

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

Chemical Imaging and Proteomics Approaches for Probing Copper and Histidine Biology

### Permalink

<https://escholarship.org/uc/item/6pn1x867>

### Author

Jia, Shang

### Publication Date

2019

Peer reviewed|Thesis/dissertation

Chemical Imaging and Proteomics Approaches for Probing Copper and Histidine Biology

By

Shang Jia

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Christopher J. Chang, Chair

Professor Sanjay Kumar

Professor Matthew B. Francis

Professor Michelle C. Chang

Summer 2019

Chemical Imaging and Proteomics Approaches for Probing Copper and Histidine Biology

© 2019

by Shang Jia

## **Abstract**

Chemical Imaging and Proteomics Approaches for Probing Copper and Histidine Biology

by

Shang Jia

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Christopher J. Chang, Chair

Chemistry provides valuable tools to probe and influence the molecular processes in biology. Such processes are made available by both large biomolecules and small molecule species. As an example of a transition metal nutrient, copper plays essential roles in respiration, electron transfer, antioxidant defense and oxidation of metabolites, while aberrant copper homeostasis occurs with severe diseases states including neurodegenerative disorders, diabetes and cancer. On the other side, proteins, as large biomolecules, are composed of simple amino acid building blocks. Histidine is an important amino acid, with its imidazole side chain frequently used at catalytic centers and metal-binding sites. This dissertation summarizes my efforts to probing copper and histidine in their biology contexts. Specifically, a series of copper fluorescent sensors with varying colors were designed for visualizing this metal in cells. For probing histidine, we firstly developed bio-inspired thiophosphorodichloridate reagents for its chemoselective labeling on proteins. The labeling of this amino acid provides a starting point for probing histidine by proteomics methodology. Further optimization of the reagents as well as the protocol enables the profiling of histidine in a proteome context.

## Dedication

*To my parents, for their unwavering love and support  
And Lei Guo, my best friend, my soul mate, and the love of my life.*

## Table of Contents

Acknowledgement.....	iii
Chapter 1 Synthetic tools to study metals and proteins by affinity-based and activity-based sensing.....	1
Chapter 2 Development of a color palette of fluorescent copper sensors through heteroatom substitution at rhodol cores .....	20
Chapter 3 Development of thiophosphorodichloridate reagents for chemoselective histidine bioconjugation.....	55
Chapter 4 Optimization of thiophosphorodichloridate labeling strategy for histidine profiling .....	93
Appendix 1 Imaging of cellular hydrogen peroxide by a microtubule-localizing H <sub>2</sub> O <sub>2</sub> fluorescent sensor.....	125
Appendix 2 Development of propargyl pyrocarbonate and its analogs for protein profiling.....	144
Appendix 3 Fast sample preparation of digested proteins by immobilized trypsin.....	160

## Acknowledgement

First and foremost, thank you Prof. Christopher J. Chang for serving as my Ph.D. advisor. He is extraordinarily supportive, encouraging, and is always calm and gentle. He cares about his students inside and outside the laboratory and I am so blessed to have spent my five years of graduate life in his laboratory. Prof. Chang has set a remarkable example for me as a principle investigator and will be always my mentor throughout my career. I am additionally grateful for the valuable support from Prof. Peng Chen, my undergraduate mentor at Peking University.

I am also indebted to Vivian Lin, Shixian Lin and Jie Li for being nice and supportive mentors throughout my years in the laboratory. They taught me from the basic lab skills to the understanding of scientific projects, guiding me to grow from a newbie to an actual researcher in chemistry. All of the Chang lab members have helped to make my time at graduate school rich and enjoyable. I especially thank Karla Ramos-Torres, Safacan Köleman, Cheri Ackerman, Diana Iovan, Dan He, Lakshmi Krishnamoorthy and Zeming Wang for their collaboration, guidance and friendship. It is my great pleasure and to have learned and worked with these excellent labmates.

Last but not least, I want to express my deepest gratitude to my parents. They have been providing me with unconditional love and support since my birth; it is not possible for me to have journeyed from a mid-sized city to the capitol of China and now overseas at the United States without their guidance, encouragement, and insights. And Lei, it is my greatest luck to meet you here at Berkeley and I have always been enjoying the time we spend together, from the day we meet until forever.

# Chapter 1

## **Synthetic tools to study metals and proteins by affinity-based and activity-based sensing**

Portions of this work were published in the following scientific journal:

Iovan, D. A.\*; Jia, S.\*; Chang, C. J. Inorganic Chemistry Approaches to Activity-Based Sensing: From Metal Sensors to Bioorthogonal Metal Chemistry. *Inorg. Chem.* **2019**, doi 10.1021/acs.inorgchem.9b01221.

(\* Denotes equal contribution.)

## 1.1 Background and motivation

The complex network of chemical processes that sustain life motivates the development of new synthetic tools to decipher biological mechanisms of action at a molecular level. In this context, fluorescent and related optical probes have emerged as useful chemical reagents for monitoring small-molecule and metal signals in biological systems, enabling visualization of dynamic cellular events with spatial and temporal resolution.

The design of fluorescent probes for small molecule sensing has widely explored the versatile reactivity of the analytes and the coordination chemistry of metals, encompassing two general mechanisms for achieving specificity: (1) binding-based sensing (BBS) and (2) activity-based sensing (ABS).<sup>1</sup> In both cases, a site that recognizes the analyte is appended to a fluorophore to effect a change in fluorescence upon interaction with the analyte of interest. The binding-based approach, which is widely-explored in sensing of metal ions, relies on a chelating ligand site that reversibly binds the metal of interest and induces a fluorescence response via mechanisms such as internal charge transfer (ICT), photoinduced electron transfer (PET), or Förster resonance energy transfer (FRET).<sup>2</sup> Using these fundamental coordination chemistry principles, fluorescent sensors have been successfully developed for detection of many endogenous metals<sup>1,3-5</sup> in biological systems, most commonly  $\text{Ca}^{\text{II}}$ <sup>6</sup> and  $\text{Zn}^{\text{II}}$ ,<sup>4,5</sup> as well as  $\text{Cu}^{\text{I}}$ <sup>7-9</sup> and  $\text{Ni}^{\text{II}}$ .<sup>10</sup> On the other hand, the ABS approach makes use of analyte-specific bond forming or cleavage reactions to modify the optical properties of the fluorophore and results in an irreversible turn-on or change in fluorescence response. Transient and reactive small molecules that would be difficult to track owing to their similar shape and size as other biological metabolites are capable of engaging in highly specific and bioorthogonal chemical reactions.<sup>11-14</sup> Representative examples from our laboratory include redox-neutral 2-aza-Cope type rearrangements for formaldehyde sensing,<sup>15,16</sup> oxidative conversion of aryl boronates for hydrogen peroxide detection,<sup>17-21</sup> and reductive organic azide to aniline transformations for hydrogen sulfide sensing.<sup>22-24</sup> Such chemical reaction event can be combined with a metal binding process to enable detection of redox-active transition metals including  $\text{Fe}^{\text{II}}$ ,  $\text{Co}^{\text{II}}$ , and  $\text{Cu}^{\text{I}}$ , with high metal and oxidation-state selectivity.<sup>1</sup>

In addition to small molecule biological analytes, large biomolecules, including proteins, can also be analyzed using small molecule probes. In this context, activity-based reaction on proteins has enabled activity-based sensing of proteins functions. Early reports of such method made use of activity-based reaction on serine proteases to track their function by in-gel fluorescence.<sup>25,26</sup> In the post-genomic era, advances in tandem mass spectrometry and separation technology are combined to provide a profiling platform for sensitive readout of the protein of interest. Such method, termed activity-based protein profiling (ABPP), are widely applied for characterization of enzyme function, as well as research of natural product targets and post-translational modifications.<sup>27-29</sup>

## 1.2 Fluorescent sensing of Copper

Copper is an essential element for biology. Owing to its redox activity, copper plays central roles in a variety of proteins covering oxidases, reductases, free radical scavenging, electron transfer and energy production, with dedicated proteins for its transport, storage and chaperonage.<sup>30</sup> Aberrant copper homeostasis, including both deficiencies and excesses, can lead to severe disease states.<sup>31</sup> The geometry of the ligand and hard-soft acid-base preferences provides guidelines for designing metal-specific ligand for *in vivo* copper detection. The following section briefly summarizes binding-based and activity-based fluorescent sensors for Cu<sup>I</sup> as an example of an essential transition metal nutrient, focusing on ligands and triggers that can be generally applied to multiple imaging modalities.

### 1.2.1 Binding-based sensing of Copper

Unlike divalent iron and cobalt, monovalent copper is the dominant oxidation state of this metal in reducing intracellular environments.<sup>32</sup> As a diamagnetic ion, Cu<sup>I</sup> does not effectively quench fluorophores and as such, fluorescence detection of this species can be read out by employing PET switch or a related charge transfer pathway. In its off-state, the electron on the receptor can be transferred onto the excited fluorophore and thus quench the fluorescence. When the receptor binds Cu<sup>I</sup>, the electron transfer pathway is inhibited, thereby the fluorescence is turned-on. The Cu<sup>I</sup>-specific receptor commonly employs a thioether-rich ligand as a soft ligand for the cuprous ion, as an inspiration from cysteine and methionine used as copper-proteins' binding sites. Chemical structures of various Cu<sup>I</sup> probes are shown Figure 1.1.

The first small molecule fluorescent sensor for Cu<sup>I</sup>, CTAP-1 was reported by Fahrni and co-workers (Figure 1.1).<sup>33</sup> CTAP-1 consists of a 1,3-diarylpyrazoline fluorophore linked to a tetrathiazacrown ether receptor (NS4). This probe exhibits a 5-fold fluorescence turn-on upon UV excitation with high specificity for Cu<sup>I</sup>. The fluorescence of CTAP-1 was found to colocalize with copper and sulfur signals by XFM imaging. Following this initial design, modification to the fluorophore and copper receptor has been introduced to achieve higher turn-on ratio,<sup>34</sup> improved aqueous solubility (CTAP-2),<sup>35</sup> and tighter Cu<sup>I</sup> binding affinity (CTAP-3) (Figure 1.1).<sup>36</sup> CTAP-2 was applied to the detection of copper center of Atox1, a copper transporter protein, by in-gel fluorescence readout, and CTAP-3 was shown to be useful in detection of copper without turn-on effect from liposomes. However, due to the short absorption and emission wavelengths and limited cellular uptake, these probes are yet to find more applications in cellular imaging experiments.

Our laboratory's research in copper chemistry and biology gave rise to CS1 as the first fluorescent Cu<sup>I</sup> sensor with excitation and emission wavelengths in the visible region (Figure 1.1). CS1 was designed by combining dipyrromethene (BODIPY)-based fluorophore with an acyclic NS4 receptor. In addition to its 10-fold turn-on response to Cu<sup>I</sup> in *in vitro* experiments with high Cu<sup>I</sup> selectivity, CS1 can detect copper supplementation or depletion in HEK293T cells by fluorescence imaging. Following a similar design, our laboratory was able to synthesize CS3, with a turn-on ratio of 75-fold and a tighter Cu<sup>I</sup> binding affinity, by replacing the electron-poor

fluorine atoms on the fluorophore with electron-rich methoxy groups.<sup>37</sup> On top of this, we introduced the same copper receptor NS4' was to rhodol fluorophores for improved hydrophilicity and longer shelf life, resulting in rhodol-based CF3<sup>38</sup> and CF4<sup>39</sup>, and silicon-rhodol based CSR1 (Figure 1.1).<sup>40</sup> Besides the copper-responsive fluorescent dyes, our laboratory also introduced their Cu<sup>I</sup>-nonresponsive analogs, with two or four sulfur atoms on the ligand moiety replaced with isosteric methylene groups that disrupts its affinity towards metals. These control compounds can be used in extra control experiments to distinguish dye-related response from copper-dependent turn-on. With their suitable hydrophilicity and excitation wavelengths, this family of fluorescent copper sensors has assisted the discovery of new copper biology in bacteria,<sup>41,42</sup> yeast,<sup>43-45</sup> plant,<sup>46</sup> worm<sup>47</sup> and mammalian<sup>48-51</sup> models.

### 1.2.2 Activity-based sensing of copper

The unique chemical properties of metals, including tunable Lewis acidity and redox activity, have been exploited toward the development of reactivity triggers for imaging bioavailable metal pools. Herein we list a few ABS probes for and copper detection, focusing on triggers that can be generally applied to multiple imaging modalities.

In early work in this area, Taki and colleagues drew inspiration from the vast literature on the O<sub>2</sub> activation by bioinorganic model complexes for copper-dependent enzymes<sup>52-54</sup> and reported **FluTPA2**,<sup>55</sup> a probe featuring a Tokyo Green derivative caged by the Cu<sup>I</sup> chelator tris[(2-pyridyl)-methyl]amine (TPA). This compound undergoes a Cu<sup>I</sup>- and O<sub>2</sub>-dependent oxidative cleavage of the TPA benzyl ether linkage with subsequent oxidation and release to generate the highly fluorescent Tokyo Green fluorophore product (Figure 1.2a). **FluTPA2** selectively detects Cu<sup>I</sup> over Cu<sup>II</sup> in the presence of glutathione (GSH) with a >100-fold turn-on and can report changes in cellular copper levels. This ABS trigger was further adapted for mitochondria-targeting (Figure 1.2c)<sup>56</sup> and separately coupled with coumarin-based fluorophores.<sup>57</sup> Our laboratory employed this motif to cage luciferin and create Copper Caged Luciferin 1 (**CCL-1**, Figure 1.2a) for *in vivo* bioluminescent copper imaging, showing in a diet-induced mouse model of non-alcoholic fatty liver disease (NAFLD) that copper deficiency occurs as early as two weeks before metabolic symptoms of the disease manifest at two months (Figure 1.2d).<sup>58</sup>

In contrast to the growing use of chemical probes for bioimaging of Cu<sup>I</sup>, detection of Cu<sup>II</sup> in biological environments is rarer. An early approach for Cu<sup>II</sup> sensing was developed by Czarnik and colleagues,<sup>59</sup> employing a hydrazide-substituted spirorhodamine fluorophore locked in a non-fluorescent ring-closed form. Binding of Cu<sup>II</sup> promotes hydrazide hydrolysis with concomitant ring-opening of the spirorhodamine, thus turning-on the fluorophore (Figure 1.2b). Depending on the fluorophore/trigger combination, hydrazide-based Cu<sup>II</sup> detection has been reported in many solvents, and two-dye systems with a hydrazide motif have been reported to enable ratiometric imaging of Cu<sup>II</sup> in living cells.<sup>60-62</sup>

### 1.3 Fluorescent sensing of H<sub>2</sub>O<sub>2</sub>

Reactive oxygen species (ROS) are naturally occurring transient oxidants generated from cellular respiration through a variety of pathways.<sup>63–65</sup> Misregulation of these processes is associated with disease states including Alzheimer's, Huntington's, Parkinson's and other neurodegenerative diseases.<sup>66–68</sup> On the other hand, controlled generation of ROS has recently connected with cell signaling pathways.<sup>69,70</sup> In particular, H<sub>2</sub>O<sub>2</sub> can lead to certain protein post-translational modification on cysteine and methionine residues to affect the function of the downstream proteins.<sup>71,72</sup> The following sections focus on detection of H<sub>2</sub>O<sub>2</sub> by two activity-based sensing mechanisms, namely boronate oxidation and metal-redox reactions.

#### 1.3.1 Detection of H<sub>2</sub>O<sub>2</sub> by boronate ester

Boronic acids and boronate esters have been extensively used in organic synthesis as chemical building blocks predominantly in the Suzuki coupling. In addition to its activity in organometallic reactions, its reduction activity can be utilized for designing chemosensors.<sup>73,74</sup> Our laboratory has harnessed this reactivity to incorporate the boronate ester as a caging group for fluorophores, which, upon selective oxidation cleavage by H<sub>2</sub>O<sub>2</sub> as a decaging process, results in free hydroxyl group that turns on the fluorophore. This family of H<sub>2</sub>O<sub>2</sub> fluorescent sensors includes freely-diffusing sensors such as PF1,<sup>21</sup> and targeting sensors including cytosol-trapping PF6-AM,<sup>75</sup> mitochondria-targeting Mito-PY1<sup>76</sup> and nuclear-localizing NucPE1,<sup>77</sup> all of which exhibit excellent selectivity towards H<sub>2</sub>O<sub>2</sub> over other competing reactive species (Figure 1.3). These fluorescent indicators have been applied to hydrogen peroxide detection in a variety of cell types for exploring redox-associated biology. Furthermore, such caging utility can also extend beyond fluorescent detection to PET imaging with PC-[<sup>18</sup>F]FLT-1<sup>78</sup> and puromycin staining with peroxympicin-1 (Figure 1.3).<sup>79</sup>

#### 1.3.2 Detection of H<sub>2</sub>O<sub>2</sub> by metal sensors

Drawing inspiration from heme and non-heme iron enzymes that utilize hydrogen peroxide as an oxidant, Kodera and colleagues developed a fluorescent H<sub>2</sub>O<sub>2</sub> indicator combining a reduced resorufin analogue and an Fe<sup>III</sup> polypyridine complex to generate **MBFh1**. Oxidation of the dye-bound ferric center with H<sub>2</sub>O<sub>2</sub> is proposed to initiate subsequent two electron ligand oxidation, resulting in C–O bond cleavage to liberate the fluorescent resorufin dye.<sup>80</sup> Subsequent modifications of the fluorophore scaffold to incorporate an *O*-alkyl resorufin derivative (**MBFh2**, Figure 1.4a) afforded a stable probe for intracellular detection of hydrogen peroxide (Figure 1.4b).<sup>81</sup> In a similar approach, Nam and co-workers employed a known Zn sensor and made its corresponding Fe<sup>II</sup> complex (**ZPIFe<sub>2</sub>**),<sup>82</sup> which is activated by H<sub>2</sub>O<sub>2</sub> to promote oxidative *N*-dealkylation and release the bis-carboxylate fluorescein (Figure 1.4a). Unlike the ferric-based sensor, this diiron probe selectively reacts with H<sub>2</sub>O<sub>2</sub> over other ROS such as O<sub>2</sub><sup>•-</sup> or <sup>t</sup>BuOOH and was found to localize at the lysosome, enabling organelle-specific hydrogen peroxide detection.

## 1.4 Probing proteins by activity-based profiling

Activity-based protein profiling (ABPP) based on tandem mass spectrometry utilizes small molecule probes to report protein function in a whole proteome context.<sup>27</sup> Typically, such small molecule probes contain a reactive “warhead” and a tag for enrichment (see Figure 1.5b for examples). The warhead is used to react with the protein to form conjugate with site or residue-specificity. The tag moiety provides a way for the pull-down of probe-labeled protein, with biotin or desthiobiotin being most frequently used. Alkyne can also serve as a handle for subsequent click reaction with biotin-containing azide. The enriched proteins can be analyzed by gel-electrophoresis, or LC-MS, or the combination of both to give the identification of the labeled protein; a generalized workflow is shown in Figure 1.5a. Notably, as a way for quantitative analysis using mass spectrometry, isotopically-labeled probes or tags can be used for labeling that enables the quantitative comparison of the same protein under two or more experimental conditions. This methodology, termed isotopic Tandem Orthogonal Proteolysis (isoTOP)-ABPP, expands the applicability of ABPP to solve challenging biology questions.<sup>83</sup> With its growing toolbox, ABPP has enabled profiling of specific enzymes including hydrolases, proteases and phosphatases,<sup>27,84,85</sup> post-translational modifications such as cysteine sulfenylation<sup>86</sup> and sulfinylation<sup>87</sup>, hyper-reactive and “ligandable” residues including lysine,<sup>88</sup> cysteine<sup>89</sup> and methionine<sup>90</sup> and so on.

## 1.5 Dissertation overview

An interest in probing and functioning biological-relevant species with small molecule tools has inspired us to design molecules with property optimized for specific biological situations, including fluorescence imaging and protein bioconjugation. The work in this dissertation focuses on the design, synthesis and application of various chemical tools for the detection of copper and the labeling of histidine.

Chapter 2 presents the development of a series of fluorescent copper(I) sensors with varying absorption and emission wavelengths. After confirming their turn-on response in *in vitro* experiment, we successfully applied the new red-shifted probes to the detection of labile-copper pool increase in cell-based models of disease as well as dual-color imaging of copper and calcium.

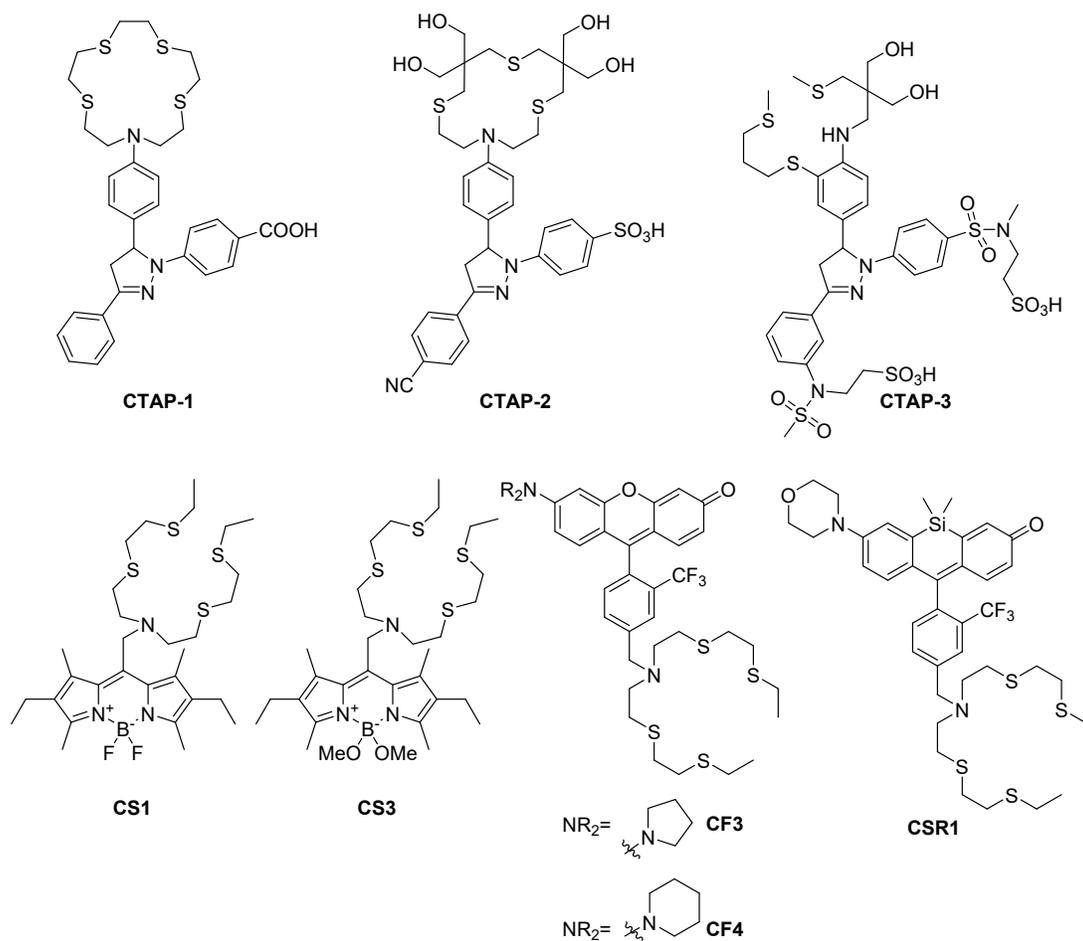
Chapter 3 describes the design of thiophosphorodichloridates as histidine-selective bioconjugate reagents. This compound exhibits fast kinetics and high selectivity towards histidine residue on small molecule and on protein models. The alkyne handle introduced by thiophosphorodichloridate modification allows for introduction of complex payloads onto the labeled histidine and we employed this reagent for protein delivery by modifying the polyhistidine tag region.

Chapter 4 details the optimization of thiophosphorodichloridates for histidine profiling in proteome. With a longer alkyne handle, the labeling reagents offer much stable bioconjugate on

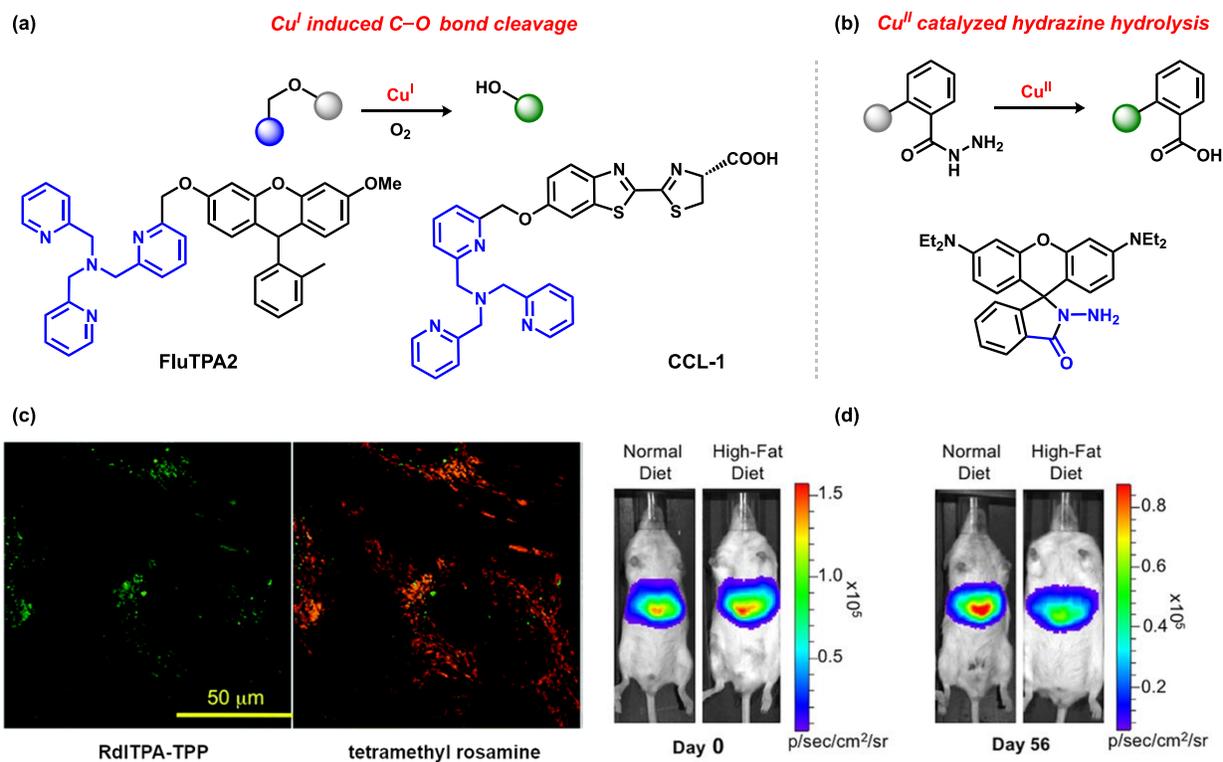
histidine. We also introduced triply-charged azide compounds for signal enhancement on mass spectrometers.

Appendices report the development of a microtubule-localizing  $\text{H}_2\text{O}_2$  probe for higher resolution imaging of cellular  $\text{H}_2\text{O}_2$  (Appendix 1), pyrocarbonate derivatives for nucleophilic amino acid profiling (Appendix 2) and protocol examples of fast proteomics sample preparation using immobilized trypsin (Appendix 3).

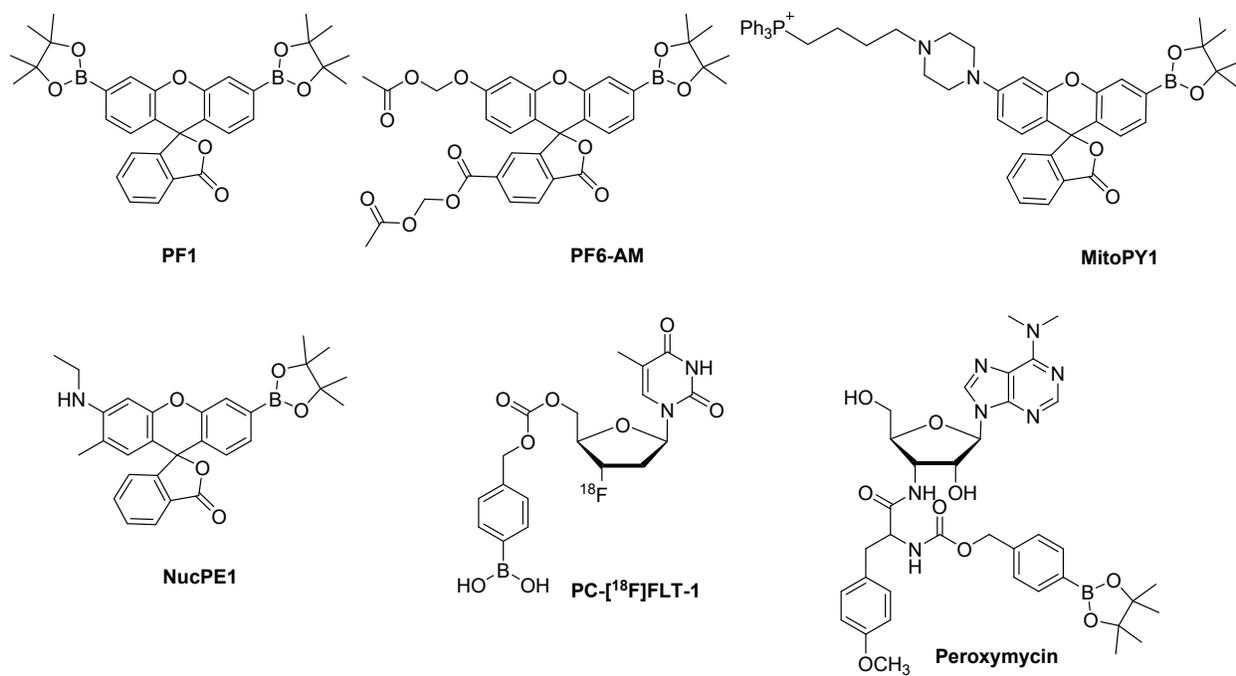
## 1.6 Figures and schemes



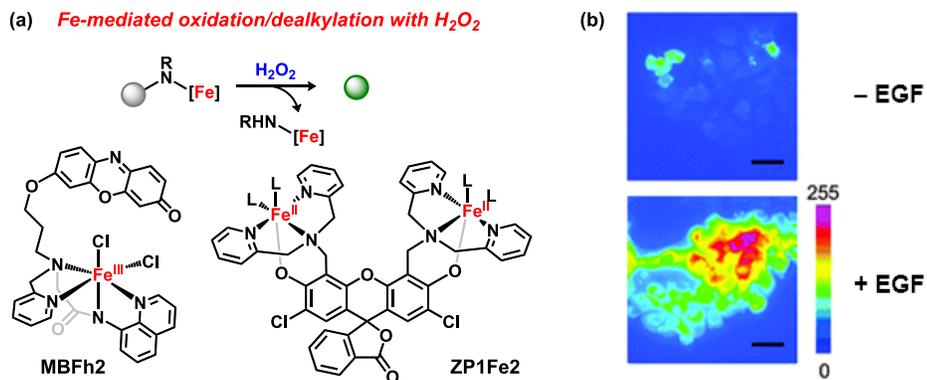
**Figure 1.1** Structures of selected binding-based Cu<sup>I</sup> probes.



**Figure 1.2** Structures and applications of selected activity-based probes for Cu detection. (a) Biomimetic  $\text{Cu}^{\text{I}}$ - $\text{O}_2$  activation, followed by C–O cleavage to release fluorophore;<sup>55,58</sup> (b)  $\text{Cu}^{\text{II}}$  catalyzed hydrazide hydrolysis.<sup>59</sup> (c) Monitoring copper levels in the mitochondria (**RdITPA-TPP**).<sup>56</sup> (d) Monitoring labile copper levels in a diet-induced mouse model of non-alcoholic fatty liver disease (**CCL-1**).<sup>58</sup>



**Figure 1.3** Structures of selected boronate-based probes for H<sub>2</sub>O<sub>2</sub> sensing.



**Figure 1.4** Structures and applications of selected H<sub>2</sub>O<sub>2</sub> sensors based on metal redox reaction. (a) Biomimetic Fe<sup>II</sup> and Fe<sup>III</sup> hydrogen peroxide activation followed by ligand oxidation to release fluorophore.<sup>81,82</sup> (b) Monitoring H<sub>2</sub>O<sub>2</sub> levels upon EGF stimulation of A431 cells (MBFh2).<sup>81</sup>



## 1.7 References

- (1) Aron, A. T.; Ramos-Torres, K. M.; Cotruvo, J. A.; Chang, C. J. Recognition- and Reactivity-Based Fluorescent Probes for Studying Transition Metal Signaling in Living Systems. *Acc. Chem. Res.* **2015**, *48* (8), 2434–2442.
- (2) Fu, Y.; Finney, N. S. Small-Molecule Fluorescent Probes and Their Design. *RSC Adv.* **2018**, *8* (51), 29051–29061.
- (3) Domaille, D. W.; Que, E. L.; Chang, C. J. Synthetic Fluorescent Sensors for Studying the Cell Biology of Metals. *Nat. Chem. Biol.* **2008**, *4* (3), 168–175.
- (4) Que, E. L.; Domaille, D. W.; Chang, C. J. Metals in Neurobiology: Probing Their Chemistry and Biology with Molecular Imaging. *Chem. Rev.* **2008**, *108* (5), 1517–1549.
- (5) Carter, K. P.; Young, A. M.; Palmer, A. E. Fluorescent Sensors for Measuring Metal Ions in Living Systems. *Chem. Rev.* **2014**, *114* (8), 4564–4601.
- (6) Minta, A.; Kao, J. P.; Tsien, R. Y. Fluorescent Indicators for Cytosolic Calcium Based on Rhodamine and Fluorescein Chromophores. *J. Biol. Chem.* **1989**, *264* (14), 8171–8178.
- (7) Cotruvo, J. A., J.; Aron, A. T.; Ramos-Torres, K. M.; Chang, C. J. Synthetic Fluorescent Probes for Studying Copper in Biological Systems. *Chem. Soc. Rev.* **2015**, *44* (13), 4400–4414.
- (8) Ramos - Torres, K. M.; Kolemen, S.; Chang, C. J. Thioether Coordination Chemistry for Molecular Imaging of Copper in Biological Systems. *Isr. J. Chem.* **2016**, *56* (9–10), 724–737.
- (9) Fahrni, C. J. Synthetic Fluorescent Probes for Monovalent Copper. *Curr. Opin. Chem. Biol.* **2013**, *17* (4), 656–662.
- (10) Dodani, S. C.; He, Q.; Chang, C. J. A Turn-On Fluorescent Sensor for Detecting Nickel in Living Cells. *J. Am. Chem. Soc.* **2009**, *131* (50), 18020–18021.
- (11) Yang, Y.; Zhao, Q.; Feng, W.; Li, F. Luminescent Chemodosimeters for Bioimaging. *Chem. Rev.* **2013**, *113* (1), 192–270.
- (12) Chan, J.; Dodani, S. C.; Chang, C. J. Reaction-Based Small-Molecule Fluorescent Probes for Chemoselective Bioimaging. *Nat. Chem.* **2012**, *4* (12), 973–984.
- (13) Lee, M. H.; Kim, J. S.; Sessler, J. L. Small Molecule-Based Ratiometric Fluorescence Probes for Cations, Anions, and Biomolecules. *Chem. Soc. Rev.* **2015**, *44* (13), 4185–4191.
- (14) Tang, Y.; Lee, D.; Wang, J.; Li, G.; Yu, J.; Lin, W.; Yoon, J. Development of Fluorescent Probes Based on Protection–Deprotection of the Key Functional Groups for Biological Imaging. *Chem. Soc. Rev.* **2015**, *44* (15), 5003–5015.
- (15) Brewer, T. F.; Chang, C. J. An Aza-Cope Reactivity-Based Fluorescent Probe for Imaging Formaldehyde in Living Cells. *J. Am. Chem. Soc.* **2015**, *137* (34), 10886–10889.
- (16) Bruemmer, K. J.; Brewer, T. F.; Chang, C. J. Fluorescent Probes for Imaging Formaldehyde in Biological Systems. *Curr. Opin. Chem. Biol.* **2017**, *39*, 17–23.

- (17) Miller, E. W.; Albers, A. E.; Pralle, A.; Isacoff, E. Y.; Chang, C. J. Boronate-Based Fluorescent Probes for Imaging Cellular Hydrogen Peroxide. *J. Am. Chem. Soc.* **2005**, *127* (47), 16652–16659.
- (18) Lippert, A. R.; Van de Bittner, G. C.; Chang, C. J. Boronate Oxidation as a Bioorthogonal Reaction Approach for Studying the Chemistry of Hydrogen Peroxide in Living Systems. *Acc. Chem. Res.* **2011**, *44* (9), 793–804.
- (19) Brewer, T. F.; Garcia, F. J.; Onak, C. S.; Carroll, K. S.; Chang, C. J. Chemical Approaches to Discovery and Study of Sources and Targets of Hydrogen Peroxide Redox Signaling Through NADPH Oxidase Proteins. *Annu. Rev. Biochem.* **2015**, *84* (1), 765–790.
- (20) Miller, E. W.; Tulyathan, O.; Isacoff, E. Y.; Chang, C. J. Molecular Imaging of Hydrogen Peroxide Produced for Cell Signaling. *Nat. Chem. Biol.* **2007**, *3* (5), 263–267.
- (21) Chang, M. C. Y.; Pralle, A.; Isacoff, E. Y.; Chang, C. J. A Selective, Cell-Permeable Optical Probe for Hydrogen Peroxide in Living Cells. *J. Am. Chem. Soc.* **2004**, *126* (47), 15392–15393.
- (22) Lippert, A. R.; New, E. J.; Chang, C. J. Reaction-Based Fluorescent Probes for Selective Imaging of Hydrogen Sulfide in Living Cells. *J. Am. Chem. Soc.* **2011**, *133* (26), 10078–10080.
- (23) S. Lin, V.; Chen, W.; Xian, M.; J. Chang, C. Chemical Probes for Molecular Imaging and Detection of Hydrogen Sulfide and Reactive Sulfur Species in Biological Systems. *Chem. Soc. Rev.* **2015**, *44* (14), 4596–4618.
- (24) Lin, V. S.; Lippert, A. R.; Chang, C. J. Cell-Trappable Fluorescent Probes for Endogenous Hydrogen Sulfide Signaling and Imaging H<sub>2</sub>O<sub>2</sub>-Dependent H<sub>2</sub>S Production. *Proc. Natl. Acad. Sci.* **2013**, *110* (18), 7131–7135.
- (25) Kam, C. M.; Abuelyaman, A. S.; Li, Z.; Hudig, D.; Powers, J. C. Biotinylated Isocoumarins, New Inhibitors and Reagents for Detection, Localization, and Isolation of Serine Proteases. *Bioconjug. Chem.* **1993**, *4* (6), 560–567.
- (26) Abuelyaman, A. S.; Hudig, D.; Woodard, S. L.; Powers, J. C. Fluorescent Derivatives of Diphenyl [1-(N-Peptidylamino)Alkyl]Phosphonate Esters: Synthesis and Use in the Inhibition and Cellular Localization of Serine Proteases. *Bioconjug. Chem.* **1994**, *5* (5), 400–405.
- (27) 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.
- (28) Mann, M.; Jensen, O. N. Proteomic Analysis of Post-Translational Modifications. *Nat. Biotechnol.* **2003**, *21* (3), 255.
- (29) Krysiak, J.; Breinbauer, R. Activity-Based Protein Profiling for Natural Product Target Discovery. In *Activity-Based Protein Profiling*; Sieber, S. A., Ed.; Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; pp 43–84.
- (30) Festa, R. A.; Thiele, D. J. Copper: An Essential Metal in Biology. *Curr. Biol.* **2011**, *21* (21), R877–R883.
- (31) Madsen, E.; Gitlin, J. D. Copper and Iron Disorders of the Brain. *Annu. Rev. Neurosci.* **2007**, *30* (1), 317–337.

- (32) Go, Y.-M.; Jones, D. P. Redox Compartmentalization in Eukaryotic Cells. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2008**, *1780* (11), 1273–1290.
- (33) Yang, L.; McRae, R.; Henary, M. M.; Patel, R.; Lai, B.; Vogt, S.; Fahrni, C. J. Imaging of the Intracellular Topography of Copper with a Fluorescent Sensor and by Synchrotron X-Ray Fluorescence Microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (32), 11179–11184.
- (34) Chaudhry, A. F.; Mandal, S.; Hardcastle, K. I.; Fahrni, C. J. High-Contrast Cu(I)-Selective Fluorescent Probes Based on Synergistic Electronic and Conformational Switching. *Chem. Sci.* **2011**, *2* (6), 1016–1024.
- (35) Morgan, M. T.; Bagchi, P.; Fahrni, C. J. Designed To Dissolve: Suppression of Colloidal Aggregation of Cu(I)-Selective Fluorescent Probes in Aqueous Buffer and In-Gel Detection of a Metallochaperone. *J. Am. Chem. Soc.* **2011**, *133* (40), 15906–15909.
- (36) Morgan, M. T.; McCallum, A. M.; Fahrni, C. J. Rational Design of a Water-Soluble, Lipid-Compatible Fluorescent Probe for Cu(I) with Sub-Part-per-Trillion Sensitivity. *Chem. Sci.* **2016**, *7* (2), 1468–1473.
- (37) Achard, M. E. S.; Stafford, S. L.; Bokil, N. J.; Chartres, J.; Bernhardt, P. V.; Schembri, M. A.; Sweet, M. J.; McEwan, A. G. Copper Redistribution in Murine Macrophages in Response to Salmonella Infection. *Biochem. J.* **2012**, *444* (1), 51–57.
- (38) Dodani, S. C.; Firl, A.; Chan, J.; Nam, C. I.; Aron, A. T.; Onak, C. S.; Ramos-Torres, K. M.; Paek, J.; Webster, C. M.; Feller, M. B.; et al. Copper Is an Endogenous Modulator of Neural Circuit Spontaneous Activity. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 16280–16285.
- (39) Xiao, T.; Ackerman, C. M.; Carroll, E. C.; Jia, S.; Hoagland, A.; Chan, J.; Thai, B.; Liu, C. S.; Isacoff, E. Y.; Chang, C. J. Copper Regulates Rest-Activity Cycles through the Locus Coeruleus-Norepinephrine System. *Nat. Chem. Biol.* **2018**, *14* (7), 655.
- (40) Krishnamoorthy, L.; Cotruvo, J. A. J.; Chan, J.; Kaluarachchi, H.; Muchenditsi, A.; Pendyala, V. S.; Jia, S.; Aron, A. T.; Vander Wal, M. N.; Guan, T.; et al. Copper Regulates Cyclic AMP-Dependent Lipolysis. *Nat. Chem. Biol.* **2016**, *12*, 586–592.
- (41) Santo, C. E.; Lam, E. W.; Elowsky, C. G.; Quaranta, D.; Domaille, D. W.; Chang, C. J.; Grass, G. Bacterial Killing by Dry Metallic Copper Surfaces. *Appl. Environ. Microbiol.* **2011**, *77* (3), 794–802.
- (42) Ladomersky, E.; Khan, A.; Shanbhag, V.; Cavet, J. S.; Chan, J.; Weisman, G. A.; Petris, M. J. Host and Pathogen Copper-Transporting P-Type ATPases Function Antagonistically during Salmonella Infection. *Infect. Immun.* **2017**, IAI.00351-17.
- (43) Quaranta, D.; Krans, T.; Santo, C. E.; Elowsky, C. G.; Domaille, D. W.; Chang, C. J.; Grass, G. Mechanisms of Contact-Mediated Killing of Yeast Cells on Dry Metallic Copper Surfaces. *Appl. Environ. Microbiol.* **2011**, *77* (2), 416–426.
- (44) Cusick, K. D.; Minkin, S. C.; Dodani, S. C.; Chang, C. J.; Wilhelm, S. W.; Sayler, G. S. Inhibition of Copper Uptake in Yeast Reveals the Copper Transporter Ctr1p As a Potential Molecular Target of Saxitoxin. *Environ. Sci. Technol.* **2012**, *46* (5), 2959–2966.

- (45) Beaudoin, J.; Ioannoni, R.; López-Maury, L.; Bähler, J.; Ait-Mohand, S.; Guérin, B.; Dodani, S. C.; Chang, C. J.; Labbé, S. Mfc1 Is a Novel Forespore Membrane Copper Transporter in Meiotic and Sporulating Cells. *J. Biol. Chem.* **2011**, *286* (39), 34356–34372.
- (46) Bernal, M.; Casero, D.; Singh, V.; Wilson, G. T.; Grande, A.; Yang, H.; Dodani, S. C.; Pellegrini, M.; Huijser, P.; Connolly, E. L.; et al. Transcriptome Sequencing Identifies SPL7-Regulated Copper Acquisition Genes FRO4/FRO5 and the Copper Dependence of Iron Homeostasis in Arabidopsis. *Plant Cell* **2012**, *24* (2), 738–761.
- (47) Chun, H.; Sharma, A. K.; Lee, J.; Chan, J.; Jia, S.; Kim, B.-E. The Intestinal Copper Exporter CUA-1 Is Required for Systemic Copper Homeostasis in Caenorhabditis Elegans. *J. Biol. Chem.* **2017**, *292* (1), 1–14.
- (48) Schrag, M.; Crofton, A.; Zabel, M.; Jiffry, A.; Kirsch, D.; Dickson, A.; Mao, X. W.; Vinters, H. V.; Domaille, D. W.; Chang, C. J.; et al. Effect of Cerebral Amyloid Angiopathy on Brain Iron, Copper, and Zinc in Alzheimer’s Disease. *J. Alzheimers Dis.* **2011**, *24* (1), 137–149.
- (49) Öhrvik, H.; Nose, Y.; Wood, L. K.; Kim, B.-E.; Gleber, S.-C.; Ralle, M.; Thiele, D. J. Ctr2 Regulates Biogenesis of a Cleaved Form of Mammalian Ctr1 Metal Transporter Lacking the Copper- and Cisplatin-Binding Ecto-Domain. *Proc. Natl. Acad. Sci.* **2013**, *110* (46), E4279–E4288.
- (50) Huang, C. P.; Fofana, M.; Chan, J.; Chang, C. J.; Howell, S. B. Copper Transporter 2 Regulates Intracellular Copper and Sensitivity to Cisplatin. *Metallomics* **2014**, *6* (3), 654–661.
- (51) Polishchuk, E. V.; Concilli, M.; Iacobacci, S.; Chesi, G.; Pastore, N.; Piccolo, P.; Paladino, S.; Baldantoni, D.; van IJendoorn, S. C. D.; Chan, J.; et al. Wilson Disease Protein ATP7B Utilizes Lysosomal Exocytosis to Maintain Copper Homeostasis. *Dev. Cell* **2014**, *29* (6), 686–700.
- (52) Elwell, C. E.; Gagnon, N. L.; Neisen, B. D.; Dhar, D.; Spaeth, A. D.; Yee, G. M.; Tolman, W. B. Copper–Oxygen Complexes Revisited: Structures, Spectroscopy, and Reactivity. *Chem. Rev.* **2017**, *117* (3), 2059–2107.
- (53) Liu, J. J.; Diaz, D. E.; Quist, D. A.; Karlin, K. D. Copper(I)-Dioxygen Adducts and Copper Enzyme Mechanisms. *Isr. J. Chem.* **2016**, *56* (9–10), 738–755.
- (54) Itoh, S. Developing Mononuclear Copper–Active-Oxygen Complexes Relevant to Reactive Intermediates of Biological Oxidation Reactions. *Acc. Chem. Res.* **2015**, *48* (7), 2066–2074.
- (55) Taki, M.; Iyoshi, S.; Ojida, A.; Hamachi, I.; Yamamoto, Y. Development of Highly Sensitive Fluorescent Probes for Detection of Intracellular Copper(I) in Living Systems. *J. Am. Chem. Soc.* **2010**, *132* (17), 5938–5939.
- (56) Taki, M.; Akaoka, K.; Mitsui, K.; Yamamoto, Y. A Mitochondria-Targeted Turn-on Fluorescent Probe Based on a Rhodol Platform for the Detection of Copper(I). *Org. Biomol. Chem.* **2014**, *12* (27), 4999–5005.

- (57) Hu, Z.; Hu, J.; Wang, H.; Zhang, Q.; Zhao, M.; Brommesson, C.; Tian, Y.; Gao, H.; Zhang, X.; Uvdal, K. A TPA-Caged Precursor of (Imino)Coumarin for “Turn-on” Fluorogenic Detection of Cu<sup>+</sup>. *Anal. Chim. Acta* **2016**, *933*, 189–195.
- (58) Heffern, M. C.; Park, H. M.; Au-Yeung, H. Y.; Bittner, G. C. V. de; Ackerman, C. M.; Stahl, A.; Chang, C. J. In Vivo Bioluminescence Imaging Reveals Copper Deficiency in a Murine Model of Nonalcoholic Fatty Liver Disease. *Proc. Natl. Acad. Sci.* **2016**, 1613628113.
- (59) Dujols, V.; Ford, F.; Czarnik, A. W. A Long-Wavelength Fluorescent Chemodosimeter Selective for Cu(II) Ion in Water. *J. Am. Chem. Soc.* **1997**, *119* (31), 7386–7387.
- (60) Yuan, L.; Lin, W.; Chen, B.; Xie, Y. Development of FRET-Based Ratiometric Fluorescent Cu<sup>2+</sup> Chemodosimeters and the Applications for Living Cell Imaging. *Org. Lett.* **2012**, *14* (2), 432–435.
- (61) Fan, J.; Zhan, P.; Hu, M.; Sun, W.; Tang, J.; Wang, J.; Sun, S.; Song, F.; Peng, X. A Fluorescent Ratiometric Chemodosimeter for Cu<sup>2+</sup> Based on TBET and Its Application in Living Cells. *Org. Lett.* **2013**, *15* (3), 492–495.
- (62) Hu, Z.; Hu, J.; Cui, Y.; Wang, G.; Zhang, X.; Uvdal, K.; Gao, H.-W. A Facile “Click” Reaction to Fabricate a FRET-Based Ratiometric Fluorescent Cu<sup>2+</sup> Probe. *J Mater Chem B* **2014**, *2* (28), 4467–4472.
- (63) Veal, E. A.; Day, A. M.; Morgan, B. A. Hydrogen Peroxide Sensing and Signaling. *Mol. Cell* **2007**, *26* (1), 1–14.
- (64) Winterbourn, C. C. Reconciling the Chemistry and Biology of Reactive Oxygen Species. *Nat. Chem. Biol.* **2008**, *4* (5), 278–286.
- (65) Dickinson, B. C.; Chang, C. J. Chemistry and Biology of Reactive Oxygen Species in Signaling or Stress Responses. *Nat. Chem. Biol.* **2011**, *7* (8), 504–511.
- (66) Lin, M. T.; Beal, M. F. Mitochondrial Dysfunction and Oxidative Stress in Neurodegenerative Diseases. *Nature* **2006**, *443* (7113), 787.
- (67) Uttara, B.; Singh, A. V.; Zamboni, P.; Mahajan, R. T. Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options. *Curr. Neuropharmacol.* **2009**, *7* (1), 65–74.
- (68) Kim, G. H.; Kim, J. E.; Rhie, S. J.; Yoon, S. The Role of Oxidative Stress in Neurodegenerative Diseases. *Exp. Neurobiol.* **2015**, *24* (4), 325–340.
- (69) Schieber, M.; Chandel, N. S. ROS Function in Redox Signaling and Oxidative Stress. *Curr. Biol. CB* **2014**, *24* (10), R453–R462.
- (70) Stone, J. R.; Yang, S. Hydrogen Peroxide: A Signaling Messenger. *Antioxid. Redox Signal.* **2006**, *8* (3–4), 243–270.
- (71) Paulsen, C. E.; Carroll, K. S. Orchestrating Redox Signaling Networks through Regulatory Cysteine Switches. *ACS Chem. Biol.* **2010**, *5* (1), 47–62.
- (72) Antunes, F.; Brito, P. M. Quantitative Biology of Hydrogen Peroxide Signaling. *Redox Biol.* **2017**, *13*, 1–7.

- (73) Kuivila, H. G.; Wiles, R. A. Electrophilic Displacement Reactions. VII. Catalysis by Chelating Agents in the Reaction between Hydrogen Peroxide and Benzeneboronic Acid<sub>1,2,3</sub>. *J. Am. Chem. Soc.* **1955**, *77* (18), 4830–4834.
- (74) Kuivila, H. G.; Armour, A. G. Electrophilic Displacement Reactions. IX. Effects of Substituents on Rates of Reactions between Hydrogen Peroxide and Benzeneboronic Acid<sub>1-3</sub>. *J. Am. Chem. Soc.* **1957**, *79* (21), 5659–5662.
- (75) Dickinson, B. C.; Peltier, J.; Stone, D.; Schaffer, D. V.; Chang, C. J. Nox2 Redox Signaling Maintains Essential Cell Populations in the Brain. *Nat. Chem. Biol.* **2011**, *7* (2), 106–112.
- (76) Dickinson, B. C.; Chang, C. J. A Targetable Fluorescent Probe for Imaging Hydrogen Peroxide in the Mitochondria of Living Cells. *J. Am. Chem. Soc.* **2008**, *130* (30), 9638–9639.
- (77) Dickinson, B. C.; Tang, Y.; Chang, Z.; Chang, C. J. A Nuclear-Localized Fluorescent Hydrogen Peroxide Probe for Monitoring Sirtuin-Mediated Oxidative Stress Responses In Vivo. *Chem. Biol.* **2011**, *18* (8), 943–948.
- (78) Carroll, V.; Michel, B. W.; Blecha, J.; VanBrocklin, H.; Keshari, K.; Wilson, D.; Chang, C. J. A Boronate-Caged [<sup>18</sup>F]FLT Probe for Hydrogen Peroxide Detection Using Positron Emission Tomography. *J. Am. Chem. Soc.* **2014**, *136* (42), 14742–14745.
- (79) Yik-Sham Chung, C.; Timblin, G. A.; Saijo, K.; Chang, C. J. Versatile Histochemical Approach to Detection of Hydrogen Peroxide in Cells and Tissues Based on Puromycin Staining. *J. Am. Chem. Soc.* **2018**, *140* (19), 6109–6121.
- (80) Hitomi, Y.; Takeyasu, T.; Funabiki, T.; Kodera, M. Detection of Enzymatically Generated Hydrogen Peroxide by Metal-Based Fluorescent Probe. *Anal. Chem.* **2011**, *83* (24), 9213–9216.
- (81) Hitomi, Y.; Takeyasu, T.; Kodera, M. Iron Complex-Based Fluorescent Probes for Intracellular Hydrogen Peroxide Detection. *Chem. Commun.* **2013**, *49* (85), 9929–9931.
- (82) Song, D.; Lim, J. M.; Cho, S.; Park, S.-J.; Cho, J.; Kang, D.; Rhee, S. G.; You, Y.; Nam, W. A Fluorescence Turn-on H<sub>2</sub>O<sub>2</sub> Probe Exhibits Lysosome-Localized Fluorescence Signals. *Chem. Commun.* **2012**, *48* (44), 5449–5451.
- (83) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative Reactivity Profiling Predicts Functional Cysteines in Proteomes. *Nature* **2010**, *468* (7325), 790–795.
- (84) Sanman, L. E.; Bogoyo, M. Activity-Based Profiling of Proteases. *Annu. Rev. Biochem.* **2014**, *83*, 249–273.
- (85) Yang, P.; Liu, K. Activity-Based Protein Profiling: Recent Advances in Probe Development and Applications. *ChemBioChem* **2015**, *16* (5), 712–724.
- (86) Pan, J.; Carroll, K. S. Chemical Biology Approaches to Study Protein Cysteine Sulfenylation. *Biopolymers* **2014**, *101* (2), 165–172.
- (87) Akter, S.; Fu, L.; Jung, Y.; Conte, M. L.; Lawson, J. R.; Lowther, W. T.; Sun, R.; Liu, K.; Yang, J.; Carroll, K. S. Chemical Proteomics Reveals New Targets of Cysteine Sulfenic Acid Reductase. *Nat. Chem. Biol.* **2018**, *14* (11), 995.

- (88) Ward, C. C.; Kleinman, J. I.; Nomura, D. K. NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem. Biol.* **2017**, *12* (6), 1478–1483.
- (89) Roberts, A. M.; Ward, C. C.; Nomura, D. K. Activity-Based Protein Profiling for Mapping and Pharmacologically Interrogating Proteome-Wide Ligandable Hotspots. *Curr. Opin. Biotechnol.* **2017**, *43*, 25–33.
- (90) Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Toste, F. D.; et al. Redox-Based Reagents for Chemoselective Methionine Bioconjugation. *Science* **2017**, *355* (6325), 597–602.

## Chapter 2

### **Development of a color palette of fluorescent copper sensors through heteroatom substitution at rhodol cores**

Portions of this work were published in the following scientific journal:

Jia, S.\*; Ramos-Torres\*, K. M.; Kolemen, S.; Ackerman, C. M.; Chang, C. J. Tuning the Color Palette of Fluorescent Copper Sensors through Systematic Heteroatom Substitution at Rhodol Cores. *ACS Chem. Biol.* **2018**, *13* (7), 1844–1852.

(\* Denotes equal contribution.)

## Abstract

Copper is an essential nutrient for sustaining life, and emerging data has expanded the roles of this metal in biology from its canonical functions as a static enzyme cofactor to dynamic functions as a transition metal signal. At the same time, loosely-bound, labile copper pools can trigger oxidative stress and damaging events that are detrimental if misregulated. The signal/stress dichotomy of copper motivates the development of new chemical tools to study its spatial and temporal distributions in native biological contexts such as living cells. Here we report a family of fluorescent copper sensors built upon carbon-, silicon- and phosphorus-substituted rhodol dyes that enable systematic tuning of excitation/emission colors from the orange to near-infrared. These probes can detect changes in labile copper levels in living cells upon copper supplementation and/or depletion. We demonstrate the ability of the carbon-rhodol based congener, Copper Carbo Fluor 1 (CCF1), to identify elevations in labile copper pools in the *Atp7a*<sup>-/-</sup> fibroblast cell model of the genetic copper disorder Menkes disease. Moreover, we showcase the utility of the red-emitting phosphorus-rhodol based dye Copper Phosphorus Fluor 1 (CPF1) in dual-color, dual-analyte imaging experiments with the green-emitting calcium indicator Calcium Green-1 to enable simultaneous detection of fluctuations in copper and calcium pools in living cells. The results provide a starting point for advancing tools to study the contributions of copper to health and disease and for exploiting the rapidly growing palette of heteroatom-substituted xanthene dyes to rationally tune the optical properties of fluorescent indicators for other biologically important analytes.

## 2.1 Introduction

Copper is an indispensable element for life.<sup>1,2</sup> The redox capacity of this transition metal is widely exploited as a catalytic and structural cofactor in proteins that spans a diverse array of fundamental processes including oxygen transport, respiration and metabolism, cell growth and differentiation, and signal transduction.<sup>1-5</sup> Conversely, copper dysregulation can lead to cellular malfunctions resulting from the aberrant production of reactive oxygen species (ROS) and subsequent oxidative damage to proteins, lipids and DNA/RNA.<sup>6,7</sup> Indeed, organisms have evolved cellular machineries to carefully regulate copper uptake, transport, storage and excretion,<sup>8-13</sup> and irregular deviations from this delicate balance have been linked to pathogenic states including neurodegenerative disorders like Alzheimer's,<sup>14-17</sup> Parkinson's,<sup>18</sup> and Huntington's<sup>19</sup> diseases and familial amyotrophic lateral sclerosis,<sup>20-23</sup> metabolic disorders such as diabetes and obesity,<sup>24-26</sup> and genetic disorders like Menkes<sup>27,28</sup> and Wilson's<sup>29-31</sup> diseases. In addition, emerging data from our laboratory and others have revealed that dynamic copper fluxes can also regulate essential physiological functions<sup>32,33</sup> spanning metabolic processes such as lipolysis,<sup>5</sup> neural processes such as spontaneous activity,<sup>34</sup> neuronal calcium signaling,<sup>35</sup> and olfaction,<sup>36,37</sup> as well as kinase pathways involved in signaling and tumorigenesis.<sup>3,4</sup>

The broad contributions of copper to health and disease motivate the development of technologies to help disentangle its disparate physiological and pathological effects. In this

context, the use of fluorescent sensors for visualizing metal fluxes has proven to be a potentially powerful strategy for studying these elements in their native biological contexts with spatial and temporal resolution.<sup>32,33,38-44</sup> This approach is well-suited for the simultaneous study of multiple biological events using different probes as long as spectral overlap between chromophores is sufficiently minimized.<sup>39,45-48</sup> With specific regard to copper, a growing toolbox of small-molecule<sup>44,49-51</sup> and macromolecular<sup>52-55</sup> fluorescent probes for this essential metal have emerged for use in cells and more complex biological specimens. Moreover, application of these chemical reagents in conjunction with other direct imaging techniques as well as supporting biochemical and cell biology studies have identified new copper biology in bacteria,<sup>56,57</sup> yeast,<sup>58-60</sup> plant,<sup>61</sup> worm<sup>62</sup> and mammalian<sup>63-66</sup> models. Included are examples of activity-dependent neuronal copper translocation,<sup>34</sup> copper-dependent antimicrobial behavior,<sup>57,67,68</sup> hyperaccumulation of copper in cuprosome organelles triggered by zinc deficiency,<sup>69</sup> and copper-regulated lipolysis.<sup>5</sup> Despite this progress, the base fluorophores for fluorescent copper detection have relied on a variety of scaffolds, ranging from UV-excitable pyrazoline<sup>70-72</sup> and naphthalene,<sup>73</sup> visible-wavelength BODIPY<sup>35,74,75</sup> and rhodol,<sup>34</sup> to far-red silicon rhodol<sup>5</sup> and near-infrared cyanine dyes,<sup>76,77</sup> which presents a unique challenge for optimizing the combination of copper-selective recognition elements exhibiting high metal and redox specificity along with dye platforms allowing for fine control of excitation/emission color profiles.

Against this backdrop, we sought to pursue an alternative strategy in which rational tuning of probe excitation/emission colors of fluorescent copper sensors could be achieved independently of the metal-responsive moiety. In particular, we were inspired by elegant studies that have greatly expanded the optical spectral window of xanthene-based fluorophores like fluorescein, rhodamine, and rhodol<sup>78-90</sup> and thus turned our attention to reports on substitution of the endocyclic oxygen atom in the xanthene core of green-fluorescent fluorescein by carbon, silicon or phosphorus to provide new red-shifted fluorophores with emission profiles from the orange to near-infrared,<sup>82,86,88</sup> yielding improved tissue penetration and minimized sample photodamage.<sup>91</sup> We now report the development of the Copper Xanthene Fluor (CXF) family of copper-responsive fluorescent indicators based on carbon (Copper Carbo Fluor 1, CCF1), silicon (Copper Silicon Fluor 1, CSF1), and phosphorus (Copper Phosphorus Fluor 1, CPF1) analogs of the rhodol-based Copper Fluor series,<sup>34</sup> along with matched control dyes that possess the same fluorophore scaffolds but features isosteric sulfur-to-carbon substitutions along the receptor portion to render them non-responsive to copper. We establish the utility of these CXF dyes to visualize changes in labile copper pools in HEK 293T cells with copper supplementation and/or chelation. Moreover, the CCF1 probe is capable of distinguishing elevations in labile copper levels in the Menkes *Atp7a*<sup>-/-</sup> fibroblast cell model from normal levels in wildtype controls. Finally, the most red-shifted probe of the series, CPF1, enables dual-color, dual-analyte imaging of copper and calcium pools in conjunction with the green-emitting dye Calcium Green-1.

## 2.2 Methods

### 2.2.1 General materials and synthetic method

All reactions utilizing air- or moisture-sensitive reagents were performed in dried glassware under an atmosphere of dry N<sub>2</sub>. THF used for anhydrous reactions was dried and stored over 4 Å molecular sieves. 3,6-dihydroxy-10,10-dimethylanthracen-9(10*H*)-one (**1a**),<sup>86</sup> 3,7-dihydroxy-5,5-dimethyldibenzo[*b,e*]silin-10(5*H*)-one (**1b**)<sup>82</sup> and 3,7-dihydroxy-5-phenyl-10*H*-acridophosphin-10-one 5-oxide (**1c**)<sup>88</sup>, *N*-(4-Bromo-3-(trifluoromethyl)benzyl)-*N,N*-bis(2-((2-(ethylthio)ethyl)thio)ethyl)amine (**5**) and *N*-(4-Bromo-3-(trifluoromethyl)benzyl)-*N,N*-dioctylamine (**6**) were synthesized according to literature procedure.<sup>34</sup> All other reagents were purchased from Sigma-Aldrich (St. Louis, MO) and used without further purification. <sup>1</sup>H NMR, <sup>13</sup>C NMR and <sup>31</sup>P NMR spectra were collected in CDCl<sub>3</sub> or CD<sub>3</sub>OD (Cambridge Isotope Laboratories, Cambridge, MA) at 25°C on AVB-400, AVQ-400 or DRX-500 spectrometers at the College of Chemistry NMR Facility at the University of California, Berkeley. All chemical shifts are reported in the standard δ notation of ppm relative to residual solvent peak (CDCl<sub>3</sub> δH=7.26, δC=77.20; CD<sub>3</sub>OD δH=3.31, δC=49.00). Splitting patterns are indicated as follows: br, broad; s, singlet; d, doublet; t, triplet; m, multiplet; dd, doublet of doublets. Low-resolution electrospray mass spectral analyses were carried out using a LC-MS (Advion expression-L Compact Mass Spectrometer). Low resolution and high resolution electron ionization mass spectral analyses were carried out at the College of Chemistry Mass Spectrometry Facility at the University of California, Berkeley. High resolution mass spectral analyses (ESI-MS) were carried out at LBNL Catalysis Facility at the Lawrence Berkeley National Laboratory (Berkeley Lab) using a UHPLC-TOF (PerkinElmer AxION® 2 TOF MS).

### 2.2.2 Synthesis of CXF1 and Ctrl-CXF1

**General procedure A: synthesis of bistriflate 2.** A solution of **1**, PhNTf<sub>2</sub> and DIPEA in anhydrous DMF was stirred at room temperature overnight. The reaction was diluted with H<sub>2</sub>O, transferred into a separatory funnel and extracted with EtOAc. The organic layer was washed with H<sub>2</sub>O (×4), brine dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated under reduced pressure. The resultant residue was purified using flash chromatography (silica gel) to give the product.

#### **9,9-Dimethyl-10-oxo-9,10-dihydroanthracene-2,7-diyl bis(trifluoromethanesulfonate) (2a).**

Following general procedure A, **1a** (4.0 g, 15.90 mmol), PhNTf<sub>2</sub> (17.04 g, 47.69 mmol) and DIPEA (16.61 mL, 95.38 mmol) was reacted in DMF (60 mL) to provide **2a** (5.69 g, 69%) as an off-white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.46 (d, *J* = 8.8 Hz, 2H), 7.58 (d, *J* = 2.3 Hz, 2H), 7.38 (dd, *J* = 8.8, 2.3 Hz, 2H), 1.78 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 180.91, 153.37, 152.68, 130.95, 129.43, 120.58, 120.01, 117.32, 38.78, 33.00. HRMS (EI<sup>+</sup>) *m/z* calcd 517.9929, found 517.9922 for C<sub>18</sub>H<sub>12</sub>F<sub>6</sub>O<sub>7</sub>S<sub>2</sub><sup>+</sup> (M<sup>+</sup>).

**5,5-Dimethyl-10-oxo-5,10-dihydrodibenzo[b,e]siline-3,7-diyl bis(trifluoromethanesulfonate) (2b).** Following general procedure A, **1b** (553 mg, 2.03 mmol), PhNTf<sub>2</sub> (2.18g, 6.10 mmol) and DIPEA (2.12 mL, 12.2 mol) was reacted in DMF (6 mL) to provide **2b** (1.09g, 100%) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 8.53 (d, 2H, J=8.8 Hz), 7.54 (d, 2H, J=2.5 Hz), 7.48 (dd, 2H, J=2.6, 8.8 Hz), 0.58 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): 185.06, 152.28, 142.15, 140.06, 133.24, 130.15, 125.58, 123.43, -1.61. HRMS (EI<sup>+</sup>): m/z calcd 533.9698, found 533.9697 for C<sub>17</sub>H<sub>12</sub>F<sub>6</sub>O<sub>7</sub>S<sub>2</sub>Si<sup>+</sup> (M<sup>+</sup>).

**5-Oxido-10-oxo-5-phenyl-10H-acridophosphine-3,7-diyl bis(trifluoromethanesulfonate) (2c).** Following general procedure A, **1c** (190 mg, 0.56 mmol), PhNTf<sub>2</sub> (605.5 mg, 1.68 mmol) and DIPEA (0.57 mL, 3.36 mmol) was reacted in DMF (2.5 mL) to provide **2c** (269 mg, 80%) as a white solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.57 (dd, J = 8.8, 5.1 Hz, 2H), 7.92 (dd, J = 13.1, 2.6 Hz, 2H), 7.68 (dd, J = 8.8, 2.5 Hz, 2H), 7.58 – 7.42 (m, 5H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): δ 180.1, 180.0, 153.0, 152.9, 136.8, 135.9, 134.9, 134.8, 133.1, 132.8, 132.7, 131.3, 130.7, 130.6, 130.2, 129.5, 129.4, 126.1, 123.9, 120.2, 117.0. <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ 2.4. LRMS (EI<sup>+</sup>): m/z calcd. 600, found 600 for C<sub>21</sub>H<sub>11</sub>F<sub>6</sub>O<sub>8</sub>PS<sub>2</sub> (M<sup>+</sup>).

**General procedure B: synthesis of morpholino-monotriflate 3.** Morpholine was added to a solution of **2** in DMSO. The solution was stirred overnight at 90°C. After cooling to room temperature, the reaction was diluted with H<sub>2</sub>O and extracted with EtOAc. The organic layer was washed with H<sub>2</sub>O (×4), brine dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated under reduced pressure. The crude residue was purified using flash chromatography (silica gel) to give unreacted starting material and the product.

**9,9-Dimethyl-7-morpholino-10-oxo-9,10-dihydroanthracen-2-yl trifluoromethanesulfonate (3a).** Following general procedure B, **2a** (1.13 g, 2.18 mmol) and morpholine (188.1 μL, 2.18 mmol) was reacted in DMSO (10.9 mL) to provide **3a** (0.49 g, 49%) as a yellow solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.44 (d, J = 8.7 Hz, 1H), 8.26 (d, J = 8.8 Hz, 1H), 7.53 (d, J = 2.4 Hz, 1H), 7.31 (dd, J = 8.7, 2.4 Hz, 1H), 6.96 (d, J = 7.8 Hz, 2H), 3.89 (t, J = 4.7 Hz, 4H), 3.39 (t, J = 4.8 Hz, 4H), 1.72 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 180.38, 154.67, 152.81, 152.33, 151.76, 130.11, 129.91, 129.60, 120.90, 120.20, 119.45, 117.00, 113.49, 110.01, 66.36, 47.29, 38.22, 32.91. LRMS (ESI<sup>+</sup>): m/z calcd 456.1, found 456.0 for C<sub>21</sub>H<sub>21</sub>F<sub>3</sub>NO<sub>5</sub>S<sup>+</sup> (M+H<sup>+</sup>).

**5,5-Dimethyl-7-morpholino-10-oxo-5,10-dihydrodibenzo[b,e]siline-3-yl trifluoromethanesulfonate (3b).** Following general procedure B, **2b** (570 mg, 1.07 mmol) and morpholine (92.2 μL, 1.07 mmol) was reacted in DMSO (2 mL) to provide **3b** (227 mg, 45%), as a yellow solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 8.54 (d, 1H, J=8.8 Hz), 8.37 (d, 1H, J=9.0 Hz), 7.49 (d, 1H, J=2.6 Hz), 7.42 (dd, 1H, J=2.6, 8.8 Hz), 7.05 (dd, 1H, J=2.7, 9.0 Hz), 7.00 (d, 1H, J=2.6 Hz), 3.89 (t, 4H, J=5.0 Hz), 3.38 (t, 4H, J=5.0 Hz), 0.51 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): δ 184.9, 153.1, 151.8, 142.6, 141.0, 140.4, 132.7, 132.4, 131.1, 125.2, 122.8, 120.5,

116.9, 116.0, 66.8, 47.4, -1.31. HRMS (ESI<sup>+</sup>): m/z calcd 472.0856, found 472.0849 for C<sub>20</sub>H<sub>21</sub>F<sub>3</sub>NO<sub>5</sub>Si<sup>+</sup> (M+H<sup>+</sup>).

### **7-Morpholino-5-oxido-10-oxo-5-phenyl-10H-acridophosphine-3-yl**

**trifluoromethanesulfonate (3c).** Following general procedure B, **2c** (269 mg, 0.45 mmol) and morpholine (39.2  $\mu$ L, 0.45 mmol) was reacted in DMSO (0.5 mL) to provide **3c** (147 mg, 61%) as a yellow, thick liquid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.52 (dd, *J* = 8.8, 5.1 Hz, 1H), 8.35 (dd, *J* = 9.1, 5.9 Hz, 1H), 7.88 (dd, *J* = 12.7, 2.6 Hz, 1H), 7.62 – 7.36 (m, 6H), 7.32 (dd, *J* = 15.1, 2.7 Hz, 1H), 7.12 (dd, *J* = 9.1, 2.7 Hz, 1H), 3.82 (t, *J* = 4.9 Hz, 4H), 3.45 – 3.33 (m, 4H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>-*d*):  $\delta$  179.6, 179.5, 153.9, 153.8, 152.3, 152.2, 137.2, 136.3, 135.8, 135.0, 134.0, 133.2, 132.4, 132.3, 132.1, 132.0, 130.6, 130.5, 129.1, 129.0, 125.4, 125.3, 125.2, 123.3, 123.2, 116.9, 114.1, 114.0, 66.3, 46.7. <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>)  $\delta$  3.9. LRMS (ESI<sup>+</sup>): m/z calcd 538.1, found 538.3 for C<sub>24</sub>H<sub>20</sub>F<sub>3</sub>NO<sub>6</sub>PS<sup>+</sup> (M+H<sup>+</sup>).

**General procedure C: Synthesis of morpholino-silyl ether.** A solution of **3** in dioxane was treated with a solution of Et<sub>4</sub>NOH in MeOH. The resultant solution was stirred at room temperature for 2 h. After the reaction was complete, the volatiles were removed under reduced pressure, and the residue was dissolved in CH<sub>2</sub>Cl<sub>2</sub>. The solution was cooled to 0 °C followed by treatment with imidazole and tertbutyldimethylsilyl chloride. After stirring for 4 h at room temperature, the reaction was diluted with CH<sub>2</sub>Cl<sub>2</sub> and washed with H<sub>2</sub>O. The organic layer was washed with brine, dried (Na<sub>2</sub>SO<sub>4</sub>), and concentrated under reduced pressure. The crude residue was purified using flash chromatography (silica gel) to give the product.

### **3-((Tert-Butyldimethylsilyloxy)-10,10-dimethyl-6-morpholinoanthracen-9(10H)-one (4a).**

Following general procedure C, **3a** (230 mg, 0.57 mmol) was reacted with Et<sub>4</sub>NOH (1.0 mL, 1.5 M in methanol, 1.5 mmol) in dioxane (5 mL) and imidazole (0.10 g, 1.53 mmol), *tert*-butyldimethylsilyl chloride (0.15 g, 1.02 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (12.8 mL) to give **4a** (200 mg, 92%) as a pale yellow solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.26 (dd, *J* = 8.7, 3.2 Hz, 2H), 7.02 (dd, *J* = 12.3, 2.4 Hz, 2H), 6.91 (ddd, *J* = 29.7, 8.8, 2.4 Hz, 2H), 3.89 (t, *J* = 4.9 Hz, 4H), 3.36 (t, *J* = 4.9 Hz, 4H), 1.68 (s, 6H), 1.01 (s, 9H), 0.26 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  181.51, 159.91, 154.22, 152.57, 152.13, 129.44, 129.27, 124.43, 122.10, 118.95, 117.34, 113.42, 110.60, 66.58, 47.78, 37.92, 33.18, 25.61, 18.25, -4.33. LRMS (ESI<sup>+</sup>): m/z calcd 438.2, found 438.1 for C<sub>26</sub>H<sub>36</sub>NO<sub>3</sub>Si<sup>+</sup> (M+H<sup>+</sup>).

### **3-((tert-Butyldimethylsilyloxy)-5,5-dimethyl-7-morpholinodibenzo[b,e]silin-10(5H)-one**

**(4b).** Following general procedure C, **3b** (149 mg, 0.32 mmol) was reacted with Et<sub>4</sub>NOH (0.42 mL, 1.5 M in methanol, 0.63 mmol) in dioxane (10 mL) and imidazole (215 mg, 3.16 mmol), *tert*-butyldimethylsilyl chloride (143 mg, 0.95 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (15 mL) to give **4b** (100 mg, 70%) as a light yellow solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  8.39 (d, *J* = 8.4 Hz, 2H), 7.05–6.98 (m, 4H), 3.89 (t, *J* = 4.8 Hz, 4H), 3.36 (t, *J* = 4.8 Hz, 4H), 1.02 (s, 9H), 0.47 (s, 6H), 0.27 (s, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>):  $\delta$  185.75, 158.75, 152.64, 141.20, 140.78, 143.80, 132.24, 131.93,

123.82, 121.85, 117.19, 115.91, 66.78, 47.74, 25.82, 18.46, -1.18, -4.13. HRMS (ESI<sup>+</sup>): *m/z* calcd 454.2228, found 454.2219 for C<sub>25</sub>H<sub>36</sub>NO<sub>3</sub>Si<sub>2</sub><sup>+</sup> (M+H<sup>+</sup>).

### **3-((*tert*-Butyldimethylsilyloxy)-7-morpholino-5-phenyl-10H-acridophosphin-10-one**

**5-oxide (4c).** Following general procedure C, **3c** (110 mg, 0.21 mmol) was reacted with Et<sub>4</sub>NOH (0.28 mL, 1.5 M) in dioxane (6 mL) and imidazole (142 mg, 2.1 mmol), *tert*-butyldimethylsilyl chloride (142 mg, 2.1 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (15 mL) to give **4c** (85 mg, 78%) as a light yellow solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.48 – 8.11 (m, 2H), 7.67 – 7.48 (m, 2H), 7.44 – 7.29 (m, 5H), 7.18 – 6.86 (m, 2H), 3.80 (t, *J* = 4.9 Hz, 4H), 3.41 – 3.29 (m, 4H), 0.93 (s, 9H), 0.18 (d, *J* = 5.0 Hz, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>-*d*): δ 180.4, 169.0, 167.3, 164.8, 160.1, 153.5, 139.2, 135.3, 131.7, 131.5, 131.3, 130.5, 130.4, 130.0, 129.6, 128.7, 128.6, 126.2, 124.0, 121.6, 121.5, 116.8, 114.2, 114.1, 66.3, 46.9, 25.5, 18.1, -4.39. <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ 4.8. LRMS (ESI<sup>+</sup>): *m/z* calcd 520.2, found 520.5 for C<sub>29</sub>H<sub>35</sub>NO<sub>4</sub>PSi (M+H<sup>+</sup>).

**General procedure D: synthesis of xanthene dye derivatives.** A flame-dried flask charged with **5** or **6** in dry THF (1 mL) was cooled to -78 °C. A solution of *tert*-butyllithium in pentane was added drop-wise under nitrogen. After stirring at the same temperature for 10 to 20 min, a solution of **4** in dry THF (2 to 4 mL) was added. The resultant solution was warmed to room temperature and stirred for 60 min. Aqueous HCl (20 mL, 1 M) was added to the reaction and stirred for an additional 60 min. The reaction was neutralized with NaHCO<sub>3</sub> and extracted with EtOAc (×3). The combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated under reduced pressure. The crude product was purified via flash chromatography to give the product.

**CCF1.** Following general procedure D, **5** (215 mg, 0.39 mmol) was reacted with *tert*-butyllithium (0.46 mL, 1.7 M in pentane, 0.78 mmol) and **4a** (85 mg, 0.20 mmol) to give **CCF1** (81 mg, 53%) as a dark magenta solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.87 (d, *J* = 1.5 Hz, 1H), 7.73 (d, *J* = 7.8 Hz, 1H), 7.25 (d, *J* = 8.3 Hz, 1H), 7.09 (d, *J* = 1.7 Hz, 1H), 6.79 (d, *J* = 9.8 Hz, 1H), 6.77 (d, *J* = 1.8 Hz, 1H), 6.65 (s, 2H), 6.30 (dd, *J* = 9.7, 1.9 Hz, 1H), 3.92 – 3.78 (m, 6H), 3.40 – 3.29 (m, 4H), 2.86 – 2.83 (m, 4H), 2.76 – 2.72 (m, 12H), 2.56 (q, *J* = 7.4 Hz, 4H), 1.73 (s, 3H), 1.57 (s, 3H), 1.25 (t, *J* = 7.4 Hz, 6H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 186.25, 156.13, 152.97, 150.01, 149.37, 138.08, 132.55, 131.80, 131.66, 129.72, 126.64, 124.61, 123.28, 122.86, 112.41, 111.32, 77.52, 77.20, 76.88, 66.68, 58.03, 54.06, 47.57, 40.29, 35.99, 32.72, 31.96, 30.61, 30.34, 26.28, 14.98. HRMS (ESI<sup>+</sup>) *m/z* calcd 777.2858, found 777.2869 for C<sub>40</sub>H<sub>52</sub>F<sub>3</sub>N<sub>2</sub>O<sub>2</sub>S<sub>4</sub><sup>+</sup> (M+H<sup>+</sup>).

**Ctrl-CCF1.** Following general procedure D, **6** (100 mg, 0.21 mmol) was reacted with *tert*-butyllithium (0.25 mL, 1.7 M in pentane, 0.42 mmol) and **4a** (46 mg, 0.11 mmol) to give **Ctrl-CCF1** (44 mg, 60%) as a dark magenta solid. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.83 (s, 1H), 7.64 (s, 1H), 7.21 (d, *J* = 7.7 Hz, 1H), 7.09 (d, *J* = 2.2 Hz, 1H), 6.79 (d, *J* = 9.7 Hz, 1H), 6.74 (d, *J* = 1.8 Hz, 1H), 6.66 (d, *J* = 8.9 Hz, 1H), 6.60 (dd, *J* = 8.9, 2.3 Hz, 1H), 6.27 (dd, *J* = 9.8, 1.8 Hz, 1H), 3.92 – 3.77 (m, 4H), 3.69 (s, 2H), 3.38 – 3.25 (m, 4H), 1.73 (s, 3H), 1.57 (s, 3H), 1.37 –

1.15 (m, 33H), 0.86 (dd,  $J = 8.8, 4.8$  Hz, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  186.50, 156.13, 152.93, 149.90, 138.13, 132.47, 131.40, 126.74, 124.73, 123.38, 123.06, 112.34, 111.40, 77.55, 77.23, 76.92, 66.73, 60.59, 58.35, 54.33, 47.63, 40.28, 36.01, 32.05, 30.61, 29.90, 29.71, 29.51, 27.62, 27.30, 22.86, 14.39, 14.30. HRMS (ESI<sup>+</sup>)  $m/z$  calcd 705.4601, found 705.4604 for  $\text{C}_{44}\text{H}_{60}\text{F}_3\text{N}_2\text{O}_2^+$  (M+H<sup>+</sup>).

**CSF1.** Following general procedure D, **5** (55 mg, 0.099 mmol) was reacted with *tert*-butyllithium (0.12 mL, 1.7 M in pentane, 0.20 mmol) and **4b** (30 mg, 0.066 mmol) to give **CSF1** (23 mg, 29%) as a dark purple solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.84 (s, 1H), 7.70 (s, 1H), 7.23 (d,  $J = 7.7$  Hz, 1H), 7.09 (d,  $J = 2.6$  Hz, 1H), 6.83 (d,  $J = 2.1$  Hz, 1H), 6.75 (d,  $J = 10.1$  Hz, 1H), 6.67 (dd,  $J = 9.1, 2.6$  Hz, 1H), 6.62 (d,  $J = 9.1$  Hz, 1H), 6.21 (dd,  $J = 10.1, 2.1$  Hz, 1H), 3.91 – 3.75 (m, 6H), 3.39 – 3.24 (m, 4H), 2.84 (s, 4H), 2.73 (s, 12H), 2.56 (q,  $J = 7.4$  Hz, 4H), 1.25 (t,  $J = 7.4$  Hz, 1H), 0.52 (s, 3H), 0.42 (s, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  184.38, 150.78, 146.95, 141.61, 140.15, 136.70, 135.48, 131.81, 131.67, 128.51, 127.20, 119.44, 114.47, 77.40, 66.67, 58.02, 54.07, 47.18, 32.74, 31.98, 30.35, 26.29, 14.98, -0.26, -2.04. HRMS (ESI<sup>+</sup>)  $m/z$  calcd 793.2627, found 793.2624 for  $\text{C}_{39}\text{H}_{52}\text{F}_3\text{N}_2\text{O}_2\text{S}_4\text{Si}^+$  (M+H<sup>+</sup>).

**Ctrl-CSF1.** Following general procedure D, **6** (47 mg, 0.099 mmol) was reacted with *tert*-butyllithium (0.12 mL, 1.7 M in pentane, 0.20 mmol) and **4b** (30 mg, 0.066 mmol) to give **Ctrl-CSF1** (18 mg, 25%) as a dark purple solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.81 (s, 1H), 7.62 (s, 1H), 7.20 (s, 1H), 7.09 (s, 1H), 6.83 (d,  $J = 2.1$  Hz, 1H), 6.76 (d,  $J = 8.3$  Hz, 1H), 6.62 (s, 2H), 6.20 (dd,  $J = 10.1, 2.1$  Hz, 1H), 3.88 – 3.81 (m, 4H), 3.68 (s, 2H), 3.36 – 3.28 (m, 4H), 2.48 (s, 3H), 1.37 – 1.12 (m, 24H), 0.86 (t,  $J = 6.2$  Hz, 6H), 0.49 (d,  $J = 16.9$  Hz, 3H), 0.42 (s, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  184.46, 150.78, 147.04, 141.74, 140.19, 136.66, 135.48, 131.75, 128.52, 127.15, 119.47, 114.35, 77.40, 66.67, 51.01, 47.18, 32.00, 29.64, 29.46, 27.53, 22.83, 14.36, -0.27, -2.05. HRMS (ESI<sup>+</sup>)  $m/z$  calcd 721.4371, found 721.4358 for  $\text{C}_{43}\text{H}_{60}\text{F}_3\text{N}_2\text{O}_2\text{Si}^+$  (M+H<sup>+</sup>).

**CPF1.** Following general procedure D, **5** (127 mg, 0.23 mmol) was reacted with *tert*-butyllithium (0.30 mL 1.7 M in pentane, 0.46 mmol) and **4c** (40 mg, 0.077 mmol) to give **CPF1** (57 mg, 38 %) as a dark blue solid.  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}-d_4$ )  $\delta$  8.06 (s, 1H), 7.86 (d,  $J = 7.7$  Hz, 1H), 7.73 – 7.40 (m, 6H), 7.37 (d,  $J = 7.8$  Hz, 1H), 7.11 (dd,  $J = 16.9, 2.0$  Hz, 1H), 7.00 (dd,  $J = 9.1, 2.8$  Hz, 1H), 6.89 (dd,  $J = 10.0, 6.8$  Hz, 1H), 6.82 (dd,  $J = 9.2, 6.3$  Hz, 1H), 6.29 (dd,  $J = 10.1, 2.1$  Hz, 1H), 3.86 (s, 2H), 3.76 (t,  $J = 4.8$  Hz, 4H), 3.45 (t,  $J = 4.9$  Hz, 4H), 2.84 – 2.73 (m, 8H), 2.71 (s, 8H), 2.54 (q,  $J = 7.4$  Hz, 4H), 1.20 (t,  $J = 7.4$  Hz, 6H).  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  183.9, 183.8, 152.4, 152.3, 151.8, 142.5, 140.6, 140.2, 139.5, 136.0, 133.9, 133.4, 132.7, 132.4, 131.7, 129.8, 129.7, 129.0, 128.9, 128.4, 128.2, 126.7, 125.6, 125.1, 123.8, 122.9, 122.6, 116.6, 115.9, 66.0, 57.3, 53.8, 46.3, 31.9, 31.4, 29.6, 25.3, 13.9.  $^{31}\text{P}$  NMR (162 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  10.4. HRMS (ESI<sup>+</sup>)  $m/z$  calcd 859.2467, found 859.2454 for  $\text{C}_{43}\text{H}_{51}\text{F}_3\text{N}_2\text{O}_3\text{PS}_4^+$  (M+H<sup>+</sup>).

**Ctrl-CPF1.** Following general procedure D, **6** (70 mg, 0.14 mmol) was reacted with *tert*-butyllithium (0.18 mL 1.7 M in pentane, 0.29 mmol) and **4c** (25 mg, 0.048 mmol) to give **CPF1** (14 mg, 37 %) as a dark blue solid. <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.98 (s, 1H), 7.78 (d, *J* = 7.1 Hz, 1H), 7.69 – 7.42 (m, 6H), 7.38 (d, *J* = 7.7 Hz, 1H), 7.11 (dd, *J* = 16.9, 2.1 Hz, 1H), 6.96 (dd, *J* = 9.3, 2.8 Hz, 1H), 6.83 (ddd, *J* = 18.9, 9.7, 6.5 Hz, 2H), 6.26 (dd, *J* = 10.0, 2.2 Hz, 1H), 3.80 (s, 2H), 3.75 (t, *J* = 4.9 Hz, 4H), 3.44 (t, *J* = 4.9 Hz, 4H), 2.55 (t, *J* = 7.2 Hz, 4H), 1.54 (dd, *J* = 10.4, 3.7 Hz, 4H), 1.33 – 1.16 (m, 20H), 0.94 – 0.75 (m, 6H). <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ 183.9, 183.7, 152.4, 152.3, 151.6, 151.6, 142.3, 140.3, 140.3, 139.4, 136.0, 135.8, 134.0, 133.3, 132.7, 132.5, 132.2, 131.7, 131.6, 129.8, 129.7, 129.0, 128.8, 128.5, 128.2, 126.7, 125.6, 123.8, 122.6, 116.7, 115.8, 65.9, 57.6, 53.7, 46.3, 31.6, 29.2, 29.1, 27.1, 26.5, 22.3, 13.0. <sup>31</sup>P NMR (162 MHz, CD<sub>3</sub>OD) δ 10.4. HRMS (ESI<sup>+</sup>) *m/z* calcd 787.4210, found 787.4202 for C<sub>47</sub>H<sub>59</sub>F<sub>3</sub>N<sub>2</sub>O<sub>3</sub>P<sup>+</sup> (M+H<sup>+</sup>).

### 2.2.3 Spectroscopic methods and materials

All spectroscopic measurements were performed at 25 °C in 25 mM HEPES buffer (pH 7.4) prepared with Millipore water unless noted. Absorption spectra were recorded using a Varian Cary 60 spectrophotometer and fluorescence spectra were recorded using a Photon Technology International Quanta Master 4 L-format scan spectrofluorometer equipped with an LPS-220B 75-W xenon lamp, A-1010B lamp housing with integrated igniter, switchable 814 photocounting/analog photomultiplier detection unit and MD5020 motor driver. Samples for absorption and emission measurements were contained in 0.35 × 1 cm quartz cuvettes (1.4-mL volume; Starna). Excitation was provided at 580 nm, 616 nm and 650 nm for CCF1/Ctrl-CCF1, CSF1/Ctrl-CSF1 and CPF1/Ctrl-CPF1, respectively. Stock solutions of [Cu(CH<sub>3</sub>CN)<sub>4</sub>]PF<sub>6</sub> in acetonitrile were used to provide Cu<sup>+</sup>. Metals used in the selectivity assays were derived from their chloride salts. The binding affinities of CCF1, CSF1 and CPF1 to Cu<sup>+</sup> were measured using thiourea as a competitive ligand to provide a buffered Cu<sup>+</sup> solution ( $\beta_{12} = 2.0 \times 10^{12}$ ,  $\beta_{13} = 2.0 \times 10^{14}$ ,  $\beta_{14} = 3.4 \times 10^{15}$ ).<sup>92</sup> For characterization of probe responses at different pH values, buffers were prepared by neutralizing the free acid solutions of 25 mM HEPES, 25 mM MES and 25 mM acetic acid with 5 M NaOH. HEK 293T cell lysates were prepared in HEPES buffer with pump-freeze-thaw cycles and handled under a nitrogen atmosphere, with their concentrations adjusted to 1 mg/mL protein as analyzed by the Bradford assay.

### 2.2.4 Preparation of cell cultures

Cells were grown in the Cell Culture Facility at the University of California, Berkeley. HEK 293T cells and MEF cells were cultured in DMEM supplemented with 10% FBS and glutamine (2 mM). Two days before imaging, cells were passed and plated on 4-well chamber slides (Lab-Tek, Thermo Fisher Scientific) coated with poly-L-lysine (50 mg/mL, Sigma-Aldrich).

### 2.2.5 Cell staining and imaging

Confocal fluorescence images were acquired with a Zeiss LSM710 laser-scanning microscope with 20x objective lens. Excitation at 543 nm for CCF1/Ctrl-CCF1, 594 nm for CSF1/Ctrl-CSF1, 633 nm for CPF1/Ctrl-CPF1 and 488 nm for Calcium Green-1 were carried out with appropriate lasers. Cells were incubated with 2  $\mu$ M CCF1/Ctrl-CCF1, CSF1/Ctrl-CSF1 and CPF1/Ctrl-CPF1 in DMEM without phenol red (Invitrogen) for 15 min at 37 °C under 5% CO<sub>2</sub>, washed and imaged in fresh DMEM without phenol red. For imaging of HEK 293T cells with copper addition or depletion, cells were treated with 50  $\mu$ M CuCl<sub>2</sub>, 500  $\mu$ M BCS, or water in DMEM without phenol red for 12 h prior to incubation with probe-containing media. For imaging of MEFs, cells were grown to the desired confluency in the growth medium before direct incubation with probe-containing media.

For dual color imaging of Cu and Ca, HEK 293T cells were treated with 100  $\mu$ M CuCl<sub>2</sub> or water in DMEM without phenol red for 12 h, incubated with 5  $\mu$ M Calcium Green 1-AM (AAT Bioquest) in HBSS without CaCl<sub>2</sub> or MgCl<sub>2</sub> (Gibco) for 30 min at 37 °C, followed by incubation in HBSS without CaCl<sub>2</sub> or MgCl<sub>2</sub> for de-esterification for 30 min at 37 °C, and incubation with 5  $\mu$ M CPF1 or Ctrl-CPF1 in HBSS without CaCl<sub>2</sub> or MgCl<sub>2</sub> for 15 min at 37 °C prior to imaging in HBSS with 1 mM CaCl<sub>2</sub> without phenol red.

For colocalization experiments, cells were stained with 5  $\mu$ M ER-Tracker Green (Thermo Fisher) in DMEM without phenol red for 15 min prior to incubation with 2  $\mu$ M CCF1, CSF1 or CPF1 in DMEM without phenol red for 15 min at 37 °C. The medium was then replaced with fresh DMEM and the cells were imaged with a 63x oil-immersion objective lens. ER-Tracker Green was excited at 488 nm and emission was collected between 493–550 nm.

### 2.2.6 Image analysis and quantification

ImageJ (National Institutes of Health) was used for image analysis. For quantification of fluorescence intensity, each image was set to 8-bit greyscale and inverted. The fluorescence intensity was estimated using non-calibrated OD function. The area of stained cells was selected by setting appropriate threshold ( $\geq 0.051$  for CCF1/Ctrl-CCF1, CSF1/Ctrl-CSF1, CPF1/Ctrl-CPF1 images and  $\geq 0.032$  for Calcium Green-1 images). The statistics of the image was then measured by the “Measure” function and the average fluorescence intensity was obtained by dividing the integrated density (IntDen) over area. For each condition, four images of different fields of cells from each biological replicate were analyzed using this process and the values were combined for statistical analysis.

### 2.2.7 Cell fractionation and Inductively Coupled Plasma (ICP)-MS analysis

Atp7a<sup>-/-</sup> and matched control MEFs were fractionated using the NE-PER kit (Thermo Fisher). Extracts were digested by adding equal volumes of concentrated nitric acid, incubated overnight at room temperature on a rotator, and boiled for 2 hours at 95 °C. Digested extracts were diluted into 2% nitric acid with an internal standard and run on an iCAP-Q ICP-MS in KED

mode. The level of each element found in the extraction reagent was subtracted from each measurement. The resulting values were normalized to protein concentration.

## 2.3 Results and discussion

### 2.3.1 Design and synthesis of a color palette of fluorescent copper probes

Inspired by emerging progress on the development of new fluorophores derived from center-atom substitution of the xanthene oxygen of classic fluorescein and rhodamine dyes, we envisioned generating a series of red-shifted fluorescent copper sensors by transforming rhodol scaffolds employed in our recently reported Copper Rhodol (CR) and Copper Fluor (CF) family (Scheme 2.1) into carbon, silicon and phosphorus-substituted analogs.<sup>5,34</sup> The synthesis and structures of the fluorescent copper indicators, CCF1, CSF1 and CPF1, along with control dyes that are not responsive to copper, are depicted in Scheme 2.2. In particular, our design makes use of a trifluoromethyl (CF<sub>3</sub>) group on the pendant aryl ring attached to the xanthene backbone, as previous work from our laboratory establishes that substitution of a methyl for trifluoromethyl at this position leads to enhancements in both the dynamic range and optical brightness of rhodol-based copper probes by decreasing the available non-radiative decay pathways from rotational motions about the aryl-aryl bond.<sup>34</sup> We also synthesized matched control dyes, Ctrl-CCF1, Ctrl-CSF1 and Ctrl-CPF1, which are non-responsive to Cu<sup>+</sup> due to the replacement of the thioether motifs by methylene units in the metal-binding domain (Scheme 2.2). These control dyes can be used in parallel imaging experiments to distinguish copper-dependent responses from potential dye-dependent variations, which could include cellular uptake, trappability and subcellular accumulation as well as changes in pH, redox, and hydrophobic/hydrophilic environments.

### 2.3.2 Spectroscopic properties of CXF copper sensors

We first sought to evaluate the optical responses of the CXF series and their control analogs to copper in spectrophotometric assays. As anticipated, all three copper probes showed turn-on responses upon treatment with Cu<sup>+</sup> in aqueous solution buffered to physiological pH, with a ca. 17-fold increase for CSF1, 5-fold increase for CCF1 and 7-fold increase for CPF1 (Figure 2.1a, c, e). We speculate that the observed differences in turn-on ratios between the three probes may result from various factors, including photoinduced electron transfer (PeT) quenching efficiency, variations in rotational freedom on the receptor upon Cu<sup>+</sup> binding, as well as changes in other non-radiative relaxation pathways that become more prominent with a smaller HOMO-LUMO gap as supported by the lower quantum yields of CSF1 and CPF1 (Table 2.1).

The copper-binding affinities of these probes are within the expected range for small-molecule fluorescent copper indicators with thioether-rich ligand sets,<sup>50,51</sup> with CPF1 having the strongest affinity and CSF1 being the weakest Cu<sup>+</sup> binder (Figure 2.2b-c, 2.3b-c and 2.4b-c). The apparent  $K_d$  values for CCF1, CSF1 and CPF1 are 0.20, 0.49 and 0.02 pM,

respectively, and all show a 1:1 Cu:dye binding ratio. Additionally, all three probes show excellent selectivity for Cu<sup>+</sup> over various biologically-relevant metal ions (Figure 2.1b, d, f) and maintain their effective turn-on responses to Cu<sup>+</sup> over a physiological pH range (Figure 2.2d, 2.3d and 2.4d). We observe a slight turn-on effect from Cu<sup>2+</sup> which most likely results from the slow reduction of Cu<sup>2+</sup> to Cu<sup>+</sup> upon interaction with the sulfur-rich receptor (Figure 2.4i). In contrast, the control dyes, Ctrl-CCF1, Ctrl-CSF1 and Ctrl-CPF1, do not exhibit any metal-dependent responses.

The spectral properties of the three probe-control pairs are listed in Table 1. As expected, all three sets show lower energy excitation and emission wavelengths compared to the rhodol-based congeners Copper Fluor 3 (CF3) and Control Copper Fluor 3 (Ctrl-CF3),<sup>34</sup> with the CCF1/Ctrl-CCF1 pair showing the smallest bathochromic shift (20-40 nm) and the CPF1/Ctrl-CPF1 pair exhibiting the most red-shifted spectra (80-100 nm) relative to the oxygen-based rhodol, which is in line with what is observed for carbon-, silicon- and phosphorus-based derivatives of fluorescein.<sup>90</sup> Because of the nature of the rhodol scaffold as a fluorescein-rhodamine hybrid, these probes exhibit a relatively broad absorption profile (full width at half maximum [fwhm] greater than 145 nm for apo-CXF and above 90 nm for copper-bound CXF indicators, Figure 2.2a, 2.3a and 2.4a), which can pose potential limitations for some multi-color, multi-analyte imaging experiments. However, the relatively narrow profiles of the emission peaks for these dyes (fwhm = 44 nm for CCF1 and CPF1, 34 nm for CSF1) balance their absorption behavior and provide an opportunity to use probes with longer emission wavelength, such as the phosphorus rhodol-based sensor CPF1, in dual-channel imaging experiments through disentangling distinct emission signals.

Finally, to further characterize the spectroscopic properties of this new CXF series of dyes, we assessed their ability to respond to copper in aqueous buffer with additives to mimic aspects of more complex biological environments. Specifically, we evaluated the response of the probes in buffer with BSA as a model protein or 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) as a model lipid, as well as whole cell lysates. As expected, both the CXF probes and their Ctrl-CXF analogs do show some background fluorescence with these *in vitro* additives. However, only the CXF indicators, but not the Ctrl-CXF dyes, respond to Cu<sup>+</sup> addition through a turn-on response over background, and more importantly, the observed copper-dependent fluorescence increases can be reversibly turned off by addition of competing copper chelators. Indeed, Ctrl-CXF control dyes display minimal response to such copper supplementation and/or depletion treatments (Figure 2.2e-h, 2.3e-h and 2.4e-h). Taken together, the data establish that the CCF1, CSF1 and CPF1 indicators, when used in conjunction with their Ctrl-CXF control analogs, are effective chemical tools for turn-on fluorescence detection of copper in aqueous buffer as well as in whole cell lysates.

### 2.3.3 Live-cell imaging of changes in labile copper pools in HEK 293T models

With these spectroscopic data in hand, we next evaluated the ability of the CXF probes to visualize changes in labile copper pools in cell culture models. To this end, we treated human embryonic kidney (HEK 293T) cells for 12 h with either 100 μM CuCl<sub>2</sub> to increase intracellular

copper concentrations or 500  $\mu$ M bathocuproine sulfonate (BCS), a membrane-impermeable copper chelator, to deplete endogenous labile copper pools, and then labeled with CXF or Ctrl-CXF probes for live-cell imaging experiments (Figure 2.5). As anticipated, HEK 293T cells stained with CCF1, CSF1 and CPF1 all showed increased intracellular fluorescence under copper supplementation conditions relative to vehicle control cells, with CPF1 having the highest turn-on signal-to-noise ratio of the three copper indicators, in line with its highest  $\text{Cu}^+$  binding affinity observed in spectroscopic assays. We observed that only CCF1 and CPF1, but not CSF1, showed a slight but statistically significant fluorescence decrease in HEK 293T cells treated with the BCS chelator that induce labile copper deficiency relative to vehicle control. The cellular distribution patterns of CXF reagents were similar for all conditions tested, and further co-staining experiments with fluorescent markers for organelles suggest overlap with the endoplasmic reticulum (ER) (Figure 2.6). In further agreement with *in vitro* spectroscopic studies, all of the Ctrl-CXF control dyes did not show significant differences in fluorescence intensity with either copper supplementation or depletion, further validating that the CXF indicators can be used to report changes in labile copper status in living cells. The results from these initial cell studies suggest that the CSF1/Ctrl-CSF1 pair in particular might be more effective in applications for detection of increases in labile cellular copper levels, whereas the CCF1/Ctrl-CCF1 and CPF1/Ctrl-CPF1 pairs may be used to visualize either depletions or accumulations of labile cellular copper pools relative to basal levels.

#### 2.3.4 Application of CCF1/Ctrl-CCF1 to detect copper increase in *Atp7a* knockout cell

We next sought to apply the CXF reagents to identify and assess aberrant changes in labile copper pools in cell-based models of disease. To this end, we utilized *Atp7a*<sup>-/-</sup> mouse embryonic fibroblasts (MEFs), a model that mimics genetic copper misregulation in Menkes disease.<sup>28,62,93</sup> *Atp7a* is a major copper transporter protein that regulates secretion and export of excess copper to maintain copper homeostasis,<sup>12</sup> and deletion and/or loss-of-function mutations result in hyperaccumulation of copper compared to wild type.<sup>94,95</sup> Indeed, we confirmed by ICP-MS measurements that MEF *Atp7a*<sup>-/-</sup> cells possess over 6 times as much total copper as its genetically matched MEF wild-type cells in the cytoplasmic extract, and 2.6 times as much in the nucleus (Figure 2.7).

Owing to its superior dynamic range in response to copper supplementation experiments within cells, we utilized CCF1 to visualize labile copper pools in *Atp7a*<sup>-/-</sup> MEFs versus wild type (WT) congeners. As shown in Figure 2.8, CCF1-stained cells displayed statistically significant elevations in intracellular fluorescence in MEF *Atp7a*<sup>-/-</sup> cells compared to WT controls. Besides the observed increases in CCF1 fluorescence in *Atp7a*<sup>-/-</sup> compared to WT cells, Ctrl-CCF1 fluorescence intensities in *Atp7a*<sup>-/-</sup> cells are slightly lower than those in WT fibroblasts, suggesting lower dye uptake in *Atp7a*<sup>-/-</sup> cells. The collective imaging data identify that in addition to total copper levels, labile copper levels in the *Atp7a* KO cells are also higher than their WT counterparts.

### 2.3.5 Dual-color imaging with CPF1 and Calcium Green-1

As an additional set of experiments to showcase the utility of an expanded color palette for the CXF family of copper sensors, we utilized the most red-shifted analog, CPF1, in dual-analyte, dual-color imaging experiments. CPF1 displays a deep red emission peak centered around 678 nm in the apo form and is essentially non-emissive below 600 nm, which makes it spectrally well-separated from common green-emitting fluorescent reporters like fluorescein and GFP. This feature of CPF1 enables simultaneous multi-color imaging of copper and another biologically relevant analytes and related targets.

To meet this goal, we performed dual-channel imaging of CPF1 alongside Calcium Green 1-AM, a fluorescein-based cell-trappable calcium indicator with green fluorescence. As expected, the green Calcium Green-1 and red CPF1 channels are spectrally well-separated (Figure 2.9). Live HEK 293T cells treated with 50  $\mu\text{M}$   $\text{CuCl}_2$  overnight exhibited a selective increase of intracellular fluorescence in the copper-sensitive CPF1 red channel but not in the Calcium Green-1 green channel (Figure 2.9a-e). In contrast, treatment of HEK 293T cells with 1  $\mu\text{M}$  of the calcium ionophore A23187, which can permeabilize the cell membrane and lead to calcium influx, resulted in enhanced fluorescence selectively in the Calcium Green-1 channel with no effect on fluorescence intensity in the CPF1 channel (Figure 2.9f-j). Similar results were obtained with dual-color imaging of Ctrl-CPF1 and Calcium Green-1, where the Ctrl-CPF1 red channel was non-responsive to either copper supplementation or triggered calcium uptake and Calcium Green-1 channel only responded to A23187-stimulated calcium influx (Figure 2.10). The results establish the red-shifted CPF1 and Ctrl-CPF1 as a valuable set of reagents to monitor changes in labile copper pools in combination with other fluorescent reporters for multi-channel, multi-analyte imaging experiments.

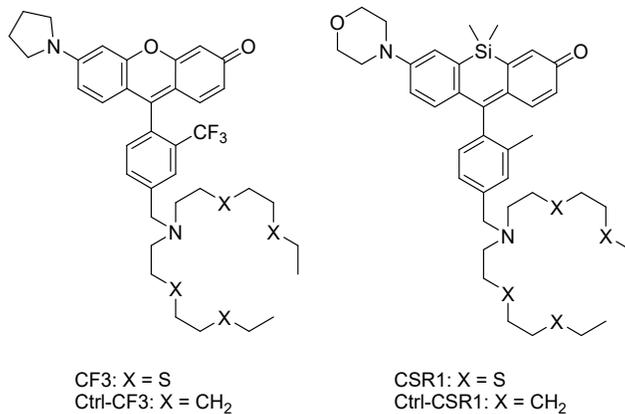
## 2.4 Conclusion

To close, we have presented the design, synthesis, *in vitro* and *in cellulo* characterization of a homologous family of fluorescent copper probes along with matched control dye compounds that feature center-atom substitution of the rhodol fluorophore. Systematic modifications to the xanthene oxygen to generate carbon, silicon, and phosphorus rhodol analogs furnished a palette of copper sensors with a range of emission profiles that span the orange to deep red, while maintaining a general copper-binding receptor to confer high metal and redox specificity. The resulting probes are capable of detecting changes in labile copper pools in living cells, as illustrated by pilot studies in HEK 293T cells with copper supplementation and/or depletion. Importantly, the imaging studies are supported by comparison with matched control Ctrl-CXF analogs that are non-responsive to copper fluctuations and can thus serve as reference compounds for potential dye-dependent responses. Moreover, imaging experiments with the carbon rhodol CCF1 and its Ctrl-CCF1 analog are consistent with elevations in labile copper pools in *Atp7a*<sup>-/-</sup> MEF cells lacking this central copper export protein compared to wild-type fibroblasts. Finally, the near-IR optical profile of the phosphorus-based CPF1 congener enables simultaneous, dual-color imaging of copper and calcium fluxes in living cells with high metal

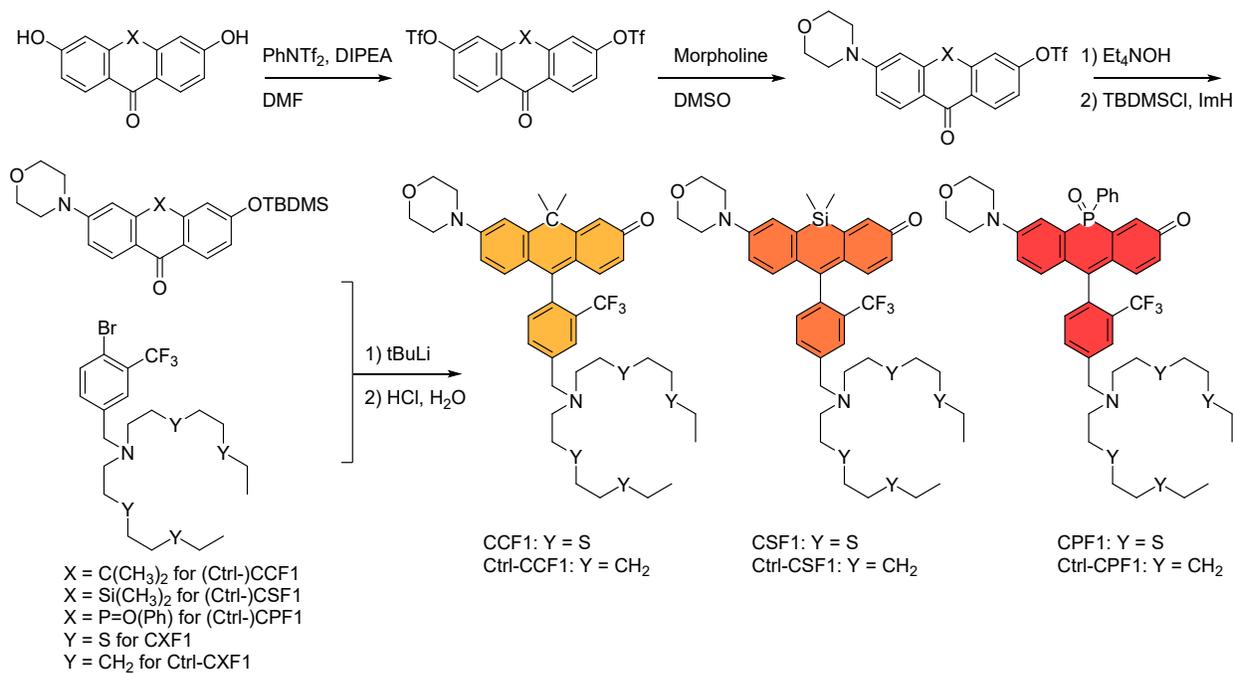
specificity. Collectively, this work shows that center-atom substitution of xanthene dyes can provide a general, rational strategy for expanding the color palette of fluorescent indicators while maintaining the same recognition/reactivity motif for analyte detection, along with providing reagents for advanced studies of copper biology.

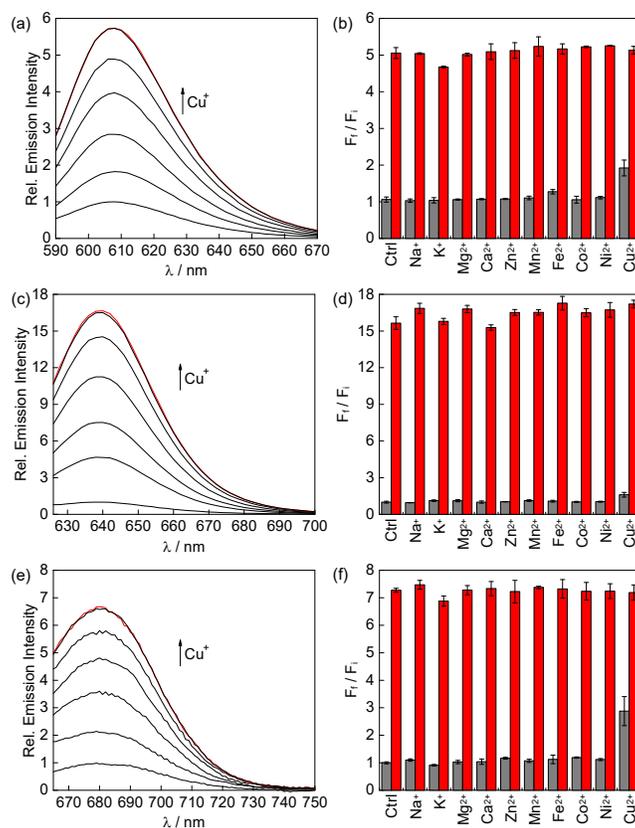
## 2.5 Figures and schemes

**Scheme 2.1** Chemical structures of previously reported copper sensors, CF3 and CSR1, along with their control analogs, Ctrl-CF3 and Ctrl-CSR1.



**Scheme 2.2** Synthesis and structures of CCF1, CSF1, CPF1 along with their control analogs Ctrl-CCF1, Ctrl-CSF1 and Ctrl-CPF1.



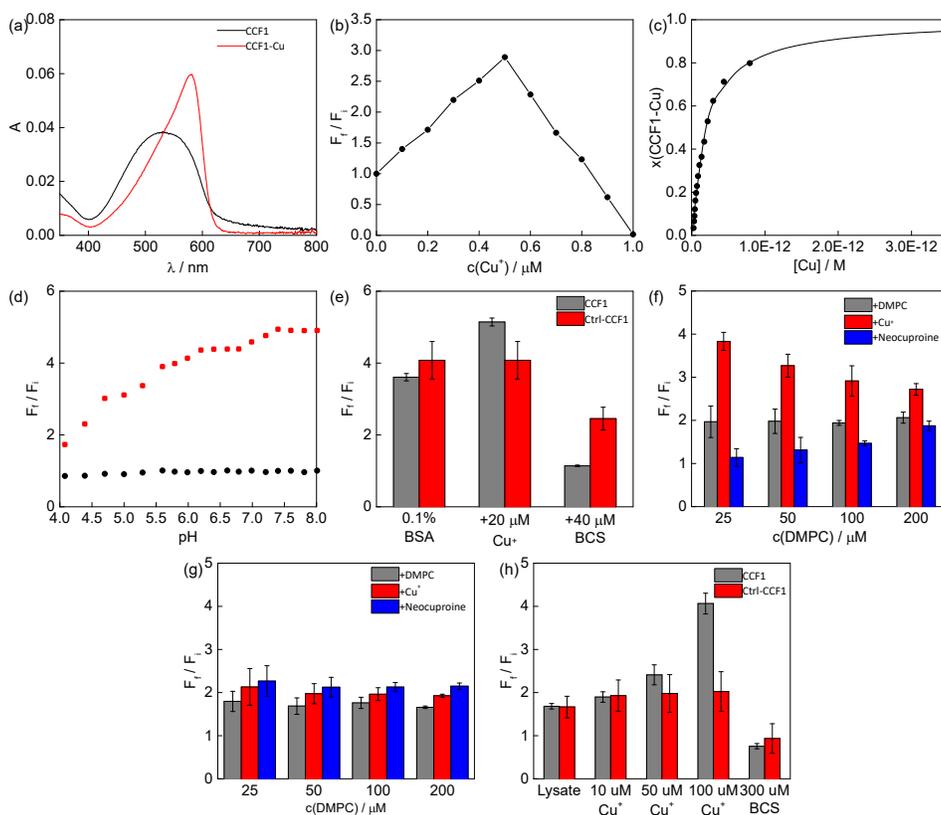


**Figure 2.1** Fluorescence turn-on responses of  $1 \mu\text{M}$  (a) CCF1, (c) CSF1, and (e) CPF1 to  $\text{Cu}^+$ . Lines represent addition of 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2  $\mu\text{M}$   $[\text{Cu}(\text{MeCN})_4]\text{PF}_6$ . Fluorescence responses of  $1 \mu\text{M}$  (b) CCF1, (d) CSF1, and (f) CPF1 to various metal ions. Black bars represent the addition of an excess of the appropriate metal ion (2 mM for  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Zn}^{2+}$ , and 50  $\mu\text{M}$  for other cations) to a  $1 \mu\text{M}$  solution of probe. Red bars represent subsequent addition of  $1 \mu\text{M}$   $[\text{Cu}(\text{MeCN})_4]\text{PF}_6$ . Bars represent the final integrated fluorescence response ( $F_f$ ) over the initial integrated emission ( $F_i$ ) shown as average  $\pm$  s.d. ( $n = 3$ ).

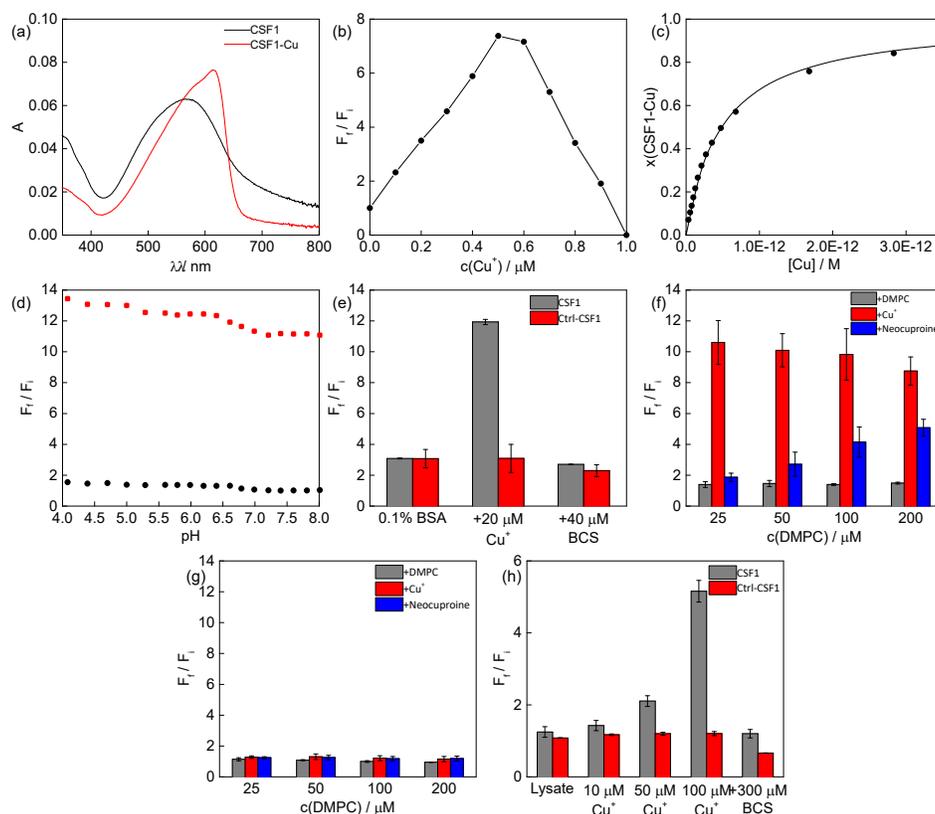
**Table 2.1 Spectral properties of CXF copper probes and their Ctrl-CXF control analogs**

	$\lambda_{\text{abs}}/\text{nm}$	$\lambda_{\text{em}}/\text{nm}$	$\epsilon/10^4 \text{ M}^{-1}\text{cm}^{-1}$	$\Phi^a$
CF3	510 and 550	557	3.3 and 3.7	0.026
CF3-Cu	534	557	6.9	0.219
Ctrl-CF3	510 and 545	557	4.7 and 4.7	0.007
CCF1	516	608	1.41	0.048
CCF1-Cu	580	608	2.05	0.30
Ctrl-CCF1	506	610	2.33	0.0083
CSF1	568	638	3.00	0.0041
CSF1-Cu	616	639	3.15	0.053
Ctrl-CSF1	531	638	1.99	0.0036
CPF1	569	679	1.76	0.00018
CPF1-Cu	654	680	1.93	0.0020
Ctrl-CPF1	568 <sup>b</sup>	674	1.72	0.00043

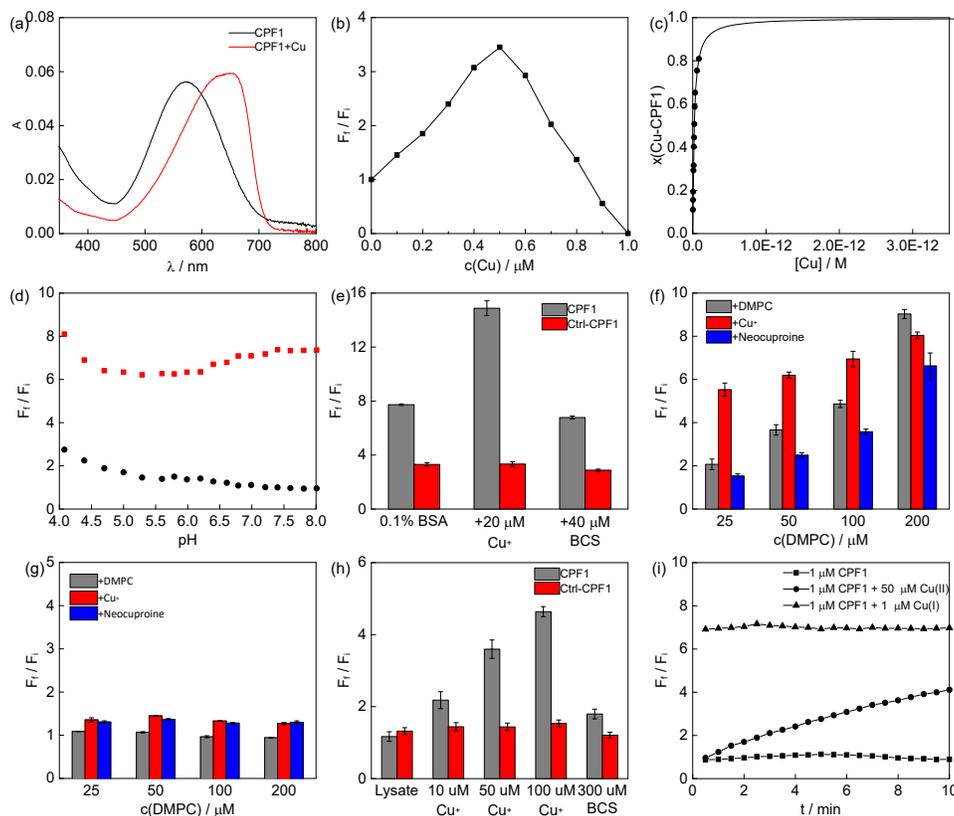
<sup>a</sup> Cresyl violet in methanol ( $\Phi = 0.54$ )<sup>92</sup> was used as a standard.



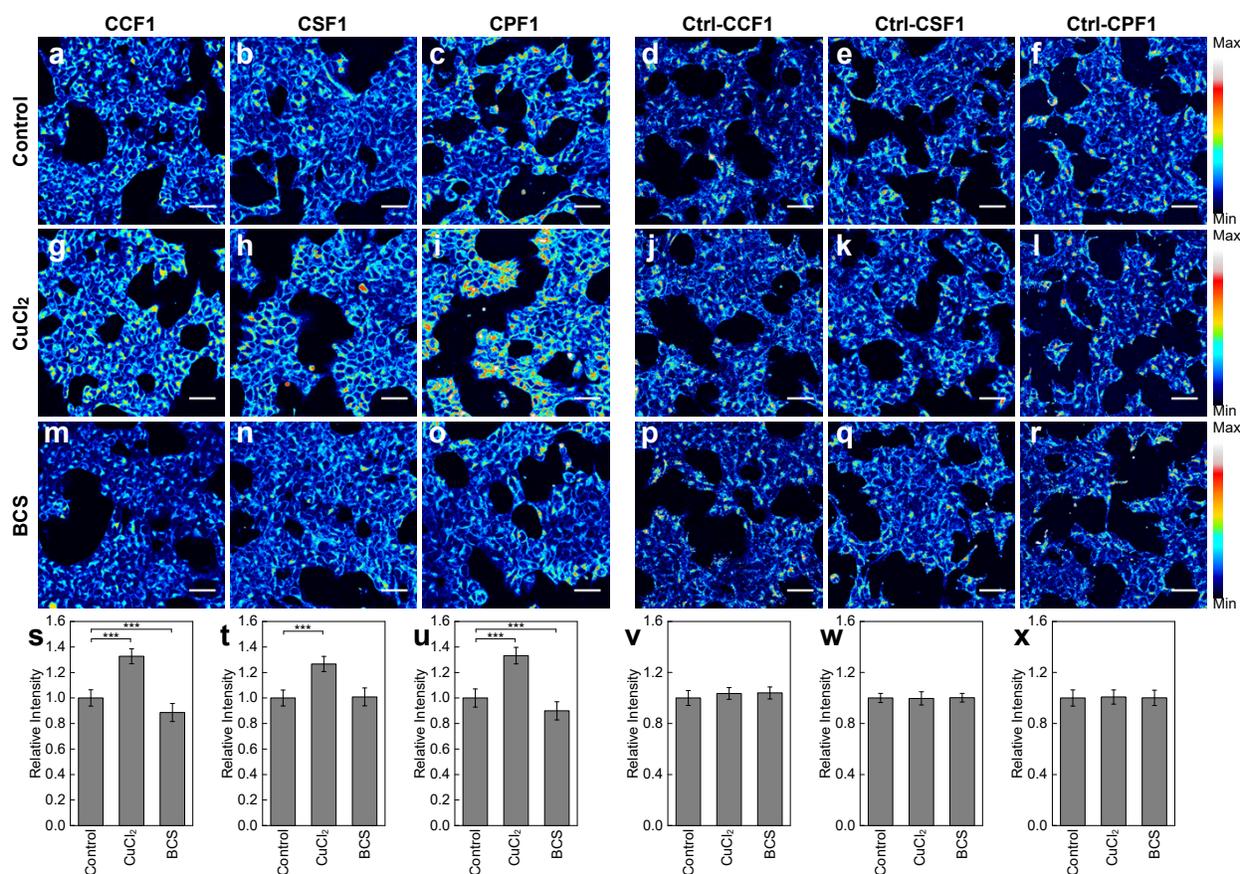
**Figure 2.2** *In vitro* characterization of CCF1 and Ctrl-CCF1. For bar graphs, bars represent the final integrated fluorescence response ( $F_f$ ) over the initial integrated emission ( $F_i$ ) of CCF1 or Ctrl-CCF1 over 590–670 nm; values are shown as average  $\pm$  s.d. ( $n = 3$ ). (a) UV-visible spectral change of 3  $\mu\text{M}$  CCF1 (black) upon addition of 3  $\mu\text{M}$   $\text{Cu}^+$  (red). (b) Job's plot of CCF1 and  $\text{Cu}^+$ . The total concentrations of CCF1 and  $\text{Cu}^+$  were kept at 1  $\mu\text{M}$ . (c) Fluorescence response of 1  $\mu\text{M}$  CCF1 to thiourea-buffered  $\text{Cu}^+$  solutions for the  $K_d$  measurement. The observed  $K_d$  value is  $2.0 \times 10^{-13}$  M. Solid line represents the calculated curve. (d) Fluorescence intensity of 1  $\mu\text{M}$  CCF1 (black) and 1  $\mu\text{M}$  CCF1-Cu (red) across a range of pH values. (e) Fluorescence response of 1  $\mu\text{M}$  CCF1 (grey) or Ctrl-CCF1 (red) to 0.1% BSA and subsequent addition of 20  $\mu\text{M}$   $\text{Cu}^+$  and 40  $\mu\text{M}$  BCS. (f) Fluorescence response of 1  $\mu\text{M}$  CCF1 in the presence of DMPC, a lipid-forming reagent (grey), with subsequent addition of 1  $\mu\text{M}$   $\text{Cu}^+$  to the solution (red), and a final addition of 10  $\mu\text{M}$  neocuproine (blue). (g) Fluorescence response of 1  $\mu\text{M}$  Ctrl-CCF1 in the presence of DMPC (grey), with subsequent addition of 1  $\mu\text{M}$   $\text{Cu}^+$  to the solution (red), and a final addition of 10  $\mu\text{M}$  neocuproine (blue). (h) Fluorescence response of 1  $\mu\text{M}$  CCF1 (grey) and Ctrl-CCF1 (red) to HEK 293T lysates (1 mg protein/mL) and subsequent addition of 10, 50 and 100  $\mu\text{M}$   $\text{Cu}^+$ , followed by addition of 300  $\mu\text{M}$  BCS.



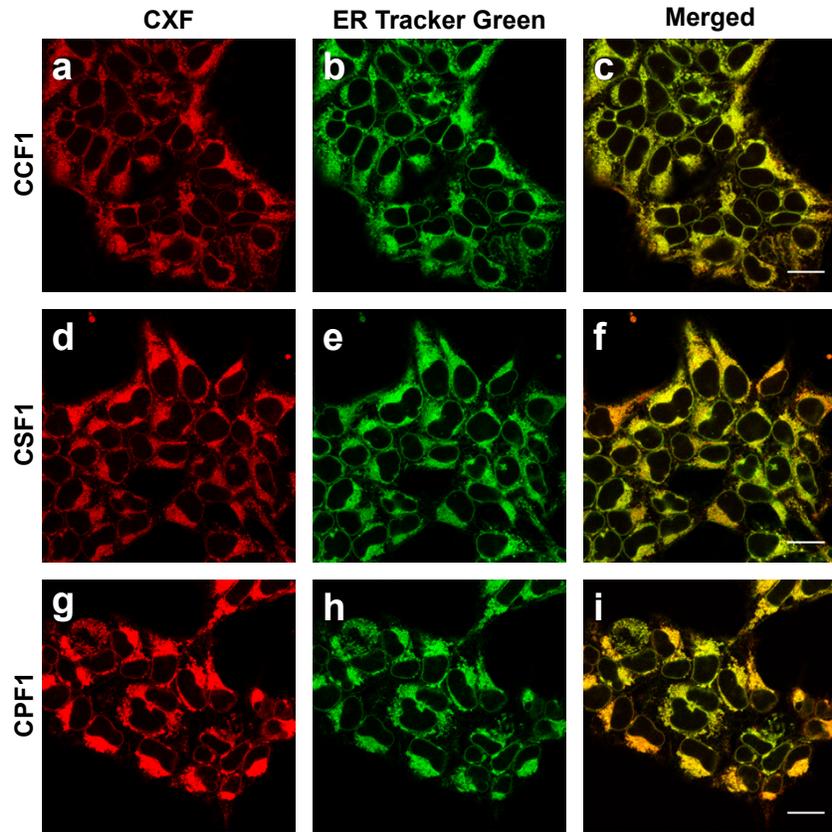
**Figure 2.3** *In vitro* characterization of CSF1 and Ctrl-CSF1. For bar graphs, bars represent the final integrated fluorescence response ( $F_f$ ) over the initial integrated emission ( $F_i$ ) of CSF1 or Ctrl-CSF1 over 626–700 nm; values are shown as average  $\pm$  s.d. ( $n = 3$ ). (a) UV-visible spectral change of 2  $\mu\text{M}$  CSF1 (black) upon addition of 2  $\mu\text{M}$   $\text{Cu}^+$  (red). (b) Job's plot of CSF1 and  $\text{Cu}^+$ . The total concentrations of CSF1 and  $\text{Cu}^+$  were kept at 1  $\mu\text{M}$ . (c) Fluorescence response of 1  $\mu\text{M}$  CSF1 to thiourea-buffered  $\text{Cu}^+$  solutions for  $K_d$  measurement. The observed  $K_d$  value is  $4.9 \times 10^{-13}$  M. Solid line represents the calculated curve. (d) Fluorescence intensity of 1  $\mu\text{M}$  CSF1 (black) and 1  $\mu\text{M}$  CSF1-Cu (red) across a range of pH values. (e) Fluorescence response of 1  $\mu\text{M}$  CSF1 (grey) or Ctrl-CSF1 (red) to 0.1% BSA and subsequent addition of 20  $\mu\text{M}$   $\text{Cu}^+$  and 40  $\mu\text{M}$  BCS. (f) Fluorescence response of 1  $\mu\text{M}$  CSF1 in the presence of DMPC (grey), with subsequent addition of 1  $\mu\text{M}$   $\text{Cu}^+$  to the solution (red), and a final addition of 10  $\mu\text{M}$  neocuproine (blue). (g) Fluorescence response of 1  $\mu\text{M}$  Ctrl-CSF1 in the presence of DMPC (grey), with subsequent addition of 1  $\mu\text{M}$   $\text{Cu}^+$  to the solution (red), and a final addition of 10  $\mu\text{M}$  neocuproine (blue). (h) Fluorescence response of 1  $\mu\text{M}$  CSF1 (grey) and Ctrl-CSF1 (red) to HEK293T lysates (1 mg protein/mL) and subsequent addition of 10, 50 and 100  $\mu\text{M}$   $\text{Cu}^+$ , followed by addition of 300  $\mu\text{M}$  BCS.



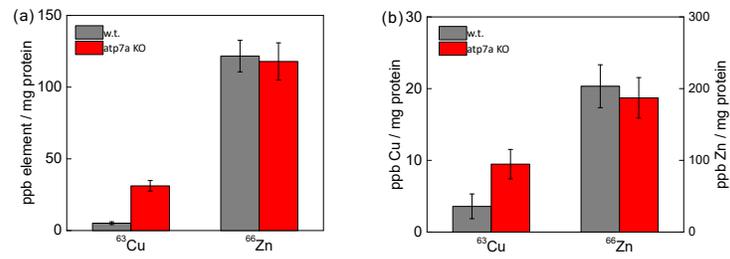
**Figure 2.4** *In vitro* characterization of CPF1 and Ctrl-CPF1. For bar graphs, bars represent the final integrated fluorescence response ( $F_f$ ) over the initial integrated emission ( $F_i$ ) of CPF1 or Ctrl-CPF1 over 660–700 nm; values are shown as average  $\pm$  s.d. ( $n = 3$ ). (a) UV-visible spectral change of 3  $\mu\text{M}$  CPF1 (black) upon addition of 3  $\mu\text{M}$   $\text{Cu}^+$  (red). (b) Job's plot of CPF1 and  $\text{Cu}^+$ . The total concentrations of CPF1 and  $\text{Cu}^+$  were kept at 1  $\mu\text{M}$ . (c) Fluorescence response of 1  $\mu\text{M}$  CPF1 to thiourea-buffered  $\text{Cu}^+$  solutions for  $K_d$  measurement. The observed  $K_d$  value is  $0.20 \times 10^{-13}$  M. Solid line represents the calculated curve. (d) Fluorescence intensity of 1  $\mu\text{M}$  CPF1 (black) and 1  $\mu\text{M}$  CPF1-Cu (red) across a range of pH values. (e) Fluorescence response of 1  $\mu\text{M}$  CPF1 (grey) or Ctrl-CPF1 (red) to 0.1% BSA and subsequent addition of 20  $\mu\text{M}$   $\text{Cu}^+$  and 40  $\mu\text{M}$  BCS. (f) Fluorescence response of 1  $\mu\text{M}$  CPF1 in the presence of DMPC (grey), with subsequent addition of 1  $\mu\text{M}$   $\text{Cu}^+$  to the solution (red), and a final addition of 10  $\mu\text{M}$  neocuproine (blue). (g) Fluorescence response of 1  $\mu\text{M}$  Ctrl-CPF1 in the presence of DMPC (grey), with subsequent addition of 1  $\mu\text{M}$   $\text{Cu}^+$  to the solution (red), and a final addition of 10  $\mu\text{M}$  neocuproine (blue). (h) Fluorescence response of 1  $\mu\text{M}$  CPF1 (grey) and Ctrl-CPF1 (red) to HEK 293T lysates (1 mg protein/mL) and subsequent addition of 10, 50 and 100  $\mu\text{M}$   $\text{Cu}^+$ , followed by addition of 300  $\mu\text{M}$  BCS. (i) Fluorescence emission of 1  $\mu\text{M}$  CPF1 (square) over time, and its response to 50  $\mu\text{M}$   $\text{CuSO}_4$  (circle) or 1  $\mu\text{M}$   $[\text{Cu}(\text{CH}_3\text{CN})_4]\text{PF}_6$  (triangle). The slow turn-on response to Cu(II) compared to the prompt and stable turn-on by Cu(I) suggests that Cu(II) may be slowly reduced to Cu(I) by interaction with the probe receptor, resulting in an attenuated turn-on response.



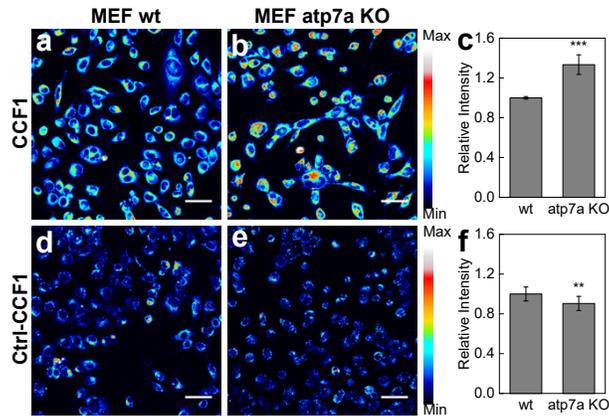
**Figure 2.5** Fluorescence imaging of labile copper pools in live HEK 293T cells with (a, g, m) CCF1, (b, h, n) CSF1, (c, i, o) CPF1, (d, j, p) Ctrl-CCF1, (e, k, q) Ctrl-CSF1 and (f, l, r) Ctrl-CPF1. Control cells (a-f) and cells incubated with (g-l) 100  $\mu$ M CuCl<sub>2</sub> or (m-r) 500  $\mu$ M BCS in the growth medium for 12 h at 37 °C were stained with 5  $\mu$ M dye for 30 min at 37 °C in DMEM. Scale-bars: 40  $\mu$ m. (s-x) Quantification of fluorescence intensity of cells stained with CCF1, CSF1, CPF1, Ctrl-CCF1, Ctrl-CSF1 and Ctrl-CPF1, respectively. Data were normalized to control cells and shown as average  $\pm$  s.d. (CCF: n = 4; Ctrl-CCF: n = 3). \*\*\* $P \leq 0.001$ ; two-tailed Student's t-test.



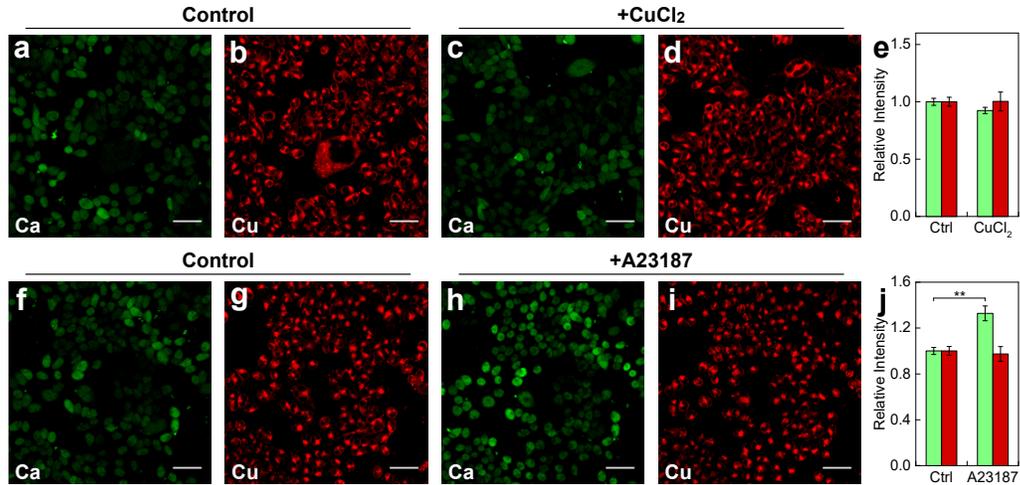
**Figure 2.6** Colocalization of CCF1, CSF1 and CPF1 with ER-Tracker Green: (a) CCF1, (b) ER-Tracker Green, (c) merge of (a) and (b). (d) CSF1, (e) ER-Tracker Green, (f) Merge of (d) and (e). (g) CPF1, (h) ER-Tracker Green, (i) merge of (g) and (h). Scale-bars: 20  $\mu\text{m}$ . Pearson's coefficients of pixel intensity spatial correlation between ER-Tracker Green and CCF1, CSF1 or CPF1 are  $0.87 \pm 0.03$ ,  $0.82 \pm 0.04$  and  $0.85 \pm 0.04$ , respectively, averaged across 3 separate fields of cells using Fiji's Coloc 2 plugin for ImageJ; error represents the standard deviation between different fields of cells.



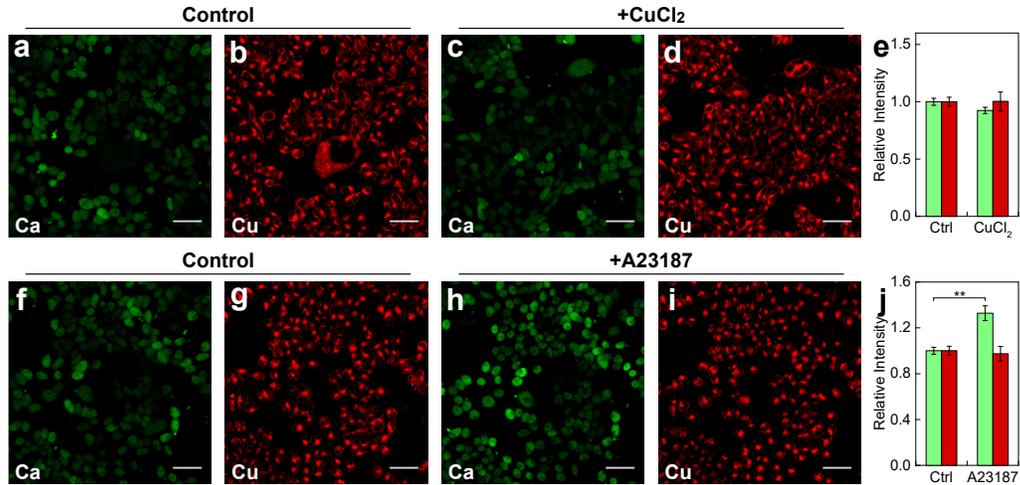
**Figure 2.7** Detection of total copper and zinc pools in (a) cytoplasmic extract and (b) nuclear extract of *Atp7a*<sup>-/-</sup> MEFs and its genetically matched controls. Values are shown as mean ± sem (n = 3).



**Figure 2.8** Fluorescence imaging of labile copper pools in MEF wildtype (WT) and *Atp 7a*<sup>-/-</sup> knockout fibroblast cells with CCF1. (a) MEF WT cells and (b) MEF *Atp7a*<sup>-/-</sup> knockout cells were stained with 2  $\mu$ M CCF1 for 10 min in DMEM and their average fluorescence intensity was (c) quantified. (d) MEF WT cells and (e) MEF *Atp7a*<sup>-/-</sup> knockout cells were stained with 2  $\mu$ M Ctrl-CCF1 for 10 min in DMEM and their average fluorescence intensity was (f) quantified. Scale-bars: 40  $\mu$ m. Data were normalized to MEF *atp7a* wt cells and shown as average  $\pm$  s.d. (CCF1: n = 4, Ctrl-CCF1: n = 3). \*\*P  $\leq$  0.01, \*\*\*P  $\leq$  0.001; two-tailed Student's t-test.



**Figure 2.9** Dual-channel imaging in HEK 293T cells with Calcium Green-1 in the green channel (a, c, f, h) and Ctrl-CPF1 in the red channel (b, d, g, i). Control cells (a, b) and cells treated with 100  $\mu$ M CuCl<sub>2</sub> for 18 h (c, d) were incubated with both dyes in HBSS and imaged; quantification is shown in (e). Cells incubated with both dyes in HBSS prior to (f, g) and after (h, i) treatment of 1  $\mu$ M calcium ionophore A23187 were imaged; quantification is shown in (j). Green bars represent the Calcium Green-1 channel and red bars are Ctrl-CPF1 channel. Scale-bars: 40  $\mu$ m. Data were normalized to controls cells and shown as average  $\pm$  s.d. (n = 4). \*\*\*P  $\leq$  0.001; two-tailed Student's t-test.



**Figure 2.10** Dual-channel imaging in HEK 293T cells with Calcium Green-1 in the green channel (a, c, f, h) and Ctrl-CPF1 in the red channel (b, d, g, i). Control cells (a, b) and cells treated with 100  $\mu$ M CuCl<sub>2</sub> for 18 h (c, d) were incubated with both dyes in HBSS and imaged; quantification is shown in (e). Cells incubated with both dyes in HBSS prior to (f, g) and after (h, i) treatment of 1  $\mu$ M calcium ionophore A23187 were imaged; quantification is shown in (j). Green bars represent the Calcium Green-1 channel and red bars are Ctrl-CPF1 channel. Scale-bars: 40  $\mu$ m. Data were normalized to controls cells and shown as average  $\pm$  s.d. (n = 4). \*\*\*P  $\leq$  0.001; two-tailed Student's t-test.

## 2.6 References

- (1) Lippard, S. J.; Berg, J. M. *Principles of Bioinorganic Chemistry*; University Science Books: Mill Valley, CA, 1994.
- (2) Tapiero, H.; Townsend, D. M.; Tew, K. D. Trace Elements in Human Physiology and Pathology. Copper. *Biomed. Pharmacother.* **2003**, *57*, 386–398.
- (3) Turski, M. L.; Brady, D. C.; Kim, H. J.; Kim, B. E.; Nose, Y.; Counter, C. M.; Winge, D. R.; Thiele, D. J. A Novel Role for Copper in Ras/Mitogen-Activated Protein Kinase Signaling. *Mol. Cell. Biol.* **2012**, *32*, 1284–1295.
- (4) Brady, D. C.; Crowe, M. S.; Turski, M. L.; Hobbs, G. A.; Yao, X.; Chaikuad, A.; Knapp, S.; Xiao, K.; Campbell, S. L.; Thiele, D. J.; et al. Copper Is Required for Oncogenic BRAF Signalling and Tumorigenesis. *Nature* **2014**, *509*, 492–496.
- (5) Krishnamoorthy, L.; Cotruvo, J. A. J.; Chan, J.; Kaluarachchi, H.; Muchenditsi, A.; Pendyala, V. S.; Jia, S.; Aron, A. T.; Vander Wal, M. N.; Guan, T.; et al. Copper Regulates Cyclic AMP-Dependent Lipolysis. *Nat. Chem. Biol.* **2016**, *12*, 586–592.
- (6) Halliwell, B.; Gutteridge, J. M. Role of Free Radicals and Catalytic Metal Ions in Human Disease: An Overview. *Methods Enzymol.* **1990**, *186*, 1–85.
- (7) Gaetke, L. M.; Chow, C. K. Copper Toxicity, Oxidative Stress, and Antioxidant Nutrients. *Toxicology* **2003**, *189*, 147–163.
- (8) Huffman, D. L.; O'Halloran, T. V. Function, Structure, and Mechanism of Intracellular Copper Trafficking Proteins. *Annu. Rev. Biochem.* **2001**, *70*, 677–701.
- (9) Luk, E.; Jensen, L. T.; Culotta, V. C. The Many Highways for Intracellular Trafficking of Metals. *J. Biol. Inorg. Chem.* **2003**, *8* (8), 803–809.
- (10) Robinson, N. J.; Winge, D. R. Copper Metallochaperones. *Annu. Rev. Biochem.* **2010**, *79*, 537–562.
- (11) Linz, R.; Lutsenko, S. Copper-Transporting ATPases ATP7A and ATP7B: Cousins, Not Twins. *J. Bioenerg. Biomembr.* **2007**, *39* (5–6), 403–407.
- (12) La Fontaine, S.; Mercer, J. F. Trafficking of the Copper-ATPases, ATP7A and ATP7B: Role in Copper Homeostasis. *Arch. Biochem. Biophys.* **2007**, *463*, 149–167.
- (13) Kim, B. E.; Nevitt, T.; Thiele, D. J. Mechanisms for Copper Acquisition, Distribution and Regulation. *Nat. Chem. Biol.* **2008**, *4*, 176–185.
- (14) Ayton, S.; Lei, P.; Bush, A. I. Metallostasis in Alzheimer's Disease. *Free Radic. Biol. Med.* **2013**, *62*, 76–89.
- (15) Barnham, K. J.; Masters, C. L.; Bush, A. I. Neurodegenerative Diseases and Oxidative Stress. *Nat. Rev. Drug Discov.* **2004**, *3*, 205–214.
- (16) Savelieff, M. G.; Lee, S.; Liu, Y.; Lim, M. H. Untangling Amyloid- $\beta$ , Tau, and Metals in Alzheimer's Disease. *ACS Chem. Biol.* **2013**, *8*, 856–865.
- (17) Matlack, K. E.; Tardiff, D. F.; Narayan, P.; Hamamichi, S.; Caldwell, K. A.; Caldwell, G. A.; Lindquist, S. Clioquinol Promotes the Degradation of Metal-Dependent Amyloid- $\beta$  (Ab) Oligomers to Restore Endocytosis and Ameliorate Ab Toxicity. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 4013–4018.

- (18) Vonk, W. I.; Kakkar, V.; Bartuzi, P.; Jaarsma, D.; Berger, R.; Hofker, M. H.; Klomp, L. W.; Wijmenga, C.; Kampinga, H. H.; van de Sluis, B. The Copper Metabolism MURR1 Domain Protein 1 (COMMD1) Modulates the Aggregation of Misfolded Protein Species in a Client-Specific Manner. *PLoS ONE* **2014**, *9*, e92408.
- (19) Xiao, G.; Fan, Q.; Wang, X.; Zhou, B. Huntington Disease Arises from a Combinatory Toxicity of Polyglutamine and Copper Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14995–15000.
- (20) Valentine, J. S.; Doucette, P. A.; Zittin Potter, S. Copper-Zinc Superoxide Dismutase and Amyotrophic Lateral Sclerosis. *Annu. Rev. Biochem.* **2005**, *74*, 563–593.
- (21) Gaggelli, E.; Kozlowski, H.; Valensin, D.; Valensin, G. Copper Homeostasis and Neurodegenerative Disorders (Alzheimer's, Prion, and Parkinson's Diseases and Amyotrophic Lateral Sclerosis). *Chem. Rev.* **2006**, *106*, 1995–2044.
- (22) Seetharaman, S. V.; Prudencio, M.; Karch, C.; Holloway, S. P.; Borchelt, D. R.; Hart, P. J. Immature Copper-Zinc Superoxide Dismutase and Familial Amyotrophic Lateral Sclerosis. *Exp. Biol. Med.* **2009**, *234* (10), 1140–1154.
- (23) Sheng, Y.; Chattopadhyay, M.; Whitelegge, J.; Valentine, J. S. SOD1 Aggregation and ALS: Role of Metallation States and Disulfide Status. *Curr. Top. Med. Chem.* **2012**, *12* (22), 2560–2572.
- (24) Burkhead, J. L.; Lutsenko, S. The Role of Copper as a Modifier of Lipid Metabolism. In *Lipid Metabolism*; Baez, R. V., Ed.; InTech: Rijeka, 2013; pp 39–60.
- (25) Huster, D.; Purnat, T. D.; Burkhead, J. L.; Ralle, M.; Fiehn, O.; Stuckert, F.; Olson, N. E.; Teupser, D.; Lutsenko, S. High Copper Selectively Alters Lipid Metabolism and Cell Cycle Machinery in the Mouse Model of Wilson Disease. *J. Biol. Chem.* **2007**, *282*, 8343–8355.
- (26) Huster, D.; Lutsenko, S. Wilson Disease: Not Just a Copper Disorder. Analysis of a Wilson Disease Model Demonstrates the Link between Copper and Lipid Metabolism. *Mol. Biosyst.* **2007**, *3*, 816–824.
- (27) Vulpe, C.; Levinson, B.; Whitney, S.; Packman, S.; Gitschier, J. Isolation of a Candidate Gene for Menkes Disease and Evidence That It Encodes a Copper-Transporting ATPase. *Nat. Genet.* **1993**, *3*, 7–13.
- (28) Kaler, S. G. ATP7A-Related Copper Transport Diseases - Emerging Concepts and Future Trends. *Nat. Rev. Neurol.* **2011**, *7*, 15–29.
- (29) Lutsenko, S. *Atp7b*<sup>-/-</sup> Mice as a Model for Studies of Wilson's Disease. *Biochem. Soc. Trans.* **2008**, *36*, 1233–1238.
- (30) Huster, D.; Hoppert, M.; Lutsenko, S.; Zinke, J.; Lehmann, C.; Mossner, J.; Berr, F.; Caca, K. Defective Cellular Localization of Mutant ATP7B in Wilson's Disease Patients and Hepatoma Cell Lines. *Gastroenterology* **2003**, *124*, 335–345.
- (31) Huster, D. Structural and Metabolic Changes in *Atp7b*<sup>-/-</sup> Mouse Liver and Potential for New Interventions in Wilson's Disease. *Ann. N. Y. Acad. Sci.* **2014**, *1314*, 37–44.
- (32) Chang, C. J. Searching for Harmony in Transition-Metal Signaling. *Nat. Chem. Biol.* **2015**, *11* (10), 744–747.

- (33) Chang, C. J. Bioinorganic Life and Neural Activity: Toward a Chemistry of Consciousness? *Acc. Chem. Res.* **2017**, *50* (3), 535–538.
- (34) Dodani, S. C.; Firl, A.; Chan, J.; Nam, C. I.; Aron, A. T.; Onak, C. S.; Ramos-Torres, K. M.; Paek, J.; Webster, C. M.; Feller, M. B.; et al. Copper Is an Endogenous Modulator of Neural Circuit Spontaneous Activity. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 16280–16285.
- (35) Dodani, S. C.; Domaille, D. W.; Nam, C. I.; Miller, E. W.; Finney, L. A.; Vogt, S.; Chang, C. J. Calcium-Dependent Copper Redistributions in Neuronal Cells Revealed by a Fluorescent Copper Sensor and X-Ray Fluorescence Microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 5980–5985.
- (36) Crabtree, R. H. Copper (I): A Possible Olfactory Binding Site. *J. Inorg. Nucl. Chem.* **1978**, *40* (7), 1453.
- (37) Duan, X.; Block, E.; Li, Z.; Connelly, T.; Zhang, J.; Huang, Z.; Su, X.; Pan, Y.; Wu, L.; Chi, Q.; et al. Crucial Role of Copper in Detection of Metal-Coordinating Odorants. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (9), 3492–3497.
- (38) Palmer, A. E.; Tsien, R. Y. Measuring Calcium Signaling Using Genetically Targetable Fluorescent Indicators. *Nat. Protoc.* **2006**, *1* (3), 1057–1065.
- (39) Domaille, D. W.; Que, E. L.; Chang, C. J. Synthetic Fluorescent Sensors for Studying the Cell Biology of Metals. *Nat. Chem. Biol.* **2008**, *4*, 168–175.
- (40) Que, E. L.; Domaille, D. W.; Chang, C. J. Metals in Neurobiology: Probing Their Chemistry and Biology with Molecular Imaging. *Chem. Rev.* **2008**, *108* (5), 1517–1549.
- (41) Nolan, E. M.; Lippard, S. J. Small-Molecule Fluorescent Sensors for Investigating Zinc Metalloneurochemistry. *Acc. Chem. Res.* **2009**, *42* (1), 193–203.
- (42) Dean, K. M.; Qin, Y.; Palmer, A. E. Visualizing Metal Ions in Cells: An Overview of Analytical Techniques, Approaches, and Probes. *Biochim. Biophys. Acta* **2012**, *1823* (9), 1406–1415.
- (43) Carter, K. P.; Young, A. M.; Palmer, A. E. Fluorescent Sensors for Measuring Metal Ions in Living Systems. *Chem. Rev.* **2014**, *114* (8), 4564–4601.
- (44) Aron, A. T.; Ramos-Torres, K. M.; Cotruvo, J. A.; Chang, C. J. Recognition- and Reactivity-Based Fluorescent Probes for Studying Transition Metal Signaling in Living Systems. *Acc. Chem. Res.* **2015**, *48*, 2434–2442.
- (45) Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. A Guide to Choosing Fluorescent Proteins. *Nat. Methods* **2005**, *2*, 905–909.
- (46) Giepmans, B. N.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. The Fluorescent Toolbox for Assessing Protein Location and Function. *Science* **2006**, *312*, 217–224.
- (47) Terai, T.; Nagano, T. Fluorescent Probes for Bioimaging Applications. *Curr Opin Chem Biol* **2008**, *12*, 515–521.
- (48) Lavis, L. D.; Raines, R. T. Bright Ideas for Chemical Biology. *ACS Chem. Biol.* **2008**, *3*, 142–155.
- (49) Fahrni, C. J. Synthetic Fluorescent Probes for Monovalent Copper. *Curr. Opin. Chem. Biol.* **2013**, *17* (4), 656–662.

- (50) Cotruvo, J. A. J.; Aron, A. T.; Ramos-Torres, K. M.; Chang, C. J. Synthetic Fluorescent Probes for Studying Copper in Biological Systems. *Chem. Soc. Rev.* **2015**, *44*, 4400–4414.
- (51) Ramos-Torres, K. M.; Kolemen, S.; Chang, C. J. Thioether Coordination Chemistry for Molecular Imaging of Copper in Biological Systems. *Isr. J. Chem.* **2016**, *56* (9–10), 724–737.
- (52) Liu, J.; Lu, Y. A DNAzyme Catalytic Beacon Sensor for Paramagnetic Cu<sup>2+</sup> Ions in Aqueous Solution with High Sensitivity and Selectivity. *J. Am. Chem. Soc.* **2007**, *129* (32), 9838–9839.
- (53) Wegner, S. V.; Arslan, H.; Sunbul, M.; Yin, J.; He, C. Dynamic Copper(I) Imaging in Mammalian Cells with a Genetically Encoded Fluorescent Copper(I) Sensor. *J. Am. Chem. Soc.* **2010**, *132* (8), 2567–2569.
- (54) Liu, J.; Karpus, J.; Wegner, S. V.; Chen, P. R.; He, C. Genetically Encoded Copper(I) Reporters with Improved Response for Use in Imaging. *J. Am. Chem. Soc.* **2013**, *135* (8), 3144–3149.
- (55) Choi, Y.-A.; Keem, J. O.; Kim, C. Y.; Yoon, H. R.; Heo, W. D.; Chung, B. H.; Jung, Y. A Novel Copper-Chelating Strategy for Fluorescent Proteins to Image Dynamic Copper Fluctuations on Live Cell Surfaces. *Chem. Sci.* **2015**, *6* (2), 1301–1307.
- (56) Santo, C. E.; Lam, E. W.; Elowsky, C. G.; Quaranta, D.; Domaille, D. W.; Chang, C. J.; Grass, G. Bacterial Killing by Dry Metallic Copper Surfaces. *Appl. Environ. Microbiol.* **2011**, *77* (3), 794–802.
- (57) Ladomersky, E.; Khan, A.; Shanbhag, V.; Cavet, J. S.; Chan, J.; Weisman, G. A.; Petris, M. J. Host and Pathogen Copper-Transporting P-Type ATPases Function Antagonistically during Salmonella Infection. *Infect. Immun.* **2017**, IAI.00351-17.
- (58) Quaranta, D.; Krans, T.; Santo, C. E.; Elowsky, C. G.; Domaille, D. W.; Chang, C. J.; Grass, G. Mechanisms of Contact-Mediated Killing of Yeast Cells on Dry Metallic Copper Surfaces. *Appl. Environ. Microbiol.* **2011**, *77* (2), 416–426.
- (59) Cusick, K. D.; Minkin, S. C.; Dodani, S. C.; Chang, C. J.; Wilhelm, S. W.; Sayler, G. S. Inhibition of Copper Uptake in Yeast Reveals the Copper Transporter Ctr1p As a Potential Molecular Target of Saxitoxin. *Environ. Sci. Technol.* **2012**, *46* (5), 2959–2966.
- (60) Beaudoin, J.; Ioannoni, R.; López-Maury, L.; Bähler, J.; Ait-Mohand, S.; Guérin, B.; Dodani, S. C.; Chang, C. J.; Labbé, S. Mfc1 Is a Novel Forespore Membrane Copper Transporter in Meiotic and Sporulating Cells. *J. Biol. Chem.* **2011**, *286* (39), 34356–34372.
- (61) Bernal, M.; Casero, D.; Singh, V.; Wilson, G. T.; Grande, A.; Yang, H.; Dodani, S. C.; Pellegrini, M.; Huijser, P.; Connolly, E. L.; et al. Transcriptome Sequencing Identifies SPL7-Regulated Copper Acquisition Genes FRO4/FRO5 and the Copper Dependence of Iron Homeostasis in Arabidopsis. *Plant Cell* **2012**, *24* (2), 738–761.
- (62) Chun, H.; Sharma, A. K.; Lee, J.; Chan, J.; Jia, S.; Kim, B.-E. The Intestinal Copper Exporter CUA-1 Is Required for Systemic Copper Homeostasis in *Caenorhabditis Elegans*. *J. Biol. Chem.* **2017**, *292* (1), 1–14.

- (63) Schrag, M.; Crofton, A.; Zabel, M.; Jiffry, A.; Kirsch, D.; Dickson, A.; Mao, X. W.; Vinters, H. V.; Domaille, D. W.; Chang, C. J.; et al. Effect of Cerebral Amyloid Angiopathy on Brain Iron, Copper, and Zinc in Alzheimer's Disease. *J. Alzheimers Dis.* **2011**, *24* (1), 137–149.
- (64) Öhrvik, H.; Nose, Y.; Wood, L. K.; Kim, B.-E.; Gleber, S.-C.; Ralle, M.; Thiele, D. J. Ctr2 Regulates Biogenesis of a Cleaved Form of Mammalian Ctr1 Metal Transporter Lacking the Copper- and Cisplatin-Binding Ecto-Domain. *Proc. Natl. Acad. Sci.* **2013**, *110* (46), E4279–E4288.
- (65) Huang, C. P.; Fofana, M.; Chan, J.; Chang, C. J.; Howell, S. B. Copper Transporter 2 Regulates Intracellular Copper and Sensitivity to Cisplatin. *Metallomics* **2014**, *6* (3), 654–661.
- (66) Polishchuk, E. V.; Concilli, M.; Iacobacci, S.; Chesi, G.; Pastore, N.; Piccolo, P.; Paladino, S.; Baldantoni, D.; van IJzendoorn, S. C. D.; Chan, J.; et al. Wilson Disease Protein ATP7B Utilizes Lysosomal Exocytosis to Maintain Copper Homeostasis. *Dev. Cell* **2014**, *29* (6), 686–700.
- (67) Achard, M. E. S.; Stafford, S. L.; Bokil, N. J.; Chartres, J.; Bernhardt, P. V.; Schembri, M. A.; Sweet, M. J.; McEwan, A. G. Copper Redistribution in Murine Macrophages in Response to Salmonella Infection. *Biochem. J.* **2012**, *444* (1), 51–57.
- (68) Kashyap, D. R.; Rompca, A.; Gaballa, A.; Helmann, J. D.; Chan, J.; Chang, C. J.; Hozo, I.; Gupta, D.; Dziarski, R. Peptidoglycan Recognition Proteins Kill Bacteria by Inducing Oxidative, Thiol, and Metal Stress. *PLoS Pathog.* **2014**, *10* (7), e1004280.
- (69) Hong-Hermesdorf, A.; Miethke, M.; Gallaher, S. D.; Kropat, J.; Dodani, S. C.; Chan, J.; Barupala, D.; Domaille, D. W.; Shirasaki, D. I.; Loo, J. A.; et al. Subcellular Metal Imaging Identifies Dynamic Sites of Cu Accumulation in Chlamydomonas. *Nat. Chem. Biol.* **2014**, *10* (12), 1034–1042.
- (70) Yang, L.; McRae, R.; Henary, M. M.; Patel, R.; Lai, B.; Vogt, S.; Fahrni, C. J. Imaging of the Intracellular Topography of Copper with a Fluorescent Sensor and by Synchrotron X-Ray Fluorescence Microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (32), 11179–11184.
- (71) Morgan, M. T.; Bagchi, P.; Fahrni, C. J. Designed To Dissolve: Suppression of Colloidal Aggregation of Cu(I)-Selective Fluorescent Probes in Aqueous Buffer and In-Gel Detection of a Metallochaperone. *J. Am. Chem. Soc.* **2011**, *133* (40), 15906–15909.
- (72) Morgan, M. T.; McCallum, A. M.; Fahrni, C. J. Rational Design of a Water-Soluble, Lipid-Compatible Fluorescent Probe for Cu(I) with Sub-Part-per-Trillion Sensitivity. *Chem. Sci.* **2016**, *7* (2), 1468–1473.
- (73) Lim, C. S.; Han, J. H.; Kim, C. W.; Kang, M. Y.; Kang, D. W.; Cho, B. R. A Copper(I)-Ion Selective Two-Photon Fluorescent Probe for in Vivo Imaging. *Chem. Commun.* **2011**, *47* (25), 7146–7148.
- (74) Zeng, L.; Miller, E. W.; Pralle, A.; Isacoff, E. Y.; Chang, C. J. A Selective Turn-On Fluorescent Sensor for Imaging Copper in Living Cells. *J. Am. Chem. Soc.* **2006**, *128* (1), 10–11.

- (75) Dodani, S. C.; Leary, S. C.; Cobine, P. A.; Winge, D. R.; Chang, C. J. A Targetable Fluorescent Sensor Reveals That Copper-Deficient SCO1 and SCO2 Patient Cells Prioritize Mitochondrial Copper Homeostasis. *J. Am. Chem. Soc.* **2011**, *133* (22), 8606–8616.
- (76) Hirayama, T.; Van de Bittner, G. C.; Gray, L. W.; Lutsenko, S.; Chang, C. J. Near-Infrared Fluorescent Sensor for in Vivo Copper Imaging in a Murine Wilson Disease Model. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (7), 2228–2233.
- (77) Cao, X.; Lin, W.; Wan, W. Development of a Near-Infrared Fluorescent Probe for Imaging of Endogenous Cu<sup>+</sup> in Live Cells. *Chem. Commun.* **2012**, *48* (50), 6247–6249.
- (78) Arden-Jacob, J.; Frantzeskos, J.; Kemnitzer, N. U.; Zilles, A.; Drexhage, K. H. New Fluorescent Markers for the Red Region. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2001**, *57*, 2271–2283.
- (79) Ohulchanskyy, T. Y.; Donnelly, D. J.; Detty, M. R.; Prasad, P. N. Heteroatom Substitution Induced Changes in Excited-State Photophysics and Singlet Oxygen Generation in Chalcogenoxanthylum Dyes: Effect of Sulfur and Selenium Substitutions. *J. Phys. Chem. B* **2004**, *108*, 8668–8672.
- (80) Calitree, B.; Donnelly, D. J.; Holt, J. J.; Gannon, M. K.; Nygren, C. L.; Sukumaran, D. K.; Autschbach, J.; Detty, M. R. Tellurium Analogues of Rosamine and Rhodamine Dyes: Synthesis, Structure, <sup>125</sup>Te NMR, and Heteroatom Contributions to Excitation Energies. *Organometallics* **2007**, *26*, 6248–6257.
- (81) Kolmakov, K.; Belov, V. N.; Bierwagen, J.; Ringemann, C.; Muller, V.; Eggeling, C.; Hell, S. W. Red-Emitting Rhodamine Dyes for Fluorescence Microscopy and Nanoscopy. *Chemistry* **2010**, *16*, 158–166.
- (82) Egawa, T.; Koide, Y.; Hanaoka, K.; Komatsu, T.; Terai, T.; Nagano, T. Development of a Fluorescein Analogue, TokyoMagenta, as a Novel Scaffold for Fluorescence Probes in Red Region. *Chem. Commun.* **2011**, *47*, 4162–4164.
- (83) Koide, Y.; Urano, Y.; Hanaoka, K.; Terai, T.; Nagano, T. Evolution of Group 14 Rhodamines as Platforms for Near-Infrared Fluorescence Probes Utilizing Photoinduced Electron Transfer. *ACS Chem. Biol.* **2011**, *6*, 600–608.
- (84) Kolmakov, K.; Wurm, C.; Sednev, M. V.; Bossi, M. L.; Belov, V. N.; Hell, S. W. Masked Red-Emitting Carbopyronine Dyes with Photosensitive 2-Diazo-1-Indanone Caging Group. *Photochem. Photobiol. Sci.* **2012**, *11*, 522–532.
- (85) Kushida, Y.; Hanaoka, K.; Komatsu, T.; Terai, T.; Ueno, T.; Yoshida, K.; Uchiyama, M.; Nagano, T. Red Fluorescent Scaffold for Highly Sensitive Protease Activity Probes. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 3908–3911.
- (86) Grimm, J. B.; Sung, A. J.; Legant, W. R.; Hulamm, P.; Matlosz, S. M.; Betzig, E.; Lavis, L. D. Carbofluoresceins and Carborhodamines as Scaffolds for High-Contrast Fluorogenic Probes. *ACS Chem. Biol.* **2013**, *8*, 1303–1310.
- (87) Lukinavicius, G.; Umezawa, K.; Olivier, N.; Honigmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Correa Jr, I. R.; Luo, Z.-G.; et al. A Near-Infrared Fluorophore for

- Live-Cell Super-Resolution Microscopy of Cellular Proteins. *Nat. Chem.* **2013**, *5*, 132–139.
- (88) Fukazawa, A.; Suda, S.; Taki, M.; Yamaguchi, E.; Grzybowski, M.; Sato, Y.; Higashiyama, T.; Yamaguchi, S. Phospha-Fluorescein: A Red-Emissive Fluorescein Analogue with High Photobleaching Resistance. *Chem. Commun.* **2016**, *52*, 1120–1123.
- (89) Zhou, X.; Lai, R.; Beck, J. R.; Li, H.; Stains, C. I. Nebraska Red: A Phosphinate-Based near-Infrared Fluorophore Scaffold for Chemical Biology Applications. *Chem. Commun.* **2016**, *52* (83), 12290–12293.
- (90) Lavis, L. D. Teaching Old Dyes New Tricks: Biological Probes Built from Fluoresceins and Rhodamines. *Annu. Rev. Biochem.* **2017**, *86* (1), 825–843.
- (91) Weissleder, R. A Clearer Vision for in Vivo Imaging. *Nat. Biotechnol.* **2001**, *19*, 316–317.
- (92) Martell, A. E.; Smith, R. M. Amides. In *Critical Stability Constants*; Critical Stability Constants; Springer US, 1982; pp 380–391.

# Chapter 3

## **Development of thiophosphorodichloridate reagents for chemoselective histidine bioconjugation**

Portions of this work were published in the following scientific journal:

Jia, S.; He, D.; Chang, C. J. Bioinspired Thiophosphorodichloridate Reagents for Chemoselective Histidine Bioconjugation. *J. Am. Chem. Soc.* **2019**, *141* (18), 7294–7301.

## Abstract

Site-selective bioconjugation to native protein residues is a powerful tool for protein functionalization, with cysteine and lysine side chains being the most common points for attachment owing to their high nucleophilicity. In this chapter we describe the strategy for histidine modification using thiophosphorodichloridate reagents that mimic post-translational histidine phosphorylation, enabling fast and selective labeling of protein histidines under mild conditions where various payloads can be introduced via copper-assisted alkyne-azide cycloaddition (CuAAC) chemistry. We establish that these reagents are particularly effective at covalent modification of His-tags, which are common motifs to facilitate protein purification, as illustrated by selective attachment of polyarginine cargoes to enhance the uptake of proteins into living cells. This work provides a starting point for probing and enhancing protein function using histidine-directed chemistry.

## 3.1 Introduction

Site-selective bioconjugation chemistry offers a versatile strategy to probe and expand the function of proteins.<sup>1-5</sup> The most common and robust chemoselective and regioselective protein bioconjugation strategies have focused on functionalization of cysteine thiol<sup>6-9</sup> and lysine amine<sup>10-12</sup> sites and related nucleophilic hydroxyl<sup>13,14</sup> and carboxyl<sup>15</sup> side chains. More recent advances in protein bioconjugation technologies have targeted less nucleophilic amino acids,<sup>16</sup> including tyrosine,<sup>17-19</sup> tryptophan<sup>20,21</sup> and methionine.<sup>22,23</sup> In contrast, selective modification of histidine, which is commonly found in enzyme active sites and metal-binding sites,<sup>24,25</sup> remains underexplored. Because the imidazole side chain of histidine is a good metal ligand, metal coordination can enable protein modification through metal-directed covalent labeling proximal to the histidine group<sup>26,27</sup> or direct non-covalent metal-histidine complexation,<sup>28-32</sup> the latter of which can be labile under biological contexts or mass spectrometry conditions. On the other hand, histidine is a useful catalytic component owing to its ability to serve as both a good nucleophile and leaving group, but this character also makes it difficult to form stable bonds with the imidazole side chain through electrophilic functionalization. Indeed, selected nucleophiles including epoxides are histidine-reactive but typically require harsh reaction conditions such as high temperatures and/or strong bases<sup>33,34</sup>, as well as sequence-dependent motifs<sup>35,36</sup> and/or an affinity-directed ligand.<sup>37-40</sup> As such, selective and direct covalent labeling of histidine remains a challenge.

In this chapter, we report the design, synthesis and application of our first generation of thiophosphorodichloridate compound, TPAC, as a chemoselective histidine bioconjugation reagent. The thiophosphorodichloridate reagent forms a biomimetic phosphohistidine modification with an alkyne handle that enables further click reaction with more complex payloads. We show that TPAC is able to label both small molecule Fmoc-His-OH and various protein models with fast kinetics and high histidine-specificity. The labeled model protein can be purified by ion exchange column and can undergo click reaction with various organic azide molecules. Finally, as an application of this bioconjugation method, we show that this reagent is

especially efficient at labeling polyhistidine tags, and after CuAAC with a polyarginine cell-penetrating peptide, enables efficient protein delivery into living cells. This study provides a starting point for probing and enhancing protein function using histidine-directed chemistry.

## 3.2 Methods

### 3.2.1 General materials and synthetic methods

All commercial reagents were used without further purification. Fmoc-Cys-OH was purchased from Chem-Impex (Wood Dale IL). All other Fmoc-protected amino acids were purchased from Ark Pharm (Arlington Heights IL). 2-(4-((bis((1-(tert-butyl)-1H-1,2,3-triazol-4-yl)methyl)amino)methyl)-1H-1,2,3-triazol-1-yl)acetic acid (BTAA) was purchased from Click Chemistry Tools (Scottsdale AZ). Custom peptides for model reactions were purchased from GenScript (Piscataway NJ). Ribonuclease A from bovine pancreas, lyzome from chicken egg white, myoglobin from equine heart and all other chemicals were purchased from Sigma-Aldrich (St. Louis MO). Calmodulin, porcine was purchased from rPeptide (Watkinsville GA). His-tagged GFP was purchased from Sino Biological (Beijing, China). His-tagged mCherry was purchased from Origene (Rockville MD). Enterokinase was purchased from New England Biolabs (Ipswich MA). TEV protease was purchased from QB3 (San Francisco CA). Enzymes were used following the protocols provided by their manufacturer. 2-Azidoethanol<sup>41</sup> and succinimidyl 3-azidopropanoate<sup>42</sup> were synthesized according to literature procedure. 3-Azidopropanoic acid was purchased from Click Chemistry Tools. H-(Arg)<sub>9</sub>-OH trifluoroacetate salt was purchased from Bachem (Bubendorf, Switzerland). All other chemicals were purchased from Sigma-Aldrich.

All reactions utilizing air- or moisture-sensitive reagents were performed in dried glassware under an atmosphere of dry N<sub>2</sub>. THF used for anhydrous reactions was dried and stored over 4 Å molecular sieves. <sup>1</sup>H NMR, <sup>13</sup>C NMR and <sup>31</sup>P NMR spectra were collected in CDCl<sub>3</sub>, MeOD or *d*<sub>6</sub>-acetone (Cambridge Isotope Laboratories, Cambridge MA) at 25 °C on AVB-400, AVQ-400 or AV-600 spectrometers at the College of Chemistry NMR Facility at the University of California, Berkeley. All chemical shifts in <sup>1</sup>H NMR and <sup>13</sup>C NMR are reported in the standard  $\delta$  notation of ppm relative to residual solvent peak (CDCl<sub>3</sub>  $\delta$ H=7.26,  $\delta$ C=77.16; MeOD  $\delta$ H=3.31,  $\delta$ C=49.00; *d*<sub>6</sub>-acetone:  $\delta$ H=2.05,  $\delta$ C=29.84), and for <sup>31</sup>P NMR 85% phosphoric acid in sealed capillary tube is used as internal standard ( $\delta$ P=0.00). Splitting patterns are indicated as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad. Low resolution electron ionization mass spectral analysis was carried out using Agilent 5975C 7890A GC/MS System. High resolution mass spectral analysis (ESI-MS and APCI-MS) were carried out at LBNL Catalysis Facility at the Lawrence Berkeley National Laboratory (Berkeley Lab) using PerkinElmer AxION® 2 TOF MS.

### 3.2.2 Synthesis of thiophosphorodichloridate reagents and related compounds

**Methyl propargyl thiophosphorochloridate (2).** To a flask containing  $\text{PSCl}_3$  (2 mL, 20 mmol) cooled in an ice/water bath was added dry methanol (2.0 mL, 49 mmol) dropwise. The mixture was stirred for 15 min on ice and excess methanol was distilled under vacuum at the same temperature to give crude methyl thiophosphorodichloridate. Sodium (0.453 g, 19.7 mmol) was dissolved in cooled propargyl alcohol (6.0 mL, 104 mmol) to form an orange, thick solution, which was added dropwise to methyl thiophosphorodichloridate cooled in an ice/water bath. The suspension was further stirred for 2 h at room temperature, diluted with  $\text{CH}_2\text{Cl}_2$ , filtered and purified by column chromatography (30:1 hexanes/EtOAc) to give product **2** as a colorless oil (2.4 g, 67%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.85 – 4.80 (m, 2H), 3.91 (d,  $J = 16.1$  Hz, 3H), 2.63 (t,  $J = 2.5$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  77.16 (s), 76.39 (d,  $J = 10.5$  Hz), 56.90 (d,  $J = 4.4$  Hz), 55.98 (d,  $J = 7.1$  Hz).  $^{31}\text{P}$  NMR (162 MHz,  $\text{CDCl}_3$ )  $\delta$  71.91. HRMS (APCI<sup>+</sup>)  $m/z$  calcd 184.9587, found 184.9586 for  $\text{C}_4\text{H}_7\text{ClO}_2\text{PS}^+$  ( $\text{M}+\text{H}^+$ ).

**Methyl propargyl thiophosphoramidate (1).** Crude **2** in propargyl alcohol and  $\text{CH}_2\text{Cl}_2$  was prepared as described above. Excess ammonia was led through this mixture to form  $\text{NH}_4\text{Cl}$  as a precipitate. The mixture was filtered, concentrated and purified by column chromatography (2:1 hexanes/EtOAc) to give product **1** as a light yellow oil (2.1 g, 65% overall yield).  $^1\text{H}$  NMR (400 MHz, MeOD)  $\delta$  4.62 (dd,  $J = 10.5, 2.5$  Hz, 2H), 3.68 (d,  $J = 13.8$  Hz, 3H), 2.94 (t,  $J = 2.5$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz, MeOD)  $\delta$  79.47 (d,  $J = 10.9$  Hz), 76.35, 55.04 (d,  $J = 2.8$  Hz), 53.72 (d,  $J = 5.3$  Hz).  $^{31}\text{P}$  NMR (162 MHz, MeOD)  $\delta$  78.25. HRMS (APCI<sup>+</sup>)  $m/z$  calcd 166.0086, found 166.0097 for  $\text{C}_4\text{H}_9\text{NO}_2\text{PS}^+$  ( $\text{M}+\text{H}^+$ ).

**Propargyl thiophosphorodichloridate (4, TPAC).** To a solution of  $\text{PSCl}_3$  (1.0 mL, 9.8 mmol) in  $\text{CH}_2\text{Cl}_2$  (10 mL) was added propargyl alcohol (0.57 mL, 9.8 mmol) and  $\text{K}_2\text{CO}_3$  (1.36 g, 9.8 mmol). After overnight stirring at room temperature, the mixture was filtered and purified by column chromatography (50:1 hexanes/EtOAc) to give product **TPAC** as a colorless oil with a pungent smell (1.1 g, 55%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.94 (dd,  $J = 15.9, 2.5$  Hz, 2H), 2.72 (t,  $J = 2.5$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  78.37, 75.48 (d,  $J = 11.1$  Hz), 58.52 (d,  $J = 7.4$  Hz).  $^{31}\text{P}$  NMR (162 MHz,  $\text{CDCl}_3$ )  $\delta$  59.68. HRMS (APCI<sup>+</sup>)  $m/z$  calcd 188.9092, found 188.9115 for  $\text{C}_3\text{H}_4\text{Cl}_2\text{OPS}^+$  ( $\text{M}+\text{H}^+$ ).

**Propargyl phosphorodichloridate (5).**  $\text{POCl}_3$  (1.0 mL, 11 mmol) and propargyl alcohol (0.62 mL, 11 mmol) was dissolved in  $\text{Et}_2\text{O}$  (20 mL) and cooled in a dry ice/acetone bath under  $\text{N}_2$ . Triethylamine (1.5 mL, 11 mmol) was dissolved in  $\text{Et}_2\text{O}$  (20 mL) and added dropwise via addition funnel to form a white suspension. The reaction mixture was warmed to room temperature and was further stirred at room temperature for 2 h. Trimethylamine hydrochloride was removed by filtration, and the solution was concentrated to give product **5** as a light yellow oil with a pungent smell (1.8 g, 99%). The product was used without further purification.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.92 (dd,  $J = 14.5, 1.0$  Hz, 2H), 2.76 (t,  $J = 1.0$  Hz

1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 78.93, 74.91 (d, *J* = 9.6 Hz), 58.53 (d, *J* = 7.2 Hz). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ 8.70. LRMS (EI<sup>+</sup>) *m/z* calcd 137.0, found 136.9 for C<sub>3</sub>H<sub>3</sub>ClO<sub>2</sub>P<sup>+</sup> (M-Cl<sup>-</sup>).

**Methyl propargyl phosphorochloridate (3).** Crude **5** with trimethylamine hydrochloride in Et<sub>2</sub>O was prepared as described above. The mixture was cooled again in a dry ice/acetone bath. Dry methanol (0.43 mL, 11 mmol) and trimethylamine (1.5 mL, 11 mmol) in Et<sub>2</sub>O (20 mL) was added dropwise via addition funnel. The slurry was further stirred at room temperature for 3 h, filtered and concentrated to give product **3** as light yellow oil with a pungent smell (1.7 g, 93%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.79 (dd, *J* = 11.5, 2.5 Hz, 2H), 3.92 (d, *J* = 13.8 Hz, 3H), 2.66 (t, *J* = 2.5 Hz, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 78.92, 76.04 (d, *J* = 9.0 Hz), 56.66 (d, *J* = 5.3 Hz), 55.97 (d, *J* = 7.0 Hz). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ 6.74. HRMS (APCI<sup>+</sup>) *m/z* calcd 168.9816, found 168.9855 for C<sub>4</sub>H<sub>7</sub>ClO<sub>3</sub>P<sup>+</sup> (M+H<sup>+</sup>).

***N*-propargyl thiophosphoramidic dichloride (6).** To a solution of PSCl<sub>3</sub> (0.30 mL, 2.9 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (5 mL) was added propargyl amine (0.19 mL, 2.9 mmol) and K<sub>2</sub>CO<sub>3</sub> (0.41 g, 2.9 mmol). After 2 h stirring at room temperature, the mixture was filtered and purified by column chromatography (30:1 hexanes/EtOAc) to give product **6** as a yellow oil with a pungent smell (0.25 g, 49%). <sup>1</sup>H NMR (400 MHz, *d*<sub>6</sub>-acetone) δ 4.09 (dd, *J* = 20.7, 2.5 Hz, 2H), 2.89 (t, *J* = 2.5 Hz, 1H). <sup>13</sup>C NMR (101 MHz, *d*<sub>6</sub>-acetone) δ 79.90 (d, *J* = 8.8 Hz), 74.13, 33.42. <sup>31</sup>P NMR (162 MHz, *d*<sub>6</sub>-acetone) δ 57.59. LRMS (EI<sup>+</sup>) *m/z* calcd 151.9, found 152.0 for C<sub>3</sub>H<sub>4</sub>ClNPS<sup>+</sup> (M-Cl<sup>-</sup>).

**Propargyl dithiophosphorodichloridate (7).** *S*-propargyl thioacetate (0.35 mg, 3.1 mmol) was stirred with NaOH (0.16 g, 4.0 mmol) in MeOH (10 mL) under N<sub>2</sub> for 30 min. The mixture was then diluted with CH<sub>2</sub>Cl<sub>2</sub> (30 mL), washed with H<sub>2</sub>O (×4) and dried (Na<sub>2</sub>SO<sub>4</sub>). To this solution was added PSCl<sub>3</sub> (0.37 mL, 3.7 mmol) and K<sub>2</sub>CO<sub>3</sub> (0.51 g, 3.7 mmol) and the mixture was stirred overnight. The mixture was filtered and purified by column chromatography (100:1 hexanes/EtOAc) to give product **7** as a light yellow oil, which degenerates into a solid mixture within minutes after concentration. We were able to obtain a crude <sup>1</sup>H NMR before degeneration and also confirmed its negligible reactivity in aqueous buffer with histidine when generated in situ. <sup>1</sup>H NMR (600 MHz, CDCl<sub>3</sub>) δ 3.75 (dd, *J* = 15.6, 2.7 Hz, 2H), 2.35 (t, *J* = 2.7 Hz, 1H).

**2-Azidoethyl thiophosphorodichloridate (8).** 2-Azidoethanol (0.43 g, 4.9 mmol) was dissolved in dry THF (15 mL) under N<sub>2</sub> and cooled in a dry ice/acetone bath. To this solution was added dropwise *n*BuLi (2.0 mL, 2.5 M in hexanes) and stirred for 20 min at room temperature to form the lithium salt. To another flask cooled in a dry ice/acetone bath was added THF (15 mL) and PSCl<sub>3</sub> (1.0 mL, 4.9 mmol) under N<sub>2</sub>. The lithium salt solution was then added dropwise at this temperature and the mixture was then stirred for 1h at room temperature. The mixture was concentrated, diluted in CH<sub>2</sub>Cl<sub>2</sub>, filtered to remove lithium chloride and purified by column chromatography (50:1 hexanes/EtOAc) to give product **8** as a light yellow oil with a pungent smell (0.60 g, 56%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.45 (dt, *J* = 11.0, 5.1 Hz, 2H), 3.63 (t, *J* =

4.4 Hz, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  69.57 (d,  $J = 9.9$  Hz), 50.11 (d,  $J = 10.6$  Hz).  $^{31}\text{P}$  NMR (162 MHz,  $\text{CDCl}_3$ )  $\delta$  59.05. HRMS (APCI $^+$ )  $m/z$  calcd 191.9202, found 191.9219 for  $\text{C}_2\text{H}_5\text{Cl}_2\text{NOPS}^+$  ( $\text{M}-\text{N}_2+\text{H}^+$ ).

**( $\pm$ )-3-butyn-2-yl thiophosphorodichloridate (9).** ( $\pm$ )-3-Butyn-2-ol (0.77 mL, 9.8 mmol) was dissolved in dry THF (10 mL) under  $\text{N}_2$  and cooled in a dry ice/acetone bath. To this solution was added dropwise  $n\text{BuLi}$  (3.9 mL, 2.5 M in hexanes) and stirred for 20 min at room temperature to form the lithium salt. To another flask cooled in a dry ice/acetone bath was added THF (20 mL) and  $\text{PSCl}_3$  (1.0 mL, 9.8 mmol) under  $\text{N}_2$ . The lithium salt solution was then added dropwise at this temperature and the mixture was then stirred for 3h at room temperature. The mixture was concentrated, diluted in  $\text{CH}_2\text{Cl}_2$ , filtered to remove lithium chloride and purified by column chromatography (100:1 hexanes/ $\text{EtOAc}$ ) to give product **9** as a colorless oil with a pungent smell (0.53 g, 26%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.46 (dq,  $J = 13.3, 6.6, 2.1$  Hz, 1H), 2.71 (d,  $J = 2.2$  Hz, 1H), 1.71 (dd,  $J = 6.6, 0.8$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  79.95 (d,  $J = 6.4$  Hz), 76.64, 69.20 (d,  $J = 8.6$  Hz), 23.18 (d,  $J = 7.6$  Hz).  $^{31}\text{P}$  NMR (162 MHz,  $\text{CDCl}_3$ )  $\delta$  58.98. HRMS (APCI $^+$ )  $m/z$  calcd 202.9249, found 202.9261 for  $\text{C}_4\text{H}_6\text{Cl}_2\text{OPS}^+$  ( $\text{M}+\text{H}^+$ ).

**Potassium *O*-propargyl ( $\text{N}^2$ -Fmoc-L-histidino)thiophosphate.** Fmoc-His-OH (0.20 g, 0.53 mmol) and  $\text{K}_2\text{CO}_3$  (1.1 g, 8.0 mmol) was dissolved in 4:6  $\text{H}_2\text{O}/\text{MeCN}$  (15 mL). To this mixture was added TPAC (0.30 g, 1.6 mmol) in MeCN (5 mL) over 1 h under vigorous stirring. The mixture was further stirred for 2 h and the desired product was separated by RP-HPLC to give the product as an off-white solid (0.22 g, 75%). This compound has the same retention time and  $m/z$  on LC/MS as the product in Figure 3.4a.  $^1\text{H}$  NMR (400 MHz, MeOD)  $\delta$  8.87 (s, 1H), 7.78 (d,  $J = 7.4$  Hz, 2H), 7.63 (d,  $J = 7.2$  Hz, 2H), 7.49 (s, 1H), 7.34 (dt,  $J = 28.2, 7.3$  Hz, 4H), 4.68 – 4.58 (m, 2H), 4.51 (dd,  $J = 9.3, 4.8$  Hz, 1H), 4.37 – 4.28 (m, 2H), 4.20 (t,  $J = 6.7$  Hz, 1H), 3.27 (d,  $J = 4.7$  Hz, 1H), 3.09 (dd,  $J = 15.1, 9.7$  Hz, 1H), 2.76 – 2.66 (m, 1H).  $^{31}\text{P}$  NMR (162 MHz, MeOD)  $\delta$  49.01. HRMS (ESI $^+$ )  $m/z$  calcd 512.1040, found 512.1092 for  $\text{C}_{24}\text{H}_{23}\text{N}_3\text{O}_6\text{PS}^+$  ( $\text{M}+\text{H}^+$ ).

**3-Azidoacetyl-(Arg) $_9$ -OH.** H-(Arg) $_9$ -OH trifluoroacetate salt (1 mg, approx. 0.5  $\mu\text{mol}$ ) was dissolved in 400  $\mu\text{L}$   $\text{H}_2\text{O}$  in an Eppendorf tube. To this solution was added HEPES buffer (100  $\mu\text{L}$ , 50 mM  $\text{pH}=7.5$ ) and succinimidyl 3-azidopropanoate (2.54 mg, 12  $\mu\text{mol}$ ) in DMSO (500  $\mu\text{L}$ ). The solution was mixed thoroughly and reacted in the dark for 3 h. The solution was then loaded onto a strong cation exchange spin column (Thermo Fisher). The column was washed with 25 mM ammonium acetate ( $2 \times 400$   $\mu\text{L}$ ) and eluted with 2 M ammonium acetate. The eluent was concentrated in a vacuum chamber overnight, and the residue was dissolved in water to a final volume of 50  $\mu\text{L}$  to give an approx. 10 mM stock solution. HRMS (ESI $^+$ )  $m/z$  calcd 507.6567, found 507.6502 for  $\text{C}_{57}\text{H}_{116}\text{N}_{39}\text{O}_{11}^{3+}$  ( $\text{M}+3\text{H}^+$ ); calcd 514.9840, found 514.9745 for  $\text{C}_{57}\text{H}_{115}\text{N}_{39}\text{NaO}_{11}^{3+}$  ( $\text{M}+2\text{H}^++\text{Na}^+$ ).

### 3.2.3 TPAC labeling on amino acids and proteins

For reaction with small molecule amino acid derivatives, Fmoc-protected amino acid was dissolved in DMSO (100 mM) and diluted to a final concentration of 0.5 mM in 25 mM HEPES buffer, pH 8.5 containing 20% MeCN unless otherwise noted. To this solution was added 5 mM of phosphorus electrophile (20× stock in MeCN). The reaction was performed at room temperature for 1 h unless otherwise noted, filtered and immediately subject to LC/MS analysis. For reaction with Fmoc-protected amino acid mixtures, 0.3 mM of each protected amino acid in 50 mM HEPES buffer, pH 8.5 containing 20% MeCN was reacted with 3 mM of TPAC.

For labeling on protein substrates, proteins were diluted into 25 mM HEPES, pH 8.5 to a concentration of 20 μM unless otherwise noted. Samples were labeled with 2 mM of TPAC (50× stock in MeCN). The reaction was performed at room temperature for 1 h. Samples were then subject to LC/MS analysis or further click reaction or LC-MS analysis.

### 3.2.3 Purification of TPAC-labeled ribonuclease A.

The reaction mixture of RNase A with TPAC was diluted in 20 mM sodium succinate, pH 4.2 and was loaded onto HiTrap SP HP cation exchange chromatography column (1 mL, GE Healthcare, Chicago IL). The protein was then separated with a linear gradient from 20 mM sodium succinate, pH 4.2 to 25 mM succinate with 1 M NaCl, pH 4.2 at a flow rate of 1.0 mL/min over 40 min. The collected fractions were concentrated and buffer-exchanged into desired buffer for click reaction or LC/MS analysis.

### 3.2.4 Click reactions on TPAC-labeled proteins.

We followed the protocol recommended by M. Finn, *et al* for click reactions.<sup>43</sup> Namely, protein solution in HEPES or PBS buffer was treated with 1 mM aminoguanidine hydrochloride (100× stock in water), 100 μM CuSO<sub>4</sub> (100× stock in water), 500 μM THPTA or BTAA (100× stock in water), 100 μM organic azide (100× stock in DMSO or water) and 5 mM sodium ascorbate (100× stock in water). For reaction on His-tagged protein 100 μM NiCl<sub>2</sub> (100× stock in water) was also added. The reaction was mixed thoroughly and placed in the dark at room temperature. After 1 h the reaction was quenched by adding 500 μM EDTA.

For click reaction with Cy3-N<sub>3</sub>, proteins were precipitated with acetone prior to click reaction to remove excess TPAC. The pellet was dissolved in PBS containing 0.1% SDS for click reaction, and precipitated again to remove unreacted dye. The pellet was then dissolved in running buffer and analyzed by SDS-PAGE.

For click reaction on TPAC-labeled, His-tagged proteins, proteins were buffer-exchanged into 25 mM HEPES, pH 7.5 by extensive ultrafiltration (Amnicon 10K, EMD Millipore, Hayward CA) prior to click reaction. After the click reaction, proteins were buffer-exchanged into PBS by extensive ultrafiltration to remove unwanted chemicals.

### 3.2.5 Cell culture and imaging

Cells were grown in the Cell Culture Facility at the University of California, Berkeley. HeLa cells were cultured in DMEM supplemented with 10% FBS and glutamine (2 mM). One day before imaging, cells were passed and plated on eight-well chamber slides (Lab-Tek, Thermo Fisher).

For imaging, cells were grown on 8-well chamber slides (LabTek, Thermo Fisher) to desired confluency, washed with PBS and incubated with 0.1 mg/mL (Arg)<sub>9</sub>-labeled GFP or mCherry in PBS for 30 min at 37 °C. The cells were then washed with PBS and stained with 1 μM Hoechst 33342 for 15 min, washed again and imaged on a Zeiss LSM710 laser-scanning microscope with a 20× objective lens for quantification and a 63× oil-immersion objective lens for images. Excitation was provided at 405 for Hoechst 33342, 488 nm for GFP and 543 nm for mCherry.

### 3.2.6 Image analysis and quantification

ImageJ (National Institutes of Health) was used for image analysis. For quantification of fluorescence intensity, each image was set to 8-bit greyscale and inverted. The fluorescence intensity was estimated using non-calibrated OD function. The area of cells were selected based on the brightfield image using edge detection and morphological operations.<sup>44</sup> The statistics of selected pixels was then measured by the “Measure” function and the average fluorescence intensity was obtained by dividing the integrated density (IntDen) over area. For each biological replicate, five images of different fields of cells were analyzed using this process and the values were combined for statistical analysis.

### 3.2.7 Enzymatic digestion of TPAC-labeled protein for LC-MS/MS analysis

Protein (0.5 mg/mL) was labeled with 2 mM TPAC following the general procedure unless otherwise noted. Labeled protein was precipitated by acetone to remove excess TPAC. The pellet was dissolved in 25 mM HEPES, pH 8.0 containing 6 M urea and subject to reduction (TCEP, 5 mM, 20 min) and alkylation (iodoacetamide, 10 mM, 40 min at 37°C). Protein was precipitated again by acetone, suspended in 25 mM HEPES, pH 8.0 and digested by trypsin (1:20, Promega, Madison WI) overnight at 37°C.

### 3.2.8 LC/MS analysis methods

Reactions on protected amino acid were analyzed by LC/MS using 1220 Infinity LC (Agilent, Santa Clara CA) coupled with Expression-L Compact Mass Spectrometer (Advion, Ithaca NY). Reaction mixtures with single protected amino acid were separated on a Zorbax rapid resolution cartridge (Agilent). Mixtures with multiple protected amino acids were separated on a Zorbax SB-phenyl column (Agilent, 4.6 × 250 mm, 5 μm). Solvent A was water + 0.05% formic acid and solvent B was methanol + 0.05% formic acid. The linear gradient employed for single protected amino acid was 25-100% B in 6.5 min and 100% B for 1.5 min; for multiple amino acids mixture was 45-100% B in 30 min and 100% B for 10 min.

Intact protein samples were analyzed using a Synapt G2-Si mass spectrometer equipped with an ionKey ESI source (Protein ionKey, C4, 1.7  $\mu\text{m}$ , 0.150  $\times$  50 mm, 300  $\text{\AA}$ ), operated in the positive ion mode, and connected in line with an Acquity M-class LC system (Waters, Milford MA). This instrumentation is located in the QB3/Chemistry Mass Spectrometry Facility at the University of California, Berkeley. The obtained mass spectra were deconvoluted using UniDec.<sup>45</sup>

Proteolytically digested protein samples were analyzed using an LTQ-Orbitrap-XL mass spectrometer equipped with an electrospray ionization (ESI) source, operated in the positive ion mode, and connected in line with an UltiMate3000 RSLCnano liquid chromatography (LC) system (Thermo Fisher). The LC system was equipped with a reversed-phase analytical column (Acclaim PepMap100, C18, 3  $\mu\text{m}$ , 0.075  $\times$  250 mm, 100  $\text{\AA}$ , Thermo Fisher). The obtained data were processed with Trans-Proteomic Pipeline using the Comet search algorithm (Institute for Systems Biology, Seattle WA).<sup>46</sup> The MS2 spectra were annotated with a mass tolerance of 15 ppm. The extracted-ion chromatograms (XIC) of MS1 were generated from the .ms1 file from RawConverter, using an in-house script that picks correctly-charged precursor ions with a mass tolerance of 20 ppm.

### 3.2.9 Expression and purification of ribonuclease A (RNaseA)

Synthetic genes for wildtype bovine pancreatic RNaseA excluding the signal peptide, and its mutant, H48Q, H48Q/H105Q, were ordered from Quintarabio (Hayward CA). The genes were cloned into pNIC28-Bsa4 containing N-terminal 6 $\times$ His-TEV cassette by Gibson assembly and confirmed by sequencing.

*E. coli* cell line Rosetta2 (DE3) pLysS harboring RNaseA-expressing plasmid were grown in 1 L LB medium containing kanamycin (50  $\mu\text{g/L}$ ) at 37 $^{\circ}\text{C}$  until OD600 reached 0.8. At this time, protein expression was induced by addition of 1 mM IPTG. After induction, the cells were shaken at 37  $^{\circ}\text{C}$  for 4.5 h and then centrifuged. The cell paste was suspended in 35 mL of 20 mM Tris-HCl, pH 8.0 and subject to sonication on ice. The lysate was centrifuged at 4 $^{\circ}\text{C}$  to discard the soluble protein. The pellet containing inclusion bodies were washed by suspending in 35 mL of 20 mM Tris-HCl pH 8.0 buffer containing 2 M urea, 0.5 M NaCl and 0.5% Triton X-100. The suspension was centrifuged again; the pellet was dissolved by mixing with 35 mL of 20 mM Tris-HCl pH 8.0 buffer containing 6 M guanidine HCl, 5 mM 2-mercaptoethanol, 0.5 M NaCl and 5 mM imidazole for 45 min at room temperature. The resulting solution was passed through a filter and loaded onto a 5 mL HisTrap FF column (GE Healthcare). The column was washed with 50 mL of 20 mM Tris-HCl pH 8.0 buffer containing 6 M guanidine HCl, 1 mM 2-mercaptoethanol, 0.5 M NaCl and 5 mM imidazole, followed by washing with 50 mL of Tris-HCl pH 8.0 buffer containing 6M urea, 0.5 M NaCl, 20 mM imidazole, 2 mM GSH and 0.5 mM GSSG. Protein refolding was performed using a linear gradient wash, starting from the previous washing buffer containing 6M urea to final washing buffer without urea (150 mL, 75

min). The column was further washed with 50 mL of buffer without urea, and the protein was eluted using a linear gradient starting with 20 mM Tris-HCl pH 8.0 buffer containing 5 mM imidazole, 0.5 M NaCl and ending with the same buffer containing 500 mM imidazole. Fractions containing RNaseA were combined, concentrated by ultrafiltration and frozen in aliquots for later processing.

For cleavage of the 6×His-tag linked by TEV-tag on RNaseA, protein solution was buffer-exchanged into 50 mM HEPES pH 8.5 by ultrafiltration. The cleavage was initiated by adding 1.5 mM DTT and 0.05 equivalent of TEV protease and incubated at 25°C for 12 h. The remaining DTT was removed by ultrafiltration and the obtained protein was used for TPAC labeling analysis.

### 3.3 Results and discussion

#### 3.3.1 Design and synthesis of electrophilic phosphorus reagents for histidine labeling

Inspired by observations of reversible histidine phosphorylation as an emerging post-translational modification in prokaryotes and eukaryotes and elegant studies to probe its biological functions,<sup>47–51</sup> we sought to develop a histidine-selective bioconjugation method that mimics this type of chemistry (Figure 3.1). We turned our attention to phosphorus-based electrophiles, based on precedent that potassium phosphoramidate can selectively phosphorylate histidine<sup>52,53</sup> and that thiophosphoryl chloride and potassium thiophosphoramidate can generate thiophosphohistidine analogs with improved aqueous stability.<sup>54–56</sup> To this end, we synthesized a series of phosphorus electrophiles with varying reactivity as suggested by their different synthesis and handling methods (Scheme 3.1). We first tested thiophosphoramidate **1** combining a phosphosulfide moiety for histidine labeling and an alkyne group for further functionalization, but this compound did not show an observable reaction with Fmoc-His-OH in buffered aqueous solution (Scheme 3.2). We then tested compound **2**, where we hypothesized that installation of chloride as a better leaving group might enhance reactivity. Indeed, thiophosphochloride **2** did show appreciable labeling of Fmoc-His-OH, but the resulting product underwent significant hydrolysis back to the unmodified histidine over prolonged incubation at neutral pH (Scheme 3.2), presumably due to the high acidity of the thiophosphoric acid diester as a leaving group.<sup>57</sup> To prevent this observed hydrolysis, we further replaced the methoxy group with another chloride that eventually hydrolyzes to the corresponding hydroxy congener to introduce a negative charge, akin to stable phosphate diester linkers found in nucleotides. The resulting compound **4**, termed thiophosphoro alkyne dichloridate (TPAC), gives higher conversion to product that is resistant to hydrolysis (Scheme 3.2, Table 3.1). The thiophosphorus electrophiles **2** and TPAC are superior histidine labeling reagents compared to their oxygen counterparts **3** and **5**, presumably owing to the greater stabilization provided by the less electronegative sulfur. On the other hand, phosphorus electrophiles with nitrogen (**6**) or sulfur (**7**) as the bridging atom between phosphorus and an alkyne show reduced reactivity towards histidine compared to TPAC.

The thiophosphodichloridate can be functionalized with other clickable handles such as azide **8** and sterically hindered alkyne **9** that show comparable reactivity. Further characterization using TPAC as a representative thiophosphodichloridate reagents shows optimal histidine labeling at more basic pH values, affording a 60% yield of bioconjugate within twenty minutes at pH 8.5 (Figure 3.2a). The short observed reaction time for TPAC-mediated histidine labeling is primarily controlled by the fast hydrolysis of TPAC itself (Figure 3.2b), which is advantageous since the reaction does not require quenching of residual TPAC to halt its reactivity and the reagent can thus be used in excess. Moreover, the thiophospho-histidine product shows reasonable stability under various conditions including high temperature, acidic, basic, reducing and alkylating environments (Table 3.1). The conjugate also shows robust stability in serum at 37°C, whereas in cell lysate at 37°C the modification undergoes efficient hydrolysis (Figure 3.3). Interestingly, the adduct is stable in cell lysate at 4°C, suggesting that cleavage at higher temperatures is possibly due to some catalyzed activity (Figure 3.3). Indeed, such behavior can be potentially useful for traceless delivery of proteins or drugs into cells. Most importantly, TPAC exhibits high selectivity for histidine, showing negligible reactivity on other nucleophilic amino acids (Figure 3.2c, 3.4). Reaction on model peptides also supports its histidine selectivity, where peptide with nucleophilic amino acids residues other than histidine shows insignificant reaction, and conversion yield increases with increasing number of histidine residues on peptide (Figure 3.5). The results collectively identify TPAC as a promising candidate for chemoselective bioconjugation to histidine.

### 3.3.2 TPAC bioconjugation on model proteins.

With these results in hand, we moved on to test TPAC labeling of histidine on intact protein substrates (Figure 3.6). We used ribonuclease A as a model protein and analyzed bioconjugation reactions by mass spectrometry. The TPAC labeling is dose-dependent, generating ca. 45% singly-modified protein and 11% doubly-modified protein at pH 8.5 (Figure 3.6b). Similar to what was observed in small-molecule amino acid models, the TPAC-ribonuclease A coupling reaction is more effective at slightly basic conditions compared to neutral pH (Figure 3.6c). We further performed LC-MS/MS of the digested protein to analyze the site of modification. Similar to other phosphopeptides, the TPAC modification undergoes significant neutral loss in collision-induced dissociation (CID), but the fragment peaks are sufficient for identifying the site of modification (Figure 3.7a, 3.8b-c). We also used electron-transfer dissociation (ETD) on peptides with ambiguous modification sites for supplementary information (Figure 3.8a). Interestingly, the LC-MS/MS of the digested protein indicates that the reaction occurs primarily on H48, a surface-exposed histidine rather than histidines at the catalytic center (H12 and H119). In addition, two other histidine sites, one at the active site (H119) and one that is not (H105), are also modified to lesser extent, as shown qualitatively by extracted precursor ion chromatogram (Figure 3.7a). The data are also in line with the negligible loss of activity of ribonuclease A after TPAC treatment (Figure 3.7b), and site-directed mutagenesis also support that H48 is the major modified residue on this protein (Figure 3.7c).

We then demonstrated that the TPAC bioconjugation method is amenable to labeling histidine residues on other proteins, including calmodulin, myoglobin and lysozyme. The yields vary for these proteins (Figure 3.6d–f), and the reaction shows excellent histidine selectivity on calmodulin and myoglobin with only small amount of side reaction on lysine and threonine (Figure 3.9, 3.10). Modest side-reactivity towards lysine and tyrosine is observed on lysozyme, but lysine labeling can be suppressed by lowering the pH to 7.5 to further block the nucleophilicity of the lysine residues by protonation, while increasing TPAC concentrations to compensate the decreased histidine reactivity (Figure 3.11). TPAC labeling in HeLa lysates also proceeds smoothly as shown by in-gel fluorescence (Figure 3.12a), and more importantly, pretreatments of lysates with competing electrophiles that block cysteine, lysine or serine residues do not strongly affect the observed TPAC signal, suggesting that the reagent does not appreciably react with these competing amino acid residues (Figure 3.12b–e). Taken together, the results show that TPAC is effective at selectively labeling native histidine residues on proteins.

We then sought to apply this histidine bioconjugation method to install clickable payloads onto proteins. Considering that TPAC converts the slightly basic histidine residue into a rather acidic thiophosphoric acid derivative, we reasoned that the labeled protein can be separated readily from the unreacted protein by its charge difference. Indeed, the separation of the reaction mixture of ribonuclease A and TPAC is effective on a strong cation exchange column buffered at pH 4.2, giving rise to three fractions (Figure 3.13a). The mass spectra of these fractions confirm our hypothesis: unreacted ribonuclease A carries more positive charge and elutes last, the doubly TPAC-modified ribonuclease A protein is less positively charged and elutes first, and the singly TPAC-modified protein elutes between these two fractions (Figure 3.13b). After the facile isolation of TPAC-modified reaction products, we further performed copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) between mono-labeled ribonuclease A and either Cy3-N<sub>3</sub> or desthiobiotin-N<sub>3</sub> payloads for detection and enrichment, respectively (Figure 3.13c). As expected, the obtained protein is uniform with only one payload per protein molecule (Figure 3.13b). Collectively, TPAC labeling with subsequent ion exchange chromatography and CuAAC chemistry provides an easy and straightforward workflow to create homogenous proteins modified on histidine residues.

### 3.3.3 TPAC bioconjugation on His-tag for protein delivery.

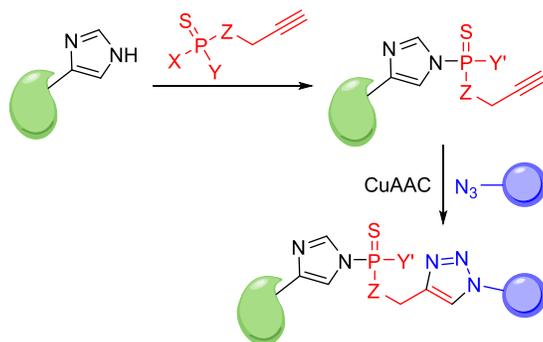
To showcase the potential merits of TPAC bioconjugation chemistry at the cellular level, we turned our attention to modification of polyhistidine-tagged proteins (i.e., His-tag). His-tag is a widely used method for purification of proteins by introducing a short polyhistidine peptide fused to the surface exposed portion of the protein of interest for subsequent resin capture and separation. We envisioned that this polyhistidine motif would greatly enhance the labeling efficiency of TPAC in this region, and we demonstrated this possibility by combining TPAC labeling with further bioconjugation by CuAAC with polyarginine, a cell-penetrating peptide, to create a general method for enabling protein delivery into living cells (Figure 3.14a). We utilized GFP bearing an N-terminal 10×His-tag as a proof-of-concept protein, where TPAC-labeling greatly reduces binding and recognition of His-tag GFP by a His-tag antibody, consistent with efficient labeling of TPAC in this His-tag region (Figure 3.14b). We further performed the

cleavage of the His-tag from the protein by enterokinase treatment to verify the labeling sites. Indeed, TPAC labeling followed by CuAAC with desthiobiotin-N<sub>3</sub> on His-tag GFP produces a significantly, stronger signal compared to GFP without the His-tag on a streptavidin blot (Figure 3.14c), further indicating that on this protein TPAC selectively reacts with histidine residues on the His-tag rather than other 10 histidine residues with 5 sites on the surface. Similarly, LC-MS analysis of the digested protein after TPAC treatment only identifies one to three modifications, all of which appear on the His-tag region (Figure 3.15). With this information in hand we performed CuAAC between TPAC-labeled His-tag GFP and N<sub>3</sub>-(Arg)<sub>9</sub>-OH to introduce cell-penetrating capabilities onto the protein. Indeed, the (Arg)<sub>9</sub>-functionalized GFP is now capable of being delivered into HeLa cells as shown by increases in intracellular green fluorescence as observed by confocal microscopy, whereas His-tagged GFP without TPAC or N<sub>3</sub>-(Arg)<sub>9</sub>-OH treatment results in low fluorescence signal over dark background (Figure 3.14d). Similar results are also observed with an mCherry construct carrying an N-terminus 6×His-tag, although on this protein histidine residues other than the His-tag are also labeled by TPAC to a smaller extent (Figure 3.16). For both cases the internalization of the modified proteins were confirmed by Z-stack images of the cells (Figure 3.17). Taken together, this line of experiments shows that TPAC can provide covalent, histidine-selective bioconjugation for protein functionalization.

### 3.4 Conclusion

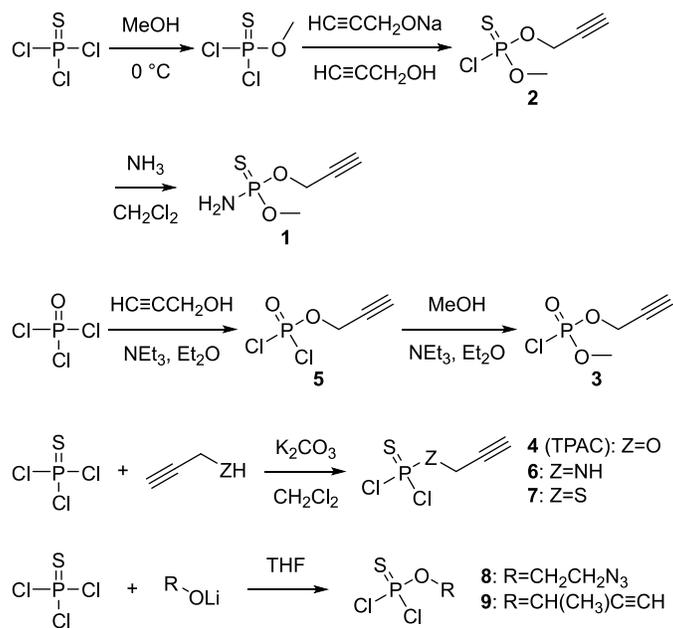
To close, we have reported thiophosphorodichloridates inspired by native histidine phosphorylation processes as reagents for histidine-selective covalent modification of proteins. With TPAC, an alkyne-tagged version of this family of probes, we demonstrated efficient bioconjugation on model proteins with excellent selectivity for histidine over other potentially reactive amino acids. TPAC-labeled small model proteins are readily purified and can undergo further coupling with bioorthogonal click chemistry to introduce a variety of payloads. In one example to illustrate the utility of this approach, TPAC labeling was applied to introduce polyarginine motifs onto His-tag proteins to endow membrane permeability and enable delivery into living cells. Owing to the central importance of histidine residues in enzyme chemistry, this work provides a starting point for probing histidine function in native contexts. Likewise, the efficient labeling of polyhistidine tags can enable versatile protein functionalization for biochemical and protein engineering studies in both fundamental and applied settings. Coupled with recent work from our laboratory on the development of selective methionine bioconjugation reagents that operate by redox-mediated nitrogen group transfer processes inspired by native oxygen atom-transfer oxidations that transform methionine to methionine sulfoxide,<sup>22</sup> mimicking histidine phosphorylation can provide a starting point for developing histidine-selective bioconjugation chemistry. We are currently expanding and applying the toolbox of activity-based sensing reagents in this direction to develop chemoselective probes for other reactive amino acid sites in proteins and proteomes.

### 3.5 Figures and schemes

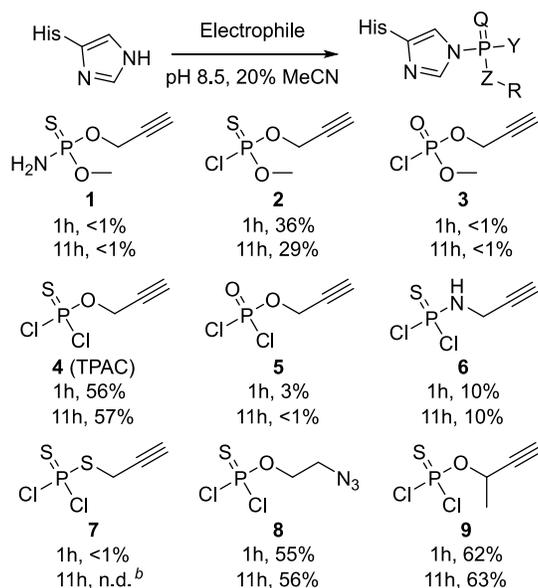


**Figure 3.1** Histidine labeling and the following click reaction for bioconjugation. X, Y=Cl, OH, OMe or NH<sub>2</sub>, Z=O, NH or S.

**Scheme 3.1** Different routes for synthesis of phosphorus electrophiles.

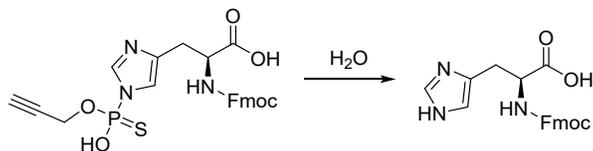


**Scheme 3.2** Reaction yields of phosphorus electrophiles on Fmoc-His-OH. <sup>a</sup>



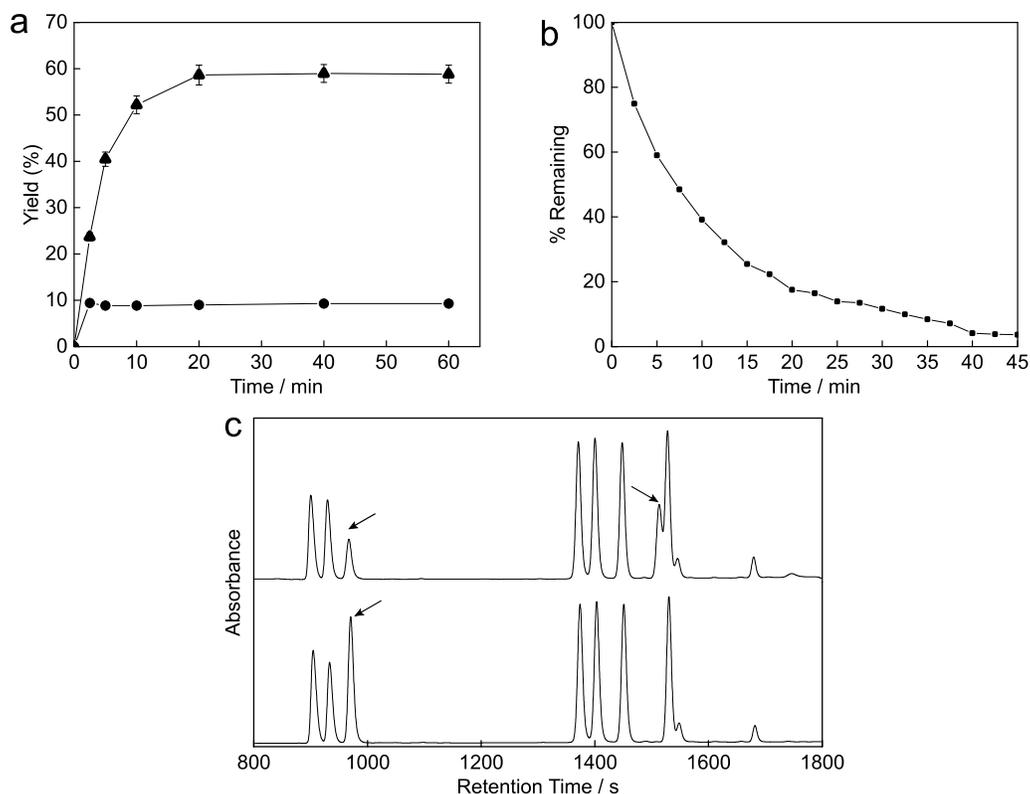
Conditions: 0.5 mM Fmoc-His-OH, 5 mM electrophile, 25 mM HEPES pH 8.5, 20% MeCN, 1h; yields were determined by HPLC. Q=S or O, Y=Cl, OH, OMe or NH<sub>2</sub>, Z=O, NH or S. <sup>b</sup> Reactant readily degenerates and forms precipitate.

**Table 3.1** Stability of TPAC-labeled Fmoc-His-OH Purified. TPAC-labeled Fmoc-His-OH (0.5 mM, 200x in DMSO) was treated under the following conditions in water (1–7) or PBS (8–12) and analyzed by HPLC after indicated period of time.

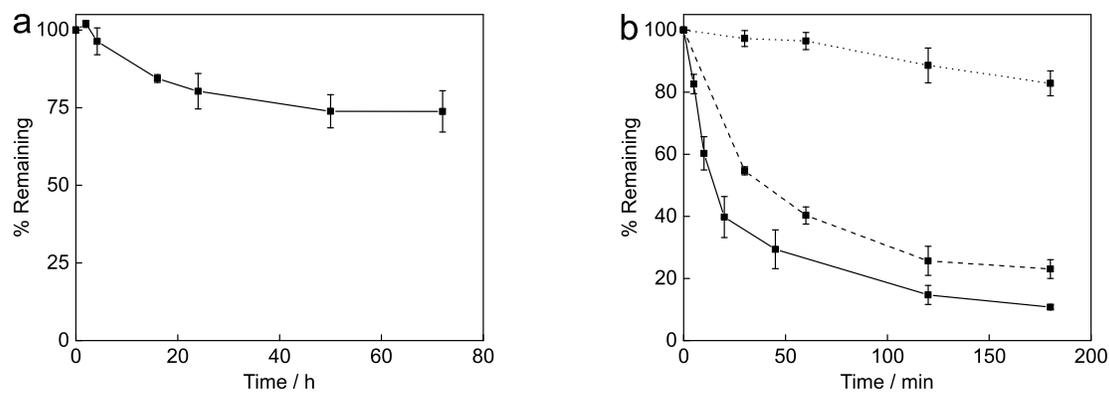


Entry	Additive (concentration)	Temperature/°C	Time/h	Hydrolysis percentage
1	Formic acid (1 M)	25	1	1%
2	Formic acid (1 M)	25	14	15%
3	HCl (0.2 M)	25	1	<1%
4	HCl (0.2 M)	25	14	20%
5	NH <sub>3</sub> ·H <sub>2</sub> O (1 M)	25	1	<1%
6	NH <sub>3</sub> ·H <sub>2</sub> O (1 M)	25	14	<1%
7	NaOH (0.2 M)	25	1	n.d. <sup>a</sup>
8	PBS	80	1	2%
9	PBS	80	14	32%
10	TCEP (5 mM) <sup>b</sup>	65	0.5	<1%
11	Dithiothreitol (5 mM)	65	0.5	1%
12	Iodoacetamide (10 mM)	37	1	2%

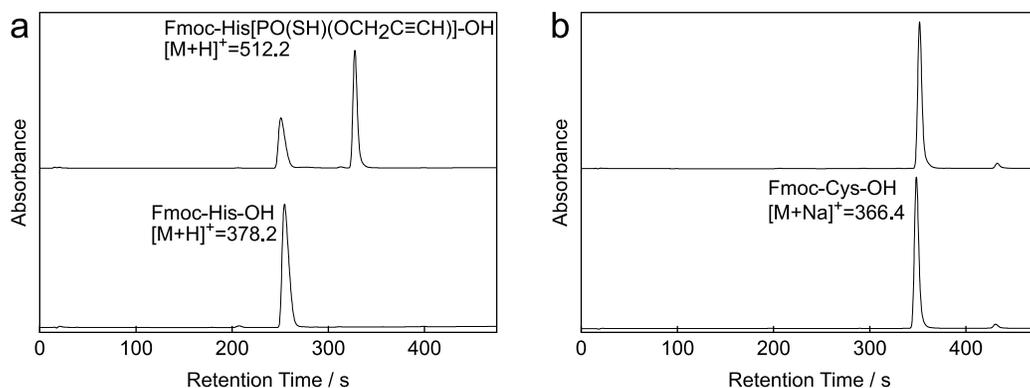
<sup>a</sup> Fmoc protection group was cleaved during treatment. <sup>b</sup> Tris(2-carboxyethyl)phosphine.



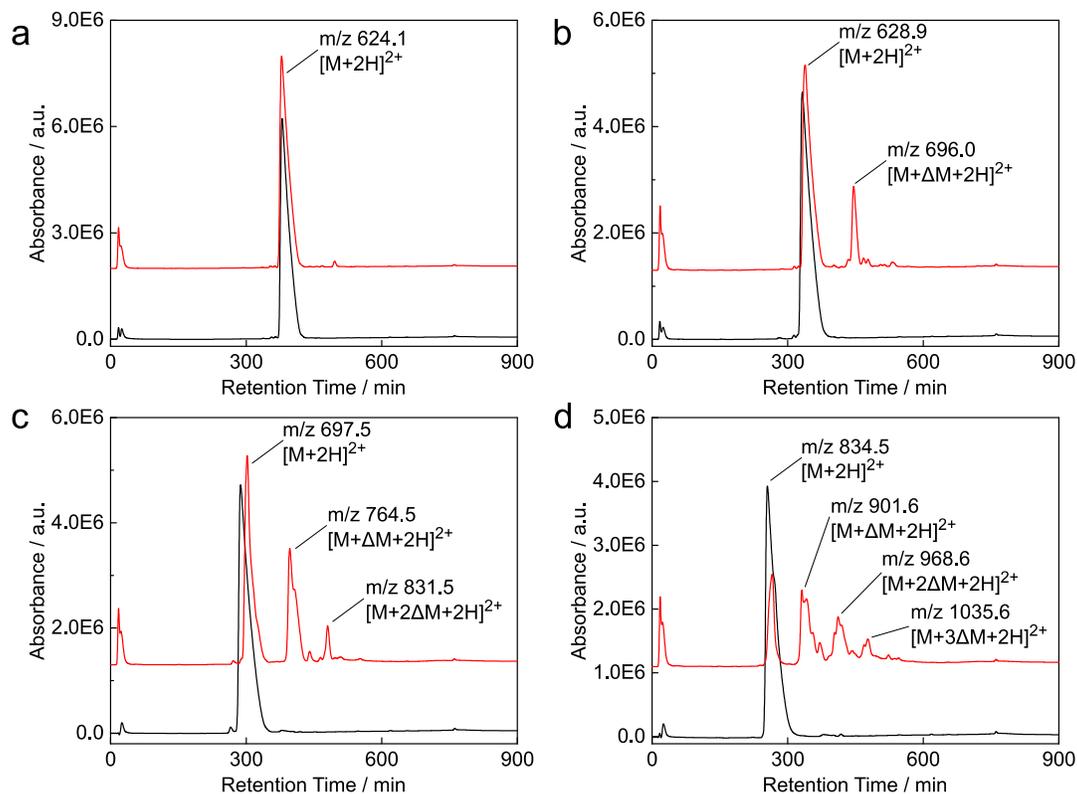
**Figure 3.2** TPAC labeling on small molecules. (a) Kinetics of reaction between TPAC and Fmoc-His-OH. Triangle: pH 8.5. Circle: pH 7.5. (n=3, average  $\pm$  s.d.). (b) Kinetics of 10 mM TPAC hydrolysis in phosphate buffered 1:1 D<sub>2</sub>O/CD<sub>3</sub>CN (100 mM K<sub>3</sub>PO<sub>4</sub>, adjusted to pD 7.0 by pH strip), monitored by <sup>1</sup>H NMR. The observed  $t_{1/2}$  is 9.2 min. (c) HPLC chromatograph showing the reaction between TPAC and nucleophilic Fmoc-amino acid mixtures. Bottom: before TPAC treatment. Top: after TPAC treatment. Left to right: Fmoc-Lys, Fmoc-Arg, Fmoc-His, Fmoc-Ser, Fmoc-Glu, Fmoc-Thr and Fmoc-Tyr. Arrows point to Fmoc-His-OH (left) and TPAC-labeled Fmoc-His-OH (right, ESI-MS m/z expected 512.1, found 512.2 for [M+H]<sup>+</sup>).



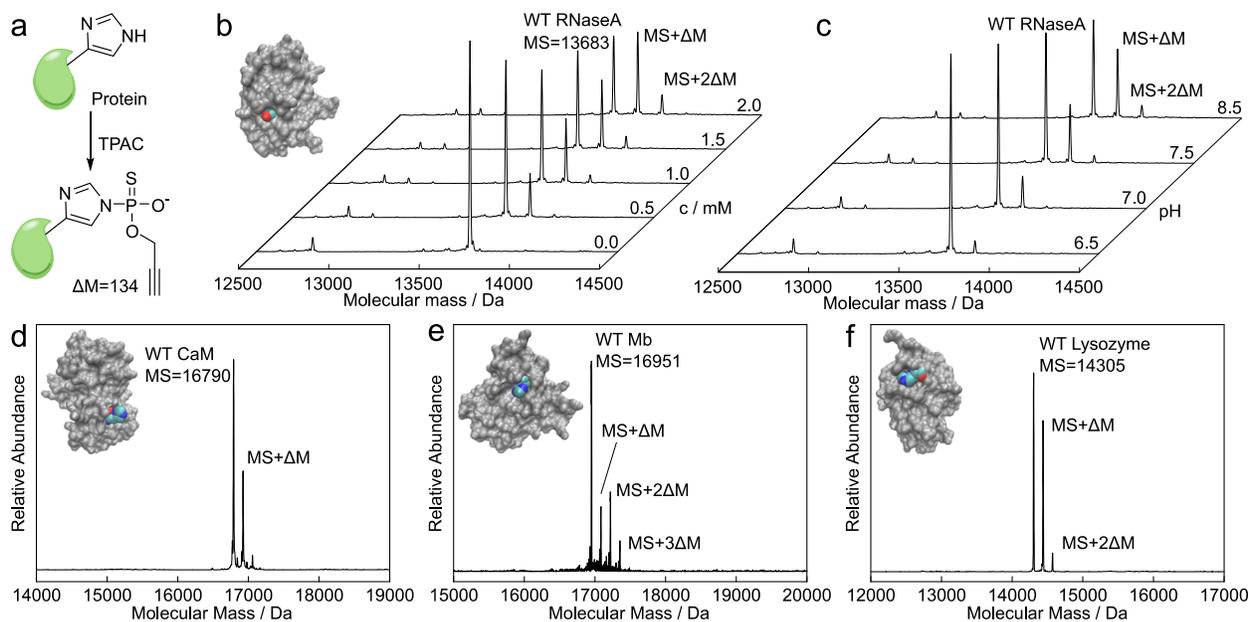
**Figure 3.3** Stability of TPAC-labeled Fmoc-His-OH in (a) fetal bovine serum and (b) HeLa cell lysate (1 mg/mL protein in 50 mM HEPES buffer, pH 7.5). Solid line: 37°C, dashed line: 25°C, dotted line: 4°C. Purified TPAC-labeled Fmoc-His-OH (0.1 mM, 100× in DMSO) was diluted and treated under each condition, protein precipitated by methanol and then analyzed by HPLC (n = 3, average ± s.d.).



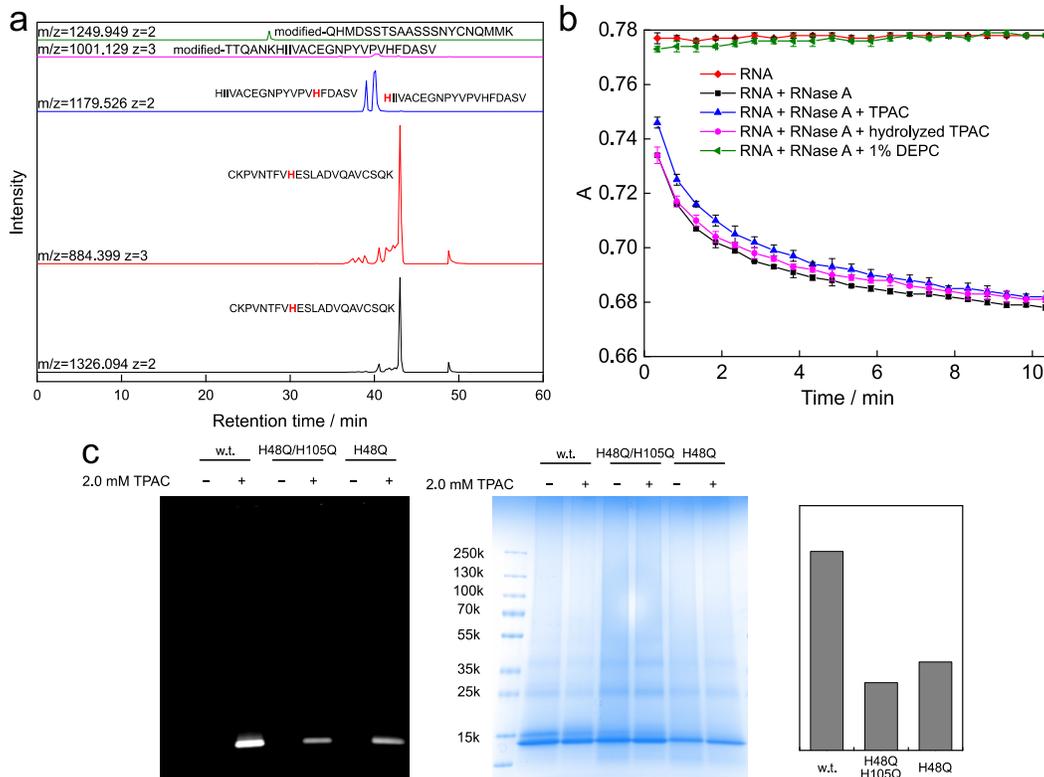
**Figure 3.4** Reactivity of TPAC toward small molecule histidine and cysteine. (a) HPLC chromatograms showing the reactivity of TPAC toward Fmoc-His-OH. Top: reaction mixture of TPAC and Fmoc-His-OH. Bottom: Fmoc-His-OH. (b) HPLC chromatograms showing TPAC is not reactive toward Fmoc-Cys-OH. Top: reaction mixture of TPAC and Fmoc-Cys-OH. Bottom: Fmoc-Cys-OH.



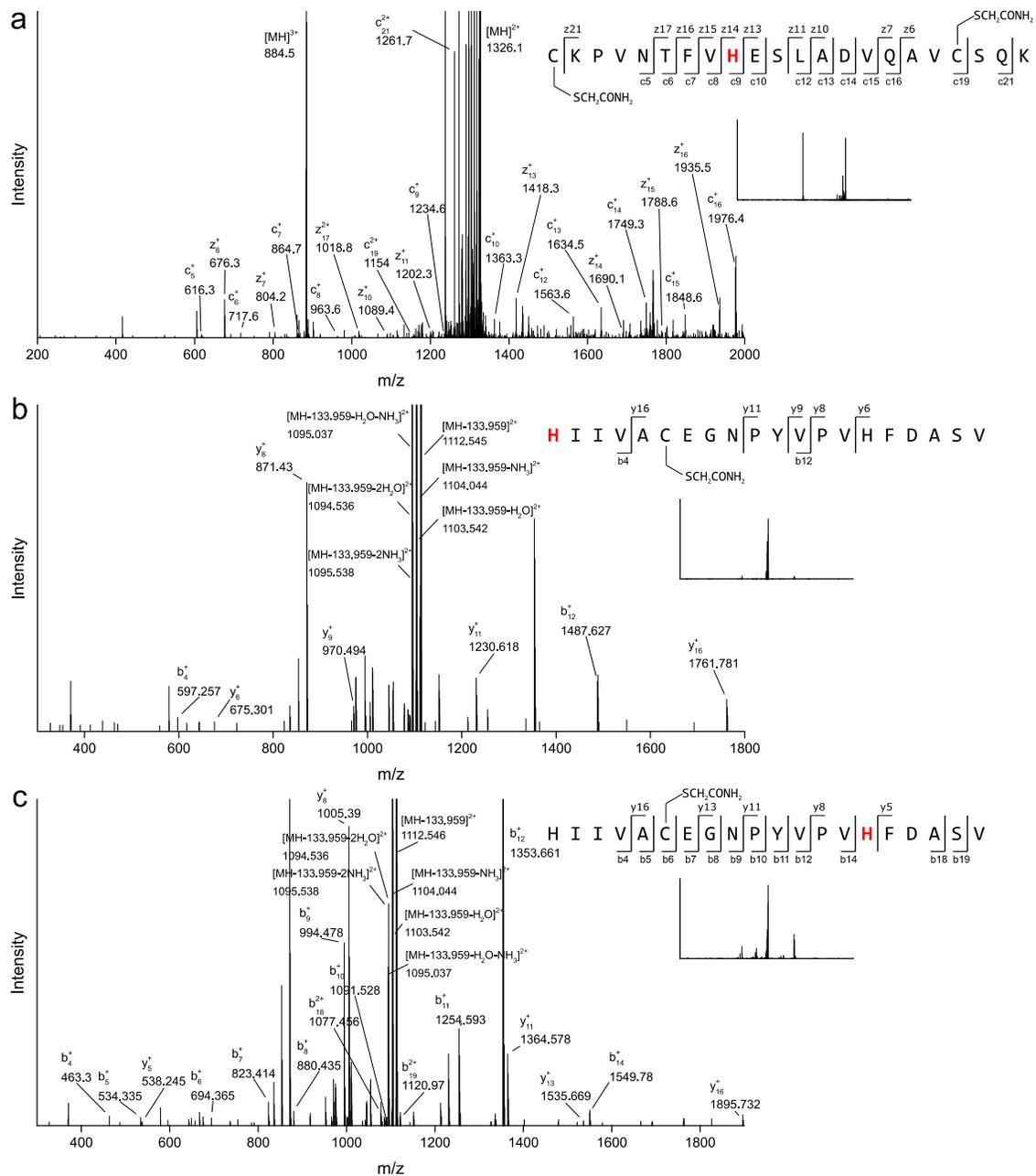
**Figure 3.5** Labeling of model peptides by TPAC. Aqueous solution buffered by 50 mM HEPES at pH 8.5 containing 0.3 mM (a) Fmoc-GRTQEKYSG, (b) Fmoc-GRTHEKYSG, (c) Fmoc-GRTHHEKYSG (d) Fmoc-GRTHHHHEKYSG was treated 5 mM TPAC for 1 h and subjected to LC-MS analysis. Black lines represent starting peptide solution; red lines represent reaction mixture.  $\Delta M=133.96$ , corresponding to the mass increase of one TPAC modification.



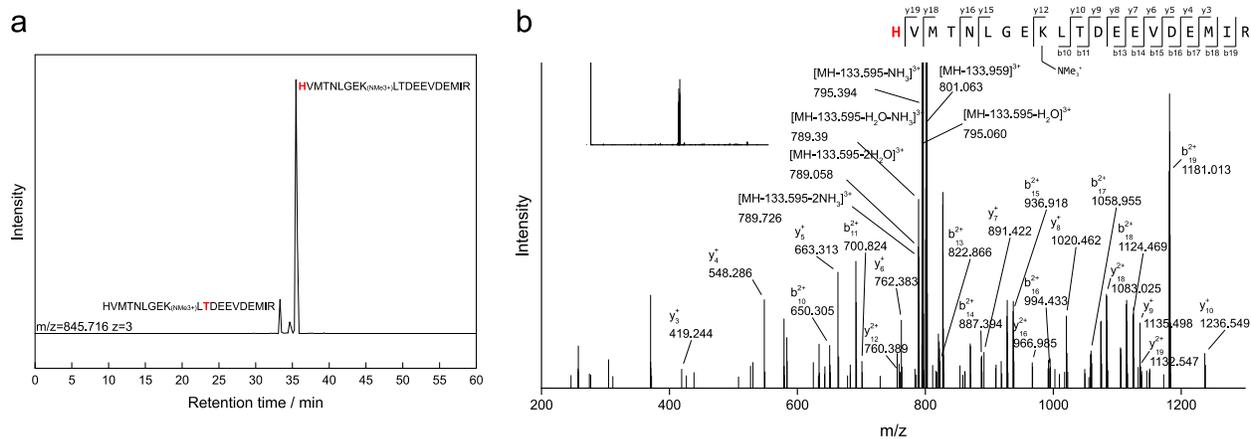
**Figure 3.6** Model protein tagging with TPAC. (a) Structure and mass difference of the histidine modification by TPAC. (b) Deconvoluted mass spectra showing ribonuclease A labeled by TPAC with concentration ranging from 0 to 2 mM. Single and double TPAC modified RNaseA: expected mass 13817, 13951 Da, found 13816, 13950 Da. Its crystal structure (PDB 1bel) is shown in the inset highlighting its major modified histidine. (c) Deconvoluted mass spectra showing ribonuclease A labeled with 1.5 mM of TPAC with different buffered pH. (d) Calmodulin, (e) myoglobin and (f) lysozyme are also labeled by TPAC as shown by their deconvoluted mass spectra. TPAC-modified calmodulin: expected mass 16924 Da, found 16924 Da; single, double and triple TPAC modified myoglobin: expected mass 17085, 17219 and 17353 Da, found 17085, 17218 and 17352 Da; single and double TPAC modified lysozyme: expected mass 14439, 14573 Da, found 14438, 14572 Da. Their crystal structures (PDB 2o60, 1bje and 1931) are shown as insets highlighting the major modified histidines. Conditions: 20  $\mu$ M protein, 2 mM TPAC in 25 mM HEPES with pH 8.5, room temperature, 1 h unless otherwise noted.



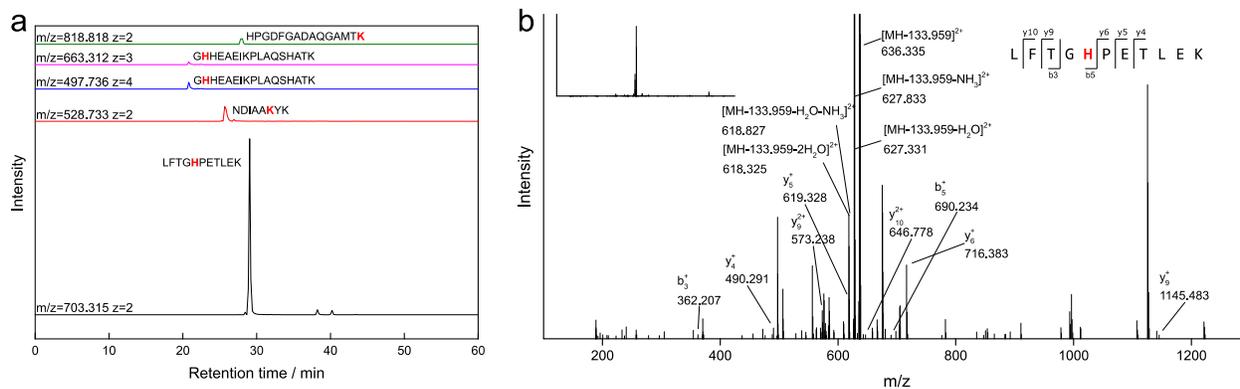
**Figure 3.7** TPAC labeling of ribonuclease A (RNaseA) (a) LC-MS analysis of digested TPAC-labeled RNaseA showing extracted precursor ion chromatograms of peptides carrying TPAC modification ( $\Delta M=133.959$ ). Identified modification sites by MS2 spectra are highlighted in red. (b) Activity assay shows TPAC does not react much with active site histidine residues on RNase A. RNase A (0.5 mg/mL) was treated with 2 mM TPAC, 2 mM hydrolyzed TPAC or 1% diethylpyrocarbonate (DEPC) before incubating with RNA in 25 mM HEPES buffered at pH 8.5. The absorbance of RNA is monitored at 300 nm. (c) Confirmation of TPAC modification site on RNaseA by site-directed mutagenesis. Wildtype RNaseA, RNaseA H48Q/H105Q and RNaseA H48Q were treated with TPAC, conjugated with Cy3-N<sub>3</sub> by CuAAC and subject to SDS-PAGE. Left: fluorescence. Middle: coomassie. Right: quantification of TPAC-labeled gel lanes (fluorescence intensity divided by coomassie intensity).



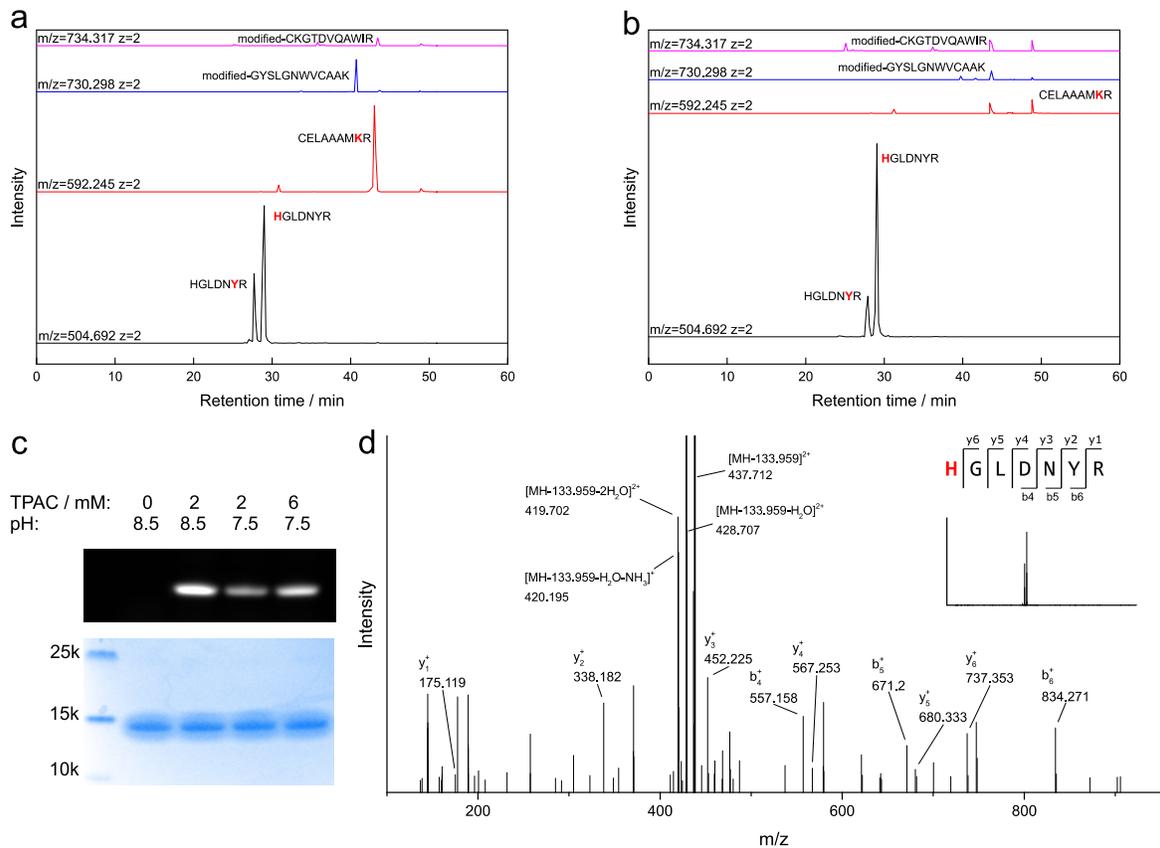
**Figure 3.8** LC-MS/MS analysis of digested TPAC-labeled RNase A. (a) ETD MS/MS spectra and (b),(c) CID MS/MS spectrum of major TPAC-modified peptides from Figure 3.7a. Y-axis is zoomed in to show fragment peaks; full spectrum is shown as inset.



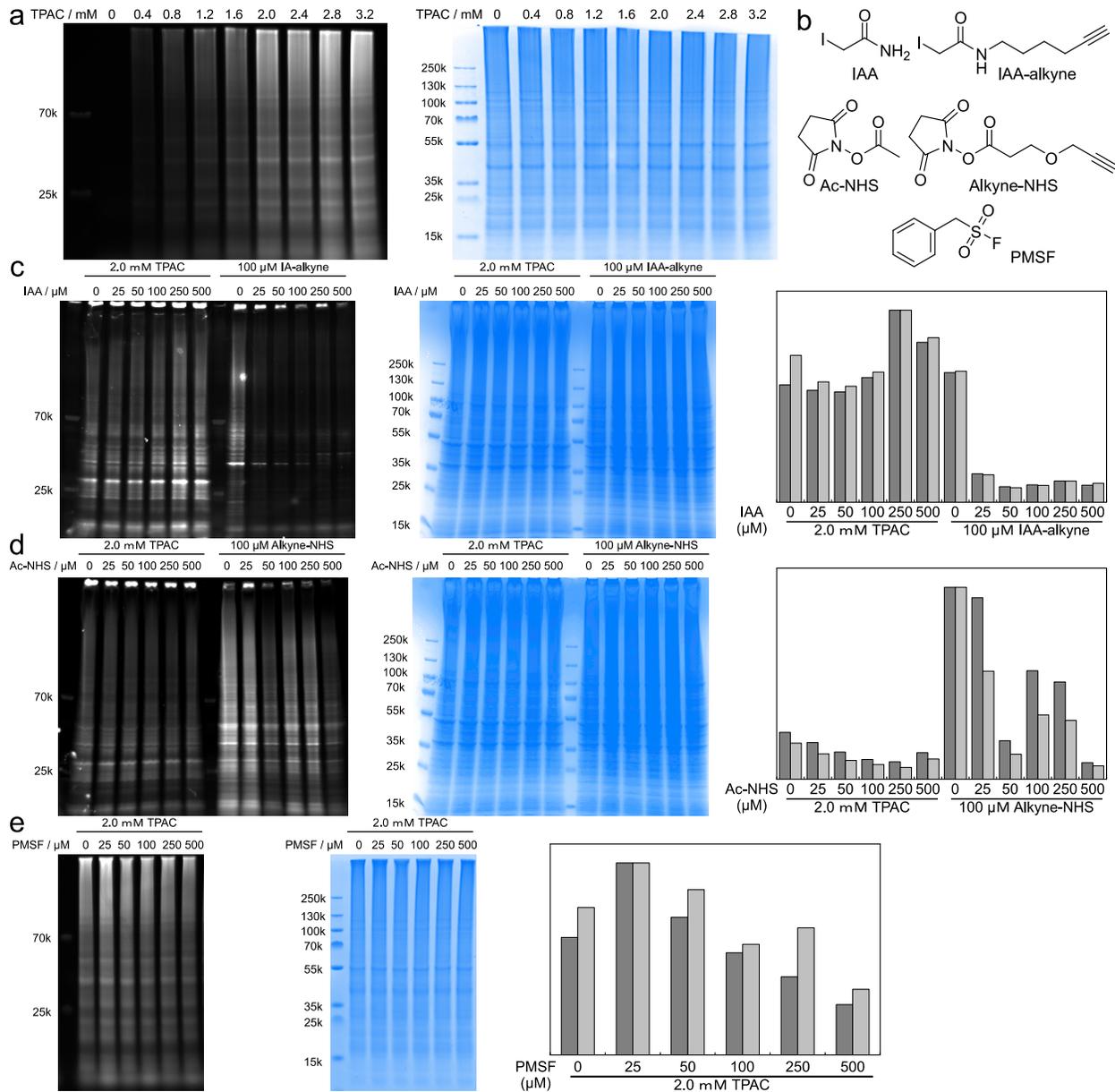
**Figure 3.9** LC-MS/MS analysis of digested TPAC-labeled calmodulin. (a) Extracted precursor ion chromatogram of peptides carrying TPAC modification ( $\Delta M=133.959$ ). Identified modification sites by MS2 spectra are highlighted in red. (b) CID MS/MS spectrum of major TPAC-modified peptide. Y-axis is zoomed in to show fragment peaks; full spectrum is shown as inset.



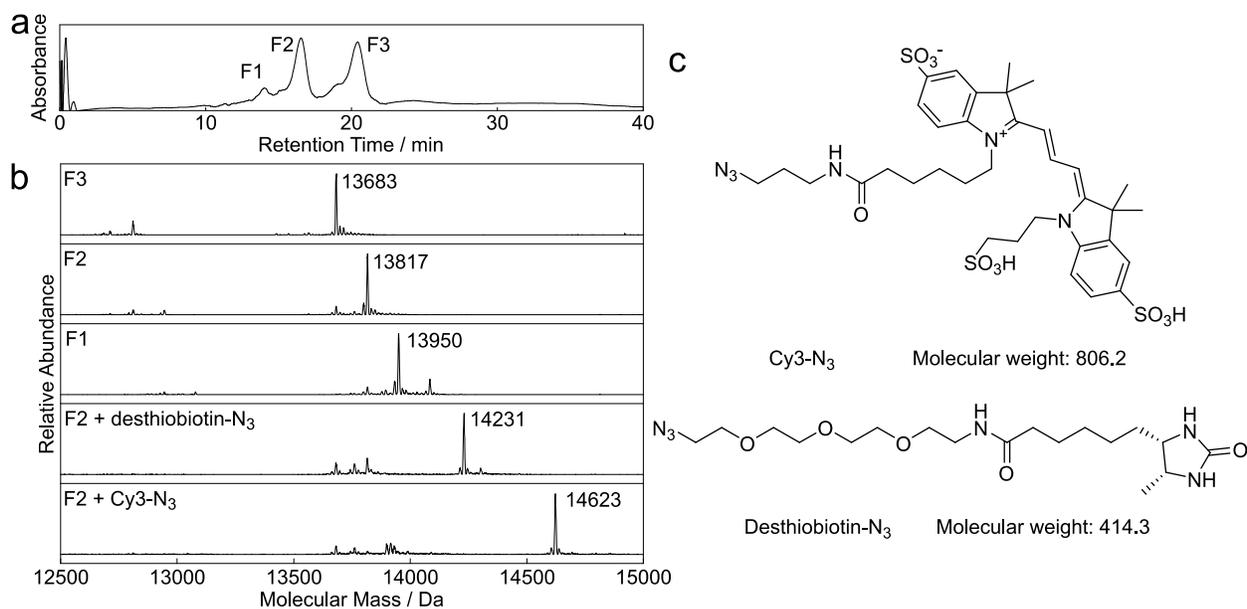
**Figure 3.10** LC-MS/MS analysis of digested TPAC-labeled myoglobin. (a) Extracted precursor ion chromatogram of peptides carrying TPAC modification ( $\Delta M=133.959$ ). Identified modification sites by MS2 spectra are highlighted in red. (b) CID MS/MS spectrum of major TPAC-modified peptide. Y-axis is zoomed in to show fragment peaks; full spectrum is shown as inset.



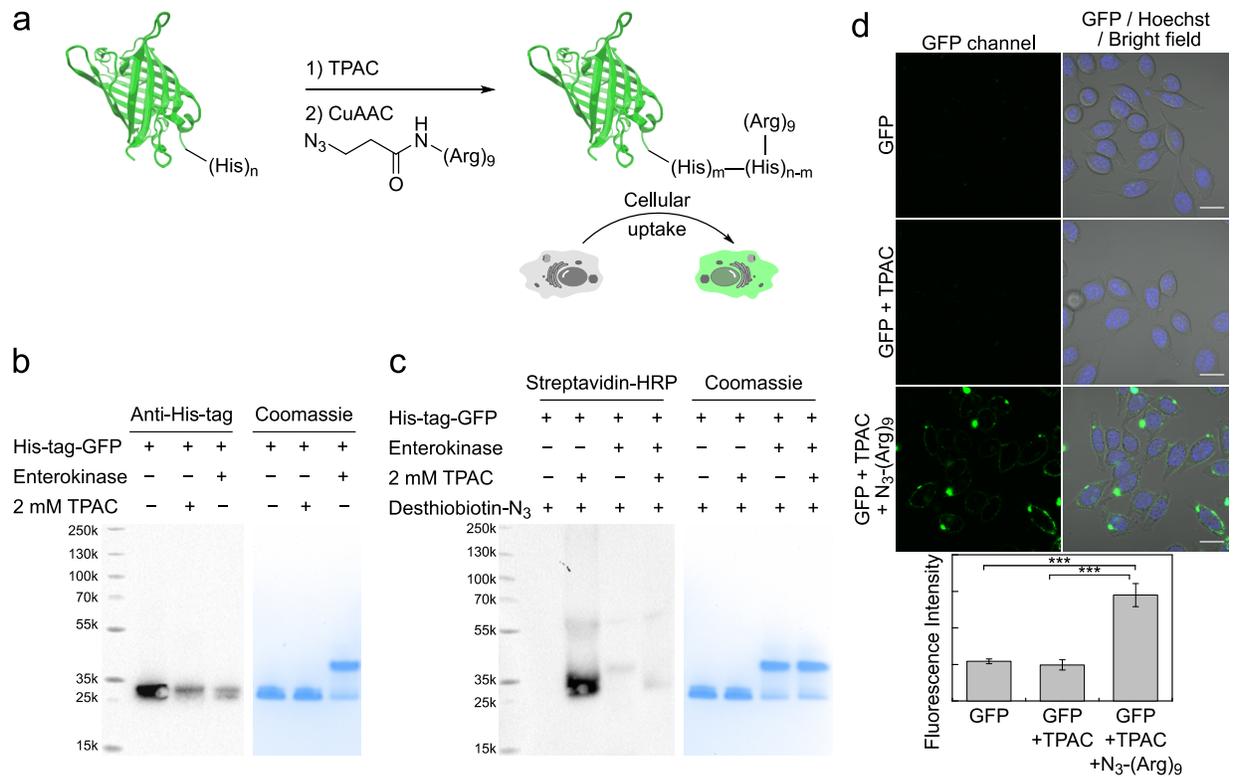
**Figure 3.11** LC-MS/MS analysis of digested TPAC-labeled lysozyme. (a) Extracted precursor ion chromatogram of peptides carrying TPAC modification ( $\Delta M=133.959$ ), digested from lysozyme treated with 2mM TPAC at pH 8.5. (b) Extracted precursor ion chromatogram of peptides carrying TPAC modification ( $\Delta M=133.959$ ), digested from lysozyme treated with 6 mM TPAC at pH 7.5. Modification sites identified by MS2 spectra are highlighted in red. (c) In-gel fluorescence shows comparable labeling efficiency between these two conditions. Lysozyme was labeled with TPAC, conjugated with Cy3-N<sub>3</sub> by CuAAC and subject to SDS-PAGE. Top: fluorescence. Bottom: coomassie. (d) CID MS/MS spectrum of major TPAC-modified peptide. Y-axis is zoomed in to show fragment peaks; full spectrum is shown as inset.



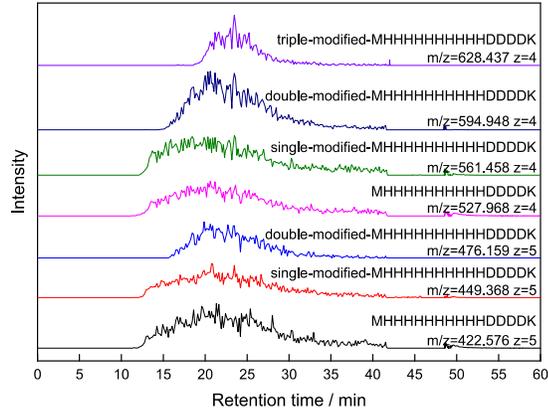
**Figure 3.12** Labeling of HeLa lysate with TPAC in 50 mM HEPES buffer, pH 7.5. (a) Dose-dependent labeling of HeLa proteome with TPAC. Lysates were treated with indicated concentration of TPAC, conjugated with Cy3-N<sub>3</sub> by CuAAC and analyzed by SDS-PAGE. Left: fluorescence. Right: coomassie. (b) Structures of competing electrophiles: IAA that reacts with cysteine, Ac-NHS that reacts with lysine, and PMSF that reacts with reactive serine. The alkyne-containing IAA-alkyne and alkyne-NHS are used to confirm the blockage of amino acid residues from IAA and Ac-NHS treatment, respectively. HeLa lysates were pretreated for 1 h with (c) IAA, (d) Ac-NHS or (e) PMSF prior to labeling with alkyne-containing electrophiles; resulting lysates were conjugated with Cy3-azide by CuAAC and analyzed by SDS-PAGE. Left: fluorescence. Middle: coomassie. Right: quantification of gel lanes (dark grey: fluorescence intensity, light grey: fluorescence intensity divided by coomassie intensity).



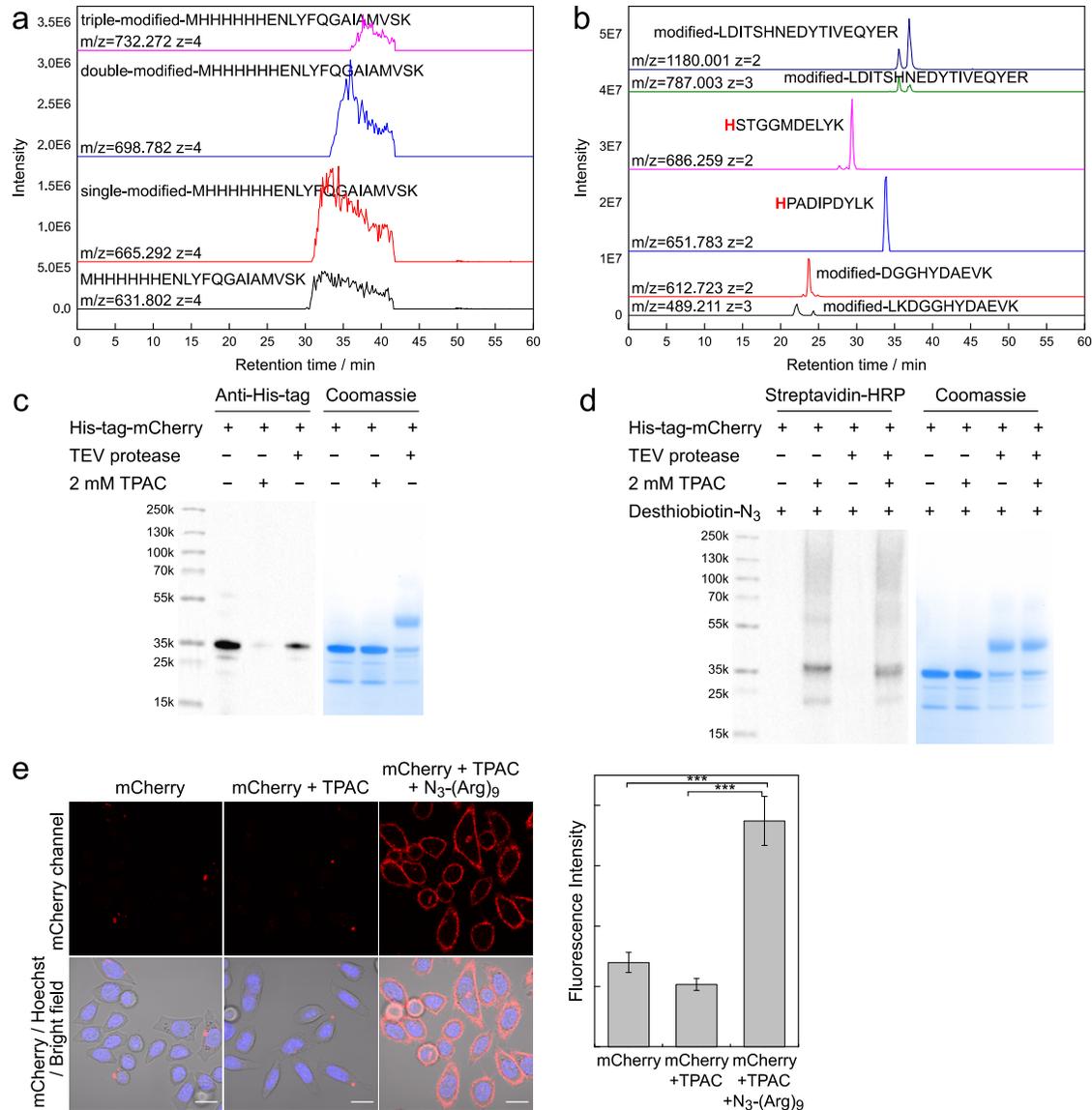
**Figure 3.13** Protein bioconjugation with TPAC. (a) Chromatogram at 280 nm showing the separation of the reaction mixture between ribonuclease A and TPAC by strong cation exchange column. Three fractions, F1-3, were collected for analysis. (b) Deconvoluted mass spectra of ribonuclease A in F1-3, and TPAC-modified ribonuclease A in F2 reacted with the model azide compounds by CuAAC. Found masses are labeled next to the peak. Expected masses for non-, single- and double-labeled ribonuclease A are 13683, 13817, 13951 Da, respectively. Expected masses for single TPAC/desthiobiotin-N<sub>3</sub> and TPAC/Cy3-N<sub>3</sub> labeled ribonuclease A are 14231 and 14623 Da, respectively. (c) Structure and molecular weight of model azide compounds: Cy3-N<sub>3</sub> and desthiobiotin-N<sub>3</sub>.



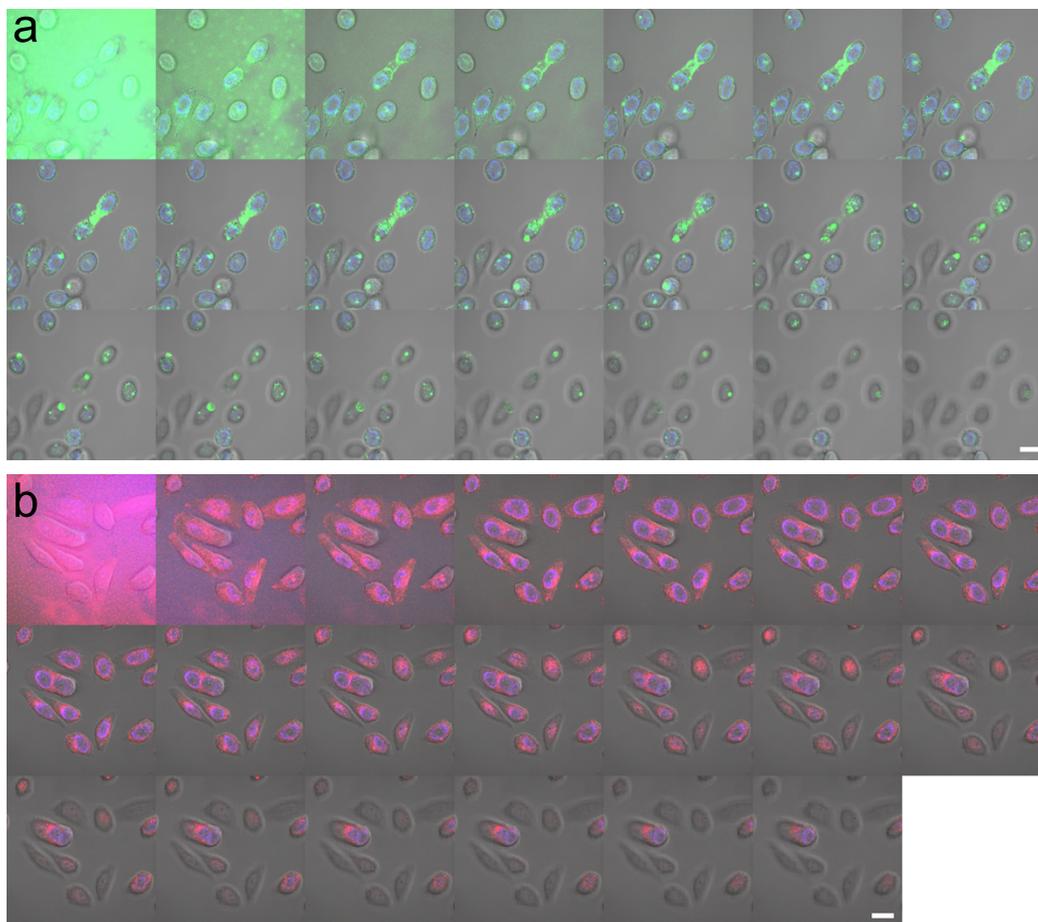
**Figure 3.14** Functionalization of His-tag on GFP to enable protein delivery. (a) Scheme of bioconjugation of polyarginine onto His-tag of a fluorescent protein using TPAC. (b) Streptavidin-HRP blot and coomassie stain showing that removal of the His-tag significantly reduces the labeling by TPAC. His-tagged GFP and native GFP prepared by enterokinase cleavage of His-tag were treated with TPAC, followed by CuAAC with desthiobiotin-N<sub>3</sub> and gel-analysis. Cleavage of His-tag leads to less coomassie staining but with similar migration on SDS-PAGE (Figure 3.18). (c) Western blot and coomassie stain showing that labeling of TPAC with His-tagged GFP significantly reduces its detection by His-tag antibody. (d) Transduction of functionalized GFP into live HeLa cells. Cells were incubated with GFP (0.1  $\mu\text{g}/\mu\text{L}$ ), stained with Hoechst 33342 and imaged; quantification is shown on the bottom ( $n = 3$ , average  $\pm$  s.d. \*\*\* $P \leq 0.001$ ; two-tailed Student's t-test.). Scale-bars: 20  $\mu\text{m}$ .



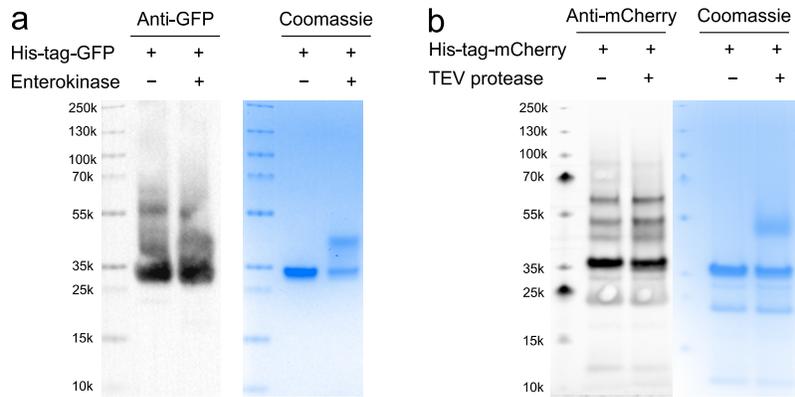
**Figure 3.15** Extracted precursor ion chromatogram of His-tag peptide and major peptides carrying TPAC modification ( $\Delta M=133.959$ ), digested from His-tagged GFP treated with 2mM TPAC at pH 8.5. For simplicity, peptides with N-terminal methionine excised and peptides with oxidized methionine are not shown.



**Figure 3.16** Functionalization of His-tag on mCherry to enable protein delivery. (a) Extracted precursor ion chromatogram of unmodified and TPAC-modified ( $\Delta M=133.959$ ) His-tag peptides, digested from His-tagged GFP treated with 2mM TPAC at pH 8.5. (b) Extracted precursor ion chromatogram of other major TPAC-modified peptides. (c) Western blot and coomassie stain showing that labeling of TPAC with His-tagged mCherry significantly reduces its detection by His-tag antibody. TEV protease was used to cleave off His-tag on this protein. Cleavage of His-tag leads to less coomassie staining but with similar migration on SDS-PAGE (Figure 3.18). (d) Streptavidin-HRP blot and coomassie stain showing that removal of the His-tag reduces the labeling by TPAC. His-tagged mCherry and native mCherry prepared by TEV protease cleavage of His-tag were treated with TPAC, followed by CuAAC with desthiobiotin-N<sub>3</sub> and gel-analysis. (e) Transduction of functionalized mCherry into live HeLa cells. Cells were incubated with mCherry (0.1  $\mu\text{g}/\mu\text{L}$ ), stained with Hoechst 33342 and imaged; quantification is shown on the right ( $n = 3$ , average  $\pm$  s.d.  $***P \leq 0.001$ ; two-tailed Student's t-test.) Scale-bars: 20  $\mu\text{m}$ .



**Figure 3.17** Z-stack imaging of protein delivery into HeLa cells. HeLa cells were treated with TPAC/ N<sub>3</sub>-(Arg)<sub>9</sub> modified (a) His-tagged GFP or (b) His-tagged mCherry and stained with Hoechst 33342. Scale-bars: 20 μm.



**Figure 3.18** Western blot and coomassie stain of GFP and mCherry. (a) His-tagged GFP and GFP. (b) His-tagged mCherry and mCherry.

### 3.6 References

- (1) Stephanopoulos, N.; Francis, M. B. Choosing an Effective Protein Bioconjugation Strategy. *Nat. Chem. Biol.* **2011**, *7* (12), 876–884.
- (2) Spicer, C. D.; Davis, B. G. Selective Chemical Protein Modification. *Nat. Commun.* **2014**, *5*, 4740.
- (3) McKay, C. S.; Finn, M. G. Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. *Chem. Biol.* **2014**, *21* (9), 1075–1101.
- (4) Koniev, O.; Wagner, A. Developments and Recent Advancements in the Field of Endogenous Amino Acid Selective Bond Forming Reactions for Bioconjugation. *Chem. Soc. Rev.* **2015**, *44* (15), 5495–5551.
- (5) Boutureira, O.; Bernardes, G. J. L. Advances in Chemical Protein Modification. *Chem. Rev.* **2015**, *115* (5), 2174–2195.
- (6) Vinogradova, E. V.; Zhang, C.; Spokoyny, A. M.; Pentelute, B. L.; Buchwald, S. L. Organometallic Palladium Reagents for Cysteine Bioconjugation. *Nature* **2015**, *526* (7575), 687–691.
- (7) Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Voorhis, T. V.; Pentelute, B. L.  $\pi$ -Clamp-Mediated Cysteine Conjugation. *Nat. Chem.* **2016**, *8* (2), 120–128.
- (8) Willwacher, J.; Raj, R.; Mohammed, S.; Davis, B. G. Selective Metal-Site-Guided Arylation of Proteins. *J. Am. Chem. Soc.* **2016**, *138* (28), 8678–8681.
- (9) Messina, M. S.; Stauber, J. M.; Waddington, M. A.; Rheingold, A. L.; Maynard, H. D.; Spokoyny, A. M. Organometallic Gold(III) Reagents for Cysteine Arylation. *J. Am. Chem. Soc.* **2018**, *140* (23), 7065–7069.
- (10) Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Guerreiro, A.; Cal, P. M. S. D.; Bertoldo, J.; Maneiro, M.; Perkins, E.; Howard, J.; Deery, M. J.; et al. Chemo- and Regioselective Lysine Modification on Native Proteins. *J. Am. Chem. Soc.* **2018**, *140* (11), 4004–4017.
- (11) Ward, C. C.; Kleinman, J. I.; Nomura, D. K. NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem. Biol.* **2017**, *12* (6), 1478–1483.
- (12) Hacker, S. M.; Backus, K. M.; Lazear, M. R.; Forli, S.; Correia, B. E.; Cravatt, B. F. Global Profiling of Lysine Reactivity and Ligandability in the Human Proteome. *Nat. Chem.* **2017**, *9* (12), 1181–1190.
- (13) Liu, Y.; Patricelli, M. P.; Cravatt, B. F. Activity-Based Protein Profiling: The Serine Hydrolases. *Proc. Natl. Acad. Sci.* **1999**, *96* (26), 14694–14699.
- (14) Kidd, D.; Liu, Y.; Cravatt, B. F. Profiling Serine Hydrolase Activities in Complex Proteomes. *Biochemistry* **2001**, *40* (13), 4005–4015.
- (15) McGrath, N. A.; Andersen, K. A.; Davis, A. K. F.; Lomax, J. E.; Raines, R. T. Diazo Compounds for the Bioreversible Esterification of Proteins. *Chem. Sci.* **2014**, *6* (1), 752–755.
- (16) deGruyter, J. N.; Malins, L. R.; Baran, P. S. Residue-Specific Peptide Modification: A Chemist’s Guide. *Biochemistry* **2017**, *56* (30), 3863–3873.

- (17) Schlick, T. L.; Ding, Z.; Kovacs, E. W.; Francis, M. B. Dual-Surface Modification of the Tobacco Mosaic Virus. *J. Am. Chem. Soc.* **2005**, *127* (11), 3718–3723.
- (18) Romanini, D. W.; Francis, M. B. Attachment of Peptide Building Blocks to Proteins Through Tyrosine Bioconjugation. *Bioconjug. Chem.* **2008**, *19* (1), 153–157.
- (19) Ban, H.; Gavriluk, J.; Barbas, C. F. Tyrosine Bioconjugation through Aqueous Ene-Type Reactions: A Click-Like Reaction for Tyrosine. *J. Am. Chem. Soc.* **2010**, *132* (5), 1523–1525.
- (20) Antos, J. M.; Francis, M. B. Selective Tryptophan Modification with Rhodium Carbenoids in Aqueous Solution. *J. Am. Chem. Soc.* **2004**, *126* (33), 10256–10257.
- (21) Antos, J. M.; McFarland, J. M.; Iavarone, A. T.; Francis, M. B. Chemoselective Tryptophan Labeling with Rhodium Carbenoids at Mild pH. *J. Am. Chem. Soc.* **2009**, *131* (17), 6301–6308.
- (22) Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Toste, F. D.; et al. Redox-Based Reagents for Chemoselective Methionine Bioconjugation. *Science* **2017**, *355* (6325), 597–602.
- (23) Taylor, M. T.; Nelson, J. E.; Suero, M. G.; Gaunt, M. J. A Protein Functionalization Platform Based on Selective Reactions at Methionine Residues. *Nature* **2018**, *562* (7728), 563–568.
- (24) Gutteridge, A.; Thornton, J. M. Understanding Nature’s Catalytic Toolkit. *Trends Biochem. Sci.* **2005**, *30* (11), 622–629.
- (25) Dokmanić, I.; Šikić, M.; Tomić, S. Metals in Proteins: Correlation between the Metal-Ion Type, Coordination Number and the Amino-Acid Residues Involved in the Coordination. *Acta Crystallogr. D Biol. Crystallogr.* **2008**, *64* (3), 257–263.
- (26) Ohata, J.; Minus, M. B.; Abernathy, M. E.; Ball, Z. T. Histidine-Directed Arylation/Alkenylation of Backbone N–H Bonds Mediated by Copper(II). *J. Am. Chem. Soc.* **2016**, *138* (24), 7472–7475.
- (27) Ohata, J.; Zeng, Y.; Segatori, L.; Ball, Z. T. A Naturally Encoded Dipeptide Handle for Bioorthogonal Chan–Lam Coupling. *Angew. Chem. Int. Ed.* **2018**, *57* (15), 4015–4019.
- (28) Waibel, R.; Alberto, R.; Willuda, J.; Finnern, R.; Schibli, R.; Stichelberger, A.; Egli, A.; Abram, U.; Mach, J.-P.; Plückthun, A.; et al. Stable One-Step Technetium-99m Labeling of His-Tagged Recombinant Proteins with a Novel Tc(I)–Carbonyl Complex. *Nat. Biotechnol.* **1999**, *17* (9), 897–901.
- (29) Kapanidis, A. N.; Ebright, Y. W.; Ebright, R. H. Site-Specific Incorporation of Fluorescent Probes into Protein: Hexahistidine-Tag-Mediated Fluorescent Labeling with (Ni<sup>2+</sup>:Nitrilotriacetic Acid)N–Fluorochrome Conjugates. *J. Am. Chem. Soc.* **2001**, *123* (48), 12123–12125.
- (30) Lata, S.; Gavutis, M.; Tampé, R.; Piehler, J. Specific and Stable Fluorescence Labeling of Histidine-Tagged Proteins for Dissecting Multi-Protein Complex Formation. *J. Am. Chem. Soc.* **2006**, *128* (7), 2365–2372.

- (31) Wang, X.; Jia, J.; Huang, Z.; Zhou, M.; Fei, H. Luminescent Peptide Labeling Based on a Histidine-Binding Iridium(III) Complex for Cell Penetration and Intracellular Targeting Studies. *Chem. – Eur. J.* **2011**, *17* (29), 8028–8032.
- (32) Lai, Y.-T.; Chang, Y.-Y.; Hu, L.; Yang, Y.; Chao, A.; Du, Z.-Y.; Tanner, J. A.; Chye, M.-L.; Qian, C.; Ng, K.-M.; et al. Rapid Labeling of Intracellular His-Tagged Proteins in Living Cells. *Proc. Natl. Acad. Sci.* **2015**, *112* (10), 2948–2953.
- (33) Li, X.; Ma, H.; Dong, S.; Duan, X.; Liang, S. Selective Labeling of Histidine by a Designed Fluorescein-Based Probe. *Talanta* **2004**, *62* (2), 367–371.
- (34) Li, X.; Ma, H.; Nie, L.; Sun, M.; Xiong, S. A Novel Fluorescent Probe for Selective Labeling of Histidine. *Anal. Chim. Acta* **2004**, *515* (2), 255–260.
- (35) Cong, Y.; Pawlisz, E.; Bryant, P.; Balan, S.; Laurine, E.; Tommasi, R.; Singh, R.; Dubey, S.; Peciak, K.; Bird, M.; et al. Site-Specific PEGylation at Histidine Tags. *Bioconjug. Chem.* **2012**, *23* (2), 248–263.
- (36) Peciak, K.; Laurine, E.; Tommasi, R.; Choi, J.; Brocchini, S. Site-Selective Protein Conjugation at Histidine. *Chem. Sci.* **2019**, *10* (2), 427–439.
- (37) Takaoka, Y.; Tsutsumi, H.; Kasagi, N.; Nakata, E.; Hamachi, I. One-Pot and Sequential Organic Chemistry on an Enzyme Surface to Tether a Fluorescent Probe at the Proximity of the Active Site with Restoring Enzyme Activity. *J. Am. Chem. Soc.* **2006**, *128* (10), 3273–3280.
- (38) Wakabayashi, H.; Miyagawa, M.; Koshi, Y.; Takaoka, Y.; Tsukiji, S.; Hamachi, I. Affinity-Labeling-Based Introduction of a Reactive Handle for Natural Protein Modification. *Chem. – Asian J.* **3** (7), 1134–1139.
- (39) Olszewska, A.; Pohl, R.; Brázdová, M.; Fojta, M.; Hocek, M. Chloroacetamide-Linked Nucleotides and DNA for Cross-Linking with Peptides and Proteins. *Bioconjug. Chem.* **2016**, *27* (9), 2089–2094.
- (40) Adusumalli, S. R.; Rawale, D. G.; Singh, U.; Tripathi, P.; Paul, R.; Kalra, N.; Mishra, R. K.; Shukla, S.; Rai, V. Single-Site Labeling of Native Proteins Enabled by a Chemoselective and Site-Selective Chemical Technology. *J. Am. Chem. Soc.* **2018**, *140* (44), 15114–15123.
- (41) Hounsou, C.; Margathe, J.-F.; Oueslati, N.; Belhocine, A.; Dupuis, E.; Thomas, C.; Mann, A.; Ilien, B.; Rognan, D.; Trinquet, E.; et al. Time-Resolved FRET Binding Assay to Investigate Hetero-Oligomer Binding Properties: Proof of Concept with Dopamine D1/D3 Heterodimer. *ACS Chem. Biol.* **2015**, *10* (2), 466–474.
- (42) Grandjean, C.; Boutonnier, A.; Guerreiro, C.; Fournier, J.-M.; Mulard, L. A. On the Preparation of Carbohydrate–Protein Conjugates Using the Traceless Staudinger Ligation. *J. Org. Chem.* **2005**, *70* (18), 7123–7132.
- (43) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. Analysis and Optimization of Copper-Catalyzed Azide–Alkyne Cycloaddition for Bioconjugation. *Angew. Chem. Int. Ed.* **48** (52), 9879–9883.
- (44) Čepa, M. Segmentation of Total Cell Area in Brightfield Microscopy Images. *Methods Protoc.* **2018**, *1* (4), 43.

- (45) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **2015**, *87* (8), 4370–4376.
- (46) Eng, J. K.; Jahan, T. A.; Hoopmann, M. R. Comet: An Open-Source MS/MS Sequence Database Search Tool. *PROTEOMICS* **13** (1), 22–24.
- (47) Puttick, J.; Baker, E. N.; Delbaere, L. T. J. Histidine Phosphorylation in Biological Systems. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2008**, *1784* (1), 100–105.
- (48) Besant, P. G.; Attwood, P. V. Detection and Analysis of Protein Histidine Phosphorylation. *Mol. Cell. Biochem.* **2009**, *329* (1–2), 93–106.
- (49) Kee, J.-M.; Muir, T. W. Chasing Phosphohistidine, an Elusive Sibling in the Phosphoamino Acid Family. *ACS Chem. Biol.* **2012**, *7* (1), 44–51.
- (50) Wilke, K. E.; Francis, S.; Carlson, E. E. Activity-Based Probe for Histidine Kinase Signaling. *J. Am. Chem. Soc.* **2012**, *134* (22), 9150–9153.
- (51) Potel, C. M.; Lin, M.-H.; Heck, A. J. R.; Lemeer, S. Widespread Bacterial Protein Histidine Phosphorylation Revealed by Mass Spectrometry-Based Proteomics. *Nat. Methods* **2018**, *15* (3), 187–190.
- (52) Medzihradzky, K. F.; Phillipps, N. J.; Senderowicz, L.; Wang, P.; Turck, C. W. Synthesis and Characterization of Histidine-Phosphorylated Peptides. *Protein Sci.* **6** (7), 1405–1411.
- (53) Hohenester, U. M.; Ludwig, K.; König, S. Chemical Phosphorylation of Histidine Residues in Proteins Using Potassium Phosphoramidate -- a Tool for the Analysis of Acid-Labile Phosphorylation. *Curr. Drug Deliv.* **2013**, *10* (1), 58–63.
- (54) Lasker, M.; Bui, C. D.; Besant, P. G.; Sugawara, K.; Thai, P.; Medzihradzky, G.; Turck, C. W. Protein Histidine Phosphorylation: Increased Stability of Thiophosphohistidine. *Protein Sci.* **8** (10), 2177–2185.
- (55) Pirrung, M. C.; James, K. D.; Rana, V. S. Thiophosphorylation of Histidine. *J. Org. Chem.* **2000**, *65* (25), 8448–8453.
- (56) Ruman, T.; Długopolska, K.; Jurkiewicz, A.; Rut, D.; Frączyk, T.; Cieśla, J.; Leś, A.; Szewczuk, Z.; Rode, W. Thiophosphorylation of Free Amino Acids and Enzyme Protein by Thiophosphoramidate Ions. *Bioorganic Chem.* **2010**, *38* (2), 74–80.
- (57) Westheimer, F. H. Why Nature Chose Phosphates. *Science* **1987**, *235* (4793), 1173–1178.

## Chapter 4

### **Optimization of thiophosphorodichloridate labeling strategy for histidine profiling**

Portions of this work were performed in collaboration with the following persons:

Synthesis of compounds **1 – 3**, their adducts with Fmoc-His-OH and characterization of CuAAC reactions were performed by Rei-Chi Chen.

Compounds **12 – 15** and **17** were synthesized by Dr. Lun An.

## Abstract

Histidine residue plays an essential role in various enzymes and metalloproteins. While it is challenging to predict active site histidine or metal-binding site histidine based on genome database, profiling histidine on proteins using proteomics methodology provides a potential way to study this residue in a high-throughput fashion. This chapter focuses on the optimization of the thiophosphorodichloridate reagent discussed in the previous chapter for its application in reactivity-based protein profiling of histidine residue. By adopting longer alkyne chains compared to propargyl group, we first circumvented metal-mediated hydrolysis of this modification and therefore boosted its stability in both lysate and in CuAAC reaction condition. For more informative tandem mass spectra using electron-based fragmentation techniques, we designed and synthesized three positively-charged azides to increase the charge state of the peptides. The thiophosphorodichloridate/charged azide modification successfully increased the charged state of the model peptide from +2 to +4, with much more fragmentation peaks after ETD fragmentation. Lastly, the higher charge state of modified peptides allows for facile enrichment using ion exchange chromatography. We have thus shown that the 2<sup>nd</sup> generation of thiophosphorodichloridates along with the charged azide provides a useful tool set to study histidine by proteomics methodology.

## 4.1 Introduction

Histidine is one of the most important amino acid residues. Its unique imidazole side chain can serve as a proton shuttle in enzymes, or metal ligand in metalloproteins. Although being overall a rare amino acid among the 20 amino acid residues in proteome (2.27% as the third rarest amino acid),<sup>1</sup> it is the most frequent amino acid residue that appears at enzyme catalytic centers<sup>2</sup> and also one of the most frequent amino acid residues at metal binding sites, especially for transitional metals.<sup>3</sup> Besides, naturally-occurring histidine phosphorylation and methylation processes suggest this amino acid also has a role of regulatory function of protein activity.<sup>4-7</sup> The oxidation of histidine residue also serve as a sensor for oxidative stress.<sup>8-10</sup> Despite of diverse roles of histidine in biology, studying such functions with conventional biochemistry approaches requires lengthy experiments and analysis. Breakthroughs in mass spectrometry-based techniques for histidine profiling would greatly push forward the study of histidine in biology.

We therefore sought to apply isotopic Tandem Orthogonal Proteolysis Activity-Based Protein Profiling (isoTOP-ABPP)<sup>11</sup> for histidine profiling. Briefly, an isoTOP-ABPP strategy makes use of residue-specific reactive probes that contain either an isotopically-heavy or an isotopically-light tag. With these two mass tags, relative residue-labeling efficiency under experimental group and control group, as readout of the degree of the modification on this residue, can be quantified by their signal strength on LC-MS/MS analysis. In order to achieve proteome-wide histidine profiling using this strategy, a reactive probe that is fast and selective towards histidine is required. Existing bioconjugation tools for histidine mostly rely on harsh conditions including high temperatures or strong bases,<sup>12,13</sup> or rely on sequence-dependent motifs

<sup>14-16</sup> or affinity-directed ligands.<sup>17-20</sup> Long incubation time, harsh conditions or site-specificity of these methods prevents their use for effective profiling of histidine residues on proteins in their native state in a high-throughput manner.

In Chapter 3 we have established thiophosphorodichloridate as a fast, efficient and histidine-selective labeling reagent. Although the copper-catalyzed azide-alkyne cycloaddition (CuAAC) proceeds as expected on labeled RNaseA as a model protein, we observed this modification is rather labile during CuAAC reaction on Fmoc-His-OH as a small molecule model, which precludes its successful application in protein profiling. The strong neutral loss character of this modification under conventional collision-based mass spectrometers also limits the sequence information provided by such modification. This chapter details our efforts to overcome these challenges by optimizing the reagents and the protocol to achieve profiling of histidine residue using thiophosphorodichloridate warheads. The 2<sup>nd</sup>-generation thiophosphorodichloridate, CuAAC azide pair and optimized procedure provides a feasible way towards histidine profiling in a proteome context.

## 4.2 Methods

### 4.2.1 General synthetic methods

Unless otherwise noted, all commercial reagents were used without further purification. All reactions utilizing air- or moisture-sensitive reagents were performed in dried glassware under an atmosphere of dry N<sub>2</sub>. 1-Pentyne-5-ol was purchased from Oakwood Chemicals (Estill, SC). Fmoc-His-OH was purchased from Ark Pharm (Arlington Heights IL). Bombesin was purchased from Alfa Aesar (Tewksbury, MA). 3-Azido-1-propanol was synthesized according to published procedure.<sup>21</sup> <sup>13</sup>C-labeled compounds were purchased from Cambridge Isotope Laboratories (Cambridge MA). All other reagents were purchased from Sigma-Aldrich. <sup>1</sup>H NMR, <sup>13</sup>C NMR and <sup>31</sup>P NMR spectra were collected in CDCl<sub>3</sub>, MeOD or *d*<sub>6</sub>-acetone (Cambridge Isotope Laboratories, Cambridge MA) at 25 °C on AVB-400 or AVQ-400 spectrometers at the College of Chemistry NMR Facility at the University of California, Berkeley. All chemical shifts in <sup>1</sup>H NMR and <sup>13</sup>C NMR are reported in the standard  $\delta$  notation of ppm relative to residual solvent peak (CDCl<sub>3</sub>  $\delta$ H = 7.26,  $\delta$ C = 77.16; MeOD  $\delta$ H = 3.31,  $\delta$ C = 49.00; *d*<sub>6</sub>-acetone:  $\delta$ H = 2.05,  $\delta$ C = 29.84), and for <sup>31</sup>P NMR 85% phosphoric acid in sealed capillary tube is used as internal standard ( $\delta$ P = 0.00). Splitting patterns are indicated as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad. Preparative HPLC was performed on 1260 Infinity LC (Agilent, Santa Clara CA) equipped with a Prep-C18 column (Agilent, 30 × 250 mm, 10  $\mu$ m). Low resolution mass spectral analysis (ESI-MS) was carried out with LC/MS using 1220 Infinity LC (Agilent, Santa Clara CA) coupled with Expression-L Compact Mass Spectrometer (Advion, Ithaca NY). High resolution mass spectral analysis (ESI-MS) was carried out at LBNL Catalysis Facility at the Lawrence Berkeley National Laboratory (Berkeley Lab) using PerkinElmer AxION® 2 TOF MS.

#### 4.2.2 Synthesis of thiophosphorodichloridates

**General procedure for thiophosphorodichloridate synthesis.** Alcohol was dissolved in dry THF under N<sub>2</sub> and cooled in an ice/water bath. To this solution was added dropwise *n*BuLi (2.5 M solution in hexanes, 1.1 equiv.) solution and stirred for 10 min. In a separate flask PSCl<sub>3</sub> (1.2 equiv.) was dissolved in dry THF under N<sub>2</sub> and cooled in a dry-ice/acetone bath. The lithium salt suspension in THF was then added dropwise. The mixture was warmed up to room temperature and stirred for 1 h. THF was evaporated and the residue was dissolved in CH<sub>2</sub>Cl<sub>2</sub>, filtered to remove LiCl and purified by column chromatography (1:100 to 1:50 ethyl acetate/hexanes) to give the product as a clear to light yellow oil.

**But-3-yn-1-yl thiophosphorodichloridate (1, 4C-TPAC).** Following the general procedure for thiophosphorodichloridate synthesis, but-3-yn-1-ol (0.62 mL, 8.2 mmol) was reacted with *n*BuLi (3.6 mL of 2.5 M in hexanes, 9.0 mmol) and PSCl<sub>3</sub> (1.0 mL, 9.8 mmol) to provide compound **1** as a clear oil (1.2 g, 71%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.41 (dt, *J* = 12.0, 6.9 Hz, 2H), 2.72 (td, *J* = 6.9, 2.6 Hz, 2H), 2.09 (t, *J* = 2.6 Hz, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 78.18, 71.39, 69.01 (d, *J* = 10.1 Hz), 20.23 (d, *J* = 10.8 Hz). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ 58.72.

**(±)-Pent-4-yn-2-yl thiophosphorodichloridate (2, 4+1C-TPAC).** Following the general procedure for thiophosphorodichloridate synthesis, (±)-buta-3-yn-1-ol (0.77 mL, 8.2 mmol) was reacted with *n*BuLi (3.6 mL of 2.5 M in hexanes, 9.0 mmol) and PSCl<sub>3</sub> (1.0 mL, 9.8 mmol) to provide compound **2** as a light yellow oil (0.7 g, 39%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.09 – 4.98 (m, 1H), 2.73 – 2.57 (m, 2H), 2.10 (t, *J* = 2.7 Hz, 1H), 1.57 (d, *J* = 6.3 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 79.29 (d, *J* = 10.0 Hz), 78.14, 72.00, 27.01 (d, *J* = 7.0 Hz), 20.47 (d, *J* = 4.6 Hz). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ 58.28.

**Pent-4-yn-1-yl thiophosphorodichloridate (3, 5C-TPAC).** Following the general procedure for thiophosphorodichloridate synthesis, pent-4-yn-1-ol (0.35 mL, 4.1 mmol) was reacted with *n*BuLi (1.8 mL of 2.5 M in hexanes, 4.5 mmol) and PSCl<sub>3</sub> (0.50 mL, 4.9 mmol) to provide compound **3** as a clear oil (0.27 g, 27%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.42 (td, *J* = 11.0, 6.0 Hz, 2H), 2.34 (td, *J* = 6.8, 2.6 Hz, 2H), 2.03 – 1.90 (m, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 81.88, 70.50 (d, *J* = 10.5 Hz), 69.92, 28.27 (d, *J* = 10.0 Hz), 14.69 (s). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ 58.25.

**3-Propargyloxy-1-propanol (4).** This compound was synthesized according to published procedure.<sup>22</sup>

**3-Propargyloxy-1-propanyl thiophosphorodichloridate (5, O6C-TPAC).** Following the general procedure for thiophosphorodichloridate synthesis, 3-propargyloxy-1-propanol (**4**, 504 mg, 4.4 mmol) was reacted with *n*BuLi (1.94 mL of 2.5 M in hexanes, 4.9 mmol) and PSCl<sub>3</sub> (0.90 mL, 8.8 mmol) to provide compound **5** as a clear oil (0.63 g, 58%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.45 (dt, *J* = 11.1, 6.2 Hz, 2H), 4.15 (d, *J* = 2.4 Hz, 2H), 3.65 (t, *J* = 5.9 Hz, 2H), 2.44

(t,  $J = 2.4$  Hz, 1H), 2.11 – 2.04 (m, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  79.52, 74.79, 69.40 (d,  $J = 10.5$  Hz), 65.26, 58.41, 29.87 (d,  $J = 9.9$  Hz).  $^{31}\text{P}$  NMR (162 MHz,  $\text{CDCl}_3$ )  $\delta$  58.67.

**Propargyl tosylate (6).** This compound was synthesized according to published procedure for the preparation of  $^{13}\text{C}_3$ -propargyl tosylate.<sup>23</sup>

**2-Propargyloxyethanol (7).** Ethylene glycol (373  $\mu\text{L}$ , 6.7 mmol) was dissolved in dry THF (10 mL) containing NaH (357 mg 60% dispersion in mineral oil, 8.9 mmol). Tiny bubbles formed while stirring for 30 min. To this mixture was added tetrabutylammonium iodide (TBAI, 165 mg, 0.45 mmol) and propargyl tosylate (**6**, 937 mg, 4.46 mmol) in THF (10 mL). The mixture was stirred at refluxing temperature for 16 h to form a light brown suspension. It was then diluted in  $\text{H}_2\text{O}$  (25 mL), saturated with NaCl and extracted with 4 $\times$ ethyl ether. The organic phase was dried ( $\text{Na}_2\text{SO}_4$ ), concentrated and purified by silica column chromatography (1:3 ethyl acetate/hexanes) to give compound **7** as a yellow oil (90 mg, 20%).  $^1\text{H}$  NMR (400 MHz,  $d_6$ -acetone)  $\delta$  4.18 (d,  $J = 2.4$  Hz, 2H), 3.68 – 3.61 (m, 3H), 3.60 – 3.54 (m, 2H), 2.94 (t,  $J = 2.4$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $d_6$ -acetone)  $\delta$  80.95, 75.65, 72.22, 61.77, 58.55.

**2-Propargyloxyethyl thiophosphorodichloridate (8, O5C-TPAC).** Following the general procedure for thiophosphorodichloridate synthesis, 2-propargyloxyethanol (**7**, 600 mg, 6.0 mmol) was reacted with  $n\text{BuLi}$  (2.88 mL of 2.5 M in hexanes, 7.2 mmol) and  $\text{PSCl}_3$  (1.22 mL, 12.0 mmol) to provide compound **8** as a yellow oil (0.86 g, 62%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.51 – 4.44 (m, 2H), 4.24 (d,  $J = 2.1$  Hz, 2H), 3.88 – 3.84 (m, 2H), 2.48 (t,  $J = 2.1$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  78.86, 75.37, 70.37 (d,  $J = 10.2$  Hz), 67.41 (d,  $J = 10.2$  Hz), 58.56.  $^{31}\text{P}$  NMR (162 MHz,  $\text{CDCl}_3$ )  $\delta$  59.67.

**$^{13}\text{C}_3$ -Propargyl tosylate (9).** This compound was synthesized according to published procedure.<sup>23</sup>

**$^{13}\text{C}_5$ -2-Propargyloxyethanol (10).**  $^{13}\text{C}_2$ -Ethylene glycol (500 mg, 7.8 mmol) was dissolved in THF (10 mL) containing NaH (415 mg 60% dispersion in mineral oil, 10.4 mmol). Tiny bubbles formed while stirring for 30 min. To this mixture was added tetrabutylammonium iodide (TBAI, 192 mg, 0.52 mmol) and  $^{13}\text{C}_3$ -propargyl tosylate (**9**, 950 mg, 4.48 mmol) in THF (10 mL). The mixture was stirred at refluxing temperature for 16 h to form a light brown suspension. It was then diluted in  $\text{H}_2\text{O}$  (25 mL), saturated with NaCl and extracted with 4 $\times$ ethyl ether. The organic phase was dried ( $\text{Na}_2\text{SO}_4$ ), concentrated and purified by silica column chromatography (1:3 ethyl acetate/hexanes) to give compound **10** as a yellow oil (21 mg, 18%).  $^1\text{H}$  NMR (400 MHz,  $d_6$ -Acetone)  $\delta$  4.17 (dm,  $J = 147.1$  Hz, 2H), 3.64 (dm,  $J = 139.2$  Hz, 2H), 3.56 (dm,  $J = 140.2$  Hz, 2H), 2.93 (ddm,  $J = 245.0, 55.1$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $d_6$ -Acetone)  $\delta$  81.08 (dd,  $J = 168.6, 71.0$  Hz), 75.48 (dd,  $J = 168.5, 13.3$  Hz), 72.23 (d,  $J = 41.7$  Hz), 61.71 (ddd,  $J = 41.9, 12.9, 3.9$  Hz), 58.55 (dd,  $J = 70.2, 13.3$  Hz).

**<sup>13</sup>C<sub>5</sub>-2-Propargyloxy-1-ethyl thiophosphorodichloridate (11, <sup>13</sup>C<sub>5</sub>-O5C-TPAC).** Following the general procedure for thiophosphorodichloridate synthesis, <sup>13</sup>C<sub>5</sub>-2-propargyloxyethanol (**10**, 84 mg, 0.81 mmol) was reacted with *n*BuLi (0.36 mL of 2.5 M in hexanes, 0.90 mmol) and PSCl<sub>3</sub> (0.17 mL, 1.62 mmol) to provide compound **8** as a yellow oil (110 mg, 57%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.48 (dm, *J* = 150.8 Hz, 2H), 4.23 (dm, *J* = 151.1 Hz, 2H), 3.86 (dm, *J* = 146.0 Hz, 1H), 2.51 (dd, *J* = 241.0, 60.1 Hz, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 79.12 (ddd, *J* = 171.9, 70.0, 1.8 Hz), 75.20 (dd, *J* = 171.8, 15.0 Hz), 70.45 (ddd, *J* = 42.7, 10.2, 4.3 Hz), 67.48 (dd, *J* = 42.6, 10.1 Hz), 58.64 (dddd, *J* = 70.2, 15.2, 4.1, 1.4 Hz). <sup>31</sup>P NMR (162 MHz, CDCl<sub>3</sub>) δ 59.53.

#### 4.2.3 Synthesis of modified Fmoc-His-OH

**General procedure for modified Fmoc-His-OH synthesis.** These alkyne-bearing phosphohistidine derivatives were synthesized according to Chapter 3. Namely, Fmoc-His-OH and K<sub>2</sub>CO<sub>3</sub> (15.0 equiv.) was dissolved in 2:3 H<sub>2</sub>O/MeCN cooled in an ice/water bath. Thiophosphorodichloridate (3.0 equiv.) was dissolved in MeCN and added by syringe pump over 1 h. The mixture was stirred vigorously for another 2 h, concentrated and purified by HPLC.

**Potassium *O*-propargyl (*N*<sup>2</sup>-Fmoc-L-histidino)thiophosphate.** This compound was synthesized according to the procedure in Chapter 3.

**Potassium *O*-but-3-yn-1-yl (*N*<sup>2</sup>-Fmoc-L-histidino)thiophosphate.** Following the general procedure for modified Fmoc-His-OH synthesis, Fmoc-His-OH (62.0 mg, 0.16 mmol) and K<sub>2</sub>CO<sub>3</sub> (340 mg, 2.5 mmol) in 5 mL of 60% MeCN was reacted with 4C-TPAC (100 mg, 0.49 mmol) to provide the titled compound as a light yellow solid (57 mg, 66%). LRMS (ESI<sup>+</sup>) calcd 526.1, found 526.3 for C<sub>25</sub>H<sub>25</sub>N<sub>3</sub>O<sub>6</sub>PS<sup>+</sup> (M+H<sup>+</sup>).

**Potassium *O*-pent-4-yn-2-yl (*N*<sup>2</sup>-Fmoc-L-histidino)thiophosphate.** Following the general procedure for modified Fmoc-His-OH synthesis, Fmoc-His-OH (62.0 mg, 0.16 mmol) and K<sub>2</sub>CO<sub>3</sub> (340 mg, 2.5 mmol) in 5 mL of 60% MeCN was reacted with 4+1C-TPAC (107 mg, 0.49 mmol) to provide the titled compound as a light yellow solid (17.4 mg, 18%). LRMS (ESI<sup>+</sup>) calcd 540.1, found 540.2 for C<sub>26</sub>H<sub>27</sub>N<sub>3</sub>O<sub>6</sub>PS<sup>+</sup> (M+H<sup>+</sup>).

**Potassium *O*-pent-4-yn-1-yl (*N*<sup>2</sup>-Fmoc-L-histidino)thiophosphate.** Following the general procedure for modified Fmoc-His-OH synthesis, Fmoc-His-OH (62.0 mg, 0.16 mmol) and K<sub>2</sub>CO<sub>3</sub> (340 mg, 2.5 mmol) in 5 mL of 60% MeCN was reacted with 4+1C-TPAC (107 mg, 0.49 mmol) to provide the titled compound as a light yellow solid (51.8 mg, 58%). LRMS (ESI<sup>+</sup>) calcd 556.1, found 556.3 for C<sub>26</sub>H<sub>27</sub>N<sub>3</sub>O<sub>7</sub>PS<sup>+</sup> (M+H<sup>+</sup>).

**Potassium *O*-(2-propargyloxyethyl) (*N*<sup>2</sup>-Fmoc-L-histidino)thiophosphate.** Following the general procedure for modified Fmoc-His-OH synthesis, Fmoc-His-OH (54.0 mg, 0.14 mmol) and K<sub>2</sub>CO<sub>3</sub> (297 mg, 2.1 mmol) in 5 mL of 60% MeCN was reacted with O5C-TPAC (100 mg,

0.43 mmol) to provide the titled compound as a light yellow solid (26.4 mg, 58%). LRMS (ESI<sup>+</sup>) calcd 540.1, found 540.2 for C<sub>26</sub>H<sub>27</sub>N<sub>3</sub>O<sub>6</sub>PS<sup>+</sup> (M+H<sup>+</sup>).

#### 4.2.4 Synthesis of triply-charged azides

##### **Tetrakis(2-carboxyethoxymethyl)methane tetraethyl ester (11).**

Tetrakis(2-cyanoethoxymethyl)methane was synthesized according to literature procedure.<sup>24</sup> Briefly, pentaerythritol (3.9 g, 28.6 mmol), acrylonitrile (9.45 mL, 144 mmol) and NaOH (0.39 mL of 40% aqueous solution) was stirred at room temperature for 20 h to form a homogenous solution, which was diluted in H<sub>2</sub>O (80 mL), extracted with ethyl acetate (2 × 100 mL), dried and concentrated to afford tetra(2-cyanoethoxymethyl)methane as a thick, light oil. The crude compound was mixed with ethanol (90 mL) and added carefully to H<sub>2</sub>SO<sub>4</sub> (30 mL, 0.56 mmol) cooled in an ice/water bath. The mixture was refluxed under N<sub>2</sub> for 30 h, during which diethyl ether was evaporated. The resulting solution was diluted in H<sub>2</sub>O and extracted with 4× ethyl acetate. The combined organic layer was washed with saturated NaHCO<sub>3</sub> and saturated NaCl, dried and concentrated to provide **10** as a light yellow syrup (12.0 g, 78%). This compound was used without further purification. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.17 – 4.08 (m, 2H), 3.69 – 3.58 (m, 8H), 2.53 (dt, *J* = 16.9, 6.6 Hz, 8H), 1.30 – 1.21 (m, 12H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 171.88, 69.63, 66.93, 60.51, 35.27, 14.36.

**Tetrakis(3-hydroxypropyloxymethyl)methane (12).** To a suspensions of LiAlH<sub>4</sub> (1.70 g, 46.5 mmol) in anhydrous THF (150 mL) at 0 °C, compound **11** (5.0 g, 9.3 mmol) in anhydrous THF (50 mL) was added dropwise. The reaction mixture was stirred at room temperature and monitored by TLC. Upon completion, the reaction mixture was cooled to 0°C and quenched by adding H<sub>2</sub>O (1.7 mL), 15% NaOH (1.7 mL) and H<sub>2</sub>O (5.1 mL) sequentially. The mixture was filtered and the filtrate was evaporated under vacuum. The residue was purified with silica column chromatography (10:1 ethyl acetate/methanol) to give **2** as a colorless oil (3.10 g, 90%). The NMR spectra of the obtained compound are consistent with the reported data.<sup>25</sup>

**Tetrakis(3-methylsulfonyloxypropyloxymethyl)methane (13).** To a stirred solution of compound **12** (3.03 g, 8.1 mmol) and triethylamine (4.10 g, 40.5 mmol) in 2:1 THF/CH<sub>2</sub>Cl<sub>2</sub> (30 mL) at 0°C was added a solution of MsCl (4.63 g, 40.5 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (10 mL). Then the reaction mixture was warmed to room temperature and stirred overnight. The mixture was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The organic layer was washed with 10% HCl, saturated NaHCO<sub>3</sub>, water and saturated NaCl, dried over anhydrous MgSO<sub>4</sub> and concentrated. The residue was purified with silica column chromatography (1:1 CH<sub>2</sub>Cl<sub>2</sub>/ethyl acetate) to give **13** as a white solid (3.40 g, 62%). The NMR spectra of the obtained compound are consistent with the reported data.<sup>25</sup>

##### **3-Azidopropyloxymethyl-tris(3-methylsulfonyloxypropyloxymethyl)methane (14).**

Compound **13** (1.36 g, 2.0 mmol) and NaN<sub>3</sub> (0.14 g, 2.4 mmol) were mixed with DMF (4 mL) under N<sub>2</sub> and stirred at 60°C for 48 h. The mixture was then concentrated and purified by silica column chromatography (10:1 CH<sub>2</sub>Cl<sub>2</sub>/ethyl acetate) to provide **14** as a white solid (0.47 g, 38%).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.31 (t,  $J = 6.3$  Hz, 6H), 3.49 (t,  $J = 5.9$  Hz, 6H), 3.46 (t,  $J = 5.9$  Hz, 2H), 3.38 – 3.34 (m, 8H), 3.01 (s, 9H), 1.98 (p,  $J = 6.1$  Hz, 6H), 1.82 (p,  $J = 6.3$  Hz, 2H), 1.25 (t,  $J = 7.1$  Hz, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  69.65, 69.57, 67.93, 67.30, 66.61, 48.51, 45.37, 37.30, 31.62, 29.42, 29.04.

### **3-Azidopropylloxymethyl-tris(3-(1*H*-imidazol-1-yl)propylloxymethyl)methane**

**methanesulfonic salt (15, trisimidazole-azide).** NaH (120 mg, 3.0 mmol) was added in small portions at 0 °C under  $\text{N}_2$  to a solution of imidazole (206 mg, 3.0 mmol) in anhydrous THF (2 mL). The solution was then stirred at room temperature for 12 h, after which, compound **14** (290 mg, 0.46 mmol) in anhydrous THF (1 mL) was added dropwise. The reaction mixture was stirred at room temperature and monitored by TLC. Upon completion, the mixture was concentrated and purified with silica gel chromatography (3:1 ethyl acetate/methanol) to give **15** as a pale yellow oil (128 mg, 51%). HRMS (ESI<sup>+</sup>) calcd 544.3355, found 544.3396 for  $\text{C}_{26}\text{H}_{42}\text{N}_9\text{O}_4^+$  (M+H<sup>+</sup>); Calcd 272.6714, found 272.6727 for  $\text{C}_{26}\text{H}_{43}\text{N}_9\text{O}_4^+$  (M+2H<sup>+</sup>).

### **3-Azidopropylloxymethyl-tris(3-dimethylaminopropylloxymethyl)methane methanesulfonic salt (16, trisamino-azide).**

Compound **14** (31 mg, 0.05 mmol) was dissolved with dimethylamine (6 mL of 1 M solution in THF, 6 mmol) in a pressure tube. The solution was stirred at 75°C for 24 h. Upon cooling dimethylammonium methylsulfonate precipitated as crystals, which was filtered off. The filtrate was concentrated to provide a yellow, thick oil. HRMS (ESI<sup>+</sup>) calcd 475.3967, found 475.3999 for  $\text{C}_{23}\text{H}_{51}\text{N}_6\text{O}_4^+$  (M+H<sup>+</sup>); calcd 238.2020, found 238.2026 for  $\text{C}_{23}\text{H}_{52}\text{N}_6\text{O}_4^+$  (M+2H<sup>+</sup>).

### **3-Azidopropylloxymethyl-tris(3-(trimethylammonio)propylloxymethyl)methane**

**methanesulfonic salt (17, trisammonium-azide).** A mixture of compound **14** (188 mg, 0.3 mmol) and  $\text{Me}_3\text{N}$  (2.40 g, 40.5 mmol) in MeOH (6 mL) was stirred at room temperature for 14 h and then refluxed for 2 h under  $\text{N}_2$ . The reaction mixture was concentrated to give compound **17** as a white solid. For further click reaction, 100 mg of the crude product was dissolved in  $\text{H}_2\text{O}$  and separated on a Biorad FPLC (Hercules, CA) equipped with a HiTrap SP HP strong cation exchange column (1 mL, GE Healthcare, Chicago Illinois) using a linear gradient from 5 mM to 2 M ammonium acetate. The purified fractions were lyophilized to provide **17** as a white solid. HRMS (ESI<sup>+</sup>) calcd 173.1527, found 173.1517 for  $\text{C}_{26}\text{H}_{59}\text{N}_6\text{O}_4^{3+}$  (M<sup>3+</sup>)

#### *4.2.5 Analysis of stability against hydrolysis of thiophosphorodichloridate-labeled adduct*

CuAAC reactions between thiophosphorodichloridate-labeled Fmoc-His-OH and 3-azido-1-propanol were carried out under the following condition: 100  $\mu\text{M}$  labeled Fmoc-His-OH, 200  $\mu\text{M}$  3-azido-1-propanol, 500  $\mu\text{M}$  THPTA, 100 mM  $\text{CuSO}_4$  and 5 mM sodium ascorbate (all added from their 100 $\times$  stock solution) in 25 mM HEPES, pH 7.5. The mixture was incubated at 37°C for 1 h. After filtration, 30  $\mu\text{L}$  of the solution was injected for immediate analysis by a LC/MS using 1220 Infinity LC (Agilent, Santa Clara CA) coupled with a SB-C18 Zorbax rapid resolution cartridge (Agilent) and an Expression-L Compact Mass

Spectrometer (Advion, Ithaca NY). Solvent A was water + 0.05% formic acid and solvent B was methanol + 0.05% formic acid. The linear gradient employed was 5-100% B in 8 min and 100% B for 4 min, with mass spectrometer connected 2 minutes after injection.

Stability of thiophosphorodichloridate-histidine adduct in HeLa cell lysate was tested by adding 100  $\mu$ M labeled Fmoc-His-OH into 1 mg/mL HeLa cell lysate in 50 mM HEPES, pH 7.5. The solution was incubated in a 37°C water bath and after indicated time, 10  $\mu$ L of the solution was added to 30  $\mu$ L of cold methanol in a PCR tube. Proteins were removed by cooling at -80°C for  $\geq$  1h followed by centrifugation. All supernatant was injected for analysis by HPLC using an Agilent 1100 series LC system, with the same column and mobile phase gradient as above.

#### 4.2.6 Analysis of CuAAC reaction between O5C-TPAC modified bombesin and azides

Bombesin (5 mg) was suspended in 1:1 DMF/MeOH (500  $\mu$ L) and stored at -80°C. For synthesis of O5C-TPAC modified bombesin, 20  $\mu$ L of the stock suspension was diluted into 180  $\mu$ L of 50 mM HEPES, pH 8.5. To this solution was added 2 mM O5C-TPAC (100 $\times$  solution in MeCN) and incubated for 1 h. The resulting solution was separated on a 1220 Infinity LC (Agilent, Santa Clara CA) coupled with a SB-C18 Zorbax rapid resolution cartridge (Agilent), using the same mobile phase gradient as in 4.2.5. The fraction containing the modified peptide was evaporated to dryness in a 4 mL vial to provide O5C-TPAC-modified bombesin with methionine oxidized sulfoxide. The solid was reconstituted in 50  $\mu$ L 25 mM HEPES pH 7.5, split into four PCR tubes for click reaction with 3-azido-1-propanol, trisimidazole-azide, trisamino-azide or trisammonium-azide, using the follow condition: 200  $\mu$ M organic azide, 500  $\mu$ M THPTA, 100 mM CuSO<sub>4</sub>, 1 mM amionguanidine hydrochloride and 5 mM sodium ascorbate (all added from their 100 $\times$  stock solution) in 25 mM HEPES, pH 7.5. The solution was mixed and incubated at 37°C for 1 h, quenched by adding 2.5 mM Na<sub>4</sub>(edta) and frozen at -80°C until analysis.

LC-MS/MS analysis using HCD fragmentation was carried out on a Q-Exactive Plus mass spectrometer (Thermo Fisher) operated with an Agilent 1260 HPLC system (Agilent, Santa Clara CA) with a self-packed column (100  $\mu$ m  $\times$  130 nm, 3  $\mu$ m Aqua C18 125 Å, Phenomenex, Torrance CA). LC-MS/MS analysis using ETD fragmentation was carried out at QB3/Chemistry Mass Spectrometry Facility at the University of California, Berkeley on a LTQ-Orbitrap-XL mass spectrometer equipped with an electrospray ionization (ESI) source, operated in the positive ion mode, and connected in line with an UltiMate3000 RSLCnano liquid chromatography (LC) system (Thermo Fisher). The LC system was equipped with a reversed-phase analytical column (Acclaim PepMap100, C18, 3  $\mu$ m, 0.075  $\times$  250 mm, 100 Å, Thermo Fisher).

#### 4.2.7 Cell lysate preparation

HeLa cells were grown on 15 cm dishes until 90% confluency. Cells were placed on ice and washed carefully with 3  $\times$  50 mM HEPES, pH 7.5 and were then scraped into 1.5 mL of 50 mM HEPES, pH 7.5 containing cOmplete, EDTA-free Protease Inhibitor (Roche). The cell suspension was transferred into 20 mL glass scintillation vial and lysed by probe sonicator on ice.

The lysate was centrifuged 16,000×g for 15 min and the precipitate was discarded. The concentration of the cell lysate was adjusted to 1 mg/mL by BCA assay. The lysate was then splitted into 1 mL aliquots and frozen at -80°C until usage.

#### 4.2.8 Lysate labeling, digestion and enrichment

The LC-MS/MS sample preparation protocol, including lysate labeling and fractionation with O5C-TPAC and trisammonium-azide is described in Appendix 3 (A3.4.3).

For blockage of cysteine and lysine, Step 3 (Alkylating cysteine) and Step 4 (Removing excess urea, IAA and TCEP) is changed to the following procedure, which is adapted from a procedure for N-terminome:<sup>26</sup>

Dissolved the dried pellet in 200 mM HEPES, pH 7.5 containing 2.5 M guanidine hydrochloride. Add 5 mM TCEP and 10 mM IAA and incubate in the dark at 37°C for 0.5 h. Dilute the resulting solution with DMSO (200 µL) and add 10 mM *N*-succinimidyl acetate (100 × in DMSO). Incubate at room temperature for 30 min in the dark. Split the protein solution into two 1.5 mL microcentrifuge tubes and precipitate the protein by adding methanol (100 µL) and acetone (800 µL) to each half. Cool the mixture in -20°C freezer for ≥ 3h. Centrifuge one tube, combine with another, centrifuge and discard the supernatant. Wash the pellet by adding 1000 µL 1:4 H<sub>2</sub>O / acetone. Centrifuge again and discard supernatant.

#### 4.2.9 LC-MS/MS analysis of digested lysate after fractionation

LC-MS/MS analysis using HCD fragmentation was performed on a Q-Exactive Plus mass spectrometer (Thermo Fisher) operated with an Agilent 1260 HPLC system (Agilent, Santa Clara CA) using vented column setup. Samples were pressure-loaded onto a self-packed column (100 µm × 130 nm, 3 µm Aqua C18 125 Å, Phenomenex, Torrance CA) with a self-pulled electrospray emitter tip. Solvent A was 5% MeCN and 0.1% formic acid in H<sub>2</sub>O; solvent B was 80% MeCN, 20% H<sub>2</sub>O with 0.1% formic acid. The linear gradient program for proteomics samples consists of 0 min: 100% A, 3 min: 100% A, 40 min: 55% A, 48 min 0% A, 62 min 0% A with a constant pump flow rate of 0.150 mL/min. HCD MS/MS spectra were recorded in the data-dependent mode using a Top 15 method analyzing precursor ions with 2<z<8.

LC-MS/MS analysis of proteomics samples using ETD fragmentation was carried out at UC Davis Proteomics Core facility, using an Orbitrap Fusion Lumos (Thermo Fisher) mass spectrometer. The acquired raw file was processed using Trans-Proteomics Pipeline (Institute for Systems Biology, Seattle WA) with the comet search algorithm, searching for modification on His, Lys, Ser, Thr and Tyr, with peptide assignments analyzed by peptideprophet, filtering off assignments with peptideprophet P<0.8.

## 4.3 Results and discussion

### 4.3.1 Design and synthesis of thiophosphorodichloridate with long alkynes

Although our first published thiophosphorodichloridate, TPAC with propargyl as the click reaction handle, is effective at labeling and provides clean CuAAC products on model proteins, this reagent shows major hydrolysis during click reaction when applied to small molecule histidine analogs (Scheme 4.1a). We hypothesize that the unexpected hydrolysis originates from the binding of copper ions to the propargyl group and nearby sulfur atom, in an analogous way palladium catalyst hydrolyzes the propargyl carbamate in biological-relevant conditions.<sup>27-29</sup> To circumvent this undesired pathway, we designed a series of thiophosphorodichloridate reagents with alkyne chains longer than three carbons, both to remove the special binding property of propargyl and to separate the sulfur and phosphorus from the metal complex in CuAAC (Scheme 4.1b). For easy recognition these new compounds are named according to the length and composition of the alkyne chain, namely, 4+1C-TPAC, 4C-TPAC and 5C-TPAC.

Aside from thiophosphorodichloridate containing hydrocarbon handles, we also designed reagents with ether-linked propargyl group to build isotopically-coded reagents for mass spectrometry quantification. Specifically, we synthesized O6C-TPAC and O5C-TPAC, whose D<sub>6</sub> or <sup>13</sup>C<sub>5</sub> counterparts can be synthesized from commercially-available isotope-labeled chemicals (Scheme 4.1b). To avoid retention time shift caused by deuterium substitution, we focused on the synthesis of <sup>13</sup>C<sub>5</sub>-O5C-TPAC. Synthesis of this compound starting from <sup>13</sup>C<sub>2</sub>-acetic acid resulted in side products that are difficult to separate by column chromatography (Figure 4.1a). We therefore switched to the synthesis using <sup>13</sup>C<sub>2</sub>-ethylene glycol as substrate, and was able to obtain clean <sup>13</sup>C<sub>5</sub>-O5C-TPAC that gave single adduct upon reaction with Fmoc-His-OH (Figure 4.1a,b). Together with light O5C-TPAC, we believe this set of compounds will enable quantitative profiling of histidine in proteomes under two different conditions.

### 4.3.2 Reactivity of the thiophosphorodichloridate reagents on Fmoc-His-OH

With these compounds in hand, we first tested to see if they all show similar reactivity towards histidine residue. We applied the thiophosphorodichloridate reagents to 50 mM HEPES buffer, pH 8.5 containing Fmoc-His-OH as a substrate and analyzed their reactions by LC/MS (Figure 4.2ab). We were able to observe improved reactivity for these reagents with longer alkynes compared to the original 3C-TPAC, probably due to the higher hydrophobicity that bring closer the reagent and the Fmoc protection group. The high reactivity of these reagents further proves that thiophosphorodichloridate is an excellent functionality for fast labeling of histidine.

We then moved on to test our hypothesis of metal-catalyzed hydrolysis of the modification on histidine. For this purpose we scaled-up the labeling reaction to obtain purified, TPAC-labeled Fmoc-His-OH as the substrate for its stability test. We then applied 3C-TPAC, 4C-TPAC, 5C-TPAC and O5C-TPAC-modified Fmoc-His-OH to CuAAC reaction with 3-azido-1-propanol as a model azide and subject to immediate LC/MS analysis (Figure 4.2a,b,c). Consistent with our previous observation, 3C-TPAC-modified Fmoc-His-OH experienced major

hydrolysis and reverted to unmodified Fmoc-His-OH during the reaction, with trace amount product only detectable on mass spectrometer. In sharp contrast, with a single extra methylene group, 4C-TPAC-modified Fmoc-His-OH showed mostly CuAAC adduct as the product, with a little starting material hydrolyzed. For longer-chained reagents, 5C-TPAC and O5C-TPAC, hydrolyzed product was even less, indicating that the replacement of propargyl and separation of the copper and sulfur/phosphorus effectively reduces such metal-catalyzed hydrolysis. To note, the CuAAC yield for 5C-TPAC and O5C-TPAC are 81% and 82%, respectively, which is comparable to the yield of 87% under the same condition for Fmoc-propargyl-Gly-OH, an analog of the widely-used alkyne handle on proteins and peptides.

With these positive results, we questioned whether the observed hydrolysis of 3C-TPAC labeled Fmoc-His-OH in cell lysate was also induced by such metal-catalyzed reaction. We thus performed the stability assay again in the presence of 1 mg/mL HeLa lysate at 37°C (Figure 4.3). In contrast to the fast hydrolysis of 3C-TPAC labeled Fmoc-His-OH, 5C-TPAC and O5C-TPAC counterparts showed negligible hydrolysis under the same condition, which suggests that the hydrolysis of the 3C-TPAC adduct indeed originates from the activity metals and/or metalloproteins present in the lysate. The high stability of the product of our newer generation of TPAC molecules enables smoother labeling and enrichment steps of the proteomics sample with less loss of such modification.

#### *4.3.3 Labeling of the thiophosphorodichloridate reagents on HeLa lysate*

After establishing the reactivity on small molecule model, we moved on to test the reactivity of the new TPAC reagents on lysate. We treated HeLa cell lysate with each thiophosphorodichloridate reagent followed by CuAAC with rhodamine-azide for fluorescence readout. The results suggest that the new TPAC reagents all possess higher labeling efficiency compared to 3C-TPAC, with linear TPAC reagents better than branched 4+1C TPAC molecule (Figure 4.4). We believe that the increase in signal partially comes from the higher hydrophobic interaction between the long alkyne chain and the protein that facilitate the reaction, and partially from the less hydrolysis during CuAAC reaction on sites that are easy to hydrolyze. The overall increase in signal suggest that these new reagents will outperform the 3C-TPAC for profiling purposes with labeling and click reaction combined to provide the final readout by mass spectrometry.

#### *4.3.4 Design and synthesis of charged azides for improved ETD fragmentation*

As mentioned in the previous chapter, the modification introduced by thiophosphorodichloridate reagents is a phosphohistidine analog that undergoes strong neutral loss on traditional tandem mass spectrometer with collision-based fragmentation (including collision-induced dissociation, CID, and Higher-energy C-trap dissociation, HCD). To avoid this issue that results in unidentifiable sites of modification, we turned our attention to newer tandem mass spectrometers with electron-based fragmentation technique, specifically electron-transfer dissociation (ETD), that does not induce neutral loss of phosphopeptides.<sup>30</sup> Although already applied in proteomics analysis, canonical ETD mass spectrometers often suffer from low

dissociation yield and thus poor MS2 information, especially for precursor ions with low cation charge ( $\leq 2$ ) density.<sup>31,32</sup> In an effort to improve the fragmentation yield, researchers have applied collision-based fragmentation following electron-based dissociation and designed EThcD technology; however, the collision step can still possibly result in neutral loss of the phosphate modification and also complicate the spectrum for proteomics analysis by giving rise to b- and y-type product ions in addition to ETD's c- and z-type ions.

In order to apply our thiophosphorodichloridate reagents to histidine profiling using more available canonical ETD mass spectrometers, we decided to introduce extra positive charges onto the modified peptide by CuAAC reaction. Previous reports have confirmed that increasing charge states of precursor ions by either chemical derivatization or tuning the mobile phase, the dissociation efficiency can be greatly enhanced, especially for +2 charge state peptides.<sup>33,34</sup> With the alkyne handle already installed on our modified peptides, it is straightforward to pull up the charge state with multiply-charged azides. After trying different designs, we finally settled on an extended pentaerythritol derivative, with one end carrying an azido group and the three other ends being either basic residues or ammonium groups (Scheme 4.2). The ether-linked arms from pentaerythritol reduce the steric hindrance during CuAAC reaction and also allows for easy substitution of the mesylate with various nucleophilic groups. As such, we successfully synthesized trisimidazole-azide as a mobile proton model, trisamion-azide as an immobilized proton moiety, and trisammonium-azide as a fixed charged fragment (Scheme 4.2).

#### 4.3.5 Evaluation of ETD fragmentation of O5C-TPAC/charged azide modified peptide

Following the successful synthesis of the charge azide compounds, we carried on to evaluate whether these compounds can increase the charge state of peptides for more fragmentation information. We used bombesin, a fourteen-peptide hormone, as a model substrate due to the lack of competing residues that can complicate the synthesis of phosphorodichloridate-modified peptide. We first labeled this peptide with O5C-TPAC and then performed CuAAC with three of our charged azide compounds, as well as 3-azido-1-propanol as a charge-neutral control, followed by LC-ETD MS/MS analysis of the obtained product (Figure 4.5 – 4.9). As we expected, these charged azides successfully raised the charged state on MS1 spectra from mainly +2 for O5C-TPAC modified TPAC and its CuAAC adduct with 3-azido-1-propanol, to +3 and +4 for CuAAC adduct with these charged azides. The fragmentation of the +2 precursor ions are, indeed, quite inefficient and results in poor sequence information on their MS2 spectra except the unmodified peptide (Figure 4.5-4.7). Although the fragmentation was only improved to a limited extent for trisimidazole-azide-derivatized peptides (Figure 4.8), the +3/+4-charged precursor ions from trisammonium-azide and trisamino-azide modification provide information-rich product ions that allow for the accurate determination of the modification site (Figure 4.9, 4.10). With this result we carried on our project with trisammonium-azide, since this molecule is most effective at increasing charge state of the peptide regardless of pH.

#### 4.3.6 Enrichment of O5C-TPAC/charged azide modified peptide by ion exchange column

After confirming that the charge increase by CuAAC with trisammonium-azide facilitates ETD analysis of peptides, we turned our attention to the enrichment of modified peptide for proteomics sample preparation. Such enrichment is necessary for proteomics sample preparation because the columns used in proteomics analysis with diameters  $< 0.1$  mm enables ultra-high sensitivity but at the same time comes with limited sample loading capacity. Moreover, the complex unenriched sample of peptides proves challenging to be separated by LC and even by their masses, resulting in two or more adjacent precursor ions sent into fragmentation with complicated MS2 spectrum. Therefore, enrichment of the modified peptides will increase the signal of the peptides of interest and at the same time reduce the complexity of the sample for easier analysis.

Although our modification with trisammonium-azide does not carry an affinity label for enrichment, we envision that its unusually high positive charge density can be utilized for enrichment by ion-exchange chromatography. For this purpose, we employed a strong-cation exchange spin column, using 300 mM NaCl buffered by 1:1 formic acid/formate to elute peptides with low positive charge, 50 mM ammonium acetate to neutralize the column and finally 3 M ammonium acetate to elute the peptides that remained after NaCl wash. We then subjected each fraction to LC-MS/MS analysis using HCD fragmentation. We were able to observe  $>8$ -fold higher signal in the NaCl wash fraction than in the elution fraction. To estimate the relative abundance of modified peptides in each fraction, we utilized the “neutral loss” character of this phosphate modification during HCD fragmentation and counted the number of MS2 spectra containing the cleaved trisammonium-thiometaphosphate ion (Figure 4.11a). As anticipated, although there were a few modified peptides detected in the wash and neutralization fractions, the elution fraction contained much more enriched peptides carrying O5C-TPAC/trisammonium modification (Figure 4.11b,c). Collectively, these results show that a simple fractionation step by strong cation exchange column can effectively enrich the O5C-TPAC/trisammonium modified peptides for proteomics sample preparation.

#### 4.3.7 Analysis of enriched O5C-TPAC/trisammonium-modified peptides by LC-ETD MS/MS

With the optimized proteomics sample preparation protocol in hand, we subject the resulting proteomics sample with enriched O5C-TPAC/trisammonium modified peptides to LC-ETD MS/MS analysis. Considering that target peptides are highly charged, we set the mass spectrometer to only analyze MS1 precursor ions with  $z \geq 3$  to enhance time-efficiency of the analysis. However, contrary to what we observed for 3C-TPAC modification on model proteins in Chapter 3, most modified peptides analyzed carry the modification on lysine residue (Figure 4.12). We hypothesize that this unexpected behavior originates from the gas-phase transfer of the thiophosphate modification from histidine to lysine residues during ETD fragmentation process. Indeed, transfer of phosphate modification has been reported from phosphorylated lysine to nearby lysine and serine residues during ETD fragmentation,<sup>35</sup> and around half the detected modified lysine residues have a nearby histidine residue. Therefore, to avoid such complications,

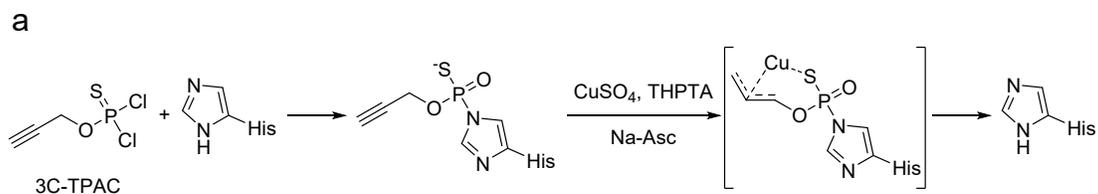
we protected lysine residues with *N*-succinimidyl acetate right after blockage of cysteine residues. Proteomics samples prepared from such protocol is currently under analysis.

#### **4.4 Conclusion**

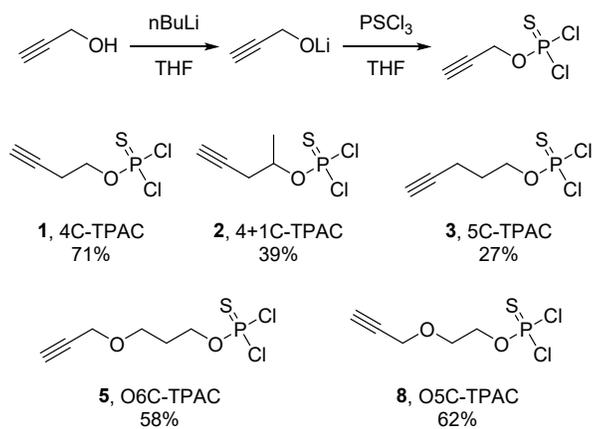
In this chapter we have optimized the reagents and protocol for histidine profiling using thiophosphorodichloridate reagent. By incorporating a longer alkyne chain rather than a propargyl group as the CuAAC handle, we were able to significantly increase the stability of the modified histidine residue against hydrolysis in the presence of metal ions; the new generation of TPAC reagents now offer much more stable adduct with histidine for further CuAAC reaction. In terms of tandem mass spectrometry, we opted to ETD-based LC-MS/MS for proteomics analysis, and to increase the signal of the tandem mass spectrum, we designed three charged azides to be coupled onto modified peptides for increase of positive charge. On our model peptide substrate, the trisammonium-azide results in increased charged state of +4/+3 on LC-MS from the unmodified +2, which greatly improved the quality ETD fragmentation spectra. The increase in charge also enables facile enrichment by ion exchange chromatography. Although initial proteomics analysis reports modifications mainly on lysine residues, we are in the process of tuning the procedure towards profiling of histidine residues, and will apply the isotopically-light and heavy pair of O5C-TPAC for quantitative analysis of histidine state under different biological conditions. The charged azide design also provides insights for signal enhancement in experiments utilizing ETD mass spectrometers.

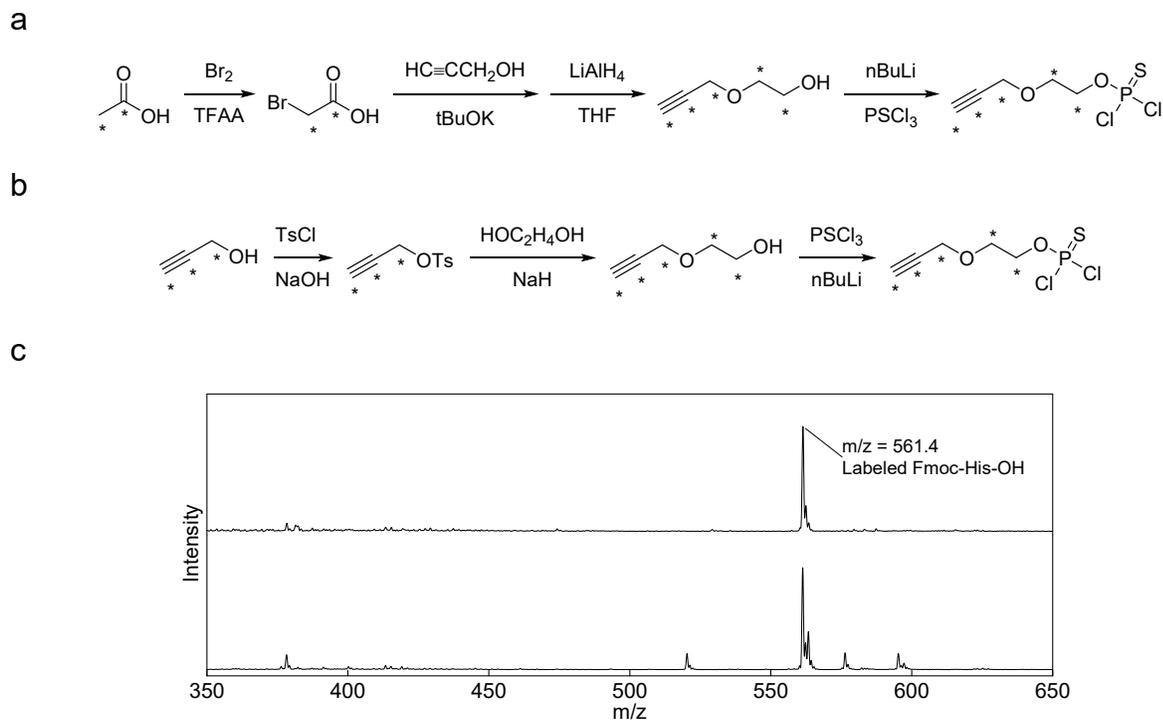
## 4.5 Figures and schemes

**Scheme 4.1** (a) Hypothesized hydrolysis pathway of 3C-TPAC modified histidine during CuAAC reaction. (b) Synthesis of newer TPAC with longer alkyne chains.

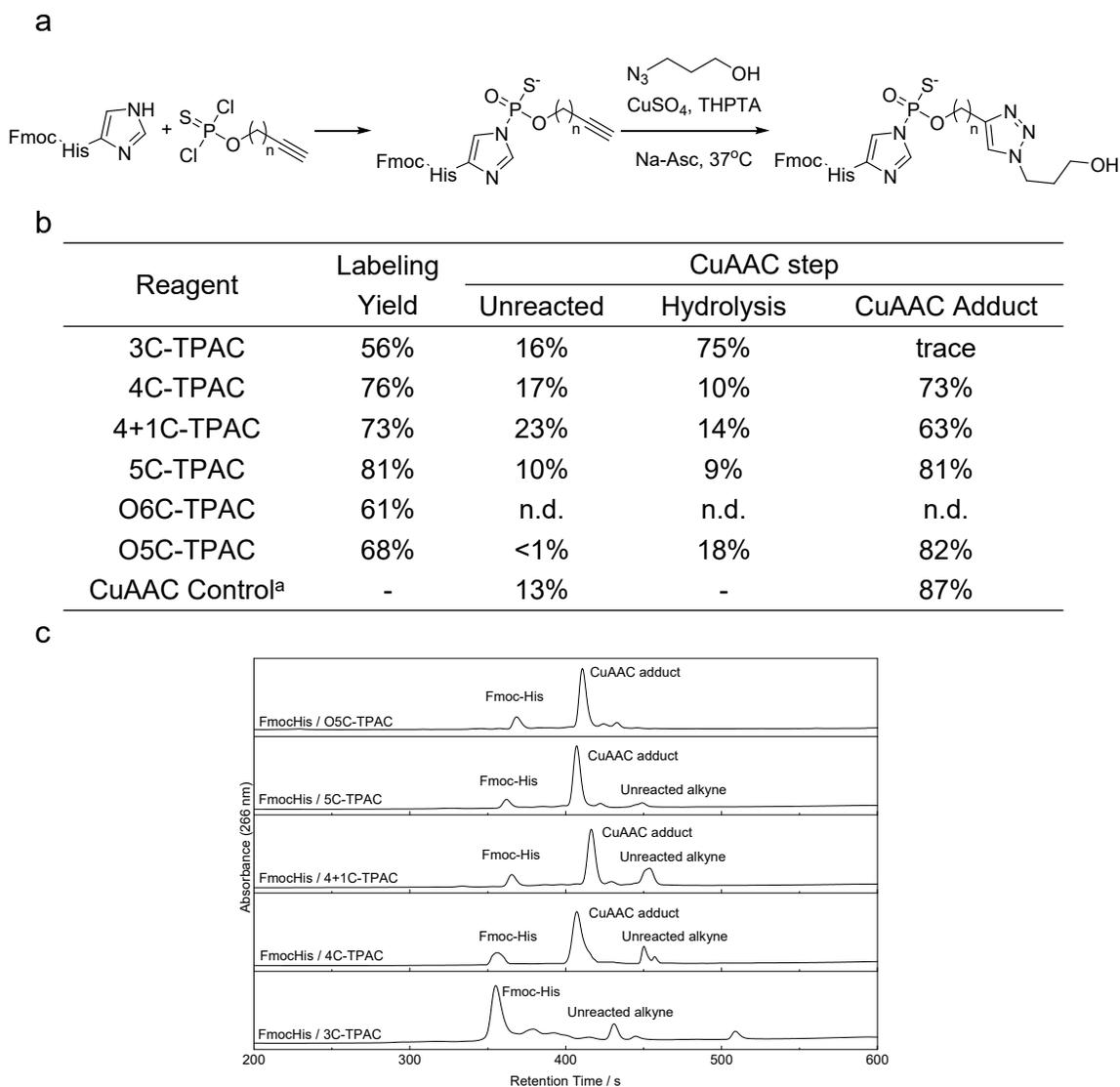


**b**

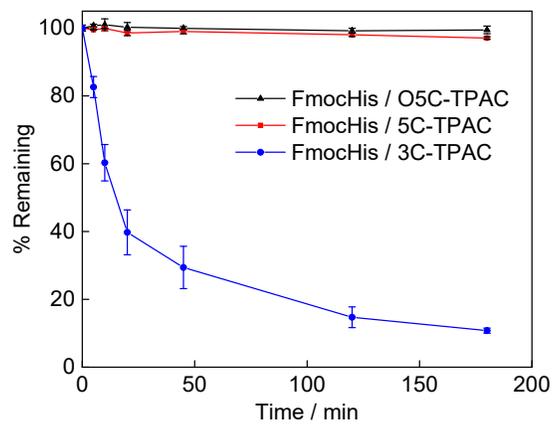




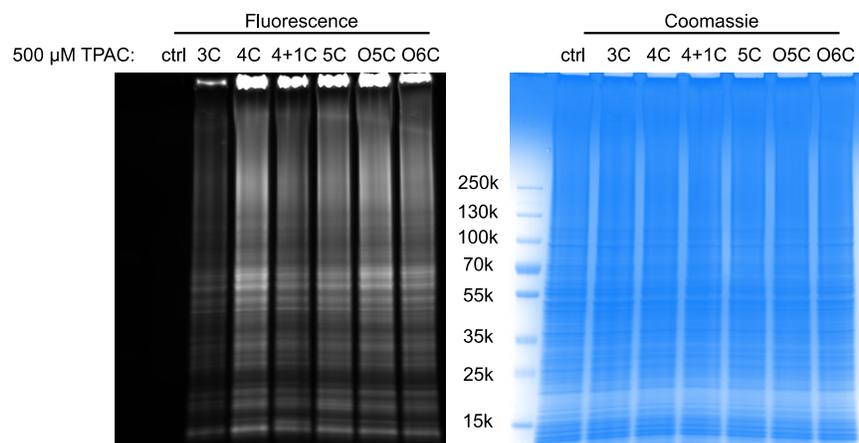
**Figure 4.1** Synthesis of isotopically-heavy O5C-TPAC. (a) Synthesis of  $^{13}\text{C}_5$ -O5C-TPAC starting from  $^{13}\text{C}_2$ -acetic acid. (b) Synthesis of  $^{13}\text{C}_5$ -O5C-TPAC starting from  $^{13}\text{C}_2$ -ethylene glycol. (c) Mass spectra of reactions between Fmoc-His-OH and  $^{13}\text{C}_5$ -O5C-TPAC synthesized from  $^{13}\text{C}_2$ -acetic acid (down) and  $^{13}\text{C}_2$ -ethylene glycol (up).



**Figure 4.2** Reactivity of thiophosphorodichloridate reagents on Fmoc-His-OH. (a) Scheme of the labeling of Fmoc-His-OH with thiophosphorodichloridates and CuAAC reaction between the adduct and 3-azido-1-propanol. (b) Labeling yield of thiophosphorodichloridate reagents, and the product distribution after CuAAC reaction. <sup>a</sup> Fmoc-propargylglycine-OH. (c) LC chromatograms of the CuAAC reaction of each thiophosphorodichloridate-histidine adduct. Chemical identities of each chromatographic peak were determined by its *m/z*.

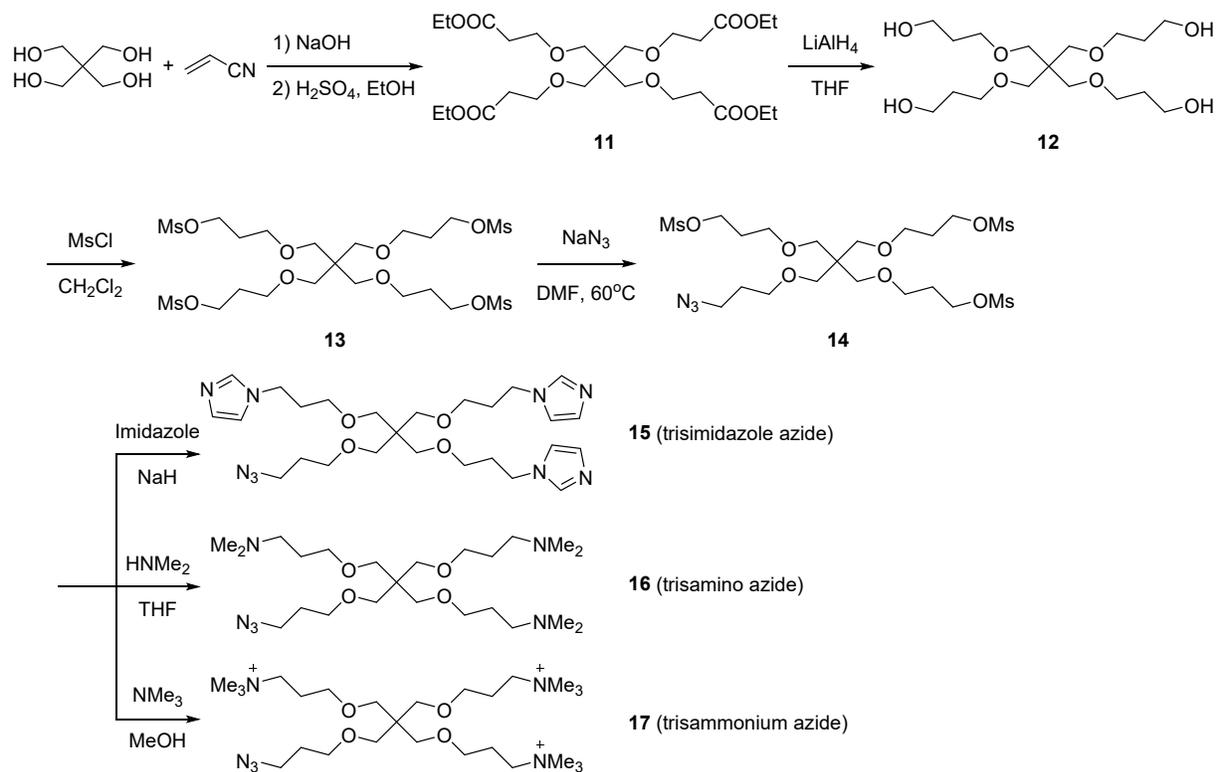


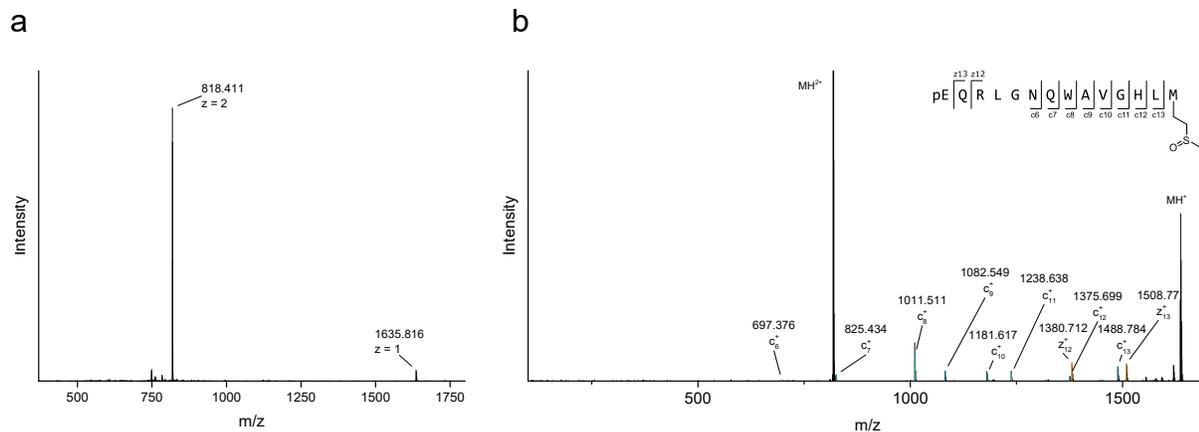
**Figure 4.3** Stability of Fmoc-His-OH labeled with 3C-TPAC, 5C-TPAC or O5C-TPAC at 37°C in 50 mM HEPES, pH 7.5 containing 1 mg/mL HeLa lysate (n = 3, average  $\pm$  s.d.).



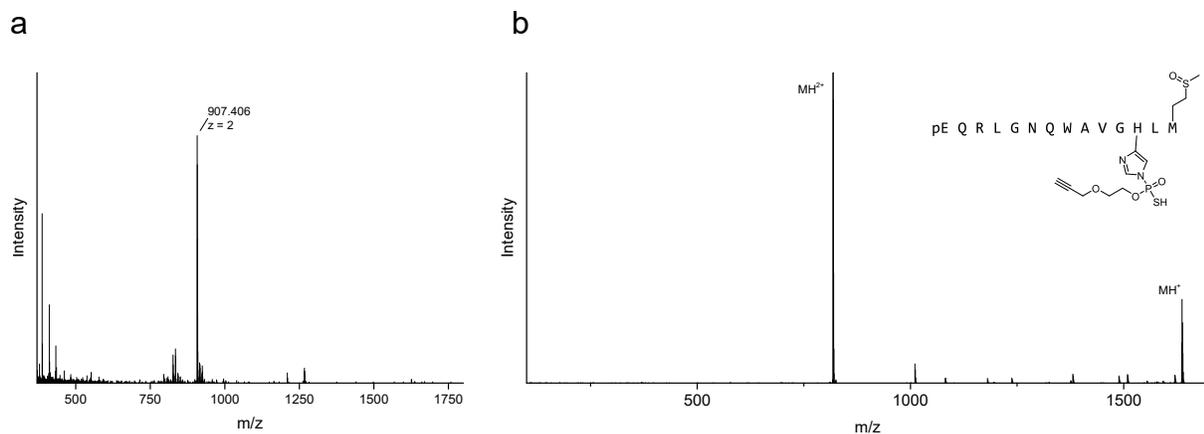
**Figure 4.4** Labeling of HeLa lysate with thiophosphorodichloridate reagents. HeLa lysate (1 mg/mL) in 25 mM HEPES buffer, pH 7.5 was labeled with 500  $\mu$ M of TPAC, clicked with rhodamine azide and analyzed by SDS-PAGE.

### Scheme 4.2 Synthesis of charged azides

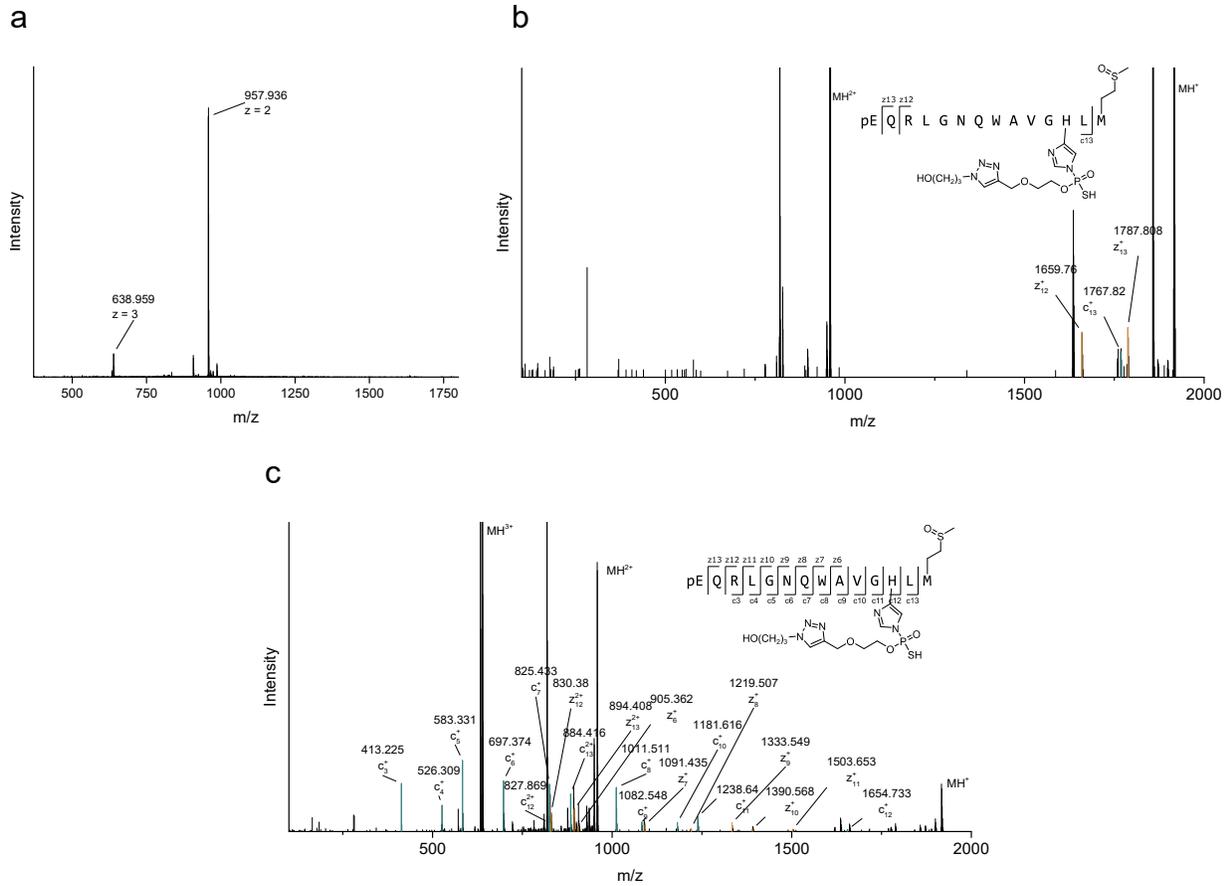




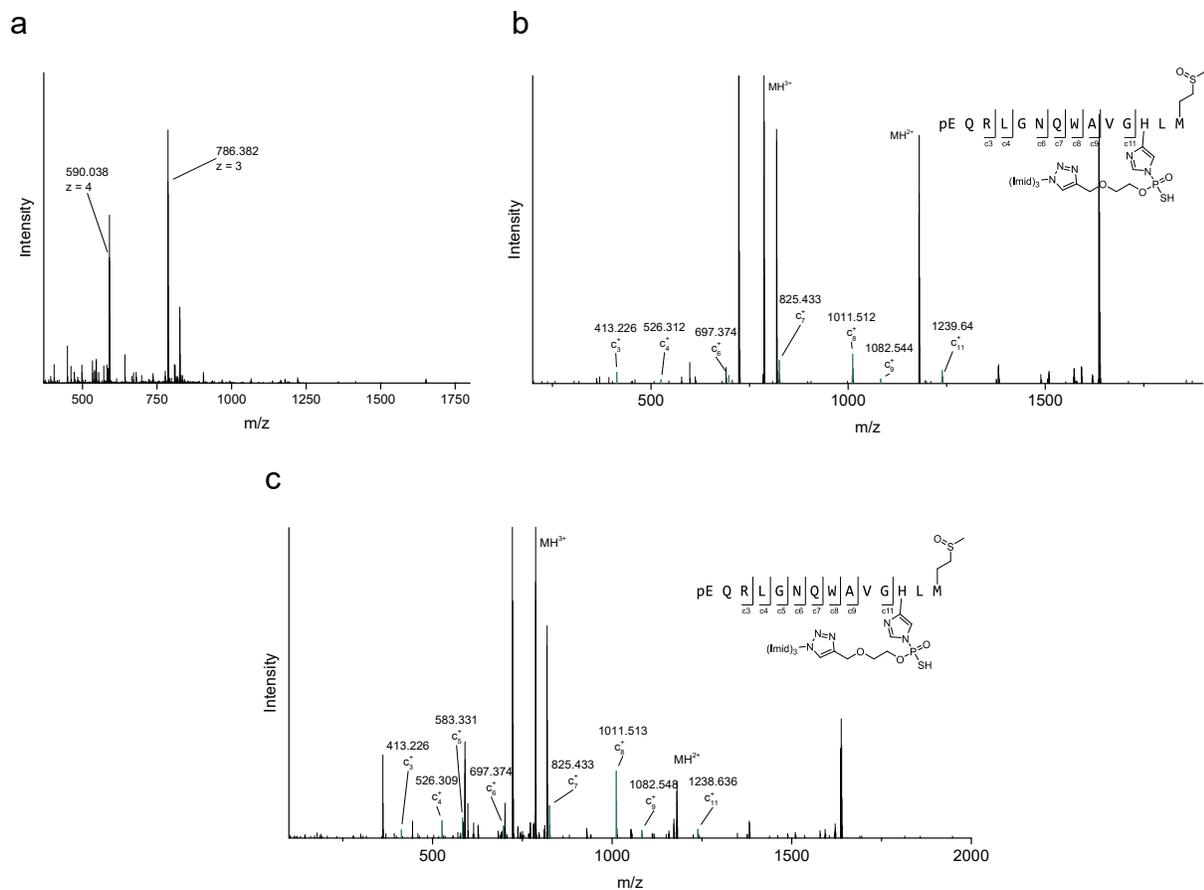
**Figure 4.5** LC-ETD MS/MS of methionine-oxidized bombesin. (a) MS1 spectra of methionine-oxidized bombesin. (b) ETD fragmentation spectrum of methionine-oxidized bombesin ( $z = 2$ ,  $m/z = 818.4$ ).



**Figure 4.6** LC-ETD MS/MS of methionine-oxidized bombesin labeled with O5C-TPAC. (a) MS1 spectra of methionine-oxidized bombesin labeled with O5C-TPAC. (b) ETD fragmentation spectrum of methionine-oxidized bombesin labeled with O5C-TPAC ( $z = 2$ ,  $m/z = 907.4$ ).



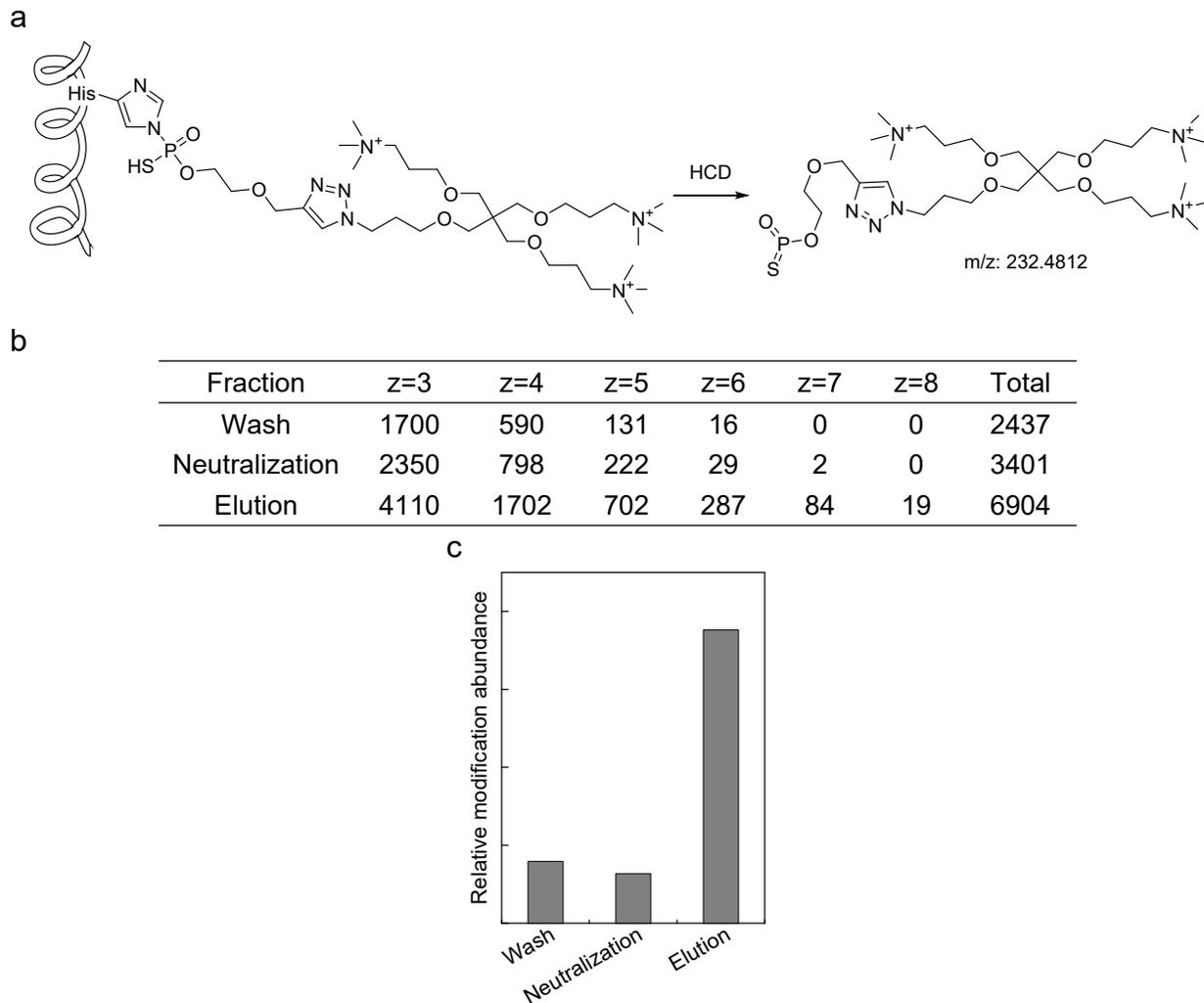
**Figure 4.7** LC-ETD MS/MS of methionine-oxidized bombesin labeled with O5C-TPAC/3-azido-1-propanol. (a) MS1 spectra of methionine-oxidized bombesin labeled with O5C-TPAC/3-azido-1-propanol. (b) ETD fragmentation spectrum of methionine-oxidized bombesin labeled with O5C-TPAC/3-azido-1-propanol ( $z = 2$ ,  $m/z = 957.9$ ). (c) ETD fragmentation spectrum of methionine-oxidized bombesin labeled with O5C-TPAC/3-azido-1-propanol ( $z = 3$ ,  $m/z = 639.0$ ).



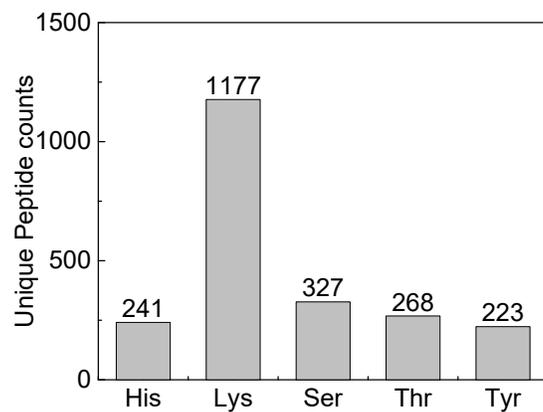
**Figure 4.8** LC-ETD MS/MS of methionine-oxidized bombesin labeled with O5C-TPAC/trisimidazole-azide. (a) MS1 spectra of methionine-oxidized bombesin labeled with O5C-TPAC/trisimidazole-azide. (b) ETD fragmentation spectrum of methionine-oxidized bombesin labeled with O5C-TPAC/trisimidazole-azide ( $z = 3$ ,  $m/z = 786.4$ ). (c) ETD fragmentation spectrum of methionine-oxidized bombesin labeled with O5C-TPAC/trisimidazole-azide ( $z = 4$ ,  $m/z = 590.0$ ).







**Figure 4.11** Enrichment of O5C-TPAC/trisammonium-azide modified peptides by cation exchange chromatography. (a) The “neutral loss” cleavage of O5C-TPAC/trisammonium-azide during HCD fragmentation, resulting in a  $m/z$  232.4,  $z=3$  product ion as a fingerprint. (b) Number of detected precursor ions carrying O5C-TPAC/trisammonium-azide modification by HCD LC-MS/MS. Precursor ion is deemed modified if its fingerprint product ion is higher than 50% intensity of the highest product ion intensity. (c) Relative modification abundance in each fraction, estimated by dividing the number of detected precursor ions with modification over the integrated total ion current.



**Figure 4.12** Number of unique peptides carrying O5C-TPAC/trisammonium modification on each amino acid, as analyzed by LC-ETD MS/MS.

## 4.6 References

- (1) UniProtKB/Swiss-Prot Release 2019\_06 statistics <https://web.expasy.org/docs/relnotes/relstat.html> (accessed Jul 4, 2019).
- (2) Gutteridge, A.; Thornton, J. M. Understanding Nature's Catalytic Toolkit. *Trends Biochem. Sci.* **2005**, *30* (11), 622–629.
- (3) Dokmanić, I.; Šikić, M.; Tomić, S. Metals in Proteins: Correlation between the Metal-Ion Type, Coordination Number and the Amino-Acid Residues Involved in the Coordination. *Acta Crystallogr. D Biol. Crystallogr.* **2008**, *64* (3), 257–263.
- (4) Fuhs, S. R.; Hunter, T. PHisphorylation: The Emergence of Histidine Phosphorylation as a Reversible Regulatory Modification. *Curr. Opin. Cell Biol.* **2017**, *45*, 8–16.
- (5) Klumpp, S.; Krieglstein, J. Reversible Phosphorylation of Histidine Residues in Proteins from Vertebrates. *Sci. Signal.* **2009**, *2* (61), pe13–pe13.
- (6) Verheij, H. M.; Volwerk, J. J.; Jansen, E. H. J. M.; Puyk, W. C.; Dijkstra, B. W.; Drenth, J.; De Haas, G. H. Methylation of Histidine-48 in Pancreatic Phospholipase A2. Role of Histidine and Calcium Ion in the Catalytic Mechanism. *Biochemistry* **1980**, *19* (4), 743–750.
- (7) Johnson, P.; Harris, C. I.; Perry, S. V. 3-Methylhistidine in Actin and Other Muscle Proteins. *Biochem. J.* **1967**, *105* (1), 361–370.
- (8) Lee, J.-W.; Helmann, J. D. The PerR Transcription Factor Senses H<sub>2</sub>O<sub>2</sub> by Metal-Catalysed Histidine Oxidation. *Nature* **2006**, *440* (7082), 363–367.
- (9) Traoré, D. A. K.; El Ghazouani, A.; Jacquamet, L.; Borel, F.; Ferrer, J.-L.; Lascoux, D.; Ravanat, J.-L.; Jaquinod, M.; Blondin, G.; Caux-Thang, C.; et al. Structural and Functional Characterization of 2-Oxo-Histidine in Oxidized PerR Protein. *Nat. Chem. Biol.* **2009**, *5* (1), 53–59.
- (10) Lin, J.-M.; Tsai, Y.-T.; Liu, Y.-H.; Lin, Y.; Tai, H.-C.; Chen, C.-S. Identification of 2-Oxohistidine Interacting Proteins Using E. Coli Proteome Chips. *Mol. Cell. Proteomics* **2016**, *15* (12), 3581–3593.
- (11) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative Reactivity Profiling Predicts Functional Cysteines in Proteomes. *Nature* **2010**, *468* (7325), 790–795.
- (12) Li, X.; Ma, H.; Dong, S.; Duan, X.; Liang, S. Selective Labeling of Histidine by a Designed Fluorescein-Based Probe. *Talanta* **2004**, *62* (2), 367–371.
- (13) Li, X.; Ma, H.; Nie, L.; Sun, M.; Xiong, S. A Novel Fluorescent Probe for Selective Labeling of Histidine. *Anal. Chim. Acta* **2004**, *515* (2), 255–260.
- (14) Cong, Y.; Pawlisz, E.; Bryant, P.; Balan, S.; Laurine, E.; Tommasi, R.; Singh, R.; Dubey, S.; Peciak, K.; Bird, M.; et al. Site-Specific PEGylation at Histidine Tags. *Bioconjug. Chem.* **2012**, *23* (2), 248–263.
- (15) Joshi, P. N.; Rai, V. Single-Site Labeling of Histidine in Proteins, on-Demand Reversibility, and Traceless Metal-Free Protein Purification. *Chem. Commun.* **2019**, *55* (8), 1100–1103.

- (16) Peciak, K.; Laurine, E.; Tommasi, R.; Choi, J.; Brocchini, S. Site-Selective Protein Conjugation at Histidine. *Chem. Sci.* **2019**, *10* (2), 427–439.
- (17) Takaoka, Y.; Tsutsumi, H.; Kasagi, N.; Nakata, E.; Hamachi, I. One-Pot and Sequential Organic Chemistry on an Enzyme Surface to Tether a Fluorescent Probe at the Proximity of the Active Site with Restoring Enzyme Activity. *J. Am. Chem. Soc.* **2006**, *128* (10), 3273–3280.
- (18) Wakabayashi, H.; Miyagawa, M.; Koshi, Y.; Takaoka, Y.; Tsukiji, S.; Hamachi, I. Affinity-Labeling-Based Introduction of a Reactive Handle for Natural Protein Modification. *Chem. – Asian J.* **3** (7), 1134–1139.
- (19) Olszewska, A.; Pohl, R.; Brázdová, M.; Fojta, M.; Hocek, M. Chloroacetamide-Linked Nucleotides and DNA for Cross-Linking with Peptides and Proteins. *Bioconjug. Chem.* **2016**, *27* (9), 2089–2094.
- (20) Adusumalli, S. R.; Rawale, D. G.; Singh, U.; Tripathi, P.; Paul, R.; Kalra, N.; Mishra, R. K.; Shukla, S.; Rai, V. Single-Site Labeling of Native Proteins Enabled by a Chemoselective and Site-Selective Chemical Technology. *J. Am. Chem. Soc.* **2018**, *140* (44), 15114–15123.
- (21) Berry, M. T.; Castrejon, D.; Hein, J. E. Oxidative Esterification of Aldehydes Using Mesoionic 1,2,3-Triazolyl Carbene Organocatalysts. *Org. Lett.* **2014**, *16* (14), 3676–3679.
- (22) Zhang, S.; Moussodia, R.-O.; Sun, H.-J.; Leowanawat, P.; Muncan, A.; Nusbaum, C. D.; Chelling, K. M.; Heiney, P. A.; Klein, M. L.; André, S.; et al. Mimicking Biological Membranes with Programmable Glycan Ligands Self-Assembled from Amphiphilic Janus Glycodendrimers. *Angew. Chem. Int. Ed.* **2014**, *53* (41), 10899–10903.
- (23) Neukom, J. D.; Perch, N. S.; Wolfe, J. P. Intramolecular Alkene Aminopalladation Reactions of (Dppf)Pd(Ar)[N(Ar1)(CH2)3CH=CH2] Complexes. Insertion of Unactivated Alkenes into Pd–N Bonds. *J. Am. Chem. Soc.* **2010**, *132* (18), 6276–6277.
- (24) Higginson, C. J.; Kim, S. Y.; Peláez-Fernández, M.; Fernández-Nieves, A.; Finn, M. G. Modular Degradable Hydrogels Based on Thiol-Reactive Oxanorbornadiene Linkers. *J. Am. Chem. Soc.* **2015**, *137* (15), 4984–4987.
- (25) Newkome, G. R.; Mishra, A.; Moorefield, C. N. Improved Synthesis of an Ethereal Tetraamine Core for Dendrimer Construction. *J. Org. Chem.* **2002**, *67* (11), 3957–3960.
- (26) Kleifeld, O.; Doucet, A.; Prudova, A.; auf dem Keller, U.; Gioia, M.; Kizhakkedathu, J. N.; Overall, C. M. Identifying and Quantifying Proteolytic Events and the Natural N Terminome by Terminal Amine Isotopic Labeling of Substrates. *Nat. Protoc.* **2011**, *6* (10), 1578–1611.
- (27) Li, J.; Yu, J.; Zhao, J.; Wang, J.; Zheng, S.; Lin, S.; Chen, L.; Yang, M.; Jia, S.; Zhang, X.; et al. Palladium-Triggered Deprotection Chemistry for Protein Activation in Living Cells. *Nat. Chem.* **2014**, *6* (4), 352–361.
- (28) Tomás-Gamasa, M.; Martínez-Calvo, M.; Couceiro, J. R.; Mascareñas, J. L. Transition Metal Catalysis in the Mitochondria of Living Cells. *Nat. Commun.* **2016**, *7*, 12538.

- (29) Völker, T.; Dempwolff, F.; Graumann, P. L.; Meggers, E. Progress towards Bioorthogonal Catalysis with Organometallic Compounds. *Angew. Chem. Int. Ed.* **2014**, *53* (39), 10536–10540.
- (30) Potel, C. M.; Lemeer, S.; Heck, A. J. R. Phosphopeptide Fragmentation and Site Localization by Mass Spectrometry: An Update. *Anal. Chem.* **2019**, *91* (1), 126–141.
- (31) Han, X.; Aslanian, A.; Yates, J. R. Mass Spectrometry for Proteomics. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 483–490.
- (32) Riley, N. M.; Coon, J. J. The Role of Electron Transfer Dissociation in Modern Proteomics. *Anal. Chem.* **2018**, *90* (1), 40–64.
- (33) Meyer, J. G.; A Komives, E. Charge State Coalescence during Electrospray Ionization Improves Peptide Identification by Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (8), 1390–1399.
- (34) Frey, B. L.; Lador, D. T.; Sondalle, S. B.; Krusemark, C. J.; Jue, A. L.; Coon, J. J.; Smith, L. M. Chemical Derivatization of Peptide Carboxyl Groups for Highly Efficient Electron Transfer Dissociation. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (11), 1710–1721.
- (35) Bertran-Vicente, J.; Schümman, M.; Hackenberger, C. P. R.; Krause, E. Gas-Phase Rearrangement in Lysine Phosphorylated Peptides During Electron-Transfer Dissociation Tandem Mass Spectrometry. *Anal. Chem.* **2015**, *87* (14), 6990–6994.

# Appendix 1

## **Imaging of cellular hydrogen peroxide by a microtubule-localizing H<sub>2</sub>O<sub>2</sub> fluorescent sensor**

Portions of this work were performed in collaboration with the following person:  
Imaging and cell toxicity experiments were carried out in collaboration with Dr. Shixian Lin.  
Neuron imaging was performed by Corey Webster.

## A1.1 Synopsis

Hydrogen peroxide ( $\text{H}_2\text{O}_2$ ) is a major reactive oxygen species (ROS) that can cause cellular stress when dysregulated, leading to aging and disease states including cancer, inflammation and neurodegeneration.<sup>1-3</sup> The amount  $\text{H}_2\text{O}_2$  is thus tightly regulated in healthy cells by peroxidases, with an estimated steady-state concentration in the submicromolar range.<sup>4</sup> In the meantime, emerging results have shown  $\text{H}_2\text{O}_2$  as an important molecule in cellular signaling pathways.<sup>5,6</sup> Controlled generation of  $\text{H}_2\text{O}_2$  is found in cellular response to growth factors, cytokines and neurotransmitters.<sup>7-9</sup> Fluorescence imaging tools provides a valuable method to elucidate the role of  $\text{H}_2\text{O}_2$  as a secondary messenger in these pathways due to the high spatial and temporal resolution. Our laboratory and others has previously developed a series of  $\text{H}_2\text{O}_2$  fluorescent indicators based on the oxidative cleavage of boronic esters, covering a wide range of color and targeting organelles.<sup>10</sup> However, these probes either diffuse freely inside or across the cytoplasm, or localize onto certain cellular compartments other than the cytoplasm, making it difficult to track the source of  $\text{H}_2\text{O}_2$  generation in these signaling pathways.

We aimed to meet this challenge by localizing our hydrogen peroxide probe onto cytoskeletons to improve its spatial resolution. Inspired by elegant work by Kai Johnsson and co-workers, who coupled silicon-rhodamine fluorophore with microtubule-binding docetaxel to achieve super-resolution imaging of microtubule in living cells,<sup>11</sup> we designed PY1-MT in a similar scaffold to anchor the  $\text{H}_2\text{O}_2$  sensing fluorophore onto the cytoskeleton to prevent free diffusion. The resulting compound retains the fluorescence turn-on response upon  $\text{H}_2\text{O}_2$  treatment in *in vitro* assays, whereas under microscope shows excellent distribution along the microtubule. We then showed that PY1-MT is able to detect  $\text{H}_2\text{O}_2$  level increase from exogenous treatment in HeLa cells while maintaining the localization. PY1-MT is also able to detect the endogenously-generated  $\text{H}_2\text{O}_2$  in EGF-stimulated A431 cells and in lipopolysaccharide-stimulated RAW264.7 cells. We lastly show that the probe is capable of imaging  $\text{H}_2\text{O}_2$  generation in neuron cells upon NMDA treatment. With microtubule as an anchoring point in cytosol, the new  $\text{H}_2\text{O}_2$  probe is a promising candidate for better detection of this analyte with higher spatial resolution. The concept of using cytoskeleton as targeting group to regulate the distribution of fluorescence indicators in the cytoplasm should also be applicable to other analytes of interest.

## A1.2 Methods

### A1.2.1 General synthetic and characterization methods

Unless otherwise noted, all commercial reagents were used without further purification. All reactions utilizing air- or moisture-sensitive reagents were performed in dried glassware under an atmosphere of dry  $\text{N}_2$ . Docetaxel was purchased from Alfa Aesar (Tewksbury, MA). All other reagents were purchased from Sigma-Aldrich.  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra were collected in  $\text{CDCl}_3$ , MeOD or  $d_6$ -acetone (Cambridge Isotope Laboratories, Cambridge MA) at 25 °C on AVB-400, AVQ-400 or AV-600 spectrometers at the College of Chemistry NMR Facility at the University of California, Berkeley. All chemical shifts in  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR

are reported in the standard  $\delta$  notation of ppm relative to residual solvent peak ( $\text{CDCl}_3$   $\delta\text{H} = 7.26$ ,  $\delta\text{C} = 77.16$ ; MeOD  $\delta\text{H} = 3.31$ ,  $\delta\text{C} = 49.00$ ;  $\text{CD}_3\text{CN}$ :  $\delta\text{C} = 118.26$ , acetone:  $\delta\text{H} = 2.05$ ,  $\delta\text{C} = 29.84$ ). Splitting patterns are indicated as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad. Preparative HPLC was performed either on 1260 Infinity LC (Agilent, Santa Clara CA) equipped with a Prep-C18 column (Agilent,  $30 \times 250$  mm,  $10 \mu\text{m}$ ), or on Waters Acquity Autopurification system at LBNL Catalysis Facility at the Lawrence Berkeley National Laboratory (Berkeley Lab) equipped with a semi-preparative C18 column. Low-resolution electrospray mass spectral analyses were carried out using a LC-MS (Agilent Technology 6130, Quadrupole LC/MS). High resolution mass spectral analyses (ESI-MS) were carried out at the College of Chemistry Mass Spectrometry Facility at the University of California, Berkeley.

### 41.2.2 Synthesis of PYI-MT

**Fmoc-piperazine rhodol boronate (1)** Known compound (1) was prepared according to literature procedure.<sup>12</sup>

**3'-Aminodocetaxel formate salt (2)** Docetaxel (105 mg, 0.13 mmol) was dissolved in formic acid (4 mL) at room temperature. The clear solution was stirred for 30 min and formic acid was removed by rotary evaporation. The residue was further dried on a vacuum pump for 3 h to provide compound (2) as a white solid. The crude product was dissolved in DMSO (2.6 mL) as a 50 mM stock solution and was used without further purification.

**3'-(5-Carboxyheptanoylamino)docetaxel (3)** Suberic acid (87 mg, 0.5 mmol) was dissolved in DMSO (2.5 mL) in a 4 mL vial. *N,N*-Diisopropylethylamine (DIPEA, 226  $\mu\text{L}$ , 1.3 mmol) and *N,N,N',N'*-tetramethyl-*O*-(*N*-succinimidyl)uronium tetrafluoroborate (TSTU, 39 mg, 0.13 mmol) were added. The mixture was stirred for 5 min before the addition of compound 2 (2 mL of 50 mM solution in DMSO, 0.1 mmol). The mixture was stirred for 3h at room temperature and separated by preparative HPLC (40% – 60% MeCN in 40 min) followed by semi-preparative HPLC. The fractions were lyophilized to give compound 3 as a white powder (23 mg, 27% over two steps).  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{CN}$ )  $\delta$  6.79 (d,  $J = 7.4$  Hz, 2H), 6.35 (t,  $J = 7.4$  Hz, 1H), 6.25 (t,  $J = 7.6$  Hz, 2H), 6.11 – 6.06 (m, 4H), 6.02 – 5.96 (m, 1H), 4.76 (t,  $J = 8.8$  Hz, 1H), 4.28 (d,  $J = 7.2$  Hz, 1H), 4.17 – 4.10 (m, 1H), 3.88 (s, 1H), 3.66 (d,  $J = 8.1$  Hz, 1H), 3.28 (d,  $J = 3.8$  Hz, 1H), 2.89 – 2.81 (m, 3H), 2.51 (d,  $J = 7.1$  Hz, 1H), 1.11 (ddd,  $J = 15.9, 9.6, 6.6$  Hz, 1H), 1.03 (s, 3H), 0.96 – 0.85 (m, 5H), 0.68 – 0.61 (m, 2H), 0.57 (s, 3H), 0.49 – 0.40 (m, 1H), 0.38 – 0.31 (m, 3H), 0.26 – 0.18 (m, 4H), -0.02 (d,  $J = 6.9$  Hz, 6H), -0.16 (s, 3H), -0.25 (s, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CD}_3\text{CN}$ )  $\delta$  211.39, 174.68, 173.69, 171.34, 166.91, 139.90, 138.72, 137.29, 134.26, 130.97, 130.81, 129.49, 129.34, 128.47, 127.95, 85.03, 81.67, 78.53, 77.02, 75.86, 75.06, 74.37, 72.34, 72.15, 58.35, 55.99, 47.24, 44.00, 37.22, 36.56, 34.41, 29.36, 29.31, 26.97, 26.38, 25.41, 22.94, 21.40, 14.29, 10.31. LRMS (ESI<sup>+</sup>)  $m/z$  calcd 864.4, found 864.6 for  $\text{C}_{46}\text{H}_{58}\text{NO}_{15}^+$  (M+H<sup>+</sup>).

**PY1-MT (4)** Fmoc-piperazine rhodol boronate (**1**, 5.3 mg, 7.2  $\mu\text{mol}$ ) was dissolved in  $\text{CH}_2\text{Cl}_2$  (1 mL) in a 1.5. mL microcentrifuge tube. Tris(2-aminoethyl)amine (TAEA, 54  $\mu\text{L}$ , 360  $\mu\text{mol}$ ) was added to obtain a yellow solution. The mixture was mixed on a revolver for 3 h in darkness. The organic phase was extracted with phosphate buffer to remove the residual TAEA (2M pH 5.5,  $3 \times 750 \mu\text{L}$ ). The resulting light pink solution was evaporated to dryness and then dissolved in DMF (150  $\mu\text{L}$ ). In a separate microcentrifuge tube, 3'-(5-Carboxyheptanoylamino)docetaxel (**3**, 5.2 mg, 6.0  $\mu\text{mol}$ ) was dissolved in DMF (500  $\mu\text{L}$ ) and treated with TSTU (2.2 mg, 7.2  $\mu\text{mol}$ ) and DIPEA (23  $\mu\text{L}$ , 133  $\mu\text{mol}$ ) for 5 min. The two solutions were combined and placed on a revolver for 3 h in darkness. The mixture was then separated by semi-preparative HPLC. The fractions were lyophilized to provide compound **4** as a light pink powder (2 mg, 25%).  $^1\text{H}$  NMR (400 MHz, MeOD)  $\delta$  8.10 (d,  $J = 7.2$  Hz, 2H), 8.02 (t,  $J = 6.8$  Hz, 1H), 7.74 (dt,  $J = 13.8, 7.3$  Hz, 2H), 7.67 – 7.61 (m, 2H), 7.55 (t,  $J = 7.6$  Hz, 2H), 7.45 – 7.36 (m, 5H), 7.27 (t,  $J = 6.9$  Hz, 1H), 7.19 (d,  $J = 7.2$  Hz, 1H), 6.86 (dd,  $J = 7.6, 2.1$  Hz, 1H), 6.76 (dd,  $J = 16.2, 8.5$  Hz, 2H), 6.69 – 6.62 (m, 1H), 6.15 (t,  $J = 8.7$  Hz, 1H), 5.64 (d,  $J = 7.1$  Hz, 1H), 5.47 (d,  $J = 4.5$  Hz, 1H), 5.26 (s, 1H), 4.99 (d,  $J = 9.5$  Hz, 1H), 4.58 (d,  $J = 4.5$  Hz, 1H), 4.26 – 4.14 (m, 3H), 3.86 (d,  $J = 7.2$  Hz, 1H), 3.74 – 3.64 (m, 4H), 2.41 (dd,  $J = 14.7, 7.1$  Hz, 3H), 2.36 – 2.18 (m, 6H), 2.01 (dd,  $J = 15.5, 9.0$  Hz, 1H), 1.89 (s, 3H), 1.86 – 1.78 (m, 1H), 1.69 (s, 3H), 1.61 (d,  $J = 6.6$  Hz, 4H), 1.36 (s, 12H), 1.20 (s, 2H), 1.18 (s, 3H), 1.12 (s, 3H). HRMS (ESI<sup>+</sup>)  $m/z$  calcd 1356.6022, found 1356.6066 for  $\text{C}_{76}\text{H}_{87}\text{BN}_3\text{O}_{19}^+$  (M+H<sup>+</sup>).

### 4.1.2.3 Spectroscopic materials and methods

Samples for *in vitro* measurements were prepared in 25 mM HEPES, pH 7.4 in 1-cm  $\times$  1-cm quartz cuvettes (1.4-mL volume, Starna). Emission spectra of samples under treatments were recorded using a Photon Technology International Quanta Master 4 L-format scan spectrofluorometer equipped with an LPS-220B 75-W xenon lamp and power supply, A-1010B lamp housing with integrated igniter, switchable 814 photocounting/analog photomultiplier detection unit and MD5020 motor driver.

For fluorescence turn-on measurements, 5  $\mu\text{M}$  PY1-MT was prepared in 1 mL of 25 mM HEPES, pH 7.4 from its 5 mM stock in DMSO, and treated with 100  $\mu\text{M}$   $\text{H}_2\text{O}_2$  (1000 $\times$  in  $\text{H}_2\text{O}$ ). The solution was incubated at 37  $^\circ\text{C}$  water bath and the emission spectrum was measured immediately after for indicated period of time.

Selectivity test of PY1-MT were performed in gas-tight cuvettes. *Tert*-butyl hydroperoxide (*t*-BuOOH), sodium hypochlorite (NaClO) and sodium nitrite ( $\text{NaNO}_2$ ) were diluted from stock solution prepared from commercial reagents. Other competing species were prepared as follows:

- Hydroxyl radical ( $\text{OH}\cdot$ ) and *tert*-butyloxy radical (*t*-BuOO $\cdot$ ) were generated *in situ* by mixing ammonium iron(II) sulfate (1mM, diluted from 100 $\times$  stock solution in  $\text{N}_2$ -purged  $\text{H}_2\text{O}$ ) and  $\text{H}_2\text{O}_2$  (100  $\mu\text{M}$ , 1000 $\times$  in  $\text{H}_2\text{O}$ ) or *t*-BuOOH (100  $\mu\text{M}$ , 1000 $\times$  in  $\text{H}_2\text{O}$ ) in the presence of superoxide dismutase (50 U/mL).

- Superoxide ( $\text{O}_2\cdot^-$ ) was generated *in situ* by xanthine oxidase, as described in Sigma-Aldrich's technical document *Enzymatic Assay of Superoxide Dismutase*. The catalytic

activity of xanthine oxidase was determined by cytochrome C reduction before test, and the amount of xanthine oxidase added to the reaction solution was adjusted to provide 1.7  $\mu\text{mol}/(\text{L}\cdot\text{min})$  to for 1h reaction.

- Nitric oxide (NO) was diluted from its solution in cooled 25 mM HEPES, pH 7.4 buffer.  $\text{NaNO}_2$  (2g) was dissolved in 4 mL  $\text{H}_2\text{O}$  and added via syringe to 30%  $\text{H}_2\text{SO}_4$  (2 mL) in a two-neck flask. The evolved gas was passed through 20% NaOH to remove residual acid and absorbed in  $\text{N}_2$ -purged buffer cooled in an ice/water bath. For concentration determination, 50  $\mu\text{L}$  of the solution was diluted into 950  $\mu\text{L}$  of 100 mM phosphate buffer, pH 7.4 containing 17 mM sulfanilamide and 0.4 mM *N*-(1-naphthyl)ethylenediamine under  $\text{N}_2$ . After thorough mixing the concentration was read out by absorption at 496 nm ( $\epsilon = 6600 \text{ M}^{-1}\text{cm}^{-1}$ ).<sup>13</sup>

- Peroxynitrite ( $\text{ONOO}^-$ ) was prepared from the reaction between  $\text{H}_2\text{O}_2$ , isoamyl nitrite and NaOH, according to reported procedure. The concentration of peroxynitrite was determined by its absorption at 302 nm ( $\epsilon = 6600 \text{ M}^{-1}\text{cm}^{-1}$ ).<sup>14</sup>

#### *A1.2.4 Cell culture and fluorescence imaging*

HeLa, A431 and RAW 264.7 cells were grown in the Cell Culture Facility at University of California, Berkeley. One day before imaging, cells were passed and plated on eight-well chamber slides (Lab-Tek, Thermo Fisher).

For imaging, cells were grown on 8-well chamber slides (LabTek, Thermo Fisher) to desired confluency, incubated with 1  $\mu\text{M}$  PY1-MT for 15-30 min and then treated with stimulant. Confocal microscopy was performed with a Zeiss LSM710 laser-scanning microscope using a 40 $\times$  water-immersion or 63 $\times$  oil-immersion objective lens. Excitation was provided at 405 nm for Hoechst 33342 and 488 nm for PY1-MT.

### **A1.3 Results and discussion**

#### *A1.3.1 Design and synthesis of microtubule-localizing $\text{H}_2\text{O}_2$ probe*

In analogous to Kai Johnsson's work on microtubule-localizing silicon-rhodol probe (Figure A1.1a), we decided to link the  $\text{H}_2\text{O}_2$ -responsive moiety to the primary amine of Boc-protected docetaxel. Our laboratory has previously published MitoPY1, a mitochondria-localizing  $\text{H}_2\text{O}_2$  sensor that connects a triphenylphosphonium moiety as a targeting group to the boronic ester-caged fluorescein (Figure A1.1b). Inspired by both works, we decided to take the same  $\text{H}_2\text{O}_2$  sensing motif in MitoPY1, and conjugate its secondary amine on the piperazine with the primary amine on Boc-protected docetaxel by suberic acid to form peroxyyellow-microtubule (PY1-MT, Figure A1.1c). The synthesis of this molecule shared the same initial steps with mito-PY1. However, the final amide formation step with TSTU was incompatible with in the presence of dibenzofulvene as a byproduct of Fmoc-deprotection. Such byproduct was removed by using TAEA as both deprotection reagent as well as scavenger,<sup>15</sup> which allows for easy purification of the  $\text{H}_2\text{O}_2$  sensing moiety with the secondary amine (Scheme A1.1).

### *A1.3.2 In vitro fluorescence turn-on of PY1-MT*

With the compound in hand, we first tested its turn-on in *in vitro* experiments. We treated and incubated the probe at 37°C at neutral pH for 1h and observed negligible fluorescence turn-on. In contrast, when we incubated the probe in the presence of 100  $\mu\text{M}$   $\text{H}_2\text{O}_2$  we were able to observe a 12-fold turn-on response after one-hour treatment (Figure A1.2a), which is comparable with our previously published boronic ester-based  $\text{H}_2\text{O}_2$  sensors. This result suggests that the introduction of docetaxel as the localizing moiety does not interfere with the interaction between  $\text{H}_2\text{O}_2$  and its sensing moiety, and that the probe should be suitable for use in cell imaging experiments with the same protocol.

We further tested its response to other competing reactive oxygen (ROS) and nitrogen (RNS) species (Figure A1.2b). The result shows that PY1-MT reacts primarily with  $\text{H}_2\text{O}_2$ , which is in agreement with other published boronic-based  $\text{H}_2\text{O}_2$  probes. The only interfering species we observed is peroxyxynitrite ( $\text{ONOO}^-$ ). Considering that this highly reactive species exist in cells with estimated concentration less than one nanomolar,<sup>16</sup> whereas  $\text{H}_2\text{O}_2$  concentration can reach up to micromolar range,<sup>4</sup> peroxyxynitrite is less likely to cause problems in actual cell imaging experiments. Nevertheless, this finding suggests that extra control experiment is necessary to rule out the possible effect of biological peroxyxynitrite.

### *A1.3.3 Localization and turn-on of PY1-MT in cells imaging*

We then investigated the behavior of PY1-MT in cell imaging experiments, first by checking whether the probe is able to localize onto microtubule. As we expected, PY1-MT showed even distribution in cytosol, forming a network of filaments that resembles the cytoskeleton (Figure A1.3a). A dividing cell also displayed a structure with the shape of a spindle apparatus (Figure A1.3b). We further performed transfection of the cells with mCherry-tubulin that gives rise to native microtubules labeled with red fluorescent mCherry protein. After incubating these cells with PY1-MT, we observe colocalization between PY1-MT in the green channel and mCherry-labeled tubulin in the red channel with a Pearson correlation coefficient of 0.65 (Figure A1.3c). These imaging results indicate that in PY1-MT the taxane moiety behaves as expected to assemble itself onto microtubule, which brings about the microtubule-localizing property of this fluorescent sensor.

We then sought to test its turn on response upon  $\text{H}_2\text{O}_2$  incubation. Indeed, after loading the cells with PY1-MT, incubation with 100  $\mu\text{M}$   $\text{H}_2\text{O}_2$  resulted in significant fluorescence enhancement inside the cells (Figure A1.4). More importantly, the probe retains its microtubule-localizing pattern after its turn-on with  $\text{H}_2\text{O}_2$ , which suggests that  $\text{H}_2\text{O}_2$  reacts with the fluorophore end rather than the microtubule-targeting docetaxel moiety. Collectively, HeLa cells imaging experiments suggest that PY1-MT meets our expectation as microtubule-localizing  $\text{H}_2\text{O}_2$  fluorescent indicator, and this anchoring effect provides higher spatial resolution for detection of  $\text{H}_2\text{O}_2$  in cytoplasm.

#### A1.3.4. Cell toxicity of PY1-MT

For fluorescence imaging experiments, ideally the fluorescent indicator should not interfere with cellular processes to probe the cell in its native state; the bottom line for successful experiment is that the probe should be no-toxic. Docetaxel, as a member of the taxane family, inserts into the microtubule and prevents its depolymerization. This process results in strong cytotoxicity with an  $IC_{50}$  of 1.87 nM.<sup>17</sup> To test whether PY1-MT as a docetaxel derivative shows strong cytotoxicity that prevents its practical usage as a fluorescent sensor, we performed MTT assay on HeLa cells (Figure A1.5). The result suggests that there is no observable toxic effect with PY1-MT concentration up to 4  $\mu$ M and incubation time up to 24 h. Similar results were obtained with A431, RAW 264.7 and neuron cells. This greatly reduced toxicity compared to docetaxel is in agreement with the tubulin localizing probe developed by Kai Jonsson and co-workers.<sup>11</sup> In comparison, for a typical imaging experiment we used 1  $\mu$ M PY1-MT incubation and a time period shorter than 2 h, which suggests that PY1-MT exhibits negligible toxicity in imaging experiments and can be safely used for  $H_2O_2$  imaging in cells.

#### A1.3.5 Fluorescence imaging of endogenously generated $H_2O_2$

After establishing that PY1-MT is suitable to detect  $H_2O_2$  in cells with high spatial resolution, we went on to test whether this sensor could respond to endogenously-generated  $H_2O_2$ . We first tried to detect  $H_2O_2$  generated during epidermal growth factor (EGF) signaling pathway. Against this backdrop, we used A431 cells as our model, which is a cancer cell line with EGF receptor overexpressed. When we treated PY1-MT loaded A431 cells with EGF (100 ng/mL), a significant turn-on of PY1-MT was observed (Figure A1.6a,b). In contrast, when we treated the cells with both EGF stimulation and L-NAME, a nitric oxide synthase inhibitor that can reduce reactive nitrogen species generation in cells, we observed a similar fluorescence intensity compared to cells treated with EGF only (Figure A1.6c). This indicates that peroxynitrite, which showed higher reactivity in *in vitro* assays, did not participate in the turn-on of this  $H_2O_2$  probe. Moreover, when we treated cells with EGF and PD15305, an EGF receptor inhibitor, we observed similar fluorescence similar with the untreated cells (Figure A1.6d), further confirming that the turn-on effect was indeed resulted from  $H_2O_2$  generated in the EGF signaling pathway. In addition, higher-magnification images showed filament-localization pattern of this sensor, correlating with the microtubule-localizing property of this probe that give rise to even and spatially-resolved distribution in cytosol (Figure A1.6e).

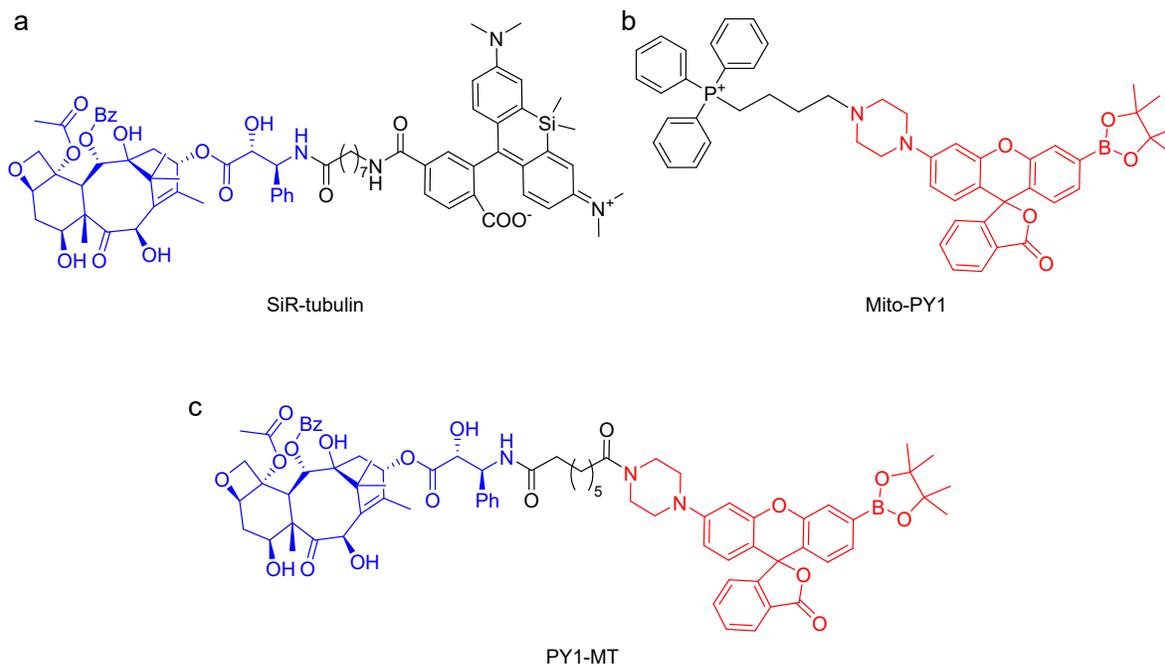
$H_2O_2$  is also produced in macrophage upon stimulation as an inflammation response to invading pathogens. To see whether PY1-MT can also detect  $H_2O_2$  generated in this process, we incubated RAW 264.7 cells with our sensor, and stimulated the cells with lipopolysaccharide (LPS) and phorbol 12-myristate-13-acetate (PMA). As we expected, stimulated cells showed approximately three-fold turn-on compared to the control group (Figure A1.7). Interestingly, different from evenly stained cells without treatment, the stimulated cells showed bright spots in the cytosol, which could be potential sites of  $H_2O_2$  generation. This exemplifies the merit of anchoring the sensor onto the cytoskeleton, in that the sensor no longer diffuses freely and thus becomes capable of detecting  $H_2O_2$  with higher spatial resolution.

Finally, we sought to investigate the possibility of imaging neurons using PY1-MT. ROS, including  $H_2O_2$ , is an important messenger in neuron signaling transduction and is under intense investigation. In particular, upon activation, *N*-Methyl-D-aspartate (NMDA) receptor can cause calcium influx to trigger a cascade of signaling events that lead to generation of ROS.<sup>18</sup> We thus tried to monitor  $H_2O_2$  increase in neurons upon NMDA activation. The results suggest that PY1-MT stained evenly inside the neuron, and upon NMDA treatment, the fluorescence increased considerably in the cell body (Figure A1.8a). Moreover, simultaneous treatment of NMDA and L-NAME only reduced the fluorescence to a small extent, suggesting that the turn-on of PY1-MT originated from  $H_2O_2$  rather than other reactive nitrogen species (Figure A1.8b). With the higher spatial resolution of cytoskeleton-immobilized PY1-MT, it should be useful in  $H_2O_2$  imaging in neurons in future experiments, especially in applications focusing on the ROS generation site and also ROS transferring between cells.

#### **A1.4 Conclusion**

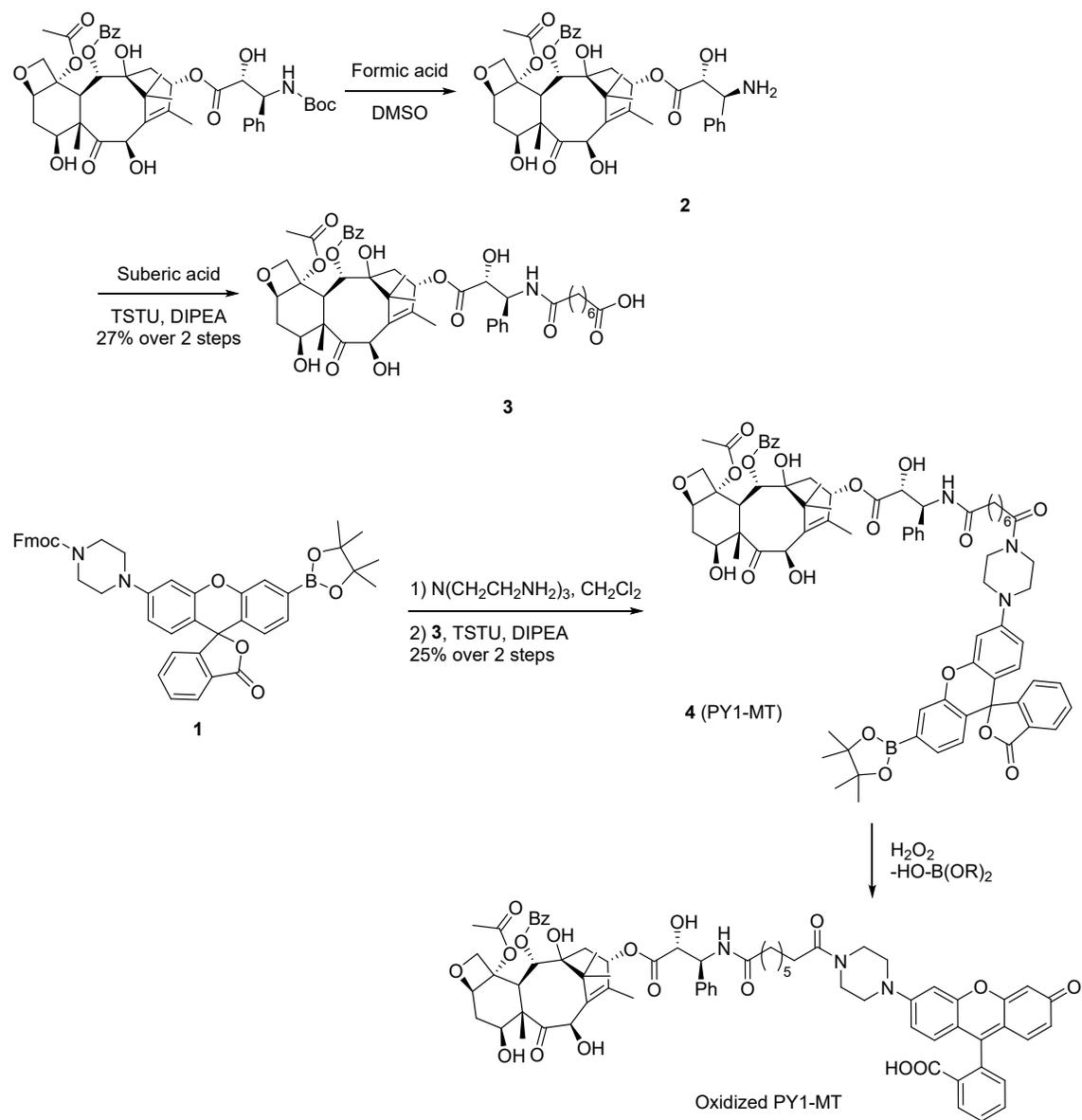
In this appendix, we have combined the fluorescent  $H_2O_2$  sensing activity with the microtubule-localizing property to obtain PY1-MT as a useful  $H_2O_2$  sensor with enhanced spatial resolution. The docetaxel targeting moiety and the  $H_2O_2$  sensing component behave independently in imaging experiments, resulting in this microtubule-anchored  $H_2O_2$  responsive probe distributed evenly in cytoplasm. We have shown that PY1-MT can successfully detects endogenously-produced  $H_2O_2$  in EGF, inflammation and neuron signaling pathways. We envision that this new generation of sensor can help to spatially-resolve the generation of  $H_2O_2$  in complex biology problems. This cytoskeleton-localization method can also be easily expanded to various fluorescent sensors for other analytes to achieve even distribution in cytosol with high spatial resolution in fluorescence imaging.

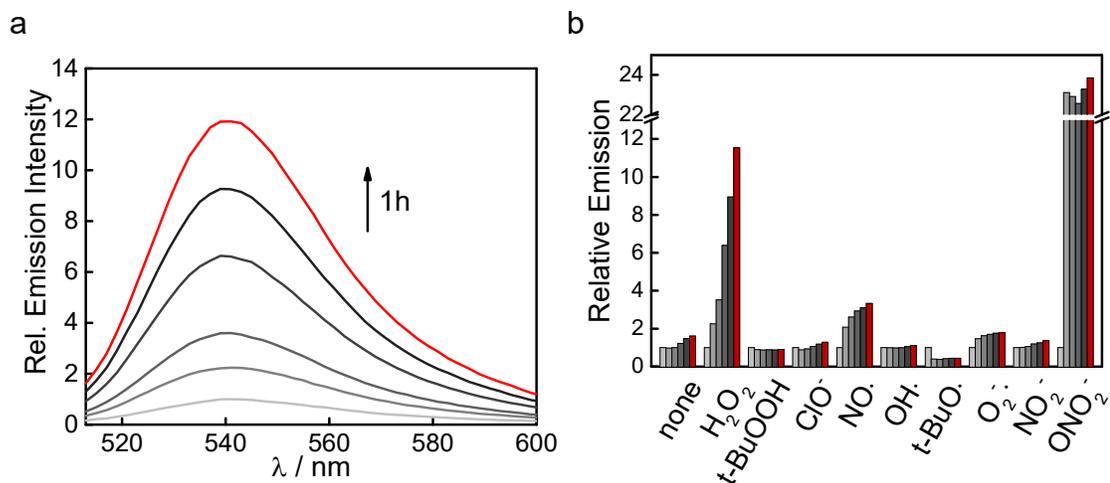
## A1.5 Figures and schemes



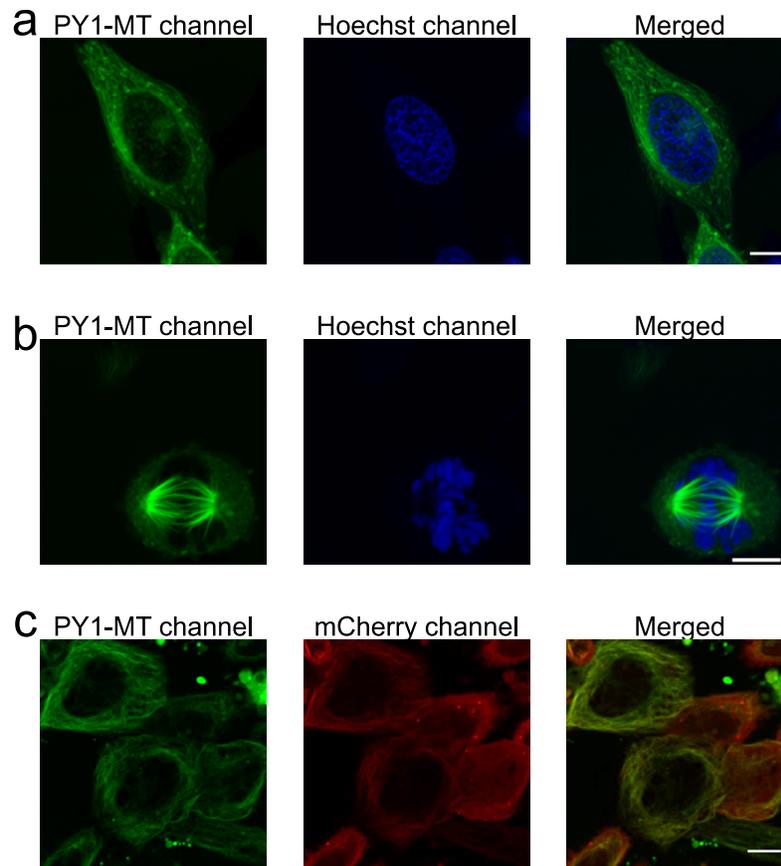
**Figure A1.1** Structures of (a) microtubule-localizing fluorescent dye, SiR-tubulin; (b) mitochondria-localizing  $\text{H}_2\text{O}_2$  sensor, Mito-PY1, and (c) microtubule-localizing  $\text{H}_2\text{O}_2$  sensor, PY1-MT.

**Scheme A1.1** Synthesis of PY1-MT and its fluorescence turn-on reaction upon  $H_2O_2$  treatment.

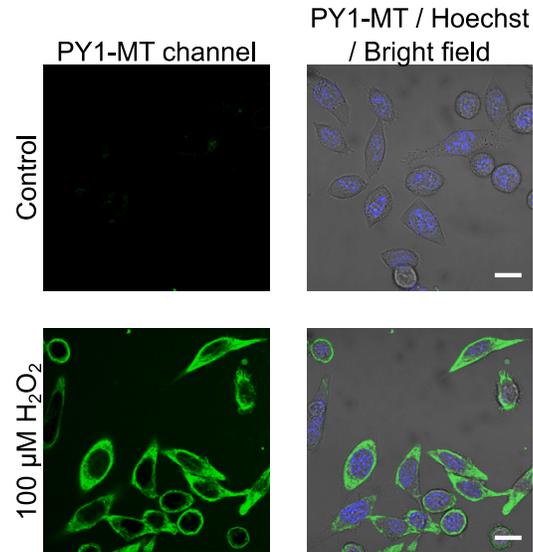




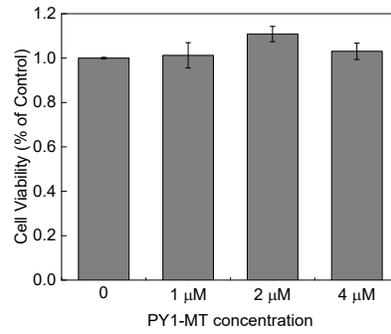
**Figure A1.2** *In vitro* characterization of PY1-MT. (a) Fluorescence emission spectrum of 5  $\mu\text{M}$  PY1-MT in 25 mM HEPES buffer pH 7.4 (bottom), and its turn-on response after treatment with 100  $\mu\text{M}$   $\text{H}_2\text{O}_2$  at 37°C for 5, 15, 30, 45 or 60 min. (b) Fluorescence response of 5  $\mu\text{M}$  PY1-MT to 100  $\mu\text{M}$  of various reactive oxygen and nitrogen species at 37°C after 5, 15, 30, 45 or 60 min of incubation.



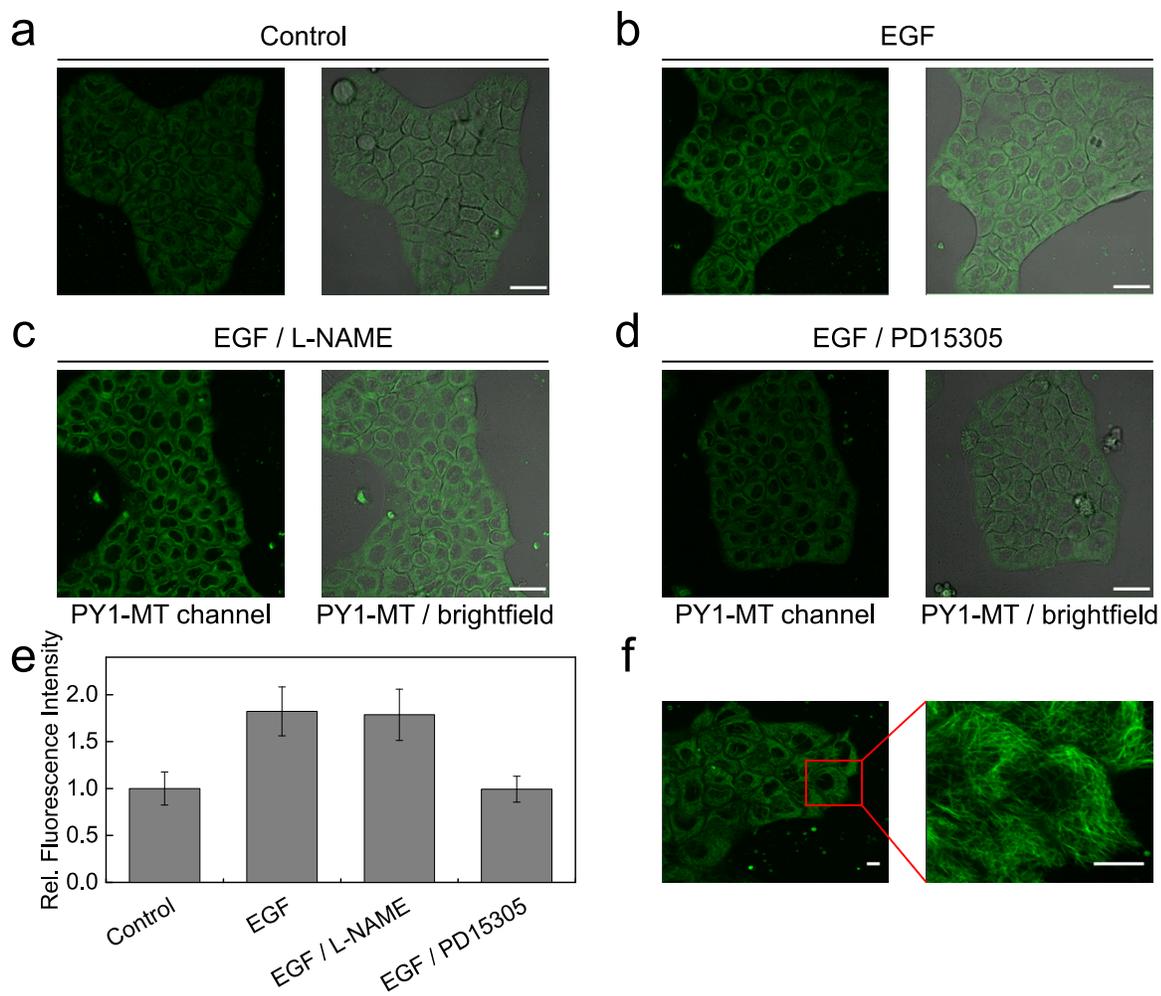
**Figure A1.3** Fluorescence confocal of PY1-MT localization in HeLa cells. (a) A HeLa cell in metaphase. (b) A HeLa cell in interphase. (c) HeLa cells expressing mCherry-tubulin were stained with PY1-MT. Scale-bars: 10 μm.



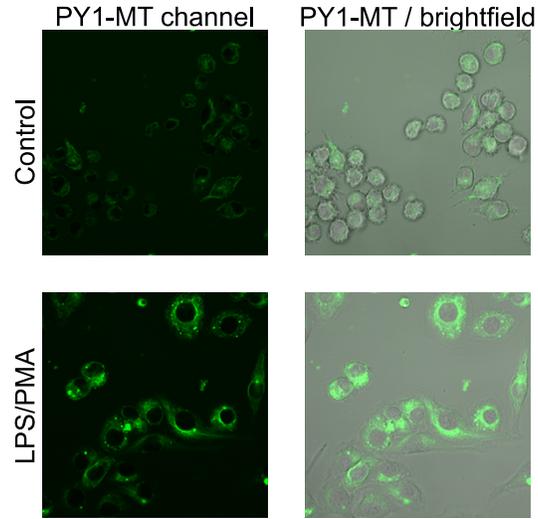
**Figure A1.4** Fluorescence turn-on of PY1-MT in HeLa cells. Cells were incubated with 1  $\mu\text{M}$  PY1-MT, followed by treatment with vehicle control or 100  $\mu\text{M}$   $\text{H}_2\text{O}_2$  at 37°C for 0.5 h. Scale-bar: 20  $\mu\text{m}$ .



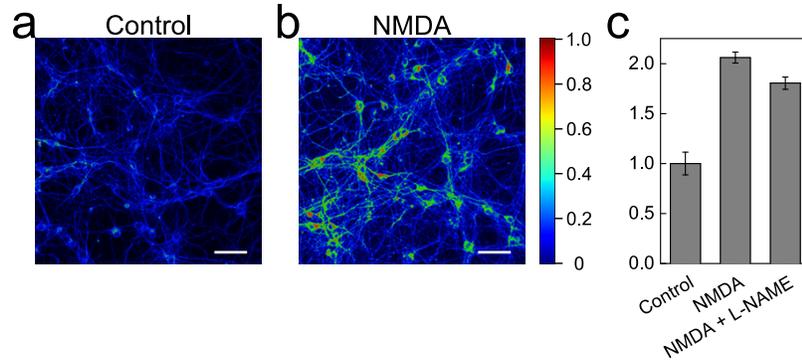
**Figure A1.5** Effects of PY1-MT on cell viability. MTT assays were performed on HeLa cells treated with 0, 1, 2 or 4  $\mu\text{M}$  PY1-MT for 24 h. Data are shown as average  $\pm$  s.d.



**Figure A1.6** Fluorescence imaging of  $H_2O_2$  generated in EGF signaling pathway in A431 cells. A431 cells loaded with PY1-MT were treated with (a) vehicle control, (b) EGF, (c) EGF and L-NAME or (d) EGF and PD15305 and imaged. Fluorescence intensity was quantified and shown in (e) as average  $\pm$  s.d. Scale-bar: 50  $\mu$ m. (f) Zoomed-in images of A431 stained with PY1-MT showing the high spatial resolution of this fluorescence probe.



**Figure A1.7** Detection of  $\text{H}_2\text{O}_2$  production in RAW 264.7 cells upon stimulation. Cells loaded with PY1-MT were treated with vehicle control (up) or LPS and PMA (bottom) and imaged by confocal microscopy.



**Figure A1.8** Detection of H<sub>2</sub>O<sub>2</sub> in hippocampal neurons upon stimulation. Neurons loaded with PY1-MT were incubated with (a) vehicle control or (b) 5  $\mu$ M NMDA and imaged by confocal microscopy. Scale-bars: 20  $\mu$ m. (c) Quantification of fluorescence intensity of control cells, NMDA-treated cells and NMDA/L-NAME-treated cells, shown as average  $\pm$  s.d.

## A1.6 References

- (1) Lin, M. T.; Beal, M. F. Mitochondrial Dysfunction and Oxidative Stress in Neurodegenerative Diseases. *Nature* **2006**, *443* (7113), 787.
- (2) Finkel, T.; Serrano, M.; Blasco, M. A. The Common Biology of Cancer and Ageing. *Nature* **2007**, *448* (7155), 767–774.
- (3) Lisanti, M. P.; Martinez-Outschoorn, U. E.; Lin, Z.; Pavlides, S.; Whitaker-Menezes, D.; Pestell, R. G.; Howell, A.; Sotgia, F. Hydrogen Peroxide Fuels Aging, Inflammation, Cancer Metabolism and Metastasis. *Cell Cycle* **2011**, *10* (15), 2440–2449.
- (4) Stone, J. R.; Yang, S. Hydrogen Peroxide: A Signaling Messenger. *Antioxid. Redox Signal.* **2006**, *8* (3–4), 243–270.
- (5) Finkel, T. Signal Transduction by Reactive Oxygen Species. *J. Cell Biol.* **2011**, *194* (1), 7–15.
- (6) Dickinson, B. C.; Chang, C. J. Chemistry and Biology of Reactive Oxygen Species in Signaling or Stress Responses. *Nat. Chem. Biol.* **2011**, *7* (8), 504–511.
- (7) Sundaresan, M.; Yu, Z.-X.; Ferrans, V. J.; Irani, K.; Finkel, T. Requirement for Generation of H<sub>2</sub>O<sub>2</sub> for Platelet-Derived Growth Factor Signal Transduction. *Science* **1995**, *270* (5234), 296–299.
- (8) Avshalumov, M. V.; Chen, B. T.; Marshall, S. P.; Peña, D. M.; Rice, M. E. Glutamate-Dependent Inhibition of Dopamine Release in Striatum Is Mediated by a New Diffusible Messenger, H<sub>2</sub>O<sub>2</sub>. *J. Neurosci.* **2003**, *23* (7), 2744–2750.
- (9) Niethammer, P.; Grabher, C.; Look, A. T.; Mitchison, T. J. A Tissue-Scale Gradient of Hydrogen Peroxide Mediates Rapid Wound Detection in Zebrafish. *Nature* **2009**, *459* (7249), 996–999.
- (10) Lin, V. S.; Dickinson, B. C.; Chang, C. J. Chapter Two - Boronate-Based Fluorescent Probes: Imaging Hydrogen Peroxide in Living Systems. In *Methods in Enzymology*; Cadenas, E., Packer, L., Eds.; Hydrogen Peroxide and Cell Signaling, Part A; Academic Press, 2013; Vol. 526, pp 19–43.
- (11) Lukinavičius, G.; Reymond, L.; D’Este, E.; Masharina, A.; Göttfert, F.; Ta, H.; Güther, A.; Fournier, M.; Rizzo, S.; Waldmann, H.; et al. Fluorogenic Probes for Live-Cell Imaging of the Cytoskeleton. *Nat. Methods* **2014**, *11* (7), 731–733.
- (12) Dickinson, B. C.; Chang, C. J. A Targetable Fluorescent Probe for Imaging Hydrogen Peroxide in the Mitochondria of Living Cells. *J. Am. Chem. Soc.* **2008**, *130* (30), 9638–9639.
- (13) Ridnour, L. A.; Sim, J. E.; Hayward, M. A.; Wink, D. A.; Martin, S. M.; Buettner, G. R.; Spitz, D. R. A Spectrophotometric Method for the Direct Detection and Quantitation of Nitric Oxide, Nitrite, and Nitrate in Cell Culture Media. *Anal. Biochem.* **2000**, *281* (2), 223–229.
- (14) Uppu, R. M.; Pryor, W. A. Synthesis of Peroxynitrite in a Two-Phase System Using Isoamyl Nitrite and Hydrogen Peroxide. *Anal. Biochem.* **1996**, *236* (2), 242–249.

- (15) Carpino, L. A.; Sadat-Aalae, D.; Beyermann, M. Tris(2-Aminoethyl)Amine as Substitute for 4-(Aminomethyl)Piperidine in the Fmoc/Polyamine Approach to Rapid Peptide Synthesis. *J. Org. Chem.* **1990**, *55* (5), 1673–1675.
- (16) Carballal, S.; Bartesaghi, S.; Radi, R. Kinetic and Mechanistic Considerations to Assess the Biological Fate of Peroxynitrite. *Biochim. Biophys. Acta* **2014**, *1840* (2), 768–780.
- (17) Peng, X.; Gong, F.; Chen, Y.; Jiang, Y.; Liu, J.; Yu, M.; Zhang, S.; Wang, M.; Xiao, G.; Liao, H. Autophagy Promotes Paclitaxel Resistance of Cervical Cancer Cells: Involvement of Warburg Effect Activated Hypoxia-Induced Factor 1- $\alpha$ -Mediated Signaling. *Cell Death Dis.* **2014**, *5* (8), e1367.
- (18) Lipton, S. A. NMDA Receptor Activity Regulates Transcription of Antioxidant Pathways. *Nat. Neurosci.* **2008**, *11* (4), 381–382.

## Appendix 2

### **Development of propargyl pyrocarbonate and its analogs for protein profiling**

## A2.1 Synopsis

In chapter 3 and 4, we have established thiophosphorodichloridates as a candidate for reactive histidine profiling. However, as stated in these chapters, the modification on histidine is labile under collision-based fragmentation conditions (CID and HCD) and mainly undergoes the neutral loss pathway, leading to unmodified peptides and their fragments on MS2 with little structural information about the site of modification on these mass spectrometers. Such pathways can be avoided in newer spectrometers equipped with electron-based fragmentation methods;<sup>1</sup> however, the limited availability of these new spectrometers compared to the traditional ones confines the wider application of this histidine profiling methodology. We are thus still seeking for histidine labeling reagents that provide stable adducts during collision-based fragmentation.

In this context, we were inspired by the commonly-known reagent, diethyl pyrocarbonate (DEPC), as a histidine labeling reagent. DEPC has been widely used to block active-site histidine residues of various ribonucleases.<sup>2</sup> Besides histidine, this molecule also modifies other nucleophilic amino acids on the protein surface.<sup>3-5</sup> To enable capture and enrichment of this modification, we replaced the ethyl groups with propargyl groups to allow for CuAAC reaction with other payloads and synthesized dipropargyl pyrocarbonate (DPPC). In order to achieve better reactivity and selectivity in aqueous buffers, we further replaced up to three oxygen atoms with sulfur to generate DPPC-S, DPPC-S2 and DPPC-S3. We observed DPPC-S3 as the mildest as well as the most selective reagent in this family of molecules, with the thiocarbamate adduct more resistant against hydrolysis. However, the histidine-DPPC-S3 adduct still showed appreciable electrophilic activity and was found to react with the amino groups on N-terminus and lysine. For DPPC-S, although the conjugate is less reactive, the transfer of the modification to amino groups was prevalent. This unexpected behavior renders (thio)pyrocarbonates not suitable for histidine profiling; nonetheless, these reagents may still be valuable for profiling protein surface amino acid residues as well as probing protein-protein interactions by such modification transfer process.

## A2.2 Methods

### A2.2.1 General synthetic and characterization methods

Unless otherwise noted, all commercial reagents were used without further purification. All reactions utilizing air- or moisture-sensitive reagents were performed in dried glassware under an atmosphere of dry N<sub>2</sub>. Thiophosgene was purchased from Oakwood Chemical (Estill, SC). All other reagents were purchased from Sigma-Aldrich. <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra were collected in CDCl<sub>3</sub>, MeOD or *d*<sub>6</sub>-acetone (Cambridge Isotope Laboratories, Cambridge MA) at 25 °C on AVB-400 spectrometer at the College of Chemistry NMR Facility at the University of California, Berkeley. All chemical shifts in <sup>1</sup>H NMR and <sup>13</sup>C NMR are reported in the standard  $\delta$  notation of ppm relative to residual solvent peak (CDCl<sub>3</sub>  $\delta$ H = 7.26,  $\delta$ C = 77.16; MeOD  $\delta$ H = 3.31,  $\delta$ C = 49.00; *d*<sub>6</sub>-acetone:  $\delta$ H = 2.05,  $\delta$ C = 29.84). Splitting patterns are indicated as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad.

### A2.2.2 Synthesis of propargyl pyrocarbonate and its sulfur-derivatives

**Dipropargyl pyrocarbonate (1, DPPC).** This compound was prepared according to a published procedure regarding carbonic anhydride preparation.<sup>6</sup> Briefly, NaOH (4.0 g, 100 mmol) was dissolved in H<sub>2</sub>O (16 mL) and cooled in an ice/water/NaCl bath. Tetrabutylammonium bromide (TBAB, 580 mg, 1.8 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and added to the NaOH solution. Propargyl chloroformate (1.76 mL, 18.1 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (50 mL) and added to the mixture by syringe pump over 30 min while stirring vigorously. The mixture was stirred for 3 h under -10 °C bath followed by quenching with acetic acid (7 mL). The organic phase was separated in a separatory funnel, washed with 2 × saturated NaHCO<sub>3</sub> and 1 × H<sub>2</sub>O, dried and concentrated to provide compound **1** as a light yellow oil (1.2 g, 73%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.76 (d, *J* = 2.5 Hz, 2H), 2.54 (t, *J* = 2.5 Hz, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 154.14, 76.76, 76.13, 55.84.

**Bis(propargyloxyformyl) sulfide (2, DPPC-S1).** This compound was prepared according to a published procedure regarding diacyl sulfide preparation.<sup>7</sup> Briefly, propargyl chloroformate (1.76 mL, 18.1 mmol) and hexadecyltributylphosphonium bromide (458 mg, 0.90 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (30 mL) and cooled in an ice/water bath. Na<sub>2</sub>S·9H<sub>2</sub>O (2.60 g, 10.8 mmol) was grinded into powder in a mortar and added portion-wise over 30 min. The mixture was stirred vigorously for another 3 h in an ice/water bath with the fluffy solid turned into thick slurry on the flask wall. TLC showed complete consumption of chloroformate. The supernatant was then passed through a pad of silica, concentrated to give compound **2** as a colorless oil (1.74 g, 97%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.88 – 4.85 (m, 1H), 2.59 (t, *J* = 2.5 Hz, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 161.80, 76.96, 75.99, 55.53.

**Propargyloxythioformyl propargyloxyformyl sulfide (3, DPPC-S2).** Propargyl alcohol (515 μL, 8.9 mmol) was dissolved in THF (20 mL) under N<sub>2</sub> cooled in an ice/water bath. To this solution was added dropwise *n*-butyllithium (3.0 mL of 2.5 M solution in hexanes, 7.5 mmol). After stirring at the same temperature for 10 min at the same temperature, carbon disulfide (539 μL, 8.9 mmol) was added. The deep yellow to orange xanthate solution was stirred for 30 min. While cooled in an ice/water bath, propargyl chloroformate (1.09 μL, 11.1 mmol) was added immediately. The resulting orange solution was further stirred for 2 h at room temperature and then concentrated. The slurry was dissolved in CH<sub>2</sub>Cl<sub>2</sub> and filtered to remove lithium chloride. The crude product was separated by silica column chromatograph (1:150 to 1:100 ethyl acetate/hexanes) to give compound **3** as a yellow solid (426 mg, 27%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.23 (d, *J* = 2.5 Hz, 1H), 4.84 (d, *J* = 2.5 Hz, 1H), 2.64 (t, *J* = 2.5 Hz, 1H), 2.59 (t, *J* = 2.4 Hz, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 201.62, 162.67, 77.48, 76.92, 76.06, 75.51, 60.95, 55.43.

**Bis(propargyloxythioformyl) sulfide (4, DPPC-S3).** Propargyl alcohol (493 μL, 8.54 mmol) was dissolved in anhydrous THF (5 mL) and cooled in a dry-ice/acetone bath. To this solution was added *n*-butyllithium (2.85 mL of 2.5 M solution in hexanes, 7.12 mmol) and the

light yellow solution was warmed to room temperature. Thiophosgene (600  $\mu$ L, 7.83 mmol) was dissolved in anhydrous THF (5 mL) cooled in a dry-ice/acetone bath. To this solution was added the lithium alcoholate solution dropwise over 10 – 15 min to form a yellow solution, which was transferred to diethyl ether (80 mL) while cold. The organic phase was washed with 2  $\times$  H<sub>2</sub>O and 1  $\times$  saturated NaCl, dried and concentrated to give the propargyl thiochloroformate as a yellow liquid, which was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (15 mL) and cooled in an ice/water bath. To this solution was added hexadecyltributylphosphonium bromide (181 mg, 0.356 mmol) and powdered Na<sub>2</sub>S $\cdot$ 9H<sub>2</sub>O (1.03 g, 4.27 mmol). The mixture was stirred vigorously for another 1 h in an ice/water bath with the fluffy solid turned into thick greenish slurry on the flask wall. TLC showed complete consumption of thiochloroformate. The supernatant was then led through a pad of silica and concentrated to give compound **4** as an orange solid (0.78 g, 96 %). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  5.21 (d,  $J$  = 2.4 Hz, 1H), 2.63 (t,  $J$  = 2.4 Hz, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  203.23, 77.50, 75.63, 61.20.

#### *A2.2.3 Characterization of DPPC-Sx reactions on amino acids*

Reactions on protected amino acid were analyzed by LC/MS using 1220 Infinity LC (Agilent, Santa Clara CA) coupled with Expression-L Compact Mass Spectrometer (Advion, Ithaca NY). Reaction mixtures with single protected amino acid were separated on a Zorbax rapid resolution cartridge (Agilent). Mixtures with multiple protected amino acids were separated on a Zorbax SB-phenyl column (Agilent, 4.6  $\times$  250 mm, 5  $\mu$ m). Solvent A was water + 0.05% formic acid and solvent B was methanol + 0.05% formic acid. The linear gradient employed for single protected amino acid was 25-100% B in 6.5 min and 100% B for 1.5 min; for multiple amino acids mixture was 45-100% B in 30 min and 100% B for 10 min.

#### *A2.2.4 Protein and amino acid labeling by DPPC-Sx*

All reactions were performed in 25 mM HEPES buffer, pH 7.0 containing 20% MeCN unless otherwise noted. Amino acid mixture (0.3 mM each, diluted from its 100 $\times$  stock solution of the mixture in DMSO) was treated with 3 mM of DPPC, DPPC-S1, DPPC-S2 or DPPC-S3 (100 $\times$  stock solution in MeCN) and reacted for 1 h followed by immediate LC/MS analysis. Protein and lysate samples were treated with 1 mM of DPPC-S or DPPC-S3 for 1h, precipitated with acetone followed by LC-MS/MS sample preparation protocol.

#### *A2.2.5 LC-MS/MS sample preparation of DPPC-Sx labeled protein or lysate*

The fast sample preparation protocol using immobilized trypsin was described in detail in Appendix 3 (A3.4.2).

#### *A2.2.6 LC-MS/MS analyses of DPPC-Sx labeled amino acids*

LC-MS/MS analyses were performed on a Q-Exactive Plus mass spectrometer (Thermo Fisher) operated with an Agilent 1260 HPLC system (Agilent, Santa Clara CA) using vented column setup. Samples were pressure-loaded onto a self-packed column (100  $\mu$ m  $\times$  190 nm, 3

$\mu\text{m}$  Aqua C18 125 Å, Phenomenex, Torrance CA) with a self-pulled electrospray emitter tip. Solvent A was 5% MeCN and 0.1% formic acid in H<sub>2</sub>O; solvent B was 80% MeCN, 20% H<sub>2</sub>O with 0.1% formic acid. The linear gradient program for proteomics samples consists of 0 min: 100% A, 5 min: 100% A, 60 min: 55% A, 70 min 0% A, 90 min 0% A with a constant pump flow rate of 0.150 ml/min. The linear gradient program for digested purified samples consists of 0 min: 100% A, 3 min: 100% A, 40 min: 55% A, 48 min 0% A, 62 min 0% A with a constant pump flow rate of 0.150 ml/min. HCD MS/MS spectra were recorded in the data-dependent mode using a Top 15 method. MS1 spectra were measured with a resolution of 70,000, an AGC target of 1e6 and a mass range from m/z 400 to 1,800. HCD MS/MS spectra were acquired with a 1.6 units of m/z isolation window, a resolution of 17,500, an AGC target of 5e4 and normalized collision energy of 27. Peptide m/z that triggered MS/MS scans were dynamically excluded from further MS/MS scans for 60 s.

The acquired raw file were processed using Trans-Proteomics Pipeline (Institute for Systems Biology, Seattle WA) with the comet search algorithm, searching for modification on N-terminus, His, Lys, Arg, Ser, Thr and Tyr.

## A2.3 Results and discussion

### A2.3.1 Design and synthesis of DPPC-Sx

DEPC has been reported to label histidine along with other nucleophilic amino acid residues.<sup>3-5,8</sup> Although DEPC hydrolyzes in aqueous condition very rapidly ( $t_{1/2}$  = 4 min at pH 6 and 9 min at pH 7 in phosphate buffer),<sup>9</sup> its reactivity with imidazole sidechains on histidine is very high and can achieve quantitative conversion; however, when excess amount of DEPC is applied, the Bamberger Cleavage will take place on the imidazole ring, resulting in ring-opened carbamate as product.<sup>10,11</sup>

Inspired by the unique reactivity of DEPC, we envisioned substituting the ethyl group with propargyl handle that can enable CuAAC between the adduct with more complex payloads. We thus synthesized dipropargyl pyrocarbonate (DPPC). Considering rapid hydrolysis of the pyrocarbonate under aqueous condition, we further replaced the center oxygen atom with sulfur to reduce the partial positive charge on the carbon atom and synthesized DPPC-S. In addition, we were able to synthesize molecules with one or two carbonyl oxygen substituted with sulfur, namely DPPC-S2 and DPPC-S3, respectively, in the hope that they are even milder labeling reagents, with DPPC-S2 giving rise to both imidazole-carbamate and imidazole-thiocarbamate and DPPC-S3 offering only imidazole-thiocarbamate (Scheme A2.1).

### A2.3.2 Reactivity of DPPC-Sx on amino acid model

With the propargyl pyrocarbonates and its derivatives in hand, we first tested their reactivity with Fmoc-protected amino acids as a small molecule model. For faster and more comparable selectivity test, we applied our compounds to the buffered solution containing seven nucleophilic amino acids (Fmoc-Lys-OH, Fmoc-Arg-OH, Fmoc-His-OH, Fmoc-Ser-OH, Fmoc-Thr-OH, Fmoc-Tyr-OH and Fmoc-Glu-OH) and analyzed the reaction by LC/MS (Figure

A2.1). Contrary to the high reactivity of DEPC, DPPC did not show observable reactivity with any of the amino acids under this condition, probably due to the higher electron-deficiency of propargyl group compared to ethyl group ( $pK_{a_{EtOH}}=16$ ,  $pK_{a_{HC=CCH_2OH}}=13$ ). In contrast, with a single oxygen replaced by sulfur, DPPC-S1 showed complete conversion of Fmoc-His-OH into its labeled carbamate counterpart. In addition, Fmoc-Lys-OH and Fmoc-Tyr-OH reacted with DPPC-S1 to a less extent. DPPC-S2 also showed complete conversion of Fmoc-His-OH and, as we expected, both carbamate and thiocarbamate appeared as product. For DPPC-S3, most Fmoc-His-OH was converted to thiocarbamate product, with a few hydrolyzed carbamate adduct. Moreover, different from DPPC-S and DPPC-S2, this three-sulfur analog showed reduced reactivity with Fmoc-Lys-OH and Fmoc-Tyr-OH, indicating that this milder reagent is more selective for Fmoc-His-OH.

We then tested the stability of the adduct against hydrolysis, as the DEPC adduct hydrolyzes slowly even at neutral pH.<sup>5</sup> For this purposes, we generated *in situ* the carbamate and thiocarbamate conjugate between Fmoc-His-OH and DPPC-S1 or DPPC-S3 and, after the complete hydrolysis of the thiopyrocarbonate reagent, monitored the decrease of labeled Fmoc-His-OH and increase of unlabeled Fmoc-His-OH under different conditions to deduce the hydrolysis percentage (Figure A2.2). The results show although the histidine adduct of DPPC-S3 still exhibits appreciable hydrolysis, the hydrolysis rate is much slower compared to that of DPPC-S1 adduct under both slight basic condition at 37°C (mimicking trypsin digestion condition;  $t_{1/2} = 2.6$  h and 6.0 h, respectively) and acidic condition at room temperature (mimicking LC/MS condition;  $t_{1/2} = 0.5$  h and 4.2 h, respectively). Considering that these conditions are both used in proteomics sample preparation and analysis, an increase in product stability would allow more modified peptides to survive and be analyzed by mass spectrometer. Collectively, we conclude that DPPC-S3 is the most promising candidate among the four reagents due to its milder reactivity, higher histidine selectivity and improved product stability.

### A2.3.3 Cell lysate labeling with DPPC-Sx

We then moved on to labeling of protein with DPPC-Sx reagents. We treated HeLa cell lysate with vehicle control or our reagents, followed by CuAAC with Cy3-N<sub>3</sub> for fluorescence readout after SDS-PAGE (Figure A2.3). All four reagents showed dose-dependent labeling on the cell lysate. In agreement with the small molecule results, DPPC-S1 showed the strongest fluorescence intensity, DPPC-S3 showed weaker signal and DPPC-S2 was in between; again, DPPC exhibited negligible reactivity on cell lysate when comparing to its sulfur analogs. The smearing gel bands suggest that a wide variety of proteins were labeled by these reagents, making these reagents promising for protein profiling purpose.

### A2.3.4 Analysis of DPPC-S1 and DPPC-S3 reactivity with lysate by LC-MS/MS

After confirming the reactivity of the reagents on small molecules and on lysate, we finally studied the selectivity of DPPC-S3 proteome-wide. Due to the hydrolysis activity of the adduct, we optimized our protocol to shorten the incubation time at 37 °C for digestion and cysteine blocking in order to retain as much modified peptides as possible (see Appendix 3 for

detail). With the revised protocol, we enriched the modified peptides by acid-cleavable biotin-azide and analyzed by LC-MS/MS profiling. After searching for the potential modifications on N-terminus, histidine, lysine, serine, threonine and tyrosine, we were surprised that only a handful of histidine residues were found to be modified; instead, N-terminus were found to be mostly modified, followed by lysine and other amino acids (Figure A2.4a). The modification on N-terminus is unlikely to happen during labeling since these amino groups were not present on an intact protein, which suggests that the modification can be transferred onto these sites after digestion. Considering that imidazole is an excellent leaving group and that scrambling of the histidine modification onto cysteine was also observed on DEPC counterpart,<sup>4</sup> it is possible that most of the thiocarbamate modification on histidine reacted with nearby amino groups on N-terminus and lysine to form a stable product (Figure 2.4c). We thus performed supplement assays on amino acids to prove our hypothesis, where we mixed DPPC-S3 labeled Fmoc-His-OH with H<sub>2</sub>N-Lys(Fmoc)-OMe as an N-terminus analog. Indeed, we observed appreciable consumption of Fmoc-His(COOCH<sub>2</sub>C≡CH)-OH and a robust appearance of HC≡CCH<sub>2</sub>OCO-Lys(Fmoc)-OH and unmodified Fmoc-His-OH after only 15 min treatment at room temperature (Figure 2.5a). Since the transfer of the modification to the N-terminus on peptide is intramolecular, it is most likely that the scrambling of the modification occurs.

With these results of histidine profiling, we switched back to DPPC-S1 as the reagent, which offers the product in closer analogy with that of DEPC. No transfer of the modification from Fmoc-His(COOCH<sub>2</sub>C≡CH)-OH to H<sub>2</sub>N-Lys(Fmoc)-OH was observed under the same condition, suggesting the imidazole carbamate less reactive than imidazole thiocarbamate (Figure 2.5b). However, when we performed the LC-MS/MS analysis of the lysate labeled with DPPC-S1, we were still only able to see a few modified histidine sites and lots of modification on N-terminus, most likely due to the faster kinetics of intramolecular transfer (Figure 2.4b). Collectively, these results suggest that both kinds of modification on imidazole of histidine are prone to both hydrolysis and most importantly, aminolysis with nearby amino groups, which suggest these reagents are not well-suited for histidine profiling.

#### *A2.3.5 Efforts towards avoiding scrambling of modification*

We then tried to see if we could prevent the transfer of this modification to amino groups. First we reasoned that the amino groups need to be deprotonated to enable the nucleophilic attack and therefore applied slightly acidic conditions for both labeling and trypsin. For protein labeling, we used pH 6.0 instead of pH 8.0, since the reaction tolerates a rather wide range of pH, although longer time is required to reach similar yield at lower pH (Figure A2.6). On the other hand, trypsin digestion was rather incomplete at pH 6.0 and we finally settled on pH 7.0 for the digestion step of proteomics sample preparation. However, the proteomics analysis of the sample prepared in this manner suggests that this approach is not sufficient to prevent the scrambling of the modification onto N-terminus.

We then attempted to use bulkier 1-methyl and 1,1-dimethyl propargyl group in place of propargyl group to slow down the transfer process. However, the synthesis of these derivatives was rather challenging. To test our hypothesis, we performed a quick analysis of the labeling of

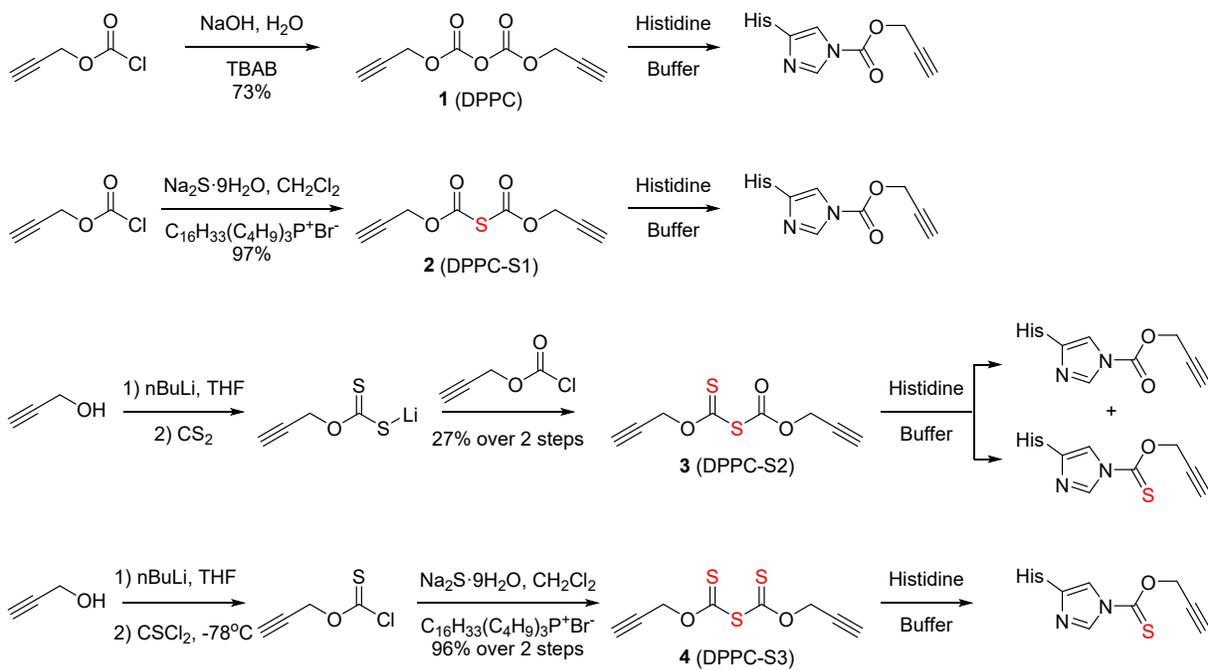
BSA with Boc<sub>2</sub>O, an analog of the target molecule that does not carry a clickable handle. This time the modified peptides showed significant neutral loss that make the modification site indistinguishable. This is not unexpected in that Boc-protection group is acid-labile with the *t*-butyl group very likely to eliminate into isobutene. The strong neutral loss behavior nullifies the purpose of developing reagents for histidine-profiling using collision-based tandem mass spectrometer and makes the synthesis not worthy.

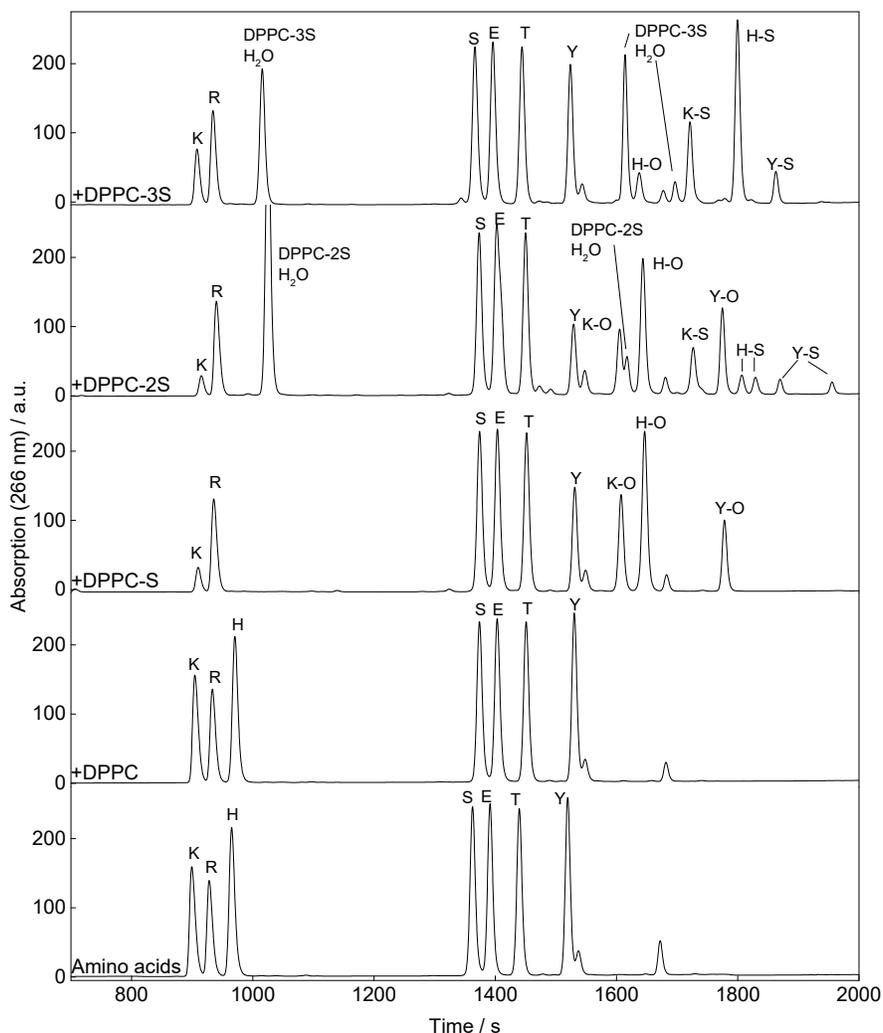
## **A2.4 Conclusion**

In this appendix we developed four DEPC analogs, with DPPC-S3 initially as a promising candidate for histidine profiling by LC-MS/MS. However, we discovered that the labeled histidine is not stable enough and modifies other amino groups during handling, which makes it impossible for studying histidine by profiling methods. Nonetheless, the newly developed DPPC-S3 provides such a reactivity intermediate that can be tailored to study protein-protein interaction, using histidine-labeled protein to transfer its modification to lysines on its closely interacting proteins. This reagent also holds promise in profiling of protein surface nucleophilic amino acids, with the labeling on histidine purposely removed by ammonia treatment. We hope this new reagent become useful in other ways in future research.

## A2.5 Figures and schemes

**Scheme A2.1** Synthesis and histidine-labeling of DPPC and its thiol-substituted derivatives.





**Figure A2.1** LC/MS analysis of the reaction between dipropargyl pyrocarbonate and its analogs (3 mM) with Fmoc-protected amino acid mixtures (0.3 mM each). Chemical identities of each chromatographic peak were determined by its  $m/z$ . From left to right:

K – Fmoc-Lys-OH

R – Fmoc-Arg-OH

H – Fmoc-His-OH

DPPC-S2/3-H<sub>2</sub>O – hydrolyzed DPPC2/3

S – Fmoc-Ser-OH

E – Fmoc-Glu –OH

T – Fmoc-Thr-OH

Y – Fmoc-Tyr-OH

K-O – Fmoc-Lys(COOCH<sub>2</sub>C≡CH)-OH

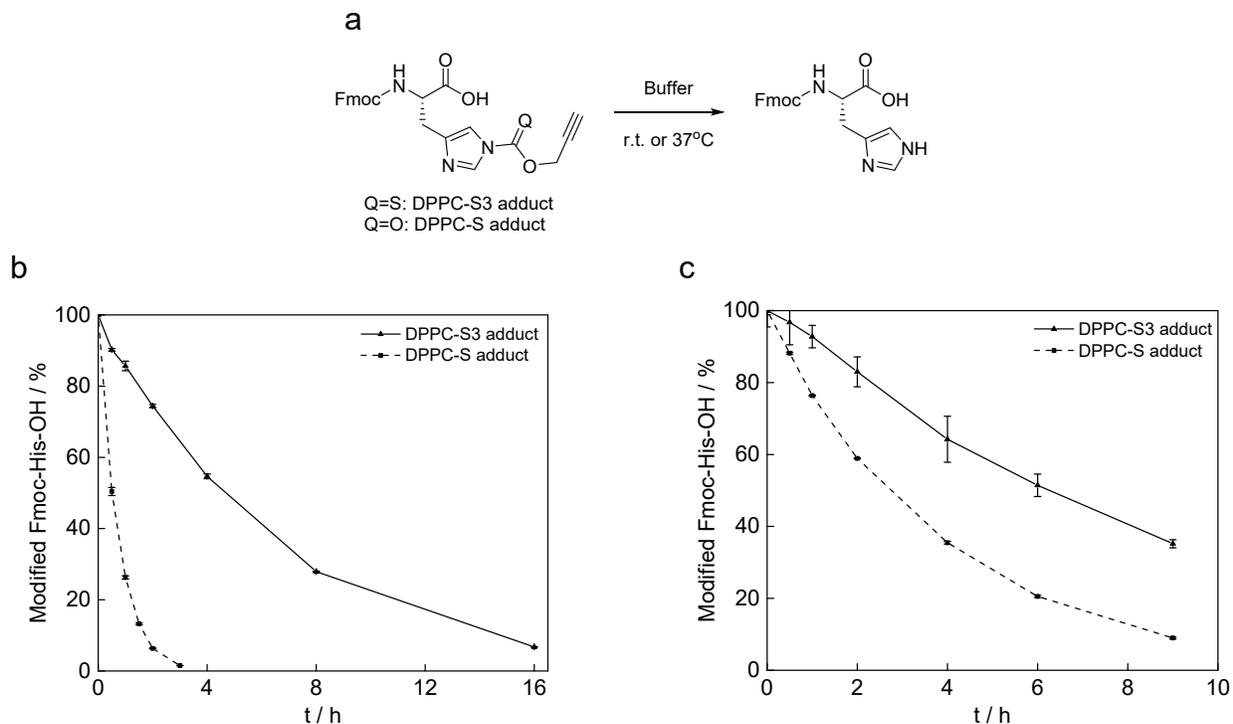
H-O – Fmoc-His(COOCH<sub>2</sub>C≡CH)-OH

K-S – Fmoc-Lys(CSOCH<sub>2</sub>C≡CH)-OH

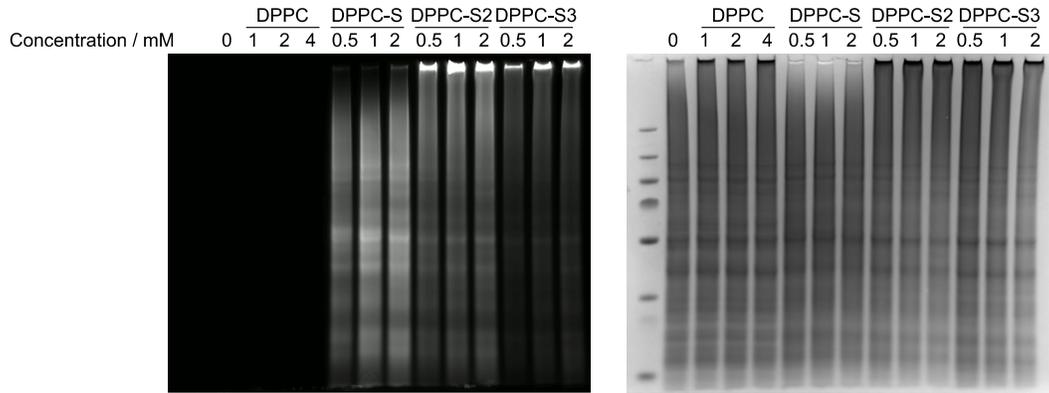
Y-O – Fmoc-Tyr(COOCH<sub>2</sub>C≡CH)-OH

H-S – Fmoc-His(CSOCH<sub>2</sub>C≡CH)-OH

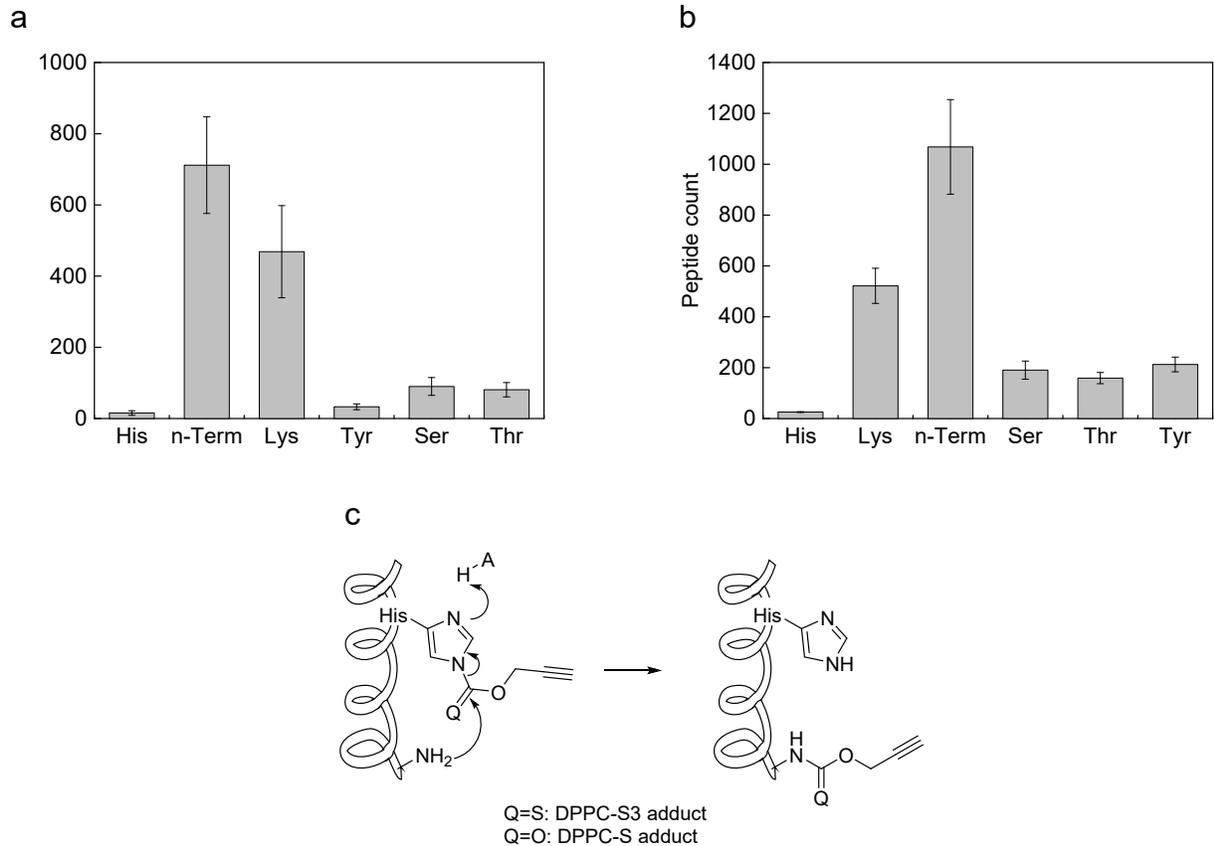
Y-S – Fmoc-Tyr(CSOCH<sub>2</sub>C≡CH)-OH



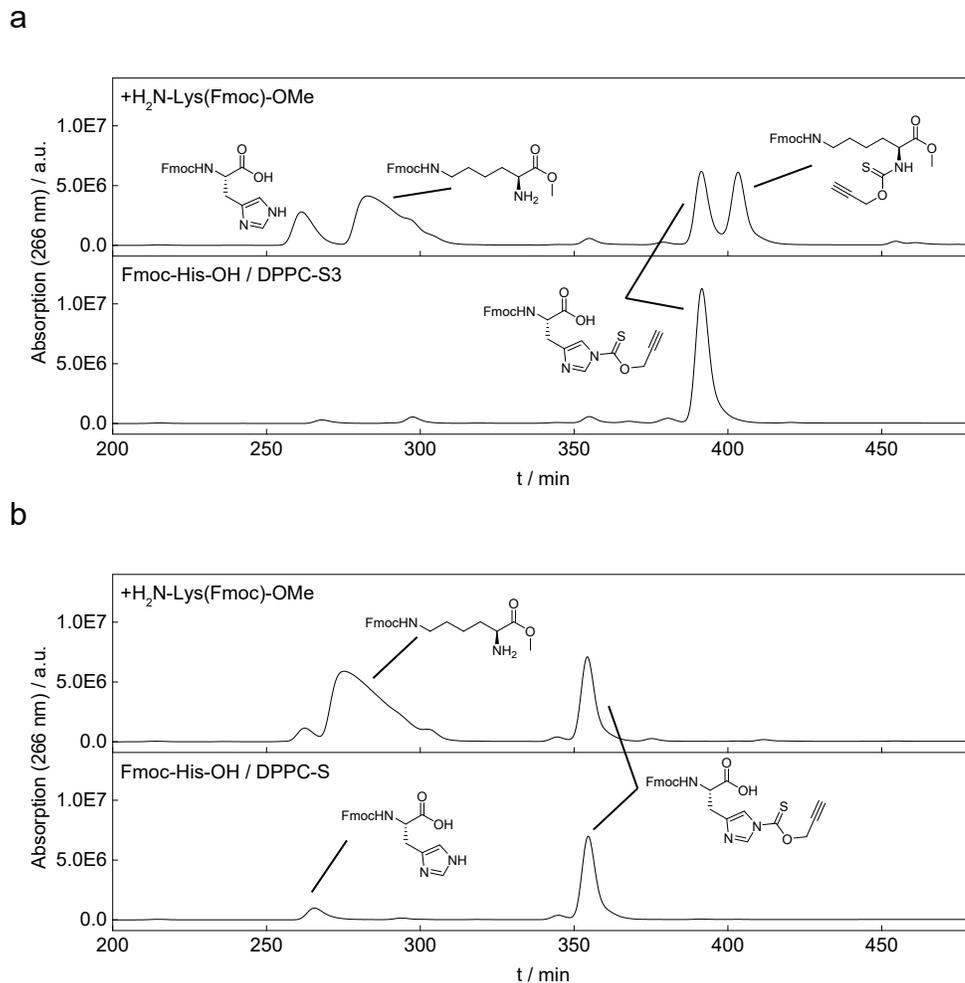
**Figure A2.2** Stability of histidine-DPPC-S1 or DPPC-S3 adduct against hydrolysis. (a) Scheme of the hydrolysis of the two adducts. (b) Hydrolysis of the two adducts in pH 8.0 buffer at 37°C (n = 3, average  $\pm$  s.d.). (c) Hydrolysis of the two adducts in the presence of 5% formic acid at room temperature (n = 3, average  $\pm$  s.d.).



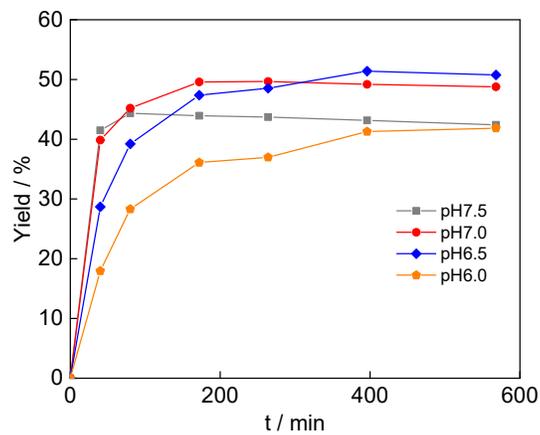
**Figure A2.3** DPPC-Sx labeling on HeLa lysate. HeLa lysate (1 mg/mL) were treated with 1 mM of DPPC, DPPC-S1, DPPC-S2 or DPPC-S3, clicked with Cy3-N<sub>3</sub> and analyzed by SDS-PAGE. Left: fluorescence. Right: coomassie.



**Figure A2.4** Modified amino acid sites detected by LC-MS/MS. (a) Number of unique DPPC-S3 labeled peptides with modification on each amino acid with ProteinProphet  $P \geq 0.8$  ( $n = 3$ , average  $\pm$  s.d.). (b) Number of unique DPPC-S3 labeled peptides with modification on each amino acid with ProteinProphet  $P \geq 0.8$  ( $n = 3$ , average  $\pm$  s.d.). (c) Proposed transfer pathway that reduces modified histidine and produces modified N-terminus and lysine.



**Figure A2.5** Transfer of the modification from Fmoc-His(COOCH<sub>2</sub>C≡CH)-OH to H<sub>2</sub>N-Lys(Fmoc)-OMe. HEPES buffer (25 mM pH 8.5) containing 0.3 mM Fmoc-His-OH was treated with (a) 1 mM DPPC-S3 or (b) 1 mM DPPC-S for 1 h, analyzed by LC/MS (top trace), treated with 0.6 mM H<sub>2</sub>N-Lys(Fmoc)-OMe for 15 min and analyzed again by LC/MS (bottom trace). Chemical identities of each chromatographic peak were determined by its m/z.



**Figure A2.6** Kinetics of the reaction between DPPC-S3 and Fmoc-His-OH under different pH. HEPES (pH 7.0 and 7.5, 25 mM) or MES (pH 6.0 and 6.5, 25 mM) buffer containing 0.5 mM Fmoc-His-OH were treated with 1 mM DPPC-S3 and the reaction was monitored by HPLC (n = 1).

## A2.6 References

- (1) Mikesch, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E. P.; Shabanowitz, J.; Hunt, D. F. The Utility of ETD Mass Spectrometry in Proteomic Analysis. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2006**, *1764* (12), 1811–1822.
- (2) Jones, G. H. On the Efficacy of Commonly Used Ribonuclease Inhibitors. *Biochem. Biophys. Res. Commun.* **1976**, *69* (2), 469–474.
- (3) Mendoza, V. L.; Vachet, R. W. Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Anal. Chem.* **2008**, *80* (8), 2895–2904.
- (4) Zhou, Y.; Vachet, R. W. Diethylpyrocarbonate Labeling for the Structural Analysis of Proteins: Label Scrambling in Solution and How to Avoid It. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (5), 899–907.
- (5) Zhou, Y.; Vachet, R. W. Increased Protein Structural Resolution from Diethylpyrocarbonate-Based Covalent Labeling and Mass Spectrometric Detection. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (4), 708–717.
- (6) Plusquellec, D.; Roulleau, F.; Lefeuvre, M.; Brown, E. A New Synthesis of Carboxylic and Carbonic Acid Anhydrides Using Phase Transfer Reactions. *Tetrahedron* **1988**, *44* (9), 2471–2476.
- (7) Juliá, S.; Tagle, G.; Vega, J. C. Synthesis of Diacyl Sulfides by Phase Transfer Catalysis. *Synth. Commun.* **1982**, *12* (11), 897–901.
- (8) Hnízda, A.; Šantrůček, J.; Šanda, M.; Strohalm, M.; Kodíček, M. Reactivity of Histidine and Lysine Side-Chains with Diethylpyrocarbonate — A Method to Identify Surface Exposed Residues in Proteins. *J. Biochem. Biophys. Methods* **2008**, *70* (6), 1091–1097.
- (9) Berger, S. L. Diethyl Pyrocarbonate: An Examination of Its Properties in Buffered Solutions with a New Assay Technique. *Anal. Biochem.* **1975**, *67* (2), 428–437.
- (10) Loosemore, M. J.; Pratt, R. F. The Irreversible Cleavage of Histidine Residues by Diethylpyrocarbonate (Ethoxyformic Anhydride). *FEBS Lett.* **1976**, *72* (1), 155–158.
- (11) Grace, M. E.; Loosemore, M. J.; Semmel, M. L.; Pratt, R. F. Kinetics and Mechanism of the Bamberger Cleavage of Imidazole and of Histidine Derivatives by Diethyl Pyrocarbonate in Aqueous Solution. *J. Am. Chem. Soc.* **1980**, *102* (22), 6784–6789.

## Appendix 3

### **Fast sample preparation of digested proteins by immobilized trypsin**

## A3.1 Synopsis

Mass spectrometry has enabled the analysis of proteins and their modifications in the context of complex proteomes.<sup>1</sup> While the technologies of the mass spectrometers as well as analytical software development evolve rapidly over the years, the results of mass spectrometry-based proteomics remains largely dependent on the quality of the sample itself. Due to the complexity of the proteome and varying research purposes, there is no standard protocol that applies to every scenario. On top of this, lengthy proteomics sample preparation has been a major concern, not only to enhance the time-efficiency of the research, but also to reduce deterioration of sample quality during its preparation. This is especially important for experiments focusing on the not-so-stable modifications on proteins in that prolonged and incorrectly harsh sample preparation procedures can possibly cause cumulative degradation of these modifications. As such, fast sample preparation procedure with least exposure to elevated temperature will greatly benefit the detection of such protein modifications, including naturally-occurring *S*- and *N*-phosphorylation, *O*- and *S*-glycosylation and acylation, and also labile artificial modifications introduced for labeling and enrichment purposes.

This appendix briefly summarizes the proteomics sample preparation procedure using immobilized trypsin for digestion with three written-up protocols as examples. Owing to the high enzyme density in immobilized trypsin, digestion of proteins can reach completion in as short as 30 minutes at 37 °C, rather than canonical overnight incubation for solution-based trypsin digestion. Moreover, the immobilized trypsin can be filtered so that immobilized trypsin will not contaminate the sample and interfere with the analysis. This appendix also puts forward some common caveats during sample preparation for readers' reference in the hope that they can be useful in future proteomics sample preparation.

## A3.2 General consideration

### A3.2.1 Protein precipitation

Protein precipitation is a useful tool to remove unwanted small molecules from a protein solution. This is especially useful in proteomics sample preparation in that a variety of reagents and materials are used sequentially which can cause interference with one another. Commonly used protein precipitation methods are summarized as follows:

1. Acetone precipitation. Add 3× volume of cold acetone to the protein solution, mix and cool the mixture at -20°C for ≥1h (-80°C will freeze the mixture), followed by centrifugation at high RPMs. Remove as much supernatant as possible. It has been reported that with higher ionic strength (≥10 mM NaCl) near full recovery of the protein is possible.<sup>2</sup> This precipitation is used in the example protocols due to its easy process.
2. Methanol precipitation. Add 3× volume of cold methanol to the protein solution, mix and cool the mixture at -20°C or -80°C for ≥1h, followed by centrifugation at high RPMs. Remove as much supernatant as possible.
3. Methanol/chloroform precipitation. This is the fastest precipitation method; however, special caution is required during separation of the layers. Briefly, add 4× volume of methanol and

1× volume of chloroform, mix thoroughly and centrifuge at high RPMs. The protein will precipitate as a flake at the interface of two layers. Remove the top layer carefully, add another 4× volume of methanol and centrifuge again. Remove as much supernatant as possible.

4. Trichloroacetic acid precipitation. This method involves the usage of a relatively strong acid and is not recommended for precipitation of protein with acid-sensitive modifications. Add 0.2 g/mL TCA into protein solution from 1 g/mL TCA stock solution, mix and incubate on ice for 30 min, centrifuge at high RPMs and discard the supernatant. Add cold acetone, mix and centrifuge at high RPMs. Remove as much supernatant as possible.

The protein pellet obtained from protein precipitation does not readily dissolve in buffer. For better reconstitution, the remaining volatile organic solvent needs to be evaporated. The dried pellet can be re-suspended in pure aqueous buffer in a bath sonicator to form a near-homogenous suspension, or dissolved in buffer containing surfactants or high concentration of chaotropic agents such as urea and guanidinium chloride.

#### *A3.2.2 Cysteine blocking with iodoacetamide*

This step involves reduction of disulfide bonds in cystine and blockage of all cysteine with a strong alkylating reagent, with iodoacetamide (IAA) mostly used in available protocols. Commonly used reducing reagents include tris(2-carboxyethyl)phosphine (TCEP) and dithiothreitol (DTT). Buffer used in this step usually contains high concentration of urea (~8 M) to denature protein for faster reaction between cysteine and IAA. Below are some notes that require attention:

1. Usage of IAA is not recommended in many reports due to its reactivity with other amino acids, including methionine, lysine, N-terminus, glutamic acid, serine, threonine and tyrosine.<sup>3,4</sup> Chloroacetamide (CAA) has been put forward as a substitute for IAA due to its lower reactivity with other amino acids and better selectivity towards cysteine; however, it has been reported to cause methionine oxidation.<sup>5</sup> We compared the blockage efficiency of IAA, CAA and acrylamide (AA) and the results suggest that while all reagents show almost complete blockage at pH 8.0, the alkylating reaction strongly favors higher pH (Figure A3.1). Therefore we used IAA at pH 6.5-7.0 rather than pH 8.0 to reduce its reactivity in the following sample preparation protocol examples.
2. Buffer containing urea in this step needs to be freshly prepared, because urea can undergo elimination upon storage to form isocyanic acid that results in protein carbamylation.<sup>6</sup> Such process can be suppressed by using ammonium buffers according to literature.<sup>7</sup> Considering that only a small amount of buffer is used in this step, for the following examples the buffer was freshly prepared from solid HEPES and urea in a microcentrifuge tube, adjusting pH with 5 M NaOH using a pH strip. For making a 50 mM HEPES buffer with 8M urea, the following empirical formulation can be used to fasten the buffer preparation, accounting for the volume of both H<sub>2</sub>O and urea:

$x \text{ mg HEPES} \Rightarrow \text{final volume } V = x / (0.050 \times 238.3) \text{ mL} \Rightarrow$

$m(\text{urea}) = (V/\text{mL}) \times 8 \times 60.06 \text{ mg} \Rightarrow V(\text{H}_2\text{O}) = [m(\text{urea})/\text{mg}] \times 1.37 \mu\text{L}$

3. TCEP is a better reductant for disulfide compared to DTT, because the reduction reaction for TCEP can take place at room temperature or 37°C, whereas DTT commonly requires 65°C. Lowered temperature is especially beneficial for studying unstable modifications on proteins.
4. Commercially available TCEP is strongly acidic when directly dissolved in water, due to the presence of phosphonium hydrochloride and the three carboxyl groups. To prepare a TCEP stock solution (500 mM), the crystals can be added to minimal amount of H<sub>2</sub>O followed by addition of NaOH to help dissolve and also adjust pH to around 7 and finally adjust the total volume to give a neutral, 500mM stock solution. This stock solution can be frozen at -80°C after splitting into use-and-discard aliquots.
5. The residual alkylating reagent and high concentration of urea can be harmful for downstream trypsin digestion. These small molecules can be easily removed by protein precipitation. To note, after addition of 4× volume of acetone into 8 M urea, the resulting suspension will freeze in -20°C freezer. To circumvent this issue, the 8 M urea solution can be first diluted into 4M with H<sub>2</sub>O before adding 8× volume of acetone.
6. This step is not only help simplifying MS spectrum analysis, but also facilitates trypsin digestion. Reduction of disulfide bonds and blockage of cysteine results in the denaturation of protein, and thus helps the following digestion step (Figure A3.2a).

### A3.2.3 Cu-click reaction

The copper-catalyzed azide-alkyne cycloaddition reaction (CuAAC) involves using a copper catalyst as well as a reducing reagent. This two species will result in drastic reactive oxygen species (ROS) generation that is harmful to the protein or peptide. To increase the yield and reduce side-reactions, we accommodates the recommendations of M.G. Finn and co-workers:<sup>8</sup>

1. A water-soluble copper ligand (THPTA or BTAA) is used instead of an insoluble one (TBTA). TBTA tends to aggregate in water with poor performance. We observed around 3-fold higher signal with Cy3-N<sub>3</sub> for THPTA and BTAA compared to TBTA. Additionally, the water-soluble ligand can be used with higher concentration as a ROS scavenger to prevent imidazole sidechains on histidine from oxidation.
2. Ascorbate is a preferred reducing reagent compared to TCEP due to its lack of Cu-binding property.
3. Aminoguanidine is added to scavenge oxidized ascorbate, which contains a triketone moiety that can form adduct with arginine residues.

A typical reaction mixture contains the following reagents, with all reagents added from their 100× stock solution:

Organic azide (100–200 μM), CuSO<sub>4</sub> (100 μM), THPTA or BTAA (500 μM) (premix before addition);

Sodium ascorbate (5 mM), aminoguanidine hydrochloride (1 mM) (at last, premix before addition).

#### A3.2.4 Digestion

Immobilized trypsin was purchased from Promega (V9012) and stored at 4°C. Addition of an extra mass spectrometry-compatible surfactant (e.g. ProteaseMax) is optional. Due to the high density of this reagent, it is exceptionally effective at protein digestion and can be used at room temperature. We incubate at 37°C for complete digestion, and were able to observe high activity even at slightly acidic conditions (Figure 3.2b).

The manufacture's protocol suggests digesting from concentrated protein solution with MeCN, which is difficult to prepare from precipitated proteins. Instead, 2 mg/mL near-homogenous suspension can be prepared in pH 8.0 aqueous buffer by sonication. This suspension will turn into a clear solution upon complete digestion, and an opaque eluent after digestion suggests incomplete digestion.

### A3.4 Examples of proteomics sample preparation protocol

(All reagents are added from their 100× stock solution in H<sub>2</sub>O, MeCN or DMSO. Centrifugation for precipitated protein is performed at 16,000×g for 5min.)

#### A3.4.1 LC-MS/MS sample preparation for RNaseA labeled with DPPC-S3

1. **Labeling:** Add 1 mM DPPC-S3 to 100 μL of 0.5 mg/mL RNaseA in 25 mM HEPES, pH 7.5 in a 0.75 mL microcentrifuge tube. Incubate for 1h.
2. **Removing excess reagent:** Precipitate the protein by adding 400 μL of cold acetone. Cool the mixture in -20°C freezer for ≥ 1h. Centrifuge and discard supernatant. Wash the pellet by adding 500 μL of cold 1:4 H<sub>2</sub>O/acetone. Centrifuge again and discard supernatant.
3. **Alkylating cysteine:** Dissolve dried pellet in 50 μL of 50 mM HEPES, pH 7.0 containing 8 M urea. Add 5 mM TCEP and 10 mM IAA and incubate in the dark at 37°C for 0.5 h.
4. **Removing excess urea, IAA and TCEP:** Dilute the sample with 50 μL of H<sub>2</sub>O and precipitate the protein using the same procedure as step 2
5. **Digestion by immobilized trypsin:** Shake the bottle of immobilized trypsin to re-suspend the beads and transfer 20 μL into a PCR tube. Centrifuge the beads for a few seconds on a table-top centrifuge and carefully remove the supernatant. Wash the beads by adding 150 μL of 25 mM HEPES pH 8.0 containing 1 mM CaCl<sub>2</sub>, vortex, centrifuge and discard the supernatant. Repeat the wash step.

Reconstitute the dried pellet from step 4 in 25  $\mu\text{L}$  of 25 mM HEPES pH 8.0 containing 1 mM  $\text{CaCl}_2$  and 0.01% ProteaseMax (added from frozen aliquots of 1% stock solution in 25 mM  $\text{NH}_4\text{CO}_3$ ) to form a near-homogeneous suspension (vortex/sonicate/pipette if necessary). Transfer the protein suspension into the PCR tube containing immobilized trypsin and vortex thoroughly. Place the PCR tube on a rotator at 37°C and incubate for 30 min. Take out the mixture and transfer onto a spin filter on a receiving microcentrifuge tube. Centrifuge for a few seconds on a table-top centrifuge. Wash the PCR tube with 20  $\mu\text{L}$  of 1:9 MeCH/H<sub>2</sub>O, transfer to the spin filter and centrifuge. This will provide a clear solution of the digested protein in approximately 50  $\mu\text{L}$  of 5% MeCN; 5  $\mu\text{L}$  is enough for each analysis. Freeze the sample at -80°C until analysis.

(Turbid filtrate indicate incomplete digestion, which requires attention.)

#### *A3.4.2 Proteomics sample preparation of DPPC-S3 labeled cell lysate using acid-cleavable biotin-azide for enrichment*

(The starting amount of protein is designed for reactive labeling reagents like DPPC-S3, i.e. with very faint fluorescent for the colored bands of the PageRuler Plus Prestained Protein Ladder compared to labeled lysate after CuAAC with Cy3-N<sub>3</sub> or Rhodamine-N<sub>3</sub>. For labeling reagent with higher or lower reactivity the amount can be adjusted accordingly.)

(For easy transfer of streptavidin-agarose beads, use 200  $\mu\text{L}$  pipette tip with its end cut to allow for ~1 mm diameter.)

1. **Labeling:** Add 500  $\mu\text{M}$  DPPC-S3 to 300  $\mu\text{L}$  of 1 mg/mL 231-mfp cell lysate in 25 mM HEPES, pH 7.5 in a 2 mL microcentrifuge tube. Incubate for 1h.
2. **Removing excess reagent:** Precipitate the protein by adding 1200  $\mu\text{L}$  of cold acetone. Cool the mixture in -20°C freezer for  $\geq$  1h. Centrifuge and discard supernatant. Wash the pellet by adding 1000  $\mu\text{L}$  of cold 1:4 H<sub>2</sub>O/acetone. Centrifuge again and discard supernatant.
3. **Alkylating cysteine:** Dissolve the dried pellet in 150  $\mu\text{L}$  of 50 mM HEPES, pH 7.0 containing 8 M urea. Add 5 mM TCEP and 10 mM IAA and incubate in the dark at 37°C for 0.5 h.
4. **Removing excess urea, IAA and TCEP:** Dilute the sample with 150  $\mu\text{L}$  of H<sub>2</sub>O and precipitate the protein using the same procedure as step 2.
5. **CuAAC with acid-cleavable biotin-azide:** Dissolve the protein pellet in 300  $\mu\text{L}$  25 mM HEPES, pH 7.5 containing 0.1% SDS by vortex, sonication and/or pipetting. Add 200  $\mu\text{M}$  acid-cleavable biotin-azide, 100  $\mu\text{M}$   $\text{CuSO}_4$  and 500  $\mu\text{M}$  THPTA, followed by 1 mM aminoguanidine and 5 mM sodium ascorbate. Place on a rotator and mix for 1 h in the dark.

Add 2.5 mM Na<sub>4</sub>edta to quench the reaction.

6. **Removing excess click reaction reagents:** Same procedure as step 2.
7. **Digestion by immobilized trypsin:** Shake the bottle of immobilized trypsin to re-suspend the beads and transfer 150  $\mu$ L into a 0.75 mL microcentrifuge tube. Centrifuge the beads for a few seconds on a table-top centrifuge and carefully remove the supernatant. Wash the beads by adding 150  $\mu$ L of 25 mM HEPES pH 8.0 containing 1 mM CaCl<sub>2</sub>, vortex, centrifuge and discard the supernatant. Repeat the wash step twice.

Reconstitute the dried pellet from step 6 in 150  $\mu$ L of 25 mM HEPES pH 8.0 containing 1 mM CaCl<sub>2</sub> and 0.01% ProteaseMax (added from frozen aliquots of 1% in 25 mM NH<sub>4</sub>CO<sub>3</sub>) to form a near-homogeneous suspension (vortex/sonicate/pipette if necessary). Transfer the protein suspension into the microcentrifuge tube containing immobilized trypsin and vortex thoroughly. Place the microcentrifuge tube on a rotator at 37°C and incubate for 30 min.

8. **Pull-down by streptavidin-argarose beads:** Shake the bottle of streptavidin-argarose beads and transfer 200  $\mu$ L to a 1.5 mL microcentrifuge tube. Centrifuge at 1000 $\times$ g for 1 min and discard the supernatant carefully. Wash the beads by adding 500  $\mu$ L of 25 mM HEPES, pH 8.0, vortex, centrifuge and discard the supernatant. Repeat the wash step twice

Take out the digested protein mixture and transfer onto a spin filter on the microcentrifuge tube with streptavidin-argarose beads. Centrifuge for a few seconds on a table-top centrifuge. Wash the digestion microcentrifuge tube with 120  $\mu$ L of 1:9 MeCN/HEPES buffer (25 mM HEPES, pH 8.0), transfer to the spin filter and centrifuge. Place the microcentrifuge tube on a rotator at 4°C and mix for 4 h.

9. **Washing the beads:** Transfer the beads from the microcentrifuge tube into a spin filter mounted on a vacuum manifold and remove the liquid. Transfer the remaining beads in the microcentrifuge tube and wash the beads with 3  $\times$  500  $\mu$ L 5% MeCN in HEPES buffer (25 mM HEPES, pH 8.0) and 3  $\times$  500  $\mu$ L 5% MeCN in H<sub>2</sub>O.
10. **Releasing the peptides:** Transfer the beads from the spin filter into a 1.5 mL microcentrifuge tube with 4  $\times$  100  $\mu$ L H<sub>2</sub>O. Treat the beads slurry with 100  $\mu$ L of 5% formic acid in MeCN. Place the mixture on a rotator and mix for 20 min. Transfer the beads slurry into a spin filter on a receiving tube and centrifuge at 1000 $\times$ g for 1 min. Transfer the remaining beads in the microcentrifuge tube and wash the beads with 2  $\times$  200  $\mu$ L 1:1 MeCN/H<sub>2</sub>O and combine the eluents in a 1.75 mL microcentrifuge tube. Lyophilize the eluent until less than 10  $\mu$ L in volume. Adjust the remainder to 20  $\mu$ L; 10  $\mu$ L is enough for each analysis. Freeze the sample at -80°C until analysis.

#### *A3.4.3 Proteomics sample preparation of O5C-TPAC labeled cell lysate using trisammonium azide for enrichment*

(This protocol features click reaction on peptides after digestion, which is reported to be more effective than that whole protein click reaction.<sup>9</sup>)

1. **Labeling:** Add 2 mM O5C-TPAC to 600  $\mu$ L of 1 mg/mL 231-mfp cell lysate in 50 mM HEPES, pH 8.5 in a 2 mL microcentrifuge tube. Incubate for 1h.
2. **Removing excess reagent:** Split the protein solution into two 2 mL microcentrifuge tubes and precipitate the protein by adding 1200  $\mu$ L of cold acetone to each half. Cool the mixture in  $-20^{\circ}\text{C}$  freezer for  $\geq$  1h. Centrifuge one tube, combine with another, centrifuge and discard the supernatant. Wash the pellet by adding 1000  $\mu$ L of cold 1:4  $\text{H}_2\text{O}$ /acetone. Centrifuge again and discard supernatant.
3. **Alkylating cysteine:** Dissolve the dried pellet in 300  $\mu$ L of 50 mM HEPES, pH 7.0 containing 8 M urea. Add 5 mM TCEP and 10 mM IAA and incubate in the dark at  $37^{\circ}\text{C}$  for 0.5 h.
4. **Removing excess urea, IAA and TCEP:** Dilute the sample with 300  $\mu$ L of  $\text{H}_2\text{O}$  and precipitate the protein using the same procedure as step 2
5. **Digestion by immobilized trypsin:** Shake the bottle of immobilized trypsin to re-suspend the beads and transfer 200  $\mu$ L into a 0.75 mL microcentrifuge tube. Centrifuge the beads for a few seconds on a table-top centrifuge and carefully remove the supernatant. Wash the beads by adding 150  $\mu$ L of 25 mM HEPES pH 8.0 containing 1 mM  $\text{CaCl}_2$ , vortex, centrifuge and discard the supernatant. Repeat the wash step twice.

Reconstitute the dried pellet from step 4 in 300  $\mu$ L of 25 mM HEPES pH 8.0 containing 1 mM  $\text{CaCl}_2$  and 0.01% ProteaseMax (added from frozen aliquots of 1% stock solution in 25 mM  $\text{NH}_4\text{CO}_3$ . This is an anionic detergent that does not interfere with cation exchange fractionation) to form a near-homogeneous suspension (vortex/sonicate/pipette if necessary). Transfer the protein suspension into the microcentrifuge tube containing immobilized trypsin and vortex thoroughly. Place the microcentrifuge tube on a rotator at  $37^{\circ}\text{C}$  and incubate for 30 min.

Take out the digested protein mixture and transfer onto a spin filter on a 1.75 mL receiving microcentrifuge tube. Centrifuge for a few seconds on a table-top centrifuge. Wash the digestion microcentrifuge tube with  $2 \times 50 \mu\text{L}$  40% MeCN in  $\text{H}_2\text{O}$ , transfer to the spin filter and centrifuge. This will provide the digested peptides as a clear solution in approximately 400  $\mu$ L HEPES buffer containing 10% MeCN.

6. **CuAAC with acid-cleavable biotin-azide:** Add 200  $\mu\text{M}$  triammonium azide, 100  $\mu\text{M}$   $\text{CuSO}_4$  and 500  $\mu\text{M}$  THPTA, followed by 1 mM aminoguanidine and 5 mM sodium ascorbate. Place on a rotator and mix for 1 h at 37°C in the dark.
7. **Enrichment by SCX spin column:** Add 4  $\mu\text{L}$  of formic acid to the click reaction solution for acidification. Load the sample onto a strong cation exchange spin column (Thermos, catalog # 90008) and spin inside a 50 mL conical centrifuge tube at 2,000 $\times$ g for 5 – 7 min until dryness using centrifuge equipped with a swinging-bucket rotor. The following fractions can be collected for comparison:

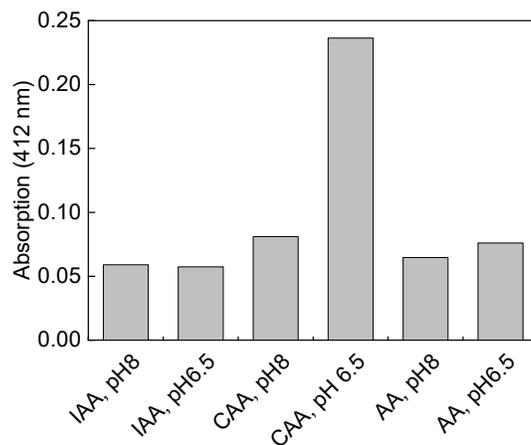
Wash: add 400  $\mu\text{L}$  of 10% MeCN solution containing 300 mM NaCl, 2.5 mM formic acid and 2.5 mM sodium formate; centrifuge.

Neutralization: add 400  $\mu\text{L}$  of 10% MeCN solution containing 50 mM ammonium acetate; centrifuge.

Elution: add 400  $\mu\text{L}$  of 10% MeCN solution containing 3 M ammonium acetate; centrifuge.

Lyophilize each fraction to dryness and reconstitute in 25  $\mu\text{L}$  of  $\text{H}_2\text{O}$  (wash fraction contains NaCl as solid after lyophilize; neutralization fraction contains less solid and elution fraction contains only a little solid residue); 1  $\mu\text{L}$  of the wash fraction, or 8  $\mu\text{L}$  of the neutralization and elution fraction is enough for each analysis. Freeze the sample at -80°C until analysis.

### A3.5 Figures



**Figure A3.1** Cysteine blockage with different alkylating conditions. HeLa lysate in buffer with indicated pH were treated with 5 mM TCEP and 10 mM alkylating reagent for 30 min at 37°C; the residual, non-blocked cysteine was quantified DTNB assay.



### A3.6 References

- (1) Han, X.; Aslanian, A.; Yates, J. R. Mass Spectrometry for Proteomics. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 483–490.
- (2) Crowell, A. M. J.; Wall, M. J.; Doucette, A. A. Maximizing Recovery of Water-Soluble Proteins through Acetone Precipitation. *Anal. Chim. Acta* **2013**, *796*, 48–54.
- (3) Nielsen, M. L.; Vermeulen, M.; Bonaldi, T.; Cox, J.; Moroder, L.; Mann, M. Iodoacetamide-Induced Artifact Mimics Ubiquitination in Mass Spectrometry. *Nat. Methods* **2008**, *5* (6), 459–460.
- (4) Müller, T.; Winter, D. Systematic Evaluation of Protein Reduction and Alkylation Reveals Massive Unspecific Side Effects by Iodine-Containing Reagents. *Mol. Cell. Proteomics* **2017**, *16* (7), 1173–1187.
- (5) Hains, P. G.; Robinson, P. J. The Impact of Commonly Used Alkylating Agents on Artifactual Peptide Modification. *J. Proteome Res.* **2017**, *16* (9), 3443–3447.
- (6) Kollipara, L.; Zahedi, R. P. Protein Carbamylation: In Vivo Modification or in Vitro Artefact? *PROTEOMICS* **2013**, *13* (6), 941–944.
- (7) Sun, S.; Zhou, J.-Y.; Yang, W.; Zhang, H. Inhibition of Protein Carbamylation in Urea Solution Using Ammonium-Containing Buffers. *Anal. Biochem.* **2014**, *446*, 76–81.
- (8) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. Analysis and Optimization of Copper-Catalyzed Azide–Alkyne Cycloaddition for Bioconjugation. *Angew. Chem. Int. Ed.* *48* (52), 9879–9883.
- (9) Yang, J.; Gupta, V.; Carroll, K. S.; Liebler, D. C. Site-Specific Mapping and Quantification of Protein S-Sulphenylation in Cells. *Nat. Commun.* **2014**, *5*, 4776.