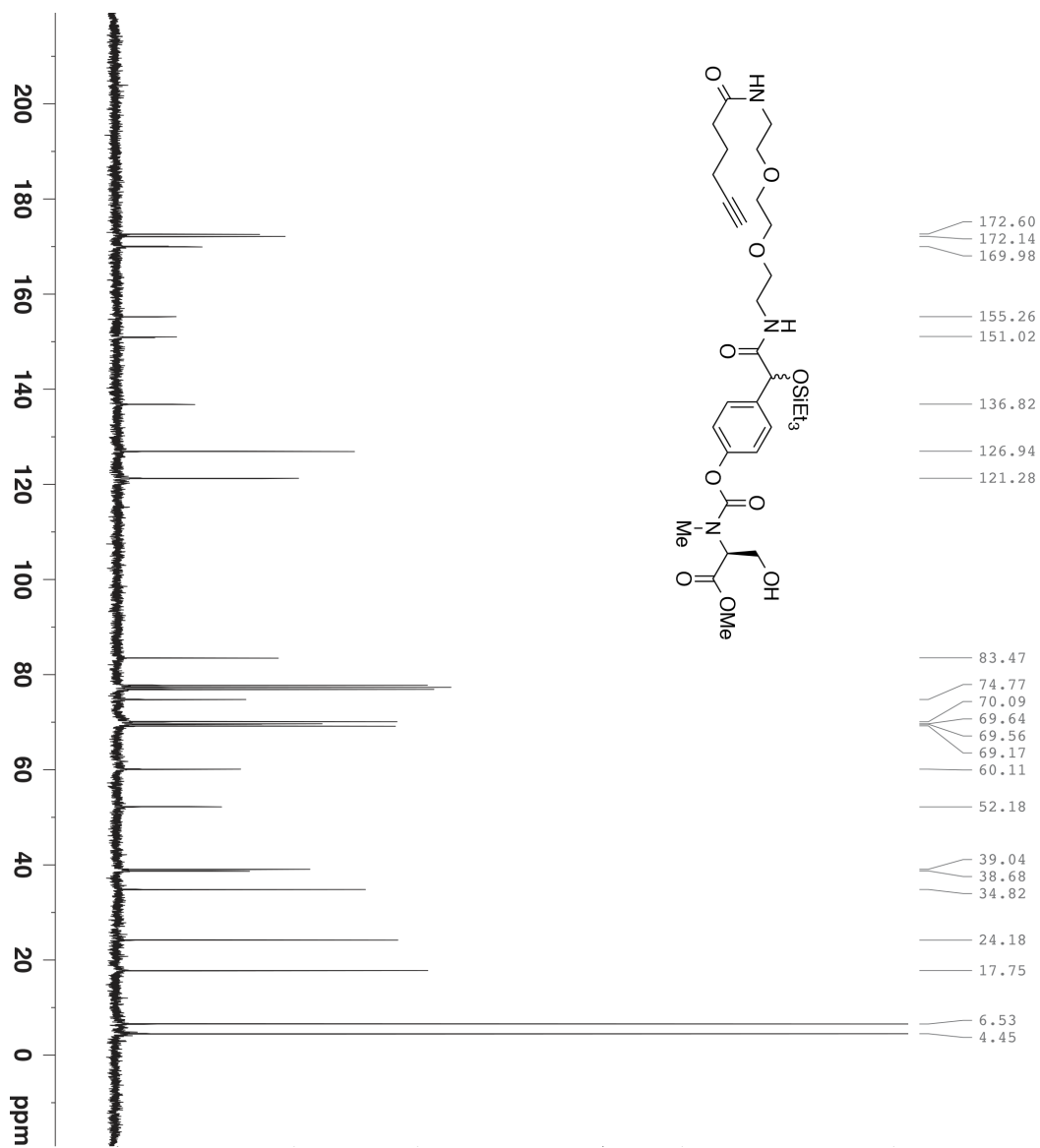
## HRMS



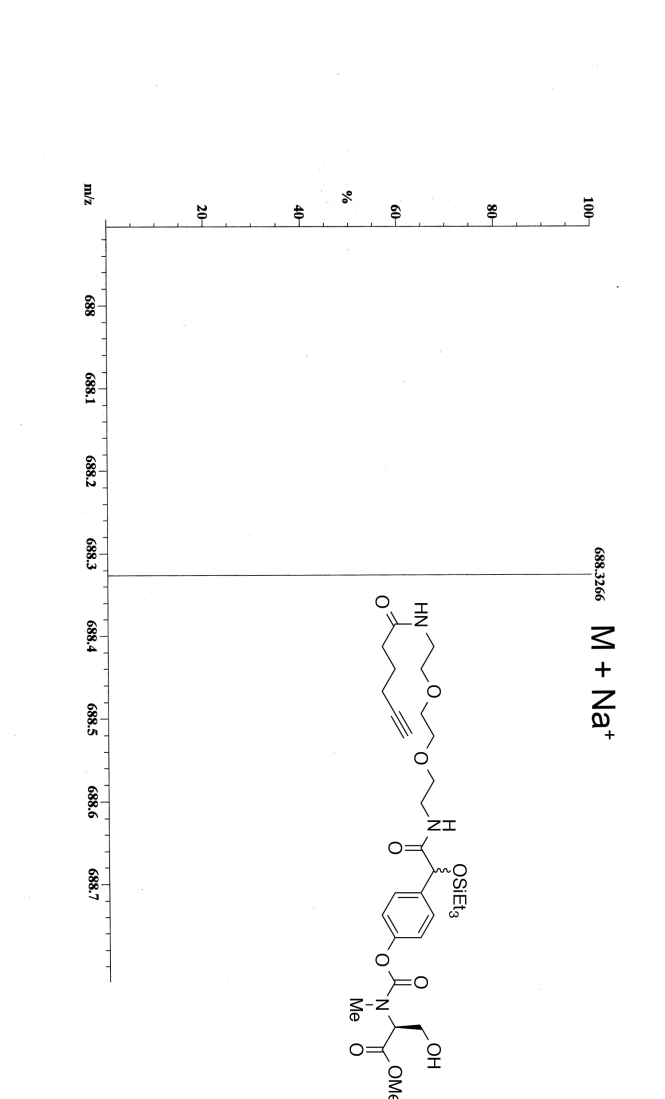
6.5. Spectra of methyl *N*-((4-(3,3-diethyl-6,17-dioxo-4,10,13-trioxa-7,16-diaza-3-siladocos-21-yn-5-yl)phenoxy)carbonyl)-*N*-methyl-*L*-serinate (**5**)

$^1\text{H-NMR}$



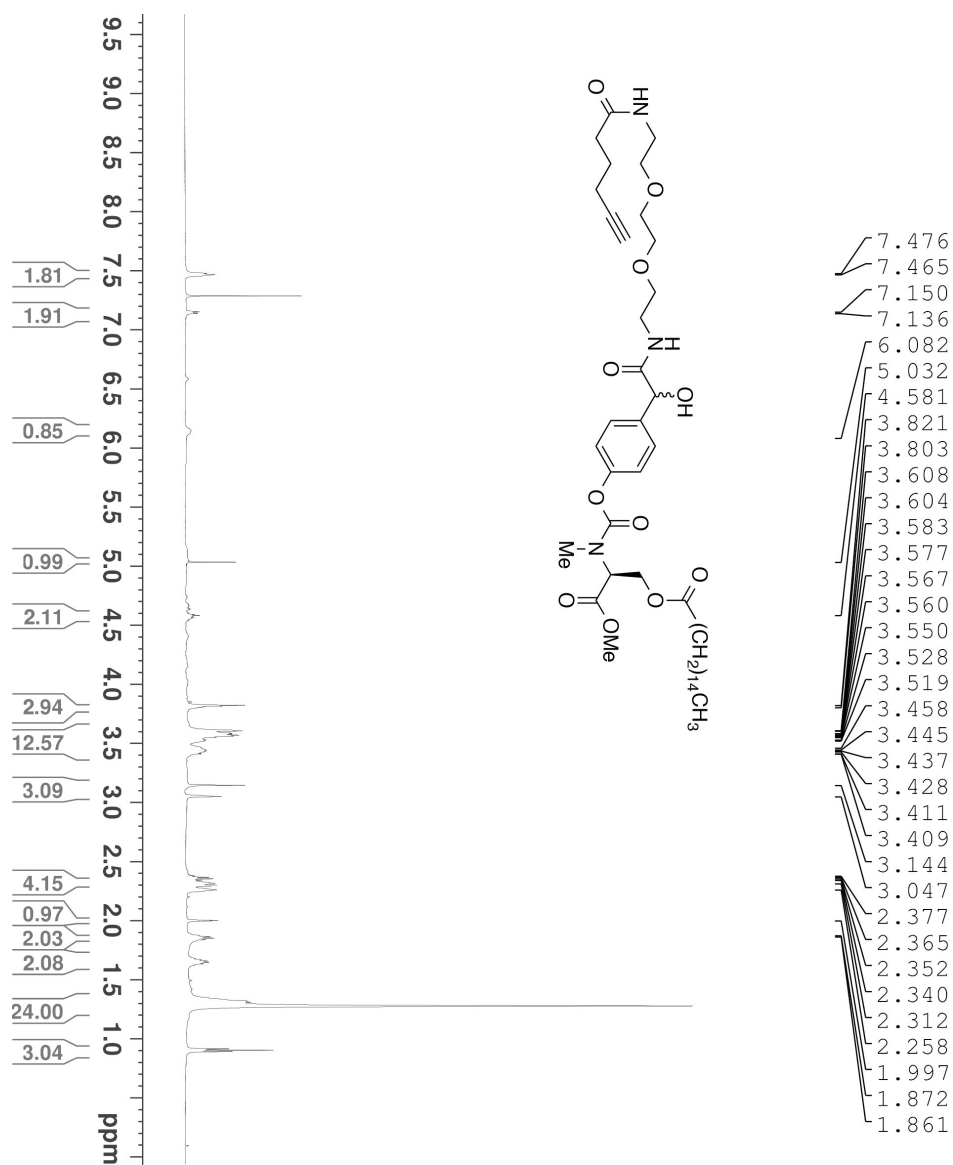
<sup>13</sup>C-NMR

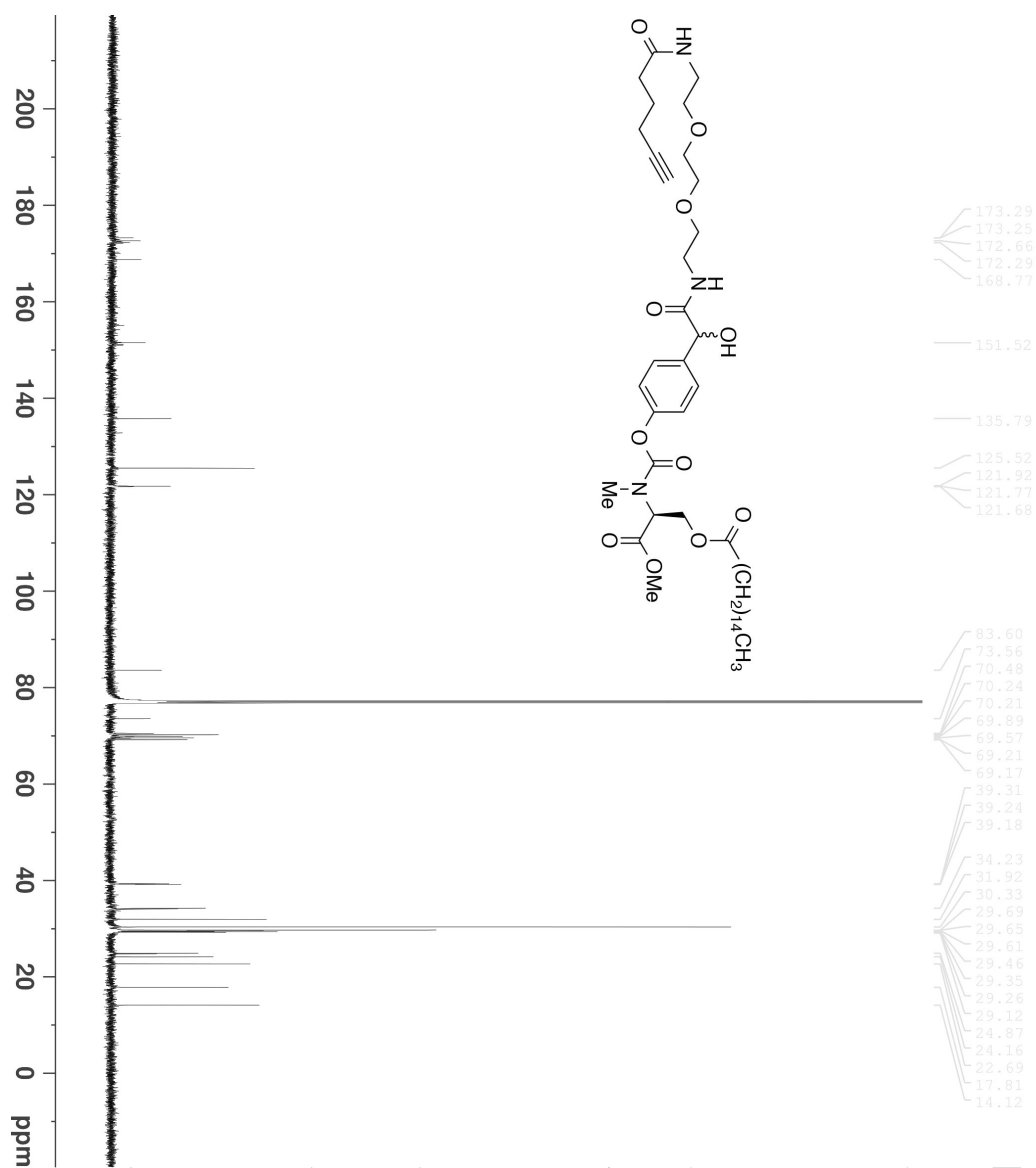
## HRMS



6.6. Spectra of (2*S*)-2-(((4-(1-hydroxy-2,13-dioxo-6,9-dioxo-3,12-diazaoctadec-17-yn-1-yl)phenoxy)carbonyl)(methyl)amino)-3-methoxy-3-oxopropyl palmitate (**6**)

<sup>1</sup>H-NMR



$^{13}\text{C}$ -NMR

812.4648  $M + Na^+$

100  
80  
60  
40  
20  
0

812.3 812.35 812.4 812.45 812.5 812.55 812.6 812.65

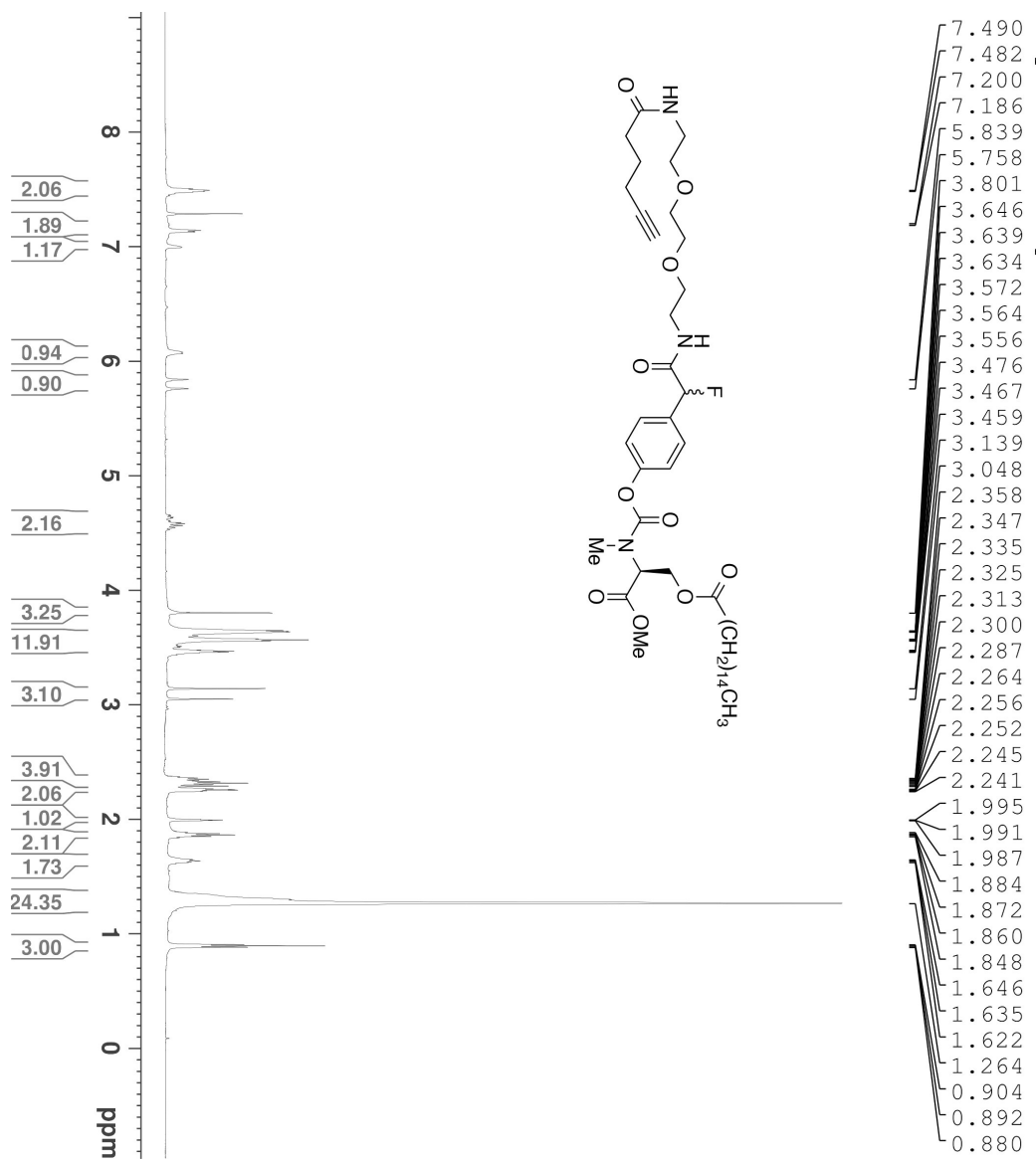
m/z

Chemical structure of compound 10:

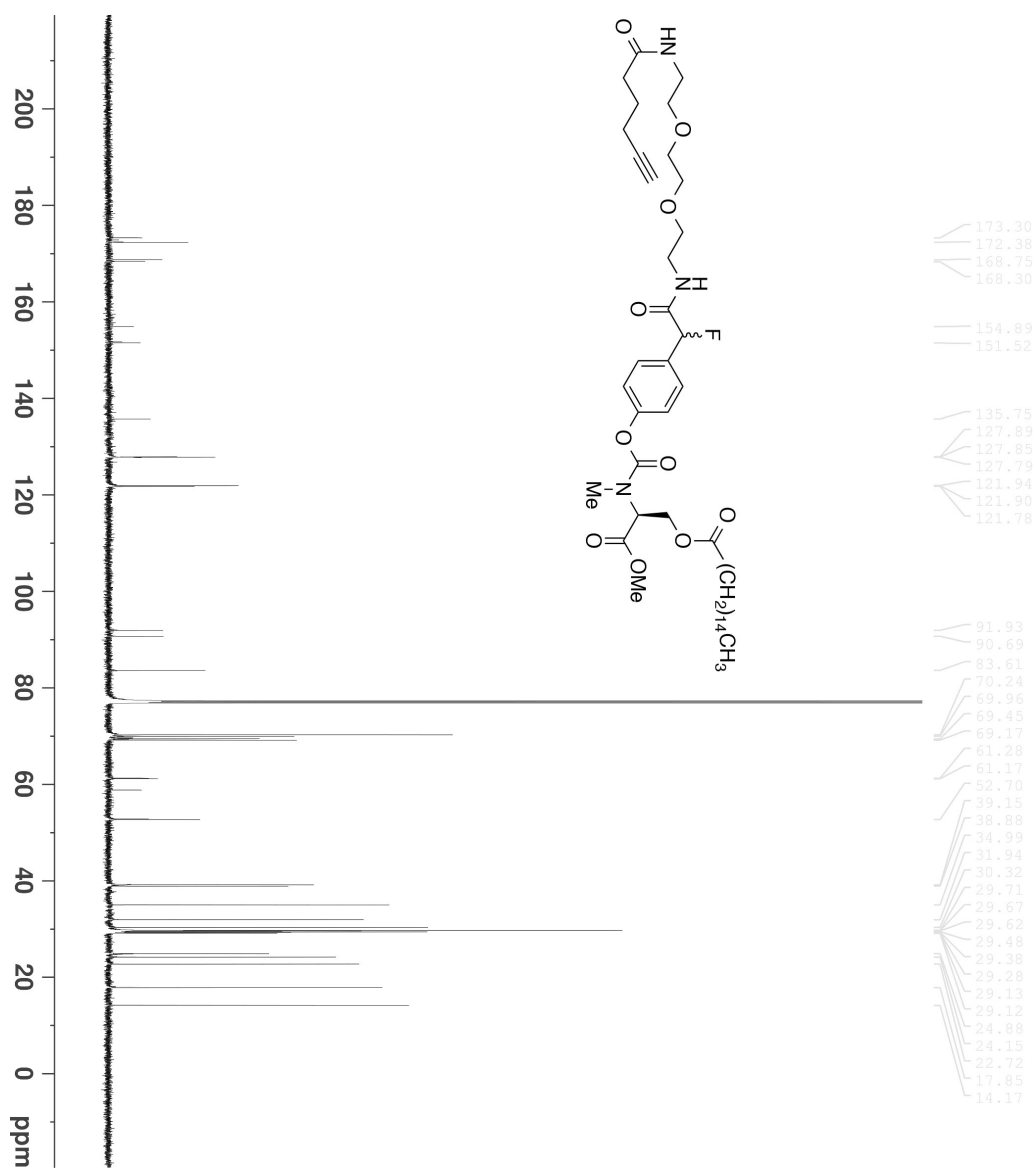
CCCCCCCCCCCCCCCC(=O)O[C@H](C)C(=O)N(C)C(=O)Oc1ccc(cc1)C(=O)C(O)C(=O)NCCOCC#CCOCCNC=O

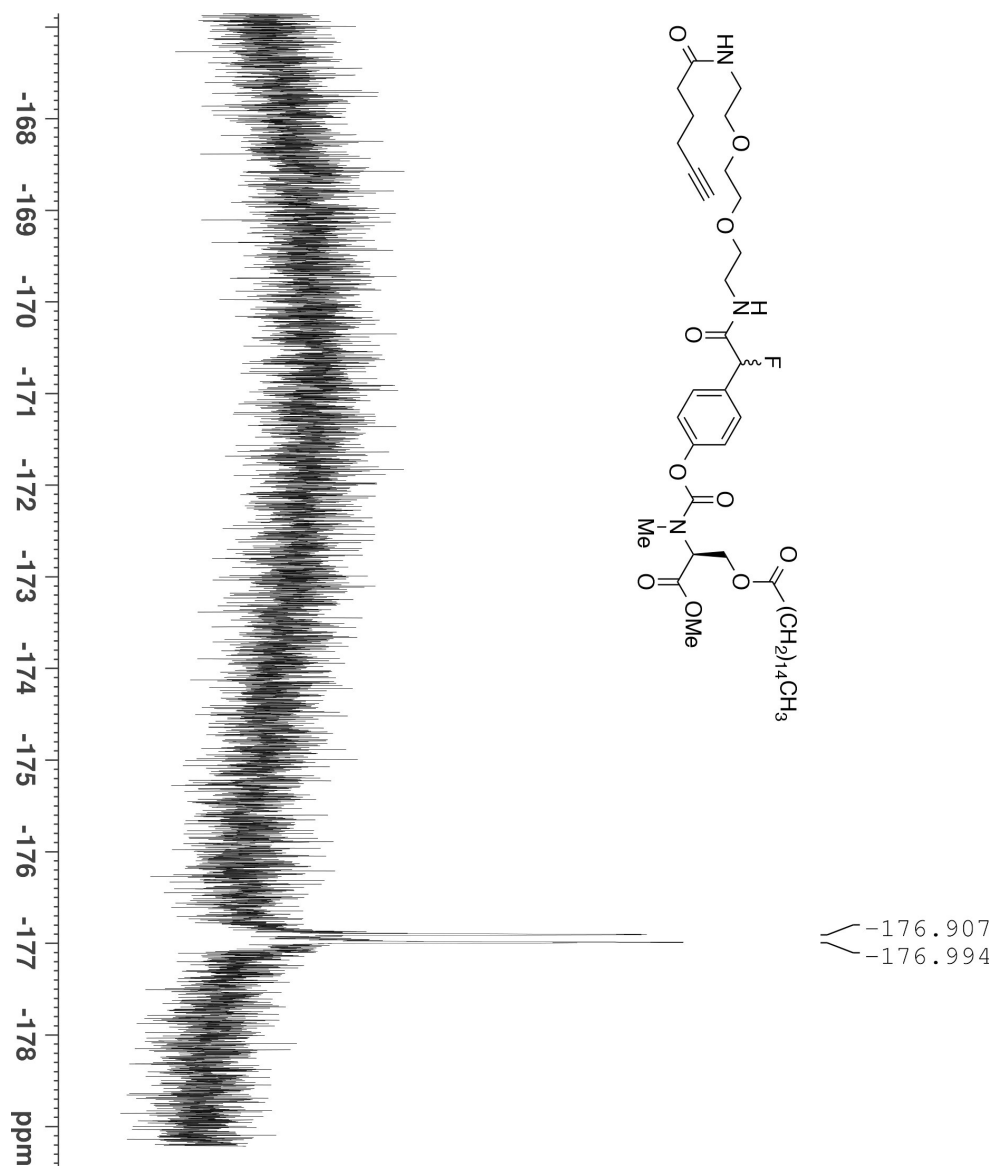
6.7. Spectra of (2*S*)-2-(((4-(1-fluoro-2,13-dioxo-6,9-dioxo-3,12-diazaoctadec-17-yn-1-yl)phenoxy)carbonyl)(methyl)amino)-3-methoxy-3-oxopropyl palmitate (**7**)

$^1\text{H-NMR}$

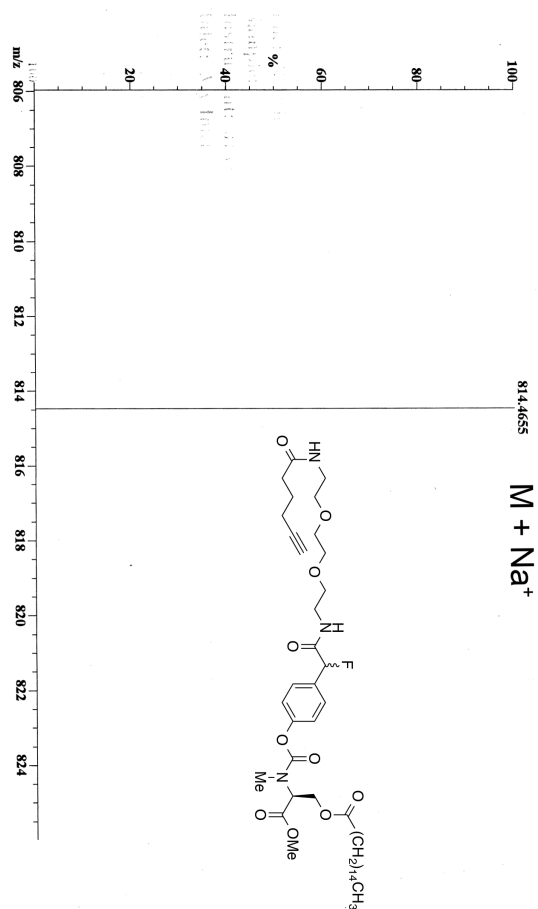


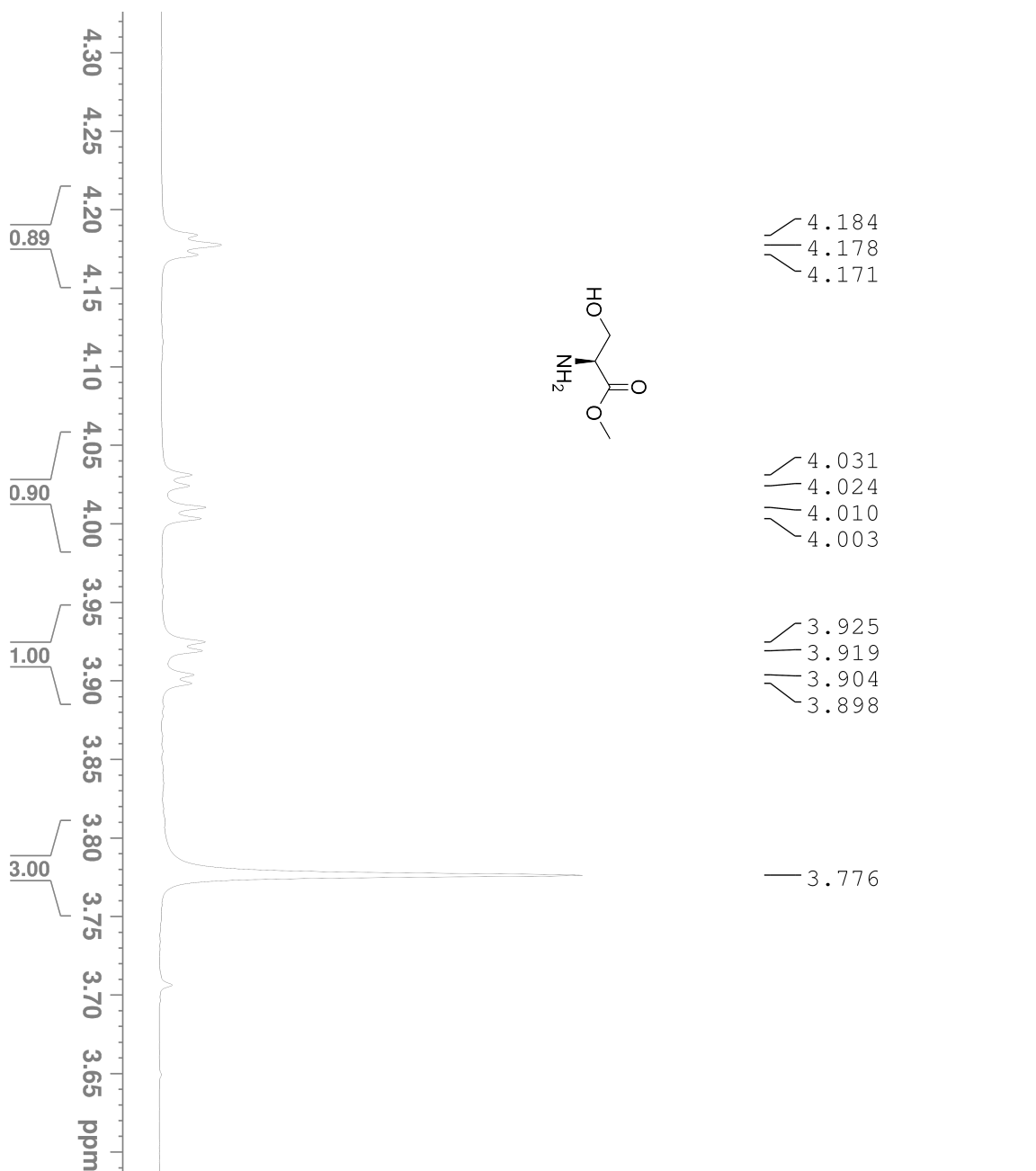


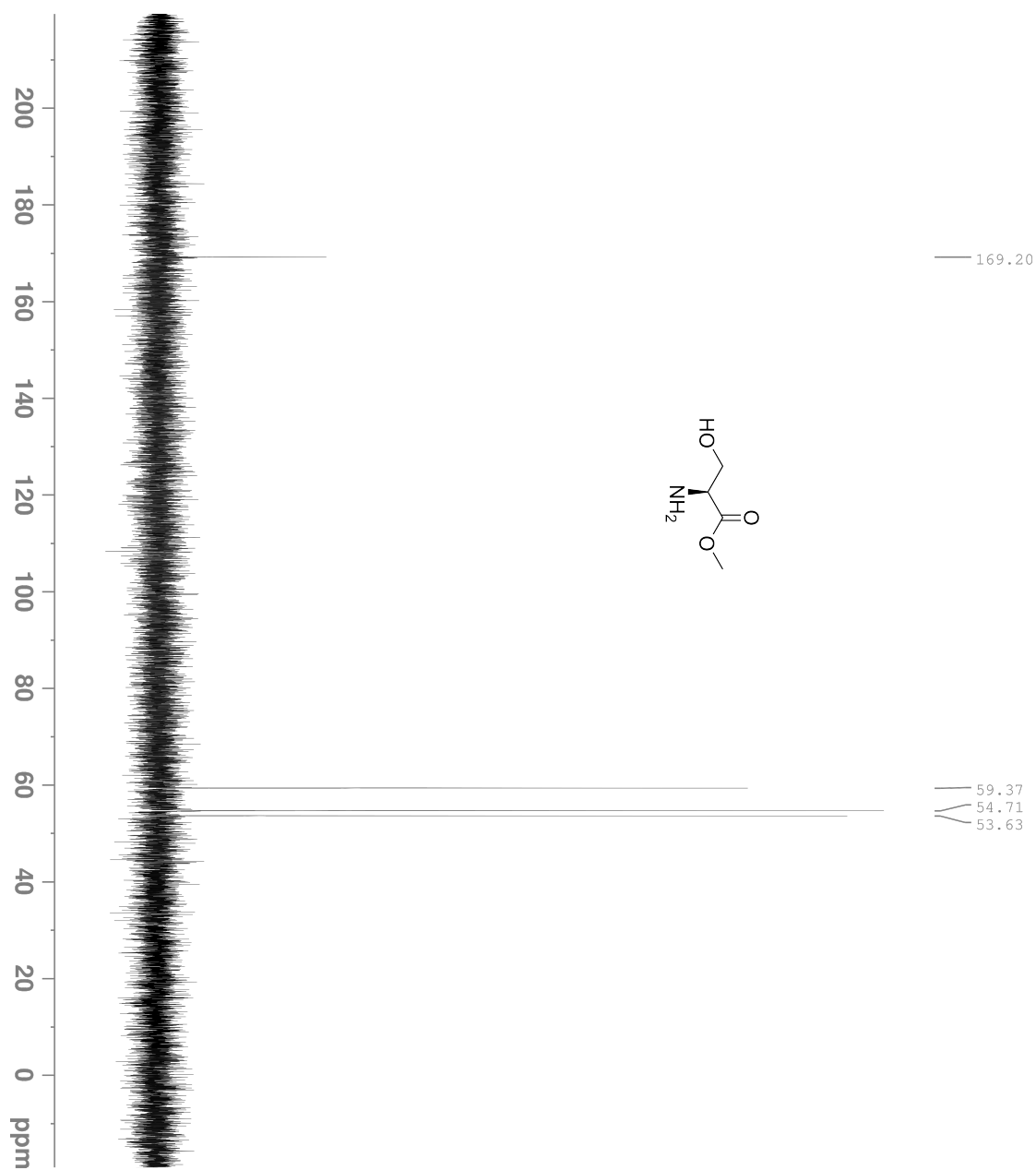
$^{13}\text{C}$ -NMR

$^{19}\text{F}$ -NMR

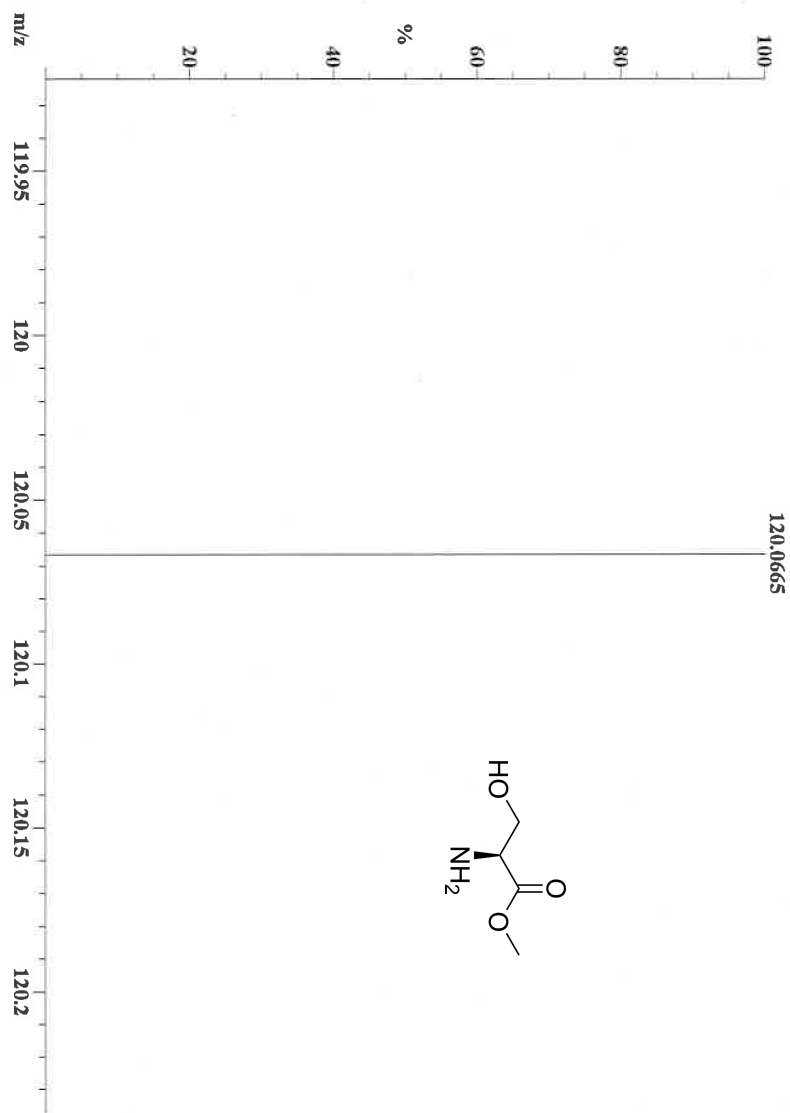
## HRMS

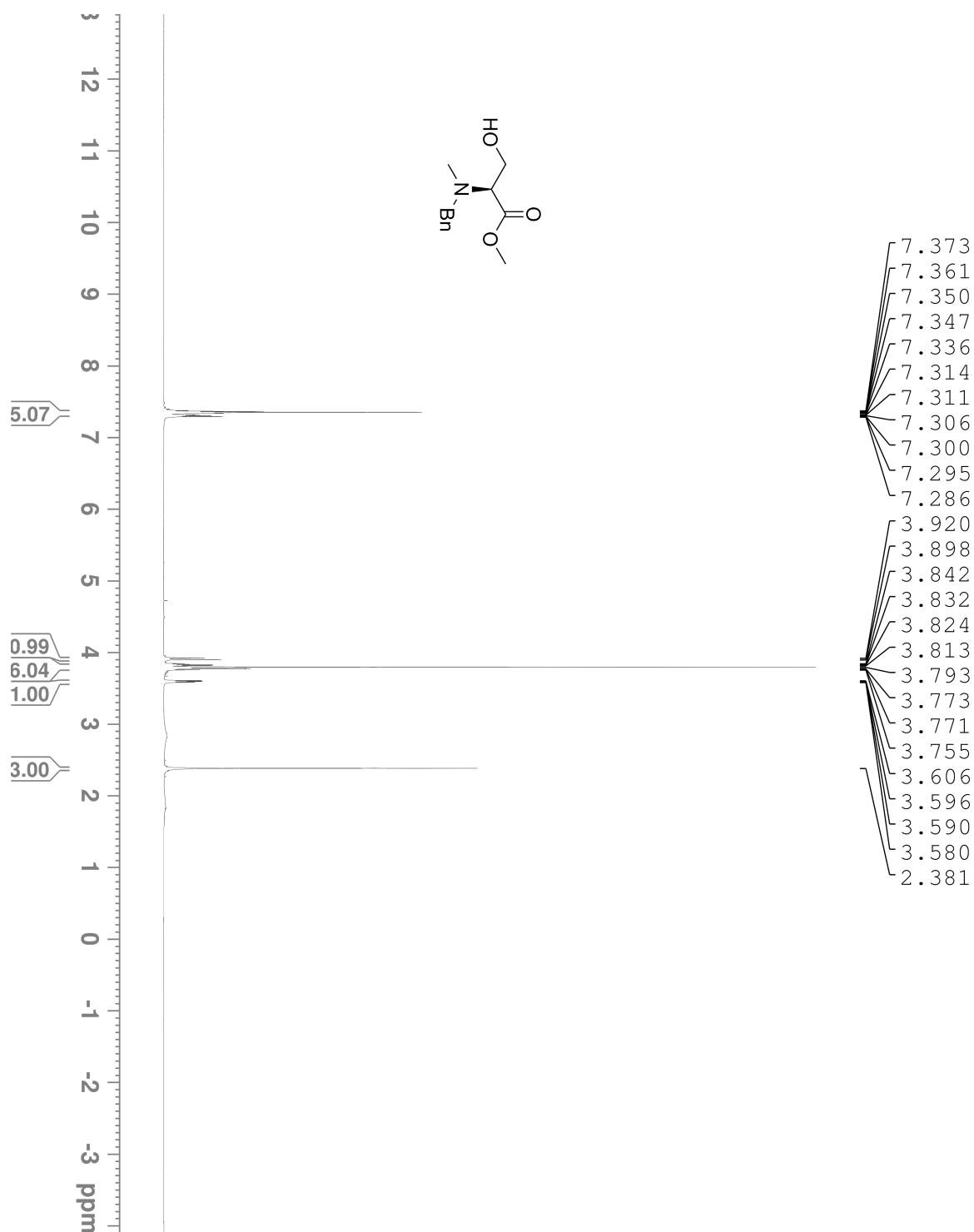


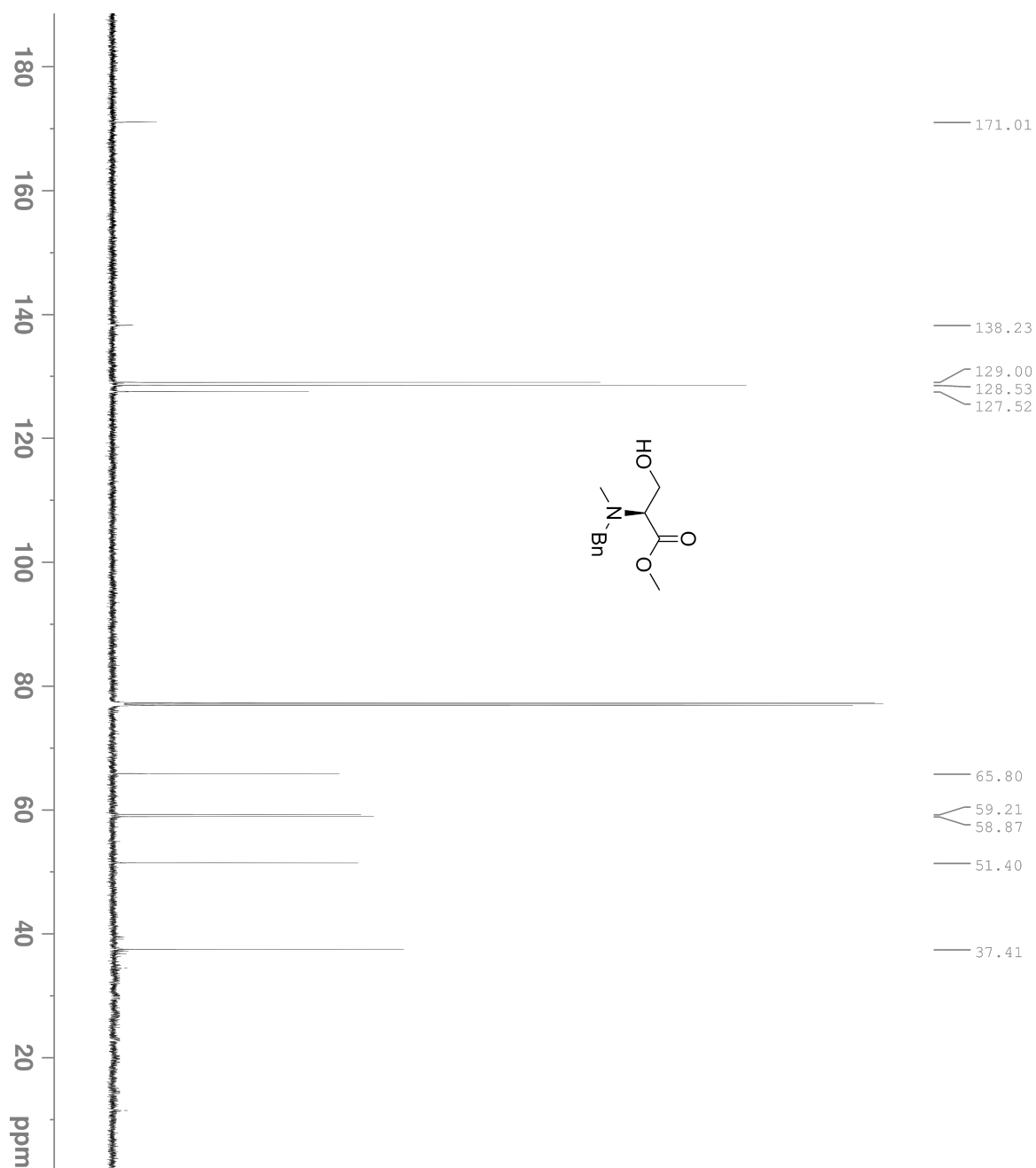
6.8. Spectra of methyl *L*-serinate (**8**) $^1\text{H}$ -NMR

$^{13}\text{C}$ -NMR

## HRMS

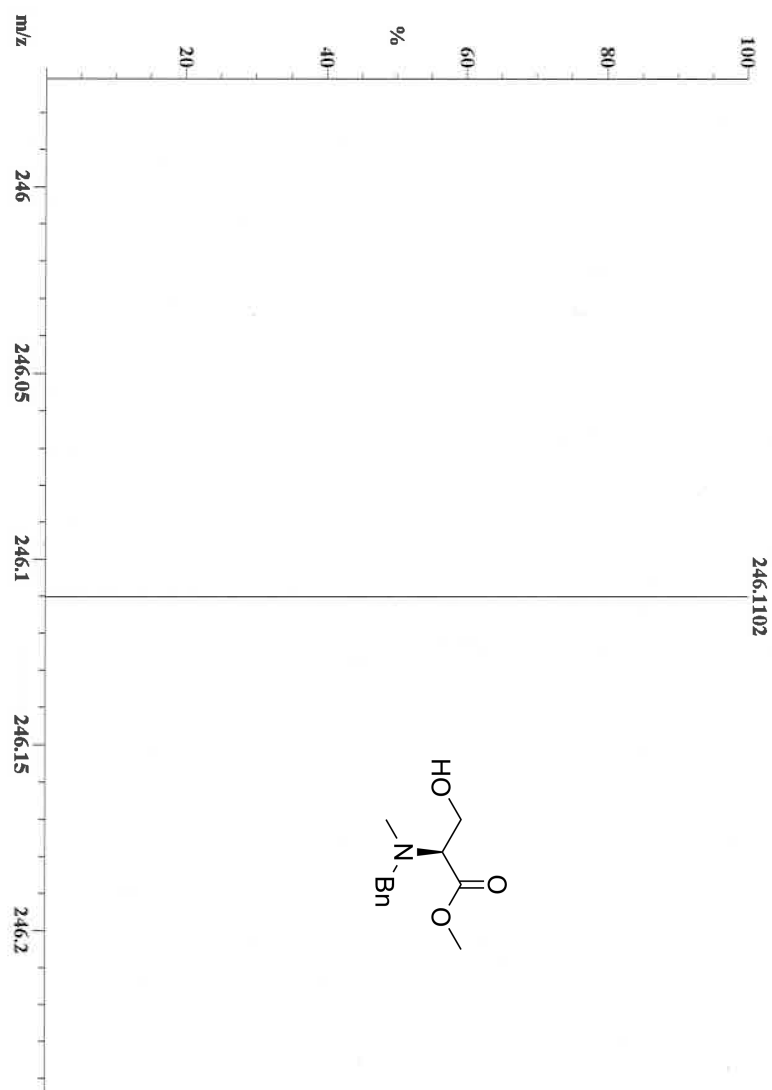


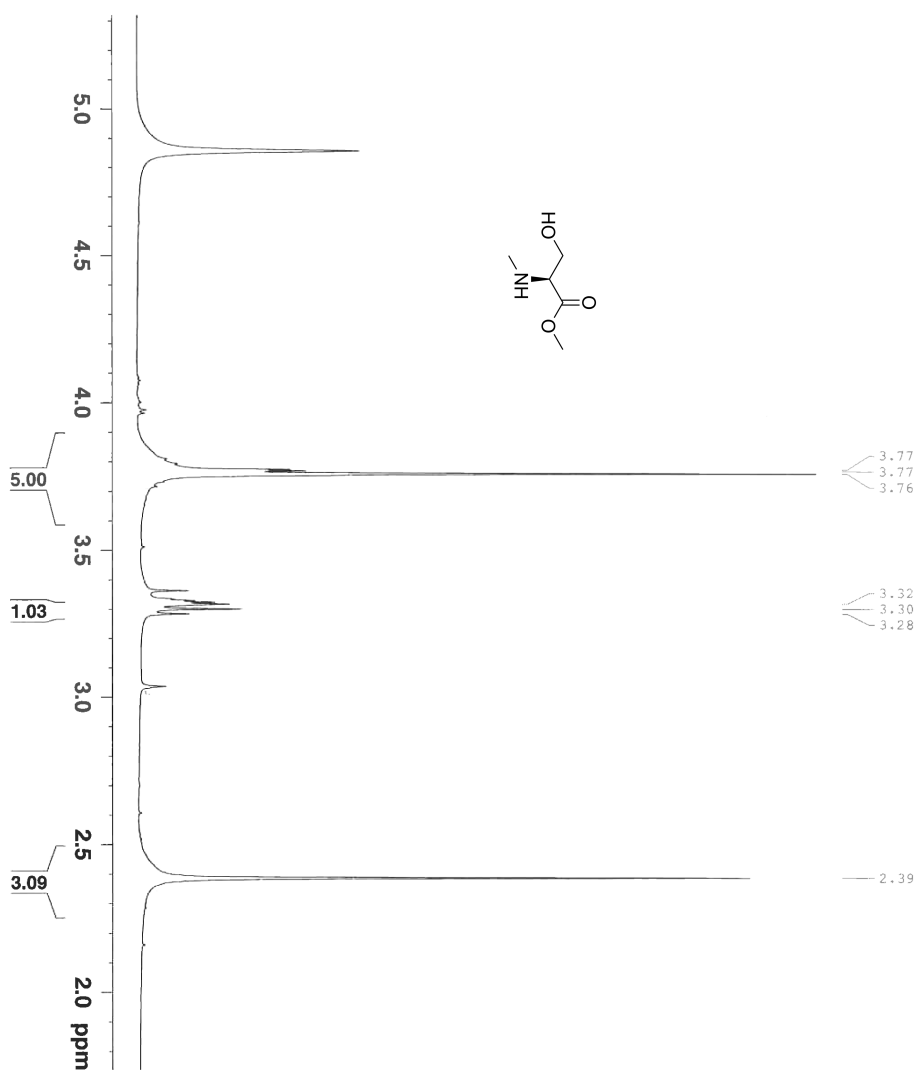
6.9. Spectra of methyl *N*-benzyl-*N*-methyl-*L*-serinate (**9**)<sup>1</sup>H-NMR

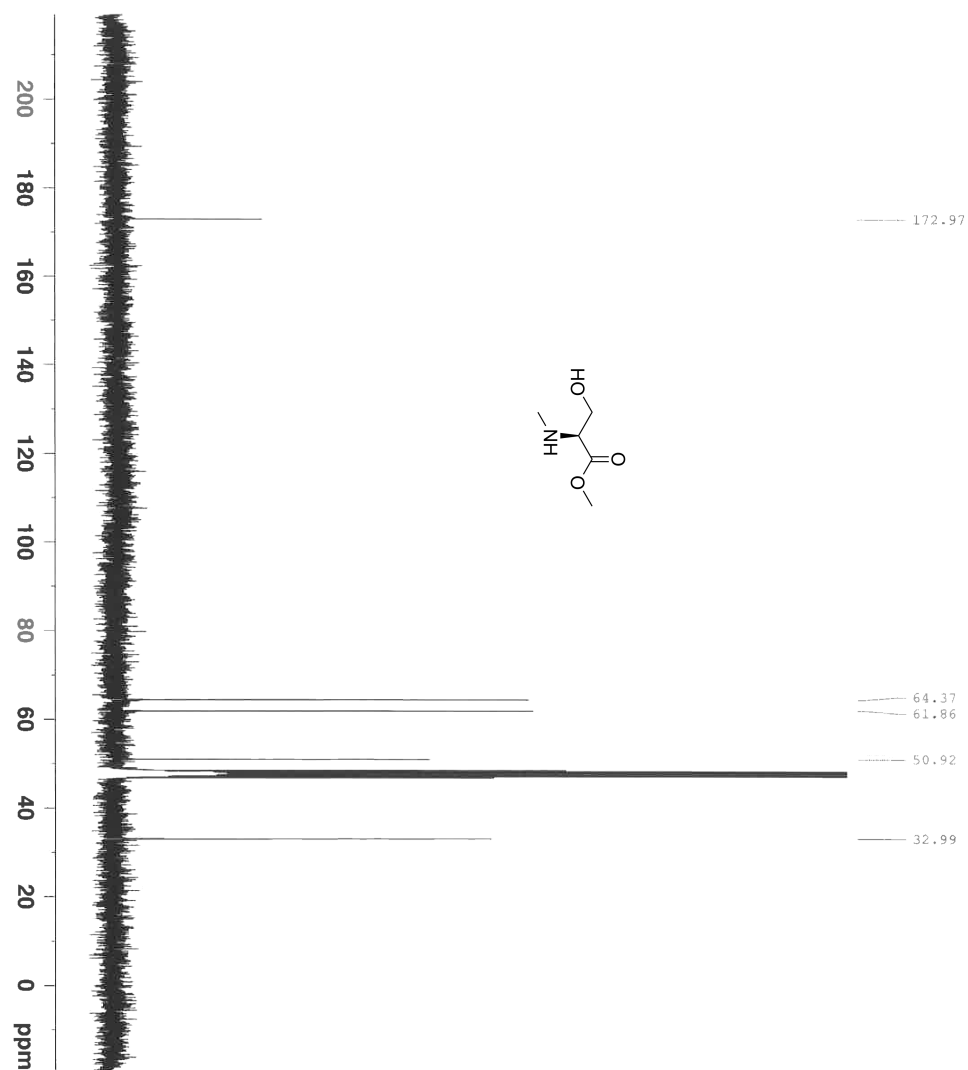
$^{13}\text{C}$ -NMR



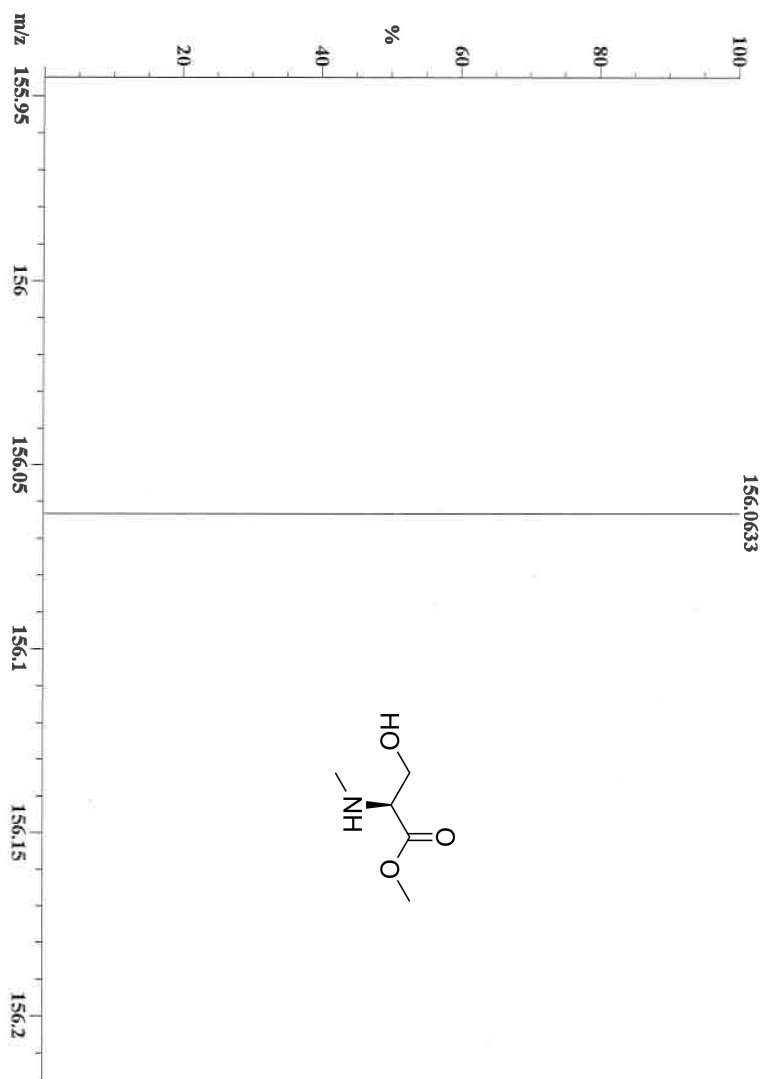
## HRMS

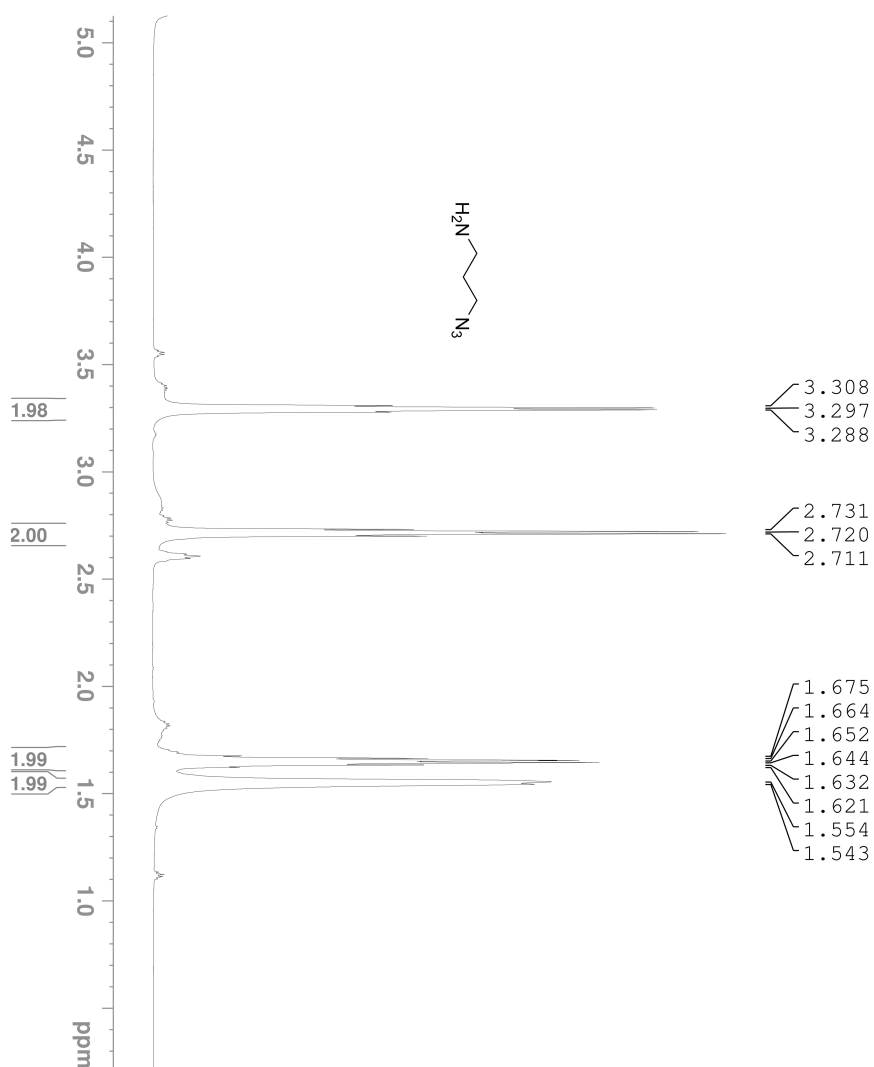


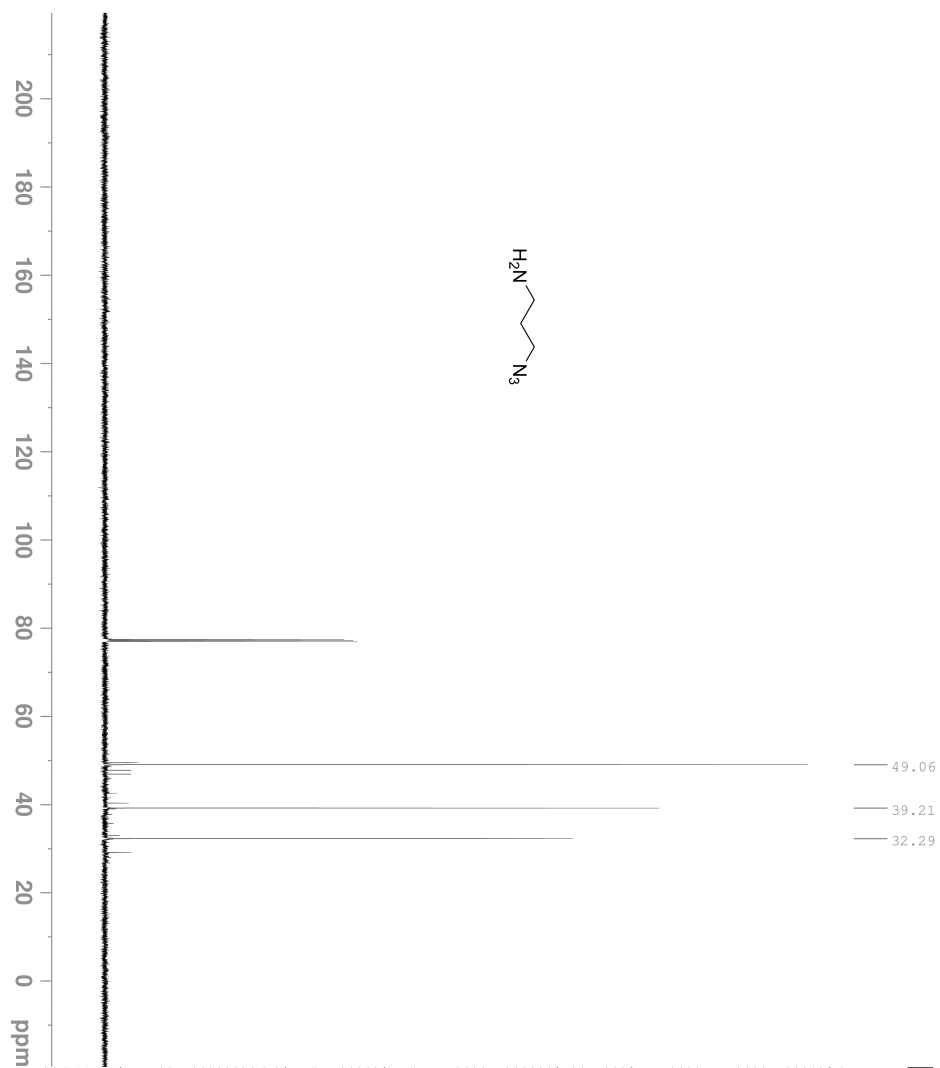
6.10. Spectra of methyl *N*-methyl-*L*-serinate (**10**) $^1\text{H}$ -NMR

$^{13}\text{C}$ -NMR

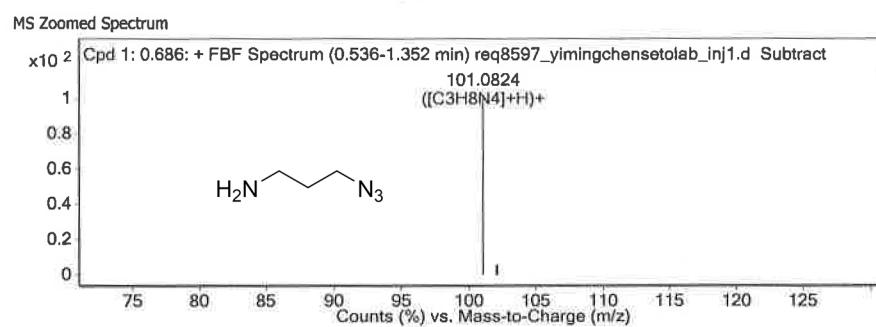
## HRMS



6.11. Spectra of 3-azidopropylamine (**11**) $^1\text{H}$ -NMR

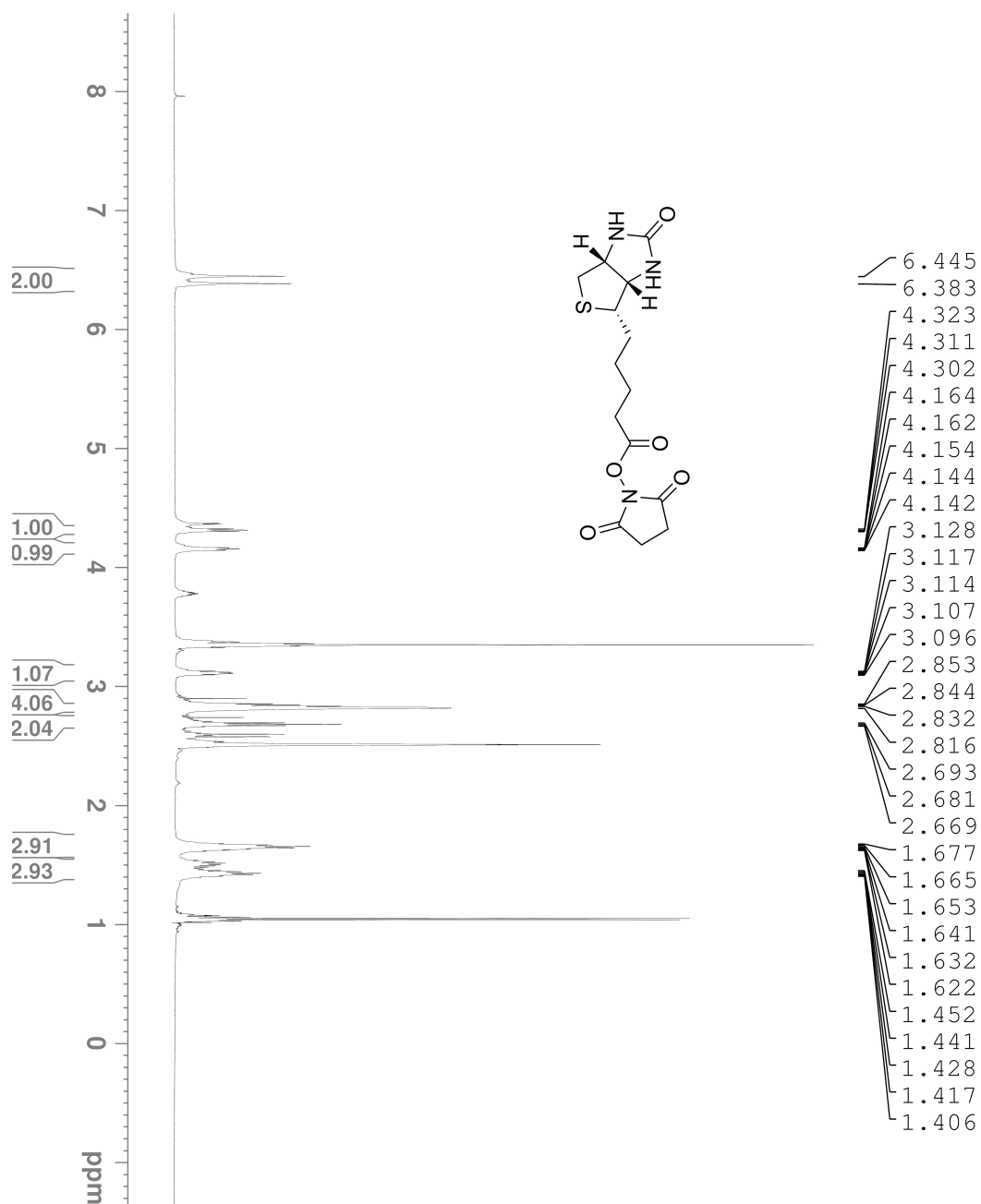
$^{13}\text{C}$ -NMR

## HRMS

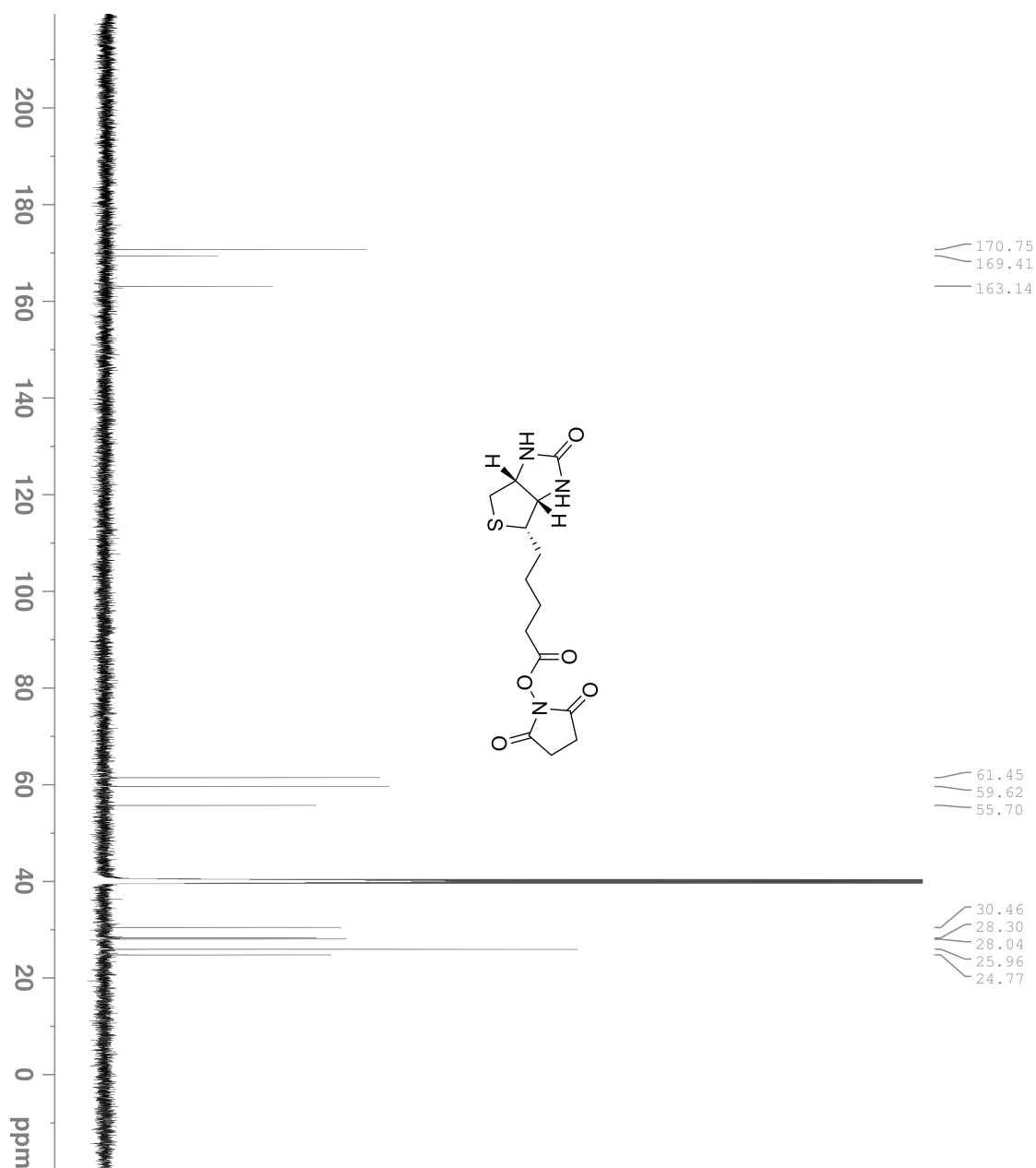


6.12. Spectra of 2,5-dioxopyrrolidin-1-yl 5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno  
[3,4-*d*]imidazol-4-yl)pentanoate (**12**)

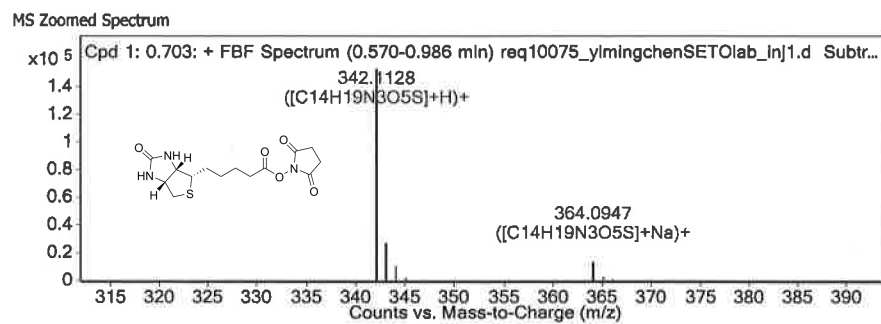
<sup>1</sup>H-NMR





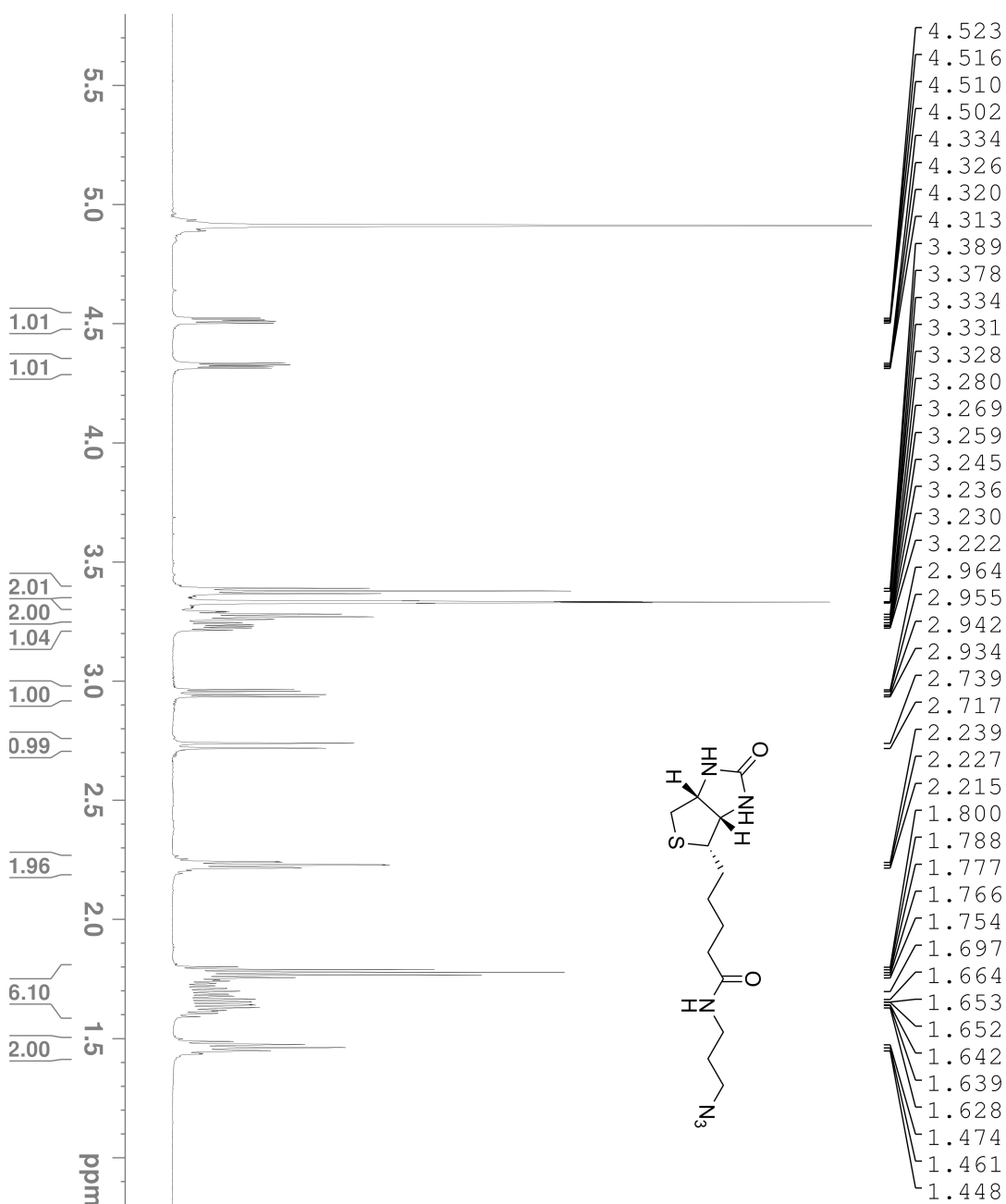
$^{13}\text{C}$ -NMR

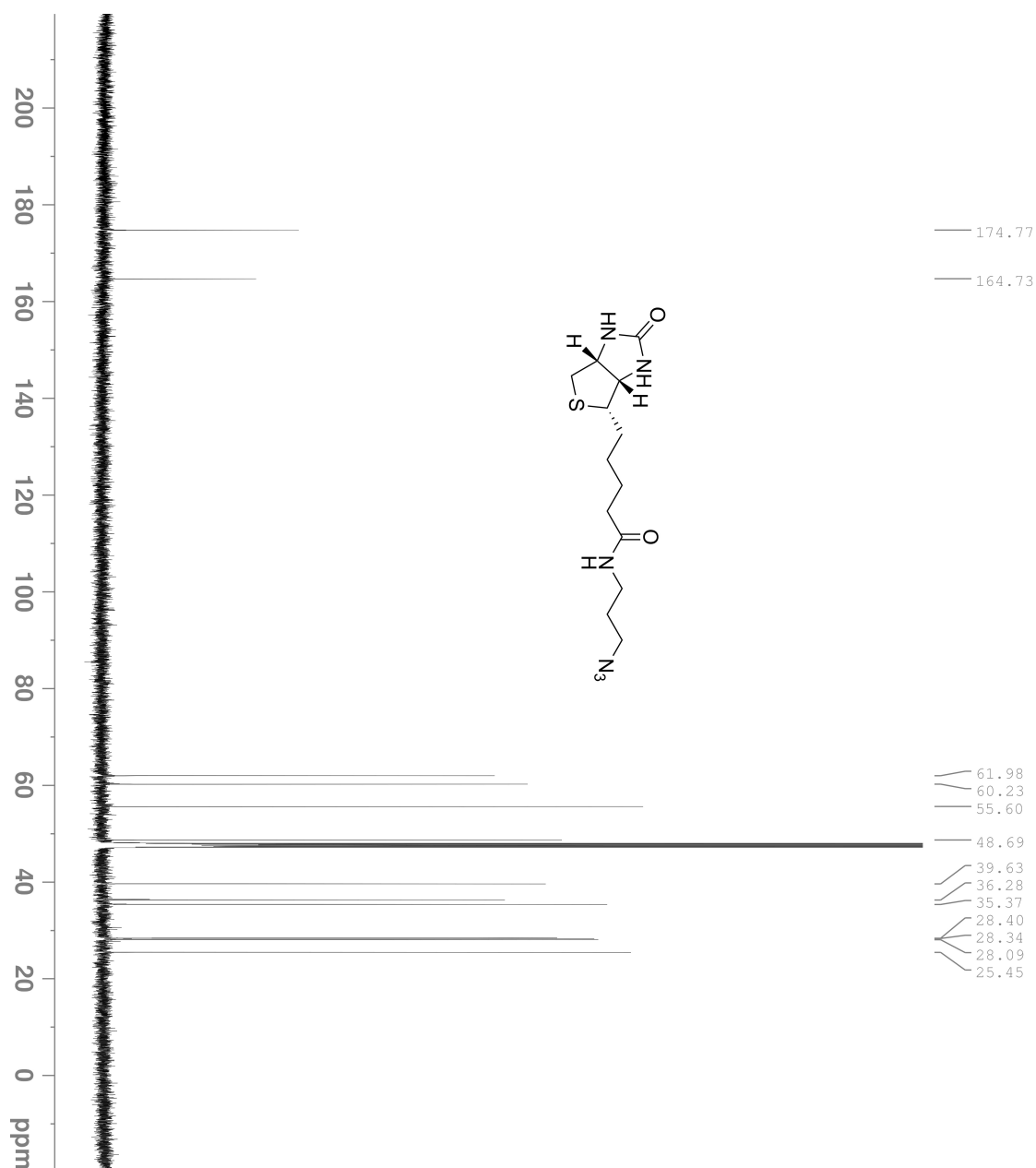
## HRMS



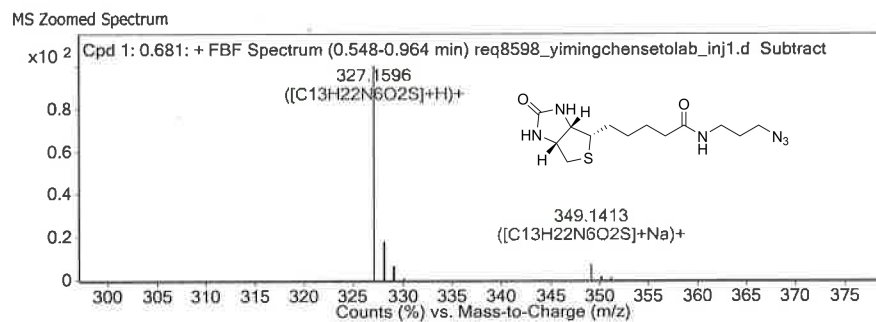
6.13. Spectra of *N*-(3-azidopropyl)-5-((3*aS*,4*S*,6*aR*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamide (**13**)

<sup>1</sup>H-NMR



$^{13}\text{C}$ -NMR

## HRMS



## 7. References

1. Xue, F. T.; Seto, C. T., Fluorogenic peptide substrates for serine and threonine phosphatases. *Org. Lett.* **2010**, *12* (9), 1936-1939.
2. (a) Kumar, S.; Zhou, B.; Liang, F. B.; Wang, W. Q.; Huang, Z. H.; Zhang, Z. Y., Activity-based probes for protein tyrosine phosphatases. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101* (21), 7943-7948; (b) Lo, L. C.; Chiang, Y. L.; Kuo, C. H.; Liao, H. K.; Chen, Y. J.; Lin, J. J., Study of the preferred modification sites of the quinone methide intermediate resulting from the latent trapping device of the activity probes for hydrolases. *Biochem. Bioph. Res. Co.* **2005**, *326* (1), 30-35.
3. (a) Rusch, M. Design, synthesis and application of small molecule acyl protein thioesterase inhibitors. Ph.D. Dissertaion, Technische Universität Dortmund, Dortmund, 2011; (b) Davda, D.; Martin, B. R., Acyl protein thioesterase inhibitors as probes of dynamic S-palmitoylation. *Medchemcomm* **2014**, *5* (3), 268-276.
4. Devedjiev, Y.; Dauter, Z.; Kuznetsov, S. R.; Jones, T. L. Z.; Derewenda, Z. S., Crystal structure of the human acyl protein thioesterase I from a single X-ray data set to 1.5 angstrom. *Structure* **2000**, *8* (11), 1137-1146.
5. (a) Trott, O.; Olson, A. J., Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455-461; (b) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P., Amber, a Package of Computer-Programs for Applying Molecular Mechanics, Normal-Mode Analysis, Molecular-Dynamics and Free-Energy Calculations to Simulate

the Structural and Energetic Properties of Molecules. *Comput. Phys. Commun.* **1995**, *91* (1-3), 1-41.

6. Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R., Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1-3), 43-56.

7. (a) Schrodinger, LLC, The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint, Version 1.0. 2010; (b) Schrodinger, LLC, The PyMOL Molecular Graphics System, Version 1.3r1. 2010.

8. Sousa da Silva, A.; Vranken, W., ACPYPE - AnteChamber PYthon Parser interfAcE. *BMC Research Notes* **2012**, *5* (1), 367.

9. White, K. N.; Konopelski, J. P., Facile synthesis of highly functionalized *N*-methyl amino acid esters without side-chain protection. *Org. Lett.* **2005**, *7* (19), 4111-4112.

10. Landi, F.; Johansson, C. M.; Campopiano, D. J.; Hulme, A. N., Synthesis and application of a new cleavable linker for "click"-based affinity chromatography. *Org. Biomol. Chem.* **2010**, *8* (1), 56-59.

11. Susumu, K.; Uyeda, H. T.; Medintz, I. L.; Pons, T.; Delehanty, J. B.; Mattoussi, H., Enhancing the stability and biological functionalities of quantum dots via compact multifunctional ligands. *J. Am. Chem. Soc.* **2007**, *129* (45), 13987-13996.

12. Baggelaar, M. P.; Janssen, F. J.; van Esbroeck, A. C. M.; den Dulk, H.; Allara, M.; Hoogendoorn, S.; McGuire, R.; Florea, B. I.; Meeuwenoord, N.; van den Elst, H.; van der Marel, G. A.; Brouwer, J.; Di Marzo, V.; Overkleeft, H. S.; van der Stelt, M., Development of an Activity-Based Probe and In Silico Design Reveal Highly Selective Inhibitors for Diacylglycerol Lipase- $\alpha$  in Brain. *Angew. Chem. Int. Ed.* **2013**, *52* (46), 12081-12085.





# **CHAPTER THREE**

## **Validation of the ABP via Kinetic Assays and In-Gel Fluorescence**



## 1. Abstract

Having successfully synthesized the activity-based probe (ABP) for APTs, we next wanted to demonstrate proof-of-principle by assaying the probe with the APT1 enzyme. This was necessary, prior to attempts at activity-based protein profiling (ABPP) with the ABP, in order to establish that reaction of the probe with APT1 would proceed by our proposed mechanism of action. Extensive efforts in targeting the APT1 enzyme have led to the development of a colorimetric assay measuring the release of *p*-nitrophenoxide (*p*NP) from *p*-nitrophenyl octanoate (*p*NPO), a substrate mimic of palmitic acid thioesters. In addition, 4-methylumbelliferyl octanoate (4-MUO) was tested as a synthetic fluorogenic substrate in kinetic assays of APT1 and APT2.

In this chapter, the Michaelis constant ( $K_m$ ) and maximal velocity ( $V_{max}$ ) of the *p*NPO substrate was first measured using APT1. Subsequently, the APT1 enzyme was incubated with the ABP at different concentrations to examine the inhibition of the APT1 activity by the ABP. We observed that the ABP does inhibit the activity of APT1 in a concentration-dependent manner, which indicated the ABP labels the APT1 enzyme irreversibly.

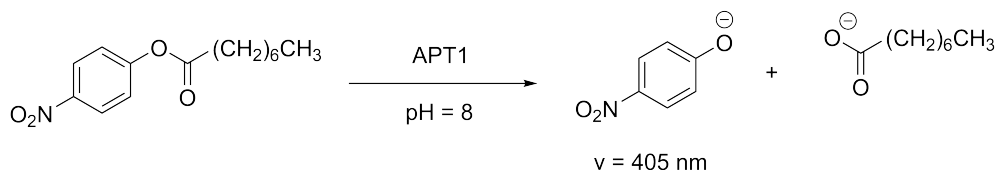
In the second part, the ABP was incubated with both recombinant APT1 and APT2 in a labeling step. These labeled APTs were then tagged by fluorescent-conjugated components through click chemistry with a functionalized azide tag. The fluorescent-conjugated labeled APTs were then resolved by SDS-PAGE and subsequently visualized by in-gel fluorescence. Results demonstrate that the ABP reacts exclusively with the active forms of APT1 and APT2. The concept of the ABP design was successfully validated.

## 2. Measurement of the $K_m$ and $V_{max}$ of APT1 using the *p*NPO Colorimetric Substrate

### 2.1. Introduction to the APT1 colorimetric assay

In enzyme kinetics, the Michaelis constant ( $K_m$ ) and maximal velocity ( $V_{max}$ ) are two important constants that describe an enzyme's activity.<sup>1</sup> In a biological system, increasing the concentration of substrate does not increase the rate of an enzyme-catalyzed reaction beyond a certain point. This point is reached when enough substrate molecules exist to completely saturate the active sites of the enzymes. The maximal velocity that can be reached, or  $V_{max}$ , is the reaction rate under these saturating conditions. Similarly, enzymes have varying tendencies to bind their substrates, so the enzyme's  $K_m$  describes the substrate concentration at which half the enzyme's active sites are occupied by substrate. Graphically, the  $K_m$  is the substrate concentration that gives the enzyme one-half of its  $V_{max}$ .

To investigate the activity of APT1, the measurement of the  $K_m$  and  $V_{max}$  of APT1 was necessary. *p*-Nitrophenyl octanoate (*p*NPO) is an ideal colorimetric substrate for this task, as it reacts with APT1 in the manner presented in Scheme 3.1.<sup>2</sup> *p*-Nitrophenyl octanoate is a fatty acid ester which is hydrolyzed by the APT1 enzyme when the pH = 8. The product of the hydrolysis, *p*-nitrophenoxide strongly absorbs light at the wavelength 405 nm. This characteristic absorbance can be utilized to monitor the concentration of *p*-nitrophenoxide (*p*NP) in the solution as a function of time, and can thus be used to calculate the reaction rate of the hydrolysis. In this assay, mouse acyl protein thioesterase 1 (APT1) activity was determined by measuring the initial hydrolysis rate of *p*-nitrophenyl octanoate with a UV micro-plate reader ( $\lambda_{abs} = 405$  nm).



**Scheme 3.1 The enzyme-catalyzed hydrolysis of *p*-nitrophenyl octanoate.** The *p*-nitrophenolate released from the reaction has characteristic strong absorbance of light at 405 nm, which can be utilized to monitor the hydrolysis reaction.

## 2.2. Experiment design

According to the Michaelis-Menten Equation (Equation 3.1), the values of  $V_{max}$  and  $K_m$  can be determined by measuring the reaction rates ( $v$ ) at different substrate concentrations ( $[S]$ ).

Thus, to determine the values of  $V_{max}$  and  $K_m$  for APT1, the enzyme was incubated with the *p*NPO at different concentrations. To measure the reaction rate  $v = \frac{d[pNP]}{dt}$ , at the initial stage of the hydrolysis reaction, it was assumed that the concentrations of the *p*NPO remained unchanged because the *p*NPO concentration is much higher than the APT1 enzyme ( $[pNPO] \gg [E_0]$ ), and we monitor the reaction to less than 10% consumption of the substrate. For this reason, the hydrolysis reaction can be assumed as a pseudo-zeroth order reaction. The linear changes of the *p*NP concentrations were monitored by UV absorbance at 405 nm, the reaction rates were extracted by fitting the *p*NP concentration data to the zeroth-order reaction rate equation.

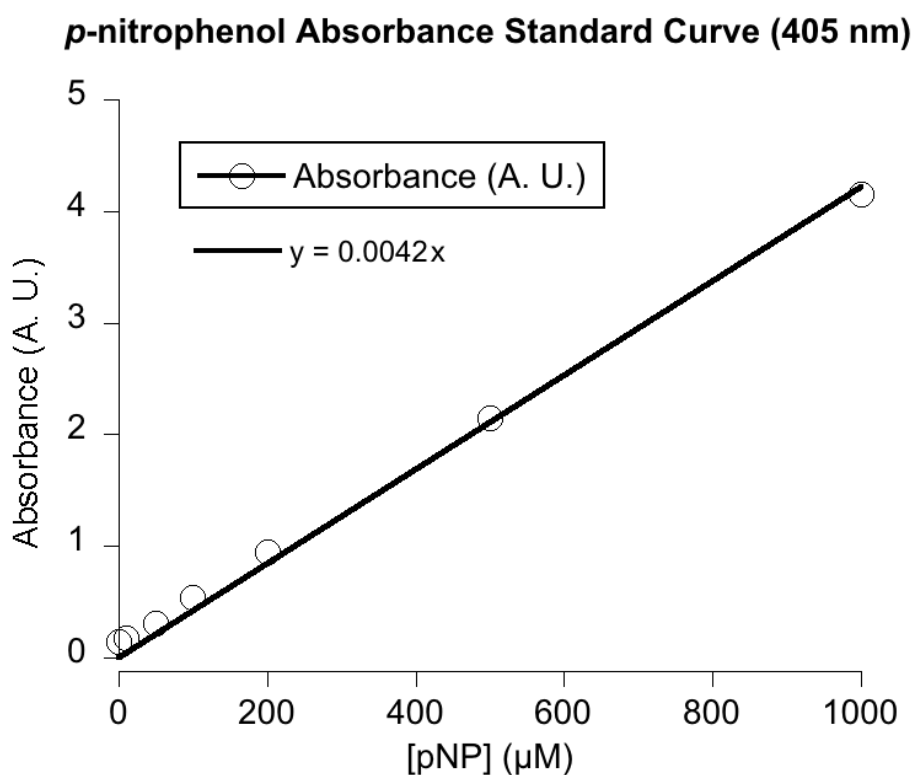
$$v = \frac{d[P]}{dt} = \frac{V_{max} [S]}{K_m + [S]}$$

**Equation 3.1 Michaelis-Menten kinetics.** In this equation,  $v$  is the reaction rate,  $[P]$  is the concentration of product,  $[S]$  is the concentration of the substrate. The values of  $V_{max}$  and  $K_m$  can be determined by measuring the reaction rates ( $v$ ) at various substrate concentrations ( $[S]$ ).

After all the reaction rate data were collected, the data were fitted to the Michalis-Menten Equation to determine the values of the  $V_{max}$  and  $K_m$ .

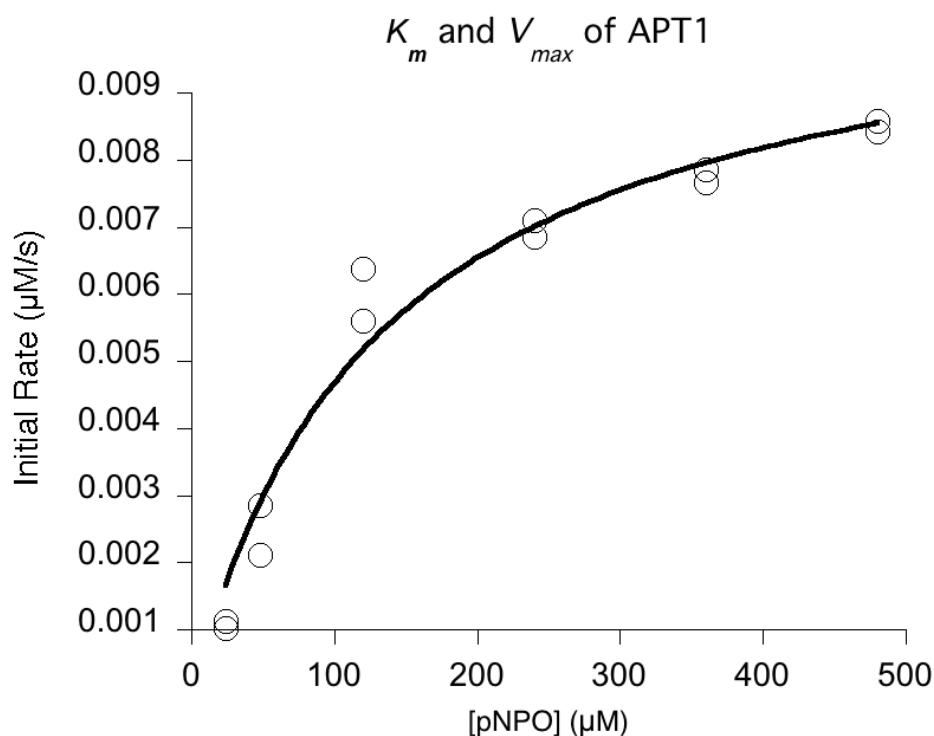
### 2.3. Results and discussion

To convert the unit of the absorbance to the actual concentration, a standard absorbance curve was constructed based on standard *p*NP solutions (Figure 3.1). The x-axis is the concentration of the standard *p*NP solution in  $\mu\text{M}$ , and the y-axis is the measured absorbance. The standard curve has the equation  $y = 0.0042x$ .



**Figure 3.1** The standard curve for *p*NP. The x-axis is the concentration of *p*NP standard solution in  $\mu\text{M}$ , the y-axis is the absorbance in atomic unit (a. u.). The linear equation is  $y = 0.0042x$ .

To measure  $V_{max}$  and  $K_m$ ,  $p$ NPO solutions were freshly prepared by vigorous stirring of  $p$ -nitrophenyl octanoate in buffer (20 mM HEPES, 150 mM NaCl, 10% glycerol, 0.08% Triton X-100, pH = 7.4). Subsequently, 30  $\mu$ L of an APT1 (250 nM) solution was added to 20  $\mu$ L of  $p$ NPO solution with varying concentration (24, 48, 120, 240, 360, 480  $\mu$ M) at room temperature (Table 3.1). The release of the  $p$ NP was monitored for 30 min at 30 sec intervals. The initial rate was calculated based on the absorbance data between 2 min and 8 min. All incubations were assembled and assayed in duplicate. By fitting the data to the Michaelis-Menten equation in Kaleidagraph (Synergy Software, Version 4.03), the Michaelis constant ( $K_m$ ) and maximum velocity ( $V_{max}$ ) were determined (Figure 3.2). The results are presented in Table 3.2.



**Figure 3.2** Data for  $p$ -nitrophenyl octanoate hydrolysis by 150 nM APT1. The solid line represents a fit of the data to the Michaelis-Menten equation.

|                         |                       |                       |                       |                       |                       |                       |
|-------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <b><i>p</i>NPO (μL)</b> | 1                     | 2                     | 5                     | 10                    | 15                    | 20                    |
| <b>Buffer (μL)</b>      | 19                    | 18                    | 15                    | 10                    | 5                     | 0                     |
| <b>APT1 (μL)</b>        | 30                    | 30                    | 30                    | 30                    | 30                    | 30                    |
| <b>[<i>p</i>NPO]</b>    | 24                    | 48                    | 120                   | 240                   | 360                   | 480                   |
| <b>Rate (au/s)</b>      | $4.32 \times 10^{-6}$ | $8.90 \times 10^{-6}$ | $2.69 \times 10^{-5}$ | $2.99 \times 10^{-5}$ | $3.23 \times 10^{-5}$ | $3.55 \times 10^{-5}$ |
|                         | $4.73 \times 10^{-6}$ | $1.20 \times 10^{-5}$ | $2.36 \times 10^{-5}$ | $2.89 \times 10^{-5}$ | $3.31 \times 10^{-5}$ | $3.62 \times 10^{-5}$ |

**Table 3.1 Conditions and results of the APT1  $K_m$  and  $V_{max}$  measurement.** All the experiments were done in duplicate.

| <b><math>K_m</math> (μM)</b> | <b><math>V_{max}</math> (nM/s)</b> |
|------------------------------|------------------------------------|
| 134 ± 25                     | 10.9 ± 0.7                         |

**Table 3.2 The  $K_m$  and  $V_{max}$  of APT1 using *p*NPO as the substrate.**

### 3. A Synthetic Fluorogenic Substrate for APT Activity Assays.

#### 3.1. Introduction

*p*NPO is an ideal colorimetric substrate for assaying APT1 activity *in vitro*. However, due to the relatively low hydrolyzing activity, the *p*NPO colorimetric assay is not sensitive enough for APT2 activity profiling. Thus, a sensitive fluorogenic substrate is highly desirable for assaying APT2 activity.

4-Methylumbelliferone (4-MU) is a fluorophore that has been widely used for enzymatic assays. Esters of 4-MU do not fluoresce unless they are cleaved to release the fluorophore. Based on this unique property, a new fluorogenic assay can be designed to measure the activities of both APT1 and APT2.



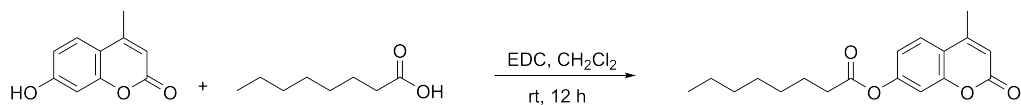
### 3.2. Synthesis of the 4-methylumbelliferyl octanoate

The 4-methylumbelliferyl octanoate (4-MUO) was synthesized according to the reaction described in Scheme 3.2. To a solution of 4-methylumbelliferone (6.0 g, 34.0 mmol) in  $\text{CH}_2\text{Cl}_2$  (50 mL) was added octanoic acid (5 g, 34.6 mmol) and  $\text{EDC}\cdot\text{HCl}$  (6.5 g, 34.0 mmol). The reaction was stirred at room temperature for 12 hr and quenched by adding 100 mL of 1M aqueous NaOH solution. The organic layer was washed by  $2 \times 100$  mL 4M HCl solution and 100 mL of brine before being concentrated *in vacuo*. The crude product was purified by flash chromatography column (0-50% EtOAc in hexanes). A white powder was obtained (9.5 g, 93%).

$^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.62 (d,  $J = 8.0$  Hz, 1H), 7.12 (d,  $J = 2.0$  Hz, 1H), 7.08-7.09 (m, 1H), 6.28 (d,  $J = 1.0$  Hz, 1H), 2.60 (t,  $J = 7.0$  Hz, 2H), 2.45 (d,  $J = 1.0$  Hz, 3H), 1.76-1.80 (m, H), 1.31-1.45 (m, 8H), 0.92 (t,  $J = 7.0$  Hz, 3H).

$^{13}\text{C}$  NMR (150 MHz,  $\text{CDCl}_3$ )  $\delta$  171.7, 160.5, 154.2, 153.2, 151.9, 125.3, 118.1, 117.8, 114.5, 110.5, 34.4, 31.6, 29.0, 28.9, 24.8, 22.6, 18.7, 14.1.

HRMS-ESI: calculated for  $\text{C}_{18}\text{H}_{22}\text{O}_4$  ( $[\text{M} + \text{H}]^+$ ) 303.1596, observed 303.1593.

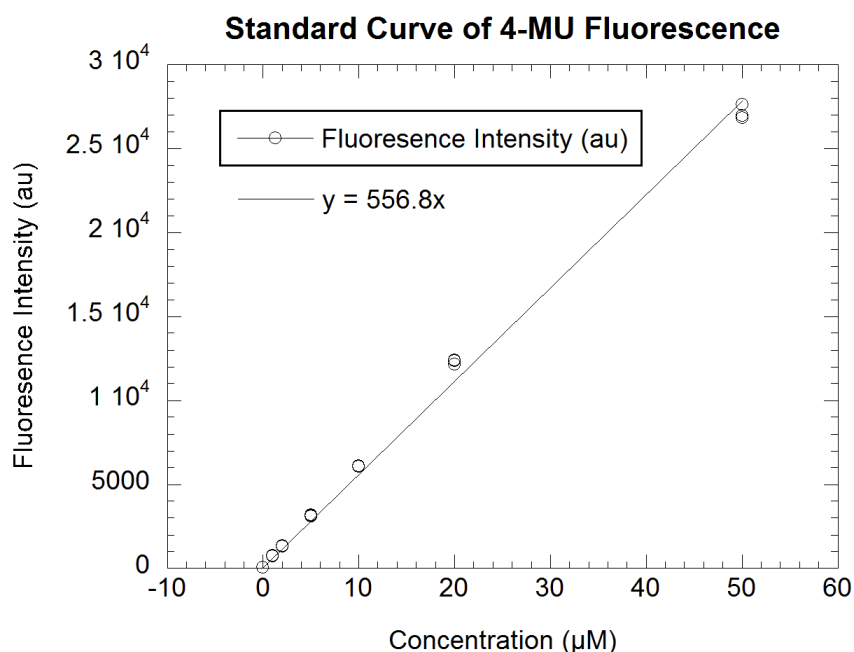


**Scheme 3.2 Synthesis of the 4-methylumbelliferyl octanoate (4-MUO).**

### 3.3. Fluorogenic kinetic assays for APT1 and APT2

Before the 4-MUO was applied in the APT activity assays, a standard curve of 4-MU was prepared in order to convert the fluorescence intensity from absorbance units to

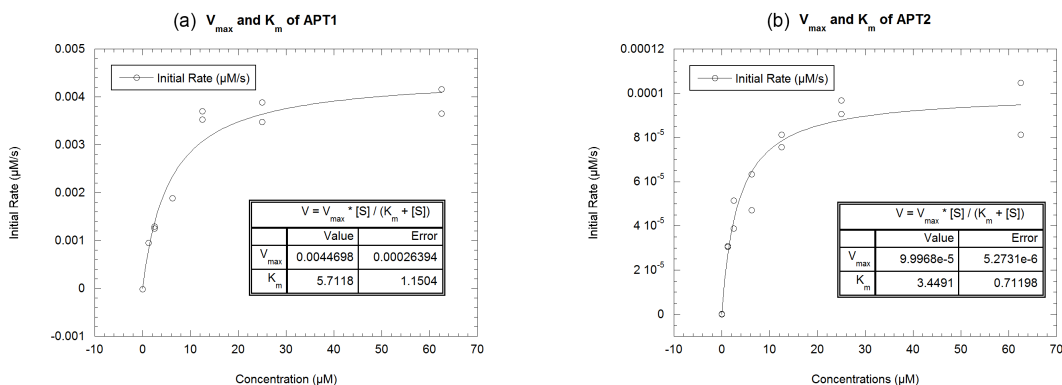
concentration (Figure 3.3). The y-axis is the fluorescence intensity in absorbance units, and the x-axis is the concentration of the 4-MUO standard in  $\mu\text{M}$ . The slope of the line is 556.8 au/ $\mu\text{M}$ .



**Figure 3.3 Standard curve of 4-MU fluorescence intensity.** The y-axis is the fluorescence intensity in absorbance units; the x-axis is the concentration of the 4-MUO standard in  $\mu\text{M}$ .

In this fluorogenic assay, the non-fluorescent substrate 4-MUO was hydrolyzed by APT1 or APT2 to release the 4-MU fluorophore. The fluorescence emitted by the 4-MU was monitored in a FLUOStar Galaxy microplate reader ( $\lambda_{\text{ex}} = 380 \text{ nm}$ ,  $\lambda_{\text{em}} = 454 \text{ nm}$ ). To measure the  $V_{\text{max}}$  and  $K_m$  of both APT1 and APT2, 4-MUO of various concentrations (0, 1.25, 2.50, 6.25, 12.5, 25, 62.5  $\mu\text{M}$ ) was incubated with APT1 or APT2 (100 nM) in Tris buffer (50 mM, pH = 8.0, 5% DMSO) at room temperature. The increasing fluorescence

intensity was fitted linearly to the zeroth-order rate equation to obtain the hydrolysis rate, which was plotted against the concentrations of 4-MUO substrate (Figure 3.4a and 3.4b). All the data points in the rate-concentration plots were fitted to the Michaelis-Menten equation (Equation 3.1) in Kaleidagraph (Synergy Software, Version 4.03) to obtain the values of  $V_{max}$  and  $K_m$  of APT1 and APT2. The results are presented in Table 3.3.



**Figure 3.4** The rate/concentration plots for APT1 and APT2. The x-axis is the concentration of the 4-MUO substrate in  $\mu\text{M}$ . The y-axis is the initial rate of the enzymatic hydrolysis reactions in  $\mu\text{M/s}$ . Data points are non-linearly fitted to the Michaelis-Menten equations to calculate the  $V_{max}$  and  $K_m$  of APT1 and APT2.

| Enzyme | $K_m$ ( $\mu\text{M}$ ) | $V_{max}$ (nM/s) |
|--------|-------------------------|------------------|
| APT1   | $5.7 \pm 1.1$           | $4.5 \pm 0.3$    |
| APT2   | $3.4 \pm 0.7$           | $0.1 \pm 0.005$  |

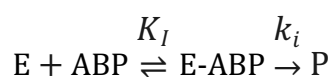
**Table 3.3** The  $V_{max}$  and  $K_m$  of APT1 and APT2 using 4-MUO as the substrate. The values of  $V_{max}$  and  $K_m$  were calculated from the initial rate and concentration data.

From the values of  $K_m$  listed in Table 3.3, we can conclude that compare to *p*NPO, 4-MUO is much more sensitive and possesses lower  $K_m$  values, which define the affinity between substrate and enzyme. In addition, the  $V_{max}$  values that we obtained demonstrate that the hydrolysis activity of APT2 is much lower than that of APT1, which suggests the different roles they play in the depalmitoylation of proteins such as Ras.

#### 4. APT1 Inhibition Kinetic Assay

##### 4.1. Introduction to the inhibition assay

The reaction between the ABP and the APT1 enzyme can be described by Equation 3.2, where the E denotes the APT1. E-ABP denotes the reversible complex formed between APT1 and the ABP. P denotes the covalently labeled APT1. The  $K_I$  and  $k_i$  denote the binding equilibrium constant and rate constants of the inactivation reaction, respectively. The double arrows between the ABP and E-ABP represent that APT1-ABP binding is reversible before the hydrolysis of the ABP.



**Equation 3.2 The reaction between APT1 and the ABP.** E denotes APT1. E-ABP is the reversible complex formed between APT1 and ABP, P is the irreversibly labeled APT1. The  $K_I$  and  $k_i$  denote the binding equilibrium constant and inhibition rate constant, respectively.

In addition, we define the initial APT1 concentration as  $[E_0] = [E] + [E-ABP] + [P]$ ; the active APT1 concentration at a time point  $t$  is  $[E_t] = [E] + [E-ABP]$ . Based on what we known, we will have Equation 3.3(a) and 3.3(b):

$$K_I = \frac{[E] \times [ABP]}{[E-ABP]} \quad (\text{Equation 3.3a})$$

$$-\frac{d([E_t])}{dt} = k_i \times [E-ABP] \quad (\text{Equation 3.3b})$$

From  $[E_t] = [E] + [E-ABP]$  and Equation 3.3a, we can get

$$[E] = \frac{K_I \times [E_t]}{K_I + [ABP]} \quad (\text{Equation 3.3c})$$

From Equation 3.3b and Equation 3.3c, we can get Equation 3.3d:

$$-\frac{d([E_t])}{dt} = k_i \frac{[E_t][ABP]}{K_I + [ABP]} \quad (\text{Equation 3.3d})$$

The solution of Equation 3.3d is

$$\ln \frac{[E_t]}{[E_0]} = - \frac{k_i}{1 + \frac{K_I}{[ABP]}} t \quad (\text{Equation 3.4})$$

For  $[ABP] \gg [E_0]$ , the concentration of the ABP can be considered unchanged. Let  $k_{app} =$

$\frac{k_i}{1 + \frac{K_I}{[ABP]}}$ , then we will have Equation 3.5a where  $k_{app}$  is the apparent rate constant of a

pseudo-first order reaction, as well as Equation 3.5b:

$$\ln \frac{[E_t]}{[E_0]} = -k_{app} t \quad (\text{Equation 3.5a})$$

$$\frac{1}{k_{app}} = \frac{1}{k_i} + \frac{K_I}{k_i} \times \frac{1}{[ABP]} \quad (\text{Equation 3.5b})$$

From Equation 3.5a, we know the decay of APT1 activity should follow a pseudo first-order rate equation, and the apparent rate constant of this exponential decay is  $k_{app}$ . When the APT1 is incubated with the ABP, the remaining active APT1 concentration  $[E_t]$  at the time point  $t$  is:

$$[E_t] = [E_0] \times e^{-k_{app}t} \quad \text{(Equation 3.6)}$$

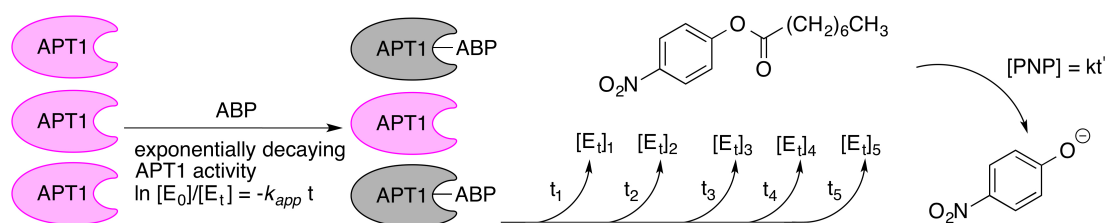
The remaining active APT1 can be assayed by hydrolyzing the colorimetric substrate *p*NPO in a secondary kinetic assay.

From Equation 3.5b, we can conclude that the values of  $K_I$  and  $k_i$  can be determined by plotting the apparent rate constants  $k_{app}$  against the ABP concentrations ( $[ABP]_s$ ). Based on these conclusions, the experiment described below was used to measure the  $K_I$  and  $k_i$  of the ABP.<sup>3</sup>

#### 4.2. Experiment design

The concept of the kinetic inhibition assay is illustrated in Figure 3.5. In the primary assay, active APT1 is incubated with the ABP at a concentration  $[ABP]_1$ . After incubation and inactivation by the ABP, the remaining activity of the APT1 should decay exponentially in a time- and concentration-dependent manner. To assay the remaining activity of the APT1 at various time points ( $t_1$  to  $t_5$ ), for each time point, an aliquot of the primary assay incubation was added to a solution of colorimetric substrate *p*NPO in a secondary assay. Since the  $[pNPO] \gg [ABP]$  and we monitor each reaction to less than 10% consumption of the  $[pNPO]$ , the hydrolysis of the *p*NPO in the secondary assay follows a pseudo zeroth-order rate equation. The remaining activities of the APT1 in the aliquots at different time

points are subsequently fitted to the first-order decay equation (Equation 3.6) to determine the apparent rate constant  $k_{app1}$  at the ABP concentration  $[ABP]_1$ . By varying the ABP concentration, different apparent rate constants will be obtained and plotted against the  $[ABP]$ . By fitting this  $k_{app}$  vs.  $[ABP]$  plot to Equation 3.5b, the values of  $K_I$  and  $k_i$  can be determined.

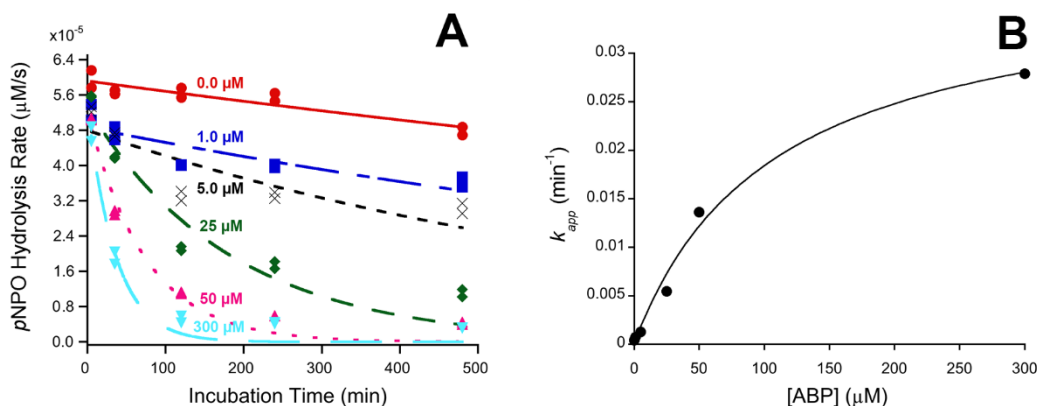


**Figure 3.5 Concept of the inhibition kinetic assay.** APT1 is first incubated with ABP in a primary assay. At time point  $t_x$ , the remaining activity of the APT1 in an aliquot of the primary assay ( $[E_t]_x$ ) incubation was measured by adding  $p$ NPO in the secondary assay. The hydrolysis of the  $p$ NPO in the secondary assay follows a pseudo zeroth-order rate equation.

#### 4.3. Results and discussion

To determine the kinetic properties of the APT1 inhibition by the ABP, APT1 (2.3  $\mu$ M) was pre-incubated in buffer (20 mM HEPES, 150 mM NaCl, 10% glycerol, 0.08% Triton X-100, pH = 7.4) with various concentrations (0, 1, 5, 25, 50, 300  $\mu$ M) of ABP at room temperature. Subsequently, to 30  $\mu$ L of  $p$ NPO (400  $\mu$ M) solution was added 10  $\mu$ L of the incubation mixture at various time points (5, 35, 120, 240, 480 min) during the incubation. The rates of  $p$ NPO hydrolysis were measured using a FLUOStar Galaxy microplate reader ( $\lambda_{abs} = 405$  nm) at each time point. The remaining APT1 activity at each time point was

fitted to an exponential decay rate expression (Equation 3.6) to determine the apparent inhibition rate constant ( $k_{app}$ ) at specific ABP concentrations [ABP] (Figure 3.6a). The  $k_{app}$  vs. [ABP] data were fit to the irreversible inhibition rate equation (Equation 3.6b) to obtain the inhibition constant ( $K_I$ ) and inactivation rate constant ( $k_i$ ), as shown in Figure 3.6b and Table 3.4.



**Figure 3.6 The inhibition of the APT1 by the ABP.** (a) The concentration- and time-dependent exponential decay of APT1 activity indicates that the ABP inhibits APT1 activity irreversibly. The solid lines represent the fit of the data to the exponential decay rate equation. The concentrations listed in the plot correspond to the concentration of ABP in each experiment. (b) Apparent inhibition rate constant ( $k_{app}$ ) vs. ABP concentration. The solid line represents the fit of the data to the irreversible inhibition rate equation.



| [ABP] ( $\mu\text{M}$ ) | $k_{app}$ of 1st order exponential decay ( $\text{min}^{-1}$ ) |
|-------------------------|----------------------------------------------------------------|
| 0.0                     | $4.05 \times 10^{-4}$                                          |
| 1.0                     | $7.28 \times 10^{-4}$                                          |
| 5.0                     | $1.28 \times 10^{-3}$                                          |
| 25                      | $5.45 \times 10^{-3}$                                          |
| 50                      | $1.36 \times 10^{-2}$                                          |
| 300                     | $2.79 \times 10^{-2}$                                          |

**Table 3.4a** The apparent rate constants obtained from the incubations. The concentration of the ABP [ABP] varies from 0  $\mu\text{M}$  to 300  $\mu\text{M}$ . Data was obtained by fitting the remaining APT1 activity ( $[E_t]$ ) data to the exponential decay equation (Equation 3.6).

| $K_I$ ( $\mu\text{M}$ ) | $k_i$ ( $\text{min}^{-1}$ ) |
|-------------------------|-----------------------------|
| $106 \pm 22$            | $0.038 \pm 0.003$           |

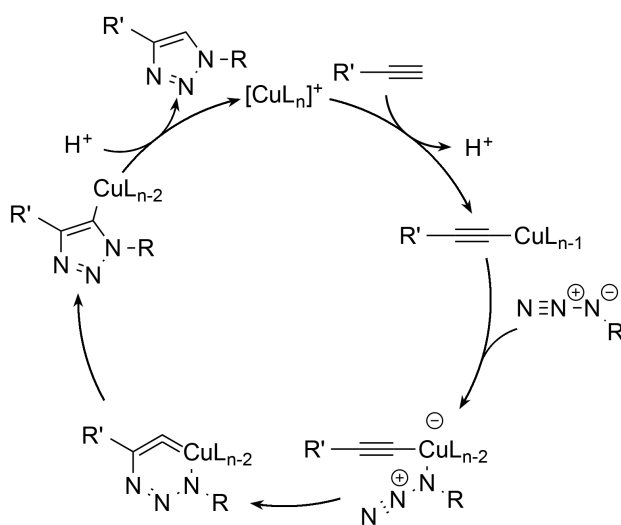
**Table 3.4b** The inhibition constant ( $K_I$ ) and inactivation rate constants ( $k_i$ ) of the ABP for APT1. Data was obtained by fitting the data in Table 3.3a to the Equation 3.5b in Kaleidagraph (Synergy Software, Version 4.03).

The binding equilibrium constant  $K_I$  is relatively low, which suggests good affinity between the ABP and the APT1 enzyme. Moreover, the exponential decay of the APT1 activity in the assay strongly suggests that the ABP labeled the APT1 in an irreversible manner as proposed.

## 5. In-gel Fluorescence of Fluorescently Labeled APTs

### 5.1. Introduction

The ABP was designed and synthesized with a terminal alkyne handle that can be reacted with an azide tag through copper (I) catalyzed click chemistry (Scheme 3.3).<sup>4</sup> This unique “tag-free ABP” strategy enables a versatile approach to profile the activity of APTs in biological system,<sup>5</sup> and avoids the bulkiness issue of the pre-conjugated tag. One of the objectives is tagging the APTs with a fluorescently-conjugated tag in order to accomplish quick and easy detection of enzymatic activity.



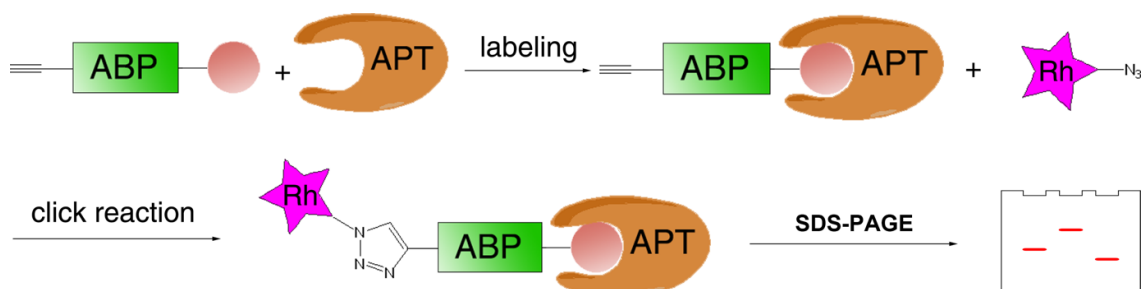
**Scheme 3.3 Proposed mechanism of the copper (I)-catalyzed click chemistry.** Under basic conditions, the terminal alkyne and azido group associate with copper (I) to form a complex. The cycloaddition step of the click chemistry features an enormous rate acceleration of  $10^7$  to  $10^8$  compared to the uncatalyzed 1,3-dipolar cycloaddition. Subsequently, the azido group inserts into the carbon-copper bond to finish the construction of triazole ring. The final step is the dissociation of the product from the copper (I) catalyst.

The classic copper (I)-catalyzed alkyne-azide cycloaddition (CuAAC) reaction requires Cu (I) salts or complexes as catalyst.<sup>4c</sup> However, there are two issues that need to be addressed with this approach. First, the solubility of the inorganic Cu (I) salts are generally poor in

aqueous solution, which will limit the application of inorganic copper (I) salts as a CuAAC reaction catalyst. Second, Cu (I) compounds are often labile to oxidants, including oxygen. To avoid these two issues, reducing Cu (II) compounds *in situ* is generally preferred to using Cu (I) directly for CuAAC reactions, especially for complicated biological systems. There are two reducing reagents commonly used to reduce the water-soluble CuSO<sub>4</sub> salt, sodium ascorbate (NaAsc) or tris(2-carboxyethyl)phosphine (TCEP).<sup>6</sup> While TCEP performs decently as a reducing agent, it can bind to copper at high concentrations and thus slow the reaction rate, or disable the reaction through an undesired Staudinger reaction by reducing the –N<sub>3</sub> to –NH<sub>2</sub>. In contrast, the high yields obtained using sodium ascorbate have established it as a very good reductant for the CuAAC reaction. Moreover, polytriazole ligands can improve the oxidation-resistance of copper (I) compounds. Among the common polytriazole ligands, tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) has been established as the most successful ligand for copper-catalyzed reactions in aerobic, aqueous environments.<sup>7</sup> This superiority is principally due to TBTA's tetradentate binding nature. The ligand fully surrounds CuI and occupies all binding sites, which prevents possible destabilizing interactions. Use of TBTA raises the Cu (I)/Cu (II) redox potential by approximately 300 mV, which increases the yield of air-exposed copper-catalyzed reactions from 10% to over 80%.

CuAAC is also a bioorthogonal conjugation reaction, meaning that the CuAAC reaction can occur in the context of a biological system without interfering with native biochemical processes. Moreover, the CuAAC reaction is rapid, efficient, and regioselective at room temperature. It has proven to be a powerful tool to conjugate biological macromolecules with synthetic chemical tags without affecting biological function. To demonstrate this

concept (Figure 3.7), after the APT enzyme is covalently labeled by the ABP, the APT-ABP complex (labeled with free terminal alkyne) is subjected to reaction with a fluorescently-conjugated azide tag under the click chemistry conditions. The fluorescently-conjugated tag becomes tethered to the labeled protein through a stable triazole adduct. The labeled proteins are then resolved on a sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) and the proteins having been labeled by a fluorescent tag can be visualized by using a fluorescent gel scanner.



**Figure 3.7 The concept of in-gel fluorescence detection of APT activity.** The APT enzyme is first labeled by the ABP in an irreversible manner. Subsequently, the fluorescent azide tag (rhodamine azide, Rh-N<sub>3</sub>) is conjugated to the terminal alkyne on the ABP through click chemistry. The APT-ABP-Rh complex is then analyzed by the fluorescence SDS-PAGE technique.

## 5.2. Experimental method

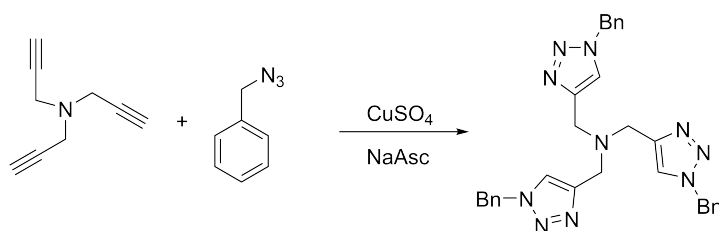
The TBTA ligand was synthesized according to the reaction described in Scheme 3.4. To a solution of benzyl azide (6.0 g, 45 mmol) in CH<sub>2</sub>Cl<sub>2</sub>/H<sub>2</sub>O (1:1, 150 mL) was added tripropargylamine (1.97 g, 15 mmol). CuSO<sub>4</sub>·5H<sub>2</sub>O (5 mol %) and NaAsc (15 mol %). The reaction mixture was stirred at room temperature for 24 hr. The organic layer was partitioned and the aqueous layer was extracted twice with 100 mL aliquots of CH<sub>2</sub>Cl<sub>2</sub>. The

organic layers were combined, dried over anhydrous  $\text{MgSO}_4$ , and concentrated *in vacuo*. The crude product was purified by flash chromatography column (5% MeOH in  $\text{CH}_2\text{Cl}_2$ ) to afford the TBTA as a white solid (5.5 g, 67%).<sup>8</sup>

$^1\text{H}$  NMR (600 MHz,  $\text{CDCl}_3$ )  $\delta$  7.70 (s, 3H), 7.34-7.37 (m, 9H), 7.26 (d,  $J = 7.0$  Hz, 6H), 5.51 (s, 6H), 3.72 (s, 6H).

$^{13}\text{C}$  NMR (150 MHz,  $\text{CDCl}_3$ )  $\delta$  144.0, 134.7, 129.1, 128.7, 128.0, 123.9, 54.1, 47.0.

HRMS-ESI: calculated for  $\text{C}_{30}\text{H}_{30}\text{N}_{10}$  ( $[\text{M} + \text{H}]^+$ ) 531.2733, observed 531.2714.



**Scheme 3.4 Synthesis of the tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA).**

The recombinant proteins were purchased from Abcam. For the APT1 and APT2 labeling experiments, 10  $\mu\text{g}$  of each active protein was incubated with 4  $\mu\text{g}$  ABP (0.25 mM) in buffer (10% glycerol, 100 mM NaCl, 20 mM Tris, 1 mM DTT, pH = 8.0) at room temperature for 16 hours. After the incubation, the small molecules were removed from the protein using a 10 kDa centrifugal filter by washing with  $2 \times 200$   $\mu\text{L}$  PBS buffer (pH = 8.0), concentrated to 20  $\mu\text{L}$ , and then dried under vacuum. To a clean eppendorf tube were added  $\text{CuSO}_4$  solution in water (5  $\mu\text{L}$ , 11 mM), a TBTA solution in DMSO (5  $\mu\text{L}$ , 1.1 mM) and a freshly prepared sodium ascorbate (NaAsc) solution in water (5  $\mu\text{L}$ , 20 mM). After centrifugation, 3  $\mu\text{L}$  of the supernatant was transferred to each modified protein sample.

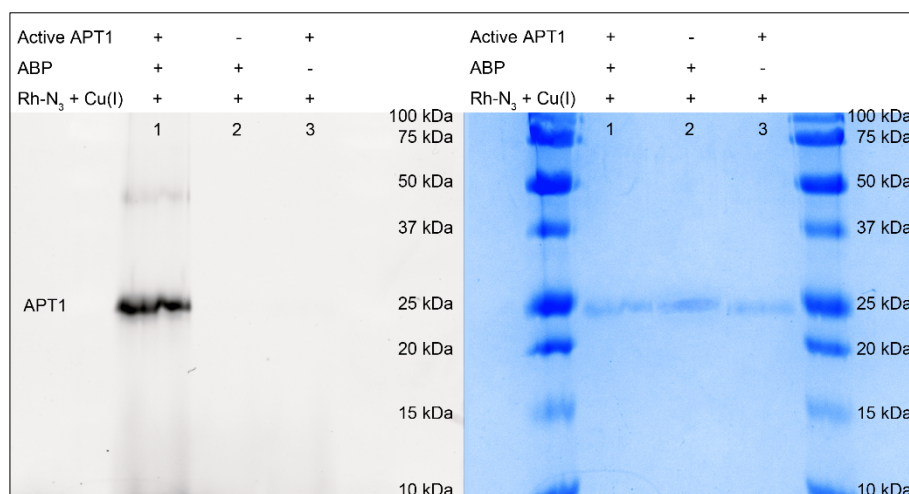
To each sample was added 5-tetramethylrhodamine azide (Rh-N<sub>3</sub>) solution in DMSO (1  $\mu$ L, 0.38 mM, Lumiprobe). The reaction stood at room temperature for 1 h before being quenched with 2  $\mu$ L of 5 $\times$  SDS-PAGE loading buffer. The protein was denatured by heating at 98 °C for 5 min. Protein samples were loaded directly onto a polyacrylamide gel (Mini-PROTEAN TGX AnyKD, Bio-Rad) for electrophoresis. As a negative control, heat-inactivated (1 min at 98 °C) APT1 or APT2 protein was also incubated with ABP and processed using the same protocol.

The polyacrylamide gel was imaged using a ChemiDoc MP imaging system (Bio-Rad). Fluorescence bands were imaged with 302 nm UV light; the Coomassie Brilliant Blue stained gel was imaged with white light. Gel images were processed with Image Lab software (version 5.1, build 8, Bio-Rad).

### 5.3. Results and discussion

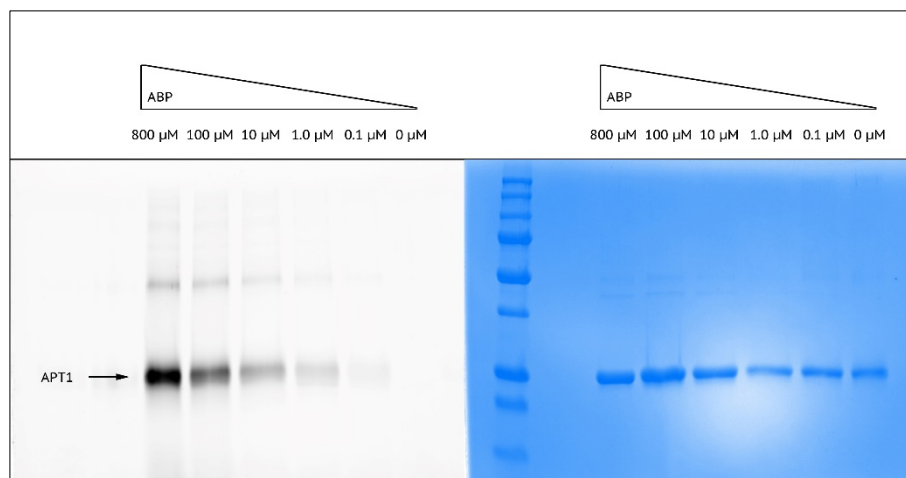
The gel images of recombinant APT1 are presented in Figure 3.8. The left panel is the fluorescence scan ( $\lambda_{\text{ex}} = 302$  nm) of the gel, the right panel is the gel stained by Coomassie Brilliant Blue under white light, which indicates the total abundance of the protein. In Lane 1, active APT1 was incubated with ABP and the fluorescent azide tag was conjugated according to previously described click chemistry protocol. As a negative control, in Lane 2 heat-inactivated APT1 was treated in the same manner as Lane 1. Lane 3 is another control reaction in which the ABP was left out in order to rule out none-specific fluorescence of the protein after the treatment. The results clearly show that the fluorescence was only observed in Lane 1 in which the active APT1 enzyme was processed. In contrast, the ABP did not label the inactivated APT1 in Lane 2 to any observable level.

However, in the Coomassie Brilliant Blue stained gel image on the right panel, APT1 bands near 25 kDa are clearly observed in all three lanes. In conclusion, the ABP labeled active APT1 enzyme only, regardless of the abundance.



**Figure 3.8 In-gel fluorescence labeling experiment result of APT1.** Left: Fluorescence scan; Right: Coomassie Brilliant Blue stained gel scan. **Lane 1.** Active APT1 was incubated with ABP. The rhodamine azide tag was conjugated via click chemistry. **Lane 2.** Heat-inactivated APT1 was processed by the same protocol. **Lane 3.** Control reaction in which the ABP was left out. MW ladders are shown in the right panel for reference.

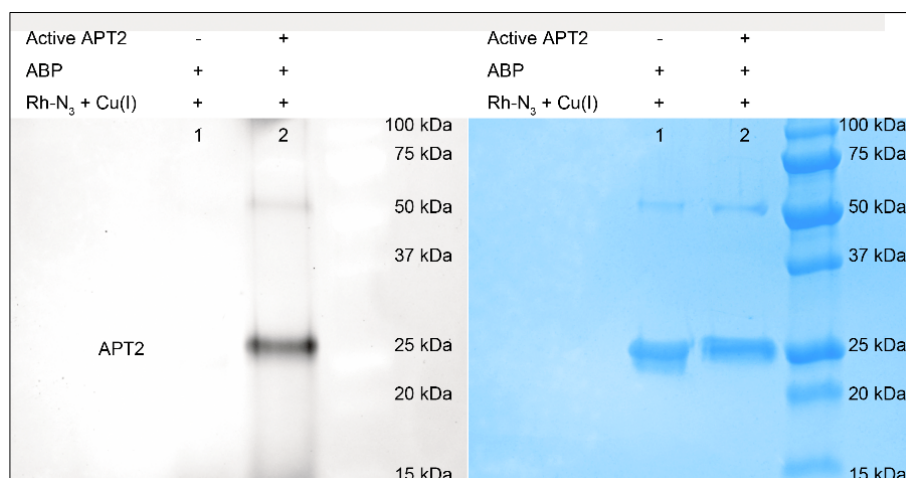
The APT1 was also treated with varying concentrations of ABP in separate incubations. Fluorescence tagging was then performed with rhodamine azide and Cu (I) catalyst, as described above. Fluorescent bands were visible on the resulting gel at an ABP concentration of just 1  $\mu$ M, demonstrating the high efficiency of labeling, and also the high sensitivity of this method. More intense labeling was observed for incubations of higher ABP concentrations (Figure 3.9). These results confirm that the ABP specifically labels active APT1 in a concentration-dependent manner.



**Figure 3.9 Concentration-dependent fluorescence labeling of APT1.** Left: Fluorescence scan; Right: Coomassie Brilliant Blue stained gel.

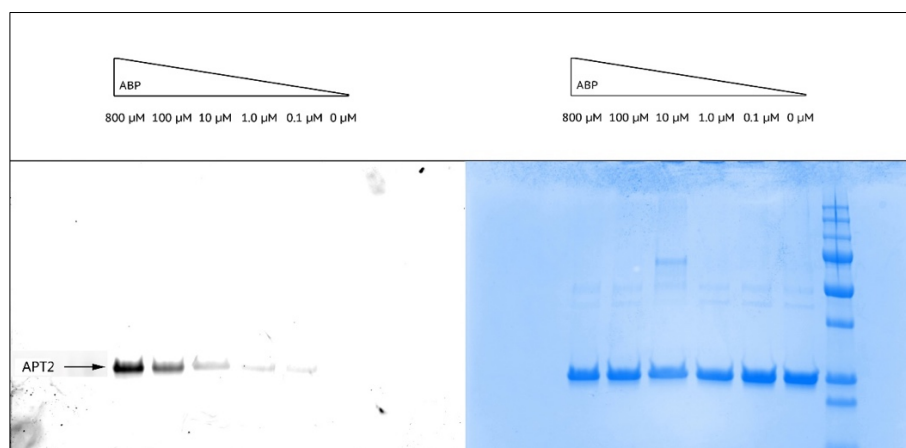
APT2 was also assayed by the same approach. ABP-treated APT2 was subjected to Cu (I)-catalyzed click chemistry with the rhodamine azide tag. These reactions were then resolved by electrophoresis on a polyacrylamide gel. In-gel fluorescence visualization was performed to discern extent of labeling, and subsequent Coomassie Brilliant Blue staining was performed to indicate overall protein abundance. Heat-inactivated APT2 underwent the same processing steps in parallel alongside active APT2. As a result, fluorescent bands were detected only in lanes containing active enzyme (Figure 3.10).





**Figure 3.10 In-gel fluorescence experiment using APT2.** Left: Fluorescence scan; Right: Coomassie Brilliant Blue stained gel. Lane 1: heat-inactivated APT2 was incubated with ABP and then tagged by click chemistry. Lane 2: active APT2 was processed using the same protocol as Lane 1. A MW ladder is shown in the right panel for reference.

A concentration-dependent APT2 labeling experiment was also performed to demonstrate the sensitivity of the ABP. The same procedure was followed to treat the APT2 with ABP of various concentrations. The resulting gel images are presented in Figure 3.11. Similarly, fluorescence can be observed within the incubation containing 1  $\mu$ M ABP, and the intensity gets stronger while the concentration of the ABP increasing to 100  $\mu$ M.



**Figure 3.11 Concentration-dependent fluorescence labeling of APT2.** Left: Fluorescence scan; Right: Coomassie Brilliant Blue stained gel.

In conclusion, the ABP only labels active APT enzymes regardless of the total protein abundances. The in-gel fluorescence experiments also validate the “tag-free ABP” strategy, and provide evidence to support the irreversible labeling of the APT by the ABP. IN addition, these experiments show that the labeling is concentration-dependent.

## 6. References

1. Johnson, K. A.; Goody, R. S., The Original Michaelis Constant: Translation of the 1913 Michaelis-Menten Paper. *Biochemistry-Us* **2011**, *50* (39), 8264-8269.
2. Dekker, F. J.; Rocks, O.; Vartak, N.; Menninger, S.; Hedberg, C.; Balamurugan, R.; Wetzel, S.; Renner, S.; Gerauer, M.; Scholermann, B.; Rusch, M.; Kramer, J. W.; Rauh, D.; Coates, G. W.; Brunsveld, L.; Bastiaens, P. I. H.; Waldmann, H., Small-molecule inhibition of APT1 affects Ras localization and signaling. *Nat. Chem. Biol.* **2010**, *6* (6), 449-456.

3. (a) Kumar, S.; Zhou, B.; Liang, F. B.; Wang, W. Q.; Huang, Z. H.; Zhang, Z. Y., Activity-based probes for protein tyrosine phosphatases. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101* (21), 7943-7948; (b) Kitz, R.; Wilson, I. B., Esters of methanesulfonic acid as irreversible inhibitors of acetylcholinesterase. *J. Biol. Chem.* **1962**, *237* (10), 3245-3249.
4. (a) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G., Bioconjugation by copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. *J. Am. Chem. Soc.* **2003**, *125* (11), 3192-3193; (b) Rodionov, V. O.; Fokin, V. V.; Finn, M. G., Mechanism of the ligand-free Cu-I-catalyzed azide-alkyne cycloaddition reaction. *Angew. Chem. Int. Ed.* **2005**, *44* (15), 2210-2215; (c) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **2002**, *41* (14), 2596-+.
5. Speers, A. E.; Adam, G. C.; Cravatt, B. F., Activity-based protein profiling in vivo using a copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. *J. Am. Chem. Soc.* **2003**, *125* (16), 4686-4687.
6. (a) Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V., Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates. *J. Am. Chem. Soc.* **2005**, *127* (1), 210-216; (b) Berg, R.; Straub, B. F., Advancements in the mechanistic understanding of the copper-catalyzed azide-alkyne cycloaddition. *Beilstein J. Org. Chem.* **2013**, *9*, 2715-2750.
7. Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V., Polytriazoles as copper(I)-stabilizing ligands in catalysis. *Org. Lett.* **2004**, *6* (17), 2853-2855.

8. Lee, B. Y.; Park, S. R.; Jeon, H. B.; Kim, K. S., A new solvent system for efficient synthesis of 1,2,3-triazoles. *Tetra. Lett.* **2006**, 47 (29), 5105-5109.

## **CHAPTER FOUR**

### **Using Mass Spectrometry to Investigate APT-ABP Labeling**



## 1. Abstract

Having successfully constructed and validated the ABP through kinetic assays and the gel-based method, we sought more direct evidence of APT labeling in our continued efforts to validate the proposed mechanism of the ABP.

To investigate labeling by the ABP, the ABP-treated APTs were analyzed by electrospray ionization mass spectrometry (ESI-MS). After deconvolution of the resulting multiply charged peaks, the mass shift induced by ABP labeling was clearly identified. As a control experiment, heat-inactivated APT was also treated and analyzed following the same protocol. Additionally, to rule out possible nonspecific labeling caused by the spontaneous hydrolysis of the ABP, active APT1 was incubated with ABP alongside cytochrome C, which does not possess known lipase activity. The results indicate that the ABP labels active APT1 and APT2 once, and not their inactive forms. This finding strongly suggests the ABP labels the APTs in an activity-based manner in their active sites, in accordance with our proposed mechanism of action.

In order to discern the labeling site of the ABP, tandem mass spectrometry was employed to analyze the peptides resulting from the proteolytic digestion of the ABP-treated APT1. By matching the fragmentation patterns with database sequence information, a Cys211 residue near the active site of APT1 was identified as the site of modification. Previously discussed molecular dynamics simulations predicted that this residue is within reasonable distance for reaction with the ABP.

## 2. Introduction to the Protein Mass Spectrometry

## 2.1. Common mass spectrometry techniques for protein analysis

The origin of mass spectrometry (MS) as an analytical technique can be dated back to the last century.<sup>1</sup> Among the common analytical technologies, MS holds a crucial position because it can measure the mass of a molecule with extremely high sensitivity and accuracy. Beginning in the last decade of the 20<sup>th</sup> century, mass spectrometry has played an increasingly important role in the biological sciences.<sup>2</sup>

The mass spectrometer requires gaseous charged molecules for analysis. However, compared to small molecules, large biological analytes such as proteins were extremely difficult to ionize and transfer into the gas phase, which limited the application of MS for protein analysis. Since then, many efforts have been devoted to develop innovative ionization sources for macromolecules. As a result, mass spectrometers equipped with electrospray ionization (ESI)<sup>3</sup> and matrix-assisted laser desorption ionization (MALDI)<sup>4</sup> techniques have been invented for biological molecules. Currently, the applications of MS in protein analysis are in two major fields. First, MS is commonly used for protein identification in large-scale proteomic investigations on complicated biological systems.<sup>5</sup> Second, because MS measures the molecular weight of a protein, it is the method of choice for the detection and characterization of post-translational modifications (PTMS) and chemical modification that alter the mass of a protein.<sup>6</sup>

## 2.2. Electrospray Ionization (ESI)

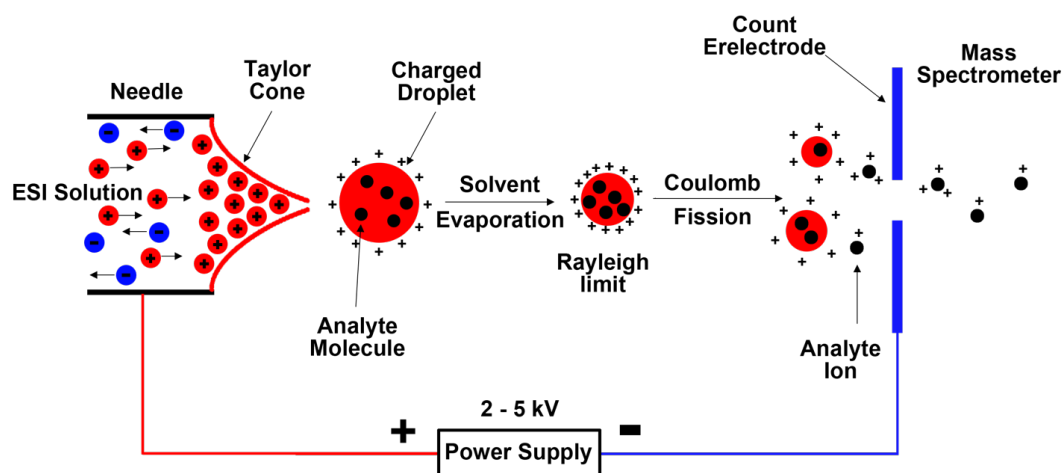
Electrospray ionization (ESI) was developed for use in biological mass spectrometry by Fenn.<sup>5a, 7</sup> Over the last decade, ESI mass spectrometry (ESI-MS) has emerged as an important analytical technique in the laboratory. Coupled with a high performance liquid



chromatograph (HPLC) for molecular fractionation prior to MS analysis, HPLC/ESI-MS has become a very powerful technique capable of analyzing both small and large molecules in complex biological samples.

The ESI technique uses a high-voltage electrical field to transfer the analyte ions from the liquid phase into the gaseous phase before they are subjected to MS analysis. Some of the neutral compounds can also be ionized by protonation or deprotonation in the liquid phase, and hence can be studied by the ESI-MS.

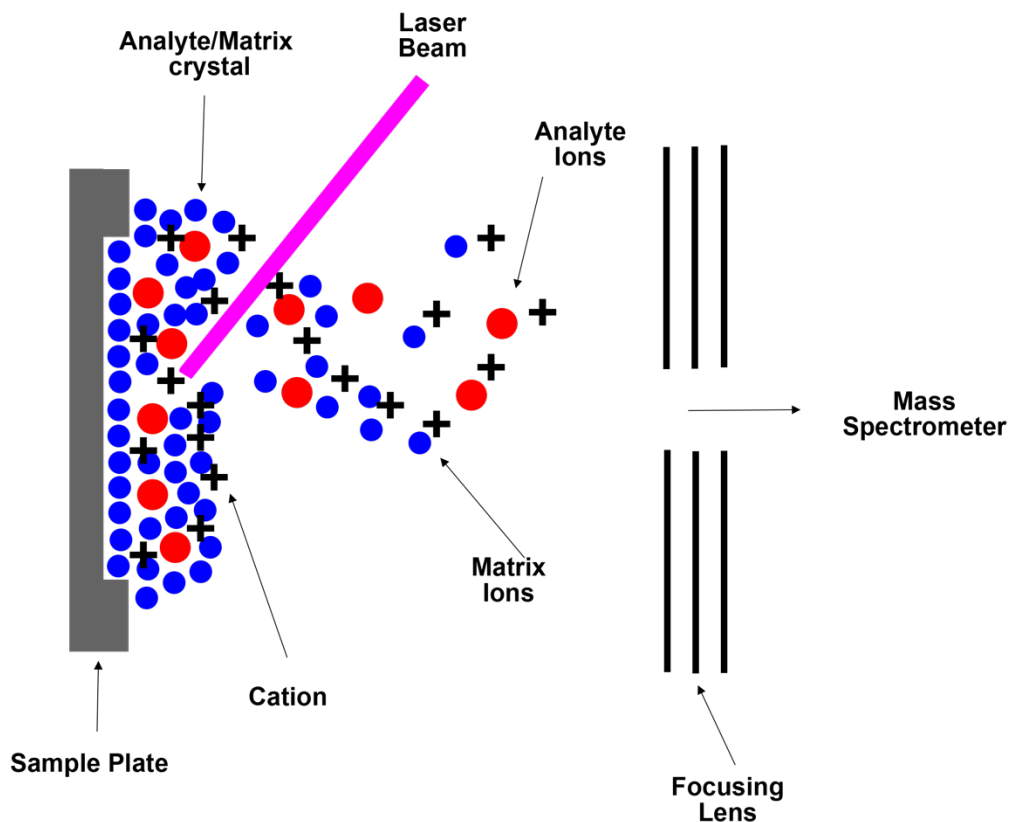
The mechanism of the ESI is illustrated in Figure 4.1. There are three major steps in the process of ESI. First, the solution containing the analyte(s) of interest is dispersed by the electrospray.<sup>8</sup> The typical voltage required to draw the liquid into a Taylor cone is 2 to 5 kV.<sup>9</sup> The resulting charged droplets from the electrospray are desolvated in the heated vacuum chamber and accelerated by the electric field. The solvent evaporates from charged droplets until they become unstable upon reaching the Rayleigh limit.<sup>10</sup> Due to the decreasing droplet volume and increasing surface charge density, the surface tension of the droplet cannot hold the droplet together. Thus, the droplets undergo Coulomb fission, whereby the droplets explode and create a few much smaller but more stable droplets. The resulting new droplets will also undergo desolvation and Coulomb fission, until only the naked charged analyte(s) remain in the gas phase. The gaseous charged analyte(s) are subsequently analyzed by the mass spectrometer. The ions observed in the mass spectrometer are quasimolecular ions created by protonation ( $[M + H]^+$ ) or addition of another cation such as a sodium ion ( $[M + Na]^+$ ). A notable characteristic feature of ESI-MSI is the observation of multiple charged analytes ( $[M + nH]^{+n}$ ), especially for large macromolecules.



**Figure 4.1 Mechanism of electrospray ionization.** The solution containing analyte(s) of interest is electrosprayed to give charged droplets. The droplet is then desolvated and accelerated until the Rayleigh limit is reached. After Coulomb fission, smaller but stable droplets will undergo the cycle again and finally the charged analyte(s) in the gas phase will be characterized by the mass spectrometer.

### 2.3. Matrix-assisted laser desorption/ionization (MALDI)

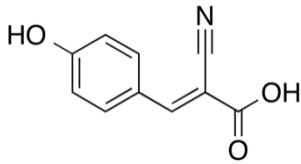
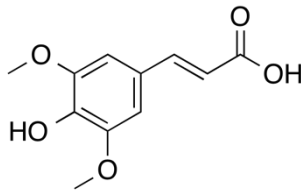
The general principle of MALDI revolves around the photo-desorption of an analyte(s) co-crystallized in a large molar-excess of a laser-absorbing matrix, followed by mass spectrometry analysis. This technique provides a non-destructive approach for analyzing both large and small biological molecules (Figure 4.2).<sup>4</sup> Although the mechanism of the ionization in MALDI is still not fully understood, several potential pathways have been proposed, including thermal ionization,<sup>11</sup> excited-state proton transfer,<sup>12</sup> ion-molecule reactions,<sup>13</sup> gas-phase photoionization,<sup>13</sup> and desorption of preformed ions.<sup>14</sup>

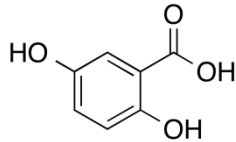
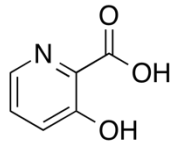
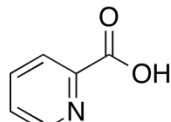


**Figure 4.2 The principle of matrix-assisted laser desorption/ionization (MALDI).** The analyte(s) crystallized with matrix on a sample plate is desorbed and transported to gas phase with the laser-absorbing matrix molecules. Subsequently, the resulting analyte ions are accelerated by electrodes and analyzed by the time-of-flight (TOF) mass spectrometer.

The selection of the laser-absorbing matrix molecule is crucial in MALDI experiments. The efficient energy transfer during a matrix-assisted desorption affords high ion yields of the analyte(s). The analyte(s) are usually co-crystallized in an excess of a solid matrix compound. The high excess of the matrix is essential since the matrix serves as the major laser absorber, and transfers the analyte(s) and itself into the gas phase. Instead of absorbing the laser radiation directly, the analyte(s) obtains energy from the excited matrix molecules, thus MALDI is considered a “soft” ionization technique that allows the analysis

of complex biological molecules without breaking covalent bonds. In Table 4.1, few of the widely used MALDI matrices are presented. For large biomolecule analysis,<sup>15</sup> the common matrices are  $\alpha$ -cyano-4-hydroxycinnamic acid (1) (CCA), 3,5-dimethoxy-4-hydroxycinnamic acid (2) (sinapinic acid or SA), 2,5-dihydroxybenzoic acid (3) (DHB), 3-hydroxypicolinic acid (4) (HPA), and pyridine-2-carboxylic acid (5) (PA). In general, CCA is primarily used for peptides and small proteins; SA is commonly used for both peptide and protein analysis, and DHB is used for small proteins and oligonucleotides. For the analysis of nucleic acids and DNA, HPA and PA are commonly used. In the application of peptide mapping (identifying proteins by matching the peptides that result from proteolytic digestion with sequence databases), the selection of the matrix plays a significant role in the analysis of the peptides. For example, using CCA in peptide mapping can yield more low-mass peptide ions (<2500 Da), while SA can provide better coverage for higher-mass peptides (>2500 Da).

| Compound                                                                                                                                        | Solvent                                        | Wavelength                 | Applications                        |
|-------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------|----------------------------|-------------------------------------|
|  <p>(1) <math>\alpha</math>-cyano-4-hydroxycinnamic acid</p> | acetonitrile<br>water<br>ethanol<br>acetone    | 337 nm<br>355 nm           | peptides<br>proteins<br>nucleotides |
|  <p>(2) 3,5-dimethoxy-4-hydroxycinnamic acid</p>             | acetonitrile<br>water<br>acetone<br>chloroform | 337 nm<br>355 nm<br>266 nm | peptides<br>proteins                |

|                                                                                                                     |                                                            |                            |                                                                 |
|---------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|----------------------------|-----------------------------------------------------------------|
| <br>(3) 2,5-dihydroxybenzoic acid  | acetonitrile<br>water<br>methanol<br>acetone<br>chloroform | 337 nm<br>355 nm<br>266 nm | peptides<br>nucleotides<br>oligonucleotides<br>oligosaccharides |
| <br>(4) 3-hydroxy picolinic acid   | ethanol                                                    | 337 nm<br>355 nm           | oligonucleotides                                                |
| <br>(5) pyridine-2-carboxylic acid | ethanol                                                    | 266 nm                     | oligonucleotides                                                |

**Table 4.1 The common matrices for analysis of large biological molecules by MALDI.**

Sample preparation greatly influences the data quality of MALDI mass spectrometry. The dried-droplet method is most frequently used for biological samples. In this protocol, the analyte solution is added to a saturated matrix solution at a ratio of about 1:5000. Subsequently, an aliquot (0.5 – 2.0  $\mu\text{L}$ ) of this mixture is distributed on a sample plate where it is allowed to dry. For peptides or proteins, the analyte(s) is first dissolved in a 0.1% aqueous trifluoroacetic acid (TFA) solution. After the analyte(s) is crystallized with the matrix compound, the mixture is subjected to laser radiation of a corresponding wavelength for analysis.

### 3. Whole Protein ESI-MS Validates the Labeling on APTs

#### 3.1. General experiment method

Recombinant Mouse APT1 and Human APT2 full-length proteins were acquired from Abcam. All mass spectrometric analyses were performed using a binary Agilent 1200 series HPLC directly coupled to an Agilent 6530 Q-TOF mass spectrometer equipped with an electrospray ionization source. Mobile phase: 5% (0 min) - 65% (10 min) - 95% (11 min) - 5% (14 min) acetonitrile with 0.1% formic acid in water. Stationary phase: Symmetry 300 3.5  $\mu\text{m}$  C<sub>18</sub> HPLC column (2.1 x 50 mm, Waters). Mass spectral data were analyzed with MassHunter Qualitative Analysis software (version B.06.00, Agilent). A maximum entropy algorithm was used to deconvolute multiply charged ions. The range of deconvoluted ions was set from 10 kDa to 67 kDa.

For the mass spectra of intact APT1 and APT2, 1  $\mu\text{g}$  of each protein was purified using a 10 kDa centrifugal filter (Amicon Ultra 0.5 mL, Millipore) by washing with  $2 \times 200$   $\mu\text{L}$  buffer (100 mM NaCl, 20 mM Tris, 1 mM DTT, pH = 8.0). 20  $\mu\text{L}$  of the protein solution was collected and directly injected into the mass spectrometer.

For the APT1 and APT2 labeling experiments, 10  $\mu\text{g}$  of each active protein was incubated with 4  $\mu\text{g}$  ABP (0.25 mM) in buffer (10% glycerol, 100 mM NaCl, 20 mM Tris, 1 mM DTT, pH = 8.0) at room temperature for 16 hours. The incubation mixture was purified using a 10 kDa centrifugal filter by washing with  $2 \times 200$   $\mu\text{L}$  Tris buffer (100 mM NaCl, 20 mM Tris, 1 mM DTT, pH = 8.0). The protein solution (5  $\mu\text{L}$ ) was injected into the mass spectrometer for analysis. Heat-inactivated APT1 and APT2 (98 °C, 5 min) was also subjected to the same incubation as a negative control.

### 3.2. Mass spectrometry of untreated APT1 and APT2

Based on the sequence of APT1 (Table 4.2), the theoretical molecular weight was calculated.

To measure the molecular mass of the untreated APT1 protein by ESI-MS, 1  $\mu$ g recombinant APT1 protein was desalted with a 10 kDa centrifugal filter and injected into the mass spectrometer.

| Name       | Sequence                                                                                                                                                                                                                                                                                     | Calculated MW |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|
| Mouse APT1 | GSSHHHHHHSSGLVPRGSHMCGNNMSA<br>PMPAVVPAARKATAAVIFLHGLGDTGH<br>GWAEAFAGIKSPHIKYICPHAPVMPVTL<br>NMNMAMPWFWDIVGLSPDSQEDESIGK<br>QAAETVKALIDQEVKNGIPSNRIILGGFS<br>QGGALSLYTALTQQKLAGVTALSCWL<br>PLRASFSQGPINSANRDISVLQCHGDCDP<br>LVPLMFGSLTVERLKALINPANVTFKIYE<br>GMMHSSCQQEMMDVKHFIDKLLPPID     | 26719.7 Da    |
| Human APT2 | GSSHHHHHHSSGLVPRGSHMCGNTMSV<br>PLLTDAA TVSGAERETA AVIFLHGLGDT<br>GHSWADALSTIRLPHVKYICPHAPRIPVT<br>LNMKMVMPSWFDLMGLSPDAPEDEAGI<br>KKAAENIKALIEHEMKNIGIPANRIVLGG<br>FSQGGALSLYTALTCPHPLAGIVALSCW<br>LPLHRAFPQAANGSAKDLAILQCHGELD<br>PMVPVRFGALTAEKLRVVTPARVQFKT<br>YPGVMHSSCPQEMAAVKEFLEKLLPPV | 26769.0 Da    |

**Table 4.2 The sequences and the calculated molecular weights of the APT1 and APT2 proteins.**

In ESI-MS, the protein will be multiply-protonated by the formic acid in the mobile phase.

If a protein with a molecular weight  $M$  was charged by  $n$  protons, the total mass of the

charged protein is  $[M+nH]^{n+} = M + n \times 1.008$ , the total charges of the protonated protein is  $n e^-$ . Thus, the mass-to-charge ratio is  $\frac{m}{z} = \frac{M+n \times 1.008}{n} = \frac{M}{n} + 1.008$ . Since the value of  $m/z$  depends on the charge state of the protein, if the protein molecules has multiple charging states ( $[M+H]^+$ ,  $[M+2H]^{2+}$ ,  $[M+3H]^{3+}$ ...  $[M+nH]^{n+}$ ), a series of peaks with different mass-to-charge ratios will be observed in the mass spectrum.

To calculate the molecular weight of the protein from the multiply-charged peaks, the charge state of each peak in the mass spectrum must first be determined. This aim can be achieved by observing the spacing pattern between the isotopic peaks of a single  $m/z$  peak. For the singly-charged ion, the spacing between the different isotopic peaks should be around  $m/z = 1 \text{ Da}/e^-$ . However, if the ion is multiply-charged, the difference of the mass-to-charge ratio will be  $m/z = 1 \text{ Da}/2e^- = 1/2 \text{ Da}/e^-$ . For the same reason, if the ion is charged by  $n$  protons, the spacing between the isotopic peaks is  $m/z = 1/n \text{ Da}/e^-$ .

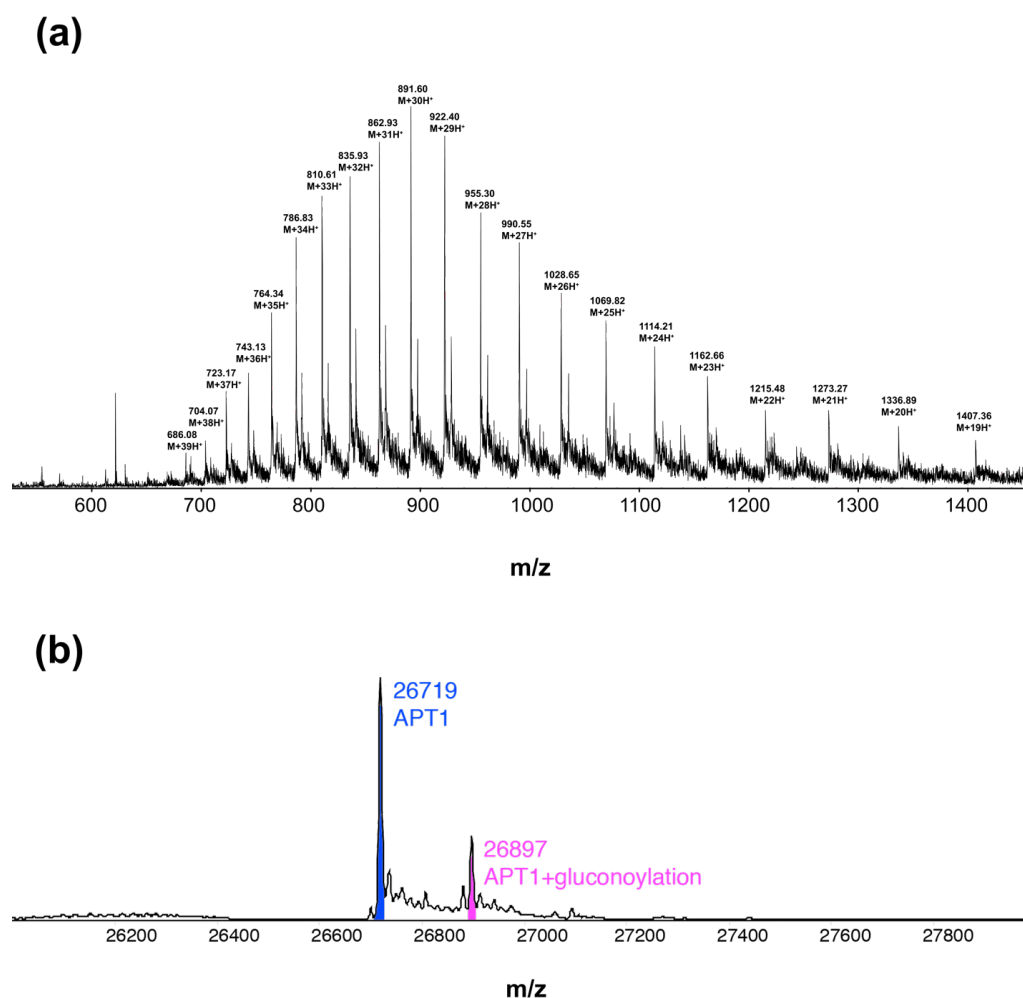
If the charge state  $n$  of a single  $m/z$  peak has been figured out, then the molecular weight (MW) of the protein can be calculated by

$$MW = \left( \frac{m}{z} - 1.008 \right) \times n$$

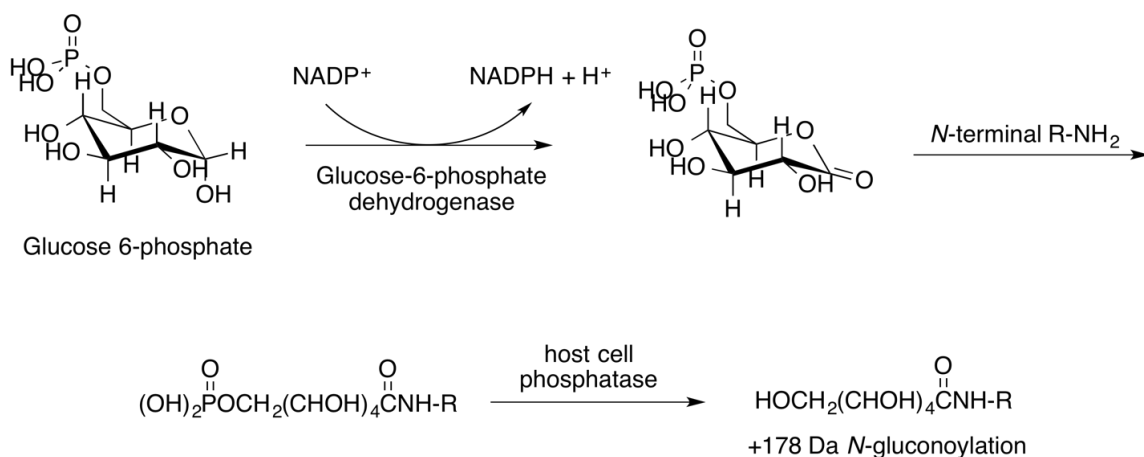
By averaging all peaks introduced by multiple charging, the error caused by artifacts and impurities will be minimized. As a result, the MS spectrum of untreated APT1 protein is presented in Figure 4.1a. All the peaks observed in the spectrum were deconvoluted using the MassHunter software to afford the molecular weight  $MW = 26719 \text{ Da}$  (Figure 4.1b), which agrees with the theoretical  $MW = 26719.7 \text{ Da}$  calculated from the sequence.



However, another deconvoluted peak at  $m/z = 26897$  Da was observed in Figure 4.2b, too. The mass shift of the peak is +178 Da, which corresponds to a gluconoylation post-translational modification (PTM). The gluconoylation (Scheme 4.1) is a common PTM on the *N*-terminal hexahistidine tag (His-tag) of recombinant proteins expressed in *Escherichia coli*.<sup>16</sup>

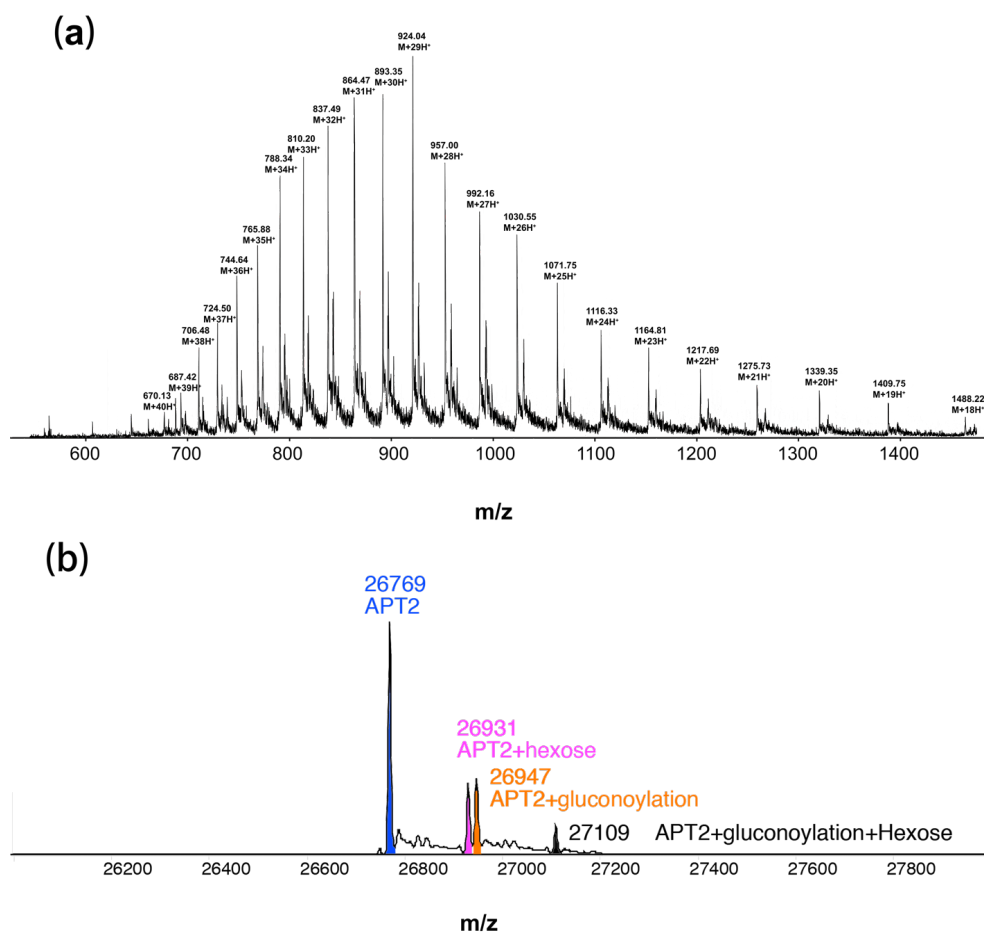


**Figure 4.3** The mass spectra of mouse recombinant APT1 protein. (a) The multiply charged peaks of the APT1 observed in the mass spectrometer (b) The deconvolution of the multiply-charged peaks in the MassHunter software package affords the molecular weight 26719 Da. The gluconoylation post-translational modification (+178 Da) was observed too.



**Scheme 4.1** The mechanism of *N*-terminal gluconoylation which introduces the mass shift of +178 Da in some recombinant proteins.

Following the same protocol, the mass spectrometric analysis of APT2 protein was also accomplished. The multiply charged peaks and deconvoluted mass spectrum are presented in Figure 4.4. Similarly, the molecular weight of APT2 acquired from deconvolution is 26769 Da, which matches the theoretical molecular weight calculated directly from the sequence. Additionally, gluconoylation PTM (+178 Da) was also observed for APT2. However, there are two new peaks in the mass spectrum; the correspondent mass shifts are +162 Da and +340 Da, respectively. The mass shift of 162 Da suggests it is the addition of a hexose molecule,<sup>17</sup> and the +340 Da peak at 27109 Da indicates a protein molecule with both gluconoylation and hexose addition.



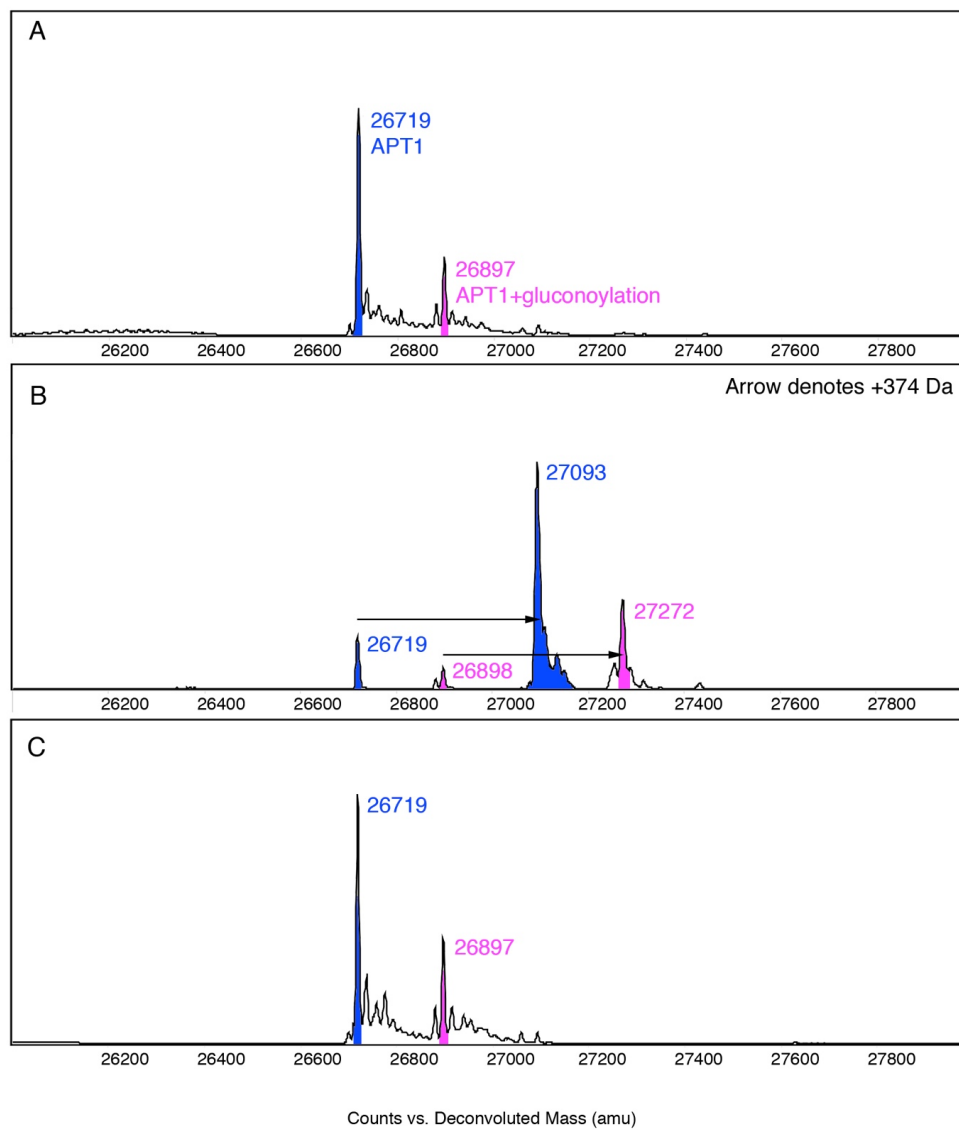
**Figure 4.4 The mass spectra of human recombinant APT2 protein.** (a) The multiply charged peaks of the APT2 observed in the mass spectrometer (b) The deconvolution of the multiply-charged peaks using the MassHunter software package affords the molecular weight 26769 Da. The gluconoylation post-translational modification (+178 Da) and addition of hexose (+162) were observed too. The protein with both post-translational modifications (+340 Da) was also observed.

In conclusion, using the ESI-MS, we successfully characterized untreated APT1 and APT2 proteins. The molecular weights obtained from the mass spectrometry experiments match the theoretical values calculated from the sequences. Additionally, post-translational

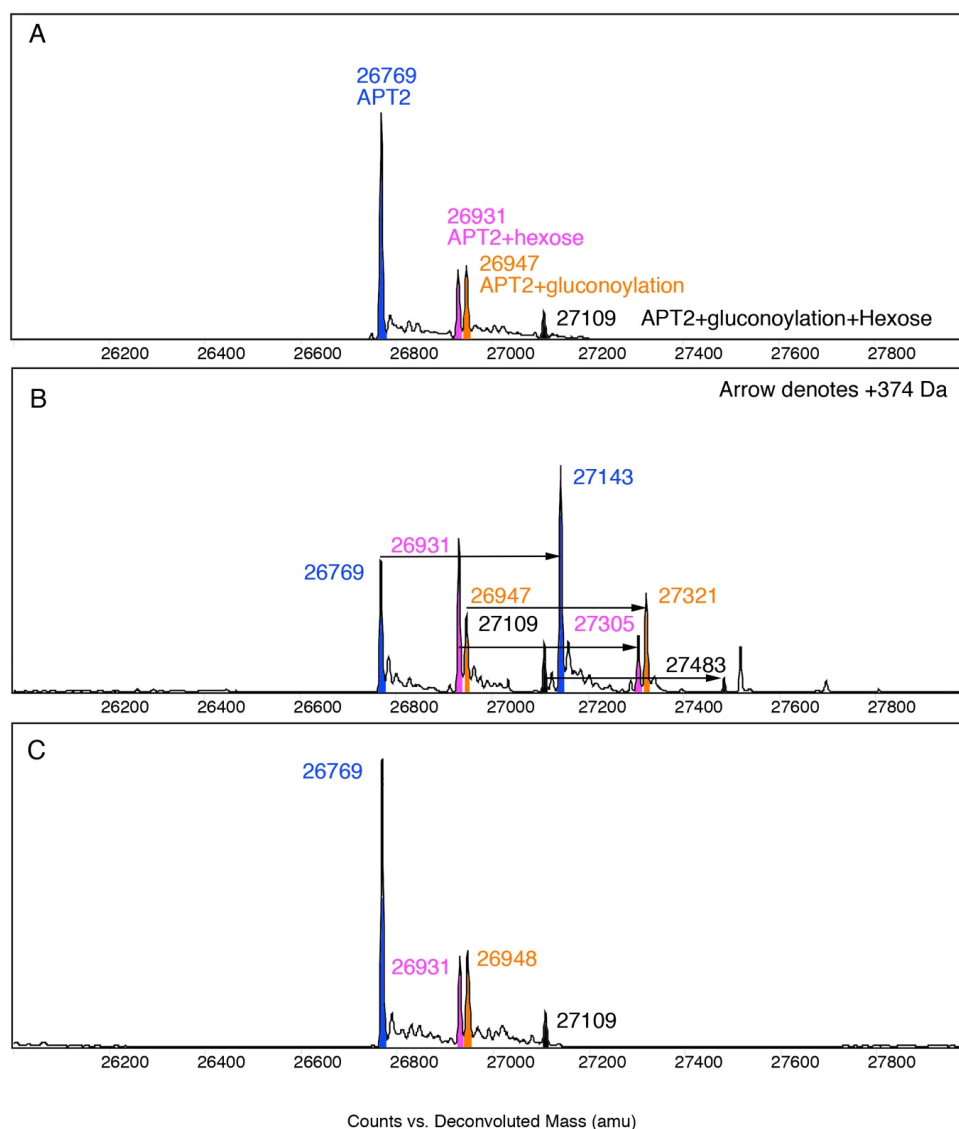
modifications on the proteins were also identified, which gave us confidence about the characterization of chemical modification on proteins introduced by the ABP.

### 3.3. Using ESI-MS to validate the chemical labeling of APTs

Seeking direct evidence that the ABP covalently labels APT1, we analyzed ABP-treated APT1 by ESI-MS. The resulting deconvoluted mass spectrum is presented in Figure 4.5. An observed mass peak at 26,719 Da agrees with the value calculated from the known sequence of unmodified APT1. An observed mass peak at 26898 Da corresponds to gluconoylation (+178 Da) of the N-terminal hexahistidine tag. A prominent mass peak observed at 27093 Da corresponds to the ABP-labeled APT1, indicated by a mass shift of +374 Da. This mass shift of +374 Da matches exactly the calculated mass difference resulting from covalent attachment of the quinone methide (+375 Da) with concomitant loss of a proton from nucleophilic residue (-1 Da). In addition, we observe a mass peak at 27272 Da, which corresponds to the gluconoylated APT1 that having been labeled. It is worth mentioning that only single labeling is observed on active APT1, which is consistent with our hypothesis that the ABP labels APT1 in an activity-dependent manner. Based on integration of these deconvoluted mass peaks, the extent of APT1 labeling was 75%, indicating high efficiency of the ABP. In a control experiment, heat-inactivated APT1 was also treated with the ABP, but only unmodified APT1 peaks are observed in MS. By MS analysis of ABP-treated APT1, we conclude that only active APT1 is labeled. This finding is in accordance with our ABP's proposed mechanism of action.



**Figure 4.5 Deconvoluted ESI-MS spectra of APT1.** Peaks corresponding to unmodified and modified APT1 are colored blue and red, respectively. The numbers above the arrows indicated the mass shifts of labeling. (A) MS of active APT1; (B) MS of ABP-treated APT1; (C) ABP-treated inactive APT1 as a control.



**Figure 4.6** (A) Deconvoluted mass spectrum of intact APT2. Calculated: 26769 Da; observed: 26769 Da. Gluconoylation of the *N*-terminal hexahistidine tag (+178 Da) was also observed at 26947 Da. The peak at 26931 Da (+162 Da) corresponds to the addition of a hexose unit to APT2. The peak at 27109 Da (+340 Da) would then correspond to APT2 with both gluconoylation and hexose addition. (B) Four new peaks were observed after incubating active APT2 with ABP. All the mass shifts were +374 Da, indicating that the four species observed in spectrum A were all labeled by a single quinone methide in spectrum B. (C) Heat-inactivated APT2 was not labeled after incubating with ABP using the same protocol.

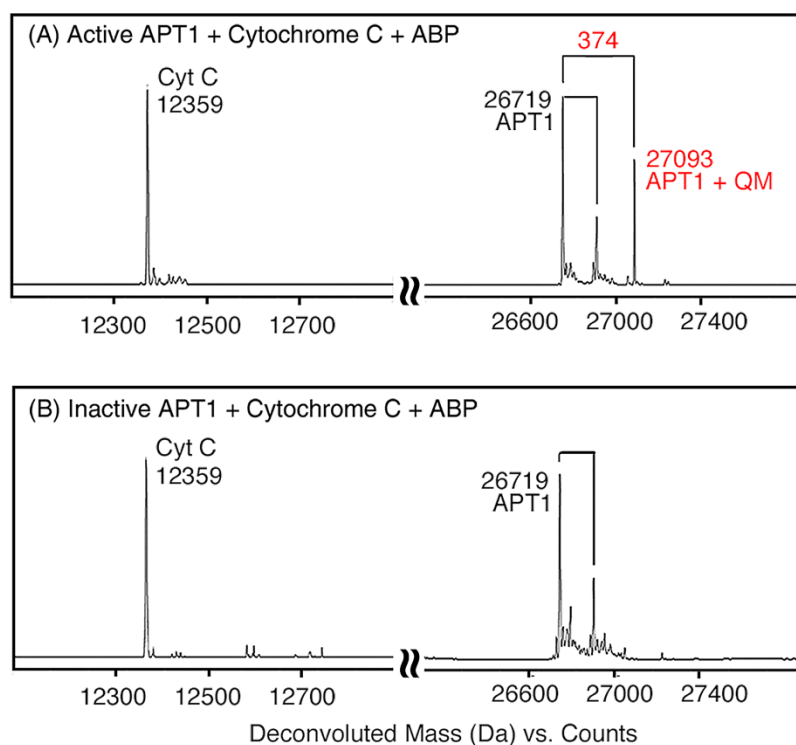
This analysis was also performed with ABP-treated APT2. The resulting deconvoluted mass spectrum is given in Figure 4.6. The results are similar with the APT1 labeling experiment. There are four peaks in the MS of intact APT2: the base peak with the mass of 26769 Da is consistent with the mass calculated from the protein sequence. Two post-translationally modified peaks are also observed: gluconoylation and the addition of a single hexose molecule (+162 Da). Intact APT2 with either or both modifications are observed in the MS. After treatment with the ABP, mass shifts of +374 Da are observed only on active APT2. No multiple labeling was observed on APT2, either.

In conclusion, ESI-MS was successfully used to observe the mass shifts induced by chemical labeling by the ABP on the APTs. The +374 Da mass shifts corresponding to the quinone methide (QM) generated *in situ* were only observed on the active APT1 and APT2, but not the heat-inactivated proteins. More importantly, we only observed single labeling introduced by the ABP, which is consistent with the idea that the ABP labels APTs in an activity dependent manner.

### 3.4. Ruling out non-specific labeling caused by quinone methide diffusion

In order to test the possibility of non-specific labeling induced by diffusion of the quinone methide from active site, the ABP was incubated with cytochrome C in the presence of APT1, either active or inactive. Cytochrome C does not possess any thioesterase or esterase activity, and is used here as a spectator protein. The two proteins were co-incubated in a 1:1 (w/w) ratio with ABP. Only the active APT1 was labeled by the ABP (27093 Da in Figure 4.7); neither cytochrome C nor inactive APT1 was modified.

In this experiment, bovine serum albumin (BSA) from bovine serum was also utilized to spike the solution of APTs. After treatment with ABP, the proteins were subjected to MS analysis. There were no obvious mass shifts observed on BSA. However, tens of peaks were observed in the deconvoluted MS spectrum of BSA alone due to multiple post-translational modifications on the mature BSA protein. These PTMs complicate the identification of the mass peaks and make it difficult to observe any possible +374 Da peaks caused by reaction of BSA with the ABP. Thus, cytochrome C was more suitable for the control experiment with a bystander protein.



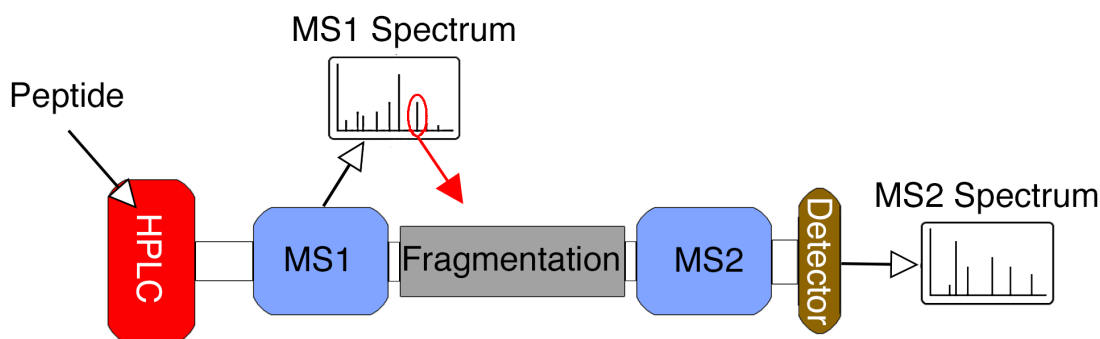
**Figure 4.7 Co-incubation of the ABP with cytochrome C and APT1.** Quinone methide labeling (+374 Da) was observed on active APT1, but not cytochrome C or inactive APT1.

#### 4. Using Tandem Mass Spectrometry to Reveal the Labeling Site



#### 4.1. Introduction to proteomic tandem mass spectrometry

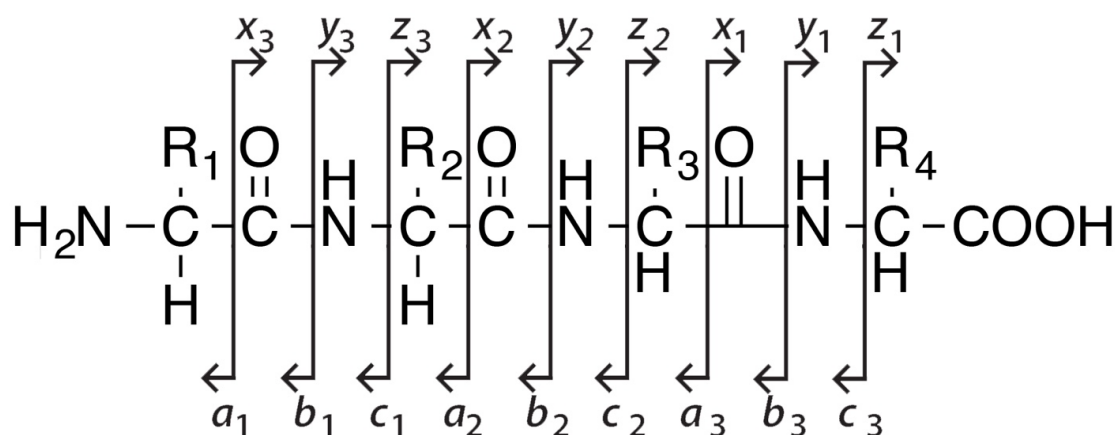
Tandem mass spectrometry, also called MS/MS or MS<sup>2</sup>, involves multiple steps of mass spectrometry filtering, with fragmentation stages for selected ions. In this approach, ions of interest are selected based on the  $m/z$  from the first round of MS and subsequently are fragmented by one of a series of dissociation methods, such as collision-induced dissociation (CID)<sup>18</sup> or higher energy collision dissociation (HCD).<sup>19</sup> Some other methods of ion fragmentation are also available, including electron-transfer dissociation (ETD)<sup>20</sup> and electron-capture dissociation (ECD).<sup>21</sup> The fragments produced from the selected ion are then analyzed in another round of MS to offer the structural and sequential information about the parent ion (Figure 4.8).



**Figure 4.8 Principle of tandem mass spectrometry in peptide sequencing.** The peptide is introduced into the mass spectrometer via the coupled HPLC. The ion of interest (parent ion) is selected in the first round of MS (MS) and transported into fragmentation device. The fragments of the parent ion are then analyzed in the MS<sup>2</sup> to afford sequence information about the peptide parent ion.

MS/MS is commonly used to sequence proteins,<sup>22</sup> as the fragmentation of peptides follows specific patterns and the fragments can be matched with predicted peptide sequences.

Several common peptide backbone fragmentation patterns of a tetrapeptide are presented in Figure 4.9. The product ions are referred to by the letters a, b, c and x, y, z. If the charge is retained on the N-terminal fragment, the ion is named as either a, b or c. If the charge is retained on the C-terminus, the ion type is either x, y or z. To specify the internal cleavage position on the backbone, a subscripted number is added to the ion letter designation. The types of fragment ions observed in MS/MS depend on factors such as peptide sequence, the internal energy applied for fragmentation, and charge state. For example, in the low energy CID fragmentation, b-ion and y-ion are the most common fragment ions observed in MS<sup>2</sup> spectra.



**Figure 4.9 Common fragmentation patterns observed for a tetrapeptide.** The a, b, c-ions are the ions carrying charge on the N-terminus; x, y, z-ions are the ions carrying charge on the C-terminus. The b-ions and y-ions are the most common fragment ions observed in low energy CID MS<sup>2</sup> spectra.

To identify a protein, the whole protein needs to be digested into peptides by proteases, which hydrolyzes the amide bonds of a protein at specific positions. The most widely-used proteases for protein mass spectrometry include trypsin,<sup>23</sup> chymotrypsin, Arg-C, Asp-N, elastase, pepsin, Lys-C, and so on. The resulting peptides are then subjected to tandem MS

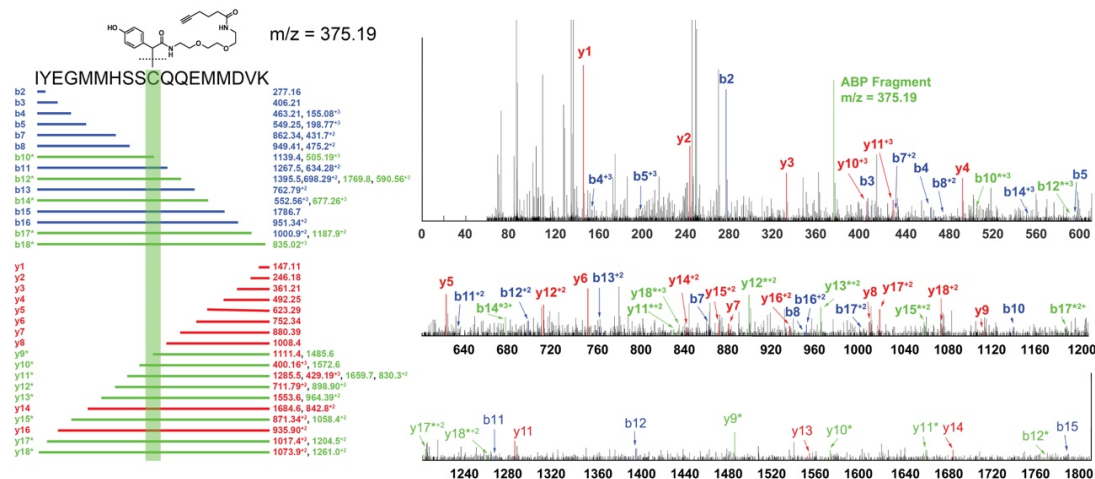
to be sequenced. By matching the fragment ions and the fragment patterns of the peptides to the predicted fragments based on the protein sequence, the protein can be identified.

For analyzing the labeling site of the ABP, the APT1 protein was incubated with the ABP to be labeled, and subsequently cleaved by trypsin, which predominantly cleaves proteins at the carboxyl side of the amino acids lysine and arginine, except when either is bound to a C-terminal proline. The resulting peptides are then injected into ESI-MSI for tandem MS analysis. In MS1 spectrum, the peptide which contains the ABP-labeled amino acid residue will show a mass shift of +374 Da corresponding to the quinone methide. In addition, the MS2 spectrum from the parent peptide ion will also present mass shifts on fragment ions that contain the modification, which will help us to locate the labeling site of the ABP.

#### 4.2. Experiment and result

10 µg APT1 protein was incubated with the ABP according to the protocol described previously. Labeled APT1 protein was re-suspended in 50 µL of 50 mM Tris buffer (pH = 8.0). To this solution was added 0.5 µg of a trypsin/Lys-C mixture (1 µL of a 0.5 µg/µL stock, Promega V5072). The digest was incubated at 37 °C for 16 h, and then terminated by acidification with 1% formic acid. Subsequently, Ziptip C<sub>18</sub> pipette tips (EMD Millipore) were utilized to partially purify the digest. The eluent was injected into the Agilent 6530 Q-TOF mass spectrometer equipped with an HPLC-chip cube for analysis. Tandem MS data were analyzed manually with a mass match tolerance of 40 ppm using ProteoWizard (version 3.0.6002)<sup>24</sup> to identify sites of covalent modification.

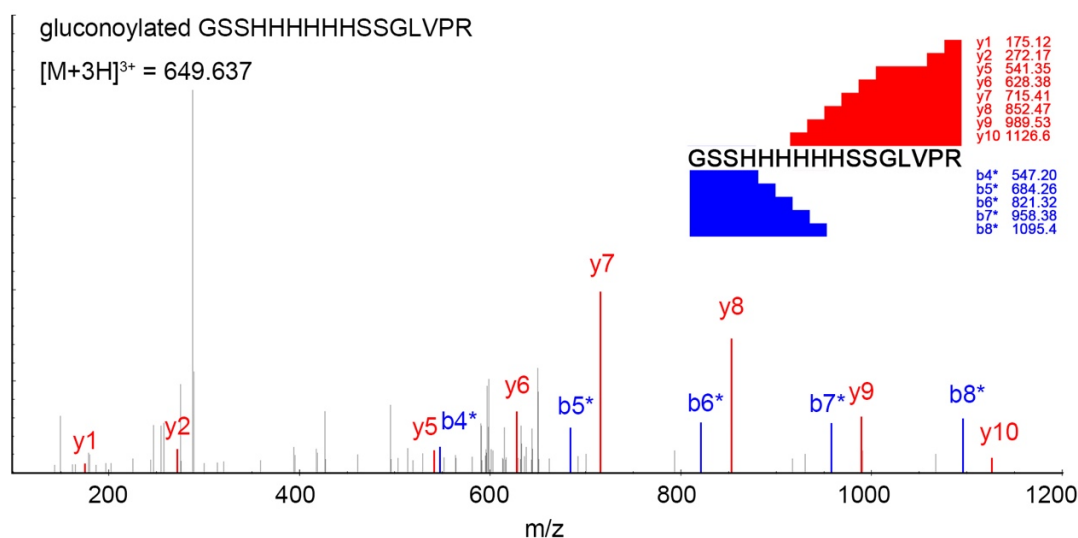
By matching the mass of the precursor and fragment ions, only one labeled peptide was observed. The precursor ion ( $m/z = 841.023$ ,  $M+3H^+$ ) corresponds to the ABP labeled peptide IYEGMMHSSCQQEMMDVK (underline denotes the modified amino acid residue), which contains the His208 catalytic residues. In the tandem mass spectrum, we observed peaks corresponding to a number of labeled and unlabeled peptide fragments which, when considered together, indicate that the site of labeling is Cys211. It is worth mentioning the benzylic cation ( $m/z = 375.19$ , which results from breakage of the probe-peptide bond) was also observed in the tandem spectrum, which further supports the idea that the ABP has labeled this peptide. Results of this analysis indicate that the ABP modified the Cys211 close to the catalytic His208 residue near the active site (Figure 4.10), which is in agreement with the prediction made by our molecular dynamics simulation as previously discussed.



**Figure 4.10** The tandem mass spectrum of the ABP-labeled peptide of APT1. The y-ion and b-ions are labeled in red and blue, respectively. The y/b-ions with the ABP label (+374 Da) are labeled in green. The sequence and mass peaks for each observed y/b-ions are listed in the left panel.

### 4.3. Tandem MS proves *N*-terminal gluconoylation of APT1

While seeking evidence of the ABP labeling site, the tandem MS also lends strong evidence which proves the hypothesis that the *N*-terminal hexahistidine tag of APT1 was the site of gluconoylation (Figure 4.11). A series of b-ions and y-ions produced from the peptide GSSHHHHHHSSGLVPR were matched with the calculated  $m/z$  values with ProteoWizard software. The result indicates that the gluconoylation which introduced the +178 Da mass shift in MS of APT1 and APT2 occurred on the *N*-terminal His-tag.



**Figure 4.11** The tandem mass spectrum of the *N*-terminal peptide from His-tagged APT1. The y-ion and b-ions are labeled in red and blue, respectively. An asterisk denotes gluconoylated (+178 Da) fragments.

## 5. References

1. Griffiths, J., A brief history of mass spectrometry. *Anal. Chem.* **2008**, 80 (15), 5678-5683.

2. (a) Craig, R. D.; Bateman, R. H.; Green, B. N.; Millington, D. S., Mass-Spectrometry Instrumentation for Chemists and Biologists. *Philos. T. R. Soc. A* **1979**, 293 (1400), 135-155; (b) Baker, M., Mass spectrometry for biologists. *Nat. Methods* **2010**, 7 (2), 157-161.
3. (a) Dole, M.; Mack, L. L.; Hines, R. L., Molecular Beams of Macroions. *J. Chem. Phys.* **1968**, 49 (5), 2240-&; (b) Dole, M.; Mack, L. L.; Kralik, P., Molecular Beams of Macroions .2. *B Am Phys Soc* **1969**, 14 (3), 383-&; (c) Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M., Molecular Beams of Macroions .2. *J. Chem. Phys.* **1970**, 52 (10), 4977-&; (d) Clegg, G. A.; Dole, M., Molecular Beams of Macroions .3. Zein and Polyvinylpyrrolidone. *Biopolymers* **1971**, 10 (5), 821-&.
4. (a) Karas, M.; Bachmann, D.; Hillenkamp, F., Influence of the Wavelength in High-Irradiance Ultraviolet-Laser Desorption Mass-Spectrometry of Organic-Molecules. *Anal. Chem.* **1985**, 57 (14), 2935-2939; (b) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T., Protein and polymer analyses up to  $m/z$  100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, 2 (8), 151-153.
5. (a) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray Ionization for Mass-Spectrometry of Large Biomolecules. *Science* **1989**, 246 (4926), 64-71; (b) Ho, C. S.; Lam, C. W. K.; Chan, M. H. M.; Cheung, R. C. K.; Law, L. K.; Lit, L. C. W.; Ng, K. F.; Suen, M. W. M.; Tai, H. L., Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications. *Clin. Biochem. Rev.* **2003**, 24 (1), 3-12; (c) Pitt, J. J., Principles and Applications of Liquid Chromatography-Mass Spectrometry in Clinical Biochemistry. *Clin. Biochem. Rev.* **2009**, 30 (1), 19-34.

6. (a) Witze, E. S.; Old, W. M.; Resing, K. A.; Ahn, N. G., Mapping protein post-translational modifications with mass spectrometry. *Nat. Methods* **2007**, *4* (10), 798-806; (b) Silva, A. M. N.; Vitorino, R.; Domingues, M. R. M.; Spickett, C. M.; Domingues, P., Post-translational Modifications and Mass Spectrometry Detection. *Free Radical Biol. Med.* **2013**, *65*, 925-941.
7. (a) Mann, M., John Bennett Fenn, 1917-2010 Obituary. *J. Am. Soc. Mass Spectr.* **2011**, *22* (3), 602-603; (b) Arpino, P., John Bennett Fenn (1917-2010). *Actual. Chimique* **2011**, (350), 54-54; (c) Muddiman, D. C., John Bennett Fenn (1917-2010). *Science* **2011**, *331* (6014), 160-160.
8. Zeleny, J., The Electrical Discharge from Liquid Points, and a Hydrostatic Method of Measuring the Electric Intensity at Their Surfaces. *Phys. Rev.* **1914**, *3* (2), 69-91.
9. Wilson, C. T. R.; Taylor, G. I., The bursting of soap-bubbles in a uniform electric field. *Math. Proc. Cambridge* **1925**, *22* (05), 728-730.
10. Rayleigh, L., XX. On the equilibrium of liquid conducting masses charged with electricity. *Philos. Mag., Ser. 5* **1882**, *14* (87), 184-186.
11. (a) Bae, Y. J.; Kim, M. S., A Thermal Mechanism of Ion Formation in MALDI. *Annu. Rev. Anal. Chem.* **2015**, *8* (1), 41-60; (b) Molin, L.; Seraglia, R.; Czarnocki, Z.; Maurin, J. K.; Pluciński, F. A.; Traldi, P., On the Primary Ionization Mechanism(s) in Matrix-Assisted Laser Desorption Ionization. *J. Anal. Methods Chem.* **2012**, *2012*, 8.
12. Breuker, K.; Knochenmuss, R.; Zenobi, R., Proton transfer reactions of matrix-assisted laser desorption/ionization matrix monomers and dimers. *J. Am. Soc. Mass Spectr.* **1999**, *10* (11), 1111-1123.

13. Witt, M.; Fuchser, J.; Baykut, G., In-source H/D exchange and ion-molecule reactions using matrix assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry with pulsed collision and reaction gases. *J. Am. Soc. Mass Spectr.* **2002**, *13* (4), 308-317.
14. Knochenmuss, R., Ion formation mechanisms in UV-MALDI. *Analyst* **2006**, *131* (9), 966-986.
15. Tyan, Y. C.; Yang, M. H.; Liao, P. C.; Liao, J. D.; Jong, S. B.; Liu, C. Y.; Wang, M. C.; Grunze, M., MALDI-TOF MS sample preparation by using alkanethiolate self-assembled monolayers: A preliminary application for protein sample analysis. *Int. J. Mass spectrom.* **2007**, *262* (1-2), 67-72.
16. Geoghegan, K. F.; Dixon, H. B. F.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G., Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in *Escherichia coli*: The cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **1999**, *267* (1), 169-184.
17. Frolov, A.; Hoffmann, P.; Hoffmann, R., Fragmentation behavior of glycated peptides derived from D-glucose, D-fructose and D-ribose in tandem mass spectrometry. *J. Mass Spectrom.* **2006**, *41* (11), 1459-1469.
18. Wells, J. M.; McLuckey, S. A., Collision-induced dissociation (CID) of peptides and proteins. *Biol. Mass Spectrom.* **2005**, *402*, 148-185.
19. Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M., Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* **2007**, *4* (9), 709-712.



20. Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *P. Natl. Acad. Sci. U.S.A.* **2004**, *101* (26), 9528-9533.
21. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W., Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265-3266.
22. (a) Frese, C. K.; Altelaar, A. F. M.; van den Toorn, H.; Nolting, D.; Griep-Raming, J.; Heck, A. J. R.; Mohammed, S., Toward Full Peptide Sequence Coverage by Dual Fragmentation Combining Electron-Transfer and Higher-Energy Collision Dissociation Tandem Mass Spectrometry. *Anal. Chem.* **2012**, *84* (22), 9668-9673; (b) Dancik, V.; Addona, T. A.; Clauser, K. R.; Vath, J. E.; Pevzner, P. A., De novo peptide sequencing via tandem mass spectrometry. *J. Comput. Biol.* **1999**, *6* (3-4), 327-342.
23. Olsen, J. V.; Ong, S. E.; Mann, M., Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **2004**, *3* (6), 608-614.



## **CHAPTER FIVE**

### **Using the ABP to Profile Thioesterase Activity in Cells**

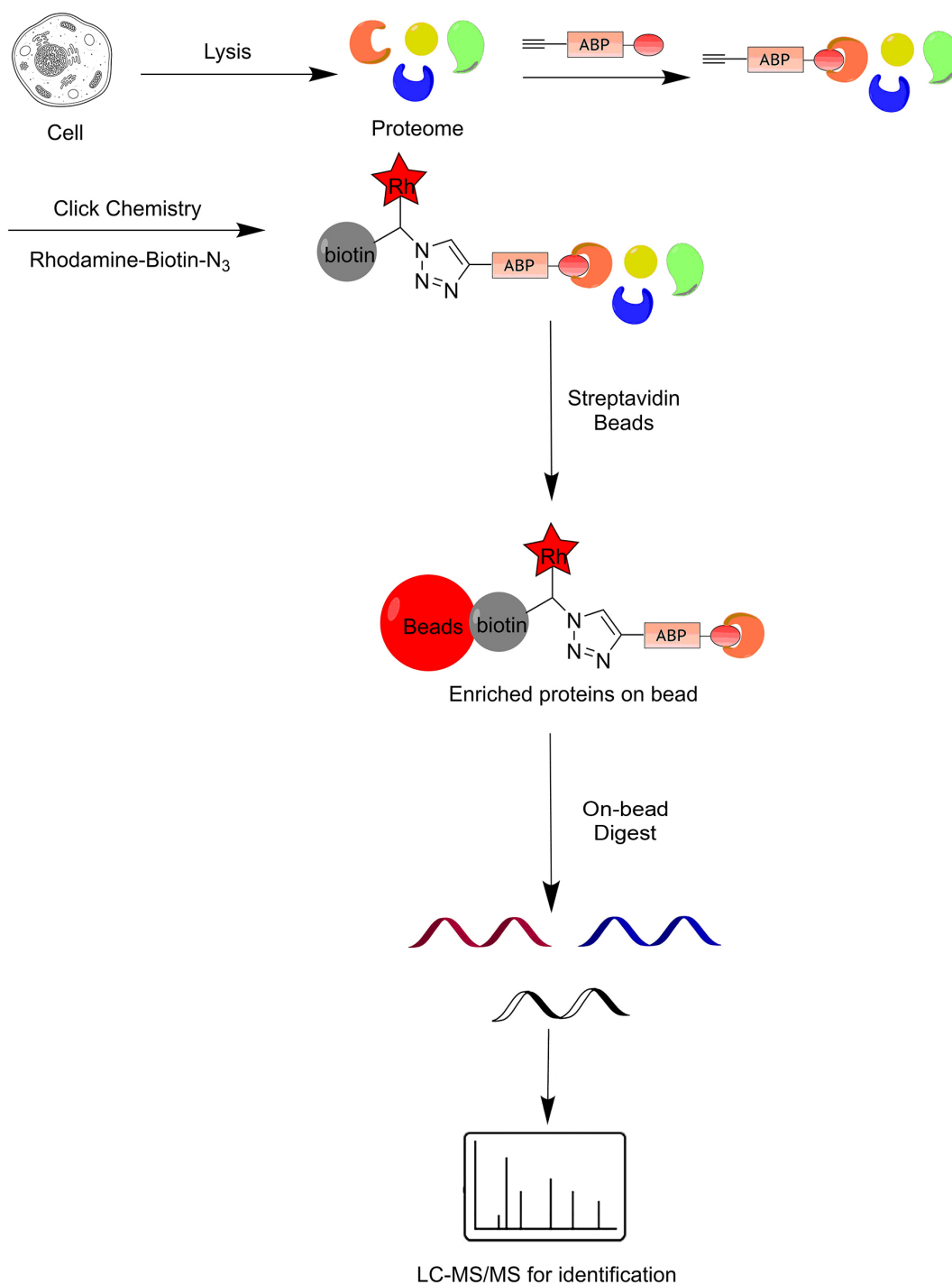


## 1. Abstract

To determine if the ABP can target APTs in a cellular environment, and to identify additional potential target proteins, activity-based proteome profiling (ABPP) experiments were performed using ABP pull down assays. In ABPP experiments, the ABP covalently labels a nucleophilic residue of a target enzyme, thereby allowing the enrichment of target proteins with streptavidin beads via affinity purification. The enriched target proteins are then digested and analyzed by proteomic tandem MS to evaluate the ABP-labeled proteins from the cell.

## 2. Experiment Design

Figure 5.1 describes the chemical proteomic strategy employed for cell lysate labeling experiments. Cells were first lysed by sonication, and then the resulting lysate was clarified by centrifugation. The soluble fraction of the lysate was incubated with ABP allowing the covalent labeling of the target proteins. Subsequently, labeled proteins were tagged through a copper (I) catalyzed alkyne azide cycloaddition (CuAAC) reaction with a rhodamine-biotin-azide bifunctional azide tag, which allowed enrichment by exploiting the biotin/streptavidin affinity. After the affinity enrichment, the APT-labeled target proteins were subjected either directly to on-bead digestion, or to SDS-PAGE separation and in-gel digestion. The resulting peptides were analyzed by proteomic tandem MS, which allowed for their identifications.



**Figure 5.1 Experiment design of the ABPP labeling experiments.** After lysis, the cell lysate is incubated with the ABP and target proteins are subsequently tagged by click chemistry with the rhodamine-biotin-N<sub>3</sub> tag. After affinity purification with streptavidin beads, the target proteins are subjected to SDS-PAGE and then in-gel digestion, or else on-bead digestion. All the resulting peptides from the target proteins are then identified via tandem mass spectrometry.

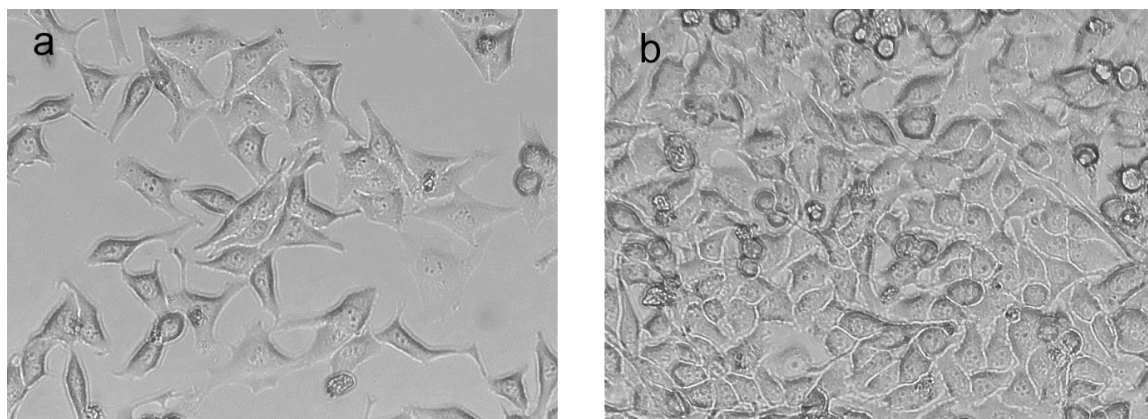
### 3. HeLa Cell Culture

#### 3.1. Cell counting

The number of cells was evaluated using a hemocytometer. To reduce the error, the cell suspension was first diluted with trypan blue by a factor 5 (1:4 v/v). The trypan blue distinguishes living cells from dead cells, as only dead cells are stained blue by trypan blue dye. The diluted cell suspension (10  $\mu$ L) was injected into the counting chamber of the hemocytometer and cells located in the counting grid (consisting of four large outer squares and one central square) were counted. The total volume of the cell suspension counted (in 5 large squares) was  $5 \times 10^{-4}$  mL. Thus, the number of cells contained in 1 mL of the original cell suspension can subsequently be calculated by multiplying this number by a factor of  $10^4$ .

#### 3.2. HeLa cell line

The HeLa cell line is the oldest and most widely used human cell line in scientific research. It was derived from cervical cancer cells taken from Henrietta Lacks during a biopsy,<sup>1</sup> a patient who eventually died of her cancer on October 4, 1951. However, the cells, named HeLa, continued to grow and divide as long as they are cultured under suitable condition. Designated as being “immortal”, this cell line has become a significant tool in biomedical research. The HeLa cell line is famous for its proliferative capacity *in vitro*. The cells multiply every 23 hours and are easy to culture in standard medium. Microscopic images of HeLa cells at low or high density are presented in Figure 5.2.



**Figure 5.2 Microscopic images of HeLa cells at different densities.** The morphologies of the cells are slightly different.

### 3.3. Instruments and materials

Before any operations, the laminar flow hood and gloves were always sprayed with 70% ethanol. DMEM + 10% FBS medium, PBS buffer and 0.05% trypsin were preheated to 37 °C in water bath. The frozen cells were thawed in a 37 °C water bath upon arrival, and transferred to 5 mL of culture medium immediately. The cell suspension was centrifuged at 800 RPM for 5 minutes to pellet the cells, and then washed again with 5 mL of PBS buffer. The cell pellet was then resuspended in 3 mL of culture medium and incubated in a T25 culture flask within the cell incubator for 10 hours. Subsequently, the cell growth medium containing a trace amount of DMSO was carefully discarded and new medium was added. The cells were then incubated in a culture flask until reaching approximately 80% confluence.

Once the cell culture reached 80% confluence, the cells were harvested and passaged. To detach the HeLa cells from the culture flasks, the culture medium was first removed and the cell layers were washed carefully with 3 mL of PBS buffer. After the removal of the



PBS buffer, the cell layers were digested in 3 mL 0.05% trypsin at 37 °C for 3 minutes. The flasks were gently tapped to help the cells to detach. Before adding 2 mL culture medium into the cell suspension to quench the digestion, the flasks were checked under a microscope to ensure that all the cells were detached. The entire cell suspension was then centrifuged at 800 RPM for 5 minutes, and the cell pellet was resuspended in medium and counted with a hemocytometer. Based on the cell suspension density, the cell suspension was split into 3 or 4 portions into new culture flasks for subculturing.

A portion of cell culture was preserved in liquid N<sub>2</sub> when every 2 generations were passaged. To freeze the cells, the cells were first detached and washed according to the protocol described previously. The cell pellet was then resuspended in the medium with 5% DMSO in a 2.0 mL cryogenic tube and frozen in Nalgene Mr. Frosty at -80 °C for 4 hours before transferred to the vapor phase of the liquid N<sub>2</sub> preservation tank.

#### 4. Cell Proteome Labeling Experiment

##### 4.1. Cell lysis procedure

The cells were harvested and counted following the procedure described in section 3.4. The cell pellet containing about 3 million cells was resuspended in 3 mL of Tris buffer (50 mM, pH = 8.0) and lysed by sonication (30% power, 3 × 10 s at 10 s interval) at 0 °C. The cell lysate was then centrifuged at 4000 × g for 1 hour at 4 °C to pellet the insoluble fraction.<sup>1</sup>

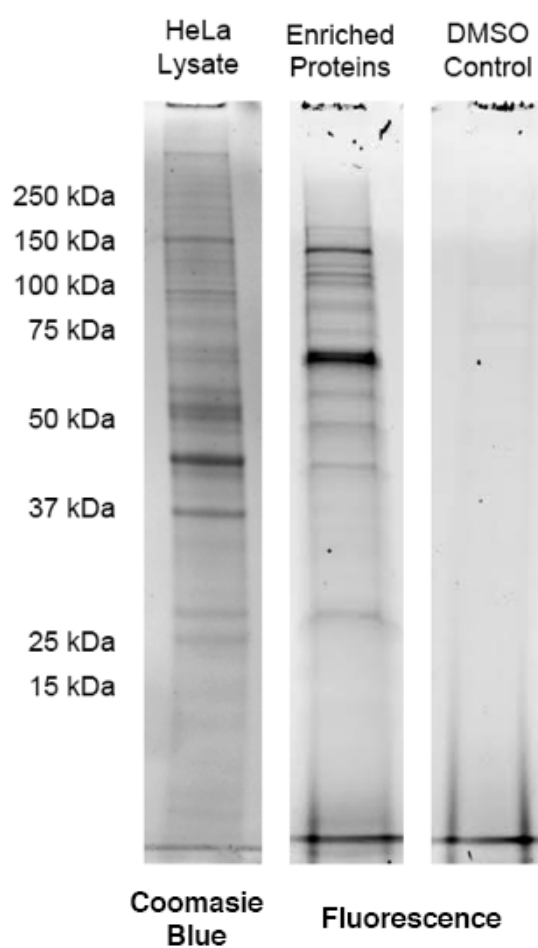
##### 4.2. Proteome labeling and enrichment experiment

Total protein concentration of the clarified lysate was determined using the bicinchoninic acid (BCA) assay. The absorbance of the copper complex was measured at 562 nm. The lysate was then diluted to a protein concentration of 2 mg/mL with 50 mM Tris buffer. ABP was then added (as a 1000× stock solution (100 mM) in DMSO) to 1 mL of lysate to give a final concentration of 100  $\mu$ M ABP. The lysate was allowed to incubated with ABP at room temperature for 16 hours before being washed 3 × with 400  $\mu$ L of PBS buffer in a 3 kDa molecular weight cutoff (MWCO) filter. The ABP-labeled proteome was then resuspended in 100  $\mu$ L PBS buffer (pH = 8.0), to which was added 3  $\mu$ L CuSO<sub>4</sub> aqueous solution (11 mM), 3  $\mu$ L sodium ascorbate aqueous solution (110 mM) and 3  $\mu$ L TBTA solution in DMSO (1.1 mM) following standard click chemistry conditions. The rhodamine-biotin-N<sub>3</sub> bifunctional tag (Click Chemistry Tools, 1048) solution in DMSO (3  $\mu$ L, 3.75 mM) was then added to initiate the click reaction. The reaction was allowed to proceed at room temperature for 4 hours before partial enrichment of the proteome using a 3 kDa MWCO filter.

To 100  $\mu$ L of the bifunctionally-tagged proteome solution was added 100  $\mu$ L streptavidin agarose resin slurry (Thermo Scientific, 20347) and the mixture was shaken on an orbital shaker at room temperature for 2 hours at 900 RPM. The resin beads were subsequently washed with PBS buffer containing 0.2 % SDS (0.5 mL) three times, twice with 6 M urea aqueous solution (0.5 mL), and then three times with PBS buffer (0.5 mL). The target proteins enriched by the streptavidin bead were subjected to either SDS-PAGE or on-bead tryptic digestion for analysis.

#### 4.3. In-gel fluorescence analysis of labeled proteome

After the affinity enrichment on streptavidin-beads, the fluorescently-tagged cellular proteome was resolved using SDS-PAGE. The resulting gel was scanned using an in gel imaging instrument to detect fluorescent protein bands (Figure 5.3). In Lane 1, enriched lysate incubated with ABP was loaded. As a negative control, the lysate that was treated with DMSO alone (no ABP) was loaded. As a result, several fluorescent bands were detected in the Lane 1, but there was no fluorescence visible in Lane 2.



**Figure 5.3 In-gel fluorescence analysis of the ABP-labeled cellular proteome.** Several fluorescent bands were detected in the Lane 2, but no visible fluorescence was observed in Lane 3.

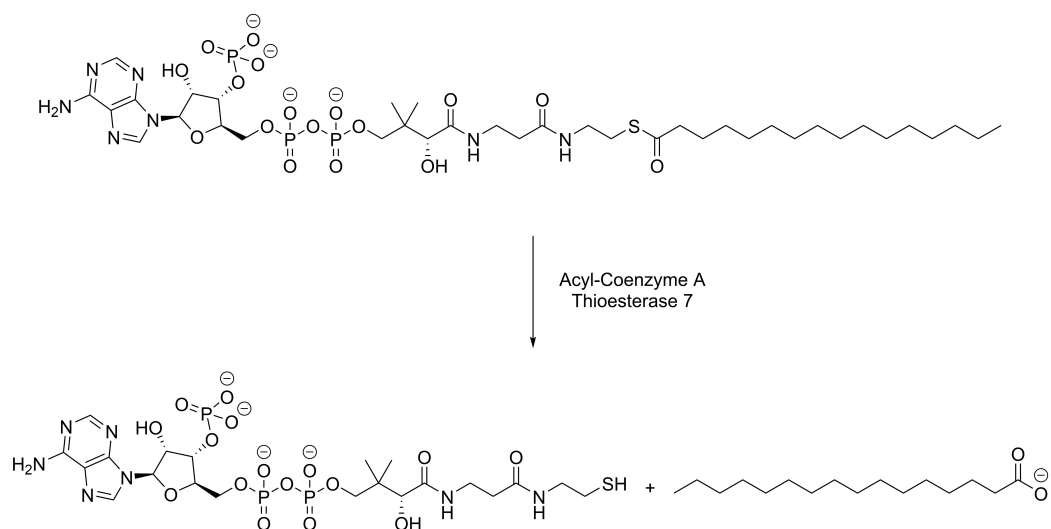
#### 4.4. On-bead digestion and mass spectrometry analysis

Target proteins enriched using the streptavidin beads were resuspended in 100  $\mu$ L Tris buffer (50 mM, pH = 8.0) containing 6 M urea. 10  $\mu$ g of a trypsin/Lys-C mixture in 20  $\mu$ L resuspension buffer was added, and the mixture was incubated at 37 °C for 3 hours. 700  $\mu$ L of Tris buffer (50 mM, pH = 8.0) was added to the digestion mixture to dilute the urea. The reaction was then allowed to continue incubating at 37 °C for 10 more hours.<sup>3</sup>

At the end of the digestion period, 5  $\mu$ L of formic acid was added to terminate the digestion. Streptavidin beads were removed using a spin filter and the filtrate was concentrated in a vacuum centrifuge at 60 °C. The resulting peptides in the urea solution was diluted by a factor 4 (1:3) with 20% acetonitrile in aqueous solution containing 2% trifluoroacetic acid (TFA). The resulting solution containing urea and Tris was desalted on a C<sub>18</sub> solid phase extraction column (Pierce, 89870) following a standard procedure. After washing the column with 5% acetonitrile containing 0.5% TFA, peptides bound to the C<sub>18</sub> resin were eluted with 40  $\mu$ L of 70% acetonitrile containing 0.1 % formic acid. The enriched peptides were concentrated by vacuum centrifugation and reconstituted in HPLC water. The concentration of the peptides in this solution was determined by protein A280 assay on a NanoDrop 2000 UV-Vis spectrometer. 10  $\mu$ g of peptides from the target proteins in 10  $\mu$ L HPLC water was submitted to the COBRE Center for Cancer Research Development at Rhode Island Hospital for proteomic MS analysis.

The peptide sequence data obtained were of very high quality with a 1% false discovery rate (FDR) confidence level for the sequence assignments. The proteins returned as results were imported into the MySQL database (Version 5.7.8.0) for analysis. All the identified proteins with at least two unique peptides scoring above 40 are presented in Appendix I.

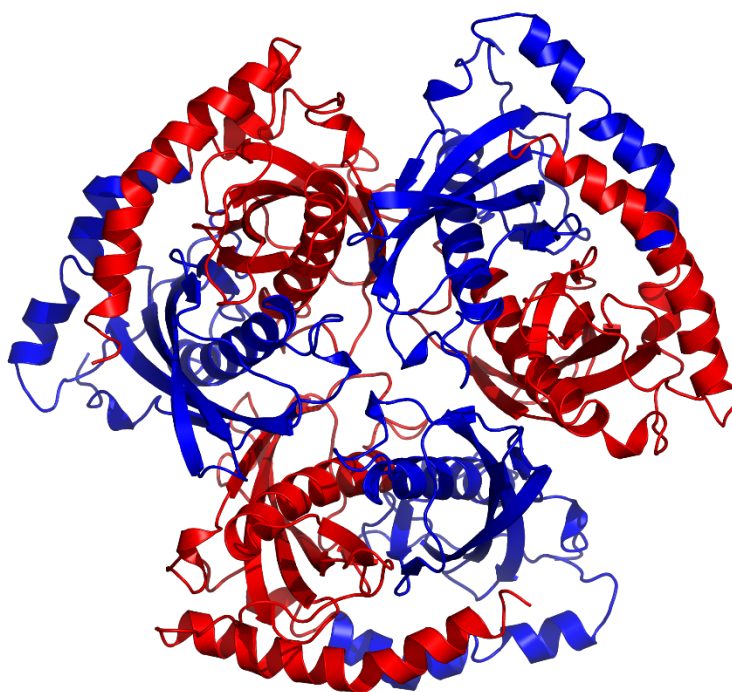
In this list of the target proteins, thioesterases and esterases were identified as being labeled by the ABP. This finding supports the idea that the ABP labels proteins with similar activities in accordance with its proposed mechanism of action. One of the thioesterases identified in this analysis was cytosolic acyl coenzyme A thioester hydrolase (acyl-CoA thioesterase 7 or ACOT7), which is a 44 kDa protein that hydrolyzes the palmitoyl coenzyme A thioester to palmitic acid and coenzyme A (Scheme 5.1).<sup>4</sup> ACOT7 protein plays an important role in a series of biological pathways.<sup>5</sup> For example, in human brain cytosol, ACOT7 accounts for 75% of the palmitoyl-CoA hydrolase activity found in the brain tissue.<sup>6</sup>



**Scheme 5.1 Acyl-Coenzyme A thioesterase 7 (ACOT7) catalyzes the hydrolysis of palmitoyl Coenzyme A thioester.**

The crystal structures of both the N- and C-terminal domains of the mouse ACOT7 have been investigated.<sup>2</sup> The quaternary arrangement in ACOT7 features a trimer of hotdog fold

dimers (Figure 5.4). Both N- and C-domains of are required for the ACOT7 thioesterase activity. Asn-24 and Asp-213 (from the N- and C-domains, respectively) were identified as essential residues for the thioesterase activity through site directed mutagenesis. However, the hydrolysis mechanism of ACOT7 is still unclear.

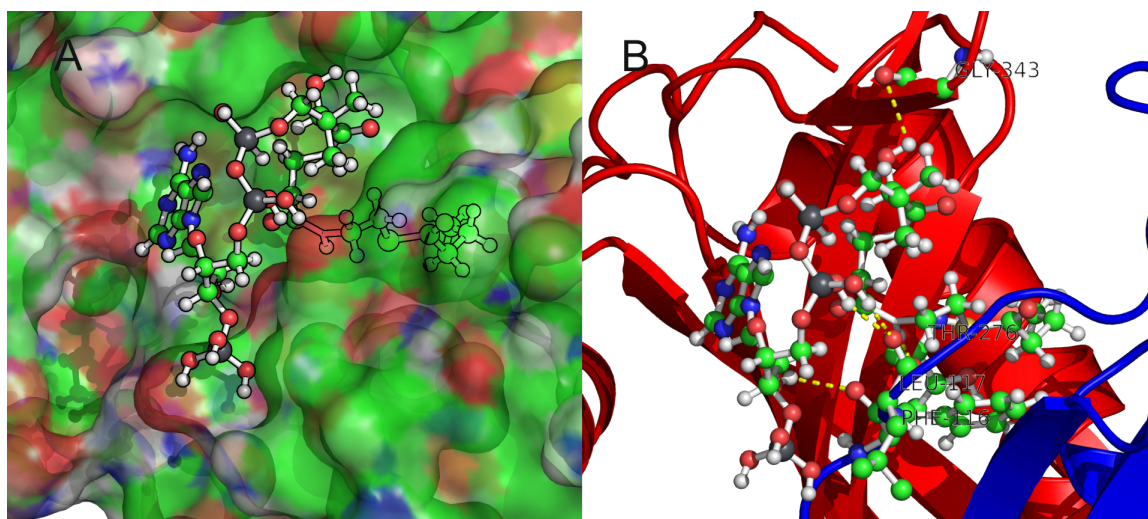


**Figure 5.4 Quaternary structure of ACOT7.** The *N*-terminal domains are presented in blue and C-terminal domains are red. PDB code: 4ZV3.

To elucidate the binding of the substrate to the ACOT7 active site, molecular dynamics (MD) simulation was employed. The butyryl Coenzyme A was first docked into the active site of ACOT7 with AutoDock Vina (Version 1.1.2),<sup>3</sup> the complex obtained from docking was used as the initial structure for the MD simulation in Gromacs (Version 5.0.2).<sup>4</sup>

In the simulation, the GROMOS96 43a1 force field and SPC water model were used. The topology of the ligand was generated by PRODRG 2.5.<sup>5</sup> First, the energy of the complex was minimized by using the steepest descent minimization algorithm in 50000 steps. Second, a 100 ps position-restrained dynamics was performed under NVT ensemble to relax the water molecules. Third, another 100 ps position-restrained dynamics was performed under NPT ensemble. Finally, a full 1000 ps dynamics run was performed. All the calculations were carried out in the Center for Computation and Visualization at Brown University.

Figure 5.5a and 5.5b illustrate how Coenzyme A binds to the active site of the ACOT7. The phosphate and thiol groups of the Coenzyme A forms hydrogen bonds with several amino acid residues in the binding pocket of the ACOT7, such as Phe-116, Phe-117 on the *N*-terminal domain, and Thr-276 and Gly-343 on the *C*-terminal domain. However, the catalytic activity of the ACOT7 still cannot be explained by this binding model.



**Figure 5.5 Binding of the Coenzyme A within the active site of ACOT7.** A. Coenzyme A in the binding pocket of the ACOT7. B. Coenzyme A forms a series of hydrogen bonds with amino acid residues on *N*-terminal domain (blue) and *C*-terminal domain (red).

Besides the ACOT7, a few of other enzymes with thioesterase activity were labeled and identified, including the ubiquitin carboxyl-terminal hydrolases 5/7/14/47 (USP5, USP7, USP14, and USP47, respectively), acyl-coenzyme A thioesterase 9 (ACOT9), *S*-formylglutathione hydrolase (ESD), ubiquitin thiolesterase L5 (UCHL5), ubiquitin carboxyl terminal esterase L3 (UCHL3), ubiquitin thioesterase (OTUB1). The results are listed in Table 5.1.

| Protein Name                             | Peptides Found                                                                                                                                            | Score                                              | Accession |
|------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|-----------|
| Ubiquitin carboxyl-terminal hydrolase 5  | K.IFQNAPTDPTQDFSTQVAK.L<br>K.LDVSIEMPEELDISQLR.G<br>R.DRVTSAVEALLSADSASR.K<br>K.LGHGLLSGEYSKPVPESGDGERVPEQK.E<br>R.YFDGSGGNNHAVEHYR.E<br>K.IVILPDYLEIAR.D | 90.93<br>45.27<br>36.25<br>25.41<br>20.71<br>20.45 | P45974    |
| Ubiquitin carboxyl-terminal hydrolase 7  | K.NSSLAEFVQSLSQTMGFQDQIR.L<br>K.LSEVLQAVTDHDIPQQLVER.L<br>K.EVFGTFGIPFLLR.I<br>R.LNTDPMQLQFFK.S                                                           | 88.88<br>74.10<br>39.67<br>20.78                   | B7Z815    |
| Ubiquitin carboxyl-terminal hydrolase 14 | K.FEGVELNTDEPPMVFK.A<br>R.LPAYLTIQMVR.F<br>R.SSSSGHYVSWVK.R                                                                                               | 68.89<br>38.29<br>17.26                            | A6NJA2    |
| Ubiquitin carboxyl-terminal hydrolase 47 | K.QHLEPFVGVLSHFV.K<br>R.ELEQHIQTSDPENFQSEER.S<br>K.LDPFQEVVLESSSVDEL.R                                                                                    | 30.37<br>28.83<br>14.91                            | E9PM46    |
| Ubiquitin carboxyl-terminal hydrolase L3 | R.WLPLEANPEVTNQFLK.Q<br>R.VTHETSAHEGQTEAPSIDEK.V                                                                                                          | 59.79<br>43.88                                     | P15374    |
| Ubiquitin thiolesterase L5               | R.FNLMAIVSDRK.M<br>R.LDTIFFAK.Q<br>K.TLAHQQLIPLVEK.A                                                                                                      | 34.16<br>20.66<br>16.41                            | Q5LJA5    |
| Ubiquitin thioesterase OTUB1             | K.EYAEDDNIYQQK.I<br>R.LELSVLYK.E<br>R.GEGGTTNPHIFPEGSEPK.V                                                                                                | 41.98<br>18.48<br>17.21                            | J3KR44    |
| <i>S</i> -formylglutathione hydrolase    | R.MYSYVTEELPQLINANFPVDPQR.M<br>K.SYPGSQLDILIDQK.D<br>K.SGYHQSAHEGLVVIAPDTSPR.G                                                                            | 88.66<br>59.83<br>35.16                            | P10768    |
| Acyl-coenzyme A thioesterase 9           | K.MSPLSIVTALVDK.I                                                                                                                                         | 23.83                                              | Q9Y305    |

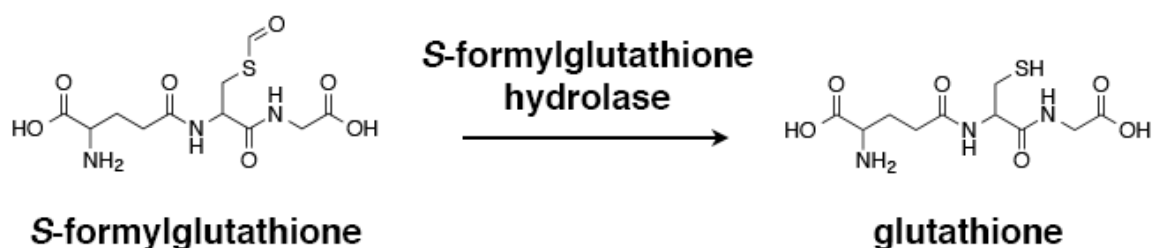
**Table 5.1 The enzymes with known thioesterase activity which were identified as being labeled by the ABP by tandem mass spectrometry.**

Ubiquitin carboxyl-terminal hydrolases can cleave a wide variety of products from the C-terminal glycine residue of ubiquitin. This class of enzymes were first identified as



thioesterases since they can hydrolyze dithiothreitol thioester of ubiquitin, and they were thought to also hydrolyze ubiquitin-glutathione and other ubiquitin thioesters.<sup>6</sup> Later investigations shown that ubiquitin carboxyl-terminal hydrolases hydrolyze amides and other groups from ubiquitin.<sup>7</sup> These enzymes belong to a larger class of peptidases called deubiquitinating enzymes that hydrolyze lysine-glycine amide bonds in ubiquitinated proteins.<sup>8</sup>

Another class of small molecule thioesterase, *S*-formylglutathione hydrolases, which hydrolyze *S*-formylglutathione into formate and glutathione (Scheme 5.2),<sup>9</sup> are also discovered in proteomic mass spectrometry results. The *S*-formylglutathione is responsible for most of the hydrolysis of *S*-formylglutathione in human liver, which is in charge of formaldehyde detoxification.<sup>10</sup>



**Scheme 5.2** The hydrolysis reaction of *S*-formylglutathione catalyzed by *S*-formylglutathione hydrolases.

In conclusion, the ABP was successfully applied to a mammalian cell lysate to profile thioesterase activities. The concept that the ABP labels proteins in a mechanism-based fashion based on their catalytic activities- and not necessarily their protein structure - was demonstrated. A series of thioesterases were labeled and identified via proteomic mass

spectrometry, and the results obtained included most of the common thioesterase families. This information will be helpful to identify other potential thioesterases that can process palmyoylated Ras as a substrate. In addition, the ABP can also be used as a tool to investigate potential off-target proteins bind to thioesterase inhibitors.

## 5. References

1. Rusch, M.; Zimmermann, T. J.; Burger, M.; Dekker, F. J.; Gormer, K.; Triola, G.; Brockmeyer, A.; Janning, P.; Bottcher, T.; Sieber, S. A.; Vetter, I. R.; Hedberg, C.; Waldmann, H., Identification of Acyl Protein Thioesterases 1 and 2 as the Cellular Targets of the Ras-Signaling Modulators Palmostatin B and M. *Angew. Chem. Int. Ed.* **2011**, *50* (42), 9838-9842.
2. Forwood, J. K.; Thakur, A. S.; Guncar, G.; Marfori, M.; Mouradov, D.; Meng, W.; Robinson, J.; Huber, T.; Kellie, S.; Martin, J. L.; Hume, D. A.; Kobe, B., Structural basis for recruitment of tandem hotdog domains in acyl-CoA thioesterase 7 and its role in inflammation. *P. Natl. Acad. Sci. U.S.A.* **2007**, *104* (25), 10382-10387.
3. Trott, O.; Olson, A. J., Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455-461.
4. Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R., Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1-3), 43-56.
5. Schuttelkopf, A. W.; van Aalten, D. M. F., PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D* **2004**, *60*, 1355-1363.

6. Rose, I. A.; Warms, J. V. B., An Enzyme with Ubiquitin Carboxy-Terminal Esterase-Activity from Reticulocytes. *Biochemistry-Us* **1983**, 22 (18), 4234-4237.
7. Pickart, C. M.; Rose, I. A., Ubiquitin Carboxyl-Terminal Hydrolase Acts on Ubiquitin Carboxyl-Terminal Amides. *J. Biol. Chem.* **1985**, 260 (13), 7903-7910.
8. Amerik, A. Y.; Hochstrasser, M., Mechanism and function of deubiquitinating enzymes. *Bba-Mol. Cell Res.* **2004**, 1695 (1-3), 189-207.
9. Uotila, L.; Koivusal.M, Purification and Properties of S-Formylglutathione Hydrolase from Human Liver. *J. Biol. Chem.* **1974**, 249 (23), 7664-7672.
10. Harms, N.; Ras, J.; Reijnders, W. N. M.; vanSpanning, R. J. M.; Stouthamer, A. H., S-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: A universal pathway for formaldehyde detoxification? *J. Bacteriol.* **1996**, 178 (21), 6296-6299.



# **APPENDIX I**

## **Identified Cellular Proteins Labeled by ABP**



**Table S1 The proteins from Mascot search with at least two unique peptides scoring above 40.0.**

| Protein Name                                            | Peptide Sequence Found                                                                                                                                | Score                                             | Accession |
|---------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|-----------|
| 14-3-3 epsilon                                          | K.VAGMDVELTVEER.N<br>R.IISSIEQKEENK.G                                                                                                                 | 71.76<br>49.5                                     | B4DJF2    |
| 14-3-3 Eta                                              | K.QAFDDAIAELDTLNEDSYK.D<br>K.AVTELNEPLSNEDR.N<br>K.DSTLIMQLLR.D                                                                                       | 84.08<br>59.2<br>52.8                             | Q04917    |
| Aconitase, soluble                                      | K.IETVNESWNLATPSDK.L<br>R.SNLVGMGVIPLEYLPGENADALGLTGQER.Y<br>R.ANYLASPPVIAIYAIAGTIR.I<br>K.QAPQTIHLPSGEILDVFDAAER.Y<br>K.SPPFFENLTLDLQPPK.S           | 81.63<br>78.6<br>63.14<br>62.46<br>49.22          | Q9HBB2    |
| Actin related protein 1A                                | K.DQLQTFSEEHPVLLTEAPLNPR.K<br>K.AQYYLPDGSTIEIGPSR.F<br>R.TLFSNIVLSGGSTLFK.G                                                                           | 46.66<br>46.04<br>42.1                            | P61163    |
| Actin-like 6A                                           | R.STGLILDSGATHHTTAIPVHDGYVLQQGIVK.S<br>K.NGMVEDWDSFQAILDHTYK.M                                                                                        | 72.13<br>48.1                                     | O96019    |
| Acyl-CoA thioesterase 7                                 | K.TNIVTASVDAINFHDK.I<br>R.IMRPDDANVAGNVHGGTILK.M                                                                                                      | 74.2<br>41.28                                     | O00154    |
| Acylpeptide hydrolase                                   | K.ALDVSASDDEIAR.L<br>R.QVLLSEPEEAAALYR.G                                                                                                              | 60.52<br>46.69                                    | P13798    |
| Adenylosuccinate synthase                               | R.VGIGAFPTEQDNEIGELLQTR.G<br>R.DGVYFLYEALHGPPK.K                                                                                                      | 115.42<br>40.55                                   | P30520    |
| ADP ribosylation factor 1                               | K.LGEIVTTIPTIGFNVETVEYK.N<br>K.QDLPNAMAISEMTDK.L                                                                                                      | 70.14<br>69.62                                    | P18085    |
| Aldehyde dehydrogenase 5                                | K.VAFTGSTEVGHLLQK.A<br>K.VAEQTPLSALYLASLIK.E<br>K.VGNPFELDTQQGPQVDKEQFER.V                                                                            | 64.37<br>55.36<br>44.89                           | P30837    |
| Aldehyde dehydrogenase 7 family, member A1              | R.EENEGVYNGSWGGR.G<br>R.QASVADYEETVK.K                                                                                                                | 62.83<br>41.14                                    | P49419    |
| Aldo keto reductase family 1, member C3                 | K.LNDGHFMPVLGFGTYAPPEVPR.S<br>R.HIDSAHLYNNEEQVGLAIR.S                                                                                                 | 84.85<br>47.58                                    | P42330    |
| Annexin IV                                              | R.ISQTYQQYGR.S<br>K.SETSGSFEDALLAIVK.C                                                                                                                | 55.81<br>53.9                                     | P09525    |
| Basic leucine zipper and W2 domain-containing protein 1 | K.QYSPLLAFFTQGGSELTLKK.I<br>K.KEELVAEQAIIK.H                                                                                                          | 58.05<br>57                                       | Q7L1Q6    |
| C14orf166 protein                                       | K.NAEPLINLDVNNPDFK.A<br>K.AGVMALANLLQQR.H<br>K.HILGFDTGDAVLNEAAQILR.L                                                                                 | 74.94<br>61.68<br>49.91                           | Q9Y224    |
| Capping protein, muscle Z line, alpha 1                 | K.IIENAENEYQTAISENYQTMSDTTFK.A<br>K.IEGYEDQVLITEHGDLSNR.F<br>K.DVQDSLTVSNEAQTAKE<br>K.FITHAPPGEFNEVFNDVR.L<br>K.EASDPQPEEADGGLK.S                     | 109.24<br>68.54<br>65.73<br>58.52<br>46.24        | P52907    |
| Carbamoyl phosphate synthetase I                        | K.TVLMNPNIASVQTNEVGLK.Q<br>K.EPLFGISTGNLITGLAAGAK.T<br>K.AQTAHIVLEDGTMK.M<br>K.MEYDGILIAGGPGNPALAEPLIQNVR.K<br>K.VSGLLVLDYSK.D<br>K.IEFEGQPVDVFDPNK.Q | 109.34<br>92.3<br>60.77<br>58.7<br>54.91<br>41.03 | Q5R211    |
| Carbonyl reductase 1                                    | K.EYGGLDVLVNNAGIAFK.V<br>K.SPEEGAETPVYLLALLPPDAEGPHGQFVSEK.R<br>K.FRSETITEEELVGLMKNK.F                                                                | 99.18<br>56.29<br>45.73                           | P16152    |
| CBP80                                                   | K.LQPGSLPQVLAQATEMLYMR.L                                                                                                                              | 81.41                                             | Q09161    |

|                                                          |                                                                                                                                         |                                                    |        |
|----------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|--------|
|                                                          | R.MFDYTDDPEGPVMPGSHSVER.F<br>R.NYNFGGEFVEAMIR.Q                                                                                         | 69.2<br>50.01                                      |        |
| CDC37                                                    | R.LGPGGLDPVEVYESLPEELQK.C<br>R.QYMEGFNDELEAFK.E                                                                                         | 70.21<br>51.75                                     | Q16543 |
| CGI-128 protein                                          | R.SGERPVTAGEEDEQVPDSIDAR.E<br>R.SINDPEHPLTLEELNVVEQVR.V                                                                                 | 52.45<br>45.43                                     | Q9Y3D0 |
| Chaperonin containing T-complex polypeptide 1 subunit 6A | K.VLAQNSGFDLQETLVK.I<br>K.VHAELADVLTEAVVDSILAIAK.K                                                                                      | 135.77<br>113.48                                   | P40227 |
| Chromosome segregation 1like protein                     | K.FLESVEGNQNYPLLLTLLEK.S<br>K.ANIVHMLMLSSPEIQK.Q<br>R.FQSGDFHVINGVLR.T<br>K.ALTLPGSSENEYIMK.A<br>K.LVLDAFALPLTNLFK.A<br>K.QLSDAISIIGR.E | 118.78<br>78.8<br>73.52<br>54.96<br>54.87<br>43.16 | P55060 |
| Cisplatin resistance associated overexpressed protein    | R.STTSTIESFAAQEK.Q<br>K.IDVLLQQIEELGSEGK.V                                                                                              | 83.76<br>63.99                                     | O95232 |
| Clathrin, heavy polypeptide                              | K.FDVNTSAVQVLIHIGNLDR.A<br>K.LHIEVGTPPTGNQFPK.K<br>K.IYIDSNNNPER.F<br>R.ISGETIFVTAPHEATAGIIGVNR.K                                       | 93.05<br>92.15<br>73.89<br>65.42                   | Q00610 |
| Clusterin                                                | K.LFSDSPITVTVPEVSR.K<br>R.ASSIIDELFQDR.F                                                                                                | 88.51<br>64.75                                     | E7ETA7 |
| Coatomer protein complex, subunit beta 2                 | K.AAESLADPTEYENLFPGLK.E<br>K.VWQLGSSSPNFTLEGHEK.G                                                                                       | 68.65<br>54.95                                     | P35606 |
| Cofilin 2                                                | K.QILVGDIGDTVEDPYTSFVK.L<br>K.HEWQVNLDDIKDR.S                                                                                           | 117.7<br>54.56                                     | Q9Y281 |
| Collapsin response mediator protein 2                    | K.GTVVYGEPIASLTGDSHYWSK.N<br>R.NLHQSGFSLSGAQIDNIPR.R                                                                                    | 58.74<br>49.25                                     | Q16555 |
| Collapsin response mediator protein 4                    | R.ISVGSDDLVIWDPDAVK.I<br>R.NLHQSGFSLSGTQVDEGVR.S<br>K.DNFTAIEGTNGVEER.M                                                                 | 107.62<br>70.41<br>49.04                           | Q14195 |
| COP9 signalosome subunit 4                               | R.LYLEDDDPVQAEAYINR.A<br>K.LYNNITFEELGALLEIPAAK.A                                                                                       | 100.45<br>42.92                                    | Q9BT78 |
| Culin 4B                                                 | R.EDSLDSVLFLK.K<br>R.SLLGMLSDLQVYK.D                                                                                                    | 50.28<br>50.18                                     | Q13619 |
| Cullin 2                                                 | K.FVQLINTVLNGDQHFMALDK.A<br>K.LMVEPLQAILIR.M                                                                                            | 49.04<br>43.94                                     | Q13617 |
| Damage-specific DNA binding protein 1, 127kDa            | K.TEPATGFIDGLIESFLDISRPK.M<br>R.ALYYLQIHPQELR.Q                                                                                         | 60.52<br>59.16                                     | Q16531 |
| Dihydrolipoamide S-acetyltransferase                     | K.VPLPSLSPTMQAGTIAR.W<br>K.GVETIANDVVSLATK.A                                                                                            | 74.47<br>47.27                                     | P10515 |
| DNA dependent protein kinase catalytic subunit           | K.TVSLLDENNVSSYLSK.N<br>K.QITQSALLAEAR.S                                                                                                | 72.96<br>43.03                                     | P78527 |
| Dynamin related protein 1                                | K.SSLDDLLTESEDMAQR.R<br>K.ALQGASQIAEIR.E                                                                                                | 112.12<br>70.14                                    | G8JLD5 |
| eIF2 alpha                                               | R.LERENAEVDGDDDAEMEAK.A<br>R.TEGLSVLSQAMAVIK.E<br>K.HAVSDPSILDSLNLNEDER.E                                                               | 61.91<br>55.89<br>55.87                            | P05198 |
| EIF3S4                                                   | K.LPGELEPVQATQNK.T<br>K.ELAEQLGLSTGEK.E                                                                                                 | 62.41<br>55.8                                      | O75821 |
| EIF4A1                                                   | K.GVAINMVTEEDKR.T<br>R.KGVAINMVTEEDKR.T                                                                                                 | 73.18<br>41.25                                     | P60842 |
| Electron transfer flavoprotein, alpha polypeptide        | R.AAVDAGFVPNDMQVGQTGK.I<br>K.TIVAINKDPEAPIFQVADYGIVADLFK.V                                                                              | 77.16<br>52.16                                     | H0YK49 |
| Eukaryotic translation elongation factor 1 gamma         | R.KLDPGSEETQTLVR.E<br>K.YSNEDTLSVALPYFWEHFDK.D<br>K.ALIAAQYSGAQVR.V<br>R.ILGLLDAYLK.T                                                   | 67.83<br>55.64<br>53.12<br>48                      | P26641 |



|                                                       |                                                                                                 |                                   |        |
|-------------------------------------------------------|-------------------------------------------------------------------------------------------------|-----------------------------------|--------|
|                                                       | K.LDPGSEETQTLVR.E                                                                               | 40.45                             |        |
| Eukaryotic translation initiation factor 3, subunit 1 | K.ETFGVNNAVYCIDAMNPSSR.D<br>K.DNWDDDDDEK.K                                                      | 112.22<br>41.95                   | O75822 |
| Eukaryotic translation termination factor 1           | K.FHTEALTALLSDDSK.F<br>K.TELSQSDFMDQR.L                                                         | 108.79<br>66.58                   | Q96CG1 |
| F actin capping protein alpha 2 subunit               | R.EGAAHAFAQYNLDQFTPVK.I<br>K.FTITPSTTQVVGILK.I                                                  | 52.38<br>46.01                    | P47755 |
| FACTp140                                              | K.VTMINAIPVASLDPIK.E<br>K.EELEFEVPR.D                                                           | 48.3<br>47.1                      | Q9Y5B9 |
| Flavin reductase                                      | K.TVAGQDAVIVLLGTR.N<br>R.NDLSPTTVMSEGR.N                                                        | 70.49<br>58.46                    | P30043 |
| GART                                                  | K.IEFVVVGPEAPLAAGIVGNLR.S<br>R.VFSWLQQEGHLSSEEMAR.T<br>R.ENLISALEEK.K<br>K.VDLGGFAGLFDLK.A      | 63<br>56.29<br>50.79<br>50.45     | P22102 |
| Gelsolin                                              | K.NWRDPDQTDGLGLSYLSSHIANVER.V<br>K.QTQVSVLPEGGETPLFK.Q<br>K.AGALNSNDAFVLK.T                     | 67.24<br>45.84<br>45.68           | P06396 |
| Glucose-6-phosphate dehydrogenase                     | K.KPGM#FFNPEESELDTYGNR.Y<br>R.GSTTATFAAVVLYVENER.W                                              | 61.94<br>49.71                    | E9PDV8 |
| Glutamine dependent carbamoyl phosphate synthase      | R.LSLDLLQR.L<br>R.EELSALVAPAFHTSQVLVDK.S                                                        | 66.81<br>53.82                    | P27708 |
| Glutamyl-prolyl-tRNA synthetase                       | R.WFGFLEAQAFQSVGK.W<br>R.VATTAGLYGSNLMHEIDEHWLEFSATK.L<br>K.THMVVANTMEDFQK.I                    | 88.25<br>44.83<br>42.86           | P07814 |
| Glutathione S transferase omega 1                     | K.EDPTVSALLTSEK.D<br>K.GSAPPGPVPEGSIR.I<br>K.VPSLVGSFIR.S                                       | 66.63<br>42.87<br>40.14           | P78417 |
| Glycosyltransferase 25 domain containing 1            | K.MLPVDEFLPVMFDK.H<br>K.EEYGLFPVPLR.A                                                           | 70.85<br>43.49                    | Q8NBJ5 |
| Heat shock 70 kDa protein 8                           | K.SINPDEAVAYGAAVQAAILSGDK.S<br>R.FEELNADLFR.G<br>K.SQIHDIIVLVGGSTR.I                            | 126.52<br>67.25<br>59.98          | A8K7Q2 |
| Heterogeneous nuclear ribonucleoprotein C             | R.MIAGQVLDINLAAEPK.V<br>K.GFAFVQYVNER.N                                                         | 87.86<br>58.65                    | P07910 |
| Histone deacetylase 2                                 | K.TFNLPLLMLGGGGYTIR.N<br>R.DGIDDESYGQIFKPIISK.V<br>K.YGEYFPGTGDLR.D<br>K.LHISPSNMTNQNTPEYMEK.I  | 76.21<br>66.49<br>57.6<br>42.03   | B3KRS5 |
| Hsc70 interacting protein                             | K.VAAIEALNDGELQK.A<br>R.AIEINPDSAQPYK.W<br>K.QDPSVLHTEEMR.F                                     | 59.22<br>54.12<br>53.24           | Q3KNR6 |
| IMP dehydrogenase 2                                   | K.LPIVNEDDELVAIAR.T<br>K.REDLVVAPAGITLK.E                                                       | 141.08<br>59.95                   | P12268 |
| Importin 11                                           | R.AGLITNFNEPINQIATQIAVLIK.V<br>K.LSALALLSLPSDNSVIQDK.F<br>R.VLLQNTSFFSSLLNEMAHK.F               | 55.23<br>45.39<br>45.04           | Q9UI26 |
| Importin 9                                            | K.EALVDLTGILSPVQEV.R.A<br>K.YSNDPVVASLAQDIFK.E<br>R.EVTDQTQMLVAPVILPEMYK.I                      | 83.09<br>69.44<br>40.88           | Q96P70 |
| Interleukin enhancer binding factor 2                 | K.ILPTLEAVAALGNK.V<br>R.NQDLAPNSAEQASILSLVTK.I<br>K.VLQSALAAIR.H                                | 70.2<br>56.52<br>50.52            | F4ZW62 |
| Karyopherin alpha 2                                   | K.GINSSNVENQLQATQAAR.K<br>R.NNQGTWNWSVDDIVK.G<br>K.EATWTMSNITAGR.Q<br>R.QDQIQQVNVHGLVPFLVSLSK.A | 145.58<br>79.38<br>67.52<br>60.38 | P52292 |
| KIAA0153 protein                                      | K.FNQTYQLAHGTAEK.M<br>R.AFALNDLDDYEK.H                                                          | 47.27<br>45.25                    | Q14166 |

|                                           |                                                                                                                                                                                                  |                                                                     |                   |
|-------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|-------------------|
| Kinesin family member 5B                  | K.SLSALGNVISALAEGSTYVPYR.D<br>R.NTIQWLENELNR.W                                                                                                                                                   | 85.69<br>45.17                                                      | P33176            |
| Leucine rich PPR motif containing protein | K.TVQLTSSELESTLETLK.A<br>R.SEAANGNLDVLSFLK.S<br>R.AGDMENAENILTVMR.D<br>R.LANQFAIYKPVTDFLQLVDAGK.V<br>K.NNNIDAAIENIENMLTSENK.V<br>K.MVFINNIALAQIK.N<br>K.VIEEQLEPAVEK.I<br>K.DLPVTEAVFSALVTGHAR.A | 96.7<br>87.75<br>65.52<br>51.61<br>47.04<br>46.96<br>44.05<br>43.64 | E5KNY5            |
| Leucine rich repeat containing 47         | R.NALGPGLSPELGPLPALR.V<br>R.APGPGLAQGLPQLHSLVLR.R                                                                                                                                                | 49.59<br>43.54                                                      | Q8N1G4            |
| Leukotriene B4 12 hydroxy dehydrogenase   | R.TGPLPPGPPPEIIVYQELR.M<br>K.GTIVLASPGWTTTHSISDGK.D                                                                                                                                              | 60.98<br>42.57                                                      | Q14914            |
| Macrophage capping protein                | K.YQEGGVESAFHK.T<br>R.EVQGNESDLFMSYFPR.G                                                                                                                                                         | 51.61<br>42.56                                                      | P40121            |
| Mitochondrial 3-oxoacyl-CoA thiolase      | K.VSPETVDSVIMGNVLQSSSDAIYLAR.H<br>K.DFTATDLSEFAAK.A                                                                                                                                              | 74.71<br>60.97                                                      | P42765            |
| NOL1/NOP2/Sun domain family, member 2     | K.WTTLNSLQLHGLQLR.I<br>K.FHQFLVSETESGNISR.Q                                                                                                                                                      | 46.86<br>45.28                                                      | Q08J23            |
| NS1 associated protein 1                  | K.DLEGENIEIVFAKPPDQK.R<br>K.VTEGLTDVILYHQPDCK.K<br>R.NLANTVTEEILEK.A<br>R.DLFEDELVPLFEK.A<br>R.LMMDPLTGLNR.G<br>K.VTEGLTDVILYHQPDCK.K                                                            | 84.27<br>67.15<br>62.93<br>55.89<br>55.85<br>52.58                  | O60506            |
| Nuclear mitotic apparatus protein 1       | K.GEVLGDVLQLETLK.Q<br>R.QDHAQQLATAAEER.E                                                                                                                                                         | 61.19<br>44.02                                                      | Q14980            |
| Nucleoside diphosphate kinase A           | R.VMLGETNPADSKPGTIR.G<br>K.EISLWFKPEELVDYK.S                                                                                                                                                     | 89.1<br>59.84                                                       | F6XY72            |
| Nudix type motif 5                        | K.EQTADGVAVIPVLR.T<br>K.QYIISEELISEGK.W                                                                                                                                                          | 70.22<br>46.3                                                       | Q9UKK9            |
| ORP150                                    | K.VLQLINDNTATALSYGVFR.R<br>K.LGNTISSLFGGGTTTPDAK.E<br>R.VESVFETLVEDSAEEESTLTK.L                                                                                                                  | 127.47<br>80.25<br>55.29                                            | Q9Y4L1            |
| Osmotic stress protein 94                 | R.SGGIETIANEYSDR.C<br>K.QNLEGDHSAPMETETSFK.N                                                                                                                                                     | 71.82<br>54.12                                                      | IPI009646<br>84.1 |
| PA2G4                                     | R.ITSGFEPDLYK.S<br>K.TIIQNPTDQOK.K<br>K.SEMEVQDAELK.A                                                                                                                                            | 54.49<br>51.92<br>48.64                                             | H0Y3X3            |
| Peroxiredoxin 3                           | R.DYGVLLLEGSLALR.G<br>K.HLSVNDLPVGR.S                                                                                                                                                            | 44.37<br>40.54                                                      | P30048            |
| Peroxiredoxin 5                           | K.VGDAIPAVEVFEGEPGNK.V<br>R.FSMVVQDGIVK.A                                                                                                                                                        | 67.07<br>66.99                                                      | P30044            |
| PGK2                                      | K.LGDVYVNDAFGTAHR.A<br>R.AHSSM#VGVNLPQK.A                                                                                                                                                        | 74.89<br>48.22                                                      | B4E1H9            |
| Phenylalanyl-tRNA synthetase alpha chain  | K.SLQALGEVIEAELR.S<br>R.RLEASDGGGLSAELAAELGMEHQAVVGAVK.S                                                                                                                                         | 70.82<br>70.12                                                      | K7ER00            |
| Phosphoglucosyltransferase 1              | K.ADNFEYSDPVDGSISR.N<br>K.IALYETPTGWK.F<br>K.INQDPQVMLAPLISIALK.V<br>K.VFQSSANYAENFIQSIISTVEPAQR.Q                                                                                               | 66.43<br>49.37<br>42.23<br>42.08                                    | P36871            |
| Phosphorybosylaminoimidazole carboxylase  | K.AEYEGDGIPTVFVAVAGR.S<br>K.DDANNDPQWSEEQLIAAK.F                                                                                                                                                 | 92.94<br>84.56                                                      | P22234            |
| Poly(A) binding protein 4                 | R.SLGYAYVNFQPPADAER.A<br>K.NFGEEVDDESLKELFSQFGK.T<br>R.ALDTMNFDFVIK.G                                                                                                                            | 88.88<br>58.41<br>47.96                                             | Q4VC03            |
| Polypyrimidine tract binding protein 1    | R.IAIPGLAGAGNSVLLVSNLNP.V<br>R.IIVENLFYPVTLVDLHQLFSK.F                                                                                                                                           | 110.39<br>41.24                                                     | P26599            |

|                                                          |                                                                                                                         |                                          |        |
|----------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|------------------------------------------|--------|
| Proteasome 26S subunit ATPase 5                          | K.EVIELPVKHPELFEALGIAQPK.G<br>K.IEELQLIVNDK.S                                                                           | 52.64<br>49.81                           | J3QSE0 |
| Proteasome 26S subunit, non-ATPase, 13                   | R.QWLIDTLYAFNSGNVER.F<br>K.TAWGQQPDLAANEAQLLR.K<br>K.YYQTIGNHASYYK.D<br>K.ETIEDVEEMLNLPGVTSVHSR.F<br>R.VNPLSLVEILHVVR.Q | 69.41<br>64.81<br>63.4<br>58.69<br>48.43 | J3KNQ3 |
| Protein disulfide isomerase                              | K.QFLQAAEAIDDIPFGITSNSDVFSK.Y<br>R.EADDIVNWLK.K                                                                         | 99.86<br>50.52                           | P07237 |
| Protein transport protein Sec24C                         | R.AVITSLLDQIPEMFADTR.E<br>K.TLFQPQTGAYQTLAK.E                                                                           | 47.62<br>41.25                           | G5EA31 |
| Protein tyrosine phosphatase, non-receptor type 11       | R.WFHPNITGVEAENLLLTR.G<br>R.FDSLTDLVEHYK.K                                                                              | 56.1<br>48.51                            | Q06124 |
| PRPF8                                                    | R.EAVVNTQELLDLLVK.C<br>K.ATEPQMVLFLNLYDDWLK.T                                                                           | 89.61<br>40.14                           | Q6P2Q9 |
| RAN, member RAS oncogene family                          | K.SNYNFEKPFLWLAR.K<br>K.YVATLGVEVHPLVFHTNR.G                                                                            | 57.13<br>55.11                           | P62826 |
| Rho GDP dissociation inhibitor alpha                     | R.AEEYEFLTPVEEAPK.G<br>K.SIQEIQELDKDDESLR.Y<br>K.SIQEIQELDKDDESLR.K                                                     | 46.73<br>42.72<br>42.55                  | J3KRE2 |
| Ribosomal phosphoprotein large P2                        | R.YVASYLLAALGGNSSPSAK.D<br>K.ILDSVGIEADDDRNLK.V                                                                         | 110.24<br>90.74                          | P05387 |
| Ribosomal protein L5                                     | K.GAVDGGLSIPHSK.R<br>R.YLMEEDEDAYKK.Q                                                                                   | 51.57<br>42.69                           | P46777 |
| RuvB like 2                                              | R.TQGFLALFSGDTGEIK.S<br>K.TTEMETIYDLGTK.M<br>R.ALESDMAPVLIMATNR.G                                                       | 67.11<br>54.56<br>46.61                  | Q9Y230 |
| S adenosylhomocysteine hydrolase                         | K.DGPLNMILDDGGDLTNLIHTK.Y<br>R.GISEETTTGVHNLK.M                                                                         | 97.29<br>95.69                           | F5H737 |
| SEC23 related protein A                                  | K.IDMNLTDLLGELQR.D<br>R.MVVPVAALFTPLK.E                                                                                 | 86.58<br>45.84                           | Q15436 |
| Sequestosome 1                                           | K.NYDIGAALDTIQYSK.H<br>K.LAFPSPFGHLSEGFHSR.W                                                                            | 72.07<br>47.35                           | Q13501 |
| Small nuclear ribonucleoprotein polypeptide D1           | R.YFILPDSLPLDTLLVDVEPK.V<br>K.NREPVQLETLSIR.G                                                                           | 69.5<br>40.93                            | J3QLI9 |
| Solute carrier family 9, isoform A3, regulatory factor 1 | R.AQEAPGQAEPAAAAEVQGAGNENEPR.E<br>R.LLVVDPETDEQLQK.L                                                                    | 69.18<br>41.65                           | O14745 |
| Splicing factor 3B subunit 1                             | R.SLVEIIEHGLVDEQK.V<br>K.DYIYAVTPLLEDALMDR.D                                                                            | 75.45<br>71.56                           | O75533 |
| T Complex 1                                              | R.SLLVIPNTLAVNAAQDSTDVAK.L<br>K.WIGLDLSNGKPR.D                                                                          | 87.58<br>55.17                           | E7EQR6 |
| T Complex 1                                              | K.MLVDDIGDVTITNDGATILK.L<br>R.YINENLIVNTDELGR.D                                                                         | 109.36<br>84.8                           | E7ERF2 |
| Thioredoxin reductase 1                                  | R.MIEAVQNHIGSLNWGYR.V<br>K.VMVLDFVTPTPLGTR.W                                                                            | 60.97<br>60.57                           | B7Z904 |
| Translocated promoter region                             | R.GQNLLLTNLQTIQILERS<br>K.QEQIINTMTQDLR.G<br>K.QHLSNMEVQVASQSSQR.T                                                      | 97.65<br>62.74<br>46.38                  | P12270 |
| Transportin 1                                            | K.TLAQAMLNNAQPDQYEAPDK.D<br>R.LLPHMHNIVEYMLQR.T                                                                         | 104.14<br>42.91                          | Q92973 |
| TRIP13                                                   | R.NVQSVSIIDTELK.V<br>R.VVNAVLTQIDQIK.R<br>K.SHLLDYVMTLLFSDK.N                                                           | 55.02<br>49.02<br>45.66                  | Q15645 |
| Tryptophanyl tRNA synthetase                             | K.ISFPAIQAAPSFNSFPQIFR.D<br>K.MSASDPNSSIFLTD TAK.Q<br>K.DLTLDQAYSAYAVENAK.D                                             | 110.63<br>106.07<br>63.02                | P23381 |
| Tubulin, beta 6                                          | R.AALVDLEPGTMDSVR.S<br>R.INVYYNESSSQK.Y                                                                                 | 51.88<br>49.52                           | K7EJ64 |
| Tubulin, beta 6                                          | K.MASTFIGNSTAIQELFK.R                                                                                                   | 96.77                                    | Q9BUF5 |

|                                            |                               |       |        |
|--------------------------------------------|-------------------------------|-------|--------|
|                                            | K.LTTPTYGDLNHLVSATMSGVTTSLR.F | 62.37 |        |
| Tyrosyl-tRNA synthetase                    | K.DFAAEVVHHPGDLK.N            | 49.8  | P54577 |
|                                            | K.VDAQFGGIDQR.K               | 42.76 |        |
|                                            | K.NSEPEEVIPSR.L               | 40.75 |        |
| U5 200kD                                   | R.LIGLSATLPNYEDVATFLR.V       | 89.34 | O75643 |
|                                            | R.WTELGALDILQMLGR.A           | 66.46 |        |
| Ubiquitin carboxyl terminal<br>esterase L3 | R.WLPLEANPEVTNQFLK.Q          | 59.79 | P15374 |
|                                            | R.VTHETSAHEGQTEAPSIDEK.V      | 43.88 |        |
| Ubiquitin specific protease 5              | K.IFQNAPTDPTQDFSTQVAK.L       | 90.93 | P45974 |
|                                            | K.LDVSIEMPEELDISQLR.G         | 45.27 |        |

