

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

**New Reactivities and Functionalities Introduced by
Mimics of Protein Post-Translational Modifications**

Haoyu Chen

aus

Banjing, Provinz Jiangsu, Volksrepublik China

2024

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn *Dr. Pavel Kielkowski* betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

Haoyu Chen 14.10.2024

Dissertation eingereicht am: 02.09.2024

1. Gutachter: Dr. Pavel Kielkowski

2. Gutachter: Prof. Dr. Thomas Carell

Mündliche Prüfung am: 09.10.2024

Was mich nicht umbringt, macht mich stärker.

Twilight of the Idols, Friedrich Nietzsche

Publications (Submitted)

H. Chen, P. Kielkowski, Self-Assembled PROTACs Enable Target Glycoproteins Degradation in the Living Cells, *J. Am. Chem. Soc.*, 2024, **currently under review**.

H. Chen, D. Wu, J. Holzinger, R. Götz, D. Didier, A. K. Schütz, S. Schneider, P. Kielkowski, Aryl Radicals Generated from Aryl Pinacol Boronates to Modify Peptides and Proteins, *Chem. Eur. J.*, 2024, **currently under review**.

Conference Contributions

Poster “Reactivity of aryl pinacol boronates towards disulfide bonds”, 24th Tetrahedron Symposium, 18 – 21 June 2024, Montpellier, France.

Oral presentation “Click-released PROTACs enable target protein degradation in the living cells”, Chemical Biology Ireland Conference 2024, 22 – 23 July 2024, Galway, Ireland; fully funded by Boehringer Ingelheim Foundation.

Acknowledgment

This journey in the past 3 years and 9 months has been an extraordinary experience for me, filled with unforgettable challenges and triumphs, exciting growth and discovery. It would not be possible to this ending point without the support and encouragement of many wonderful individuals nearby, to whom I am eternally grateful.

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Pavel Kielkowski, whose unwavering help, insightful advice, and comprehensive guidance have been invaluable for me. Your belief in my capabilities and your dedication to my success have inspired me to persevere even in the most difficult times, especially when I was not secured with a funding at the beginning of doctoral study. Thank you for being a remarkable supervisor and a constant source of motivation.

I am profoundly grateful to the members of my dissertation committee, including Prof. Dr. Thomas Carell, Prof. Dr. Knud Jørgen Jensen, Prof. Dr. Ivan Huc, Prof. Dr. Konstantin Karaghiosoff, and Dr. Sabine Schneider. A special thank you goes to Prof. Dr. Knud Jørgen Jensen from University of Copenhagen, who was the supervisor for my master thesis project, to present in LMU Munich and witness another historical moment in my career for the second time. I will always engrave such fabulous destiny between us in the future.

I would also like to express my gratitude to my colleagues and friends at the Institute of Chemical Epigenetics at LMU Munich, including Dmytro Makarov, Andreas Wiest, Laura Hoffman, Rubaba Abanti, Manuel Hertwig, and Tobias Becker. Your camaraderie, collaboration, and support have made this journey more enjoyable and rewarding. I am thankful for the stimulating discussions, the shared laughs, and the countless moments of solidarity.

I would also like to acknowledge the financial support from LMU-CSC Scholarship Program 2021, whose generosity has made this research possible. Your investment in my work is deeply appreciated. I also thank Deutsche Forschungsgemeinschaft (DFG) for providing daily consumables and advanced instrumentation for performing all experiments smoothly.

I then extend my gratitude to all the participants and collaborators who contributed to this research, including Dongqing Wu, Dr. Sabine Schneider, Dr. Julian Holzinger, Prof. Dr. Anne Schütz, Dr. Dorian Didier, and Rainer Götz. Your willingness to share your time, knowledge, and experiences has been crucial to the success of this project.

To my parents, Mr. Bing Chen and Mrs. Xiaoming Ma, words cannot express how grateful I am. Thank you for instilling in me the values of hard work, perseverance, and the importance of education. I am so proud to be the first PhD graduate from Chen family in my hometown, and grateful for your courage to exclude all difficulties to get the Schengen visa for presenting physically in my doctoral defense after an 11-hours flight.

This dissertation is dedicated to everyone who believed in me, supported me, and walked with me on this journey. Thank you for making this achievement possible.

With heartfelt thanks,

Haoyu Chen

Table of Contents

Summary.....	I
Zusammenfassung.....	III
1. Introduction.....	1
1.1 Protein Post-Translational Modifications (PTMs).....	1
1.2 Protein O-GlcNAcylation and O-GalNAcylation.....	2
1.3 Click Chemistry.....	5
1.4 Protein Bioorthogonal Labelling <i>via</i> Suzuki Cross-Coupling Reaction.....	6
1.5 Aryl Radicals: Generation and Applications.....	8
1.6 Organoboranes in Radical Initiation or Termination.....	9
1.7 Cysteine S-Arylation Chemistry.....	11
1.8 Proteolysis Targeting Chimeras (PROTACs).....	14
1.8.1 Protein Ubiquitinylation and Principles of PROTAC.....	16
1.8.2 Structures of PROTACs.....	16
1.8.3 PROTAC Preparation and Implementation with Click Chemistry.....	16
1.9 Chemical Proteomics.....	20
1.9.1 Single-Pot, Solid-Phase-Enhanced Sample Preparation (SP3) for Proteomics Experiments.....	20
1.9.2 Orbitrap-Based Mass Spectrometry.....	21
1.9.3 Data Acquisition.....	22
2. Aim of the Thesis.....	24
3. Results and Discussion.....	25
3.1 Suzuki Cross-Coupling Reaction for Protein Labelling in the Living Cells.....	25
3.2 Design and Synthesis of Iodinated Probes for Pulling Down AMPylated or Tyrosinated Proteins.....	26
3.2.1 Synthesis of Iodinated AMPylation Probes.....	27
3.2.2 Synthesis of Phosphorochloridate.....	28
3.2.3 Synthesis of Tyrosination Probes.....	28
3.3 Chemical Proteomics for Identifying AMPylated Proteins.....	29
3.3.1 Workflow Description.....	29
3.3.2 Data Analysis.....	30
3.4 Unknown Reactivities between Aryl Pinacol Boronates and Cell Lysates.....	31
3.5 Reactivity Studies on Single Proteins.....	33
3.6 Reactivity Studies on Short Peptides: Disulfide S-Arylation with Aryl Pinacol Boronates.....	34
3.7 Studies of Substrate Scopes and Mechanistic Pathway.....	35
3.8 Further Attempts Using Iodinated Tyrosination Probes.....	37
3.9 GlyTAC Approach for Selectively Degrading Glycoproteins.....	38
3.10 GlyTAC - Establishment.....	39
3.10.1 Synthesis of DBCO-Thal and Cytotoxicity Testing.....	39
3.10.2 GlyTAC - Proof-of-Principle.....	40
3.11 GlyTAC - Characterization.....	40
3.12 GlyTACs - Mechanism of Action.....	43

3.13 GlyTACs - Proteome-Wide Metabolic Alternation	44
3.14 GlyTACs - Screening in Cancer Cell Lines	46
4. Bibliography.....	47
5. List of Abbreviations	57
6. List of Figures and Tables.....	60
7. Experimental Sections.....	61
7.1 Organic Synthesis.....	61
7.1.1 Methods for Organic Synthesis	61
7.1.2 Synthesis of AMPylation Probes 1 and 2.....	62
7.1.3 Synthesis of AMPylation Probe 3	63
7.1.4 Synthesis of AMPylation Probe 4	65
7.1.5 Synthesis of Phosphorochloridate	66
7.1.6 Synthesis of Tyrosination Probes 5 and 6	67
7.1.7 Synthesis of Biotinylated Boronic Acid Pinacol Ester 21	68
7.1.8 Synthesis of Palladium Catalyst.....	69
7.1.9 Synthesis of TAMRA-Bpin 22	69
7.1.10 Synthesis of GSSG, KEHSPC Disulfide and RWTPCD Disulfide	71
7.1.11 Synthesis of S-arylated Glutathione, KEHSPC and RWTPCD	72
7.1.12 Synthesis of DBCO-Thal	74
7.1.13 Synthesis of Ac ₄ GalNAz	76
7.2 Biological Methods and Analysis.....	78
7.2.1 Cell Culture, Harvesting, Lysates Preparation and Protein Concentration Measurement.....	78
7.2.2 ThioredoxinA (wild type, C33S, C36S) Expression and Purification	78
7.2.3 Methods for Cell Lysates or Single Protein Labelling	80
7.2.4 In-gel Fluorescence Analysis.....	81
7.2.5 Treatments of Two-Components GlyTACs <i>in cellulo</i>	81
7.2.6 MTT Cytotoxicity Assay	82
7.2.7 Western Blot	83
7.2.8 MS Sample Preparations	84
7.2.9 Mass Spectrometry Measurements.....	85
7.2.10 Data Acquisition and Calculation	86
7.2.11 Data Visualization.....	87
8. Appendix	89
8.1 Supplementary Figures.....	89
8.2 Supplementary Table	115
8.3 Uncropped Images	116
8.4 Mass Analysis.....	120
8.5 NMR Spectrums	127

Summary

Protein post-translational modifications (PTMs) play crucial roles in regulating protein function, stability, localization, and interactions. These modifications, including phosphorylation, glycosylation, ubiquitinylation, acetylation, and methylation, among others, are able to activate or deactivate enzymatic activities, dictate subcellular localization, mediate protein-protein interactions, and label proteins for degradation. By dynamically altering protein properties, PTMs enable cells to respond rapidly to environmental changes and maintain homeostasis, thereby contributing to processes such as signal transduction, immune response, cell cycle control, and apoptosis.

To study protein PTMs, scientists employ a variety of methods such as high-resolution tandem mass spectrometry that allows for the precise identification and quantification of PTMs by analyzing peptide fragments. Western blotting, using specific antibodies recognizing modified residues, is another common method to detect and analyze PTMs. These methods, combined with bioinformatics tools, offer comprehensive approaches to understanding the complex roles of PTMs in protein regulation.

We started to investigate a newly discovered PTM called adenylation (AMPylation) by developing halogen-modified AMPylation mimics for subsequent cross-coupling via Suzuki-Miyaura reactions in the living cells. However, the downstream chemical biology analysis using functionalized aryl pinacol boronates revealed an unknown reactivity. For example, incubation between cell lysates and a fluorescent aryl pinacol boronate has brought significant labelling regardless of the addition of artificial AMPylation probes or palladium catalysts. To our delight, we found that the unknown reactivity was due to selective protein labelling (*S*-arylation on disulfides) triggered by ammonium persulfate (APS). Then oxidized glutathione and two more short peptides were applied to verify the *S*-arylation product formation, which was further fragmented in high-resolution MS2 to identify diagnostic ions. The radical-based mechanistic pathway of this reaction was confirmed by 5,5-dimethyl-1-pyrroline *N*-oxide (DMPO) spin trap experiments, and the reaction was found to proceed only when aryl moieties were substituted with electron-donating groups, which was consistent with competition studies in cell lysates. This method was the first study to perform selective disulfide *S*-arylation with aryl radicals in mild aqueous conditions, and there was no need to use metal catalyst or photocatalysis.

On the other hand, we took advantage of protein glycosylation and ubiquitinylation to develop a two-component proteolysis targeting chimeras (PROTACs) strategy, which was proved to be able to selectively target *O*-GalNAcylated and *O*-GlcNAcylated proteins for proteasomal degradation. As a result, the critical metabolic and signaling pathways governed by glycoproteins were heavily perturbed, triggering severe cytotoxicity in human cancer cell lines. The approach termed GlyTACs leveraged from metabolic incorporation of easily accessible and cell-permeable peracetylated *N*-acetylglucosamine (GlcNAc) or *N*-acetylgalactosamine (GalNAc) mimics bearing an azide group into glycoproteins. In the living cells, the azido-modified glycoproteins served as covalent anchors for the introduction of thalidomide moiety by strain-promoted azide-alkyne cycloaddition (SPAAC) to recruit E3 ligase cereblon, resulting in stepwise ubiquitinylation of 'sensitized' proteins and their degradation by proteasome. The efficiency of the GlyTAC system was shown in a series of human cancer cell lines and the mechanistic pathway was verified by performing control experiments at each stage of the process. Given the characteristic features of cancer cells including fast nutrient turnover, and overall increase of protein

glycosylation, as well as the low cytotoxicity of the individual components, this GlyTAC approach may open a feasible strategy in cancer therapy.

In summary, these studies have discovered unexplored reactivities and functionalities of protein PTMs, helping chemical biologists utilize bioorthogonal reactions by avoiding unwanted side reactions and introducing therapeutic potentials.

Zusammenfassung

Protein-Posttranslationale Modifikationen (PTMs) spielen entscheidende Rollen bei der Regulierung der Proteinfunktion, -stabilität, -lokalisierung und -interaktionen. Diese Modifikationen, einschließlich Phosphorylierung, Glykosylierung, Ubiquitinierung, Acetylierung und Methylierung, unter anderem, können enzymatische Aktivitäten aktivieren oder deaktivieren, subzelluläre Lokalisierung bestimmen, Protein-Protein-Interaktionen vermitteln und Proteine zur Degradation markieren. Durch die dynamische Veränderung der Proteineigenschaften ermöglichen PTMs den Zellen, schnell auf Umweltveränderungen zu reagieren und die Homöostase aufrechtzuerhalten, wodurch sie zu Prozessen wie Signaltransduktion, Immunantwort, Zellzyklus-Kontrolle und Apoptose beitragen.

Um Protein-PTMs zu untersuchen, verwenden Wissenschaftler eine Vielzahl von Methoden wie hochauflösende Tandem-Massenspektrometrie, die die präzise Identifizierung und Quantifizierung von PTMs durch die Analyse von Peptidfragmenten ermöglicht. Western Blotting, bei dem spezifische Antikörper verwendet werden, die modifizierte Reste erkennen, ist eine weitere gängige Methode zur Detektion und Analyse von PTMs. Diese Methoden, kombiniert mit bioinformatischen Werkzeugen, bieten umfassende Ansätze zum Verständnis der komplexen Rollen von PTMs in der Proteinregulation. Wir begannen, eine neu entdeckte PTM namens Adenylylierung (AMPylierung) zu untersuchen, indem wir halogenmodifizierte AMPylierungs-Mimetika für anschließende Kreuzkupplungen über Suzuki-Miyaura-Reaktionen in lebenden Zellen entwickelten. Die nachfolgende Chemisch-biologische Analyse unter Verwendung funktionalisierter Aryl-Pinacol-Boronate zeigte jedoch eine unbekannt Reaktivität. Beispielsweise führte die Inkubation zwischen Zelllysaten und einem fluoreszierenden Aryl-Pinacol-Boronat zu einer signifikanten Markierung, unabhängig von der Zugabe künstlicher AMPylierungssonden oder Palladiumkatalysatoren. Zu unserer Freude stellten wir fest, dass die unbekannt Reaktivität auf eine selektive Proteinmarkierung (Disulfid-S-Arylierung) durch Ammoniumpersulfat (APS) zurückzuführen war. Anschließend wurden oxidiertes Glutathion und zwei weitere kurze Peptide verwendet, um die Bildung des S-Arylierungs-Produkts zu verifizieren, das anschließend in hochauflösendem MS2 fragmentiert wurde, um diagnostische Ionen zu identifizieren. Der radikalbasierte Mechanismus dieses Reaktionswegs wurde durch Experimente mit dem Spin-Trap 5,5-Dimethyl-1-pyrrolin-*N*-oxid (DMPO) bestätigt, und die Reaktion verlief nur, wenn Arylteilstrukturen mit elektronenspendenden Gruppen substituiert waren, was mit Wettbewerbsstudien in Zelllysaten übereinstimmte. Diese Methode war die erste Studie, die eine selektive Disulfid-S-Arylierung mit Arylradikalen unter milden wässrigen Bedingungen durchführte, ohne dass ein Metallkatalysator oder Photokatalyse erforderlich war.

Andererseits machten wir uns die Vorteile der Glykosylierung und Ubiquitinierung von Proteinen zunutze, um eine Zwei-Komponenten-Proteolyse-Zielchimaären-(PROTACs)-Strategie zu entwickeln, die nachweislich in der Lage war, O-GalNAcylierte und O-GlcNAcylierte Proteine selektiv für den proteasomalen Abbau zu Ziel zu nehmen. Infolgedessen wurden die kritischen Stoffwechsel- und Signalwege, die von Glykoproteinen gesteuert werden, stark gestört, was zu starker Zytotoxizität in menschlichen Krebszelllinien führte. Der Ansatz, der als GlyTACs bezeichnet wird, nutzte den metabolischen Einbau von leicht zugänglichen und zellgängigen peracetylierten *N*-Acetylglucosamin (GlcNAc) oder *N*-Acetylgalactosamin (GalNAc) Mimetika, die eine Azidgruppe tragen, in Glykoproteine. In lebenden Zellen dienten die azidmodifizierten Glykoproteine als kovalente Anker für die Einführung der

Thalidomidteilstruktur durch spannungsgeförderte Azid-Alkin-Cycloaddition (SPAAC), um die E3-Ligase Cereblon zu rekrutieren, was zu einer schrittweisen Ubiquitinierung der 'sensibilisierten' Proteine und deren Abbau durch das Proteasom führte. Die Effizienz des GlyTAC-Systems wurde in einer Reihe von menschlichen Krebszelllinien gezeigt und der mechanistische Weg wurde durch Kontrollversuche in jedem Stadium des Prozesses verifiziert. Angesichts der charakteristischen Merkmale von Krebszellen, einschließlich eines schnellen Nährstoffumsatzes und einer insgesamt erhöhten Proteinglykosylierung sowie der geringen Zytotoxizität der einzelnen Komponenten, könnte dieser GlyTAC-Ansatz eine machbare Strategie in der Krebstherapie eröffnen.

Zusammenfassend haben diese Studien unerforschte Reaktivitäten und Funktionalitäten von Protein-PTMs entdeckt und helfen chemischen Biologen, bioorthogonale Reaktionen zu nutzen, indem sie unerwünschte Nebenreaktionen vermeiden und therapeutisches Potenzial einführen.

1. Introduction

1.1 Protein Post-Translational Modifications (PTMs)

As the major performer of all biological activities, proteins are one of the most crucial components in living organisms. The collection of all proteins expressed by a genome, cell or tissue at defined conditions are termed as proteome, consisting of more than one million proteins (**Figure 1**).¹ Initially, human genome with approx. 20,000 coding genes are transcribed to roughly 100,000 transcripts through diverse and regulated mRNA splicing, which are subjected to subsequent translation process in ribosome to generate primary proteome containing more than 100,000 proteins (**Figure 1**).² Afterwards protein post-translational modifications (PTMs) continue to significantly expand the proteome complexity by covalently attaching modifying groups to one or more amino acids in proteins and by proteolysis.^{3,4} PTMs are proved to greatly influence the structures, properties and dynamics of proteins, including protein folding regulation, protein-protein interactions, receptor activation, cell signaling and trafficking, protein solubility, stability, localization and degradation, and other related metabolic processes.⁵⁻⁹ PTMs take place in various organelles such as nucleus, cytoplasm, endoplasmic reticulum (ER) and Golgi apparatus.^{6,9}

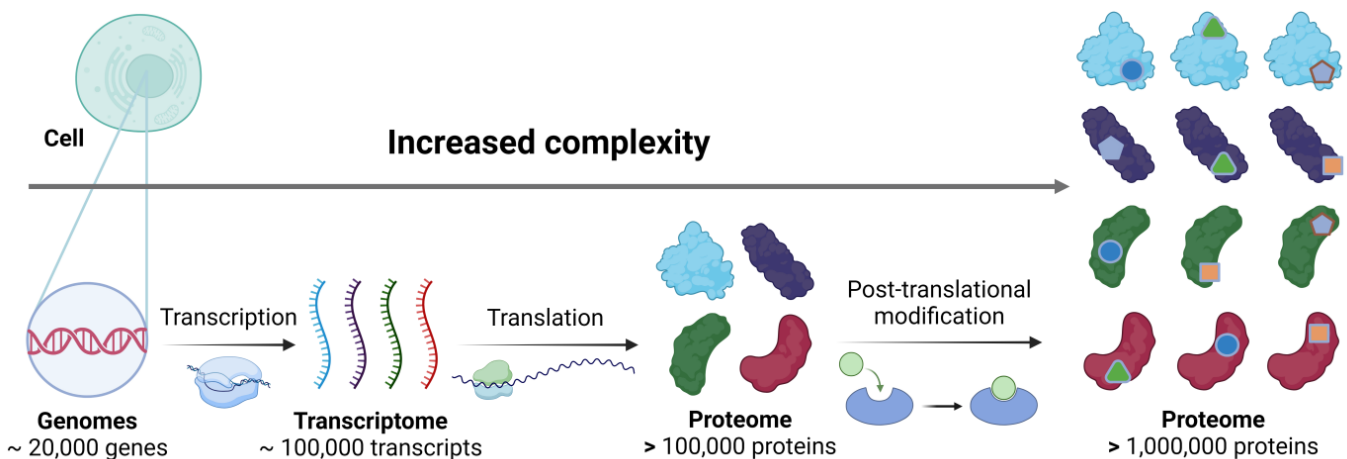


Figure 1. The proteome complexity by protein PTMs.

Generally, endogenous enzymes are mainly responsible for PTMs' occurrence and elimination, named as writer and eraser, respectively.^{10,11} Considering the chemical natures of protein PTMs, they are classified into reversible or irreversible modifications (**Figure 2**). Most common reversible modifications include phosphorylation,^{12,13} acetylation,^{14,15} methylation^{16,17} and sulfation,^{18,19}. More complex molecules such as lipids, oligosaccharides and adenosine monophosphate (AMP) could be also attached to target proteins to regulate their functions.²⁰⁻²⁴ PTMs involving small proteins like ubiquitinylation and SUMOylation (small ubiquitin-like modifier) are important for triggering downstream signaling pathway to maintain proper cell cycles.²⁵⁻²⁹ While a few examples of irreversible PTMs are disulfide formation between two cysteines,^{30,31} deamidation on asparagine or glutamine,³² and proteolysis to recycle amino acids building blocks.^{33,34} Of the above mentioned PTMs, O-glycosylation and ubiquitinylation are the two modifications of our interests, which are described in detail in § 1.2 and 1.7, respectively.

PTMs are crucial for regulating protein functions and interactions within cells. Abnormalities in PTMs are associated with a wide range of human diseases, including cancer, neurodegenerative disorders, and cardiovascular diseases.^{5,8,9} These aberrant modifications can disrupt normal cellular processes by altering protein activity, stability, or signaling pathways, thereby contributing to the onset and progression of human diseases.^{5,9} Understanding the role of PTMs in these conditions is key to identifying new biomarkers and therapeutic targets for treatment.

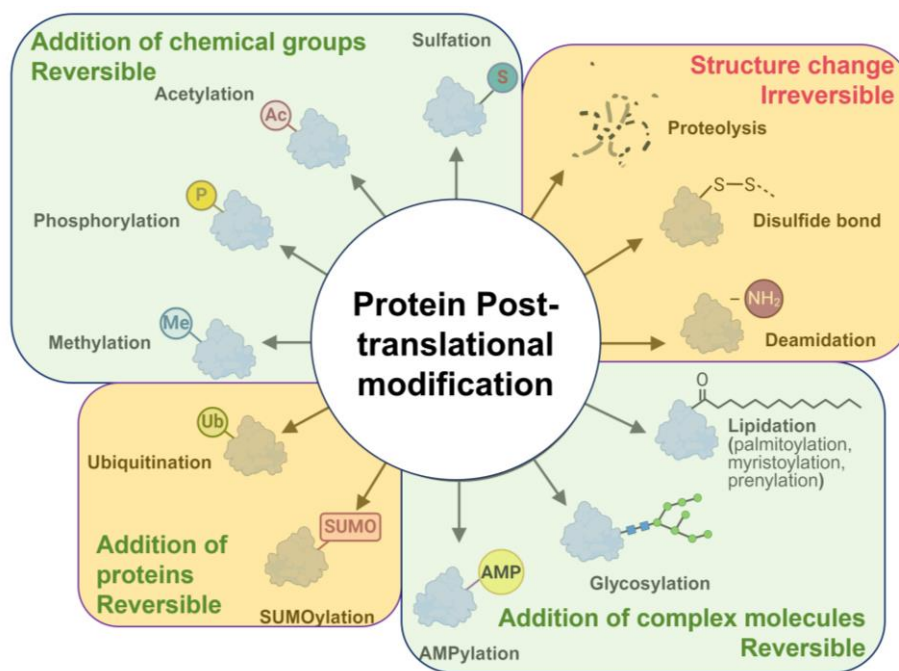


Figure 2. “PTMs clock” classified based on their chemical natures and reversibilities.

1.2 Protein O-GlcNAcylation and O-GalNAcylation

O-linked β -*N*-acetyl-glucosamine glycosylation (O-GlcNAcylation) is a post-translational and reversible modification on nucleocytoplasmic proteins to regulate their functions, localization, and stabilities.^{35–37} Dysregulation of O-GlcNAcylation cycle is involved in multiple human diseases, including diabetes, Alzheimer’s disease, and cancer.³⁶ The modification is characterized by the formation of a β -glycosidic bond and cannot be elongated for glycan biosynthesis.^{38,39} Specifically, the monosaccharide *N*-acetylglucosamine (GlcNAc) is attached or removed onto serine or threonine hydroxyl residues by a unique pair of endogenous enzymes O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA), respectively (**Figure 3A**).^{38–40} OGT utilizes uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) which is transformed from GlcNAc, as the substrate to target proteins, eliminating uridine diphosphate (UDP) (**Figure 3A**).^{38–40} Several small molecules, including Ac₄S-GlcNAc and OSMI, were developed to inhibit the global protein O-GlcNAcylation level (**Figure 3B**).^{41,42}

In order to target and detect O-GlcNAcylated proteins, a general two-step chemical biology method has been established.^{43,44} Azide-functionalized monosaccharides (**Figure 3C**) were firstly metabolically labelled in the living organisms, followed by pulling down modified proteins with azide-reactive reporters via bioorthogonal reactions. Subsequent readouts could be achieved through functional moieties inside reporters, including biotin, fluorescent tags and polyethylenglycol (PEG) units. More precisely, mass

spectrometry (MS) is able to offer definitive evidence of substoichiometric protein *O*-GlcNAcylation, based on prior enrichment steps realized by lectins, antibodies, or the above-mentioned chemical tagging. For example, M. Boyce et. al. has developed a phosphine-FLAG-His₆ probe to label azide-containing *O*-GlcNAcylated proteins in the cell lysates, followed by sequential anti-FLAG immunoprecipitation and immobilized metal affinity chromatography.⁴⁵ The enriched proteins were then subjected to proteolysis and tandem mass spectrometry measurements.

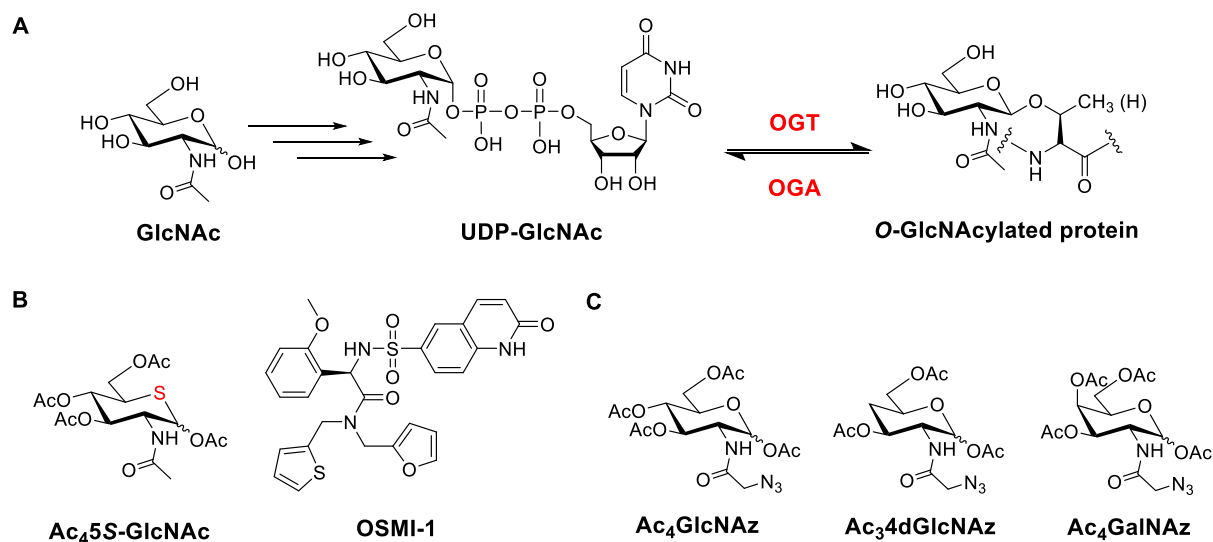


Figure 3. General scheme of protein *O*-GalNAcylation as well as its inhibitors and mimics. **A** Mechanism of protein *O*-GlcNAcylation on serine or threonine residues. **B** Structures of *O*-GlcNAcylation small molecule inhibitors. **C** Structures of *O*-GlcNAcylation mimics through endogenous metabolic labelling.

It has been proved that several cancers, including breast, colon, pancreas, liver, and lung cancers are associated with elevated *O*-GlcNAcylation.³⁶ For example, receptor tyrosine kinase (RTK) downstream pathway, such as KRAS can be activated by the upregulation of protein *O*-GlcNAcylation, leading to genomic alterations and KRAS mutations in pancreatic ductal adenocarcinoma (PDAC).⁴⁶ Moreover, several studies have shown that prostate and colorectal cancers, featuring high *O*-GlcNAcylation are correlated with worse prognosis.⁴⁷ Inhibiting *O*-GlcNAcylation in M2-type macrophage may also help to overcome its immune suppressive tumor microenvironment.⁴⁸ All these evidences point to that targeting *O*-GlcNAcylation could be a potential research interest within cancer therapy.

Similarly with protein *O*-GlcNAcylation, *O*-linked β -*N*-acetyl-galactosamine glycosylation (*O*-GalNAcylation) refers to the covalent and reversible linkage between monosaccharide *N*-acetylgalactosamine (GalNAc) and serine or threonine hydroxyl residues from target proteins.⁴⁹ The modification is initiated by the GalNAc-T enzyme families (ppGalNAc T1-T14, GALNT), recognizing uridine diphosphate *N*-acetylgalactosamine (UDP-GalNAc) as the substrate.⁴⁹

For targeting and detecting *O*-GalNAcylated proteins, H. C. Hang et. al. have designed and synthesized the envisaged azide-containing probe *N*-azidoacetylgalactosamine-tetraacetylated (Ac₄GalNAz, **Figure 3C**), which is transformed to corresponding UDP-GalNAc through enzymatic cascade.⁵⁰ Surprisingly, Ac₄GalNAz treatment resulted in more robust and abundant *O*-GlcNAcylation labelling compared with its counterpart *N*-azidoacetylglucosamine-tetraacetylated (Ac₄GlcNAz, **Figure 3C**), which might be explained from three aspects.⁴⁵ Firstly, the conversion from GlcNAc-1-P to UDP-GlcNAc by human pyrophosphorylase (AGX1/2) in the GlcNAc salvage pathway is rate-limiting for UDP-GlcNAc

biosynthesis, which collaborated with the observed relatively low labelling of protein O-GlcNAcylation by Ac₄GlcNAz (**Figure 4A**).⁵¹ Conversely, AGX1/2 transformed GalNAc-1-P to UDP-GalNAc with much better efficiency (**Figure 4A**).⁴⁵ The last reason is that UDP-GalNAc could be reversibly interconverted to its C₄ epimer, UDP-GlcNAc by the enzyme UDP-galactose 4'-epimerase (GALE) (**Figure 4A**).⁵² The GalNAc and GlcNAc salvage pathways overlap with Leloir pathway and hexosamine biosynthesis pathway, indicating complex sugar metabolism and the importance to design suitable probes for characterizing protein O-GlcNAcylation and O-GalNAcylation (**Figure 4A**).³⁷

Distinguishing O-GlcNAcylation and O-GalNAcylation is challenging for MS analysis, due to same molecular weight of GlcNAc and GalNAc. Chemoenzymatic methods using galactose oxidase (GAO) to selectively oxidize GalNAc to aldehyde, followed by imine formation with methoxyamine on immobilized solid support can effectively retain O-GalNAcylated proteins only (**Figure 4B**).⁴⁹ Then MS would be able to recognize and analyze these proteins of interests.

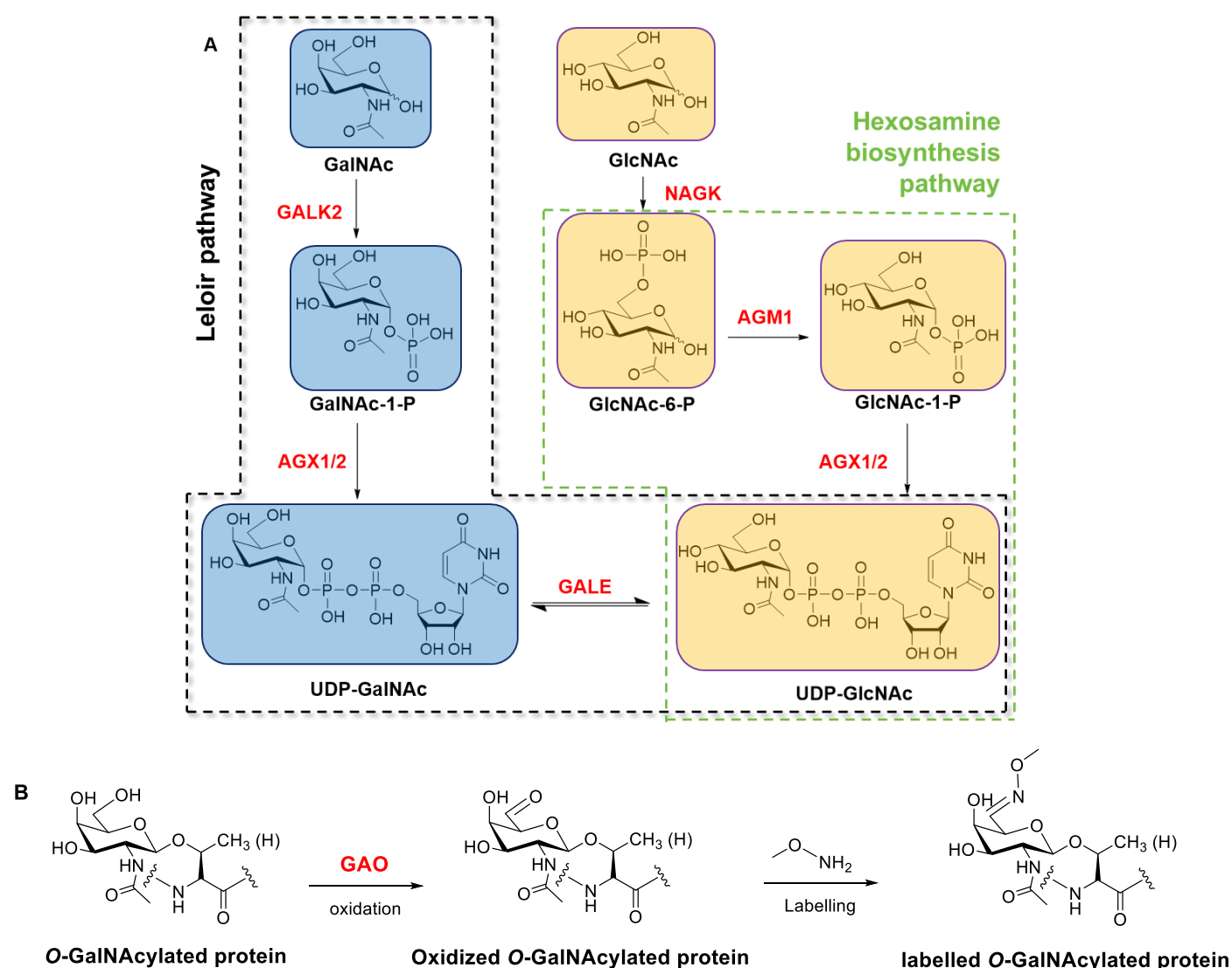


Figure 4. The GalNAc and GlcNAc salvage pathways; chemoenzymatic methods for MS analysis. **A** The two salvage pathways with enzymes shown in red. GalNAc and GlcNAc species are highlighted in blue and yellow, respectively. Leloir pathway and hexosamine biosynthesis pathway are enclosed in black and green dashed lines, respectively. **B** GAO oxidizes only O-GalNAcylated but not O-GlcNAcylated proteins, resulting in differential labeling between two modifications for MS identification.

O-GalNAcylation is also called mucin-type glycosylation because most mucins are heavily modified with GalNAc, which is related with many adenocarcinomas, including pancreatic, lung, and breast cancers.⁴⁹ Non-further modified GalNAc α 1-Serine group is termed as Tn antigen, and its strong upregulation is often associated with multiple oncogenic features, including extensive cell proliferation, decreased apoptosis, increased migration and invasion.⁴⁹

Taken together, protein O-GlcNAcylation and O-GalNAcylation are crucial PTMs for maintaining cell homeostasis for many downstream signaling pathways. That's why eliminating abnormal upregulation of these two PTMs promptly is regarded as a promising research area in cancer research.

1.3 Click Chemistry

The Nobel Prize in Chemistry 2022 was awarded to three chemists, including Carolyn R. Bertozzi, Morten Meldal and K. Barry Sharpless to commend their contributions "for the development of click chemistry and bioorthogonal chemistry". The most well-known click reaction, copper(I)-catalyzed azide-alkyne cycloaddition (**CuAAC, Figure 5A**), was firstly reported by Morten Meldal and K. Barry Sharpless and renowned for its high yield and regioselectivity for triazole formation.^{53,54} Shortly after Carolyn R. Bertozzi proposed a similarly simple and fast "**Staudinger ligation**" feasible in living organisms (**Figure 5A**), using azide with triarylphosphine to form an amide bond.^{55,56}

To avoid the cytotoxicity of copper catalysts *in vivo*, Carolyn R. Bertozzi has developed the strain-promoted azide-alkyne cycloaddition (**SPAAC, Figure 5A**), in which the triazole was generated by releasing the cyclooctyne (OCT) ring tension in the presence of a dipole (usually a azide group).⁵⁷⁻⁵⁹ The alkyne moiety is often designed in a strained bicyclo[2.1.0]nonyne (BCN), difluorooctyne (DIFO), dibenzocyclooctyne (DBCO) or biarylazacyclooctynone (BARAC) (**Figure 5B**).⁶⁰⁻⁶² On the other hand, dipoles other than azide, including nitrones, diazos and nitrilimines have been found to show much faster reaction rates in similar cycloadditions, promoting many successful applications in the living systems (**Figure 5C**).^{63,64}

In 2008, the fastest click reaction so far was reported, which is carried out between a diene 1,2,4,5-tetrazine and a strained alkene or alkyne dienophile (usually a transcyclooctene, TCO), termed as **tetrazine ligation (Figure 5A)**.^{65,66} The mechanism is regarded to be firstly an inverse-electron demand Diels-Alder reaction (IEDDA), followed by retro-Diels-Alder reaction to eliminate nitrogen gas as the only by-product (**Figure 5D**).⁶⁷ By placing electron-withdrawing groups on tetrazine or finely adjusting the size or reactivity of dienophiles (**Figure 5E**), the tetrazine ligation can be highly efficient with an observed rate exceeding $k_2 = 10^6 \text{ M}^{-1}\text{s}^{-1}$.^{68,69} One particular application is to establish strong fluorescence of tetrazine upon cycloaddition with dienophiles, leading to growing interest to improve signal-to-noise ratio in live-cell microscopy.^{70,71}

Besides the above-mentioned click chemistry, the other reactions developed for bioorthogonal applications include oxime and hydrazone ligations,^{72,73} native chemical ligation,^{74,75} photoinduced click reaction⁷⁶ and our envisaged palladium-catalyzed cross-couplings described in § 1.4. Overall most of these click reactions share common advantages such as the use of safer starting compounds and solvents, blockage of hazardous waste and by-products, minimization of derivatization and energy consumption, better selectivity, stability and tolerance, etc.⁷⁷

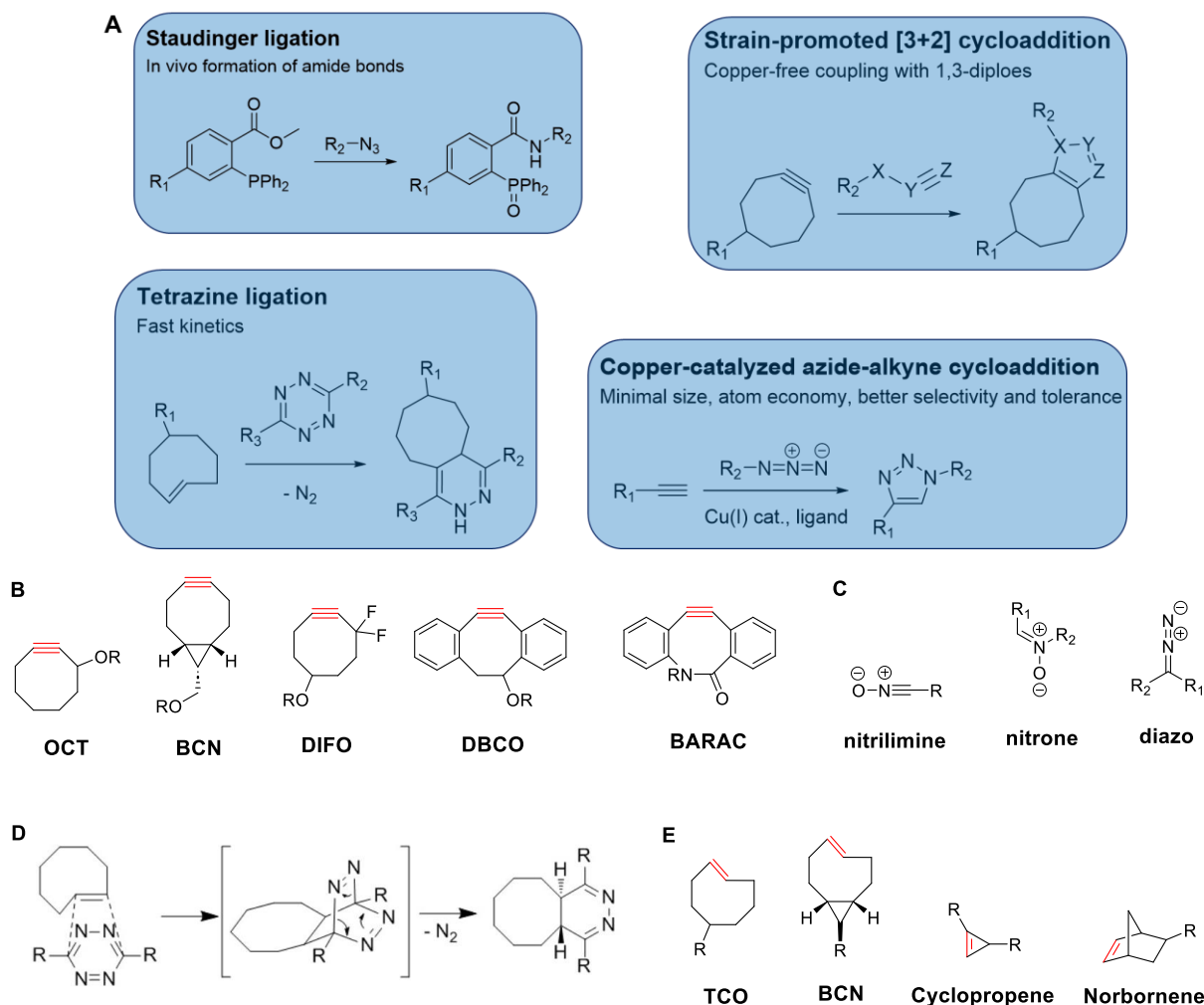


Figure 5. General schemes for click chemistry and variations. **A** Most-frequently-used click chemistry reactions. **B** Examples of OCTs' structures used in SPAAC reactions. **C** Examples of dipoles' structures used in SPAAC reactions. **D** Mechanism of tetrazine ligation. **E** Examples of dienophiles' structures used in tetrazine ligation.

1.4 Protein Bioorthogonal Labelling via Suzuki Cross-Coupling Reaction

The Suzuki cross-coupling is a palladium-catalyzed reaction between aryl halides (R^1-X) and aryl organoborons (R^2-BY_2), leading to a new carbon-carbon single bond formation (**Figure 6A**).⁷⁸ Since its discovery in 1979 by Suzuki Akira, the reaction has greatly expanded the toolbox for synthesizing natural products and functional materials due to its broad substrate tolerance. The reaction is initiated from rate-determining oxidative addition (OA) step from palladium⁽⁰⁾ to palladium^(II) in the presence of halides (R^1-X), followed by transmetalation (TM) with organoborons (R^2-BY_2) to afford transient organopalladium species in basic conditions (**Figure 6A**).⁷⁹⁻⁸¹ Afterwards reductive elimination (RE) takes place to release the desired product (R_1-R_2) and restore the palladium⁽⁰⁾ catalysts, finalizing the catalytic cycle (**Figure 6A**).^{79,81} The bulky and electron-rich ligands such as phosphines and *N*-heterocyclic carbenes (NHCs) are used as ligands.^{82,83}

Besides its huge contributions in organic chemistry, Suzuki reaction has been successfully applied for

biomolecule functionalization and labeling in the living cells. Benjamin G. Davis group has developed a water-soluble (2-amino-4,6-dihydropyrimidine)-based palladium catalyst system ($[\text{Pd}(\text{OAc})_2(\text{ADHP})_2]$) (**Figure 6B**), allowing almost full conversion of a broad range of aryl-/vinyl-boronic acids with aryl iodide-modified proteins.⁸⁴ Recently they took advantage of this homogeneous catalyst to tag the bacterial pore protein OmpC by staining iodophenylalanine-incorporated cell-surface protein with a fluorescent aryl boronic acid.⁸⁵ Similarly, W. Chen group reported a hydrophilic NHC-stabilized palladium catalyst (**Figure 6C**), labelling surface proteins bearing unnatural arylhalide-modified amino acids in mammalian cells.⁸⁶ While M. Bradley group has creatively encapsulated catalytic palladium nanoparticles within inert polystyrene microspheres, which were then transported inside the cells without notable cytotoxic effects (**Figure 6D**).⁸⁷ The catalytic activity was confirmed by a mitochondrial fluorophore and an anti-cancer reagent both synthesized *in cellulo*.⁸⁸

The other homogeneous palladium catalyst was prepared in K. M. Backus group, by mixing palladium acetate with very bulky and electron-rich phosphine ligand, which contains a hydrophilic sodium sulfonate group to largely increase complexes' solubility in aqueous medium (**Figure 6E**).^{89,90} The catalyst was proved to be effective for triggering expected Suzuki cross-coupling reaction in PBS buffer, and was compatible with bioadditives like bovine serum albumin (BSA) and HEK293T cell lysates.⁸⁹ Moreover, it was further applied to pull down iodo-functionalized proteins with biotinylated phenyl boronic acid enrichment handles. The chemical proteomic results indicated that the proteins identified overlapped quite well with commonly-used copper(I)-catalyzed azide-alkyne cycloaddition methods (CuAAC, see § 1.3).⁸⁹ These exciting discoveries have inspired and motivated us to explore more applications of Suzuki cross-coupling reaction in chemical biology areas.

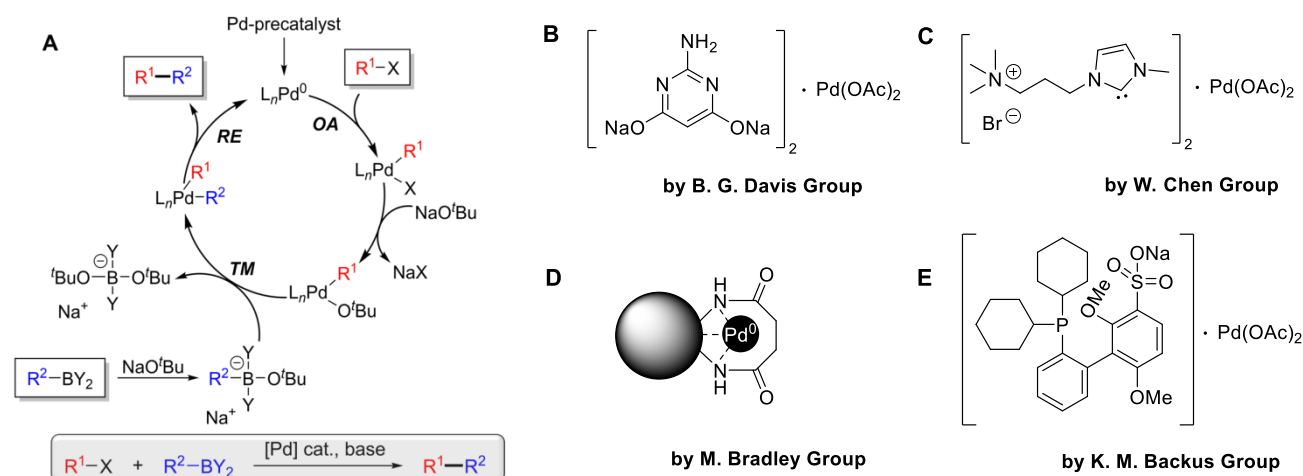


Figure 6. Suzuki cross-coupling reaction and applications in biomolecule functionalization and labeling. **A** Mechanism of Suzuki cross-coupling reaction. Sodium *tert*-butoxide (NaO^tBu) was shown as the base involved in transmetalation step. *OA*: oxidative addition, *TM*: transmetalation, *RE*: reductive elimination. **B** Homogeneous $[\text{Pd}(\text{OAc})_2(\text{ADHP})_2]$ catalyst. **C** Hydrophilic NHC-stabilized palladium catalyst. **D** Heterogeneous palladium nanoparticles trapped in polymers. **E** Homogeneous palladium catalyst for chemical proteomic applications.

1.5 Aryl Radicals: Generation and Applications

In modern organic chemistry, the aryl radicals have been proved to be powerful and versatile synthons in functional group conversions.⁹¹ Previous methods for aryl radical generation such as Meerwein arylation and Sandmeyer reaction, however, relied on stoichiometric transition metal catalysts and therefore were not efficient enough.⁹² Contemporary methods including photo- or electrochemistry have largely promoted the feasibility and applicability of aryl radicals from many substrates with broad scopes (**Figure 7A**).^{93,94} The generated aryl radicals could be functionalized with halogenation, reduction, borylation, and further arylation, etc (**Figure 7A**).

The Li group applied single electron transfer (SET) under UV irradiation between the soft iodide anion and aryl triflate, leading to the desired C-O bond cleavage (**Figure 7B**).⁹⁵ The methodology overcame the unwanted S-O bond cleavage and opened up the possibility for efficient synthesis of aryl pinacol boronates bearing electron-donating (EDGs) or electron-withdrawing groups (EWGs).⁹⁶ Another example was to perform Sandmeyer halogenation through electrochemistry way, which was developed by F. Mo group. When using platinum material for cathode and anode, the expected bromination went through very well with *N*-bromosuccinimide (NBS) as the halogen source (**Figure 7C**).⁹⁷ They have further performed electro-driven iodination and challenging chlorination in the presence of diiodomethane (CH₂I₂) and lithium chloride (LiCl), respectively.⁹⁷ C. Stephenson et. al. have used strongly reductive iridium based photocatalyst *fac*-Ir(ppy)₃ to selectively reduce vinyl and aryl iodides (**Figure 7D**).⁹⁸ The reaction was carried out through SET between iridium catalyst and aryl iodides, generating aryl radicals which were captured and quenched by hydrogen from tributylamine *via* intermolecular hydrogen atom transfer (HAT) process.⁹⁸ The other major approach to generate aryl radicals is to form electron donor-acceptor (EDA) complex, exhibiting absorption band not featured for either donor or acceptor. One typical example was from B. König group, who have solubilized aryl halides (electron acceptor) and derivatives from aniline, phenol, and thiophenol (electron donor) in water (**Figure 7E**).⁹⁹ The overall solubility of organic compounds were significantly improved by the addition of a pharmaceutical excipient meglumine (**Figure 7E**).¹⁰⁰⁻¹⁰² The SET was then triggered upon irradiation at 385 nm, generating radical cation and anion pair; while the latter one eliminated a halide anion, resulting in a neutral aryl radical and subsequent conjugate formation from two radical intermediates (**Figure 7E**).⁹⁹ It is notable that the appropriate functionalities within both donor and acceptor could be adjusted to let their orbitals overlap better for higher yields.⁹⁹

Although aryl radicals have been commonly used for organic synthesis, its application in chemical biology, for example, protein functionalization targeting specific residues was not well studied so far. Therefore, it would be very useful to further investigate such possibility to enlarge the protein modification toolbox of biochemists.

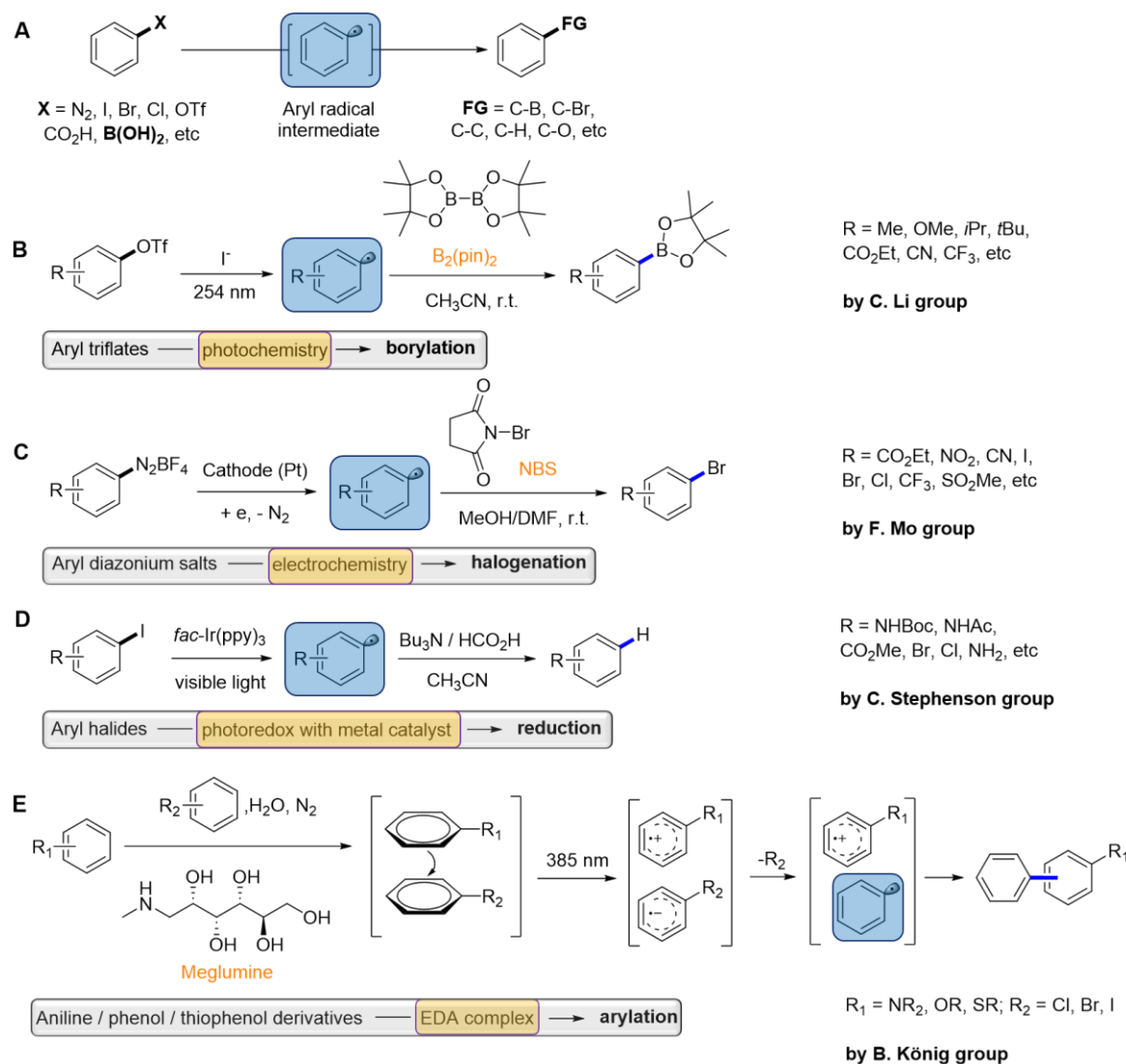


Figure 7. Contemporary methods for aryl radical generation and application. **A** General scheme for aryl radical generation and application. **B** Aryl borylation from aryl triflates. **C** Aryl halogenation from aryl diazonium salts. **D** Aryl reduction from aryl halides. **E** Aryl arylation via EDA complex formation. Aryl radicals and methods description are highlighted in blue and yellow, respectively.

1.6 Organoboranes in Radical Initiation or Termination

Organoboranes are an important class of compounds in organic synthesis due to the characteristic vacant p -orbital on boron, making them tunable Lewis acids and easily reactive with nucleophiles.¹⁰³ The reactivity of boranes was reported by Davies and Roberts as the following order: $R_3B > R_2BOR > RB(OR)_2$ with $R = \text{alkyl}$.^{104,105} The trialkylboranes are especially active and thus can react with compounds which are not commonly regarded as nucleophiles, such as oxygen.^{105,106} The formed boronate complexes are tetracoordinated with a sp^3 -hybridized central boron and are crucial for downstream reactivities. For example, trialkylboranes can go through autooxidation in the presence of oxygen or other oxygen-centered radicals, generating alkyl radicals via concerted homolytic substitution (S_H2) mechanisms (**Figure 8A**).¹⁰⁶ The alkyl radicals would further react with its surrounding species, realizing many functionality conversions, such as the Add-and-Done Desulfurization (ADD) protocol developed by X. Li group.¹⁰⁷ Later P. Renaud group proved that B -alkylcatecholboranes (R -Bcats) are able to form

stable perboryl radical intermediates due to electron delocalization onto the aromatic ring (**Figure 8B**).¹⁰⁸⁻¹¹⁰ Therefore, R-Bcats have been considered to be a highly-efficient resource of alkyl radicals. However, *B*-alkylpinacolboranes (R-Bpins), featured by better stability and processability are regarded as radical-inactive and not suitable for generating alkyl radicals. To overcome such problem, they have recently reported an *in situ* conversion from R-Bpin to R-Bcat by boron-transesterification with a substoichiometric amount of catechol methyl borate.¹¹⁰ It was also reported that *N*-alkylpyridinium salts are capable to complex with bis-(catecholato)diboron (B_2cat_2) to offer an EDA complex, enabling photoinduced SET for further chain transfer and fragmentation (**Figure 8C**).¹¹¹ V. K. Aggarwal group also pointed out that the EDA complex was externally stabilized by *N,N*-dimethylacetamide (DMA), releasing a carbon-centered radical which was subsequently borylated with B_2cat_2 (**Figure 8C**).^{111,112} This methodology was highlighted by efficient deaminative borylation at catalyst-free conditions, but still relying on electron delocalization on catechol moieties.

On the other hand, organoboranes masked by aromatic diols were not always tending to undergo B-C bond cleavage given a stronger bond dissociation energy; in this case the enhanced stability granted the possibility to trap the existed radicals, leading to termination of chain propagation. For example, D. A. Pratt group has created an aryl naphthenyl boronate with higher thermostability and thus less propensity to induce a B-C bond homocleavage (**Figure 8D**).¹¹³ The species were proved to be able to trap peroxy radicals ($ROO\cdot$), which were supposed to be reactive to release carbon-centred radicals $R\cdot$ (**Figure 8D**).¹¹³ The strategy has further expanded the versatilities of organoboranes to function as a new class of radical-trapping antioxidants. Moreover, aryl radicals, as intermediates, were capable to convert organoboron prodrugs to anti-tumor compounds *in vivo*. This proof-of-concept was studied by T. Zou group, who have photo-activated boronic acid-caged iridium(III) complex into bioactive $IrNH_2$ to introduce damage to mitochondrial DNA in cancer cell lines (**Figure 8E**).¹¹⁴⁻¹¹⁷ Mechanistic study demonstrated that the conversion was initiated by a photosensitizer with red absorption (methylene blue), leading to an aryl radical formation and subsequent oxygen capture to afford arylperoxy radicals.^{114,118} Some other aryl pinacol boronate-caged prodrugs were prepared and photoactivated in the same way, indicating the utility of aryl radicals also in biochemistry area (**Figure 8E**).^{114,119,120} Recently, aryl potassium trifluoroborate salts have been transformed into aryl phosphine through pulsed electrosynthesis (**Figure 8F**).¹²¹ The method, developed by C. A. Malapit group has successfully overcome common challenges in electrosynthesis such as aryl radical homocoupling, overoxidation, and decomposition.^{122,123} The authors have further extended its broad utilities by synthesizing C-S, C-Te, and C-Se bonds.¹²¹

Taken together, organoboranes can form carbon-centered alkyl or aryl radicals via autooxidation, EDA complex, photochemistry, or electrochemistry processes. Such versatilities have granted organoboranes great potentials as radical precursors to trigger diverse conversions and functionalizations onto target molecules.

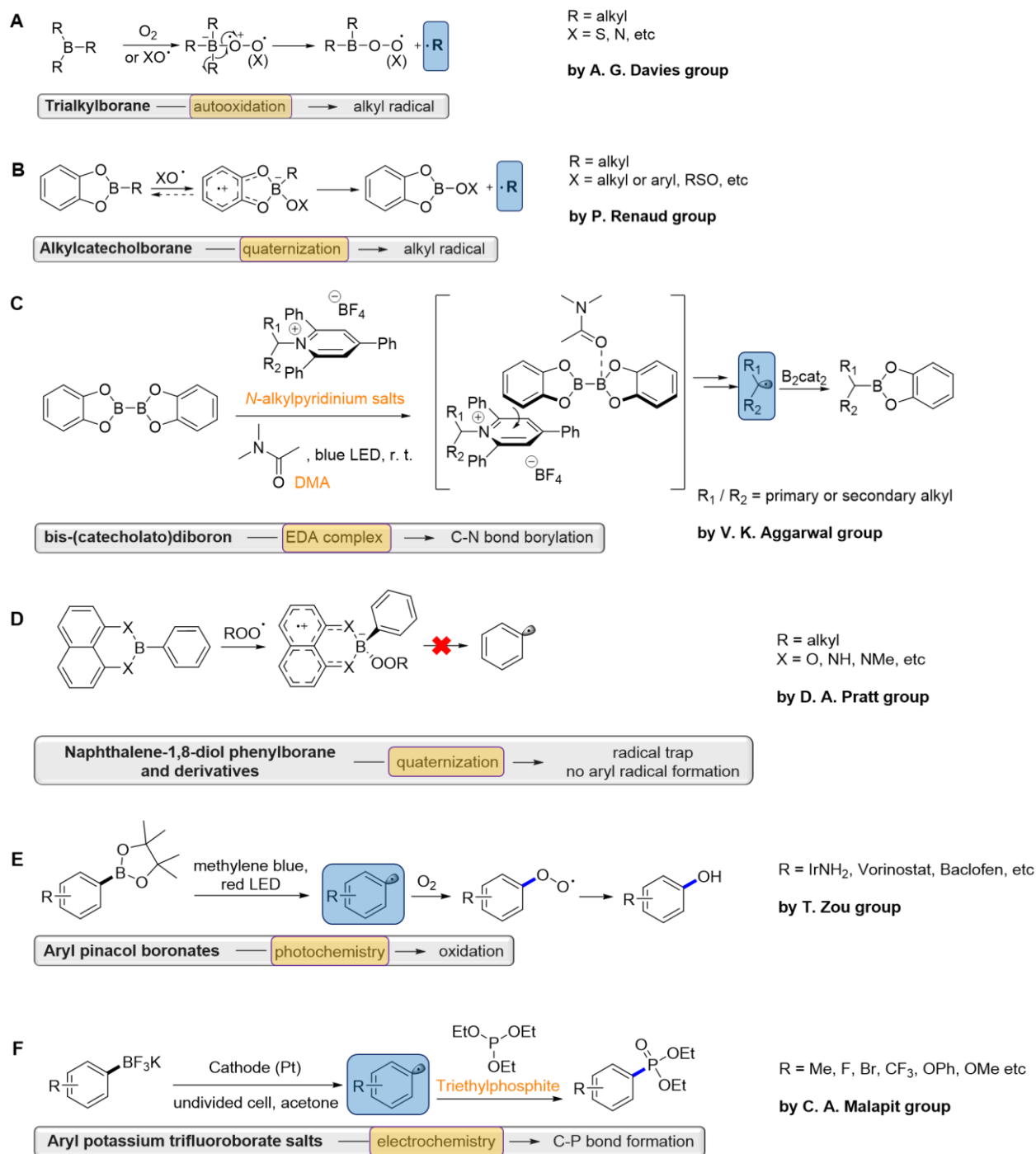


Figure 8. Radical reactions of alkyl- and aryl-substituted boron. **A** Carbon-centered alkyl radicals from electrophilic trialkylboranes by oxygen autooxidation. **B** Carbon-centered alkyl radicals from R-Bcats under mild conditions. **C** Borylation of *N*-alkylpyridinium salts by DMA-stabilized EDA complex formation. **D** Organoboranes as radical-trapping antioxidants. **E** Photo-activation of R-Bpins to offer arylperoxyl radicals for drug release. **F** Pulsed electrochemistry of aryl phosphine from aryl potassium trifluoroborate salts.

1.7 Cysteine S-Arylation Chemistry

Within 20 canonical amino acids, cysteine is the only one containing thiol group (-SH), bringing the possibility of selectively modification in the presence of other residues.¹²⁴ Besides the very rare selenocysteine, thiol in cysteine is the most acidic residue compared with hydroxyl groups in serine or

tyrosine, and ϵ -amine of lysine (**Figure 9A**).¹²⁵ In combination with the moderate electronegativity and higher polarizability of sulfur, thiol becomes the most reactive nucleophile in protein primary structures.^{124,126} Therefore, chemical biologists have developed many reagents and methods to realize selective cysteine S-arylation, generating a sp^2 -hybridized C-S bond and forming functionalized protein modulators, macrocyclic peptides, antibody-drug conjugates (ADCs), and other hybrid protein-based constructs.^{127,128}

The nucleophilic aromatic substitution (S_NAr) has been widely applied for cysteine S-arylation. The reaction takes place when incubating proteins or peptides with arylation reagent (often aryl halides) containing an EWG such as nitro to enhance its electrophilicity and stabilize the Meisenheimer intermediate complex (**Figure 9B**).¹²⁹ Afterwards halogen anion leaves and the aromaticity is restored (**Figure 9B**).¹²⁹

The E. Weerapana group has designed and synthesized a highly cysteine-selective *p*-chloronitrobenzene, which can label proteins at low micromolar concentrations (**Figure 9B**).¹³⁰ The alkyne handle has allowed chemical proteomic analysis to identify labeling sites. M. Xian et. al. have developed a selective and efficient thiol-blocking reagent, methylsulfonyl benzothiazole (MSBT), without labelling serine, tyrosine, lysine, histidine, methionine or tryptophan (**Figure 9B**).¹³¹ pH-Dependent studies showed the blocking reaction could be finished within just 5 minutes at pH 9. Thus the reagent might be considered to be an alternative for the commonly-used iodoacetamide and maleimide.¹³¹ In 2013, commercial decafluorobiphenyl reagents were firstly found to react rapidly and selectively with thiol groups in polar organic solvents (**Figure 9B**).¹³² The reaction has been applied to create macrocyclic or stapled peptides with disubstituted thiols, or peptide homodimers featuring perfluoroaryl linkers.^{132–134} These modified peptides usually possess better biological properties such as stability against proteases and binding affinity to the target.¹³⁴ The R. Derda group has further addressed the poor water solubility of perfluoroaryl reagents by inserting sulfonyl moiety into decafluorobiphenyl, and performed similar S-arylation modification in aqueous medium (**Figure 9B**).¹³⁵

Another method for carrying out cysteine S-arylation was based on delicate design of transition metal catalysts, such as palladium organometallic complexes with biarylphosphine ligands (**Figure 9C**). For example, B. L. Pentelute group and S. L. Buchwald group have developed solid and storable oxidative-addition complexes [(RuPhos)Pd(Ar)X] from aryl halide and palladium-tolyl precursors, allowing fast, selective and efficient cysteine S-arylation in benign conditions (**Figure 9C**).^{136,137} More studies showed that the formed S-(hetero)arylated products were more stable under basic, acidic, and oxidative conditions, when compared with maleimide and acetamide analogues.¹³⁶ The oxidative-addition complexes could be comprehensively functionalized by introducing fluorescent dyes, affinity tags, drug molecules and bioconjugation handles into aryl moieties, which will be conjugated with biomolecules afterwards (**Figure 9C**).¹³⁶ Moreover, S. Messaoudi et. al. proved that the selective cysteine S-arylation could be realized between short peptides and aryl iodides, in the presence of catalytic amount (2 mol %) of XantPhos-based aminobiphenyl mesylate palladium pre-catalysts (Pd-G₃-XantPhos, **Figure 9C**).¹³⁸ Although there were not many examples, transition metal-mediated S-arylation is still a complementary approach to S_NAr reaction.

In 2015, B. L. Pentelute group disclosed an umpolung strategy for selective Se-arylation in unprotected peptides (**Figure 9D**).¹³⁹ Due to high acidity of selenols and greater polarizability of selenium, selenocysteine is considered to be an attractive target for protein modification.¹⁴⁰ Within their study, the

nucleophilic selenocysteine was firstly oxidized with disulfides to get selenoelectrophile, which was exposed to oxidative addition when subjected to copper catalysts with ligands.¹³⁹ The complex then went through transmetalation with boronic aryl nucleophiles, followed by reductive elimination to release the Se-arylated products.¹³⁹ The boronic acids gave very satisfying yields regardless of the substituents, showing broad substrate tolerances.¹³⁹ This strategy could be inspiring for constructing more creative methodologies regarding selective and efficient protein S-arylation.

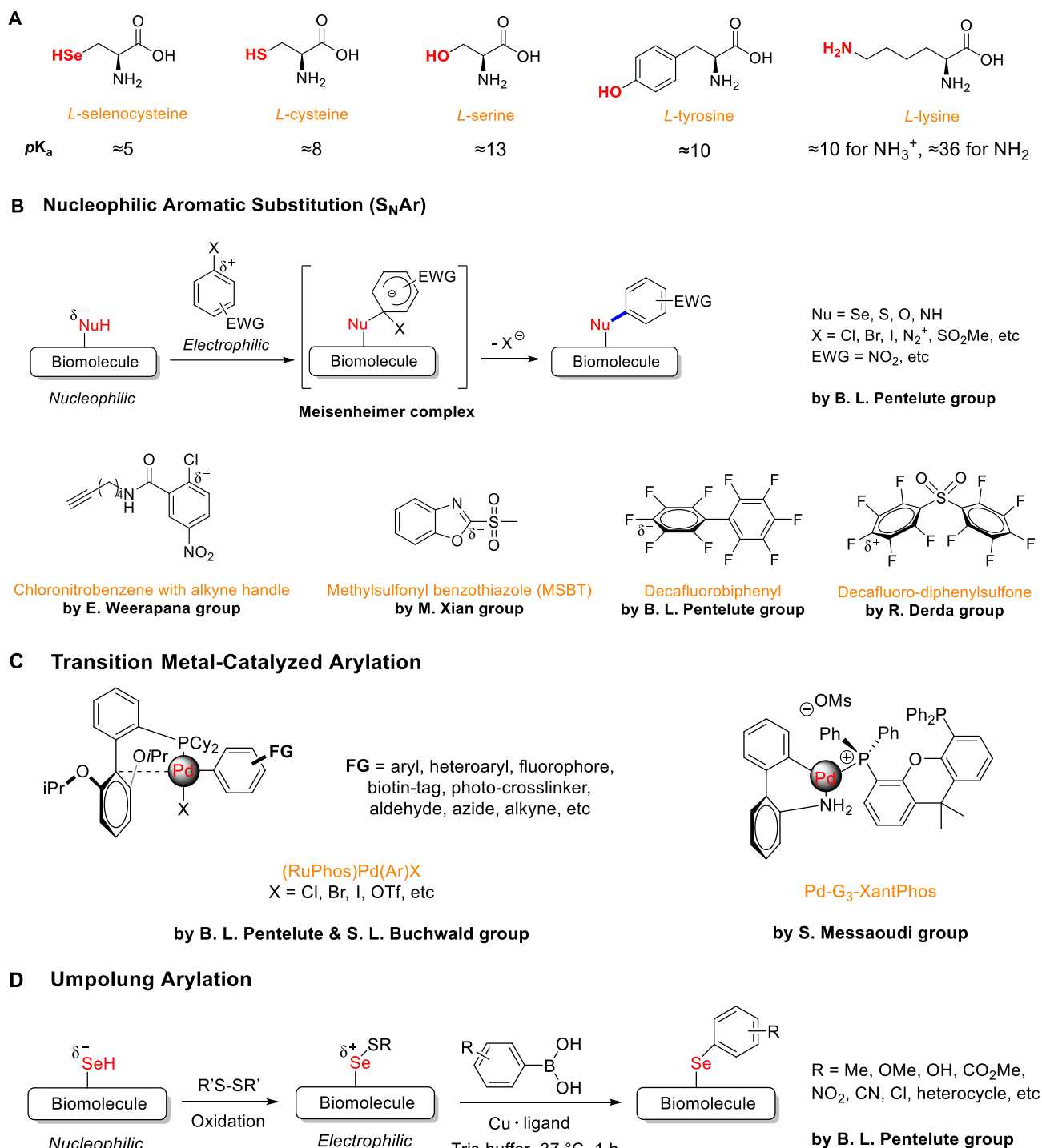


Figure 9. Cysteine S-arylation chemistry. **A** List of pK_a values of amino acids' residues. **B** Mechanism of S_NAr reaction and electrophile structures used for selective cysteine S-arylation in biomolecules. **C** Structures of oxidation-addition complex and palladium pre-catalysts for selective cysteine S-arylation in biomolecules. **D** Umpolung strategy for selective selenocysteine Se-arylation in biomolecules.

1.8 Proteolysis Targeting Chimeras (PROTACs)

A **proteolysis targeting chimera** (PROTAC) is a heterobifunctional molecule comprised of two active domains and a linker in between, capable of degrading specific unwanted proteins in the living cells.^{34,141,142} As an emerging drug candidate for therapeutic purposes, PROTAC was firstly described by Kathleen Sakamoto, Craig Crews and Ray Deshaies in 2001, and has been applied for targeting various oncoproteins.^{141,142}

Nowadays the core of most cancer therapies relies on the occupancy-based mechanism of action on oncoproteins by small molecule inhibitors, often requiring high administrative dosages and thus leading to undesired off-target toxicities.¹⁴³ What is more challenging comes from the natures of many oncoproteins, including nonenzymatic functions, lack of hydrophobic pockets, and high affinities for abundant substrates.^{144,145} Therefore the PROTAC strategy based on target protein degradation (TPD) was proposed to circumvent these therapeutic impasses.

1.8.1 Protein Ubiquitinylation and Principles of PROTAC

Protein degradation is mostly performed by lysosome or proteasome and is crucial in many physiological processes such as protein level regulation, cell cycle control, and cellular stress response.³⁴ For proteasome-involved degradation, the proteins of interest (POIs) are firstly labelled covalently with ubiquitin, a highly conserved protein containing 76 amino acids (8.6 kDa).¹⁴⁶ The ubiquitinylation process is efficiently carried out through a cascade of three enzymes. To start with, a ubiquitin is mono-adenylated by an E1 activation enzyme in the presence of ATP and magnesium ion, followed by the thioesterification to an E1 enzyme-ubiquitin intermediate (**Figure 10A**).^{147,148} Then ubiquitin is then transferred to the catalytic cysteine of an E2 conjugation enzyme through a transthioesterification reaction (**Figure 10A**).^{147,148} Subsequently, the ubiquitin bound to E2 enzyme is covalently connected to the target proteins within the E2 enzyme-E3 ligase-target protein ternary complex through an isopeptide bond between the ubiquitin C-terminus (glycine) and the primary amine from a lysine side chain (ϵ -NH₂) on the target proteins' surfaces.¹⁴⁷⁻¹⁴⁹ The target protein can be either a POI (mono-ubiquitinylation / ubiquitin chain initiation) or another ubiquitin at K6, K11, K27, K29, K33, K48, or K63 sites (ubiquitin chain elongation), of which K48-linked polyubiquitinylated proteins are recognized by the 19S regulatory subunit of the 26S proteasome.^{150,151} Finally, POIs are transported into the 20S core catalytic subunit, being unfolded, deubiquitinylated, and proteolytically degraded.^{148,149} These overall processes are termed as ubiquitin-proteasome system (UPS), functioning as the major game player in degrading and eliminating unneeded proteins. Depending on the E3 ligase, ubiquitin can be transferred directly (RING) from E2 enzyme to the target protein or sequentially (HECT) from E2 enzyme to E3 ligase through a thioester, and then to the target protein (**Figure 10A**).^{148,150}

The PROTAC strategy hijacks the UPS to degrade specific POIs by chemically decreasing proximity between the POIs and E3 ligase, while the linker in between coordinates to an appropriate geometry of the ternary complex composed of the PROTAC, the POIs, and the recruited E3 ligase (**Figure 10B**).¹⁴¹⁻¹⁴³ The ubiquitin tags are then transferred from an E2 enzyme to the POIs as described above, which are degraded by proteasome afterwards (**Figure 10B**). Another advantage of PROTAC attributes to its

catalytic properties at substoichiometric concentrations, since the intact PROTAC molecules would be released from polyubiquitinated POIs, allowing iterative target degradation and diminishing off-target side effects.^{141,143}

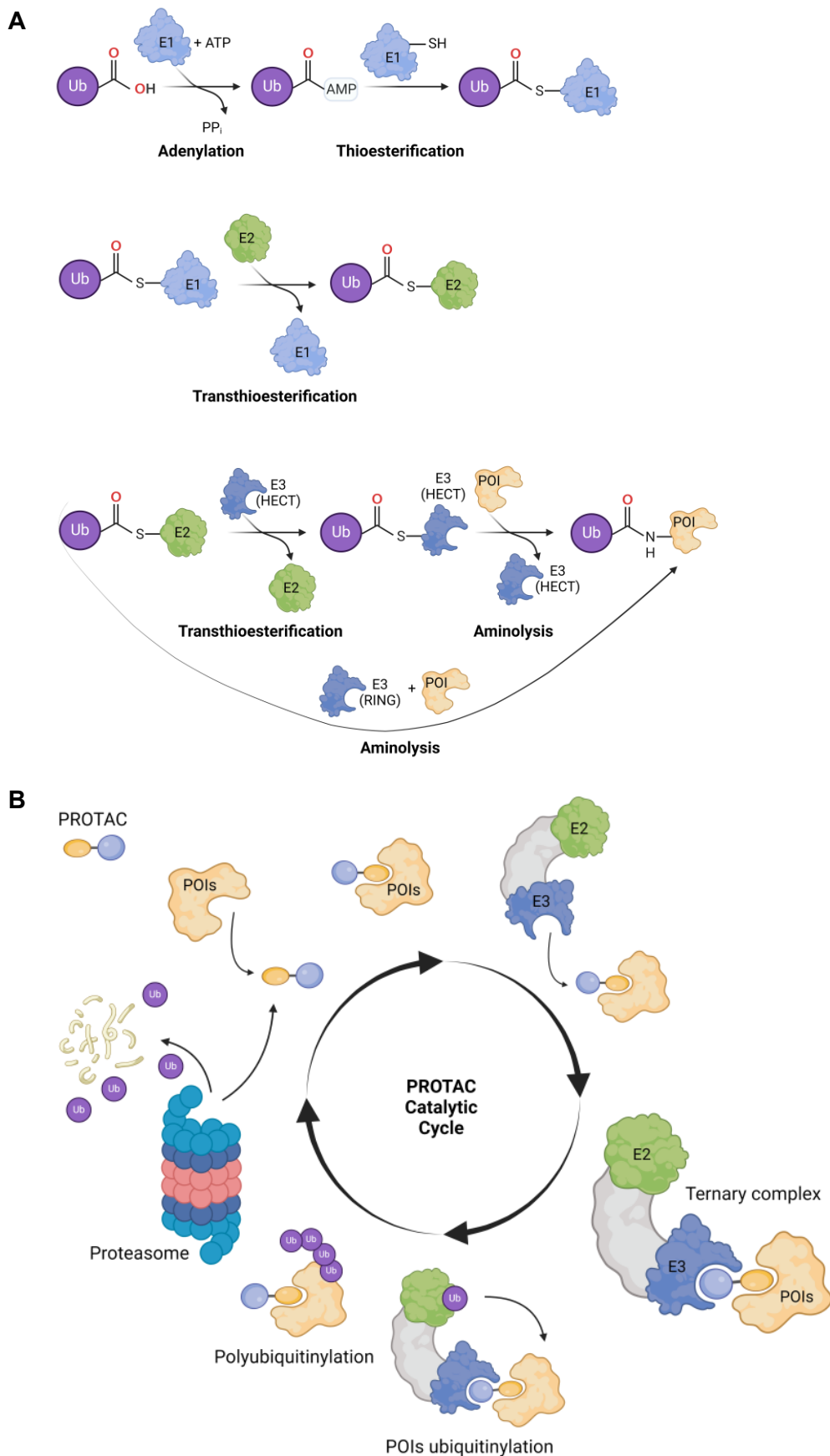


Figure 10. A Ubiquitinylation cascade by three enzymes. **B** Principles of PROTAC.

1.8.2 Structures of PROTACs

The rational designs of PROTACs largely rely on the ligand approaching to the E3 ligase, the nature of the linker and a proper chemical integration between them (**Figure 11**).¹⁴¹ So far most reported PROTACs contain either a phthalimide derivative or a modified peptide fragment from HIF-1 α , recruiting cereblon (CRBN) or Von Hippel-Lindau (VHL) E3 ligase, respectively (**Figure 11A** and **11C**).¹⁴³ Although many variations of these ligands have been designed for improving the druggabilities, incorporating *N*-methyl or decarbonyl congeners of the glutarimide ring within CRBN ligand (**Figure 11B**) and epimerization of the chiral hydroxyl group in the proline ring within VHL ligand (**Figure 11D**) will completely vanish the affinity to the corresponding E3 ligases, generating negative controls of CRBN- or VHL-dependent PROTACs.¹⁴³ On the other hands, a suitable linker plays a vital role to achieve the subtle balance of target degradation, cellular permeability, and aqueous solubility (**Figure 11E**).^{152–154} The position of linker connecting to ligands should be firstly confirmed in order not to disturb binding affinity. Then the optimal linker length should be determined, providing the basis for future modulation of linkers' chemical composition as well as its flexibility, polarity, and metabolic stability (**Figure 11E**).^{153,154} Interestingly, the two PROTACs under clinical trails (Bavdegalutamide, ARV-110 and Vepdegestrant, ARV-471) both utilized piperazine-piperidine conjugate (**Figure 11F**) as the linker, showing preferences of rigidifying and basic moieties to improve overall properties.^{155–157}

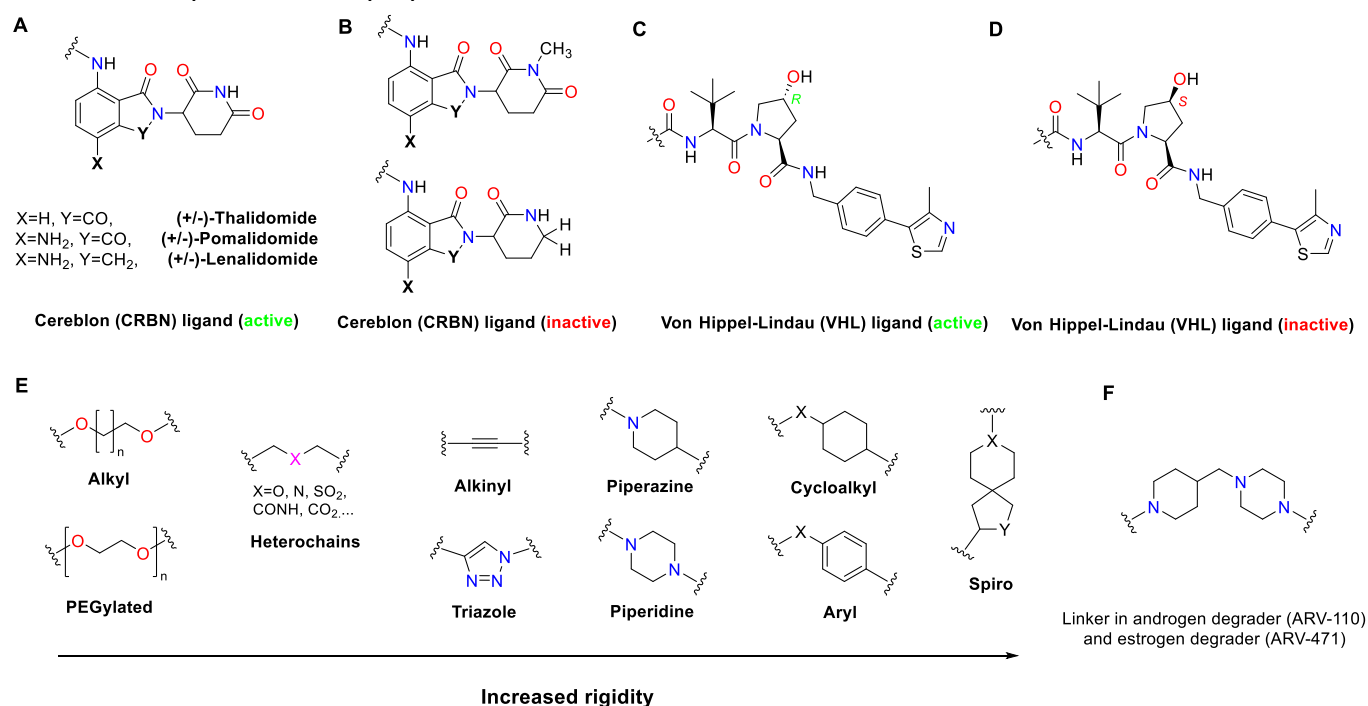
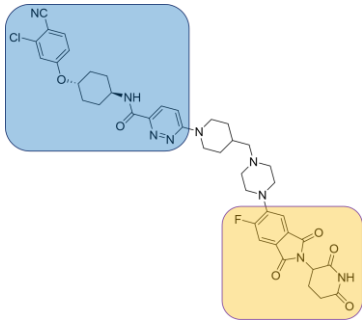
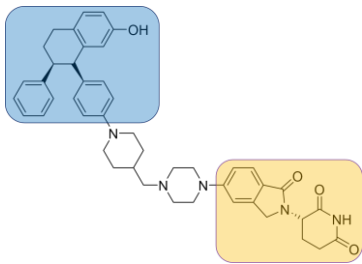
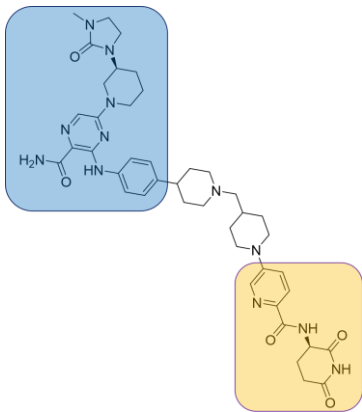
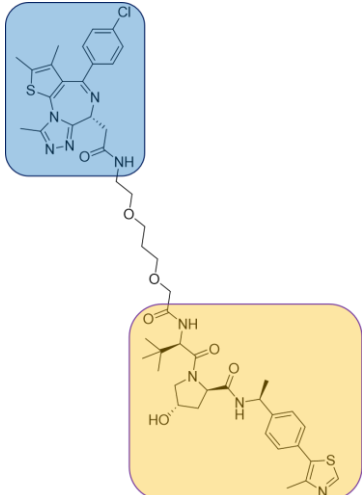


Figure 11. Structures of PROTACs. **A** Structures of prototypical cereblon (CRBN) ligands, including thalidomide, pomalidomide, and lenalidomide. **B** Inactive forms of cereblon ligands, including *N*-methyl and decarbonyl congeners (physicochemical controls). **C** Structures of prototypical Von Hippel-Lindau (VHL) ligand. **D** Inactive form of VHL ligand with a *S*-proline ring (physicochemical controls). **E** List of linkers used in known PROTAC molecules with various rigidity, polarity and length. **F** Linker used in two PROTACs under clinical trails (Bavdegalutamide and Vepdegestrant) from 2019.

PROTACs are especially promising for selectively degrading oncoproteins, and a few examples of PROTACs targeting oncoproteins (including nuclear hormone receptors, transcriptional regulators,

tyrosine kinases, etc) are listed within **Table 1**.

Target oncoproteins with description	Cancer	Degrader structure, name (if applicable); DC ₅₀ / nM	Current situation
Androgen, nuclear hormone receptor, 919 AA ^{156,157}	Prostate cancer	 Bavdegalutamide (ARV-110); < 1	Phase II clinical trial
Estrogen, nuclear hormone receptor, ~ 475-600 AA ¹⁵⁵	ER+/HER2- metastatic breast cancer	 Vepdegestrant (ARV-471); 2	Phase III clinical trial
Bruton' s tyrosine kinase (BTK), non-receptor tyrosine kinase, 659 AA ¹⁵⁸	Hematological malignancies	 NX-5948; < 0.05	Phase I clinical trial
Bromodomain-containing protein 4 (BRD4), transcriptional regulator, 548 AA ^{159,160}	Hematological malignancies		Preclinical

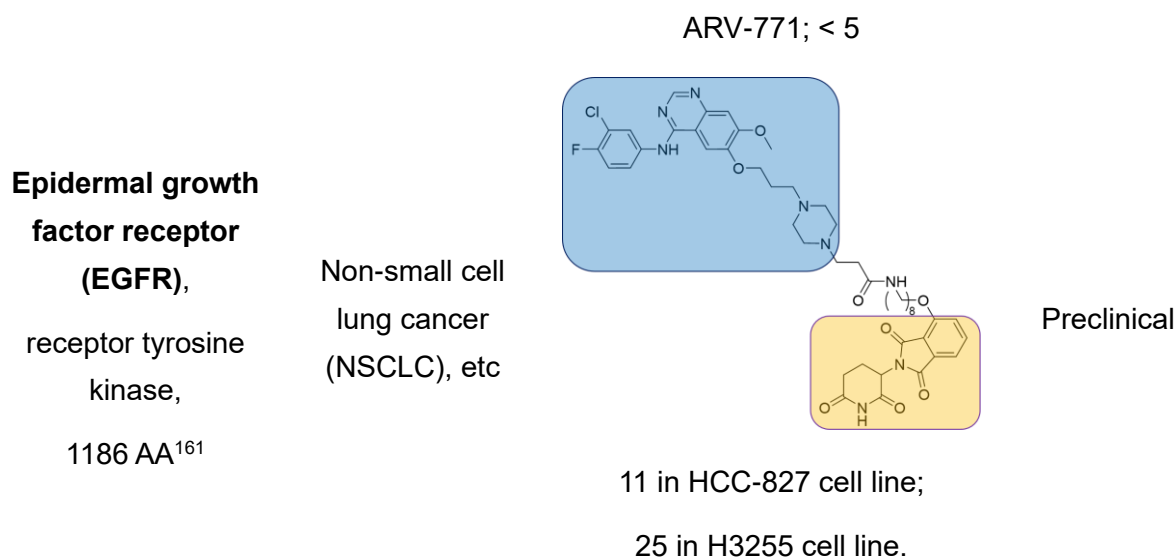


Table 1. Examples of PROTACs targeting on oncoproteins. Targeting ligands and E3 ligase ligands are highlighted in blue and yellow, respectively.

As an emerging technology in both academia and biotech industry, PROTACs still suffer from low cell permeability and challenging oral bioavailability due to its large molecular weights. Another problem to be urgently solved is to expand the library of E3 ligase - ligand interaction, since there are only very few E3 ligases (VHL, CRBN, cIAPs, and MDM2) used in the design of PROTACs, while the human genome encodes more than 600 E3 ligases.^{143,162} Understanding the molecular design, degradation activity, selectivity, and off-target effects (based on different targets, different cell lines, and different animal models) also remain unclear. For these reasons, the proteomic studies should always be performed after PROTAC treatment to evaluate expected or off-target degradation before any clinical trials.

1.8.3 PROTAC Preparation and Implementation with Click Chemistry

Click chemistry, especially the flagship CuAAC has been intensively applied for the final step synthesis of many PROTACs, generating a triazole ring as the rigid linker in between.^{163–165} However, there are only few examples to combine the biocompatibility and selectivity merits of click chemistry to undergo a covalent and fast self-assembly between two precursors to form PROTACs *in situ*, partially because of the inevitable addition of cytotoxic copper ions in the living cells. While this idea could in principle be fascinating to potentially overcome some natural issues of PROTACs such as scarce cell permeability, low bioavailability, and poor water solubility.^{141,162} One of the most promising example was proposed by Astex Pharmaceuticals, in which tetrazine-tagged thalidomide derivatives and transcyclooctene-tagged JQ1 (JQ1-TCO) were designed and synthesized as smaller and inactive PROTAC precursors to penetrate into cells easily (**Figure 12**).¹⁶⁶ Afterwards the tetrazine ligation between two molecules occurred to release the intact and active PROTAC, targeting on the well-known bromodomain-containing protein 4 (BRD4) and leading to ultimate proteasome degradation.¹⁶⁶ The same strategy was then extended to selectively address extracellular-regulated protein kinase 1/2 (ERK1/2, **Figure 12**).¹⁶⁶ Notably, the intact PROTACs (prepared *in vitro* through organic chemistry way) in both cases were not able to introduce any proteasome degradation, probably due to the lack of cell permeability.¹⁶⁶

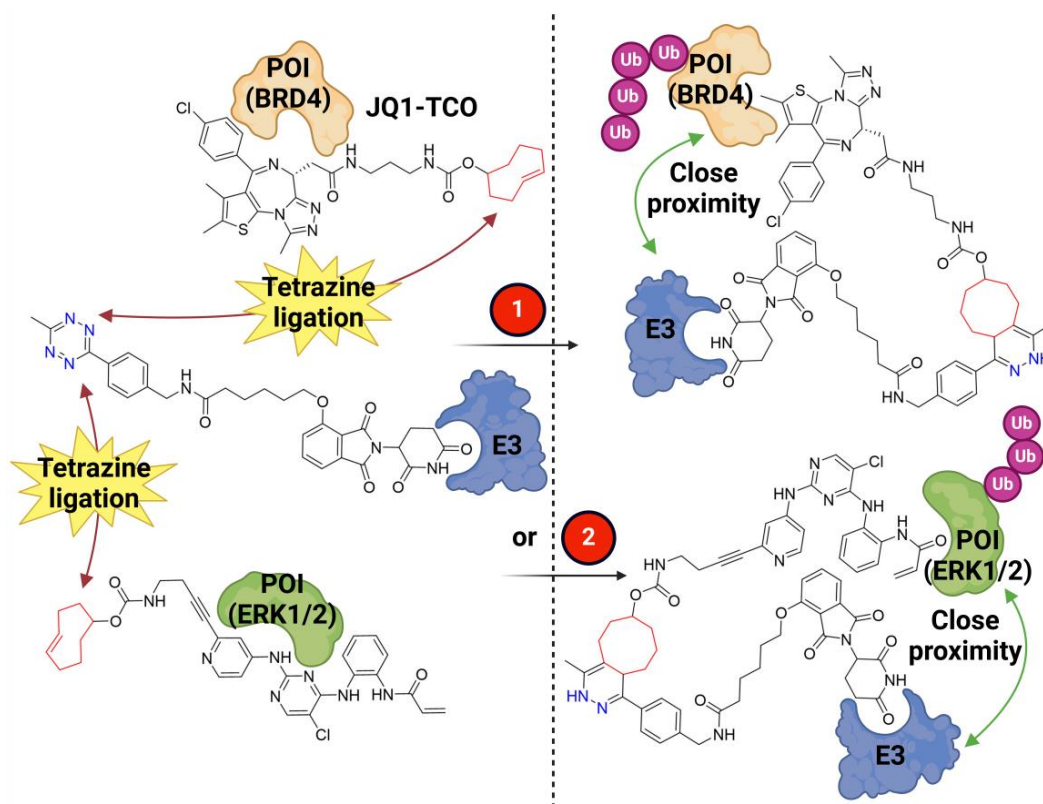


Figure 12. *In situ* self-assembly of PROTACs through tetrazine ligation targeting on BRD4 or ERK1/2 protein.

The other pioneer work done by Liu et. al. has utilized SPAAC to form intact PROTAC from VHL ligand-tagged BCN (VHLL-X-BCN) and commercially available azide-modified DNA oligomer (N3-ODN) (**Figure 13**).¹⁶⁷ The reaction was carried out under physiological conditions (PBS, 37 °C) to generate 18 PROTACs with different linker compositions, among which 5 were demonstrated to reduce p65 transcription protein level successfully.¹⁶⁷ This discovery was of high interest as transcriptional factors are important biological targets involved in many diseases, especially cancers, but they are regarded to be 'undruggable'. The SPAAC-involved PROTAC strategy has shown perspectives to largely expand the boundary of therapeutic methods, becoming a winning combination for chemists.

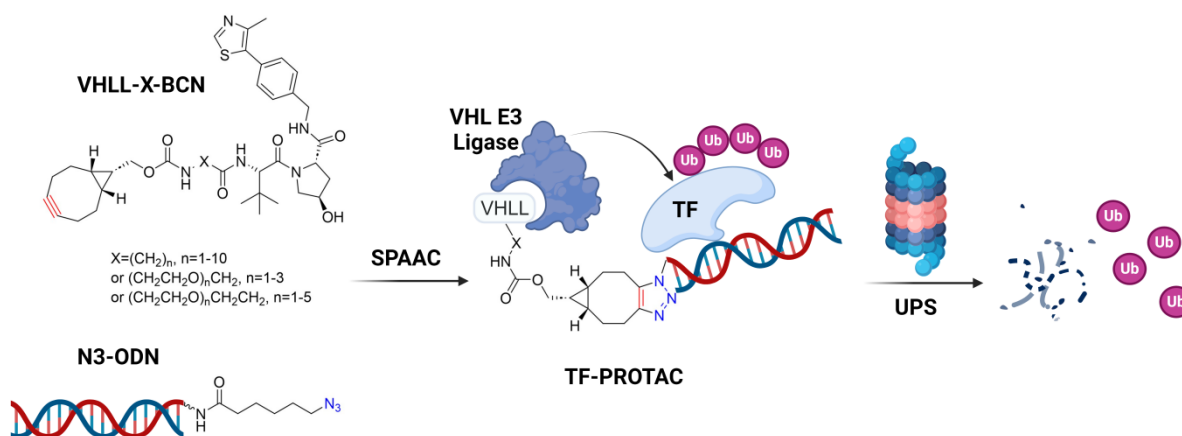


Figure 13. TF-PROTACs synthesized by SPAAC targeting on transcription factors.

1.9 Chemical Proteomics

1.9.1 Single-Pot, Solid-Phase-Enhanced Sample Preparation (SP3) for Proteomics Experiments

In 2019, J. Krijgsveld et. al. have reported a streamlined SP3 (single-pot solid-phase-enhanced sample preparation) technique for sample preparation in proteomics, designed to enhance efficiency and reproducibility.¹⁶⁸ This approach integrates multiple sample processing steps into a single container, significantly improving the workflow from protein extraction to digestion.¹⁶⁸ SP3 technology leverages the solid-phase adsorption properties of magnetic beads to isolate proteins from complex biological samples. The process involves mixing the protein samples with beads under conditions that promote protein binding (such as specific pH and salt concentrations).¹⁶⁸ Magnetic separation is then used to divide the protein-bound beads from other components. Subsequent washing steps remove non-specifically bound substances, resulting in highly purified proteins.

In more detail, cells or tissue samples were firstly lysed to release proteins using appropriate lysis buffers.¹⁶⁸ The lysates were mixed with beads in organic solvent such as ethanol to facilitate protein binding to the beads.¹⁶⁸ Then a magnetic rack was used to separate the protein-bound beads from the solution, discarding the supernatant. Afterwards the beads were washed multiple times with ethanol in water to remove non-specifically bound contaminants including lipids, nucleic acids, detergents, salts, metabolites, etc. Bounded proteins were digested directly on the beads using enzymes like trypsin, breaking them down into peptides. The digested peptides were eluted from the beads by altering the solution conditions (such as pH), ready for subsequent mass spectrometry analysis.

By consolidating multiple steps into a single-pot process, SP3 significantly reduces sample preparation time compared to traditional methods. The use of solid-phase adsorption and magnetic separation minimizes sample loss and contamination, ensuring high reproducibility. SP3 can also effectively remove impurities, improving the sensitivity and accuracy of mass spectrometry analysis.^{168,169}

SP3 technology has broad applications in proteomics, from basic research involving protein identification and quantification to clinical studies focused on biomarker discovery and disease mechanism elucidation. As the technique continues to be refined and optimized, it is poised to become increasingly pivotal in proteomic research, offering a robust and efficient solution for sample preparation.

Recently the SP3 protocol has been developed and optimized to more advanced SP2E protocol, which is highlighted by pulling down PTM mimics-modified proteins with extra enrichment steps.¹⁷⁰ The excess bioorthogonal reagents and non-protein components of the cell lysates have to be firstly removed through intensive washing steps. The biotinylated proteins are then enriched on magnetic streptavidin-coated beads. The SP2E method can achieve higher specificity and selectivity due to the strong interaction between streptavidin and biotin, which is one of the strongest non-covalent interactions in nature.¹⁷¹ This allows for more selective enrichment of the target molecules, and is particularly advantageous when facing samples with low protein concentrations (e.g. proteins modified by low-abundance PTMs), reducing background noise and increasing the signal-to-noise ratio.¹⁷⁰

1.9.2 Orbitrap-Based Mass Spectrometry

A mass spectrometer is an instrument used to determine the relative molecular mass of chemical substances. Although there are many types of mass spectrometers, their basic structures are essentially the same. As shown in **Figure 14A**, a mass spectrometer can be divided into five main compartments: an inlet, an ion source, a mass analyzer, a detector, and a data analysis system.¹⁷² Pure substances or samples with simple compositions can be directly injected to the inlet (such as the top-down approach for purified proteins). However, complex mixtures are usually pre-separated by high-performance liquid chromatography (HPLC) before being introduced into the inlet (such as the bottom-up approach for proteomics studies) resulting in LC-MS.¹⁷²⁻¹⁷⁴ Afterwards samples are ionized at the ion source, where electrons, ions, or photons convert them into charged ions in gas-phase (either cations or anions, depending on the sample's nature). The charged ions then enter the mass analyzer for mass-to-charge (m/z) ratio measurement. With physical forces such as electric or magnetic fields, the trajectory of the ions varies, and the detector converts these variations into electrical signals, which are processed and stored in an affiliated computer.¹⁷⁴ The data is then transformed into a mass spectrum for recording different ions' mass-to-charge ratios, thereby calculating the molecular mass of the analyte. Additionally, the mass analyzer and detector require a high-vacuum system to maintain a low-pressure environment, ensuring that ions do not lose energy through collisions or have their measured mass-to-charge ratios skewed.^{172,174}

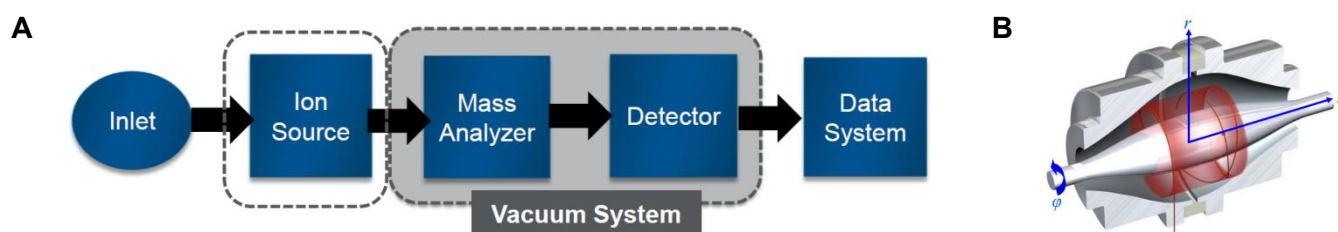


Figure 14. Mass spectrometer and advancement with Orbitrap. **A** General scheme of a mass spectrometer. **B** Description of ion movements in the Orbitrap using three characteristic frequencies: rotational movement (ω_ϕ), radial movement (ω_r), axial oscillations (ω_z). Only ω_z is used for mass analysis.

Thanks to the advancement of instrumentation technology in the past few years, Thermo Fisher scientific has developed Orbitrap Eclipse Tribrid where a linear ion trap is conjugated with quadrupole and orbitrap in front.¹⁷⁵ The samples, usually digested peptides, are ionized through nanoelectrospray ionization (nESI) under atmospheric pressure, then selected by quadrupole based on their m/z values. These ions are transported and collected within Curved Linear Trap (C-trap) and then released at orbitrap to record MS1 spectra.¹⁷⁵ Afterwards each peptide has to be fragmented, for example, in a higher energy collisional dissociation (HCD) cell to determine its identity in a MS2 scan by the linear ion trap with a much faster rate. As a result, the parallelization of the measurement improves the scanning rate of a data-dependent proteomic experiment to approx. 40 Hz, which is a major advance compared to Q Exactive where MS1 and MS2 spectra are both measured within the orbitrap.¹⁷⁶ In general, this instrument has reached the

optimal balance between scanning rate and MS spectra resolution. Future efforts would be put to parallelize the measurement with fast and highly resolving mass analyzers.

The orbitrap uses a DC (direct current) electric field to confine the ions in the ion trap and applies the fast Fourier Transform technique to convert the time signal into frequency signal, which further reflects m/z signals of ions.^{177,178} The electrode within orbitrap is a quadrupole field that produces a spindle-shaped central electric field, allowing the ions to move in a circular orbital fashion, which can be described in combination of three frequencies: rotational movement (ω_ϕ), radial movement (ω_r), axial oscillations (ω_z) (**Figure 14B**).^{177,178} Only axial oscillations (ω_z) is independent of ion energy and position (but proportional to m/z), so this frequency is the only one considered for mass analysis.

1.9.3 Data Acquisition

Generally two modes have been developed for the MS/MS data acquisition, including Data-dependent acquisition (DDA) and Data-independent acquisition (DIA).^{179,180} DDA involves a survey scan (MS1) to detect all precursor ions present in the sample, followed by the selection of the most intense ions for further analysis. These selected precursor ions are isolated and fragmented (typically through HCD), generating MS2 spectra used for peptide and protein identification (**Figure 15**). In contrast, DIA segments the mass range into smaller windows and fragments all ions within each window simultaneously, resulting in highly multiplexed MS2 spectra (**Figure 15**).^{179,180}

The main difference between the two methods lies in ion selection and fragmentation: DDA selects precursor ions based on their intensity in the survey scan and fragments only those selected ions, whereas DIA fragments all ions within defined mass windows without prior selection.¹⁸⁰ Consequently, DDA generates clean, isolated MS2 spectra for specific precursors, making data analysis more straightforward, while DIA produces complex MS2 spectra that require advanced computational tools for deconvolution and analysis.¹⁸⁰

Despite these differences, both methods share similarities such as high-resolution analysis, quantitative capability, and sophisticated instruments for data acquisition. DDA provides high-quality MS2 spectra for selected precursors with simpler data analysis, while it is regarded to be biased towards more abundant ions, and has limited throughput as only a subset of ions is selected for MS2 analysis.¹⁸¹ On the other hand, DIA offers comprehensive coverage of all ions within the defined mass windows, making it better suited for detecting low-abundance ions and especially post-translational modifications. However, due to the complexity of MS2 spectra generated from DIA, advanced deconvolution algorithms are needed for data interpretation; otherwise the DIA measurement will depend entirely on the quality and depth of the pre-measured DDA data set (which can be very time-consuming and not surely reliable).^{181–184} To address this problem, Tsou et al. have developed DIA-Umpire, a software capable of quantifying DIA results without the need of a DDA based spectral library. DIA-Umpire takes advantage of MS1 and MS2 spectra from DIA to generate pseudo-MS/MS spectra, which are subsequently used for standard database searches. The latest method leverages machine learning to create spectral libraries *in silico* on the basis of deep-learning algorithms, predicting both MS/MS spectra and their retention times. DIA-Umpire is now integrated into the freely available software DIA-NN, marking a significant advancement in the analysis of DIA data.^{185,186}

In real chemical proteomic studies, the choice between DDA and DIA depends on the specific goals of the experiment, sample complexity, and available computational resources.

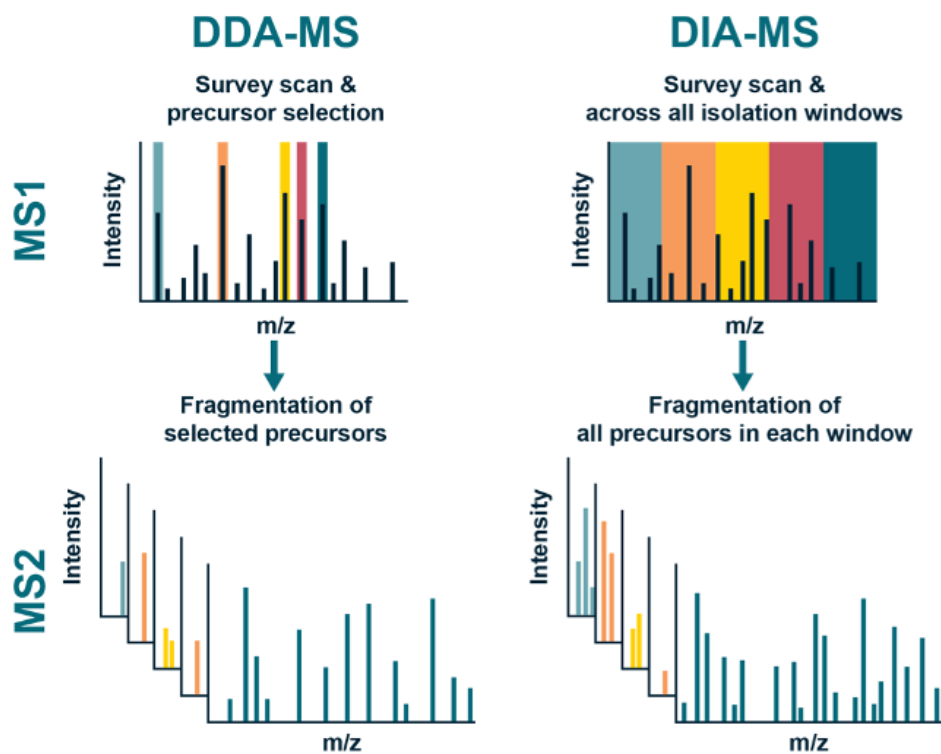


Figure 15. Schematic illustration of basic working principles of DDA and DIA methods.

2. Aim of the Thesis

The overall aim of this thesis is to explore reactivities and functionalities of protein PTMs, establishing protein PTM mimics as putative drug-like molecules with novel mode of action (MOA). The thesis can be divided into two main parts: 1) the unknown reactivity between protein and aryl pinacol boronates, and 2) the selective degradation of glycosylated proteins in the living cancer cell lines.

In more detail, the initial goal of the first part was to find an alternative methodology for commonly-used CuAAC-based protein profiling, in which a biotin- or fluorophore-tagged azide or alkyne handle was clicked to proteins bearing an artificial PTM molecule mimicking its natural counterpart. The functional moieties then allow biological readouts including in-gel fluorescence analysis or LC-MS/MS measurements to study protein PTMs even at low abundance level. We tried to substitute CuAAC with envisaged Suzuki cross-coupling reactions, but found significant side reactions caused by aryl pinacol boronates even in the absence of any metal catalysts. The research then switched to answer what is the side reaction behind and can it be useful for other biological applications. To address this question, we tested various substrates from simple peptides, single proteins to cell lysates. Validation tools such as competition assays, radical trapping experiments, and electron paramagnetic resonances (EPR) are applied to consolidate the discoveries. The newly explored reactivities would be helpful for designing more rational pathway to avoid potential side reactions when aryl pinacol boronates are involved in expected bioorthogonal labelling within living cells or lysates.

The aim of another part is to combine click chemistry and target protein degradation for synergistic effects in suppressing cell proliferation. As a proof-of-concept, we successfully established a platform (GlyTAC) to selectively degrade glycosylated proteins through ubiquitin-proteasome pathway. To accomplish this goal, a galactose-based probe containing azide group was incubated within cancer cell lines to introduce heavy glycosylation, then thalidomide-linked dibenzocyclooctyne was 'clicked' to azide-modified glycoproteins via strain-promoted azide-alkyne cycloaddition in the living cells. With covalent bonds anchoring at modified proteins, the clicked-released PROTAC is proved to be highly competent for glycoproteins degradation with fast kinetics, thus disrupting cell cycle and triggering cell death. The results have further collaborated with other quantitative measurements like whole proteomics and cell viability assays.

In conclusion, this thesis aims to expand the current therapeutic boundaries of protein PTMs, not only to identify an unknown reactivity, but also to apply PTMs as targets for potential therapeutic purposes. Since protein PTMs have largely enriched proteome complexity, it is believed that there are still large amounts of novel reactivities, interactions or applications of PTMs remain uninvestigated, and this thesis aims to address this area.

3. Results and Discussion

3.1 Suzuki Cross-Coupling Reaction for Protein Labelling in the Living Cells

Incubating synthetic analogues of biological macromolecules in the living cells has led to many breakthroughs in cell biology and in drug research. While bioorthogonal reactions have facilitated the modification of proteins and nucleic acids within living cells. However, using protein PTMs with reactive groups for further functionalization to tailor their roles in living cells remains underexplored. Notably, initial investigations into Suzuki cross-coupling reactions within living cells suggest the potential for significantly expanding the range of accessible chemical modifications.

Our preliminary goal was to develop halogen-modified protein PTMs for subsequent functionalization via Suzuki reactions in living cells (**Figure 16A**). By introducing a natural PTM mimic with halogen to target proteins through endogenous enzymatic metabolism, they can be further embellished with other functional modalities (e.g. fluorescent or biotinylated tags) using Suzuki reactions. The resulting technology was expected to advance the study of protein adenylation (AMPylation, **Figure 16B**) or tyrosination (**Figure 16C**), particularly their roles in neurodegeneration or microtubule disorder, respectively. This interdisciplinary strategy, merging organic synthesis, molecular biology, and mass spectrometry, is promising for producing useful tools in chemical biology and innovations in drug design.

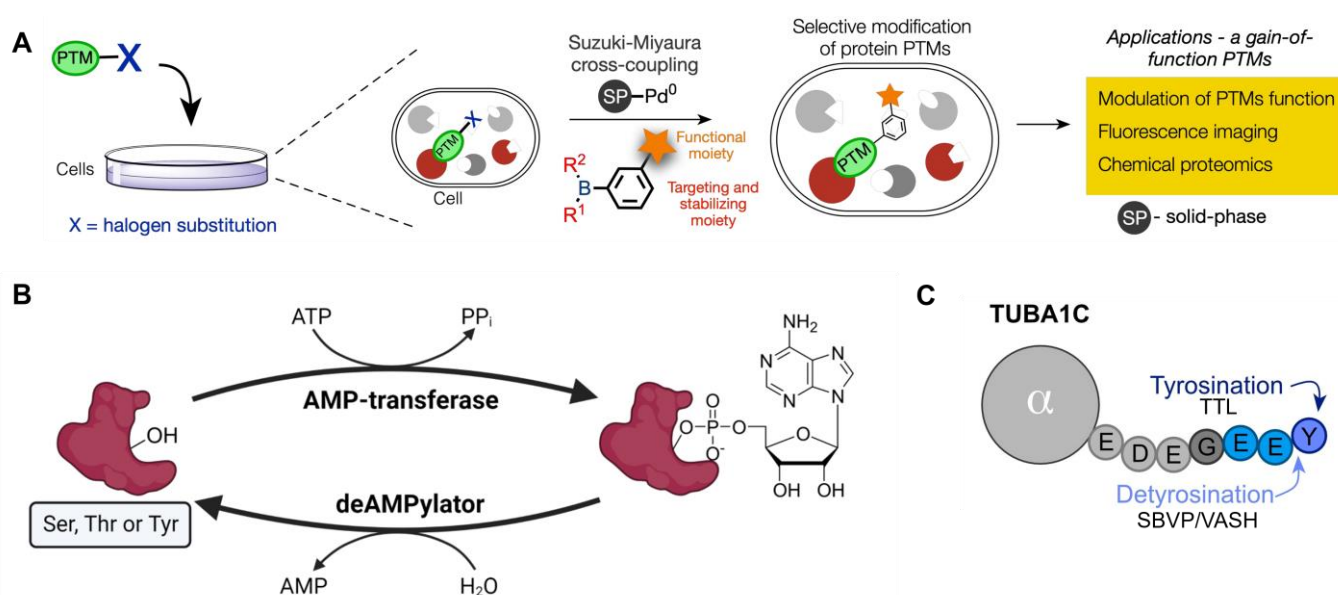


Figure 16. **A** Suzuki cross-coupling reactions for protein labelling in the living cells. **B** General (de)AMPylation scheme. Amino acids including serine (Ser), threonine (Thr) or tyrosine (Tyr) can be modified with an AMP moiety catalyzed by an AMP-transferase using ATP as a co-substrate. The modification is reversible by a deAMPylator to remove AMP with hydrolysis. **C** General (de)tyrosination scheme at α -tubulin C-terminus. Tubulin carboxypeptidases (TCPs, VASH1 or VASH2 associated with SVBP) remove the C-terminal Tyr to generate detyrosinated tubulin. Detyrosinated tubulin can be re-tyrosinated by tubulin tyrosine ligase (TTL) that adds a tyrosine residue at the C-terminal glutamate.

3.2 Design and Synthesis of Iodinated Probes for Pulling Down AMPylated or Tyrosinated Proteins

Following the objective for performing Suzuki cross-coupling reactions in the living cells, the AMPylation or tyrosination probes containing halogen were firstly designed and synthesized. Within three commonly used halogens (chlorine, bromine and iodine) used for coupling reactions, we have selected iodination to modify probe structures based on a comprise of oxidative addition reactivity and atom size to minimize structural complexity. Although iodine (133 pm) possesses the biggest covalent radius compared with bromine (114 pm) and chlorine (99 pm), it's still much smaller than propargyl (> 226 pm) or azide (> 230 pm) groups, which are key elements routinely used in CuAAC-based pull-down experiments. In such case, the application of iodinated probes could in principle be more acceptable by corresponding transferases or ligases, revealing yet not described substrate selectivity and identifying unknown protein AMPylation or tyrosination. The positions for placing iodine were either directly referred or rationally designed from literature research, shown in **Figure 17**.

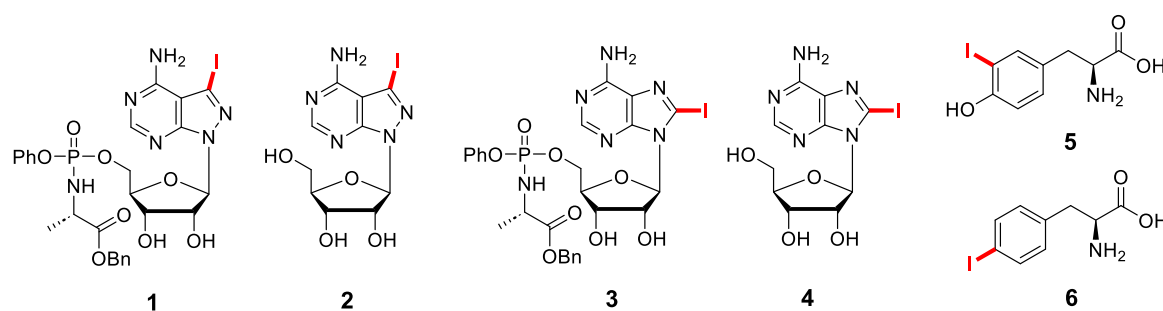


Figure 17. Structures of potential protein AMPylation or tyrosination probes.

The two AMPylation probes (**1** and **3**) were synthesized in the forms of phosphoramidates for two reasons: 1) mitigating cytotoxicity while increasing the cell permeability of iodinated adenosine analogues; 2) the iodinated adenosine analogues must be metabolically activated to corresponding adenosine triphosphate (ATP) derivatives and then be accepted as substrates by AMP transferases. For compound **1**, the core structure of compound (7-deaza-8-azapurine) was same with the iodine-containing substrate employed for *in situ* synthesis of an anti-cancer reagent in E. Indrigo et al. study, thus offering possibilities for the intended Suzuki cross-coupling reactions done on AMPylated proteins. While compound **3** contained the same core structure backbone with natural adenosine, and in principle can be transformed to active ATP analogues. Compounds **2** and **4** are adenosine analogue intermediates respectively for **1** and **3**, which were used as controls for further cell experiments.

On the other hand, the two tyrosination probes (**5** and **6**) synthesized were based the study of D. Schumacher et al., which have systematically investigated the substrate tolerance of tyrosine derivatives when subjected to TTLs. They found that TTL possess unexpected broad substrate tolerance, especially when tyrosine derivatives carry a functional group at the *ortho*-position. These iodinated probes were expected to be incorporated into the living cells and prepared for downstream labelling reactions.

3.2.1 Synthesis of Iodinated AMPylation Probes

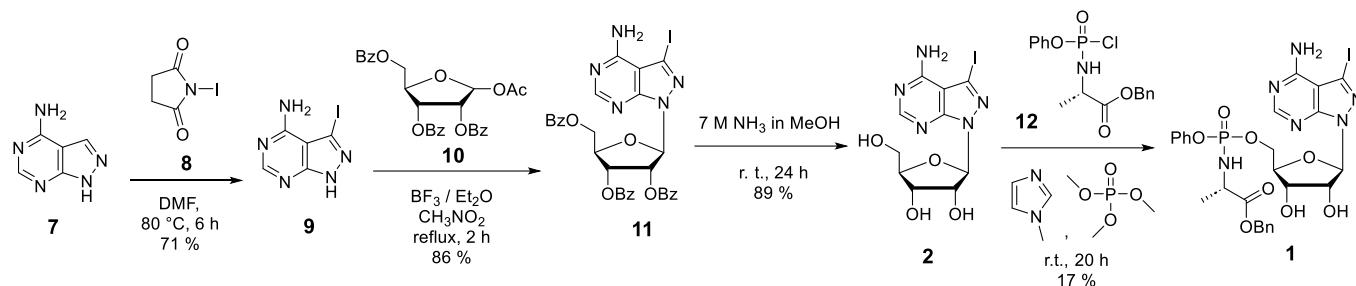


Figure 18. Synthesis of AMPylation probe **1** and **2**.

Starting from adenine analogue **7**, the first step was to introduce iodination accomplished by *N*-iodosuccinimide **8**. The reaction was carried out in DMF, a polar aprotic solvent to facilitate the formation of iodonium ion electrophile. The reaction mixture was then concentrated and triturated with ammonium hydroxide to give the crude products **9**. Afterwards compound **10**, a tribenzoyl-protected ribofuranose, was dissolved together within nitromethane and heated up to reflux to construct the adenosine analogue backbone **11**, featured with a glycosidic bond between adenine analogue (N_9) and ribofuranose (C_1). The three benzoyl groups were eliminated using concentrated ammonia in methanol, generating compound **2**. On the other hand, the pre-synthesized phosphorochloridate **12** was added dropwise into trimethyl phosphate containing **2**, and the substitution on the primary alcohol was finalized by applying *N*-methylimidazole as the deprotonation base.

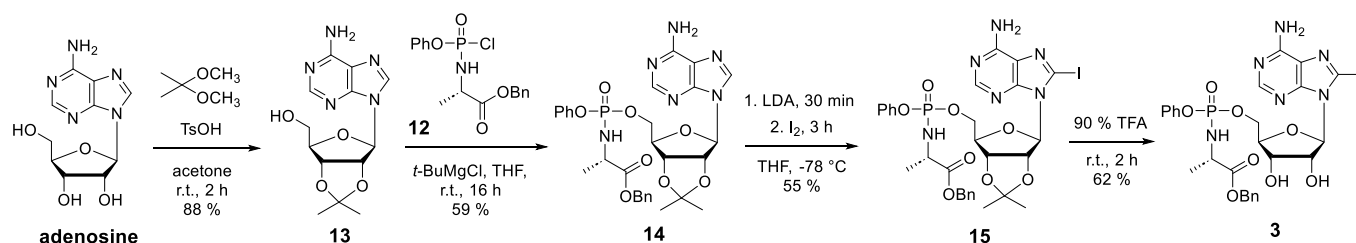


Figure 19 Synthesis of AMPylation probe **3**.

The synthesis of compound **3** has adopted a different strategy highlighted by direct phosphorylation of acetal-protected adenosine **13**. This step was more efficient when Grignard reagent was added as the base, and this was also the reason to perform iodination afterwards (otherwise iodine moiety would be eliminated through Grignard reagent). Afterwards the iodination step at C_8 position on adenine core was achieved by freshly-prepared lithium diisopropylamide (LDA) deprotonation and the subsequent slow addition of iodine. The obtained phosphoramidate compound **15** would then be subjected to deprotection by TFA to afford probe **3**.

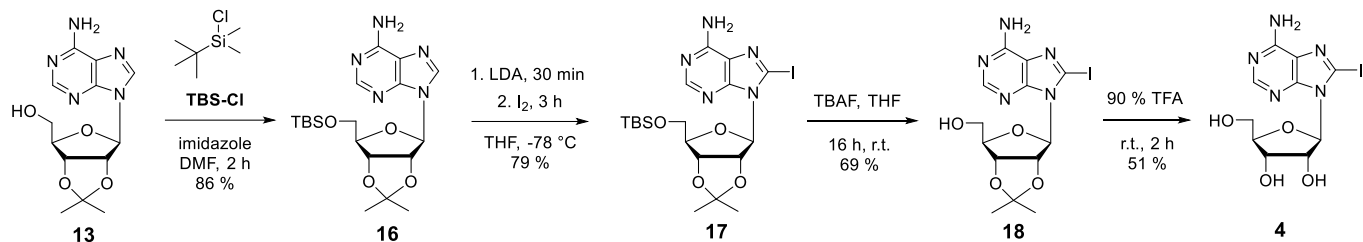


Figure 20. Synthesis of AMPylation probe **4**.

The primary alcohol in acetal-protected adenosine **13** was further protected by *tert*-butyldimethylsilyl chloride (TBS-Cl) to generate compound **16**. The iodination was performed similarly for synthesizing intermediate **15**. Compound **4** was finally obtained through a two-step deprotection reaction, including silyl deprotection by tetra-*n*-butylammonium fluoride (TBAF) and acetal deprotection by TFA to expose all three alcohols.

3.2.2 Synthesis of Phosphorochloridate

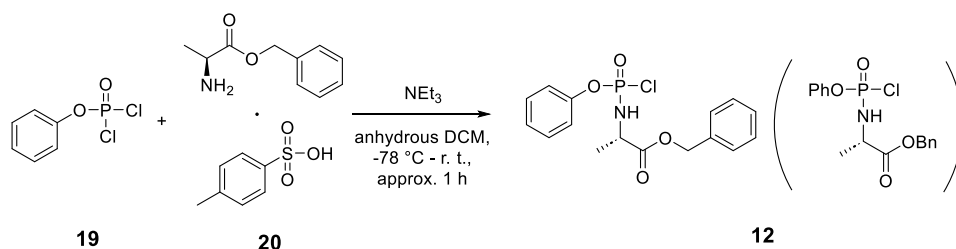


Figure 21. Synthesis of intermediate **12**.

As a key intermediate for synthesizing both **1** and **3**, phosphorochloridate **12** was a challenge in the synthetic route. Two starting materials, phenyl dichlorophosphate **19** and *L*-alanine benzyl ester *p*-toluenesulfonate salt **20** were mixed at strict equimolar in dichloromethane at low temperature, followed by slow warm up to room temperature. The phosphorochloridate must not be over substituted, in order to retain one reactive site for the subsequent alcohol substitution. The product was simply triturated and sent immediately for the next step.

3.2.3 Synthesis of Tyrosination Probes

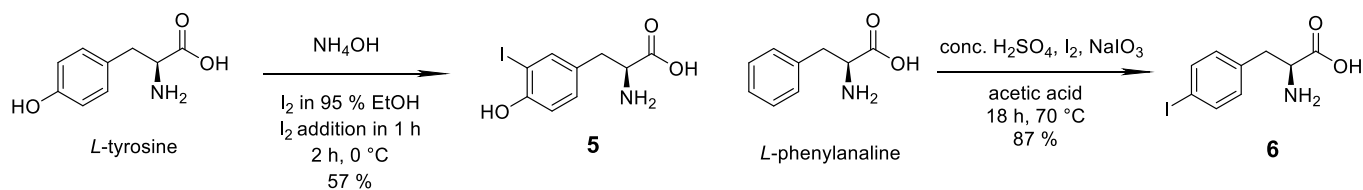


Figure 22. Synthesis of tyrosination probes **5** and **6**.

The iodination on phenyl usually follows electrophilic substitution mechanisms. Since iodine has the least electrophilicity within halogens, harsh conditions (e.g. strong oxidative reagents) are required to generate iodonium ion as an electrophile.

3.3 Chemical Proteomics for Identifying AMPylated Proteins

After synthesizing AMPylation probes **1**, **2**, **3**, and **4**, they were incubated in HeLa cells to trigger protein AMPylation and thus introduce iodohandle into these modified proteins. Due to the future perspective of performing Suzuki cross-coupling reaction *in vivo*, compounds **1** and **2**, who were demonstrated to be compatible within living cells, were selected for initial attempts.

3.3.1 Workflow Description

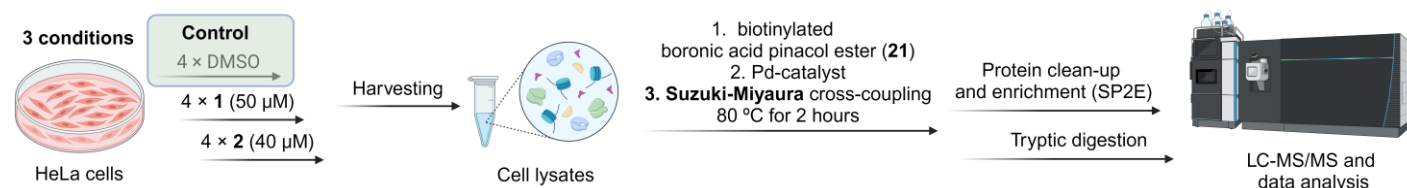


Figure 23 Workflow for identification of protein AMPylation using Suzuki cross-coupling reaction.

As shown in Figure 23, the HeLa cells were treated with compound **1** or **2** (or DMSO as a vehicle) at a constant concentration for 24 hours. The treatment concentrations of compound were determined based on their cell viability assay. Each conditions included four biological replicates. Afterwards the cells were harvested to get the cell lysates and protein concentrations were measured. The cell lysates were then treated sequentially with biotinylated boronic acid pinacol ester **21** (biotin-Bpin, **Figure 24A**) and the palladium catalyst (**Figure 24B**) to allow the labelling and biotin attachment onto the AMPylated proteins. The concentrations of both biotinylated boronic acid pinacol ester and palladium catalyst, as well as the reaction time and temperature, were optimized based on literature researches.⁸⁹ The reaction was quenched by adding highly concentrated urea, followed by protein clean-up (with carboxylated magnetic beads) and enrichment (with streptavidin-coated magnetic beads) procedures. The bound proteins were digested at 37 °C overnight by adding trypsin in a suitable buffer. Finally the peptides after digestion were eluted, desalted, and reconstituted in acidified water for LC-MS/MS measurement and analysis. The incorporation efficiencies of these iodinated probes and Suzuki cross-coupling reaction were assessed by competition with the benchmark pro-N6pA probe.^{170,187}

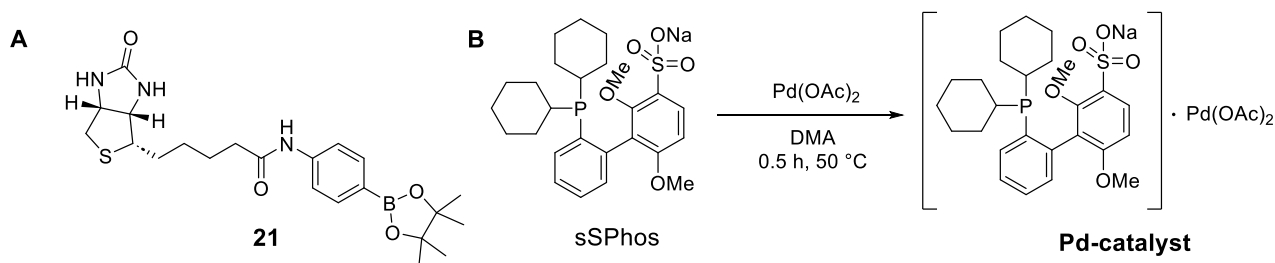


Figure 24 Reactive species used in the workflow. **A** Structure of biotinylated boronic acid pinacol ester **21**. **B** Palladium catalyst used for pulling down AMPylated proteins.

3.3.2 Data Analysis

Two iodinated probes **1** or **2** were expected to trigger protein AMPylation and thus introduce the halogen handle into these modified proteins. Subsequent pulling-down with biotinylated boronic acid pinacol ester **21** in the presence of Pd-catalyst would in theory reflect the difference between AMPylated proteins among all other backgrounds. However, as shown in **Figure 25**, the enriched proteins were rather limited compared to pro-N6pA-treated samples.^{170,187} Although there were one or two known AMPylated proteins being identified and enriched (such as CTSB and CTSZ from cathepsin families), many other characteristic AMPylation targets such as HSPA5, PPME1, ABHD6, and ACP2 were not identified. Proteins including PFKP and SQSTM1 were even down-regulated in probe-treated samples. These can be probably explained from several aspects. First, iodinated probes were not metabolically incorporated by target proteins as expected (or only being incorporated at lower level). Incubation with higher concentrations was unfortunately not possible due to cytotoxicity reasons. Second, the reactivity from Suzuki cross-coupling reaction was disguised by more dominating side reactions from added biotinylated boronic acid pinacol esters **21**, which might be the most important reason for unsatisfactory enrichment results. Lastly, the workflow for protein pulling down was mostly simulated from CuAAC-based procedures, and was not tailored to Suzuki cross-coupling chemistry. Since we have found another interesting labelling pattern caused by aryl pinacol boronates, the research soon switched to investigate what was this exact reactivity behind. Therefore, the initial attempt trying to enrich the AMPylated proteins with an alternative bioorthogonal chemistry was abandoned and discontinued.

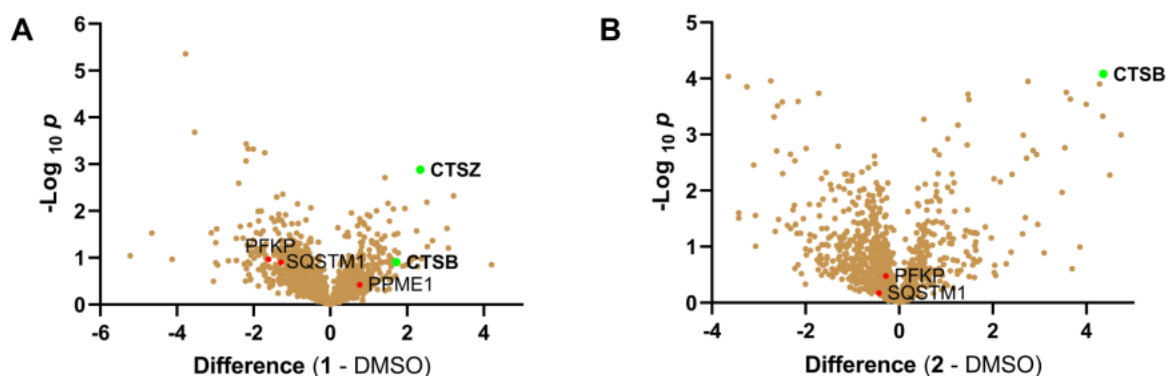


Figure 25 Volcano plot for **1**- or **2**-treated samples (versus DMSO vehicle) showing enriched AMPylated proteins; $n = 4$. Green dots highlighted AMPylated proteins also identified by benchmark pro-N6pA probe, while red dots showed unAMPylated proteins which were supposed to be targeted by iodinated AMPylation probes **1** or **2**. **A** Difference between probe **1** and DMSO vehicle. **B** Difference between probe **2** and DMSO vehicle.

3.4 Unknown Reactivities between Aryl Pinacol Boronates and Cell Lysates

Due to the failure of using chemical proteomics to directly identify AMPylated proteins, we turned our focus to employ a fluorescent aryl pinacol boronate **22** for in-gel fluorescence analysis, attempting to visualize proteins labelled with probes **1**, **2**, **3** and **4**. However, the application of tetramethyl rhodamine aryl pinacol boronate **22** in HeLa cell lysate revealed an unknown reactivity, which was surprising (**Figure 26A**). We started to explore this serendipitous reactivity of aryl pinacol boronates using the general setup shown in **Figure 26A**. The whole cell lysate in phosphate buffer saline (PBS) was treated with 100 μ M **22** (TAMRA-Bpin, **Figure 26B**) with different conditions and was incubated for 2 hours at 50 °C (Figure S1). In the initial experiments to carry out Pd-catalyzed cross-coupling, the protein labelling was not affected by decreasing Pd-catalyst concentration (Figure S2). Furthermore, the reaction proceeded similarly in presence of a metal ion chelator, ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA), thus we concluded that the reaction is not Pd-catalyzed or any metal ion-dependent (Figure S3). DMA (*N,N*-dimethylacetamide) was used as a solvent for Pd-catalyst in the former reactions. To confirm metal-free mechanism, we used plain DMA instead of the Pd-catalyst (**Figure 26C**). Therefore, the cell lysates were reacted in presence of TAMRA-Bpin **22**, with or without DMA. The proteins were subsequently acetone-precipitated and separated on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The fluorescence scan demonstrated that DMA is sufficient for productive coupling between proteins and aryl pinacol boronate **22**. In order to test the influence of the buffer composition, different lysis buffers were used for the cell lysate preparation, in which the subsequent reaction with **22** was carried out. The conditions included PBS buffer supplemented with urea, SDS, HEPES, Tris, NaCl, Tween 20, Triton-X or sodium citrate (Figure S4). The reaction was proved to be compatible with most of the additives (Figure S4). Next, we tested, whether the aryl pinacol boronate moiety was solely responsible for the observed reactivity with proteins. Therefore, we used the commercially available TAMRA-N₃ reagent, which was utilized routinely for Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) instead of aryl pinacol boronate **22**.^{54,188} The direct comparison of both reagents showed that only **22** yielded strong fluorescence protein labelling and hence that the aryl pinacol boronate moiety in presence of DMA is responsible for the reactivity (**Figure 26D**).¹¹¹ Inspired by the work on radical reactions of pinacol boronate scaffolds from Wu *et al.*, in which the EDA complex was formed in presence of DMA, we tested other coupling reagents, which might possibly substitute DMA (**Figure 26E**).¹¹¹ From the selected reagents only *N*-methyl-2-pyrrolidone (NMP) showed comparable reactivity to DMA. Next, the DMA concentration-dependent experiment showed that 1.25 % of DMA (v/v) was already sufficient to efficiently proceed with the reaction (**Figure 26F**). In the following experiments, 2.5 % DMA (v/v) was used in the reaction mixture as a standard setup. Next, the time-dependent labelling was carried out, showing the reaction reaches the plateau within 40 minutes (Figure S5).

Aryl and alkyl boronic acids were reported to react with protein post-translational modifications, in particular, with common diols of glycosylated proteins forming corresponding diesters.^{189,190} To exclude that the reactivity came from the partial hydrolysis of aryl pinacol boronate **22** and subsequent diol formation or transesterification reaction, the cell lysates were further treated with excess amounts of phenylboronic acid after the incubation with **22** (Figure S6). The transesterification with phenylboronic

acid would putatively lead to protein de-labelling in case of transesterification with any potential PTMs containing diols. However, this was not the case, therefore we concluded it was some other type of reaction. Next, the water stability of *N*-phenylacetamide and phenyl pinacol boronate esters in the presence of APS (vide infra) was estimated, and the half-life time was around 150 and 77 minutes, respectively (Figure S7 and S8).

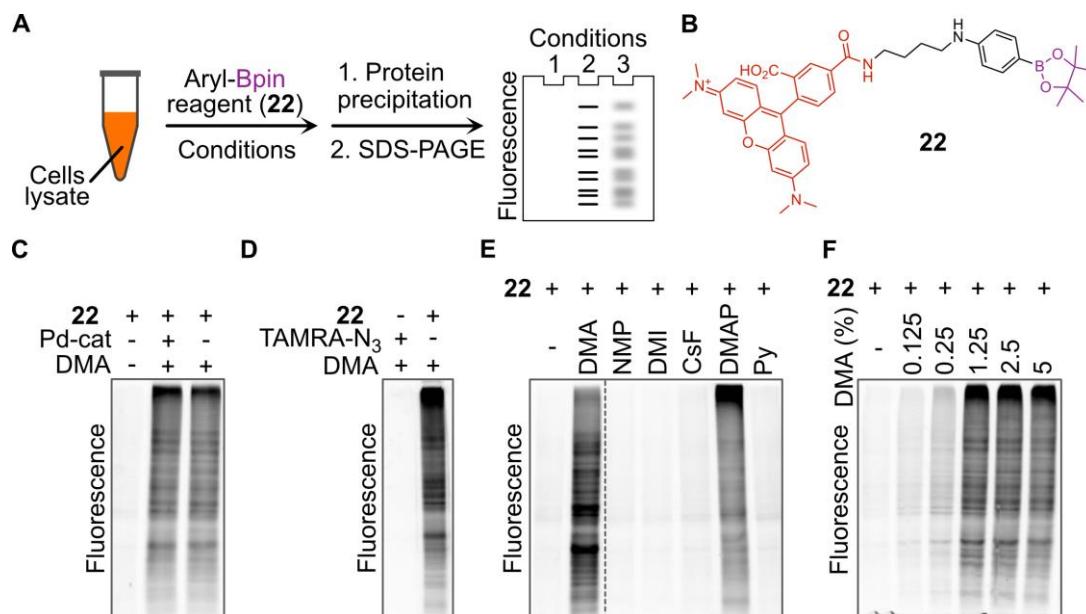


Figure 26. The metal-free reaction of aryl pinacol boronate with proteins in whole cell lysate. **A** General reaction set up for evaluation of aryl pinacol boronate **22** with proteins in whole cell lysate. All reactions were carried out at 50 °C for 2 h in 1.5 mL one-way polypropylene (PP) tubes. **B** Structure of tetramethyl rhodamine-aryl pinacol boronate conjugate (TAMRA-Bpin, **22**) used in the study. **C** The confirmation of Pd-independent reactivity of the aryl pinacol boronate **22** with proteins. The fluorescence scan of the SDS-PAGE separated proteins shows the necessity of DMA for productive coupling between **22** and proteins. **D** Comparison of reactivity between commercial TAMRA-N₃ or **22** with whole proteome lysate reveals the critical reactivity of aryl pinacol boronate. **E** Screening of the possible coupling reagent for the reaction between aryl pinacol boronate **22** and proteins. The DMA, NMP and DMF were left air-exposed from previous experiments. **F** The DMA concentration-dependent reactivity.

Based on the putative DMA-catalyzed radical reaction mechanism, which might be triggered by oxygen species present in the lysate, we decided to remove the oxygen by the described enzymatic reaction. We used a scavenging system consisting of glucose-oxidase, catalase (GODCAT) and glucose.¹⁹¹ Indeed, after oxygen removal, the reaction was fully inhibited (**Figure 27A**). To further corroborate this finding, we repeated the reaction in cell lysate with TAMRA-Bpin **22** with fresh and air-exposed reagents (DMA, NMP, DMF). While the ‘fresh’ reagents did not yield any labelling, the air-exposed ones (DMA and NMP) showed productive coupling (Figure S9). To test if an addition of a radical initiator changes the reaction outcome in lysate, we complemented the reaction with APS and azobisisobutyronitrile (AIBN) with or without DMA (Figure S10). However, either weaker or no labeling was observed (Figure S10).

3.5 Reactivity Studies on Single Proteins

To investigate the reactivity on the protein level (**Figure 27B**), we first used bovine serum albumin (BSA) as the protein substrate. The reaction was set up under both conditions either with DMA or APS. The TAMRA- N₃ was used as a control. The results are shown in **Figure 27C**. The BSA was labelled by TAMRA-Bpin **22** under both DMA- and APS-catalyzed conditions, with stronger labelling observed with APS. Surprisingly, the labelling was observed as well without DMA and APS, suggesting yet another reaction mechanism. Of note, BSA contains 17 disulfide bonds and one free cysteine. Since we suspected that preferably disulfide bonds might react, we used *Escherichia coli*-expressed thioredoxin (TrxA), a 13 kDa protein containing two cysteines linked by a disulfide bond. TrxA can be produced recombinantly by *E. coli*. in large amounts and purified to homogeneity (Table S2, S3 and Figure S11 and Figure S12). To test the reactivity of a disulfide bond, two TrxA carrying a Cys to Ser point mutations were prepared (TrxA^{C33S} and TrxA^{C36S}) and expressed purified analogously (Table S2, S3 and Figure S12). The intact protein MS confirmed the formation of disulfide bond within wild-type TrxA and the composition of two Cys to Ser point mutants (Figure S13). However, the in-gel fluorescence analysis of the reaction between TrxA and aryl pinacol boronate **22** yielded comparably intense signals for all tested conditions, including two mutants (**Figure 27D**) and TrxA dimer which was partially formed through disulfide bonds between two wild type TrxA monomers (Figure S12). The same conditions were also applied for labelling of the anti-biotin antibody (Figure S14). To visualize the amounts of modified proteins by aryl pinacol boronates, a model protein, myoglobin from equine skeletal muscle was let reacted with Biotin-Bpin **21**, followed by enrichment with streptavidin-coated magnetic beads. The modified proteins were then fluted from beads at high temperature, and were resolved using 15 % SDS-PAGE gel. Afterwards the gel was firstly stained with normal Coomassie Brilliant Blue, but there was no protein bands visible. We assumed that this was due to the very low amounts of modified proteins by biotin-Bpin. Then the gel was again stained with silver staining, shown in **Figure 27E**. To our delight, there was a weak but clear band at around 17 kDa, indicating the modification would only occur when DMA and aryl pinacol boronate species were both present. The protein band at around 15 kDa could be attributed to monomer of streptavidin, which is composed of four homogeneous subunits. It was worthy noting that equine myoglobin contains all natural amino acids except cysteine, meaning the complexity of observed reactivity in cell lysates. For future perspectives, more efforts would be done to disassemble all the possible reactivities contributing to the overall labelling in cell lysates. These observations pointed to the conclusion that on a protein level, the labelling is dependent on presence of oxygen species in the reaction mixture as well as the coupling reagent to stabilize intermediates necessary for the generation of aryl radicals.

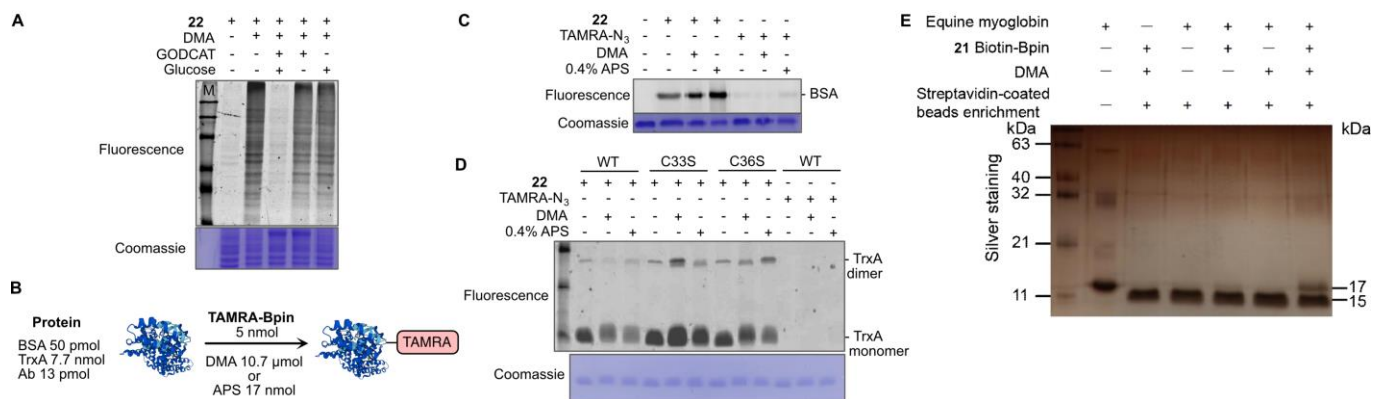


Figure 27. **A** Scavenging of the oxygen from the reaction mixture containing whole cell lysate and **22** using the GODCAT enzyme. M – Fluorescence molecular weight marker. **B** Reactivity between **22** and single proteins including BSA, TrxA and anti-biotin antibody. **C** Reaction of **22** or TAMRA-N₃ with BSA in presence of DMA or APS. **D** Reaction of **22** or TAMRA-N₃ with TrxA in presence of DMA or APS. **E** Incubation between equine myoglobin and **21**, followed by streptavidin-coated magnetic beads enrichment and visualization by silver-staining.

3.6 Reactivity Studies on Short Peptides: Disulfide S-Arylation with Aryl Pinacol Boronates

To further explore the reactivity, we set up a model reaction between *N*-phenylacetamide aryl pinacol boronate **23a** and short peptides. Since we suspected that disulfide bonds might be responsible for the observed reactivity, as noticed on the whole proteome level after disulfide bonds reduction (Figure S11), we let oxidized form of glutathione (**GSSG**, Figure S15) react with **23a** (**Figure 28A**). The LC-MS analysis of the reaction mixture confirmed the corresponding S-arylated product (**24a**) formation in presence of APS but not with DMA (**Figure 28A**). The control reaction starting from reduced glutathione (GSH) showed no product formation as well as **GSSG** without APS addition (**Figure 28B** and Figure S16). The successful formation of alkyl aryl sulfide product **24a** was further corroborated by high-resolution MS analysis and MS2 with higher-energy collisional dissociation (HCD) fragmentation (**Figure 28C**, Table S1, and Figure S17). The same reactivity was observed for *N*-propionyl-4-aminophenyl pinacol boronate (**23b**) and product **24b** formation was also confirmed by LC-MS analysis and high-resolution MS2 fragmentation (**Figure 28C**, Figure S18 and S19). Variation of catalytic amount of APS final concentration between 0.1 and 4 %, which corresponds to 0.11-4.4 mM in comparison to starting pinacol ester, yielded the desired product **24a** at APS concentration between 0.4 and 1 % (Figure S20). In parallel, the reaction was reproduced on two more complex oxidized forms of hexapeptides KEHSPC and RWTPCD (Figure S21 and S22). The structures of the products were similarly as before confirmed by LC-MS and high-resolution MS2 analysis (Figure S23-S26). Taken together, we show that aryl pinacol boronates can be utilised for selective S-arylation of peptides under mild aqueous conditions without necessity for neither photo- nor metal-catalysis, tolerating side-chain hydroxyl, thiol and amino groups, carboxylic acid, guanidine, indole and imidazole. However, given relatively low conversion, further insights into mechanism or establishment of flow process will be necessary to fully benefit from the approach.^{192,193}

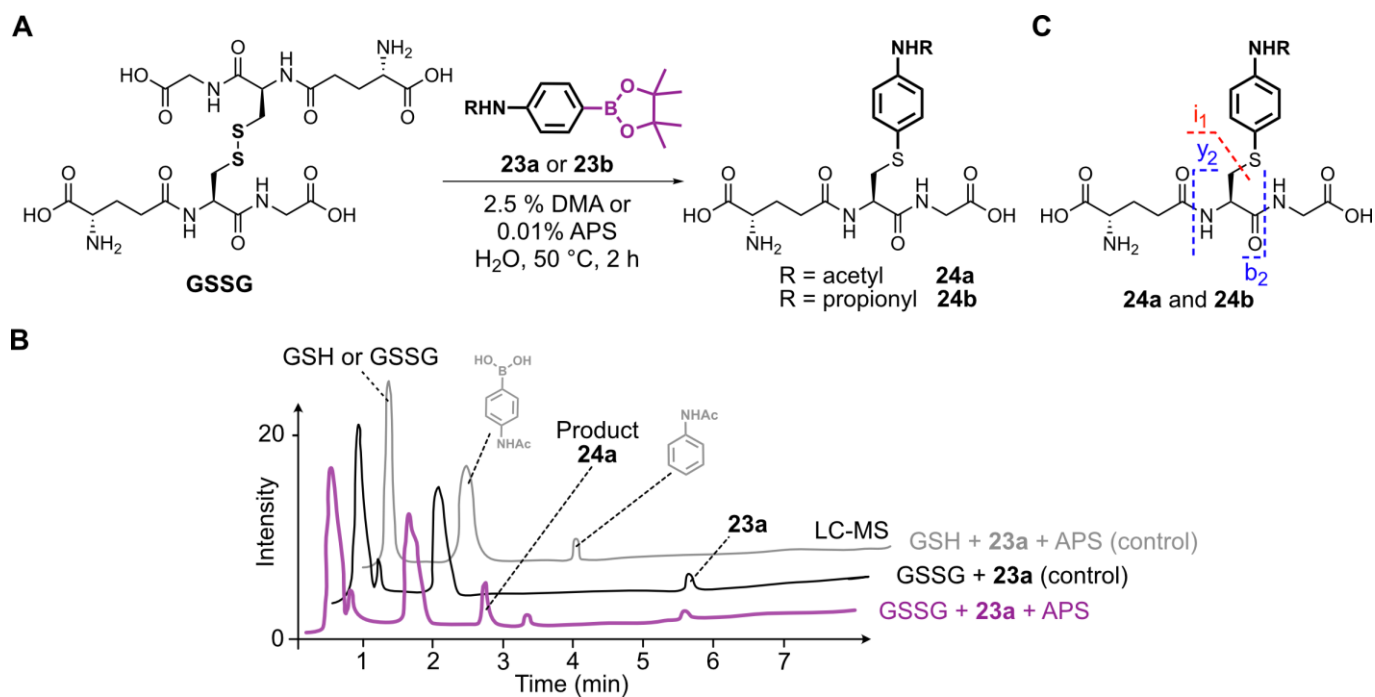


Figure 28. Characterization of aryl pinacol boronate reaction with oxidized glutathione. **A** Reaction of oxidized glutathione (GSSG) with aryl pinacol boronate **23a** or **23b** catalyzed either by APS. Conditions: 0.01% APS (w/v), 12.5 mM **23a** or **23b**, and 2.5 mM GSSG (or GSH) in a total volume of 40 mL water. The reaction mixture was incubated at 50 °C, 2 h, 750 rpm in darkness. **B** LC-MS analysis of reaction mixtures showed the selectivity of **23a** towards disulfide moiety. **C** Identified MS2 fragments of product **24a** or **24b** using HCD fragmentation. All fragments from **24a** and **24b** are listed in Table S1.

3.7 Studies of Substrate Scopes and Mechanistic Pathway

To investigate putative radical mechanism of the reaction, we carried out radical-trapping experiment. The application of APS in contrast to DMA-catalyzed reaction enabled to trap the aryl radical with 5,5-dimethyl-1-pyrroline *N*-oxide (DMPO, **Figure 29A**). The corresponding product **25** was identified by high-resolution MS analysis (**Figure 29A** and Figure S27). Next, we investigated the substituent scope of aryl pinacol boronates yielding corresponding aryl radicals in the presence of APS (**Figure 29B** and Figure S28). While the EDGs at the para position of pinacol boronates showed productive aryl-DMPO adduct formation, the unsubstituted phenyl pinacol boronate and other derivatives bearing EWGs did not react (**Figure 29B**). The yields of aryl-DMPO adduct for each pinacol boronate species correlated well with substituents' activation magnitudes, which were consistent with their competition competence against **22** within cell lysates (**Figure 29B** and Figure S29). Aryl pinacol boronates with no para-substituents or containing EWGs neither generated DMPO spin adduct nor showed any competitive effect when co-incubated with **22**, i.e., the fluorescence was not affected for cell lysates (**Figure 29B** and Figure S29). To further confirm the formation of the aryl radicals, the aryl-DMPO adducts were analyzed by electron paramagnetic resonance (EPR) showing the corresponding signal only for tested phenol pinacol boronate incubated with DMPO in presence of APS (**Figure 29C**). There was no signal belonging to a radical species observed neither for the controls nor in the reaction with DMA instead of APS (**Figure 29C**).

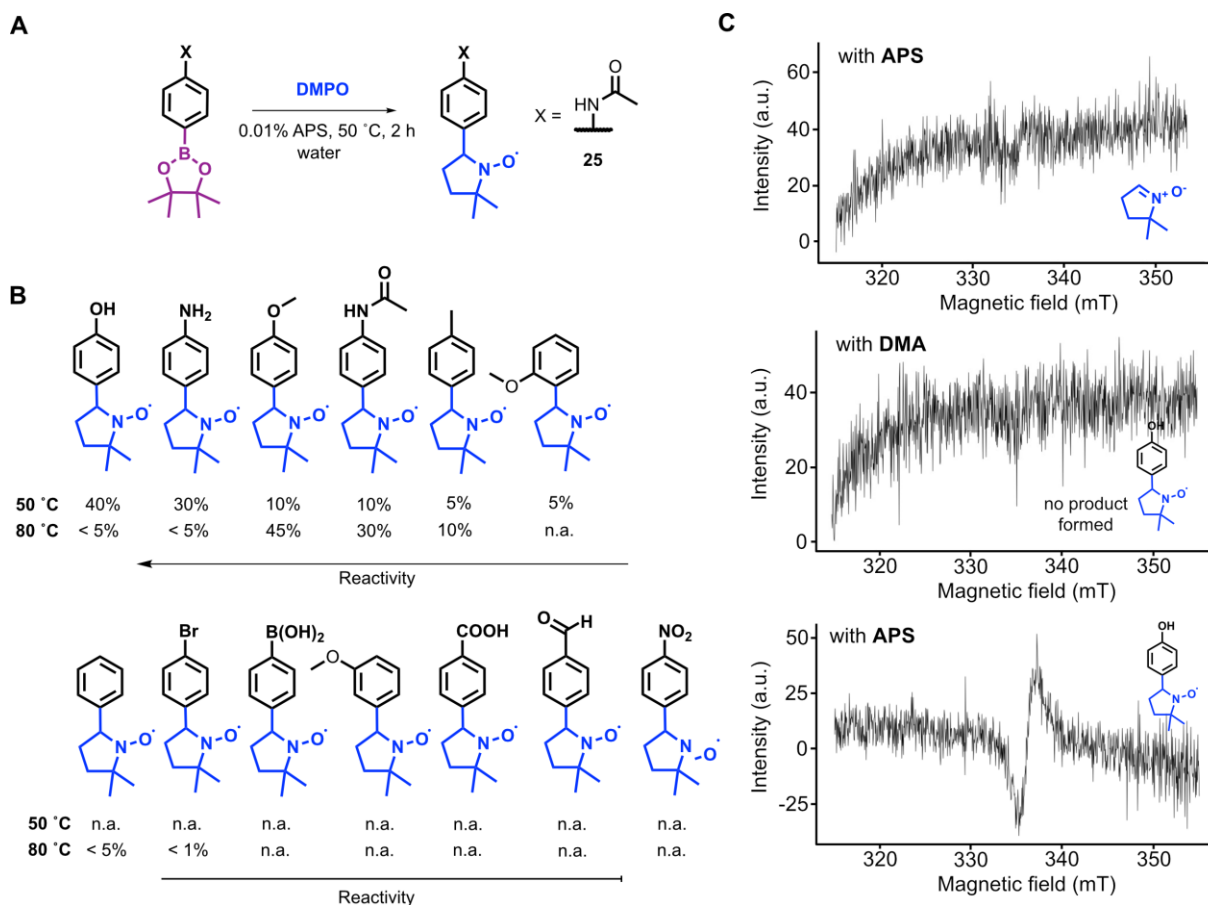


Figure 29. A General scheme of the reaction of aryl pinacol boronate derivatives with DMPO. **B** Scope of the aryl radical formation from aryl pinacol boronates. The yields were estimated by LC-MS (see Figure S21). **C** EPR measurement of the aryl-DMPO adduct.

The described reaction of aryl pinacol boronates with peptides and proteins proceeds via multiple distinct mechanisms. The first type of reaction will proceed in presence of DMA and oxygen, while the second type is catalyzed by APS. In **Figure 30A**, we propose the structure of a plausible radical intermediate with APS inevitably leading to the formation of aryl radical and we compare its stability with other organoborane compounds. Interestingly, in attempt to carry out the NMR kinetics measurement in D₂O, we found that the reaction is strongly retarded by the isotopic effect from deuteration of the solvent, leading to significantly slower kinetics revealed by LC-MS (Figure S30).¹⁹⁴ Finally, we examined another two examples of the boronates with more electron-donating substituents instead of pinacol, showing that unfortunately these do not accelerate the reactivity of the tested phenyl boronate (**Figure 30B**).

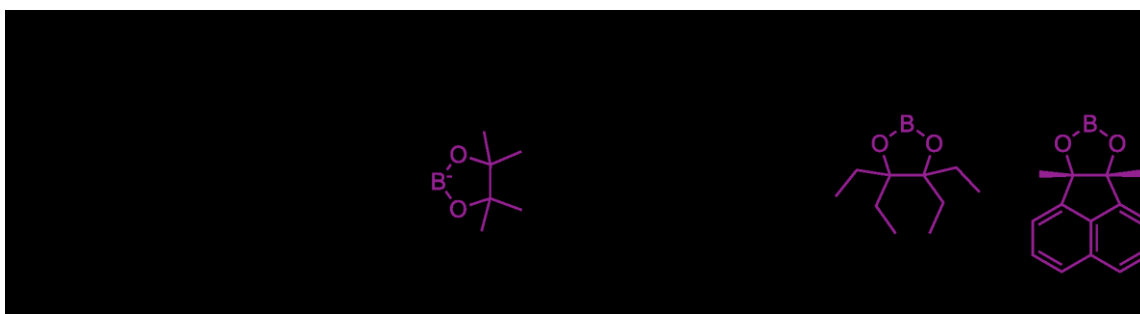


Figure 30. A Comparison of the stability of proposed intermediates. **B** Replacement of the pinacol group with more electron-donating groups did not lead to acceleration of the reaction.

In conclusion, a serendipitous observation of aryl pinacol boronates conjugation with proteins in cell lysates uncovered their previously unknown reactivity, which further widened the scope of transformations of widely available aryl pinacol boronates. The reaction is distinct from previously described transformations as it does not require photoinduction or metal catalyst to proceed, and can be carried out in aqueous conditions, while aryl boronic acids were proved to be unproductive. Though the mechanism remains to be further investigated, the oxygen scavenging, radical-trapping and spectroscopic measurements suggest aryl radical formation. On peptide level, the reaction offers mild conditions for the preparation of *S*-arylated peptides from disulfide -linked dimers.

3.8 Further Attempts Using Iodinated Tyrosination Probes

The two iodinated tyrosination probes (**5** and **6**), however, remained uninvestigated. For verifying the incorporation and labelling efficiency of two tyrosination probes, they were respectively co-incubated with the benchmark Tyr-O-Alk probe (**Figure 31A**) in SH-SY5Y cell line.¹⁹⁵ The concentration of Tyr-O-Alk probe was kept constant while the concentrations of probe **5** and **6** varied (**Figure 31B** and **31C**). The cell lysates were treated with standard conditions of CuAAC click chemistry developed by our group, introducing a fluorophore onto proteins (mainly microtubules) modified merely by Tyr-O-Alk probe.¹⁹⁵ The labelling intensity would in principle reflect the competition of two iodinated tyrosination probes against Tyr-O-Alk. As shown in **Figure 31B** and **31C**, the two probes (especially probe **6**) have indeed compete out Tyr-O-Alk to some extent at α -tubulin region (approx. 55 kDa). The competition was concentration-dependent and thus deserved a further investigation. However, as mentioned above in § 3.3.2, since the research switched to discover the reactivity arising from aryl pinacol boronates, the initial idea using Suzuki cross-coupling chemistry was not further carried out.

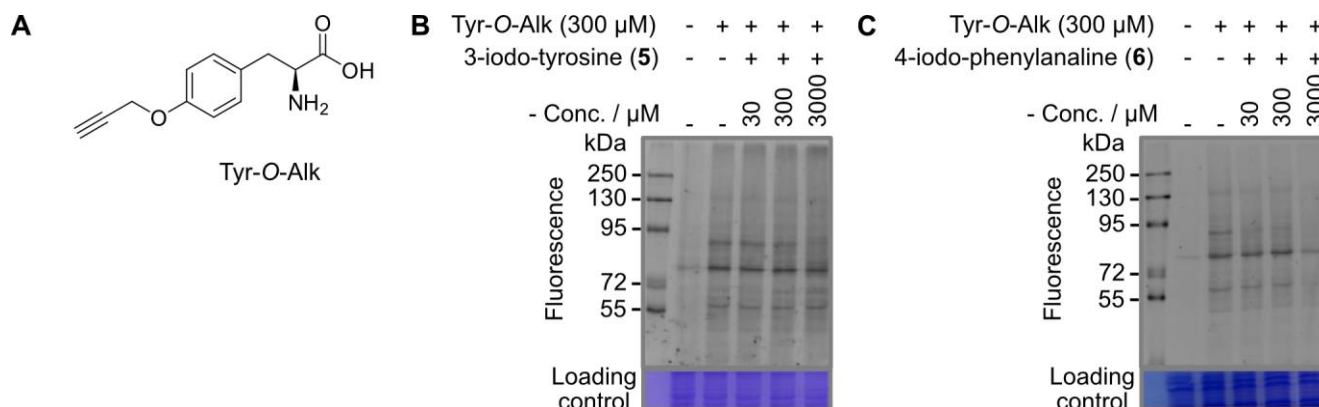


Figure 31. Competition between Tyr-O-Alk and iodinated tyrosination probes. **A** Structure of Tyr-O-Alk. **B** In-gel fluorescence analysis showing competition between **5** and Tyr-O-Alk. **C** In-gel fluorescence analysis showing competition between **6** and Tyr-O-Alk.

To conclude, we have designed and synthesized several iodinated mimics of protein PTMs, and tried to functionalize POIs in the living cells, such as monitoring dynamic PTMs with fluorescent tags. However, unexpected reactivities introduced by pinacol boronates have prevented such efforts. On the other hand, this discovery could still be helpful to better understand the full scenarios of so-called bioorthogonal reactions, and avoid the undesirable side reactivities. Instead, we have applied mimics of protein PTMs to

a more intriguing field, aiming for therapeutic potentials by selectively eliminating pathogenic proteins in the living cells. This fascinating strategy has been combined with PROTAC modalities, a rising star in drug discovery researches. We believe that a flexible manipulation of protein PTMs will always be beneficial regarding expanding the toolbox of chemical biologists.

3.9 GlyTAC Approach for Selectively Degrading Glycoproteins

PROTACs are typically developed to degrade selectively one protein target. However, the selected POI might be present in cells, tissues, and organisms in different proteoforms, which possess different metabolic or signaling functions resulting in complex responses. More specifically, protein post-translational modifications (PTMs) such as phosphorylation, acetylation and glycosylation significantly extend the number of proteoforms present in the different cell types and may switch POIs' catalytic activities, cellular localization and downstream signaling. Focusing on protein glycosylation, which is fundamental for many signaling pathways in cellular metabolism, there are currently no available platforms for selective degradation of glycoproteins.¹⁹⁶ Furthermore, protein glycosylation is significantly altered in cancer cells, featured by typically an overexpression of glycosyl-transferases, leading to overall increased glycosylation levels and *de novo* glycosylation.^{196–198} The high metabolic rate of cancer cells requires enhanced nutrient uptake and triggers profound metabolic reprogramming which contributes to the pathophysiology of the disease.¹⁹⁹ The accelerated turnover of the nutrients in cancer cells can be hijacked for diagnosis purposes, such as using 2-deoxy-2-(fluoro-¹⁸F)-D-glucose as contrast agent in positron emission tomography (PET).²⁰⁰ We report the two-component system, in which the first component 'sensitizes' specific protein targets in the cells, while the second component prompts selective protein degradation (**Figure 32A**). In the first step, the sugar GalNAc or GlcNAc analogue (**Figure 32B**) is metabolically introduced as PTMs into glycoproteins, which are subsequently self-assembled through SPAAC with thalidomide analogue to form the 'active' PROTAC in the second step (**Figure 32A**). The combination of two non-toxic precursors induces proteome-wide changes, degradation of 'tagged' glycoproteins and leads inevitably to cell death. The GlyTAC, as we named the strategy, leverages the specific metabolic activity of glycosyl-transferases and hence prime cytotoxicity in cancer cell lines.

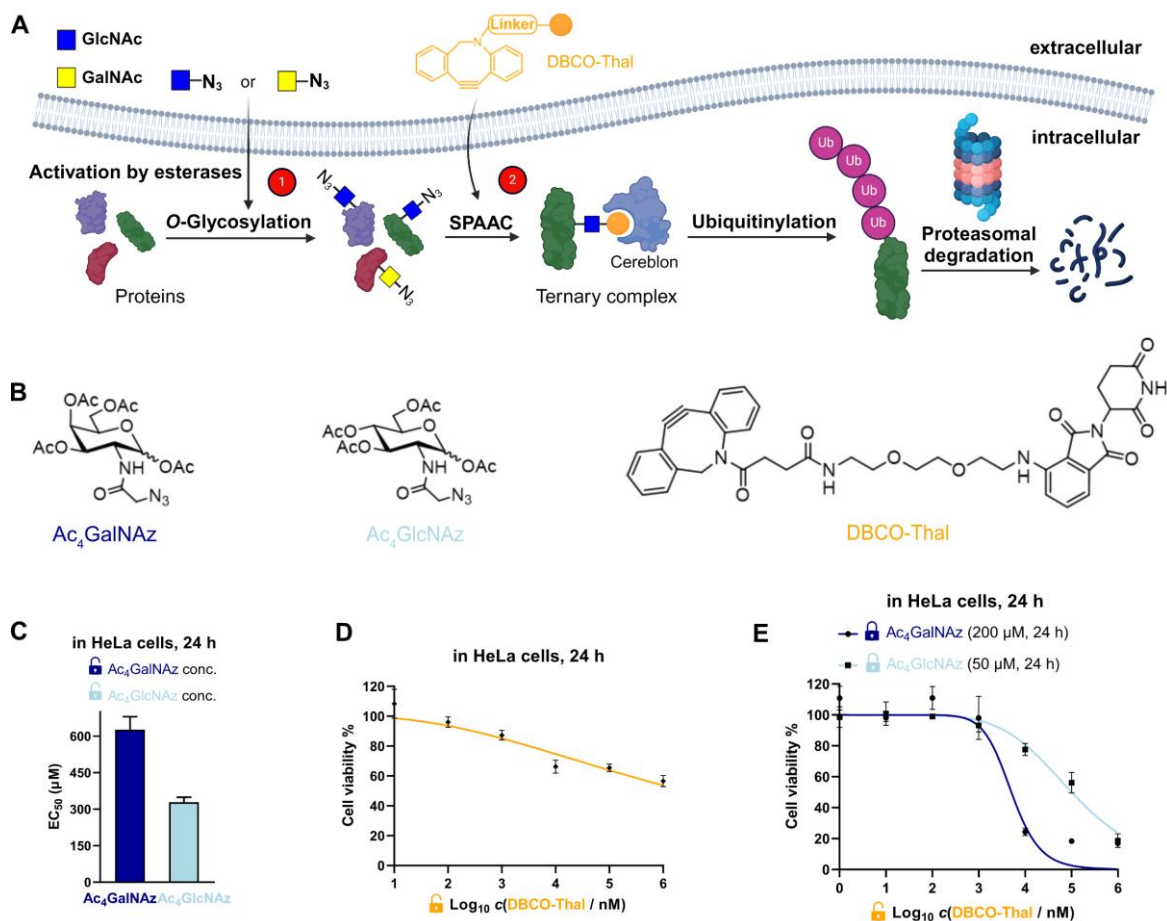


Figure 32. Overview of the GlyTAC strategy. **A** Scheme of the GlyTAC strategy. **B** Structures of Ac₄GalNAz, Ac₄GlcNAz and DBCO-thalidomide (**DBCO-Thal**) used in the study. **C** Bar plot comparing cytotoxicity of GalNAc and GlcNAc analogues in HeLa cells after 24 h incubation, $n = 3$. **D** Concentration-dependent cytotoxicity of DBCO-Thal in HeLa cells after 24 h incubation. Concentration range between 10 nM and 1 mM, $n = 3$. **E** Proof-of-concept experiment of two-component GlyTAC-induced cytotoxicity. DBCO-Thal concentration-dependent cytotoxicity after incubation with Ac₄GalNAz (200 µM, 24 h) or Ac₄GlcNAz (50 µM, 24 h) in HeLa cells, $n = 3$. **C – E** cytotoxicity was determined by MTT assay; concentrations listed were final concentration of each species in complete cell culture medium.

3.10 GlyTAC - Establishment

3.10.1 Synthesis of DBCO-Thal and Cytotoxicity Testing

To test the validity of two-component GlyTAC strategy, we selected two commonly used and commercially available glycosylation analogues Ac₄GalNAz and Ac₄GlcNAz (**Figure 32B**). These two probes were proved to be well-accepted by corresponding enzymatic writers and efficiently incorporated into O-glycosylated proteins (**Figure 32B**).^{201–204} Next, we have designed and synthesized the second component containing dibenzocyclooctyne-amine (DBCO-amine) conjugated with thalidomide to facilitate copper-free ‘click’ reaction and thus to assemble active PROTAC for CRBN recruitment (**Figure 32B**).^{57,205} The DBCO and thalidomide are linked through short ethylene glycol unit to improve the overall

solubility and leave necessary spatial distance, which allows the formation of ternary complex between thalidomide moiety and CRBN. The DBCO-Thal was synthesized sequentially by nucleophilic aromatic substitution (S_NAr) on 4-fluoro-thalidomide, followed by Boc-deprotection to release primary amine of the linker, and finally amide bond formation with readily available DBCO-NHS ester (§ 7.1.12). The overall isolated yield was above 64%.²⁰⁶ Next, MTT cell viability assay was used to evaluate the cytotoxicity of individual GlyTAC components including $Ac_4GalNAz$, $Ac_4GlcNAz$ and DBCO-Thal. After 24 h treatment in HeLa cells, the $Ac_4GalNAz$ gave EC_{50} value at 625 μM and $Ac_4GlcNAz$ at 328 μM , with no cytotoxicity up to 300 μM and 150 μM , respectively (**Figure 32C** and Figure S31). The DBCO-Thal was not cytotoxic up to 1 μM and showed only minor cytotoxicity up to 1 mM in cell culture medium. (**Figure 32D**). Analogous cell viability assays were carried out in Chinese hamster ovary (CHO) cells providing comparable results and corroborating low toxicity of all components (Figure S32). The metabolic incorporation efficiency of $Ac_4GalNAz$ was tested by in-gel fluorescence analysis after SPAAC with DBCO-TAMRA (Figure S33).²⁰⁷ Together, the cell viability assays showed wide concentration tolerance available for the application of proposed two-component system and hence the feasibility of GlyTAC approach.

3.10.2 GlyTAC - Proof-of-Principle

Given the complexity and dynamics of protein glycosylation, we next tested if the GlyTAC self-assembly via SPAAC would be sufficient to induce proteasomal degradation of glycoproteins. The degradation of glycoproteins in turn may induce cytotoxicity due to broad dysregulation of corresponding metabolic and signaling pathways. The HeLa cells were firstly incubated with $Ac_4GalNAz$ at non-toxic concentration of 200 μM for 24 h to reach saturation of azido-containing proteins, then the cells were exchanged with fresh medium containing increasing concentrations of DBCO-Thal. As shown in **Figure 32E**, such GlyTAC approach displayed rapid increase in cytotoxicity above 1 μM of DBCO-Thal and EC_{50} at 5 μM (**Figure 32E**). The parallel experiment with 50 μM $Ac_4GlcNAz$ treatment in the first step led to significantly higher EC_{50} of DBCO-Thal (approx. 100 μM), shown in a similar concentration-dependent but flattened cytotoxicity pattern when compared with $Ac_4GalNAz$ (**Figure 32E**). Together, the strong increase in cytotoxicity of DBCO-Thal after pre-treatment with $Ac_4GalNAz$ demonstrates promising functionality of GlyTAC approach.

3.11 GlyTAC - Characterization

To better understand and characterize the dynamic range of our two-component system, we tested the concentration- and time-dependencies of both glycosylation analogues and DBCO-Thal. First, at the fixed EC_{50} concentration and incubation time of DBCO-Thal, $Ac_4GalNAz$ and $Ac_4GlcNAz$ was titrated in HeLa cells respectively to show higher efficiency of $Ac_4GalNAz$ precursor (**Figure 33A**). While EC_{50} was achieved with $Ac_4GalNAz$ at 30 μM when coupled with 5 μM DBCO-Thal, $Ac_4GlcNAz$ showed much lower toxicity. Second, treatment time of DBCO-Thal was optimized with $Ac_4GalNAz$ and $Ac_4GlcNAz$ pre-incubation to display a major difference between two glycosylation analogues, requiring 14.5 h- and 27 h-treatment to reach EC_{50} (**Figure 33B**), respectively. Next, we evaluated the cell type-dependent cytotoxicity of DBCO-Thal in $Ac_4GlcNAz$ pre-treated HeLa and CHO cells, observing around two-fold lower EC_{50} in HeLa cells, which can be explained by lower glycosylation levels in CHO cells (**Figure 33C**).

Since many glycosylated proteins are located at plasma membrane with a glycan oriented towards extracellular space, we have considered to utilize a 4-deoxy GlcNAz analogue (1,3,6-tri-*O*-acetyl-4-deoxy-*N*-azidoacetyl-glucosamine, Ac₃4dGlcNAz, **Figure 3C**), which is mainly incorporated into intracellular proteins as the vacant C-4 hydroxy group restricts the prolongation of glycosylation chain (**Figure 33E-F**). Although Ac₃4dGlcNAz analogue showed promising efficacy, it was not enhancing activity enough to counterweight its synthetic accessibility. In comparison with synthetic routes to obtain Ac₄GalNAz (4 steps, § 7.1.13)^{208–211} and Ac₄GlcNAz (1 step) from cheap starting building blocks (~85 USD for 5 g), the published Ac₃4dGlcNAz synthetic route required 11 steps with rather low overall yield.^{211,212} Taken together, pre-treatment of Ac₄GalNAz at 200 μM for 24 h followed by conjugation with DBCO-Thal at 5 μM for 24 h indicated the most promising cytotoxicity effect in combination with the best accessibility.

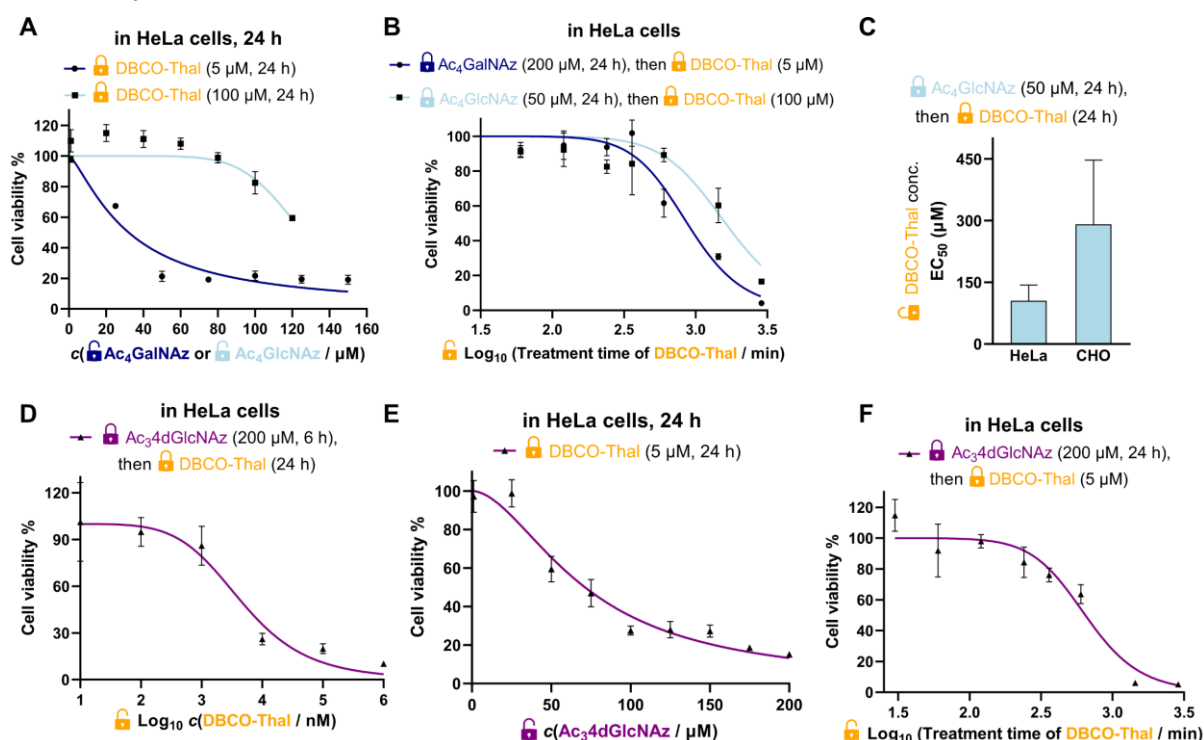


Figure 33. Characterization and optimization of GlyTAC system. **A)** Ac₄GalNAz and Ac₄GlcNAz concentration-dependency with fixed treatment conditions of DBCO-Thal in HeLa cells, $n = 3$. **B)** DBCO-Thal time-dependency in HeLa cells, $n = 3$. **C)** Cell type-dependency of the GlyTACs with fixed treatment conditions of Ac₄GlcNAz in HeLa or CHO cells, $n = 3$. **(D)** DBCO-Thal concentration-dependency with fixed treatment conditions of Ac₃4dGlcNAz in HeLa cells, $n = 3$, **(E)** Ac₃4dGlcNAz concentration-dependency with fixed treatment conditions of DBCO-Thal in HeLa cells, $n = 3$, **(F)** Time-dependency of DBCO-Thal with fixed treatment conditions of Ac₃4dGlcNAz in HeLa cells, $n = 3$.

3.11.1 GlyTACs - Proteome-Wide Ubiquitinylation and Nup62 Degradation

The anticipated mechanism of action is that the glycoproteins are firstly modified with azide-containing GalNAc analogue, which are subsequently conjugated via SPAAC with thalidomide to recruit CRBN and trigger downstream ubiquitinylation. The critical step leading to efficient degradation of tagged proteins is therefore adequate ubiquitinylation. Given the predicted proteome-wide modification of glycoproteins

with thalidomide in proximity, we expected a global increase in ubiquitinylation. Indeed, immunoblotting with anti-polyubiquitin antibody showed marked ubiquitinylation increase in concentration- and time-dependent manner of both components (**Figure 34A-C**). Next, to validate our approach, we selected well-described nuclear pore glycoprotein p62 (Nup62) as immunoblotting target, which is heavily O-GlcNAcylated together with other nuclear proteins.²¹³ As expected, Nup62 demonstrated a significant decrease in a similar concentration- and time-dependent manner when subjected to two-component GlyTACs treatment (**Figure 34D-F**). This could be explained by C-4 hydroxyl group epimerization of Ac₄GalNAz. The 4'-epimerase GALE efficiently transforms UDP-(uridine diphosphate)-GalNAz substrate to UDP-GlcNAz, which is accepted by O-GlcNAc transferase (OGT) to initiate protein O-GlcNAcylation.⁴⁵ Following the concentration- and time-dependent degradation of Nup62 using Ac₄GalNAz demonstrates the significant decrease of the protein caused by GlyTAC approach (**Figure 34D-F**). Gratifyingly, there was no observable change in Nup62 levels when treated solely with the Ac₄GalNAz or DBCO-Thal (**Figure 34D-F**). Together, the significant Nup62 degradation and global increase in ubiquitinylation validated GlyTACs mechanism of action based on protein glycosylation and CRBN recruitment.

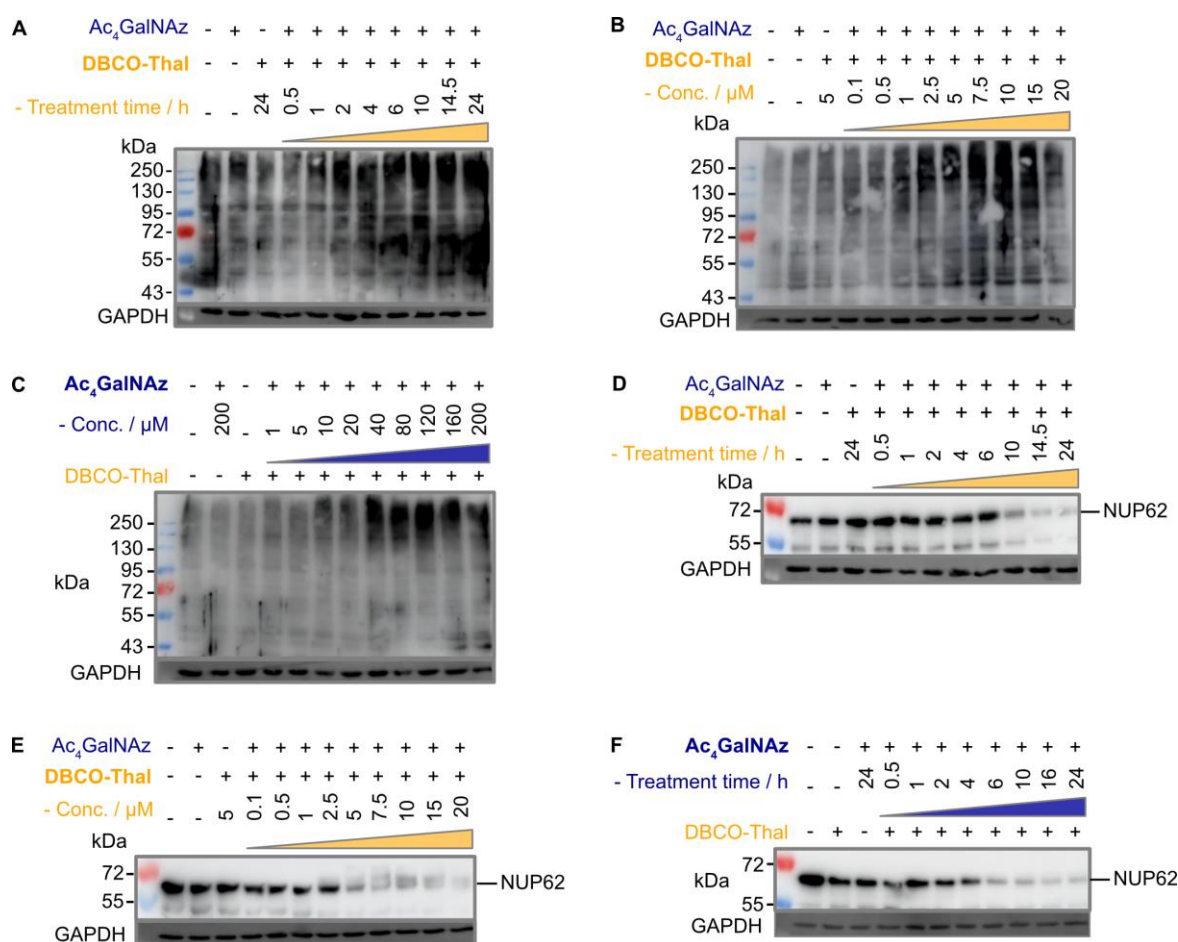


Figure 34. GlyTACs increase protein ubiquitinylation and Nup62 degradation. **A–C** Western blots using anti-polyubiquitin antibody of **A**) DBCO-Thal time-dependent protein ubiquitinylation, Ac₄GalNAz was constant at 200 μM, **B**) DBCO-Thal concentration-dependent protein ubiquitinylation, Ac₄GalNAz was constant at 200 μM and **C**) Ac₄GalNAz concentration-dependent protein ubiquitinylation, DBCO-Thal was constant at 5 μM. **D - F** Western blot using anti-Nup62 antibody of **D**) DBCO-Thal time-dependent, Ac₄GalNAz was constant at 200 μM **E**) DBCO-Thal concentration-dependent, Ac₄GalNAz was constant at 200 μM and **F**) Ac₄GalNAz concentration-dependent, DBCO-Thal was constant at 5 μM.

3.12 GlyTACs - Mechanism of Action

To further investigate the GlyTACs mechanism of action, we functionally probed each step of the process. The first critical step in GlyTAC system is the incorporation of Ac₄GalNAz or Ac₄GlcNAz into glycoproteins. In the case of Ac₄GlcNAz, the incorporation is catalyzed by OGT, which can be inhibited by a small molecule inhibitor OSMI-1.⁴² To test whether the OSMI-1 abolishes GlyTAC cytotoxicity, HeLa cells were pretreated with the inhibitor before addition of Ac₄GlcNAz. As expected, subsequent treatment with DBCO-Thal did not result in cytotoxic activity due to OGT inhibition (**Figure 35A**). In contrast, OSMI-1 treatment had no impact on Ac₄GalNAz-based GlyTAC as the O-GalNAcylation is not catalyzed by OGT, but O-GalNAc transferases (**Figure 35B**). Although the epimerization of GalNAc analogue occurred, as we showed for Nup62 degradation, overall, it was not yielding enough O-GlcNAz modified proteins to retard cytotoxicity with OSMI-1 treatment.

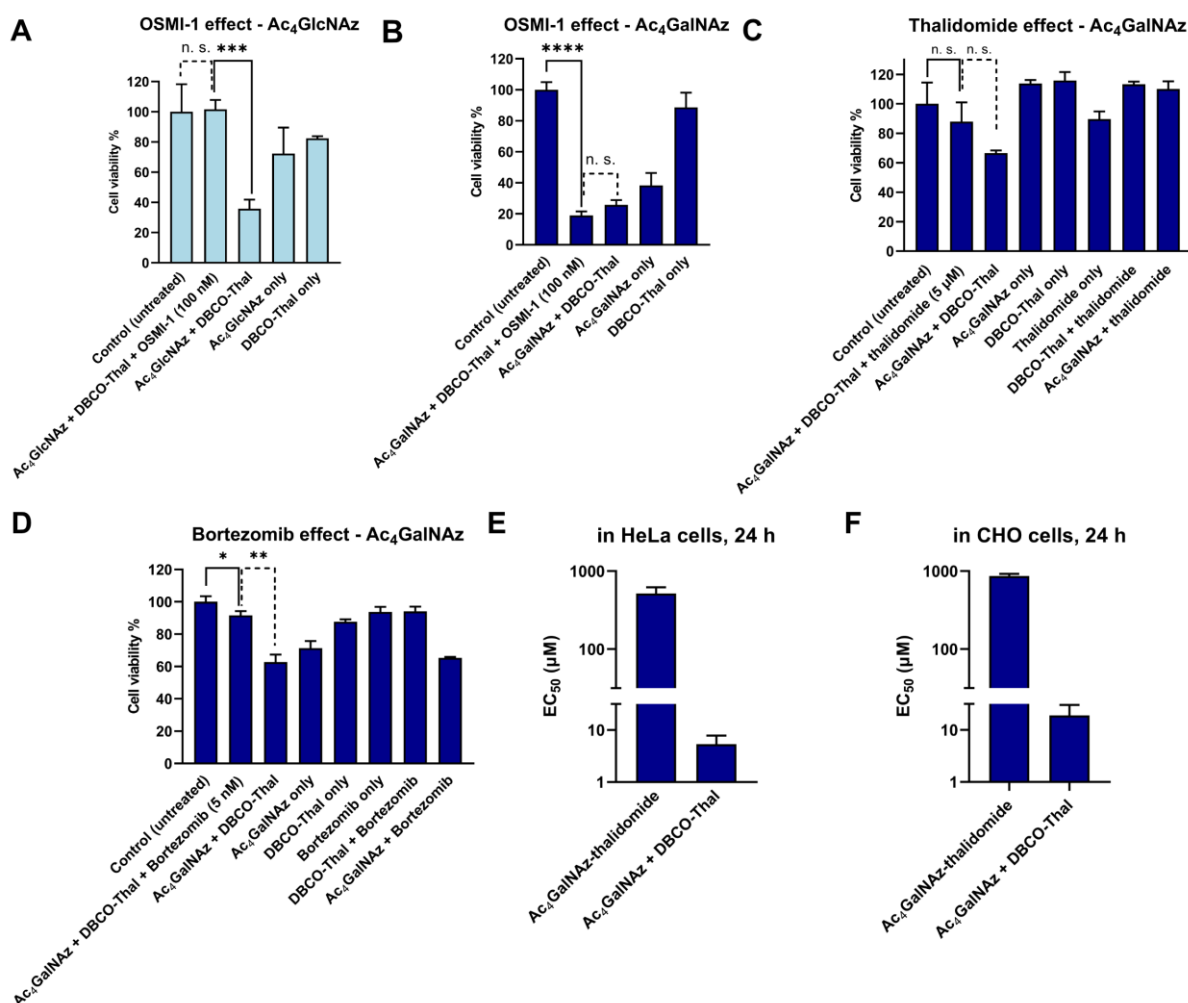


Figure 35. GlyTACs mechanism of action. **A)** OSMI-1 induced O-GlcNAcylation inhibition rescued GlyTAC-triggered cell death. **B)** OSMI-1 did not rescue O-GalNAcylation-induced cell death. **C)** Competition between DBCO-Thal and thalidomide suggested CRBN-involved mechanism of action. **D)** Proteasome inhibition by bortezomib rescued GlyTAC-triggered cell death. **E** and **F)** The sequential treatment of two components was required to induce GlyTAC-triggered cell death as the cells treated with Ac₄GalNAz-thalidomide conjugate showed negligible cytotoxicity in HeLa (**E**) and CHO (**F**) cells. All data were acquired in triplicates (**A, B** and **D-F**) or six replicates (**C**).

After SPAAC between azido-modified glycoproteins with DBCO-Thal, the next step in the GlyTAC strategy was the recruitment of CRBN to form a ternary complex. The interaction should be in principle outcompeted by the pretreatment of thalidomide. The parallel incubation of Ac₄GalNAz-treated cells with DBCO-Thal and additional 5 μM thalidomide led to partial rescue of GlyTAC cytotoxicity (**Figure 35C**). However, thalidomide competition only functioned within a limited time range (approx. 8 – 10 h), as the cytotoxicity after standard treatment time of DBCO-Thal (24 h) has always brought significant cell death. After the CRBN recruitment to glycoproteins, they are targeted for degradation by UPS. The proteasomal protein degradation can be inhibited by bortezomib, an anti-cancer small molecule drug.^{214,215} To probe that the proteasomal activity is responsible for protein degradation in the GlyTAC system, we used 5 nM bortezomib to rescue the cytotoxicity of Ac₄GalNAz-based GlyTAC (**Figure 35D**). Furthermore, it might be in principle possible to directly treat the cells with a Ac₄GalNAz-thalidomide conjugate to achieve the GlyTAC cytotoxicity. Therefore, we synthesized the Ac₄GalNAz-thalidomide linked through a triazole ring (§ 7.1.14) and used it for cell treatment, which resulted in virtually no cytotoxicity in both HeLa and CHO cells (**Figure 35E-F**). Taken together, the functional manipulation of enzymatic writer OGT, thalidomide competition for CRBN, and proteasome inhibition by bortezomib provide insight in GlyTAC mechanism of action. The comparison between efficacy of Ac₄GalNAz-thalidomide (single PROTAC compound) and sequential treatment in cells shows advantage of the two-component GlyTAC approach.

3.13 GlyTACs - Proteome-Wide Metabolic Alternation

The broad incorporation of Ac₄GalNAz into glycoproteins induces significant protein degradation in living cells after DBCO-Thal addition. To characterize changes on whole proteome level triggered by GlyTACs, a series of analyses were performed using quantitative mass spectrometry-based proteomics (**Figure 36A**).^{168,185,216–218} In the control experiments where Ac₄GalNAz or DBCO-Thal was separately treated, we observed only minor proteome changes. However, application of both components with increasing concentration of DBCO-Thal showed anticipated concentration-dependent protein down-regulation (**Figure 36B**). Further analysis revealed that from total 7187 identified proteins, 109 proteins were down-regulated after 5 μM DBCO-Thal treatment (**Figure 36C**). Increasing DBCO-Thal concentration to 20 μM led to more pronounced protein down-regulation of 433 proteins (**Figure 36C**). Comparison of both showed overlap of 91 proteins (**Figure 36C**). The proteomics also confirmed down-regulation of Nup62 and three other nuclear pore proteins including NUP214, NUP88 and NUP98 (**Figure 36D**). To find how many down-regulated proteins were known to be O-GlcNAcylated, which may result from GALE epimerase activity during Ac₄GalNAz treatment, we compared the list of the down-regulated proteins with O-GlcNAc database (**Figure 36E**). The search revealed overall down-regulation of 311 known O-GlcNAcylated proteins with 20 μM DBCO-Thal treatment (**Figure 36E**). The KEGG pathway analysis mainly showed up-regulated proteins associated with carbohydrates metabolism and HIF-1 signaling pathways, suggesting dysregulation of glycoproteins deepening hypoxia stress (**Figure 36F and G**). On the other hand, we observed up-regulation of small portion of 14 proteins including OGT and CRBN (**Figure 36D**). The proteomic analysis regarding variation of other treatment conditions, including treatment time of DBCO-Thal, the treatment time or concentration of Ac₄GalNAz, are available in Figure S34-36. Together, the proteomics analysis suggests major down-regulation of glycoproteins.

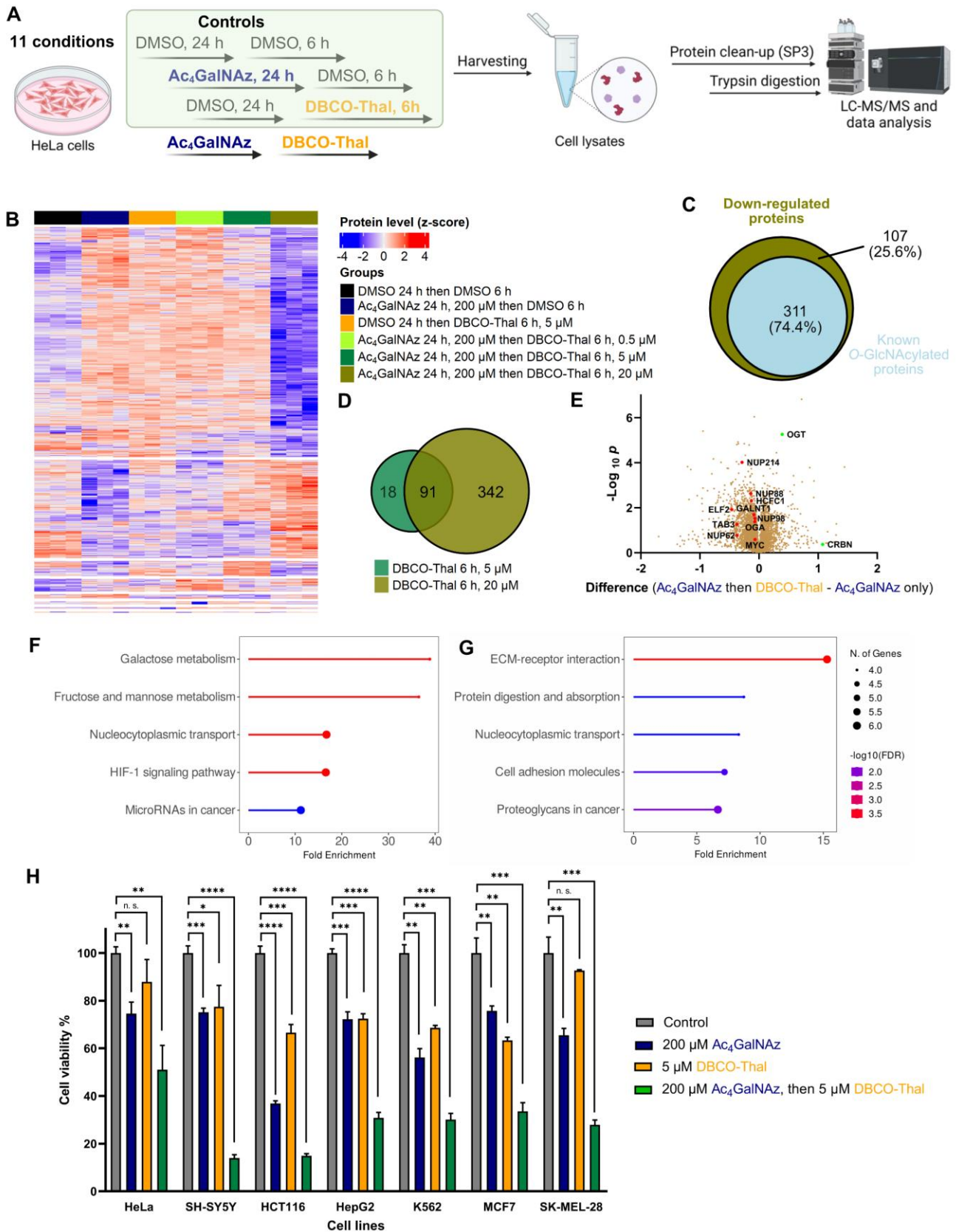


Figure 36. GlyTACs induced whole proteome changes and cancer cell screening. **A)** Experimental design of proteomics analysis. **B)** Heatmap visualizing up- and down-regulated proteins in all conditions using ANOVA-significant test. **C)** Venn diagram showing the overlap of down-regulated proteins in 5 μ M and 20 μ M DBCO-Thal treatment. **D)** Volcano plot comparing GlyTAC (200 μ M Ac₄GalNAz then 5 μ M DBCO-Thal) against separate Ac₄GalNAz treatment as a control. **E)** Venn diagram showing the percentage of known O-GlcNAcylated protein contained in significantly down-regulated proteins after GlyTAC treatment. **F, G)**

KEGG pathway analysis of significantly up-regulated (**F**) and down-regulated (**G**) proteins. **H**) Screening of 7 human cancer lines after GlyTAC treatment.

3.14 GlyTACs - Screening in Cancer Cell Lines

To better understand susceptibility of different cancer cell lines to sequential GlyTAC treatment, seven human cancer cell lines were screened using fixed concentrations of Ac₄GalNAz (200 μM) and DBCO-Thal (5 μM, **Figure 36H**). The cell lines included HeLa, SH-SY5Y, HCT116, HepG2, K562, MCF7 and SK-MEL-28. Overall, in all tested cancer cell lines, significant cytotoxicity of two-component GlyTACs was observed in comparison to blank control and treatments with Ac₄GalNAz or DBCO-Thal separately (**Figure 36H**). The most pronounced effect was determined in neuroblastoma cells (SH-SY5Y). In other cell lines, relatively high concentration of Ac₄GalNAz at 200 μM showed somewhat increased cytotoxicity, while DBCO-Thal remained in the same or slightly lower cytotoxicity range. The treatment window might be further extended in stressed or hypoxic cells, which show increased proteasome activity.²¹⁹ Optimization of treatment conditions in each cancer cell line would likely improve the efficiency of GlyTACs. Together, the presented testing of the GlyTAC strategy in different human cancer cell lines demonstrates feasibility of the approach.

We have established the two-component GlyTAC approach, which induces severe cytotoxicity in human cancer cell lines, while treatment with each component separately shows only minor effect. The system takes advantage of glycoproteins accessibility via metabolic incorporation of a PTM mimics containing an azido group, which is used in the second step to covalently anchor DBCO-Thal through SPAAC and subsequently leads to proteasomal degradation of such proteins. Although the covalent bonds between the degrader moiety and POIs prevent the GlyTACs from proceeding catalytically like other PROTAC modalities, we concluded that low cytotoxicity of the individual components overcomes this problem. More importantly, it largely improves the shortcomings of single PROTAC molecules, especially poor aqueous solubility and cell permeability. The efficiency of GlyTACs stems from wide-spread glycoproteins with azido group and self-assembly with thalidomide moiety by SPAAC. Protein PTMs enable specific subset regulation of the total protein pool, often in cell-type dependent manner, providing thus far unused lever to manipulate physiological process. Further understanding of substrate selectivity of PTM writers is needed to target selected proteins. Ongoing studies in our laboratory are testing this idea.

Overall, protein PTMs play a crucial role in regulating protein function, stability, and interactions. The study of PTM analogues, particularly small molecule mimics, is essential for understanding the underlying mechanisms of related human diseases. These analogues serve as valuable tools for investigating the causes and processes involved in pathological conditions and provide insights into potential therapeutic strategies. By mimicking or inhibiting specific PTMs, researchers can develop targeted treatments aimed at correcting the abnormal protein functions or selectively degrading pathogenic proteins, offering new avenues for precision medicine.

4. Bibliography

1. Aebersold, R. *et al.* How many human proteoforms are there? *Nat. Chem. Biol.* **14**, 206–214 (2018).
2. Tress, M. L., Abascal, F. & Valencia, A. Alternative Splicing May Not Be the Key to Proteome Complexity. *Trends Biochem. Sci.* **42**, 98–110 (2017).
3. Bugyi, F. *et al.* Influence of Post-Translational Modifications on Protein Identification in Database Searches. *ACS Omega* **6**, 7469–7477 (2021).
4. Dunphy, K., Dowling, P., Bazou, D. & O’Gorman, P. Current methods of post-translational modification analysis and their applications in blood cancers. *Cancers (Basel)*. **13**, (2021).
5. Keenan, E. K., Zachman, D. K. & Hirschey, M. D. Discovering the landscape of protein modifications. *Mol. Cell* **81**, 1868–1878 (2021).
6. Bobalova, J., Strouhalova, D. & Bobal, P. Common Post-translational Modifications (PTMs) of Proteins: Analysis by Up-to-Date Analytical Techniques with an Emphasis on Barley. *J. Agric. Food Chem.* **71**, 14825–14837 (2023).
7. Ramazi, S. & Zahiri, J. Post-translational modifications in proteins: Resources, tools and prediction methods. *Database* **2021**, 1–20 (2021).
8. Lee, J. M., Hammarén, H. M., Savitski, M. M. & Baek, S. H. Control of protein stability by post-translational modifications. *Nat. Commun.* **14**, 1–16 (2023).
9. Zhong, Q. *et al.* Protein posttranslational modifications in health and diseases: Functions, regulatory mechanisms, and therapeutic implications. *MedComm* **4**, 1–112 (2023).
10. Leutert, M., Entwisle, S. W. & Villén, J. Decoding post-translational modification crosstalk with proteomics. *Mol. Cell. Proteomics* **20**, 0–11 (2021).
11. Suskiewicz, M. J. The logic of protein post-translational modifications (PTMs): Chemistry, mechanisms and evolution of protein regulation through covalent attachments. *BioEssays* **46**, 1–25 (2024).
12. Ubersax, J. A. & Ferrell, J. E. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* **8**, 530–541 (2007).
13. Bilbrough, T., Piemontese, E. & Seitz, O. Dissecting the role of protein phosphorylation: a chemical biology toolbox. *Chem. Soc. Rev.* **51**, 5691–5730 (2022).
14. Shang, S., Liu, J. & Hua, F. Protein acylation: mechanisms, biological functions and therapeutic targets. *Signal Transduct. Target. Ther.* **7**, (2022).
15. Narita, T., Weinert, B. T. & Choudhary, C. Functions and mechanisms of non-histone protein acetylation. *Nat. Rev. Mol. Cell Biol.* **20**, 156–174 (2019).
16. Murn, J. & Shi, Y. The winding path of protein methylation research: Milestones and new frontiers. *Nat. Rev. Mol. Cell Biol.* **18**, 517–527 (2017).
17. Wu, Q., Schapira, M., Arrowsmith, C. H. & Barsyte-Lovejoy, D. Protein arginine methylation: from enigmatic functions to therapeutic targeting. *Nat. Rev. Drug Discov.* **20**, 509–530 (2021).
18. Yang, Y. S. *et al.* Tyrosine sulfation as a protein post-translational modification. *Molecules* **20**, 2138–2164 (2015).
19. Johnson, R. Controlling sulfation. *Nat. Chem. Biol.* **19**, 1172 (2023).
20. Jiang, H. *et al.* Protein Lipidation: Occurrence, Mechanisms, Biological Functions, and Enabling Technologies. *Chem. Rev.* **118**, 919–988 (2018).
21. Tate, E. W., Soday, L., de la Lastra, A. L., Wang, M. & Lin, H. Protein lipidation in cancer: mechanisms, dysregulation and emerging drug targets. *Nat. Rev. Cancer* **24**, 240–260 (2024).
22. Prabhu, H. *et al.* Protein feature engineering framework for AMPylation site prediction. *Sci. Rep.* **14**, 1–17 (2024).
23. Schjoldager, K. T., Narimatsu, Y., Joshi, H. J. & Clausen, H. Global view of human protein

- glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **21**, 729–749 (2020).
24. Reily, C., Stewart, T. J., Renfrow, M. B. & Novak, J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* **15**, 346–366 (2019).
 25. Sahin, U., de Thé, H. & Lallemand-Breitenbach, V. Sumoylation in Physiology, Pathology and Therapy. *Cells* **11**, 1–24 (2022).
 26. Celen, A. B. & Sahin, U. Sumoylation on its 25th anniversary: mechanisms, pathology, and emerging concepts. *FEBS J.* **287**, 3110–3140 (2020).
 27. Huang, C. H., Yang, T. T. & Lin, K. I. Mechanisms and functions of SUMOylation in health and disease: a review focusing on immune cells. *J. Biomed. Sci.* **31**, 1–14 (2024).
 28. Sun, M. & Zhang, X. Current methodologies in protein ubiquitination characterization: from ubiquitinated protein to ubiquitin chain architecture. *Cell Biosci.* **12**, 1–17 (2022).
 29. Damgaard, R. B. The ubiquitin system: from cell signalling to disease biology and new therapeutic opportunities. *Cell Death Differ.* **28**, 423–426 (2021).
 30. Fu, J., Gao, J., Liang, Z. & Yang, D. PDI-regulated disulfide bond formation in protein folding and biomolecular assembly. *Molecules* **26**, (2021).
 31. Khan, H. A. & Mutus, B. Protein disulfide isomerase a multifunctional protein with multiple physiological roles. *Front. Chem.* **2**, 1–9 (2014).
 32. Jin, Y., Yi, Y. & Yeung, B. Mass spectrometric analysis of protein deamidation – A focus on top-down and middle-down mass spectrometry. *Methods* **200**, 58–66 (2022).
 33. Rogers, L. D. & Overall, C. M. C. Proteolytic post-translational modification of proteins: Proteomic tools and methodology. *Mol. Cell. Proteomics* **12**, 3532–3542 (2013).
 34. Zhao, L., Zhao, J., Zhong, K., Tong, A. & Jia, D. Targeted protein degradation: mechanisms, strategies and application. *Signal Transduct. Target. Ther.* **7**, (2022).
 35. Hu, W. *et al.* Recent development of analytical methods for disease-specific protein O-GlcNAcylation. *RSC Adv.* **13**, 264–280 (2022).
 36. Lee, J. B., Pyo, K. H. & Kim, H. R. Role and function of o-glcnaacylation in cancer. *Cancers (Basel)*. **13**, 1–15 (2021).
 37. Yang, X. & Qian, K. Protein O-GlcNAcylation: Emerging mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* **18**, 452–465 (2017).
 38. Bond, M. R. & Hanover, J. A. A little sugar goes a long way: The cell biology of O-GlcNAc. *J. Cell Biol.* **208**, 869–880 (2015).
 39. Hart, G. W., Housley, M. P. & Slawson, C. Cycling of O-linked β -N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **446**, 1017–1022 (2007).
 40. Hart, G. W., Slawson, C., Ramirez-Correa, G. & Lagerlof, O. Cross Talk between O-GlcNAcylation and phosphorylation: Roles in signaling, transcription, and chronic disease. *Annu. Rev. Biochem.* **80**, 825–858 (2011).
 41. Gloster, T. M. *et al.* Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. *Nat. Chem. Biol.* **7**, 174–181 (2011).
 42. Ortiz-Meoz, R. F. *et al.* A Small Molecule That Inhibits OGT Activity in Cells. *ACS Chem. Biol.* **10**, 1392–1397 (2015).
 43. Pedowitz, N. J., Zaro, B. W. & Pratt, M. R. Metabolic Labeling for the Visualization and Identification of Potentially O-GlcNAc-Modified Proteins. *Curr. Protoc. Chem. Biol.* **12**, (2020).
 44. Lauehlin, S. T. & Bertozzi, C. R. Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via staudinger ligation. *Nat. Protoc.* **2**, 2930–2944 (2007).
 45. Boyce, M. *et al.* Metabolic cross-talk allows labeling of O-linked β -N- acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3141–3146 (2011).
 46. Hu, C. M. *et al.* High Glucose Triggers Nucleotide Imbalance through O-GlcNAcylation of Key

- Enzymes and Induces KRAS Mutation in Pancreatic Cells. *Cell Metab.* **29**, 1334-1349.e10 (2019).
47. Mi, W. *et al.* O-GlcNAcylation is a novel regulator of lung and colon cancer malignancy. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1812**, 514–519 (2011).
 48. Chang, Y. H., Weng, C. L. & Lin, K. I. O-GlcNAcylation and its role in the immune system. *J. Biomed. Sci.* **27**, 1–15 (2020).
 49. Yue, S. *et al.* Deciphering Protein O-GalNAcylation: Method Development and Disease Implication. *ACS Omega* **8**, 19223–19236 (2023).
 50. Hang, H. C., Yu, C., Kato, D. L. & Bertozzi, C. R. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 14846–14851 (2003).
 51. Wang-Gillam, A., Pastuszak, I. & Elbein, A. D. A 17-amino acid insert changes UDP-N-acetylhexosamine pyrophosphorylase specificity from UDP-GalNAc to UDP-GlcNAc. *J. Biol. Chem.* **273**, 27055–27057 (1998).
 52. Daude, N. *et al.* Molecular-Cloning, characterization, and mapping of a full-length cDNA encoding human UDP-Galactose 4'-Epimerase. *Biochemical and Molecular Medicine* vol. 56 1–7 at <https://doi.org/10.1006/bmme.1995.1048> (1995).
 53. Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-Triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **67**, 3057–3064 (2002).
 54. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew. Chemie - Int. Ed.* **41**, 2596–2599 (2002).
 55. Shah, L., Laughlin, S. T. & Carrico, I. S. Light-Activated Staudinger-Bertozzi Ligation within Living Animals. *J. Am. Chem. Soc.* **138**, 5186–5189 (2016).
 56. Saxon, E. & Bertozzi, C. R. Cell Surface Engineering by a Modified Staudinger Reaction. *Science (80-.)*. **287**, 2007–2010 (2000).
 57. Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126**, 15046–15047 (2004).
 58. Baskin, J. M. *et al.* Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16793–16797 (2007).
 59. Dommerholt, J., Rutjes, F. P. J. T. & van Delft, F. L. Strain-Promoted 1,3-Dipolar Cycloaddition of Cycloalkynes and Organic Azides. *Top. Curr. Chem.* **374**, 1–20 (2016).
 60. Ning, X., Guo, J., Wolfert, M. A. & Boons, G. J. Visualizing metabolically labeled glycoconjugates of living cells by copper-free and fast Huisgen cycloadditions. *Angew. Chemie - Int. Ed.* **47**, 2253–2255 (2008).
 61. Jewett, J. C., Sletten, E. M. & Bertozzi, C. R. Rapid Cu-free click chemistry with readily synthesized biarylazacyclooctynones. *J. Am. Chem. Soc.* **132**, 3688–3690 (2010).
 62. Debets, M. F., Van Der Doelen, C. W. J., Rutjes, F. P. J. T. & Van Delft, F. L. Azide: A unique dipole for metal-free bioorthogonal ligations. *ChemBioChem* **11**, 1168–1184 (2010).
 63. McGrath, N. A. & Raines, R. T. Diazo compounds as highly tunable reactants in 1,3-dipolar cycloaddition reactions with cycloalkynes. *Chem. Sci.* **3**, 3237–3240 (2012).
 64. Gutmiedl, K., Wirges, C. T., Ehmke, V. & Carell, T. Copper-free 'click' modification of DNA via nitrile oxide-norbornene 1,3-dipolar cycloaddition. *Org. Lett.* **11**, 2405–2408 (2009).
 65. Blackman, M. L., Royzen, M. & Fox, J. M. Tetrazine ligation: Fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J. Am. Chem. Soc.* **130**, 13518–13519 (2008).
 66. Devaraj, N. K., Weissleder, R. & Hilderbrand, S. A. Tetrazine-based cycloadditions: Application to pretargeted live cell imaging. *Bioconjug. Chem.* **19**, 2297–2299 (2008).

67. Knall, A. C. & Slugovc, C. Inverse electron demand Diels-Alder (iEDDA)-initiated conjugation: A (high) potential click chemistry scheme. *Chem. Soc. Rev.* **42**, 5131–5142 (2013).
68. Taylor, M. T., Blackman, M. L., Dmitrenko, O. & Fox, J. M. Design and synthesis of highly reactive dienophiles for the tetrazine-trans-cyclooctene ligation. *J. Am. Chem. Soc.* **133**, 9646–9649 (2011).
69. Darko, A. *et al.* Conformationally strained trans-cyclooctene with improved stability and excellent reactivity in tetrazine ligation. *Chem. Sci.* **5**, 3770–3776 (2014).
70. Devaraj, N. K., Hilderbrand, S., Upadhyay, R., Mazitschek, R. & Weissleder, R. Bioorthogonal turn-on probes for imaging small molecules inside living cells. *Angew. Chemie - Int. Ed.* **49**, 2869–2872 (2010).
71. Carlson, J. C. T., Meimetis, L. G., Hilderbrand, S. A. & Weissleder, R. BODIPY-tetrazine derivatives as superbright bioorthogonal turn-on probes. *Angew. Chemie - Int. Ed.* **52**, 6917–6920 (2013).
72. Kalia, J. & Raines, R. T. Hydrolytic stability of hydrazones and oximes. *Angew. Chemie - Int. Ed.* **47**, 7523–7526 (2008).
73. Rideout, D. Self-Assembling Cytotoxins. *Science (80-.)*. **233**, 561–563 (1986).
74. Agouridas, V. *et al.* Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev.* **119**, (2019).
75. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. H. Synthesis of Proteins by Native Chemical Ligation. *Science (80-.)*. **266**, 776–779 (1994).
76. Lim, R. K. V. & Lin, Q. Photoinducible bioorthogonal chemistry: A spatiotemporally controllable tool to visualize and perturb proteins in live cells. *Acc. Chem. Res.* **44**, 828–830 (2011).
77. Scinto, S. L. *et al.* Bioorthogonal chemistry. *Nat. Rev. Methods Prim.* **1**, 30 (2021).
78. Martin, R. & Buchwald, S. L. Palladium-catalyzed suzuki-miyaura cross-coupling reactions employing dialkylbiaryl phosphine ligands. *Acc. Chem. Res.* **41**, 1461–1473 (2008).
79. Farhang, M., Akbarzadeh, A. R., Rabbani, M. & Ghadiri, A. M. A retrospective-prospective review of Suzuki–Miyaura reaction: From cross-coupling reaction to pharmaceutical industry applications. *Polyhedron* **227**, 116124 (2022).
80. George B. Smith, George C. Dezeny, David L. Hughes, Anthony O. King, and T. R. V. Mechanistic Studies of the Suzuki Cross-Coupling Reaction. *J. Am. Chem. Soc.* **59**, 8151–8156 (1994).
81. D’Alterio, M. C. *et al.* Mechanistic Aspects of the Palladium-Catalyzed Suzuki-Miyaura Cross-Coupling Reaction. *Chem. - A Eur. J.* **27**, 13481–13493 (2021).
82. Szilvási, T. & Veszprémi, T. Internal catalytic effect of bulky NHC ligands in Suzuki-Miyaura cross-coupling reaction. *ACS Catal.* **3**, 1984–1991 (2013).
83. Kaloğlu, N. & Özdemir, İ. PEPPSI-Pd-NHC catalyzed Suzuki-Miyaura cross-coupling reactions in aqueous media. *Tetrahedron* **75**, 2306–2313 (2019).
84. Chalker, J. M., Wood, C. S. C. & Davis, B. G. A convenient catalyst for aqueous and protein Suzuki-Miyaura cross-coupling. *J. Am. Chem. Soc.* **131**, 16346–16347 (2009).
85. Spicer, C. D., Triemer, T. & Davis, B. G. Palladium-Mediated Cell-Surface Labeling. *J. Am. Chem. Soc.* **134**, 800–803 (2012).
86. Ma, X., Wang, H. & Chen, W. N-heterocyclic carbene-stabilized palladium complexes as organometallic catalysts for bioorthogonal cross-coupling reactions. *J. Org. Chem.* **79**, 8652–8658 (2014).
87. Yusop, R. M., Unciti-Broceta, A., Johansson, E. M. V., Sánchez-Martín, R. M. & Bradley, M. Palladium-mediated intracellular chemistry. *Nat. Chem.* **3**, 239–243 (2011).
88. Indrigo, E., Clavadetscher, J., Chankeshwara, S. V., Lilienkampf, A. & Bradley, M. Palladium-mediated in situ synthesis of an anticancer agent. *Chem. Commun.* **52**, 14212–14214 (2016).

89. Cao, J. *et al.* Multiplexed CuAAC Suzuki-Miyaura Labeling for Tandem Activity-Based Chemoproteomic Profiling. *Anal. Chem.* **93**, 2610–2618 (2021).
90. Kadarauch, M., Whalley, D. M. & Phipps, R. J. sSPhos: A General Ligand for Enantioselective Arylative Phenol Dearomatization via Electrostatically-Directed Palladium Catalysis. *J. Am. Chem. Soc.* **145**, 25553–25558 (2023).
91. Kvasovs, N. & Gevorgyan, V. Contemporary methods for generation of aryl radicals. *Chem. Soc. Rev.* **50**, 2244–2259 (2021).
92. Kindt Markus R., S. H. Recent Advances in Meerwein Arylation Chemistry. *Synthesis (Stuttg)*. **48**, 1597–1606 (2016).
93. Romero, N. A. & Nicewicz, D. A. Organic Photoredox Catalysis. *Chem. Rev.* **116**, 10075–10166 (2016).
94. Shaw, M. H., Twilton, J. & MacMillan, D. W. C. Photoredox Catalysis in Organic Chemistry. *J. Org. Chem.* **81**, 6898–6926 (2016).
95. Liu, W., Yang, X., Gao, Y. & Li, C. J. Simple and Efficient Generation of Aryl Radicals from Aryl Triflates: Synthesis of Aryl Boronates and Aryl Iodides at Room Temperature. *J. Am. Chem. Soc.* **139**, 8621–8627 (2017).
96. Jolly, P. I. *et al.* Reactions of triflate esters and triflamides with an organic neutral super-electron-donor. *Org. Biomol. Chem.* **10**, 5807–5810 (2012).
97. Liu, Q. *et al.* A general electrochemical strategy for the Sandmeyer reaction. *Chem. Sci.* **9**, 8731–8737 (2018).
98. Nguyen, J. D., D'Amato, E. M., Narayanam, J. M. R. & Stephenson, C. R. J. Engaging unactivated alkyl, alkenyl and aryl iodides in visible-light-mediated free radical reactions. *Nat. Chem.* **4**, 854–859 (2012).
99. Tian, Y. M. *et al.* Enforced Electronic-Donor-Acceptor Complex Formation in Water for Photochemical Cross-Coupling. *Angew. Chemie - Int. Ed.* **62**, (2023).
100. Frézard, F. *et al.* Enhanced oral delivery of antimony from meglumine antimoniate/ β -cyclodextrin nanoassemblies. *Int. J. Pharm.* **347**, 102–108 (2008).
101. Aloisio, C., de Oliveira, A. G. & Longhi, M. Solubility and release modulation effect of sulfamerazine ternary complexes with cyclodextrins and meglumine. *J. Pharm. Biomed. Anal.* **100**, 64–73 (2014).
102. Sravya, G. *et al.* A meglumine catalyst-based synthesis, molecular docking, and antioxidant studies of dihydropyrano[3, 2-b]chromenedione derivatives. *J. Heterocycl. Chem.* **57**, 355–369 (2020).
103. Sivaev, I. B. & Bregadze, V. I. Lewis acidity of boron compounds. *Coord. Chem. Rev.* **270–271**, 75–88 (2014).
104. Jehan A. Baban, Nicholas J. Goodchild, and B. P. R. Electron spin resonance studies of radicals derived from 1,3,2-benzodioxaboroles. *J. Chem. Soc. Perkin Trans. 2* 157–161 (1986).
105. Renaud, P. *et al.* Boron: A key element in radical reactions. *Pure Appl. Chem.* **79**, 223–233 (2007).
106. Davies, A. G. Homolytic reactions of organoboranes. *Pure Appl. Chem.* **39**, 497–503 (1974).
107. Sun, Z. *et al.* Superfast desulfurization for protein chemical synthesis and modification. *Chem* **8**, 2542–2557 (2022).
108. Ollivier, C. & Renaud, P. B-alkylcatecholboranes as a source of radicals for efficient conjugate additions to unsaturated ketones and aldehydes. *Chem. - A Eur. J.* **5**, 1468–1473 (1999).
109. Ollivier, C. & Renaud, P. A convenient and general tin-free procedure for radical conjugate addition. *Angew. Chemie - Int. Ed.* **39**, 925–928 (2000).
110. André-Joyaux, E., Kuzovlev, A., Tappin, N. D. C. & Renaud, P. A General Approach to Deboronative Radical Chain Reactions with Pinacol Alkylboronic Esters. *Angew. Chemie - Int. Ed.* **59**, 13859–13864 (2020).
111. Wu, J., He, L., Noble, A. & Aggarwal, V. K. Photoinduced Deaminative Borylation of Alkylamines. *J.*

- Am. Chem. Soc.* **140**, 10700–10704 (2018).
112. Fawcett, A. *et al.* Photoinduced decarboxylative borylation of carboxylic acids. *Science (80-)*. **357**, 283–286 (2017).
113. Wu, Z., Vlaming, R., Donohoe, M. & Pratt, D. A. Interrupted Homolytic Substitution Enables Organoboron Compounds to Inhibit Radical Chain Reactions Rather than Initiate Them. *J. Am. Chem. Soc.* **146**, 1153–1166 (2024).
114. Liu, M. *et al.* Photoactivation of Boronic Acid Prodrugs via a Phenyl Radical Mechanism: Iridium(III) Anticancer Complex as an Example. *J. Am. Chem. Soc.* **145**, 10082–10091 (2023).
115. Lo, K. K. W. *et al.* New Luminescent Cyclometalated Iridium(III) Diimine Complexes as Biological Labeling Reagents. *Inorg. Chem.* **42**, 6886–6897 (2003).
116. Ma, D. L., Chan, D. S. H. & Leung, C. H. Group 9 organometallic compounds for therapeutic and bioanalytical applications. *Acc. Chem. Res.* **47**, 3614–3631 (2014).
117. Liu, Z. & Sadler, P. J. Organoiridium complexes: Anticancer agents and catalysts. *Acc. Chem. Res.* **47**, 1174–1185 (2014).
118. Pitre, S. P., McTiernan, C. D., Ismaili, H. & Scaiano, J. C. Mechanistic insights and kinetic analysis for the oxidative hydroxylation of arylboronic acids by visible light photoredox catalysis: A metal-free alternative. *J. Am. Chem. Soc.* **135**, 13286–13289 (2013).
119. Cousins, M. S., Roberts, D. C. S. & Wit, H. de. GABAB receptor agonists for the treatment of drug addiction: A review of recent findings. *Drug Alcohol Depend.* **65**, 209–220 (2002).
120. Marks, P. A. & Breslow, R. Dimethyl sulfoxide to vorinostat: Development of this histone deacetylase inhibitor as an anticancer drug. *Nat. Biotechnol.* **25**, 84–90 (2007).
121. Boudjelel, M. *et al.* Electrochemical generation of aryl radicals from organoboron reagents enabled by pulsed electrosynthesis. *Angew. Chemie Int. Ed.* (2024) doi:10.1002/anie.202406203.
122. Suzuki, J., Tanigawa, M., Inagi, S. & Fuchigami, T. Electrochemical Oxidation of Organotrifluoroborate Compounds. *ChemElectroChem* **3**, 2078–2083 (2016).
123. Lennox, A. J. J., Nutting, J. E. & Stahl, S. S. Selective electrochemical generation of benzylic radicals enabled by ferrocene-based electron-transfer mediators. *Chem. Sci.* **9**, 356–361 (2018).
124. Chalker, J. M., Bernardes, G. J. L., Lin, Y. A. & Davis, B. G. Chemical modification of proteins at cysteine: Opportunities in chemistry and biology. *Chem. - An Asian J.* **4**, 630–640 (2009).
125. Bischoff, R. & Schlüter, H. Amino acids: Chemistry, functionality and selected non-enzymatic post-translational modifications. *J. Proteomics* **75**, 2275–2296 (2012).
126. Gunnoo, S. B. & Madder, A. Chemical Protein Modification through Cysteine. *ChemBioChem* **17**, 529–553 (2016).
127. Xue, L., Karpenko, I. A., Hiblot, J. & Johnsson, K. Imaging and manipulating proteins in live cells through covalent labeling. *Nat. Chem. Biol.* **11**, 917–923 (2015).
128. Cobo, I., Li, M., Sumerlin, B. S. & Perrier, S. Smart hybrid materials by conjugation of responsive polymers to biomacromolecules. *Nat. Mater.* **14**, 143–149 (2015).
129. Bunnett, J. F. & Zahler, R. E. Aromatic Nucleophilic Substitution Reactions. *J. Am. Chem. Soc.* **49**, 273–412 (1951).
130. Shannon, D. A. *et al.* Investigating the proteome reactivity and selectivity of aryl halides. *J. Am. Chem. Soc.* **136**, 3330–3333 (2014).
131. Zhang, D., Devarie-Baez, N. O., Li, Q., Lancaster, J. R. & Xian, M. Methylsulfonyl benzothiazole (MSBT): A selective protein thiol blocking reagent. *Org. Lett.* **14**, 3396–3399 (2012).
132. Spokoyny, A. M. *et al.* A Perfluoroaryl-Cysteine SNA_r Chemistry Approach to Unprotected Peptide Stapling. *J. Am. Chem. Soc.* **135**, 5946–5949 (2013).
133. Zou, Y. *et al.* Convergent diversity-oriented side-chain macrocyclization scan for unprotected polypeptides. *Org. Biomol. Chem.* **12**, 566–573 (2014).
134. Ngambenjawong, C., Pineda, J. M. B. & Pun, S. H. Engineering an Affinity-Enhanced Peptide

- through Optimization of Cyclization Chemistry. *Bioconjug. Chem.* **27**, 2854–2862 (2016).
135. Kalhor-Monfared, S. *et al.* Rapid biocompatible macrocyclization of peptides with decafluoro-diphenylsulfone. *Chem. Sci.* **7**, 3785–3790 (2016).
 136. Vinogradova, E. V., Zhang, C., Spokoyny, A. M., Pentelute, B. L. & Buchwald, S. L. Organometallic palladium reagents for cysteine bioconjugation. *Nature* **526**, 687–691 (2015).
 137. Rojas, A. J., Pentelute, B. L. & Buchwald, S. L. Water-Soluble Palladium Reagents for Cysteine S-Arylation under Ambient Aqueous Conditions. *Org. Lett.* **19**, 4263–4266 (2017).
 138. Al-Shuaeeb, R. A. A. *et al.* Palladium-Catalyzed Chemoselective and Biocompatible Functionalization of Cysteine-Containing Molecules at Room Temperature. *Chem. - A Eur. J.* **22**, 11365–11370 (2016).
 139. Cohen, D. T., Zhang, C., Pentelute, B. L. & Buchwald, S. L. An Umpolung Approach for the Chemoselective Arylation of Selenocysteine in Unprotected Peptides. *J. Am. Chem. Soc.* **137**, 9784–9787 (2015).
 140. Arnér, E. S. J. Common modifications of selenocysteine in selenoproteins. *Essays Biochem.* **64**, 45–53 (2019).
 141. Li, K. & Crews, C. M. PROTACs: past, present and future. *Chem. Soc. Rev.* **51**, 5214–5236 (2022).
 142. Liu, Z. *et al.* An overview of PROTACs: a promising drug discovery paradigm. *Mol. Biomed.* **3**, (2022).
 143. Kelm, J. M. *et al.* PROTAC'ing oncoproteins: targeted protein degradation for cancer therapy. *Mol. Cancer* **22**, 1–43 (2023).
 144. Dang, C. V., Reddy, E. P., Shokat, K. M. & Soucek, L. Drugging the 'undruggable' cancer targets. *Nat. Rev. Cancer* **17**, 502–508 (2017).
 145. Pathmanathan, S., Grozavu, I., Lyakisheva, A. & Stagljar, I. Drugging the undruggable proteins in cancer: A systems biology approach. *Curr. Opin. Chem. Biol.* **66**, 102079 (2022).
 146. Pickart, C. M. & Eddins, M. J. Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta - Mol. Cell Res.* **1695**, 55–72 (2004).
 147. Zhang, W. & Sidhu, S. S. Development of inhibitors in the ubiquitination cascade. *FEBS Lett.* **588**, 356–367 (2014).
 148. Park, J., Cho, J. & Song, E. J. Ubiquitin–proteasome system (UPS) as a target for anticancer treatment. *Arch. Pharm. Res.* **43**, 1144–1161 (2020).
 149. Ciechanover, A. & Schwartz, A. L. The ubiquitin-proteasome pathway: The complexity and myriad functions of proteins death. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 2727–2730 (1998).
 150. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).
 151. Komander, D. & Rape, M. The ubiquitin code. *Annu. Rev. Biochem.* **81**, 203–229 (2012).
 152. Bemis, T. A., Clair, J. J. La & Burkart, M. D. Unraveling the Role of Linker Design in Proteolysis Targeting Chimeras. *J. Med. Chem.* **64**, 8042–8052 (2021).
 153. Zeng, M. *et al.* Exploring Targeted Degradation Strategy for Oncogenic KRASG12C. *Cell Chem. Biol.* **27**, 19-31.e6 (2020).
 154. Goldman, S. D. B., Funk, R. S., Rajewski, R. A. & Krise, J. P. Mechanisms of Amine Accumulation in, and Egress from, Lysosomes. *Bioanalysis* **1**, 1445–1459 (2009).
 155. Hamilton, E. P. *et al.* ARV-471, an estrogen receptor (ER) PROTAC degrader, combined with palbociclib in advanced ER+/human epidermal growth factor receptor 2–negative (HER2–) breast cancer: Phase 1b cohort (part C) of a phase 1/2 study. *J. Clin. Oncol.* **40**, TPS1120–TPS1120 (2022).
 156. Gao, X. *et al.* Phase 1/2 study of ARV-110, an androgen receptor (AR) PROTAC degrader, in metastatic castration-resistant prostate cancer (mCRPC). *J. Clin. Oncol.* **40**, 17 (2022).
 157. Neklesa, T. *et al.* ARV-110: An oral androgen receptor PROTAC degrader for prostate cancer. *J. Clin. Oncol.* **37**, 259 (2019).

158. Searle, E. *et al.* Initial Findings from a First-in-Human Phase 1a/b Trial of NX-5948, a Selective Bruton's Tyrosine Kinase (BTK) Degradator, in Patients with Relapsed/Refractory B Cell Malignancies. *Blood* **142**, 4473 (2023).
159. Raina, K. *et al.* PROTAC-induced BET protein degradation as a therapy for castration-resistant prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7124–7129 (2016).
160. Deng, Y. *et al.* ARV-771 Acts as an Inducer of Cell Cycle Arrest and Apoptosis to Suppress Hepatocellular Carcinoma Progression. *Front. Pharmacol.* **13**, 1–13 (2022).
161. Cheng, M. *et al.* Discovery of Potent and Selective Epidermal Growth Factor Receptor (EGFR) Bifunctional Small-Molecule Degradators. *J. Med. Chem.* **63**, 1216–1232 (2020).
162. Gao, H., Sun, X. & Rao, Y. PROTAC Technology: Opportunities and Challenges. *ACS Med. Chem. Lett.* **11**, 237–240 (2020).
163. Wang, L. *et al.* Discovery of a first-in-class CDK2 selective degrader for AML differentiation therapy. *Nat. Chem. Biol.* **17**, 567–575 (2021).
164. Crews, C. M. Inducing Protein Degradation as a Therapeutic Strategy. *J. Med. Chem.* **61**, 403–404 (2018).
165. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
166. Lebraud, H., Wright, D. J., Johnson, C. N. & Heightman, T. D. Protein degradation by in-cell self-assembly of proteolysis targeting chimeras. *ACS Cent. Sci.* **2**, 927–934 (2016).
167. Liu, J. *et al.* TF-PROTACs Enable Targeted Degradation of Transcription Factors. *J. Am. Chem. Soc.* **143**, 8902–8910 (2021).
168. Hughes, C. S. *et al.* Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat. Protoc.* **14**, 68–85 (2019).
169. Mikulášek, K. *et al.* SP3 Protocol for Proteomic Plant Sample Preparation Prior LC-MS/MS. *Front. Plant Sci.* **12**, (2021).
170. Becker, T. *et al.* Transforming Chemical Proteomics Enrichment into a High-Throughput Method Using an SP2E Workflow. *JACS Au* **2**, 1712–1723 (2022).
171. Wong, J., Chilkoti, A. & Moy, V. T. Direct force measurements of the streptavidin-biotin interaction. *Biomol. Eng.* **16**, 45–55 (1999).
172. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics: Abstract: *Nature* **422**, 198–207 (2003).
173. Shuken, S. R. An Introduction to Mass Spectrometry-Based Proteomics. *J. Proteome Res.* **22**, 2151–2171 (2023).
174. Han, X., Aslanian, A. & Yates, J. R. Mass spectrometry for proteomics. *Curr. Opin. Chem. Biol.* **12**, 483–490 (2008).
175. Senko, M. W. *et al.* Novel parallelized quadrupole/linear ion trap/orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. *Anal. Chem.* **85**, 11710–11714 (2013).
176. Yan, X., Zhang, Y., Zhou, Y., Li, G. & Feng, X. Technical Overview of Orbitrap High Resolution Mass Spectrometry and Its Application to the Detection of Small Molecules in Food (Update Since 2012). *Crit. Rev. Anal. Chem.* **52**, 593–626 (2022).
177. Xian, F., Hendrickson, C. L., Blakney, G. T., Beu, S. C. & Marshall, A. G. Automated broadband phase correction of fourier transform ion cyclotron resonance mass spectra. *Anal. Chem.* **82**, 8807–8812 (2010).
178. Lange, O., Damoc, E., Wiegand, A. & Makarov, A. Enhanced Fourier transform for Orbitrap mass spectrometry. *Int. J. Mass Spectrom.* **369**, 16–22 (2014).
179. Guo, T. & Aebersold, R. Recent advances of data-independent acquisition mass spectrometry-based proteomics. *Proteomics* **23**, 20–22 (2023).

180. Li, J., Smith, L. S. & Zhu, H. J. Data-independent acquisition (DIA): An emerging proteomics technology for analysis of drug-metabolizing enzymes and transporters. *Drug Discov. Today Technol.* **39**, 49–56 (2021).
181. Bruderer, R. *et al.* Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell. Proteomics* **14**, 1400–1410 (2015).
182. Tabb, D. L. *et al.* Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **9**, 761–776 (2010).
183. Schubert, O. T. *et al.* Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nat. Protoc.* **10**, 426–441 (2015).
184. Gillet, L. C. *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **11**, 1–17 (2012).
185. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).
186. Tsou, C. C. *et al.* DIA-Umpire: Comprehensive computational framework for data-independent acquisition proteomics. *Nat. Methods* **12**, 258–264 (2015).
187. Kielkowski, P. *et al.* A Pronucleotide Probe for Live-Cell Imaging of Protein AMPylation. *ChemBioChem* **21**, 1285–1287 (2020).
188. Meldal, M. & Tomøe, C. W. Cu-catalyzed azide - Alkyne cycloaddition. *Chem. Rev.* **108**, 2952–3015 (2008).
189. Hiller, N. de J. *et al.* Arylboronic Acids and their Myriad of Applications Beyond Organic Synthesis. *European J. Org. Chem.* **2020**, 4841–4877 (2020).
190. Mollner, T. A. *et al.* Post-translational insertion of boron in proteins to probe and modulate function. *Nat. Chem. Biol.* **17**, 1245–1261 (2021).
191. Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* **94**, 1826–1835 (2008).
192. Lin, X., Haimov, E., Redko, B. & Vigalok, A. Selective Stepwise Arylation of Unprotected Peptides by PtlV Complexes. *Angew. Chemie - Int. Ed.* **61**, (2022).
193. Bottecchia, C. *et al.* Visible-Light-Mediated Selective Arylation of Cysteine in Batch and Flow. *Angew. Chemie - Int. Ed.* **56**, 12702–12707 (2017).
194. Li, W. *et al.* Scalable and selective deuteration of (hetero)arenes. *Nat. Chem.* **14**, 334–341 (2022).
195. Makarov, D. & Kielkowski, P. Chemical Proteomics Reveals Protein Tyrosination Extends Beyond the Alpha-Tubulins in Human Cells. *ChemBioChem* **23**, 1–6 (2022).
196. Peixoto, A., Relvas-Santos, M., Azevedo, R., Lara Santos, L. & Ferreira, J. A. Protein glycosylation and tumor microenvironment alterations driving cancer hallmarks. *Front. Oncol.* **9**, 1–24 (2019).
197. Wang, H., Ramakrishnan, A., Fletcher, S., Prochownik, E. V & Genetics, M. Protein glycosylation in cancer. *Annu Rev Pathol* **2**, 473–510 (2015).
198. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer* **15**, 540–555 (2015).
199. Altea-Manzano, P., Cuadros, A. M., Broadfield, L. A. & Fendt, S. Nutrient metabolism and cancer in the in vivo context: a metabolic game of give and take. *EMBO Rep.* **21**, 1–22 (2020).
200. E., J. M. & D., C. B. Positron-Emission Tomography and Assessment of Cancer Therapy. *N. Engl. J. Med.* **354**, 496–507 (2006).
201. Hashimoto, A., Suenaga, K., Gloter, A., Urita, K. & Iijima, S. Direct evidence for atomic defects in graphene layers. *Nature* **430**, 870–873 (2004).
202. Vocadlo, D. J., Hang, H. C., Kim, E. J., Hanover, J. A. & Bertozzi, C. R. A chemical approach for

- identifying O-GlcNAc-modified proteins in cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9116–9121 (2003).
203. Hang, H. C., Yu, C., Pratt, M. R. & Bertozzi, C. R. Probing Glycosyltransferase Activities with the Staudinger Ligation. *J. Am. Chem. Soc.* **126**, 6–7 (2004).
204. Teo, C. F. & Wells, L. Monitoring protein O-linked β -N-acetylglucosamine status via metabolic labeling and copper-free click chemistry. *Anal. Biochem.* **464**, 70–72 (2014).
205. Ito, T. *et al.* Identification of a Primary Target of Thalidomide Teratogenicity. *Science (80-.)*. **327**, 1345–1350 (2010).
206. Brownsey, D. K., Rowley, B. C., Gorobets, E., Gelfand, B. S. & Derksen, D. J. Rapid synthesis of pomalidomide-conjugates for the development of protein degrader libraries. *Chem. Sci.* **12**, 4519–4525 (2021).
207. Wojdyla, K. & Rogowska-Wrzesinska, A. Differential alkylation-based redox proteomics - Lessons learnt. *Redox Biol.* **6**, 240–252 (2015).
208. Carmona, J. A. *et al.* Asymmetric organocatalytic synthesis of tertiary azomethyl alcohols: key intermediates towards azoxy compounds and α -hydroxy- β -amino esters. *Org. Biomol. Chem.* **15**, 2993–3005 (2017).
209. Agarwal, K. *et al.* Inhibition of Mucin-type O-glycosylation through metabolic processing and incorporation of N-thioglycolyl-d-galactosamine peracetate (Ac 5GalNTGc). *J. Am. Chem. Soc.* **135**, 14189–14197 (2013).
210. Metier, C. *et al.* Profiling of Haemophilus influenzae strain R2866 with carbohydrate-based covalent probes. *Org. Biomol. Chem.* **19**, 476–485 (2021).
211. Tomaszewska, J., Kowalska, K. & Koroniak-Szejn, K. Glucosamine- and galactosamine- based monosaccharides with highly fluorinated motifs. *J. Fluor. Chem.* **191**, 1–13 (2016).
212. Li, J. *et al.* An OGA-resistant probe allows specific visualization and accurate identification of O-GlcNAc-modified proteins in cells. *ACS Chem. Biol.* **11**, 3002–3006 (2016).
213. Wulff-Fuentes, E. *et al.* The human O-GlcNAcome database and meta-analysis. *Sci. Data* **8**, 1–11 (2021).
214. Fernandes, G. F. S., Denny, W. A. & Dos Santos, J. L. Boron in drug design: Recent advances in the development of new therapeutic agents. *Eur. J. Med. Chem.* **179**, 791–804 (2019).
215. Zhao, J., Zhai, B., Gygi, S. P. & Goldberg, A. L. mTOR inhibition activates overall protein degradation by the ubiquitin proteasome system as well as by autophagy. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 15790–15797 (2015).
216. Müller, T. *et al.* Automated sample preparation with SP3 for low-input clinical proteomics. *Mol. Syst. Biol.* **16**, 1–19 (2020).
217. Messner, C. B. *et al.* Mass spectrometry-based high-throughput proteomics and its role in biomedical studies and systems biology. *Proteomics* **23**, 1–15 (2023).
218. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).
219. Sahu, I. *et al.* The 20S as a stand-alone proteasome in cells can degrade the ubiquitin tag. *Nat. Commun.* **12**, (2021).
220. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

5. List of Abbreviations

Ac₃4dGlcNAz	1,3,6-Tri- <i>O</i> -acetyl-4-deoxy- <i>N</i> -azidoacetyl-glucosamine
Ac₄GalNAz	<i>N</i> -azidoacetylgalactosamine-tetraacylated
Ac₄GlcNAz	<i>N</i> -azidoacetylglucosamine-tetraacylated
ADCs	Antibody-drug conjugates
ADD	Add-and-Done Desulfurization
AGC	Automatic Gain Control
AIBN	Azobisisobutyronitrile
AMP	Adenosine monophosphate
AMPylation	Adenylylation
APS	Ammonium persulfate
ATP	Adenosine triphosphate
BARAC	Biarylazacyclooctynone
BCN	Bicyclo[2.1.0]nonyne
Biotin-Bpin	4-Biotinamidophenyl pinacol boronate
BRD4	Bromodomain-containing protein 4
BSA	Bovine serum albumin
CHO	Chinese hamster ovary
C-trap	Curved linear trap
CuAAC	Copper(I)-catalyzed azide-alkyne cycloaddition
CV	Compensation voltage
DBCO	Dibenzocyclooctyne
DC	Direct current
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
DIFO	Difluorooctyne
DMA	<i>N,N</i> -dimethylacetamide
DMPO	5,5-Dimethyl-1-pyrroline <i>N</i> -oxide
DTT	Dithiothreitol
EDA	Electron donor-acceptor
EDGs	Electron-donating groups
EDTA	Ethylenediaminetetraacetic acid disodium salt dihydrate
EPR	Electron paramagnetic resonance
ER	Endoplasmic reticulum
ERK1/2	Extracellular-regulated protein kinase 1/2
EWGs	Electron-withdrawing groups
FAIMS	High field asymmetric waveform ion mobility spectrometry
GalNAc	<i>N</i> -acetylgalactosamine
GAO	Galactose oxidase
GlcNAc	<i>N</i> -acetylglucosamine
GODCAT	Glucose oxidase/catalase

GSH	Glutathione
GSSG	Glutathione disulfide
HAT	Hydrogen atom transfer
HCD	Higher energy collisional dissociation
HPLC	High-performance liquid chromatography
HRP	Horseradish peroxidase
IEDDA	Inverse-electron demand Diels-Alder reaction
LDA	Lithium diisopropylamide
MOA	Mode of action
MS	Mass spectrometry
MSBT	Methylsulfonyl benzothiazole
MTT	3-(4,5-Dimethyl-2-thiazolyl)-2,5-diphenyl-2 <i>H</i> -tetrazolium bromide
NBS	<i>N</i> -bromosuccinimide
nESI	Nanoelectrospray ionization
NHCs	<i>N</i> -heterocyclic carbenes
NMP	<i>N</i> -methyl-2-pyrrolidone
NUP	Nuclear pore glycoprotein
OA	Oxidative addition
OCT	Cyclooctyne
OGA	O-GlcNAcase
O-GalNAcylation	O-linked β - <i>N</i> -acetyl-galactosamine glycosylation
O-GlcNAcylation	O-linked β - <i>N</i> -acetyl-glucosamine glycosylation
OGT	O-GlcNAc transferase
PBS	Phosphate-buffered saline
PDAC	Pancreatic ductal adenocarcinoma
PEG	Polyethyleneglycol
PET	Positron emission tomography
POIs	Proteins of interest
PROTAC	Proteolysis targeting chimera
PTMs	Post-translational modifications
R-Bcats	<i>B</i> -alkylcatecholboranes
R-Bpins	<i>B</i> -alkylpinacolboranes
RE	Reductive elimination
RTK	Receptor tyrosine kinase
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SET	Single electron transfer
S_NAr	Nucleophilic aromatic substitution
SP3	Single-pot, solid-phase-enhanced sample preparation
SPAAC	Strain-promoted azide-alkyne cycloaddition
TAMRA-Bpin	Tetramethyl rhodamine aryl pinacol boronate
TBAF	tetra- <i>n</i> -Butylammonium fluoride

TBS	<i>tert</i> -Butyldimethylsilyl
TCEP	Tris(2-carboxyethyl)phosphine
TCO	Transcyclooctene
TCPs	Tubulin carboxypeptidases
TLC	Thin layer chromatography
TM	Transmetalation
TPD	Target protein degradation
TrxA	Thioredoxin
TTL	Tubulin tyrosine ligase
UDP	Uridine diphosphate
UDP-GalNAc	Uridine diphosphate <i>N</i> -acetylgalactosamine
UDP-GlcNAc	Uridine diphosphate <i>N</i> -acetylglucosamine
UPS	Ubiquitin-proteasome system
VHL	Von Hippel-Lindau

6. List of Figures and Tables

- Figure 1.** The proteome complexity by protein PTMs.
- Figure 2.** “PTMs clock” classified based on their chemical natures and reversibilities.
- Figure 3.** General scheme of protein O-GalNAcylation as well as its inhibitors and mimics.
- Figure 4.** The GalNAc and GlcNAc salvage pathways; chemoenzymatic methods for MS analysis.
- Figure 5.** General schemes for click chemistry and variations.
- Figure 6.** Suzuki cross-coupling reaction and applications in biomolecule functionalization and labeling.
- Figure 7.** Contemporary methods for aryl radical generation and application.
- Figure 8.** Radical reactions of alkyl and aryl substituted boron.
- Figure 9.** Cysteine S-arylation chemistry.
- Figure 10.** Protein ubiquitinylation and principles of PROTAC.
- Figure 11.** Structures of PROTACs.
- Figure 12.** *In situ* self-assembly of PROTACs through tetrazine ligation targeting on BRD4 or ERK1/2 protein.
- Figure 13.** TF-PROTACs synthesized by SPAAC targeting on transcription factors.
- Figure 14.** Mass spectrometer and advancement with Orbitrap.
- Figure 15.** Schematic illustration of basic working principles of DDA and DIA methods.
- Figure 16.** Suzuki cross-coupling reactions for protein labelling in the living cells, and general schemes for protein (de)AMPylation and (de)tyrosination.
- Figure 17.** Structures of potential protein AMPylation or tyrosination probes.
- Figure 18.** Synthesis of AMPylation probe **1** and **2**.
- Figure 20.** Synthesis of AMPylation probe **4**.
- Figure 21.** Synthesis of intermediate **12**.
- Figure 22.** Synthesis of tyrosination probes **5** and **6**.
- Figure 23.** Workflow for identification of protein AMPylation using Suzuki cross-coupling reaction.
- Figure 24.** Reactive species used in the workflow.
- Figure 25.** Volcano plot for **1**- or **2**-treated samples (versus DMSO vehicle).
- Figure 26.** The metal-free reaction of aryl pinacol boronate with proteins in whole cell lysate.
- Figure 27.** single protein labelling with TAMRA-Bpin.
- Figure 28.** Characterization of aryl pinacol boronate reaction with oxidized glutathione.
- Figure 29.** DMPO spin trap experiment.
- Figure 30.** Reaction substrate scope.
- Figure 31.** Competition between Tyr-O-Alk and iodinated tyrosination probes.
- Figure 32.** Overview of the GlyTAC strategy.
- Figure 33.** Characterization and optimization of GlyTAC system.
- Figure 34.** GlyTACs increase protein ubiquitinylation and Nup62 degradation.
- Figure 35.** GlyTACs mechanism of action.
- Figure 36.** GlyTACs induced whole proteome changes and cancer cell screening.

Table 1. Examples of PROTACs targeting on oncoproteins.

7. Experimental Sections

7.1 Organic Synthesis

7.1.1 Methods for Organic Synthesis

Reagents and solvents were purchased from commercial suppliers, including abcr GmbH (Karlsruhe, Germany), Acros Organics (Thermo Fisher Scientific, USA), AppliChem GmbH (Darmstadt, Germany), Carbolution Chemicals GmbH (St. Ingbert, Germany), Grüssing GmbH (Filsum, Germany), Key Organics (Cornwall, UK), Sigma-Aldrich (St. Louis, USA), TCI Deutschland GmbH (Eschborn, Germany), Thermo Scientific (Loughborough, UK), VWR (Leuven, Belgium) and they were used without any further purification. Of note, the bovine serum albumin (BSA) was obtained from Thermo Scientific (Rockford, USA; REF # 23209); biotin polyclonal antibody was obtained from Invitrogen (Thermo Fisher Scientific, USA; Catalog # 31852); TAMRA-NHS ester (**31**) was obtained from abcr GmbH (Karlsruhe, Germany; Article ID # AB464086); TAMRA-N₃ was obtained from baseclick GmbH (Neuried, Germany; Catalog # BCFA-008-10); *N*-azidoacetylglucosamine-tetraacylated (Ac₄GlcNAz) was obtained from Jena Bioscience GmbH (Jena, Germany; Catalog # CLK-1085-5); DBCO-PEG₄-5-tetramethylrhodamine (DBCO-PEG₄-5-TAMRA) was obtained from Jena Bioscience GmbH (Jena, Germany; Catalog # CLK-A131N-1); 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2*H*-tetrazolium bromide (MTT) was obtained from Merck KGaA (Darmstadt, Germany; Product No. # M2128); DBCO-NHS ester (**36**) was obtained from abcr GmbH (Karlsruhe, Germany; Article ID # AB549646); (+/-)-thalidomide was obtained from abcr GmbH (Karlsruhe, Germany; Article ID # AB347541); bortezomib was obtained from TCI Deutschland GmbH (Eschborn, Germany; Article # B5741). 1,3,6-tri-*O*-acetyl-4-deoxy-*N*-azidoacetyl-glucosamine (Ac₃4dGlcNAz) was synthesized in Prof. Dr. Anja Hoffmann-Röder's lab in LMU Munich. The two short hexapeptides (NH₂-KEHSPC-COOH & NH₂-RWTPCD-COOH) were custom-made from Peptide Specialty Laboratories GmbH (Heidelberg, Germany).

TLC (thin layer chromatography) was performed to monitor the reaction progress, which was done on pre-coated silica gel plates (60 F-254, 0.25 mm, from Merck KGaA, Darmstadt, Germany) as the stationary phase, with detection by UV lights ($\lambda = 254$ and/or 366 nm).

Flash chromatography was performed by Pure Chromatography Instruments (Pure-C815 Flash) from BÜCHI Labortechnik GmbH (Essen, Germany) on pre-fabricated column (FlashPure silica gel, 40 μ m irregular, 12 g) with the indicated eluent.

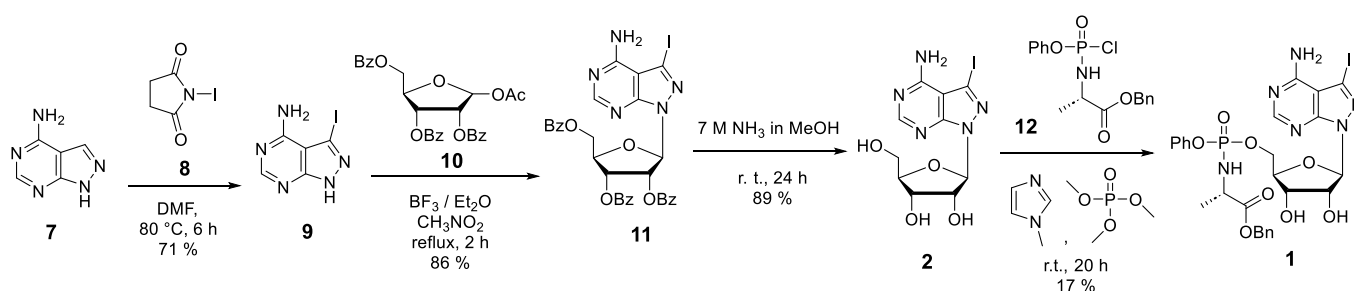
¹H and proton-decoupled ¹³C NMR spectra for compound characterization in deuterated chloroform (CDCl₃) were acquired at 298K on a Bruker Avance Neo 500 spectrometer (11.7 T), using a nitrogen-cooled Prodigy BBO probe. All chemical shifts are reported in delta (δ) units in parts per million (ppm) relative to distinguished solvent signals as an internal reference. Coupling constant *J* are indicated in Hertz (Hz). Splitting patterns for peak assignments are indicated as the following abbreviations: br, broad; s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublets; dt, doublet of triplets. Spectra were analyzed using MestReNova (version 14.3.3, Mestrelab Research, Santiago de Compostela, Spain).

For characterizing 4-biotinamidophenyl pinacol boronate (Biotin-Bpin, **21**), the 2D ¹H-¹H COSY

and ^{13}C -edited ^1H - ^{13}C HSQC NMR spectra were acquired at 298K on a Bruker Avance III 800 spectrometer (18.8 T), using a cryogen-cooled TCI probe.

LC-MS chromatograms were obtained from Thermo Fisher LC-MS system composed of a DIONEX UltiMate 3000 HPLC system (pump, auto sampler, column compartment, and diode array detector) and an ESI(electrospray ionization)-MS based MSQ Plus single quadrupole mass spectrometer; the latter one was used for ESI-MS measurements with direct injection. Reversed phase column chromatographic methods use a Hypersil Gold C18 selectivity column (100 × 2.1 mm) with indicated injection volume, eluent gradients and flow rates.

7.1.2 Synthesis of AMPylation Probes 1 and 2



Iodinated adenine analogue (9). 4-amino-1*H*-pyrazolo[3,4-*d*]pyrimidine (**7**, 2.9 g, 21.7 mmol, 1.0 equiv.) was suspended in dry DMF (75 mL), and *N*-iodosuccinimide (**8**, 5.9 g, 26 mmol, 1.2 equiv.) was added. The mixture was heated with stirring at 80 °C for 6 hours, evaporated to dryness in vacuo, and triturated with ethanol (50 mL) to yield 5.0 g (88 %) of crude product. This material was purified by suspending in 50 % aqueous ethanol (100 mL), heating near boiling, and slowly adding concentrated ammonium hydroxide (50 mL) to affect solution. The resulting solution was decolorized with activated charcoal, filtered, and boiled to remove most of the ammonia. The resulting white solid was collected and dried to yield 4.1 g (71 %) product (**9**) and used for the next step without further purification.

ESI-MS (*m/z*) calculated mass for product (**9**) $\text{C}_5\text{H}_5\text{N}_5\text{I}$ [$\text{M}+\text{H}$] $^+$: 261.95; observed mass: 262.07.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 13.80 (s, 1H, NH), 8.17 (s, 1H, H_6).

Protected iodinated adenosine analogue (11). 1-*O*-Acetyl-2,3,5-tri-*O*-benzoyl- β -D-ribofuranose (**10**, 1740 mg, 3.45 mmol, 1.5 equiv.) was dissolved in dry nitromethane (CH_3NO_2), and 4-amino-3-iodo-1*H*-pyrazolo[3,4-*d*] pyrimidine (iodinated adenine analogue, **9**, 600 mg, 2.3 mmol, 1 equiv.) was added. The resulting suspension was brought to reflux temperature, and boron trifluoride etherate (426 μL , 3.45 mmol, 1.5 equiv.) was added through the condenser. The suspension became clear immediately and began to darken slowly. After 2 hours at reflux, the mixture was cooled and evaporated in vacuo to a thick, dark oil. The oil was dissolved in EtOAc and poured with stirring into sat. aqueous NaHCO_3 ; the layers were then separated, and the water layer was extracted with EtOAc. The organic layers were combined and washed once with water, and the organic layer was dried over MgSO_4 and evaporated to provide a dry foam (**11**, 1394 mg, 86 %) which was purified by column chromatography on silica gel using 5 % acetone in dichloromethane.

ESI-MS (*m/z*) calculated mass for product (**11**) $\text{C}_{31}\text{H}_{25}\text{N}_5\text{O}_7\text{I}$ [$\text{M}+\text{H}$] $^+$: 706.07; observed mass: 706.27.

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 8.25 (s, 1H), 8.06 – 8.03 (m, 2H), 7.92 (ddd, $J = 8.4, 4.7, 1.4$ Hz, 4H),

7.72 – 7.43 (m, 10H), 6.67 (d, $J = 3.7$ Hz, 1H), 6.31 (dd, $J = 5.4, 3.6$ Hz, 1H), 6.12 (t, $J = 5.8$ Hz, 1H), 4.88 (dt, $J = 6.6, 3.5$ Hz, 1H), 4.61 (s, 2H).

Iodinated adenosine analogue (2). (2*S*,3*R*,4*R*,5*R*)-2-(acetoxymethyl)-5-(4-amino-3-iodo-1*H*-pyrazolo[3,4-*d*]pyrimidin-1-yl)tetrahydrofuran-3,4-diyl diacetate (protected iodinated adenosine analogue, **11**, 475 mg, 0.674 mmol) was dissolved in a solution of ammonia (7 M in methanol, 10 mL). After stirring at room temperature for 48 hours, the mixture was concentrated under reduced pressure. Ether was added to the resulting solid which was subsequently isolated by suction filtration to afford an off-white solid (**2**, 235.5 mg, 89 %).

ESI-MS (m/z) calculated mass for product (**2**) $C_{10}H_{13}N_5O_4I$ $[M+H]^+$: 394.03; observed mass: 394.16.

1H NMR (400 MHz, $DMSO-d_6$) δ 8.23 (s, 1H), 6.04 (d, $J = 5.0$ Hz, 1H), 5.40 (d, $J = 6.0$ Hz, 1H), 5.17 (d, $J = 5.4$ Hz, 1H), 4.85 (dd, $J = 6.2, 5.5$ Hz, 1H), 4.57 (q, $J = 5.3$ Hz, 1H), 4.17 (q, $J = 5.1$ Hz, 1H), 3.89 (q, $J = 4.8$ Hz, 1H), 3.47 – 3.40 (m, 2H).

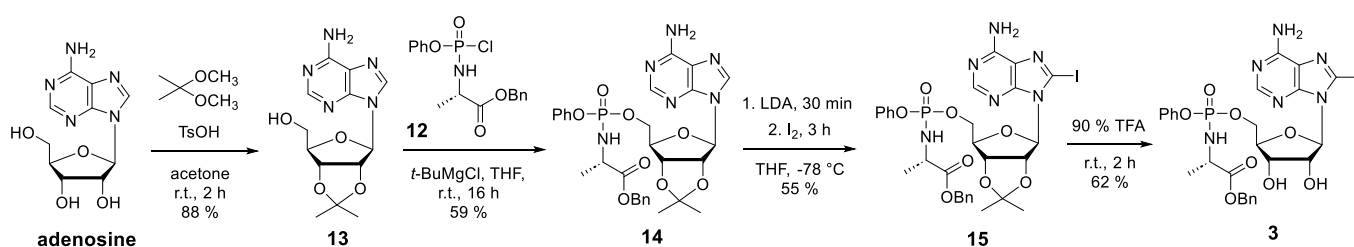
Iodinated phosphoramidate (1). (2*R*,3*R*,4*S*,5*R*)-2-(4-amino-3-iodo-1*H*-pyrazolo[3,4-*d*]pyrimidin-1-yl)-5-(hydroxymethyl)tetrahydrofuran-3,4-diol (iodinated adenosine analogue, **2**, 100 mg, 0.25 mmol, 1.0 equiv.) was dissolved in 8 mL trimethyl phosphate and then slowly added *N*-methylimidazole (121.6 μ L, 1.50 mmol, 6 equiv.) at room temperature. The pre-synthesized phenyl(benzyloxy-*L*-alaninyl) phosphorochloridate (**12**, starting from 6 equiv. phenyl dichlorophosphate and 6 equiv. *L*-alanine benzyl ester *p*-toluenesulfonate salt) in anhydrous THF was added dropwise into the mixture and stirred gently for 20 hours. Afterwards the reaction mixture was diluted with EtOAc and washed with sat. aqueous $NaHCO_3$ and $NaCl$ solution, dried over anhydrous Na_2SO_4 , filtered and concentrated under reduced pressure. The residue was subjected to silica gel chromatography eluting with 0 - 5% methanol in dichloromethane to give product **1** as a white powder (30 mg, 17 %).

ESI-MS (m/z) calculated mass for product (**1**) $C_{26}H_{29}N_6O_8PI$ $[M+H]^+$: 711.10; observed mass: 711.28.

1H NMR (400 MHz, $CDCl_3$) δ 8.34 – 8.27 (m, 1H), 7.40 – 7.27 (m, 7H), 7.15 (ddd, $J = 27.2, 17.8, 9.8$ Hz, 3H), 6.44 – 6.33 (m, 1H), 6.09 (d, $J = 29.2$ Hz, 2H), 5.19 – 5.09 (m, 2H), 4.69 (d, $J = 13.2$ Hz, 1H), 4.43 – 4.22 (m, 3H), 4.15 – 4.00 (m, 2H), 3.89 – 3.62 (m, 3H), 1.27 – 1.24 (m, 3H).

^{31}P NMR (162 MHz, $CDCl_3$) δ 3.19, 2.93.

7.1.3 Synthesis of AMPylation Probe 3



Acetal-protected adenosine (13). Adenosine (1.0 g, 3.74 mmol, 1.0 equiv.) was dissolved in 15 mL acetone, followed by adding 4.58 mL 2,2-dimethoxypropane (37.4 mmol, 10.0 equiv.) dropwise and *p*-toluenesulfonic acid (TsOH, 3.22 g, 18.7 mmol, 5.0 equiv.). The resulting mixture was stirred vigorously at room temperature for 2 hours, then stopped with sat. aqueous $NaHCO_3$ until neutralized. The reaction mixture was extracted with EtOAc (3×25 mL) and the combined organic phases were dried over

anhydrous MgSO₄. The solvents were removed under reduced pressure and the residue was purified by column chromatography on silica gel using gradient 1 - 5% methanol in dichloromethane to give the product as a pale yellow solid (**13**, 1010 mg, 88 %).

ESI-MS (m/z) calculated mass for product (**13**) C₁₃H₁₈N₅O₄ [M+H]⁺: 308.15; observed mass: 308.25.

¹H NMR (500 MHz, DMSO) δ 8.35 (s, 1H), 8.16 (s, 1H), 7.34 (s, 2H), 6.13 (d, *J* = 3.1 Hz, 1H), 5.35 (dd, *J* = 6.1, 3.1 Hz, 1H), 4.97 (dd, *J* = 6.2, 2.5 Hz, 1H), 4.22 (td, *J* = 4.8, 2.5 Hz, 1H), 3.55 (qd, *J* = 11.7, 4.9 Hz, 2H), 1.56 (s, 3H), 1.34 (s, 3H).

Acetal-protected phosphoramidate (14). 2',3'-*O*-isopropylideneadenosine (acetal-protected adenosine, **13**, 390 mg, 1.27 mmol) was dissolved in 8 mL anhydrous THF. Solution of *t*BuMgCl in THF (1 M, 3.16 mL, 3.175 mmol, 2.5 equiv.) was added dropwise and the reaction mixture was stirred for 15 minutes. The pre-synthesized phenyl(benzyloxy-*L*-alaninyl) phosphorochloridate (**12**, starting from 3 equiv. phenyl dichlorophosphate and 3 equiv. *L*-alanine benzyl ester *p*-toluenesulfonate salt) in anhydrous THF was added dropwise into the mixture and stirred gently at room temperature for 16 hours. Reaction was quenched by saturated NH₄Cl solution (6 mL) and mixture was extracted once with 30 mL EtOAc and then 30 mL water. Organic layer was dried over anhydrous Na₂SO₄, filtered and evaporated under reduced pressure. Clean product was obtained using column chromatography on silica gel (1 - 5% methanol in dichloromethane) as acetal-protected phosphoramidate (**14**, 561 mg, 59 %) in colourless oily compound form.

ESI-MS (m/z) calculated mass for product (**14**) C₂₉H₃₄N₆O₈P [M+H]⁺: 625.21; observed mass: 625.30.

¹H NMR (400 MHz, CDCl₃) δ 8.28 (d, *J* = 7.4 Hz, 1H), 7.98 (d, *J* = 6.2 Hz, 1H), 7.26 – 7.03 (m, 11H), 6.66 (s, 2H), 6.08 (dd, *J* = 4.5, 2.5 Hz, 1H), 4.96 – 4.76 (m, 2H), 4.46 (ddd, *J* = 13.8, 4.7, 2.6 Hz, 1H), 4.08 – 3.95 (m, 1H), 1.59 (d, *J* = 2.8 Hz, 3H), 1.34 (d, *J* = 10.8 Hz, 3H), 1.28 (t, *J* = 6.3 Hz, 3H).

Protected iodinated phosphoramidate (15). In a 100 mL round-bottom flask, diisopropylamine (*i*-Pr₂NH, 303 μL, 2.16 mmol, 4.5 equiv) was dissolved in THF (3 mL) and cooled to -78 °C. *n*-Butyllithium (*n*-BuLi, 2.5 M in hexanes, 865.1 μL, 2.16 mmol, 4.5 equiv) was added and the mixture was stirred for 30 minutes at -78 °C. Benzyl (((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methoxy)(phenoxy)phosphoryl)-*L*-alaninate (acetal-protected phosphoramidate, **14**, 300 mg, 0.48 mmol, 1.0 equiv) was dissolved in THF (10 mL) and added to the LDA solution via cannula. After stirring at -78 °C for 30 minutes, a solution of I₂ (305 mg, 1.2 mmol, 2.5 equiv) in THF (6 mL) was added dropwise via cannula over 10 min. The resulting deep orange solution was stirred at -78 °C for an additional 3 hours. The reaction was quenched by addition of sat. aqueous NH₄Cl (8 mL) and sat. aqueous Na₂S₂O₃ (15 mL) and extracted with EtOAc (3 × 30 mL). The combined organic extracts were washed with sat. aqueous NH₄Cl (2 × 50 mL) and sat. aqueous NaCl (2 × 50 mL), dried over anhydrous MgSO₄, filtered and evaporated *in vacuo*. Purification was performed by silica chromatography (2:8 hexanes / EtOAc) to yield protected iodinated phosphoramidate (**15**, 198.5 mg, 55 %) as an off-white solid.

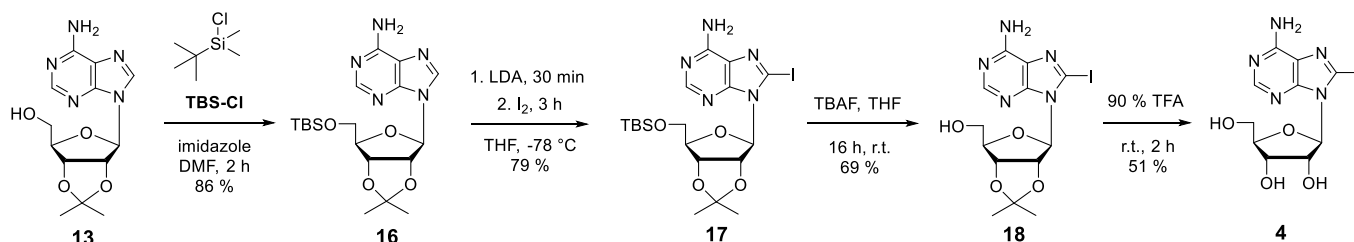
ESI-MS (m/z) calculated mass for product (**15**) C₂₉H₃₃N₆O₈PI [M+H]⁺: 751.12; observed mass: 751.25.

Iodinated phosphoroamidate (3). Benzyl (((((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-8-iodo-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methoxy)(phenoxy)phosphoryl)-*L*-alaninate (protected iodinated phosphoroamidate, **15**, 100 mg, 0.13 mmol) was dissolved in 90 % aqueous TFA (5 mL) and stirred for 2 hours at room temperature. Then the solvent was repeatedly co-evaporated with methanol under reduced pressure. Product was purified using column chromatography on silica gel (10 % methanol in dichloromethane). Several co-evaporations with chloroform gave iodinated phosphoramidate (**3**, 58.8 mg, 62 %) as a white solid.

ESI-MS (*m/z*) calculated mass for product (**3**) C₂₆H₂₉N₆O₈PI [M+H]⁺: 711.10; observed mass: 711.30.

¹H NMR (400 MHz, CDCl₃) δ 8.03 (d, *J* = 21.0 Hz, 1H), 7.28 – 7.21 (m, 5H), 7.17 – 6.97 (m, 6H), 5.83 (s, 1H), 5.68 (s, 2H), 5.12 (s, 1H), 5.04 – 4.92 (m, 2H), 4.82 (s, 1H), 4.05 (q, *J* = 7.2 Hz, 1H), 3.75 – 3.57 (m, 1H), 1.18 (s, 3H).

7.1.4 Synthesis of AMPylation Probe 4



TBS- and acetal-protected adenosine (16). Imidazole (1107.7 mg, 16.27 mmol, 2.5 equiv.) and 2',3'-*O*-isopropylideneadenosine (acetal-protected adenosine, **13**, 2.0 g, 6.51 mmol, 1.0 equiv.) were dissolved in 20 mL anhydrous DMF and *tert*-butyldimethylchlorosilane (1177.1 mg, 16.27 mmol, 2.5 equiv.) was added to the solution at 0 °C under inert atmosphere. The mixture was stirred vigorously for 2 hours at 0 °C. The reaction was quenched with methanol in 1 hour and was partitioned between EtOAc (100 mL) and sat. aqueous Na₂CO₃ (75 mL). The organic layer was washed with water (3 × 50 mL) and brine (100 mL), dried over anhydrous MgSO₄ and evaporated *in vacuo*. The product was recovered as colorless crystals (**16**, 2.36 g, 86.1 %).

ESI-MS (*m/z*) calculated mass for product (**16**) C₁₉H₃₂N₅O₄Si [M+H]⁺: 422.22; observed mass: 422.37.

¹H NMR (400 MHz, CDCl₃) δ 8.36 (s, 1H), 8.03 (s, 1H), 6.15 (d, *J* = 2.5 Hz, 1H), 5.71 (s, 2H), 5.25 (dd, *J* = 6.2, 2.5 Hz, 1H), 4.94 (dd, *J* = 6.1, 2.5 Hz, 1H), 4.41 (td, *J* = 4.0, 2.4 Hz, 1H), 3.86 (dd, *J* = 11.2, 3.9 Hz, 1H), 3.75 (dd, *J* = 11.2, 4.2 Hz, 1H), 1.62 (s, 3H), 1.39 (d, *J* = 0.7 Hz, 3H), 0.82 (s, 9H).

8-Iodinated TBS- and acetal-protected adenosine (17). In a 100 mL round-bottom flask, *i*-Pr₂NH (659 μL, 4.7 mmol, 4.5 equiv) was dissolved in THF (3 mL) and cooled to -78 °C. *n*-BuLi (2.5 M in hexanes, 1880.3 μL, 4.7 mmol, 4.5 equiv) was added and the mixture was stirred for 30 min. 9-(((3*aR*,4*R*,6*R*,6*aR*)-6-(((*tert*-butyldimethylsilyl)oxy)methyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine (TBS- and acetal-protected adenosine, **16**, 440 mg, 1.04 mmol, 1.0 equiv) was dissolved in THF (3 mL) and added to the LDA solution via cannula. After stirring at -78 °C for 30 min, a solution of I₂ (662.8 mg, 2.61 mmol, 2.5 equiv) in THF (6 mL) was added via cannula over 10 min. The resulting deep orange solution was stirred at -78 °C for an additional 3 hours. The reaction was quenched by addition of sat. aqueous NH₄Cl (8 mL) and sat. aqueous Na₂S₂O₃ (15 mL) and extracted with EtOAc (3 × 30 mL). The combined organic extracts were washed with sat. aqueous NH₄Cl (2 × 50 mL) and sat.

aqueous NaCl (2 × 50 mL), dried over anhydrous MgSO₄, filtered and evaporated *in vacuo*. Purification was performed by silica chromatography (2:8 hexanes / EtOAc) to yield 8-iodinated TBS-and acetal-protected adenosine (**17**, 451.5 mg, 79%) as an off-white solid.

ESI-MS (m/z) calculated mass for product (**17**) C₁₉H₃₁N₅O₄Si [M+H]⁺: 548.11; observed mass: 548.25.

¹H NMR (400 MHz, CDCl₃) δ 8.28 (s, 1H), 6.14 (d, *J* = 1.9 Hz, 1H), 5.99 (d, *J* = 6.4 Hz, 2H), 5.87 (dd, *J* = 6.3, 1.9 Hz, 1H), 5.21 (dd, *J* = 6.3, 3.2 Hz, 1H), 4.35 (td, *J* = 6.6, 3.2 Hz, 1H), 3.80 (dd, *J* = 10.6, 6.8 Hz, 1H), 3.68 (dd, *J* = 10.6, 6.5 Hz, 1H), 1.68 (s, 3H), 1.46 (s, 3H), 0.88 (s, 9H).

8-Iodinated acetal-protected adenosine (18). To a solution of ((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-8-iodo-9*H*-purin-9-yl)-9-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methanol (8-iodinated TBS-and acetal-protected adenosine, **17**, 335 mg, 0.61 mmol, 1.0 equiv.) in dry THF (10 mL) was added tetrabutyl ammoniumfluoride (1.0 M TBAF in THF, 3.1 mL, 3.06 mmol, 5.0 equiv.) under inert atmosphere at room temperature. After stirring overnight, sat. aqueous NH₄Cl solution was added to quench the reaction. The aqueous phase was then extracted with EtOAc (3 × 30 mL). The combined organic solutions were washed with aqueous NaHCO₃ and brine, dried over anhydrous MgSO₄, filtered and evaporated *in vacuo*. Purification was performed by silica chromatography (1 - 5 % methanol in dichloromethane) to yield 8-iodinated acetal-protected adenosine (**18**, 182.8 mg, 69%) as a white solid.

ESI-MS (m/z) calculated mass for product (**18**) C₁₃H₁₇N₅O₄I [M+H]⁺: 434.04; observed mass: 434.16.

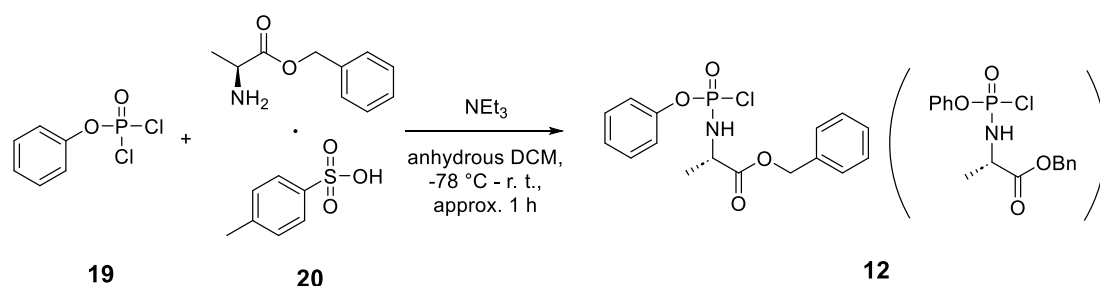
¹H NMR (400 MHz, MeOD) δ 8.15 (s, 1H), 6.07 (d, *J* = 3.5 Hz, 1H), 5.48 (dd, *J* = 6.1, 3.4 Hz, 1H), 5.05 (dd, *J* = 6.0, 2.7 Hz, 1H), 4.31 (td, *J* = 4.4, 2.6 Hz, 1H), 3.73 (dd, *J* = 12.1, 4.2 Hz, 1H), 3.65 (dd, *J* = 12.1, 4.8 Hz, 1H), 3.22 – 3.16 (m, 1H), 1.60 (s, 3H), 1.34 (s, 3H).

8-Iodinated adenosine (4). ((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-8-iodo-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methanol (8-iodinated acetal-protected adenosine, **18**, 100 mg, 0.23 mmol) was dissolved in 90 % aqueous TFA (5 mL) and stirred for 2 hours at room temperature. Then the solvent was repeatedly co-evaporated with methanol under reduced pressure. Product was purified using column chromatography on silica gel (1 - 10 % methanol in dichloromethane). Several co-evaporations with chloroform gave iodinated adenosine (**4**, 46.3 mg, 51 %) as a white solid.

ESI-MS (m/z) calculated mass for product (**4**) C₁₀H₁₃N₅O₄I [M+H]⁺: 394.06; observed mass: 394.16.

¹H NMR (500 MHz, CDCl₃) δ 8.15 (s, 1H), 6.10 (d, *J* = 7.8 Hz, 1H), 4.95 (dd, *J* = 7.8, 5.1 Hz, 1H), 4.46 (dd, *J* = 5.1, 1.1 Hz, 1H), 4.29 (d, *J* = 1.4 Hz, 1H), 3.90 (dd, *J* = 12.7, 1.9 Hz, 1H), 3.77 – 3.71 (m, 2H), 3.67 (q, *J* = 7.0 Hz, 2H), 2.41 – 2.35 (m, 2H).

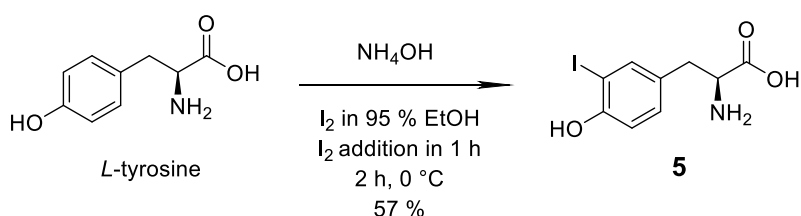
7.1.5 Synthesis of Phosphorochloridate



Glasswares used in this reaction must be fully dried (i.e. dry in the oven overnight before conducting the reaction). Anhydrous triethylamine (264 μL , 1.90 mmol, 2.0 equiv.) was added dropwise to a stirred solution of phenyl dichlorophosphate (**19**, 142 μL , 0.95 mmol, 1.0 equiv.) and *L*-alanine benzyl ester *p*-toluenesulfonate salt (**20**, 333 mg, 0.95 mmol, 1.0 equiv.) in anhydrous dichloromethane (4 mL) at -78 °C within inert atmosphere. Following the addition, the reaction mixture was stirred firstly at -78 °C for 30 minutes, then slowly warmed up to room temperature within 15 minutes. The mixture was allowed to be stirred for further 20 - 30 minutes, with reaction progress monitored by LC-MS. After this period the solvent was removed under reduced pressure and the residue was triturated with anhydrous THF. The precipitate was filtered and the solution was concentrated partially with some THF left inside. The crude product **12** was used for the future steps, including the synthesis for compounds **1** and **14**.

ESI-MS (*m/z*) calculated mass for product (**12**) $\text{C}_{16}\text{H}_{18}\text{NO}_4\text{PCl}$ $[\text{M}+\text{H}]^+$: 354.08; observed mass: 354.19 and 350.24 (chlorine substituted by methoxy group from solvent methanol).

7.1.6 Synthesis of Tyrosination Probes 5 and 6

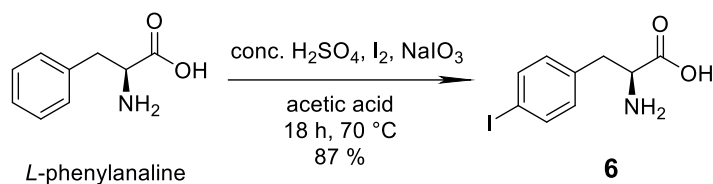


L-Tyrosine (5.00 g, 27.6 mmol, 1 equiv.) was dissolved in concentrated NH_4OH (500 mL) and cooled to 0 °C. Iodine (I_2 , 7.00 g, 27.6 mmol, 1 equiv.) was dissolved in 100 mL 95 % EtOH in water and added via a cannula needle to the *L*-tyrosine dropwise over 1 hour and stirred for additionally 2 hours. The solution was then concentrated under reduced pressure to a volume of 150 mL, at which time a precipitate began to form. The solution was acidified to pH 4.5 and cooled to 0 °C. After 1 hour at 0 °C, the crystals formed were collected by filtration under reduced pressure and air dried. The crystals were then suspended in acetone and stirred at 0 °C for 2 hours to remove residual iodine. They were again collected by suction filtration to yield pure product (*S*)-2-amino-3-(4-hydroxy-3-iodophenyl)propanoic acid (**5**) in 57 % yield (4.84 g). 307 mg compound **5** was dissolved in 1.5 mL 1 M NaOH and 8.5 mL water as a 100 mM stock solution.

ESI-MS (*m/z*) calculated mass for product (**5**) $\text{C}_9\text{H}_{11}\text{NO}_3\text{I}$ $[\text{M}+\text{H}]^+$: 307.98; observed mass: 308.11.

^1H NMR (500 MHz, MeOD) δ 7.74 – 7.62 (m, 1H), 7.16 (dt, $J = 7.9, 3.9$ Hz, 1H), 6.86 – 6.76 (m, 1H), 3.72 (dd, $J = 8.5, 4.4$ Hz, 1H), 3.19 (dd, $J = 14.7, 4.5$ Hz, 1H), 2.92 (dd, $J = 14.7, 8.5$ Hz, 1H).

^{13}C NMR (126 MHz, MeOD) δ 172.28, 156.08, 140.12, 139.71, 130.18, 130.10, 128.48, 115.32, 114.65, 83.66, 56.35, 56.14, 48.24, 48.07, 47.90, 36.01, 35.37.



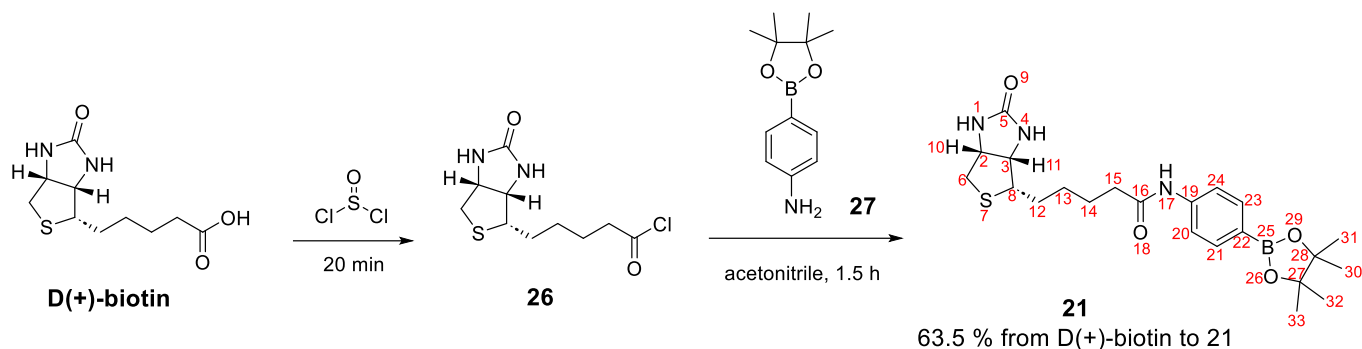
To a solution of *L*-phenylalanine (4.01 g, 24.3 mmol) in acetic acid (22 mL) were added concentrated sulfuric acid (2.9 mL, 5.14 mmol), iodine (2.47 g, 4.7 mmol) and sodium iodate (1.02 g, 5.14 mmol). The mixture was heated to 70 °C and allowed to stir at this temperature for 16 hours before an additional portion of sodium iodate (1.02 g, 5.14 mmol) was added. The reaction was left for a further 2 hours before being concentrated, dissolved in methanol (20 mL) and treated with NaOH (60 mL). The mixture was left to precipitate out of the basic solution overnight and the resulting solid was filtered by vacuum filtration to yield (*S*)-2-amino-3-(4-iodophenyl)propanoic acid (**6**, 6.15 g, 21.1 mmol, 86.9 %) as a pink solid. 291 mg compound **6** dissolved in 1.2 mL 1 M NaOH and 8.8 mL water as a 100 mM stock solution.

ESI-MS (*m/z*) calculated mass for product (**6**) $\text{C}_9\text{H}_{11}\text{NO}_2$ $[\text{M}+\text{H}]^+$: 291.99; observed mass: 292.07.

^1H NMR (500 MHz, D_2O) δ 7.67 – 7.59 (m, 2H), 7.00 – 6.89 (m, 2H), 3.93 (dd, $J = 7.7, 5.4$ Hz, 1H), 3.11 (dd, $J = 14.6, 5.5$ Hz, 1H), 2.98 (dd, $J = 14.6, 7.7$ Hz, 1H).

^{13}C NMR (126 MHz, D_2O) δ 172.99, 137.98, 134.43, 131.27, 92.56, 55.14, 35.60.

7.1.7 Synthesis of Biotinylated Boronic Acid Pinacol Ester **21**



To a vacuum-dried 25 mL round bottom flask was added 200 mg D(+)-biotin (0.819 mmol, 1 equiv.) and then placed under high vacuum for 1 hour. A volume of 2 mL of excess thionyl chloride (SOCl_2) was added to the stirring solution at room temperature. The reaction was allowed to proceed for 20 minutes and the excess thionyl chloride was removed by rotavapor. The resulting brown oil was resuspended in 25 mL chloroform and then concentrated to ensure the complete removal of SOCl_2 . The resulting residue **26** was dissolved in 2 mL of acetonitrile and a solution of 180 mg (0.819 mmol, 1 equiv.) 4-aminophenyl pinacol boronate **27** in 2 mL acetonitrile was added. The reaction mixture was stirred vigorously for 90 minutes, which was then poured into 50 mL diethylether to form a precipitate over 3 hours. The supernatant was partitioned by centrifugation and the precipitate was resuspended in 20 mL chloroform, which was washed with water (3 \times 30 mL), and dried over anhydrous MgSO_4 . The final compound (**21**, 232 mg, 63.5 %) was obtained by silica chromatography, using 1 - 10% methanol in dichloromethane ($R_f = 0.4$).

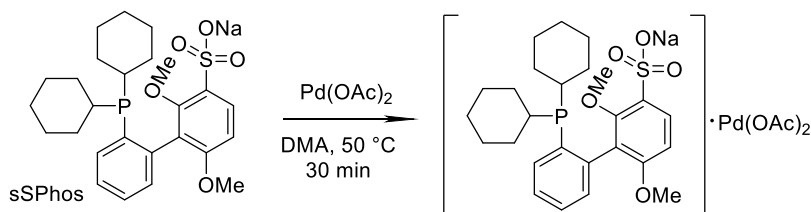
ESI-MS (*m/z*) calculated mass for product (**21**) $\text{C}_{22}\text{H}_{33}\text{BN}_3\text{O}_4\text{S}$ $[\text{M}+\text{H}]^+$: 446.23; observed mass: 446.40.

^1H NMR (500 MHz, Chloroform-*d*) δ (ppm) 8.57 (s, 1H), 7.78 – 7.69 (m, 2H, H-21, H-23), 7.62 (d, $J = 8.1$ Hz, 2H, H-20, H-24), 6.80 (s, 1H), 5.65 (s, 1H), 4.47 (dd, $J = 7.9, 4.8$ Hz, 1H, H-10), 4.26 (dd, $J = 8.0, 4.6$

Hz, 1H, H-11), 3.10 (td, $J = 7.4, 4.5$ Hz, 1H, H-8), 2.85 (dd, $J = 12.9, 4.8$ Hz, 1H, H-6a), 2.66 (d, $J = 12.8$ Hz, 1H, H-6b), 2.39 – 2.29 (m, 2H, H-15), 1.79 – 1.61 (m, 4H, H-12, H-14), 1.47 – 1.40 (m, 2H, H-13), 1.32 (s, 12H, H-30, H-31, H-32, H-33).

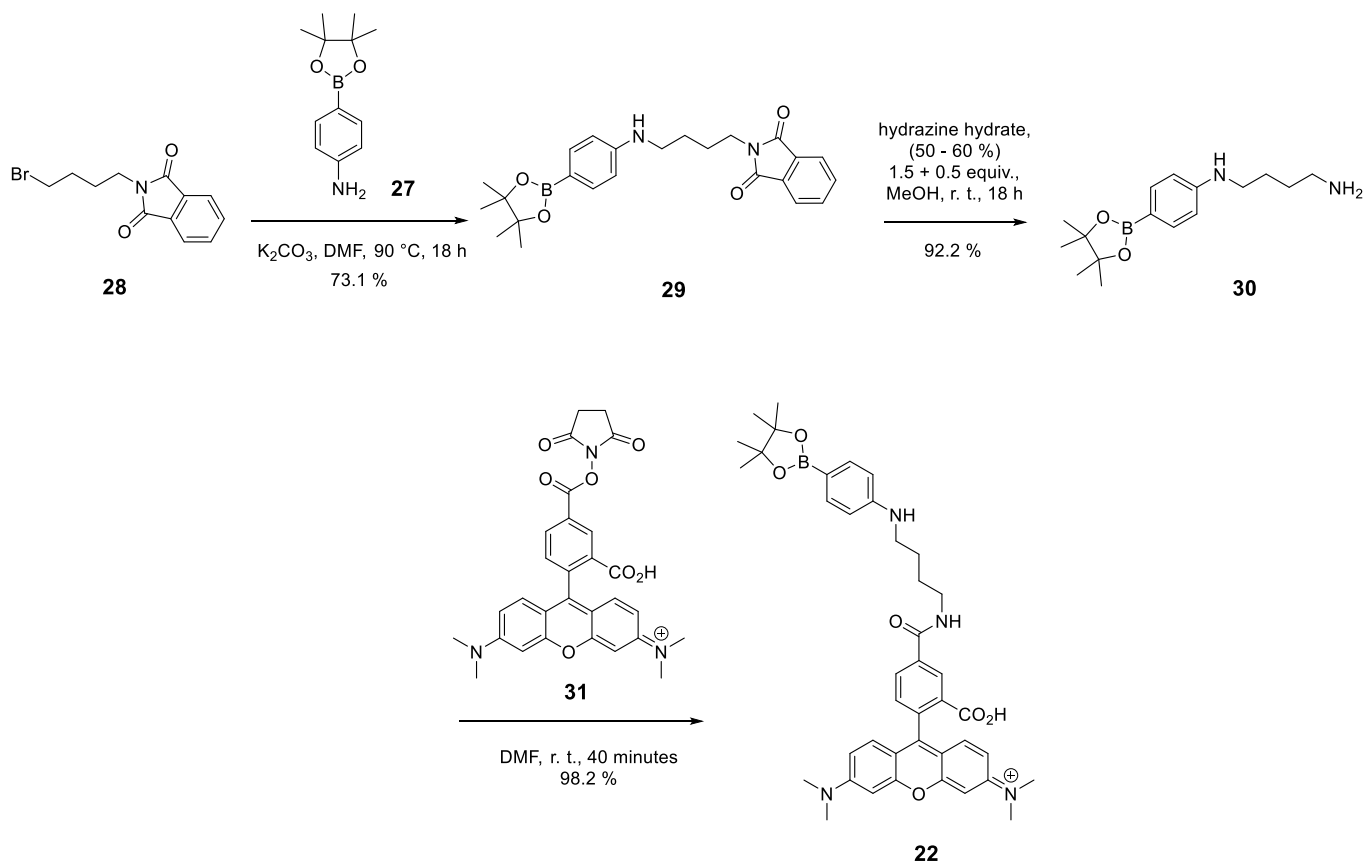
^{13}C NMR (126 MHz, Chloroform- d) δ (ppm) 172.17 (C-16), 164.35 (C-5), 141.31 (C-19), 135.75 (C-21, C-23), 118.69 (C-20, C-24), 83.74 (C-27, C-28), 61.72 (C-3), 60.25 (C-2), 55.78 (C-8), 40.57 (C-6), 37.00 (C-15), 28.22 (C-13), 28.02 (C-12), 25.68 (C-14), 24.88 (C-30, C-31, C-32, C-33).

7.1.8 Synthesis of Palladium Catalyst



4.5 mg (1 equiv.) of palladium acetate [$\text{Pd}(\text{OAc})_2$] and 10.3 mg (anhydrous basis, 1 equiv.) of sSPhos ([1,1'-biphenyl]-3-sulfonic acid, 2'-(dicyclohexylphosphino)-2,6-dimethoxy-sodium salt) were mixed within 200 μL DMA at 50 $^\circ\text{C}$ for 30 min. The resulting 100 mM stock solution of palladium catalyst was directly used for protein labelling by putatively Suzuki cross-coupling reaction.

7.1.9 Synthesis of TAMRA-Bpin 22



***N*-*n*-butylphthalimide-4-aminophenyl pinacol boronate (29).** To a stirring solution of 4-aminophenyl pinacol boronate (**27**, 207 mg, 0.945 mmol, 1 equiv.) in 3 mL DMF was added dropwise a solution of 4-bromo-*n*-butylphthalimide (**28**, 400 mg, 1.418 mmol, 1.5 equiv.) in 3 mL DMF and 326 mg potassium carbonate (K₂CO₃). When the addition was completed, the reaction mixture was heated up to 90 °C and stirred for 18 hours at this temperature. Then the solution was partitioned by pouring into water (30 mL) and extracting with ethyl acetate (40 mL). The organic phase was washed again with water (30 mL) and the combined water phase was washed once with EtOAc (40 mL). The combined organic phase was then dried with anhydrous MgSO₄ and purified by silica chromatography, using 10 - 40 % EtOAc in hexane (R_f = 0.6). The resulting greenish oil was again slowly mixed with cold methanol and then concentrated to get the final product as a white solid (**29**, 290 mg, 73.1 %).

ESI-MS (m/z) calculated mass for product (**29**) C₂₄H₃₀BN₂O₄ [M+H]⁺: 421.23; observed mass: 421.42.

¹H NMR (500 MHz, Chloroform-*d*) δ (ppm) 7.86 – 7.81 (m, 2H), 7.73 – 7.69 (m, 2H), 7.62 – 7.58 (m, 2H), 6.57 – 6.52 (m, 2H), 3.91 (s, 1H), 3.73 (t, *J* = 7.1 Hz, 2H), 3.22 – 3.15 (m, 2H), 1.79 (tt, *J* = 7.8, 6.3 Hz, 2H), 1.67 (dd, *J* = 8.4, 6.6 Hz, 2H), 1.30 (s, 12H). ¹³C NMR (126 MHz, Chloroform-*d*) δ (ppm) 207.14, 168.51, 150.71, 136.37, 134.01, 132.06, 123.28, 111.72, 83.19, 42.93, 37.65, 30.99, 26.64, 26.27, 24.84.

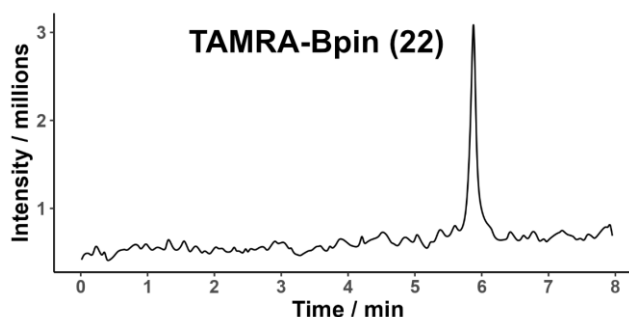
***N*-(4-amino-*n*-butyl)-4-aminophenyl pinacol boronate (30).** To a stirring solution of *N*-*n*-butylphthalimide-4-aminophenyl pinacol boronate (**29**, 85 mg, 0.202 mmol, 1 equiv.) in 2 mL methanol was added dropwise 20 μL hydrazine hydrate (50 % - 60 % in H₂O, approx. 0.303 mmol, approx. 1.5 equiv.). The mixture was stirred vigorously at room temperature. After 6 hours, 6.7 μL hydrazine hydrate (50 % - 60 % in H₂O, approx. 0.101 mmol, approx. 0.5 equiv.) was added dropwise to the reaction mixture and stirred for another 12 hours. The solvent was then evaporated and the residue was placed under high vacuum for 6 hours, and the obtained white solid (**30**, 54 mg, 0.186 mmol, 92.2 %) was directly dissolved in DMF, used for the next step without any further purification.

ESI-MS (m/z) calculated mass for product (**30**) C₁₆H₂₈BN₂O₂ [M+H]⁺: 291.22; observed mass: 291.33.

¹H NMR (500 MHz, Chloroform-*d*) δ (ppm) 7.64 – 7.58 (m, 2H), 6.60 – 6.49 (m, 2H), 3.15 (t, *J* = 7.0 Hz, 2H), 2.74 (t, *J* = 7.0 Hz, 2H), 1.68 – 1.62 (m, 2H), 1.57 – 1.52 (m, 2H), 1.31 (s, 12H).

TAMRA-Bpin (22). To a solution of 5-carboxy-tetramethylrhodamine-*N*-hydroxysuccinimide (**31**, 2.6 mg, 5 μmol, 1 equiv.) in 100 μL of DMF was added a solution of *N*-(4-amino-*n*-butyl)-4-aminophenyl pinacol boronate (**30**, 1.45 mg, 5 μmol, 1 equiv.) dissolved in 15 μL of DMF and the reaction mixture was stirred at room temperature for 40 minutes. Then this reaction mixture was frozen and lyophilized to get the final compound (**22**) as a magenta solid (3.4 mg, 98.2 %).

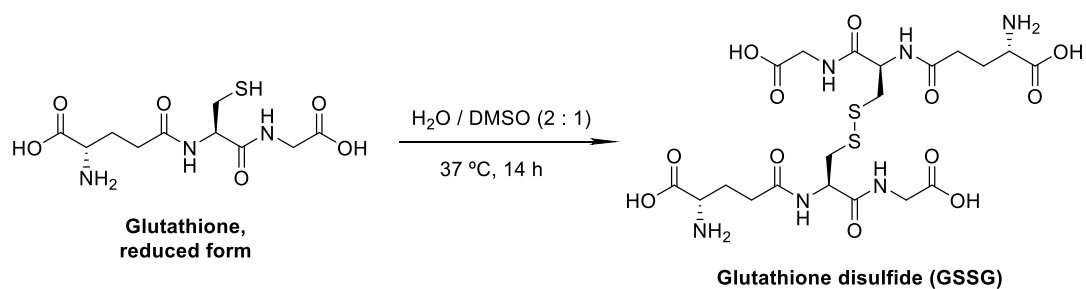
ESI-MS (m/z) calculated mass for product (**22**) C₄₁H₄₈BN₄O₆⁺ [M+H]⁺: 703.37; observed mass: 703.54.



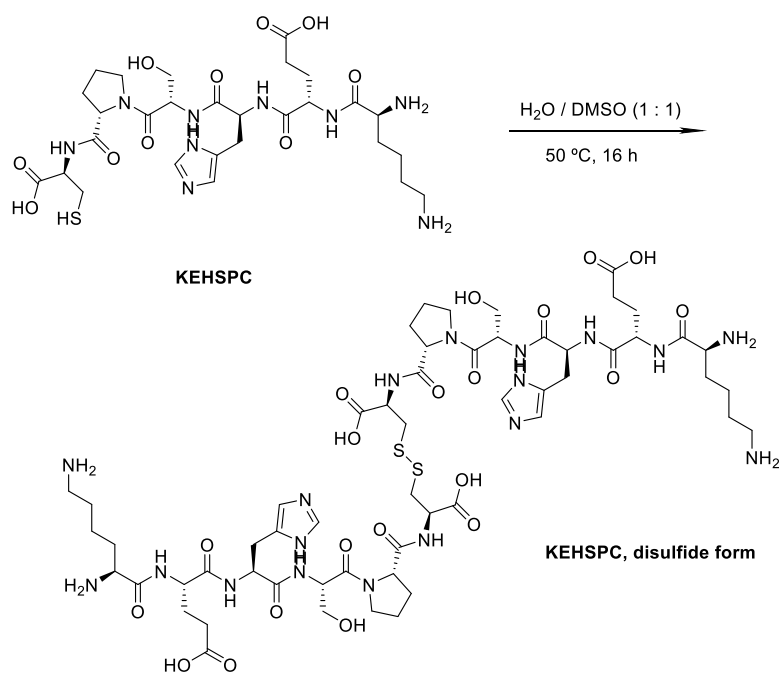
LC-MS chromatogram of TAMRA-Bpin (**22**)

7.1.10 Synthesis of GSSG, KEHSPC Disulfide and RWTPCD Disulfide

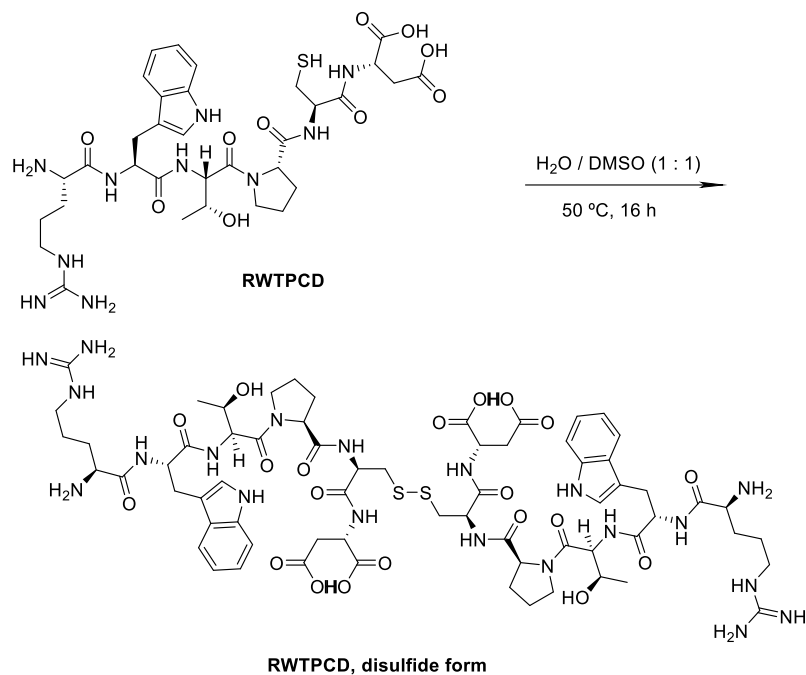
General method: a volume of 100 μL 100 mM peptide (glutathione, KEHSPC, RWTPCD) stock solution was mixed thoroughly with DMSO. The mixture was then incubated in a thermomixer. The resulting solution was lyophilized to remove excess DMSO and water, reconstituted in 50 μL Milli-Q water (100 mM stock solution).



GSSG: 50 μL DMSO added, incubate at 37 $^\circ\text{C}$ for 14 h;



KEHSPC disulfide: 100 μL DMSO added, incubate at 50 $^\circ\text{C}$ for 16 h;

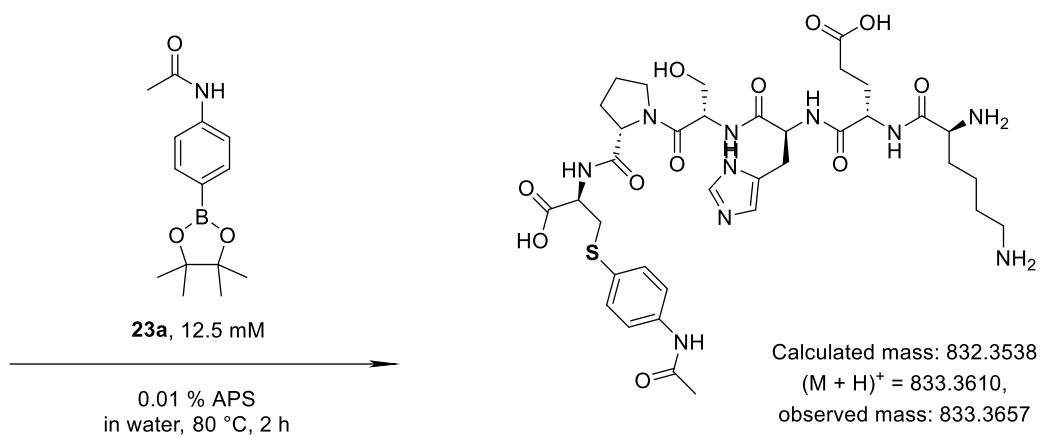
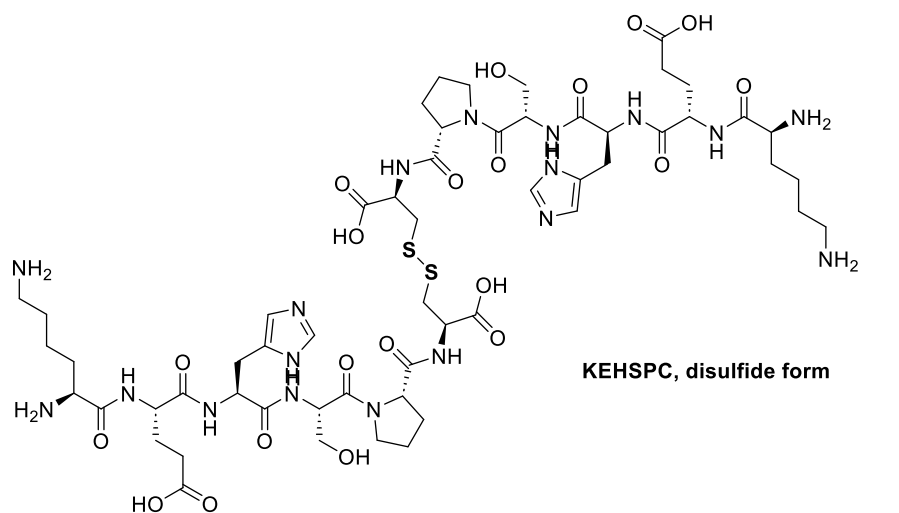
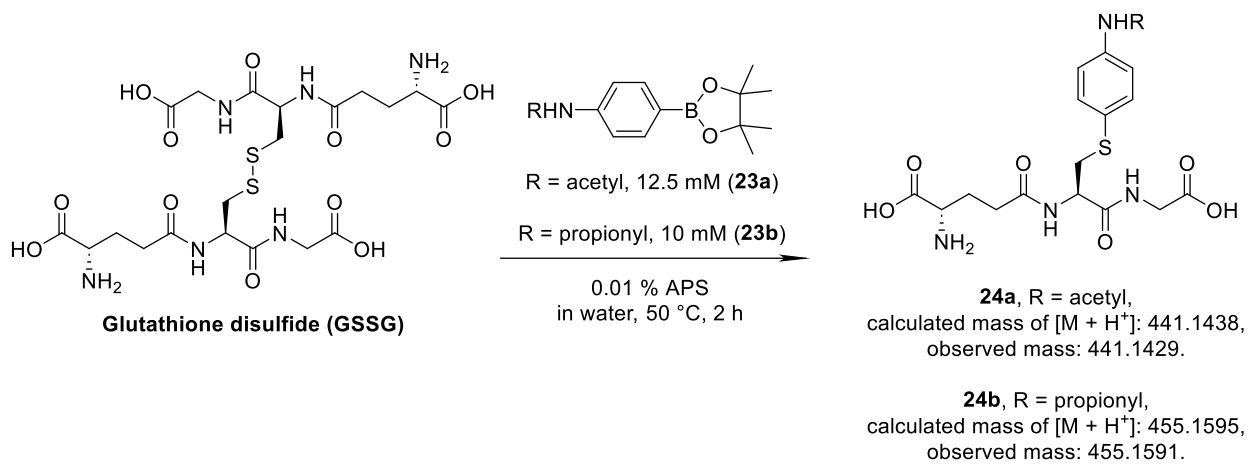


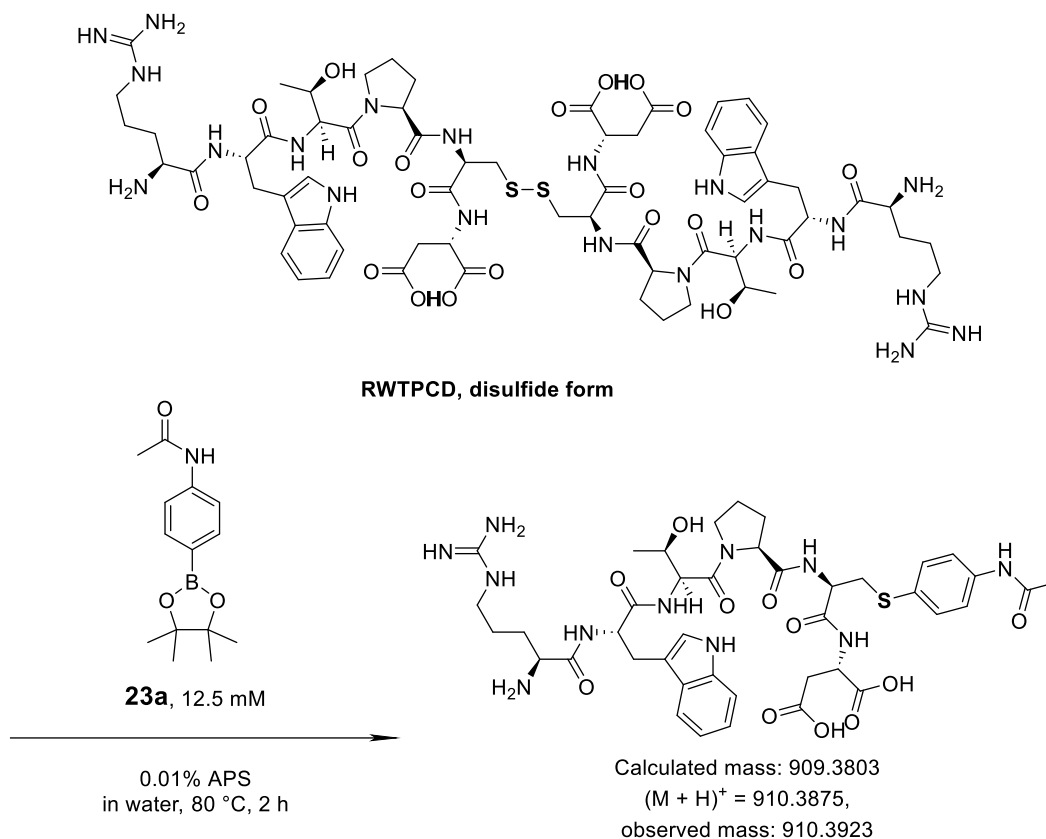
RWTPCD disulfide: 100 μL DMSO added, incubate at 50 $^\circ\text{C}$ for 16 h.

7.1.11 Synthesis of S-arylated Glutathione, KEHSPC and RWTPCD

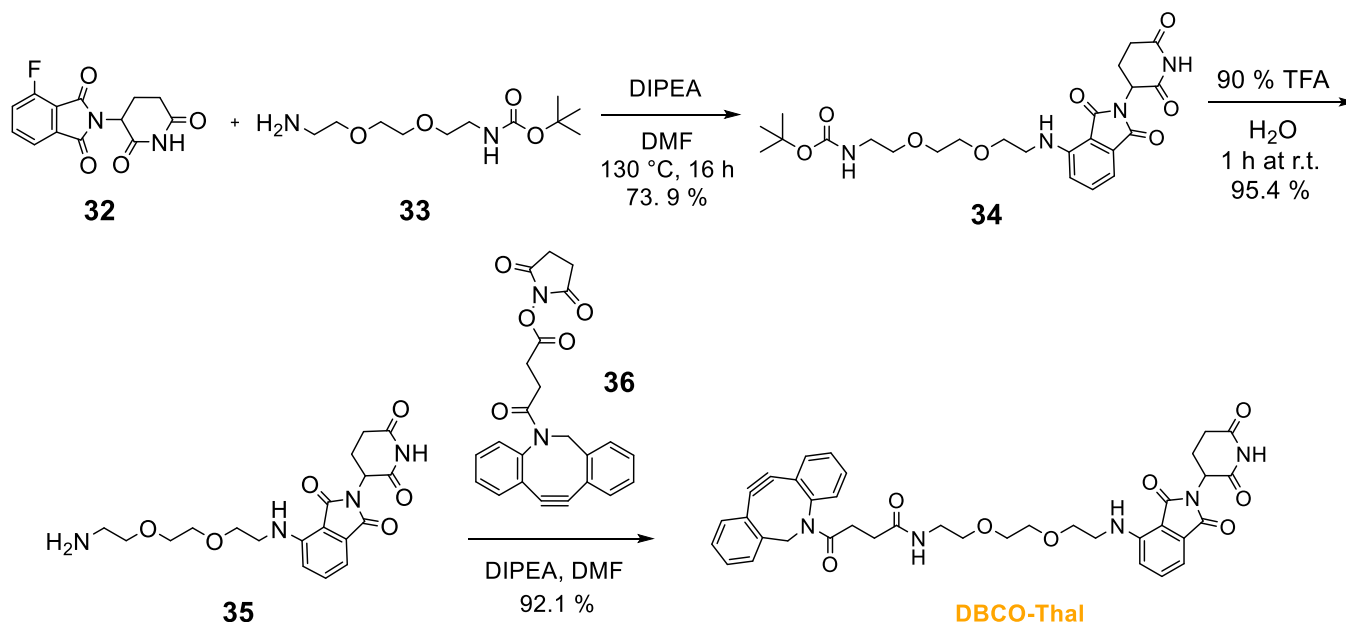
General method: A 2 μL 50 mM peptide disulfide (GSSG, KEHSPC & RWTPCD disulfide) stock solution was mixed with 1 μL 500 mM *N*-acetyl-4-aminophenyl pinacol boronate (**23a**) stock solution (three replicates in total). Afterwards the total volume was topped by MS-scale water to either 40 μL (control experiment, 1 sample) or 39 μL (2 samples). For the latter 2 samples, 1 μL DMA (2.5 % as final concentration) or 1 μL 0.4 % APS (m/v) stock solution in water (0.01 % as final concentration) were added, respectively. The samples were mixed thoroughly and incubated at 80 $^\circ\text{C}$ (for GSSG, the temperature was decreased to 50 $^\circ\text{C}$) for 2 hours with vigorous shaking, and then lyophilized to remove water. The residual white powders were reconstituted within 40 μL 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8 μL) and the rest were diluted with 200 μL 50 % acetonitrile and 0.5 % formic acid in water for intact measurement, with high-resolution MS analysis and MS2 with HCD fragmentation.

For GSSG, the same reaction setup was repeated simply by substituting *N*-acetyl-4-aminophenyl pinacol boronate (**23a**) with *N*-propionyl-4-aminophenyl pinacol boronate (**23b**) with indicated concentrations (Figure S13).





7.1.12 Synthesis of DBCO-Thal



4-(mono-*N*-Boc-diamine)thalidomide (34). 4-Fluorothalidomide (**32**, 100 mg, 0.365 mmol, 1.0 equiv.) was dissolved within 4 mL DMF, then mono-*N*-Boc-protected diamine (**33**, 95.3 μ L, 0.402 mmol, 1.1 equiv.) and DIPEA (190.7 μ L, 1.095 mmol, 3.0 equiv.) were added dropwise. The resulting mixture was continuously heated in an oil bath at 130 °C for 16 hours. Afterwards the reaction mixture was repeatedly co-evaporated with toluene and the residues were purified by flash chromatography over silica gel, eluting with EtOAc in hexane (0 - 100 %). The product, 4-(mono-*N*-Boc-diamine)thalidomide (**34**), was obtained

as a bright yellowish oil with a yield at 73.9 % (136 mg).

ESI-MS (m/z) calculated mass for product (**34**) $C_{24}H_{33}N_4O_8$ $[M+H]^+$: 505.23; observed mass: 505.43.

1H NMR (500 MHz, $CDCl_3$) δ (ppm) 8.45 (s, 1H), 7.52 – 7.43 (m, 1H), 7.12 – 7.05 (m, 1H), 6.89 (dd, $J = 13.8, 8.5$ Hz, 1H), 6.52 (d, $J = 7.1$ Hz, 1H), 5.05 (s, 1H), 4.96 – 4.85 (m, 1H), 3.71 (q, $J = 4.9$ Hz, 2H), 3.65 (qt, $J = 5.7, 2.9$ Hz, 4H), 3.58 – 3.54 (m, 3H), 3.47 (q, $J = 5.4$ Hz, 2H), 3.32 (q, $J = 5.7$ Hz, 2H), 2.87 – 2.79 (m, 1H), 2.75 (td, $J = 13.1, 6.4$ Hz, 1H), 2.16 – 2.08 (m, 1H), 1.66 (s, 1H), 1.43 (d, $J = 3.6$ Hz, 9H).

4-diaminethalidomide (35). To a solution of 4-(mono-*N*-Boc-diamine)thalidomide (**34**, 136 mg, 0.27 mmol) was added dropwise 9 mL TFA and 1 mL water. The resulting mixture was stirred at room temperature for 1 hour. Afterwards the solvents were removed by adding methanol, followed by repeatedly co-evaporation. The residues were separated by flash chromatography over silica gel, eluting with methanol in dichloromethane (0 - 10 %). The product (4-diaminethalidomide, **35**) was obtained as a yellowish solid with a yield at 95.4 % (104 mg).

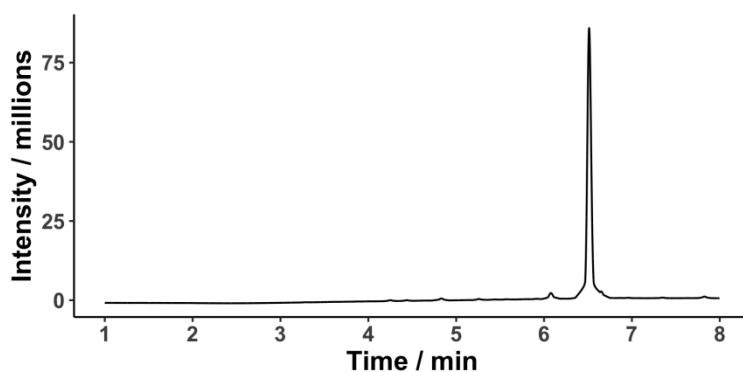
ESI-MS (m/z) calculated mass for product (**35**) $C_{19}H_{25}N_4O_6$ $[M+H]^+$: 405.18, observed mass: 405.40.

1H NMR (500 MHz, $CDCl_3$) δ (ppm) 9.66 (s, 1H), 7.47 (dd, $J = 8.5, 7.1$ Hz, 1H), 7.04 (d, $J = 7.1$ Hz, 1H), 6.86 (d, $J = 8.5$ Hz, 1H), 6.53 (s, 1H), 4.96 (dd, $J = 11.9, 5.8$ Hz, 1H), 3.84 – 3.56 (m, 8H), 3.43 (qd, $J = 7.8, 5.2$ Hz, 2H), 3.20 (s, 2H), 2.82 – 2.62 (m, 3H), 2.04 (ddd, $J = 12.4, 6.3, 4.3$ Hz, 1H), 1.35 – 1.13 (m, 1H).

To a solution of 4-diaminethalidomide (**35**, 52 mg, 0.129 mmol, 1.1 equiv.) in 3 mL DMF was added DBCO-NHS ester (**36**, 47 mg, 0.117 mmol, 1 equiv.) and DIPEA (61.1 μ L, 0.351 mmol, 3.0 equiv.) and the resulting mixture was gently stirred at room temperature for 30 minutes. Afterwards the reaction mixture was repeatedly co-evaporated with toluene and the residues were purified by flash chromatography over silica gel, eluting with methanol in dichloromethane (0 - 12 %). The product (**DBCO-Thal**) was obtained as a yellowish solid with a yield at 92.1 % (70 mg).

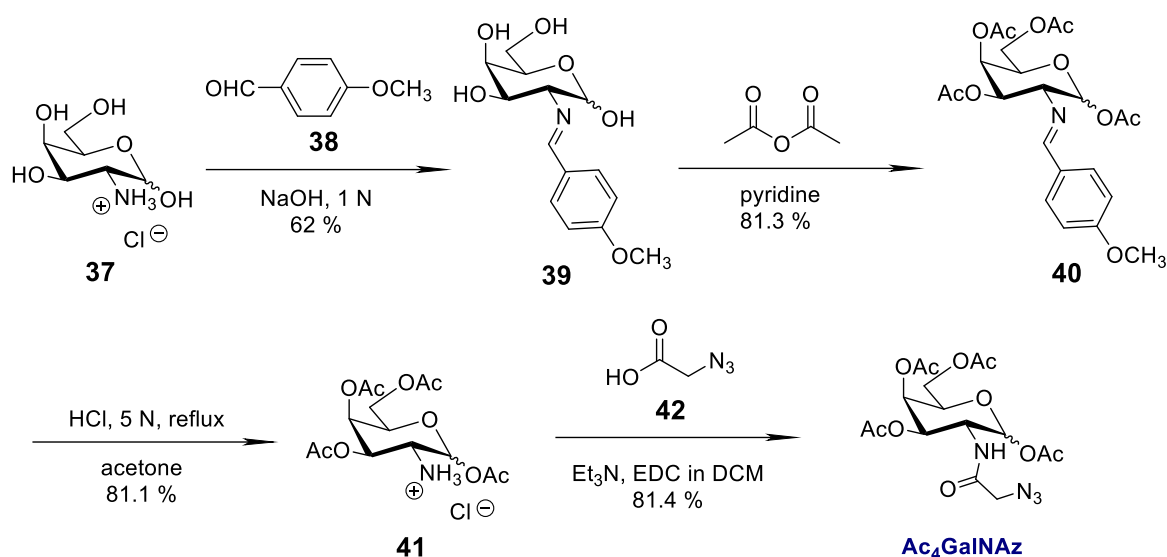
ESI-MS (m/z) calculated mass for product (**DBCO-Thal**) $C_{38}H_{38}N_5O_8$ $[M+H]^+$: 692.27, observed mass: 692.49.

1H NMR (500 MHz, $CDCl_3$) δ (ppm) 7.69 (dd, $J = 7.2, 5.0$ Hz, 1H), 7.52 – 7.43 (m, 1H), 7.42 – 7.26 (m, 4H), 7.23 (ddd, $J = 7.0, 5.4, 1.7$ Hz, 1H), 7.10 (d, $J = 7.1$ Hz, 1H), 6.87 (dd, $J = 8.5, 3.4$ Hz, 1H), 6.52 (dt, $J = 11.4, 5.5$ Hz, 1H), 6.17 (dt, $J = 39.9, 5.5$ Hz, 1H), 5.16 (dd, $J = 13.9, 9.3$ Hz, 1H), 4.92 – 4.82 (m, 1H), 3.75 – 3.49 (m, 7H), 3.47 – 3.26 (m, 5H), 2.91 – 2.65 (m, 4H), 2.54 – 2.38 (m, 1H), 2.21 – 2.05 (m, 2H), 1.92 (ddt, $J = 31.1, 16.9, 6.0$ Hz, 1H).



LC-MS chromatogram of DBCO-Thal

7.1.13 Synthesis of Ac₄GalNAz



Tetraacetylated galactosamine hydrochloride (41). D-(+)-Galactosamine hydrochloride (**37**, 400 mg, 1.855 mmol, 1.0 equiv.) was dissolved within 1 M NaOH (4 mL) at 0 °C, followed by the addition of *p*-anisaldehyde dropwise (**38**, 2.26 mL, 18.55 mmol, 10.0 equiv.). The resulting mixture was then stirred vigorously at room temperature until a white precipitate was formed (approx. 20 minutes), which was placed at 4 °C for 2 hours to complete precipitation. The white crystalline was filtered, washed with iced water and Et₂O, and dried under vacuum to get the product (**39**, 341.6 mg, 1.150 mmol, 62.0 %). This compound was directly dissolved in pyridine (5 mL) and acetic anhydride (1087 μL, 11.50 mmol, 10.0 equiv.) was added dropwise at 0 °C. The reaction mixture was then stirred vigorously at 0 °C for 30 minutes, slowly warmed up to room temperature for another 16 hours with gentle stirring. Afterwards the mixtures were all poured into iced water and kept at 4 °C for 6 hours. Then the precipitate was filtered, washed with cold water and vacuum-dried to get the tetraacetylated imine intermediate (**40**, 434.9 mg, 0.935 mmol, 81.3 %). This compound was then fully dissolved in acetone (5 mL) and heated to reflux. Hydrochloric acid (5 N, 300 μL) was quickly added into the solution and the reflux was continued for another 20 minutes. The mixture was cooled down to room temperature and poured into cold Et₂O for 2 hours. The precipitate was obtained through filtration and washing with cold Et₂O, which was further dried completely under high vacuum for 8 hours to get the product tetraacetylated galactosamine (**41**, 263.8 mg, 0.758 mmol, 81.1 %). This compound was used for the amide coupling without any further purification.

ESI-MS (*m/z*) calculated mass for product (**41**) C₁₄H₂₂NO₉ [M+H]⁺: 348.13; observed mass: 348.19.

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm) 8.77 (d, *J* = 17.9 Hz, 3H, NH₃⁺), 5.90 (dd, *J* = 8.7, 2.9 Hz, 1H, H-1), 5.32 – 5.26 (m, 2H, H-3, H-4), 4.32 – 4.28 (m, 1H, H-5), 4.04 (h, *J* = 6.7, 6.3 Hz, 2H, H-6), 3.84 – 3.66 (m, 1H, H-2), 2.17 (s, 3H), 2.13 (s, 3H), 2.00 (d, *J* = 7.8 Hz, 6H).

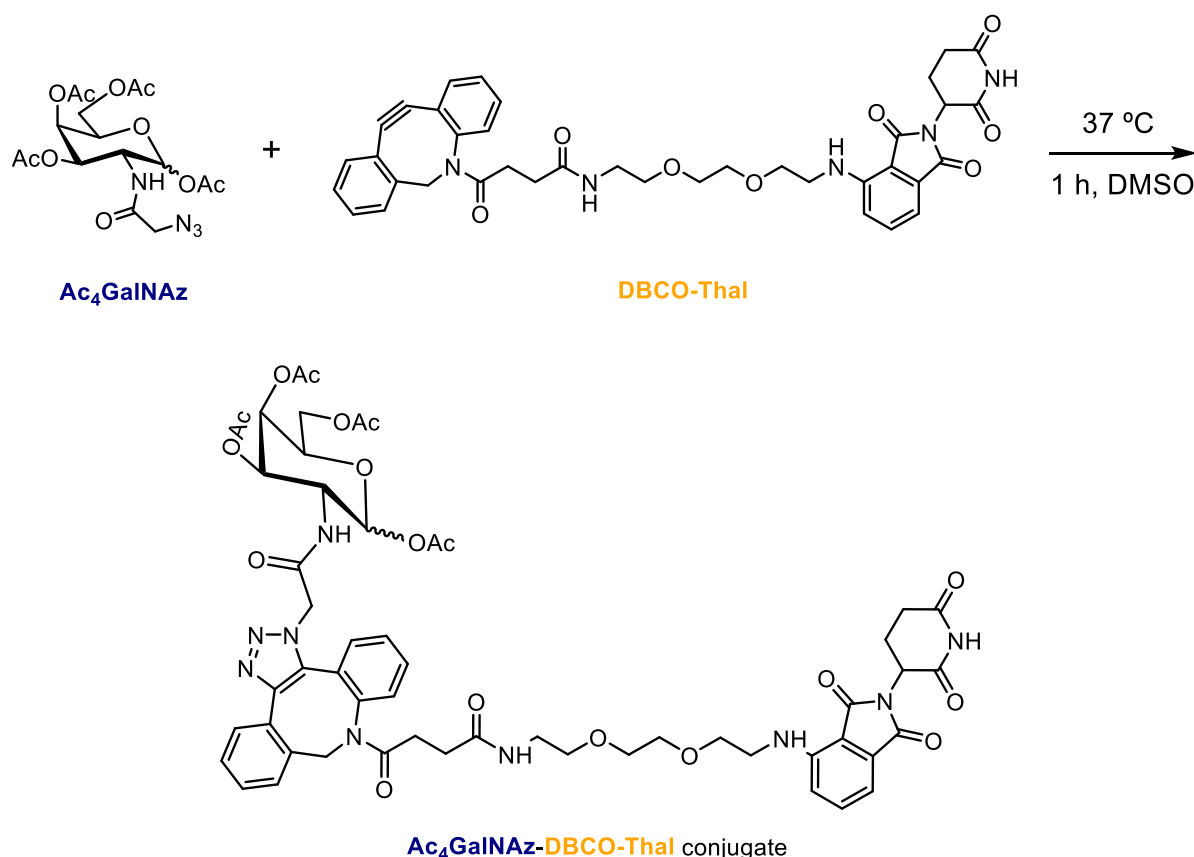
To a stirring solution of azidoacetic acid (**42**, 240.84 μL, 3.217 mmol, 5.0 equiv.) in anhydrous dichloromethane (5 mL) was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, 249.7 mg, 1.609 mmol, 2.5 equiv.) at 0 °C. This mixture was gently stirred at 0 °C for 20 minutes for active ester formation. Meanwhile a solution of tetraacetylated galactosamine (**41**, 224 mg, 0.643 mmol, 1.0 equiv.)

in anhydrous dichloromethane was mixed with anhydrous triethylamine (134.52 μL , 0.965 mmol, 1.5 equiv.) and kept stirring at 0 °C. Then the previous mixture containing azidoacetic acid and EDC was added dropwise into tetraacetylated galactoseamine solution within 5 minutes. This complete mixture was reacted firstly at 0 °C for 1 hour, then moved to room temperature for 16 hours with stirring. After that the final compound (**Ac₄GalNAz**, 225 mg, 0.523 mmol, 81.4 %) was obtained through flash chromatography by eluting with EtOAc and hexane (3:2, v/v).

ESI-MS (m/z) calculated mass for product (**Ac₄GalNAz**) C₁₆H₂₂N₄O₁₀Na [M+Na]⁺: 453.1234; observed mass: 453.1198.

¹H NMR (500 MHz, CDCl₃) δ (ppm) 6.69 (d, *J* = 9.2 Hz, 1H, NH), 5.83 (d, *J* = 8.8 Hz, 1H, H-1), 5.37 (d, *J* = 3.3 Hz, 1H, H-3), 5.26 (dd, *J* = 11.2, 3.4 Hz, 1H, H-4), 4.34 (dt, *J* = 11.3, 9.0 Hz, 1H, H-2), 4.15 – 4.08 (m, 3H, H-6a, H-6b, H-5), 3.90 (s, 2H, CH₂N₃), 2.14 (s, 3H), 2.10 (s, 3H), 2.02 (s, 3H), 1.98 (s, 3H). ¹³C NMR (126 MHz, CDCl₃) δ (ppm) 170.6, 170.5, 170.3, 169.6, 167.8, 92.5, 71.7, 70.0, 66.5, 61.4, 52.6, 50.0, 20.9, 20.6, 20.6.

7.1.14 Synthesis of Ac₄GalNAz-DBCO-Thal Conjugate



5 μL 200 mM Ac₄GalNAz in DMSO was mixed with 5 μL 200 mM DBCO-Thal in DMSO and vortexed thoroughly. The resulting mixture was then incubated at 37 °C for 1 hour, subjected to lyophilization to get the product as a white powder without any further purification. ESI-MS (m/z) calculated mass for product C₅₄H₅₉N₉O₁₈Na [M+Na]⁺: 1144.3876, observed mass: 1144.3823.

7.2 Biological Methods and Analysis

7.2.1 Cell Culture, Harvesting, Lysates Preparation and Protein Concentration Measurement

Cell culture

The HeLa cells were cultivated in a high-glucose Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % (v/v) heat-inactivated fetal bovine serum (FBS) and 2 % (v/v) L-glutamine. Cells were maintained in 10-cm cell culture dishes for adherent cells at 37 °C under constant humidity and 5 % CO₂ concentration. The same conditions were applied for culturing HCT116, HepG2, MCF7, SH-SY5Y, and SK-MEL-28 cell lines. For A549 and CHO cell lines, the medium was exchanged to Kaighn's Modification of Ham's F-12 (F-12K) Medium supplemented with 10 % (v/v) heat-inactivated fetal bovine serum and 2 % (v/v) L-glutamine. For K562 cell line, the medium was exchanged to RPMI 1640 medium supplemented with 10 % (v/v) heat-inactivated fetal bovine serum. The subcultivation ratio at approx. 1:6 was applied to all cell lines.

Harvesting HeLa cells

HeLa cells were harvested after specified treatments, or when they were grown to 90 - 95% confluence if no probe was treated. The cell culture medium was firstly removed, then the cells were washed twice with 5 mL cold PBS. After removing PBS washing buffer, cells were scraped with 1 mL PBS buffer, transferred to a 1.5 mL tube and placed at 4 °C shortly. Afterwards the cell suspension was centrifuged at 600 rpm, 4 °C for 5 minutes. The supernatant was then carefully removed, and the rest cell pellet was left for lysates preparation.

Lysates preparation

The cell pellet was reconstituted in 200 µL PBS buffer and was sonicated with an ultrasonic rod sonicator at 1s-on / 1s-off cycles at 20 % intensity for a total time of 10 seconds. The solution was clarified by centrifugation at 4°C, 14000 rpm for 15 min. The cell debris was discarded and the supernatant was then transferred to a new 1.5 mL tube and stored at -80 °C until further use.

Protein concentration measurement

Protein concentration measurement was performed with Pierce™ BCA Protein Assay Kit from Thermo Scientific.

7.2.2 ThioredoxinA (wild type, C33S, C36S) Expression and Purification

Materials

Chemicals were purchased from Merck (Sigma-Aldrich). Oligonucleotides were purchased from Biomers (Germany). Restriction enzymes, Miniprep kit, gel extraction kit, DNA purification kits and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA).

Methods

I. Cloning

To obtain the sequence encoding for the His₆-tagged fusion of *Escherichia coli* TrxA (Uniprot No. P0AA25),

a stop codon was introduced in the expression plasmid pETtrx_1a. Mutations of the two cysteine mutants, resulting in the His₆-TrxA^{C33A} and His₆-TrxA^{C36A}-expression constructs, were done by site-directed mutagenesis.

Table 2 Oligonucleotide sequences

Name	Sequence 5' -> 3'
TrxA - Fwd	CTCGAGCACCACCA
TrxA - Rev	TCATTAGGCCAGGTTAG
TrxA C33A - Fwd	GGTCCGGTCCGTGC
TrxA C33A - Rev	ACTCTGCCCAGAAATCGAC
TrxA C36A - Fwd	GTCCGTCCAAAATGATCG
TrxA C36A - Rev	CGCACCACTCTGC

Table 3 Protein sequences of the His-tagged *E. coli* TrxA (13 kDa), with the two cysteine residues C33 and C36 highlighted in red.

Name	Sequence
His ₆ -TrxA	MKHHHHHHPMSDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPCCKMIAPILDEIADEY QGKLTVAKLNIDQNPQTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA*

II. Protein expression and purification

For protein expression, the plasmids encoding for His₆-TrxA, His₆-TrxA^{C33A} or His₆-TrxA^{C36A} were transformed in *E. coli* B121 (DE3). Cells obtained from an overnight culture (Lysogeny broth (LB) medium supplemented with 50 µg / mL kanamycin) were diluted 1:100 in LB medium (kanamycin) and incubated at 37 °C, 180 rpm until an OD₆₀₀ of 0.6 was reached. Protein expression was induced with 0.5 mM isopropyl-β-D-1-thiogalactopyranoside and the cells were incubated overnight at 16 °C, 180 rpm. Cells were subsequently harvested by centrifugation (6 000 g, 20 minutes). All purification steps were performed at 4 °C or on ice. The two-step purification procedure of TrxA and TrxA-mutants comprised a His-affinity chromatography followed by a size-exclusion chromatography (SEC) column (Superdex 200 Increase 10/300 GL). The cell pellet obtained from a 1 L expression was resuspended in 40 mL lysis buffer (Dulbecco's Phosphate-Buffered saline (DPBS, Sigma-Aldrich)), supplemented with 20 mM imidazole, DNase I (AppliChem) and complete mini EDTA-free protease inhibitors (Roche). Cells were lysed by homogenization using an EmulsiFlexC5 (Avestin Inc.) and clear lysate was obtained by centrifugation (24,446 g, 30 minutes, 4 °C) and followed by filtration using a Whatman TM folded filter (Cytiva). The cleared lysate was loaded to nickel-NTA agarose resin (Qiagen). After sample application the resin was washed with wash buffer (DPBS, 20 mM imidazole), followed by elution of His₆-TrxA (DPBS, 500 mM imidazole). Elution containing the protein was concentrated using an Amicon Ultracell Centrifugal filter unit (MWCO 3 kDa, Merck Millipore) and applied onto a Superdex 200 increase (10/300) column (Cytiva), equilibrated with DPBS. Fractions containing the pure protein were pooled and aliquots flash-frozen in liquid nitrogen and stored at -80 °C.

7.2.3 Methods for Cell Lysates or Single Protein Labelling

Figure 23: Cell lysates labelling with Biotin-Bpin (21)

For each sample, cell lysates containing 600 µg proteins were aspirated from stock solution. The cell lysates were then supplemented with PBS buffer to a total volume of 182 µL. Afterwards 16 µL 100 mM biotinylated boronic acid pinacol ester (**21**, final conc. at 8 mM) and 2 µL 100 mM palladium catalyst (§ 7.1.8, final conc. at 1 mM) were added, and the resulting mixtures were incubated at 80 °C, 750 rpm for 2 hours. The labeled proteins were subjected to SP2E protocol for clean-up and enrichment.

Figure 26 C-F, 31B-C, S1-6, S9-11, S29: Cell lysates labelling with TAMRA-Bpin (22)

For each sample, cell lysates containing 100 µg proteins were aspirated from stock solution. The cell lysates were then supplemented with 1 µL 4 mM fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, final conc. at 100 µM) and / or DMA (1 µL), followed by topping up the total volume to 40 µL with PBS buffer. The mixtures were incubated in darkness at 50 °C, 750 rpm for 2 hours. The labeled proteins were resolved by 10 % SDS-PAGE gel.

Figure 27 A: Cell lysates labelling with GODCAT system then TAMRA-Bpin (22)

A glucose-oxidase (*Asp. Niger Type VII*, Sigma Aldrich) and catalase (*Asp. Niger*, Sigma Aldrich) mixture was stored in aliquots containing 6.25 µM glucose-oxidase, 400 - 800 nM catalase, 40 % glycerol and 2 mM tris(2-carboxyethyl)phosphine (TCEP) at -20 °C. For oxygen scavenging, 10 % (v/v) of the prepared mixture and/or D-glucose to a final concentration of 1 % (w/v) were added to the cell lysates and sealed to exclude air oxygen. Afterwards the lysates were incubated at 37 °C for 10 minutes, followed by treatment of fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 µL, final conc. 100 µM) in a total volume of 40 µL PBS buffer at 50 °C, 750 rpm for 2 hours. The labeled cell lysates were resolved by 10 % SDS-PAGE gel.

Figure S33: Cell lysates labelling with DBCO-TAMRA

For each sample, cell lysates containing 20 µg proteins were aspirated from stock solution. The cell lysates were then supplemented with 0.2 % SDS in PBS buffer to a total volume of 19 µL. Afterwards 1 µL 300 mM iodoacetamide in water (final conc. at 15 mM) was added to each sample and incubated at room temperature for 30 minutes in darkness. The resulting mixture was added with 1 µL approx. 0.4 mM DBCO-TAMRA in DMSO to reach the final concentration at 20 µM. The labelling was carried out also at room temperature for 30 minutes in darkness. The labeled cell lysates were resolved by 10 % SDS-PAGE gel.

Figure 27 D: ThioredoxinA labelling with TAMRA-Bpin (22)

100 µg thioredoxinA (including WT, C33S and C36S) were topped to 48 µL by PBS buffer, respectively. Afterwards they were treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 µL, final conc. 100 µM) or TAMRA-N₃ (1 µL, final conc. 100 µM), with either 1 µL DMA or 1 µL 0.4 % APS (w/v) at 50 °C, 750 rpm for 2 hours. After incubation, the reaction mixture was added with 200 µL acetone and placed in the -20 °C freezer for 2 hours, which was further spun down at 4 °C, 13,000 rpm for 10 minutes. The supernatants were then removed and the protein pellets were washed twice with cold methanol, reconstituted within 80 µL 0.2 % SDS in PBS buffer and sonicated with an ultrasonic rod sonicator at 1s-on / 1s-off cycles at 20% intensity for a total time of 10 seconds. The labeled thioredoxinA and its variants were resolved by 18 % SDS-PAGE gel.

Figure 27 C: BSA labelling with TAMRA-Bpin (22)

3.3 µg BSA was treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 µL, final conc. 100 µM) or TAMRA-N₃ (1 µL, final conc. 100 µM), with either 1 µL DMA or 1 µL 0.4 % APS (w/v) in a total volume of 40 µL PBS buffer at 50 °C, 750 rpm for 2 hours. The labeled BSA were resolved by 10 % SDS-PAGE gel.

7.2.4 In-gel Fluorescence Analysis

Figure 26C-F, 27A, 27D, 31B-C, S1-6, S9-11, S29: After incubation, the reaction mixture was added with 4 times more acetone (usually 160 µL) and placed in the -20 °C freezer for 2 hours, which was further spun down at 4 °C, 13,000 rpm for 10 minutes. The supernatants were then removed and the protein pellets were washed twice with cold methanol, reconstituted within 80 µL 0.2 % SDS in PBS buffer and sonicated with an ultrasonic rod sonicator at 1s-on / 1s-off cycles at 20% intensity for a total time of 10 seconds. The resolution of proteins was made by SDS-PAGE using 10 % (18 % for thioredoxinA and variants) acrylamide gels. Before loading onto the gel, a protein solution (16 µL, 20 µg protein in total) was mixed with 5 × Laemmli buffer (4 µL) [10% (w/v) SDS, 50% (v/v) glycerol, 25% (v/v) β-mercaptoethanol, 0.5% (w/v) bromphenol blue, 315 mM Tris/HCl, pH 6.8] and boiled at 95 °C for 5 minutes. After cooling down, the mixture was loaded onto the well within the stacking layer. BenchMark™ Fluorescent Protein Standard (Invitrogen™, Catalog # LC5928) was used as a maker. Afterwards the gel was scanned and visualized on Amersham Imager 680 (GE Healthcare). The loading control was obtained by Coomassie staining and then destained within 20 % methanol and 10 % acetic acid in water.

Figure S33: After incubation, the reaction mixture was mixed with 5 µL 5 × Laemmli buffer and boiled at 95 °C for 5 minutes. After cooling down, the mixtures in each sample were all loaded into the well within the stacking layer. The resolution of proteins was made by SDS-PAGE using 10 % acrylamide gels (same procedures as described above).

Figure 27C, S14: After incubation, the reaction mixture was mixed with 10 µL 5 × Laemmli buffer and boiled at 95 °C for 5 minutes. After cooling down, the mixtures in each sample were all loaded into the well within the stacking layer. The resolution of proteins was made by SDS-PAGE using 10 % acrylamide gels (same procedures as described above).

7.2.5 Treatments of Two-Components GlyTACs *in cellulo*

When cells were grown to approx. 70 % confluence, the old cell culture medium was removed, followed by washing with PBS buffer once. Then fresh cell culture mediums containing GlcNAc or GalNAc analogues (Ac₄GlcNAz, Ac₄GalNAz, or Ac₃4dGlcNAz) with specific concentrations were supplemented and incubated usually for 24 hours. Afterwards the medium was removed and cells were washed once with PBS buffer. Another portion of fresh cell culture mediums containing DBCO-Thal were added to reach specific concentrations and incubated usually for 24 hours.

7.2.6 MTT Cytotoxicity Assay

General methods for cell seeding and treatments

Cell concentration was measured by mixing 10 μ L cell suspension during subcultivation and 10 μ L Trypan Blue (Invitrogen, Catalog # T10282) with Countess 3 (Invitrogen). Afterwards cells were seeded at a density of 5,000 living cells per well (100 μ L from a solution of 50,000 living cells / mL) in a transparent, flat-bottom 96-well plate. Cells were grown overnight in a humidified atmosphere at 37 °C and 5 % CO₂. On the next day, the medium was exchanged with fresh medium supplemented with either DMSO (max. 1% (v/v)) or other active compounds based on experimental purposes. The cells were incubated for a certain period of time, followed by subsequent incubation with other active compound(s) in newly-exchanged medium or direct addition of MTT solution.

General methods for MTT addition and cell viability measurements

To determine metabolic activity of cells, 20 μ L MTT in sterile PBS solution (5 mg / mL) were added into each well. The resulting mixtures were pipetting up and down gently and incubated for exactly 4 hours. Then the medium was completely removed and the formazan crystals were resuspended in 200 μ L DMSO, and the absorbances at both 570 nm and the reference wavelength 630 nm were determined in an infinite F200 pro plate reader (Tecan). All data points were measured in biological triplicates. The data was normalized with respect to the DMSO control.

Figure 32C-D, 35E-F, S31, and S32A-C: For single compounds

For measuring the cytotoxicity of Ac₄GalNAz, Ac₄GlcNAz, Ac₃4dGlcNAz, DBCO-Thal, and Ac₄GalNAz-thalidomide, the fresh medium was added after overnight seeding to the listed concentrations and incubated for exactly 24 hours. Afterwards MTT solution in PBS was added and cell viability was measured. The highest, non-toxic (at least 80 % cell survival) concentration for each compound was applied for following experiments.

Figure 32E, 33A, 33C-E, and S32D: For continuous treatments

For measuring the combined cytotoxicity of Ac₄GalNAz, Ac₄GlcNAz or Ac₃4dGlcNAz and DBCO-Thal in either HeLa or CHO cells, the fresh medium containing Ac₄GalNAz, Ac₄GlcNAz or Ac₃4dGlcNAz was firstly added after overnight seeding to the listed concentrations and incubated for exactly 24 hours. Then the medium was aspirated and new medium containing DBCO-Thal was added to the listed concentrations and incubated for exactly 24 hours. Afterwards MTT solution in PBS was added and cell viability was measured. The EC₅₀ value of DBCO-Thal was thus confirmed when fixing Ac₄GalNAz or Ac₄GlcNAz concentration, and vice versa. These two concentrations were then applied as “standard conditions” for the following time-dependent experiments, mechanistic studies and cancer cell line screenings.

Figure 33B and 33F: For time-dependent experiments

For measuring half-life time ($t_{1/2}$) of HeLa cells when subjected to combined treatments of Ac₄GalNAz, Ac₄GlcNAz or Ac₃4dGlcNAz and DBCO-Thal, the fresh medium containing Ac₄GalNAz, Ac₄GlcNAz or Ac₃4dGlcNAz was firstly added after overnight seeding and incubated for exactly 24 hours. Then the medium was aspirated and new medium containing DBCO-Thal was added and incubated for the listed time. The concentrations used for both compounds were standard conc. determined from previous experiments. Afterwards MTT solution in PBS was added at each time point (0.5 h, 1 h, 2 h, 4 h, 6 h, 10 h, 24 h, 48 h) and cell viability was measured. The $t_{1/2}$ value was calculated from the plot of cell survival

curves regarding the treatment time of DBCO-Thal.

Figure 35A-B: For OSMI-1 inhibition experiment

For studying OSMI-1 effect on Ac₄GlcNAz (or Ac₄GalNAz) and DBCO-Thal treated HeLa cells, fresh medium containing 2 × listed conc. of OSMI-1 was firstly incubated alone with cells for 3 hours at non-toxic concentrations (up to 2 × 10 μM). Then new medium containing 2 × standard conc. of Ac₄GlcNAz or Ac₄GalNAz was mixed and incubated together with OSMI-1 for 24 hours more. The medium mixture was subsequently aspirated and new medium containing standard conc. of DBCO-Thal was added and incubated for exactly 24 hours. Afterwards MTT solution in PBS was added and cell viability was measured.

Figure 35C-D: For (+/-)-thalidomide / bortezomib effects experiment

For studying (+/-)-thalidomide (or bortezomib) effect on Ac₄GalNAz and DBCO-Thal treated HeLa cells, fresh medium containing standard conc. of Ac₄GalNAz was firstly added after overnight seeding and incubated for exactly 24 hours. Afterwards the medium was aspirated and new medium containing (+/-)-thalidomide (or bortezomib) was added to the listed concentrations and incubated for exactly 1 hour. Then new medium containing 2 × listed conc. of (+/-)-thalidomide (or bortezomib) was added, followed by mixing with 2 × standard conc. of DBCO-Thal. The resulting medium mixture was incubated with cells for 8.5 hours. Afterwards MTT solution in PBS was added and cell viability was measured.

Figure 36H: For cancer cell line screening

All cell concentrations were measured by mixing 10 μL cell suspension during subcultivation and 10 μL Trypan Blue (invitrogen, Catalog # T10282) with Countess 3 (invitrogen). Afterwards cells were seeded at a density of 5,000 living cells per well (100 μL from a solution of 50,000 living cells / mL) in a transparent, flat-bottom 96-well plate. Cells were grown overnight in a humidified atmosphere at 37 °C and 5 % CO₂. On next two days, fresh medium containing standard conc. of Ac₄GalNAz and then DBCO-Thal was treated with cells sequentially (24 hours incubation for each). Control experiments included incubation with medium containing DMSO-DMSO, Ac₄GalNAz-DMSO, and DMSO-DBCO-Thal.

For SK-MEL-28 cell line, the seeding was altered to 10,000 living cells per well.

For K562 cell line, the suspension cells were collected after seeding and 24-hours treatment of Ac₄GalNAz, and then centrifuged at 300 rcf to remove supernatants. Then cells were treated with standard conc. of DBCO-Thal. Afterwards MTT solution in PBS was added and cell viability was measured by centrifuging at 300 rcf to remove supernatants, resuspending in DMSO and then transferring to the 96-well plate for measurement.

7.2.7 Western Blot

Figure 34D-F: NUP62 visualization

After measuring protein concentrations, 20 μg cell lysates were taken for each sample and topped up to a total volume of 16 μL. 4 μL 5 × Laemmli buffer (same component with the buffer used in in-gel fluorescence analysis) was added and the resulting mixture was incubated at 95 °C for 5 minutes. Afterwards proteins were separated using 10 % SDS-PAGE gel (each well contained all 20 μg cell lysates). Instead of a fluorescence maker, 3 μL Color Prestained Protein Standard (10 - 250 kDa, BioLabs, # P7719S) was added into the first well as the marker. The separated proteins were then transferred onto a PVDF membrane using a blotting-buffer-moistened (48 mM Tris, 39 mM glycine, 0.0375 % (m/v) SDS,

20 % (v/v) methanol) blotting sandwich, which was composed of one extra thick blot paper, the PVDF transfer membrane (pre-incubation for 5 min in methanol), the SDS-PAGE gel and again one extra thick blot paper. The protein transfer was subsequently carried out for 30 minutes at 25 V using a Semi-Dry Blotter (Bio-Rad). In order to block non-specific binding sites, the PVDF membrane was incubated for 60 minutes in blocking solution (0.5 g non-fat milk powder in 10 mL PBST (PBS +0.5 % Tween-20)) at room temperature. Afterwards 10 μ L primary antibody against NUP62 (BD Biosciences, # 610498) was added to the blocking solution and the mixture was incubated overnight at 4 °C. The PVDF membrane was washed 3 times for 15 minutes with PBST before 1 μ L HRP-linked secondary antibody (Cell Signaling Technology, # 7076S) in 10 mL blocking solution (0.5 g non-fat milk powder in 10 mL PBST) was added. After 1-hour incubation at room temperature, the PVDF membrane was washed again 3 times for 15 minutes with PBST. Then, 400 μ L ECL Substrate and 400 μ L peroxide solution were mixed on-site and evenly distributed onto the PVDF membrane for staining. Finally, chemiluminescence signals from the membrane were captured by developing machine Amersham Imager 680 (GE Healthcare).

Figure 34A-C: Ubiquitinated protein visualization

The blocking solution was prepared by mixing 1 mL ROTI[®]Block (10 \times conc., from Carl Roth GmbH) with 9 mL water. The primary antibody (Ubiquitin Polyclonal Antibody) used was purchased from Thermo Fischer (# PA1-187) and the HRP-linked secondary antibody was bought from Thermo Fischer (# 31460), which was diluted at a 1:20,000 ratio (i.e., 1 μ L secondary antibody in 20 mL blocking solution). All other steps were identical with western blot for NUP62 visualization.

7.2.8 MS Sample Preparations

Cell lysates clean-up with SP3

The procedure was performed in a 96-well plate. Each conditions contained 3 biological replicates. 20 μ L carboxylate-coated magnetic beads (1:1 mixture of hydrophilic and hydrophobic beads) were washed manually thrice with 100 μ L MS-scale water and the last washing solution was kept within the 96-well plate. On the other hand, 20 μ g of cell lysates were placed within the 96-well plate and diluted to 50 μ L with PBS buffer. The plate was then transferred to Hamilton MicroLab PREP robot, and the last washing solution was aspirated, followed by addition of 50 μ L cell lysates onto the beads. The mixture was shaken for 1 minute at 850 rpm, room temperature. Afterwards 60 μ L absolute ethanol was added to each sample and incubated for 5 minutes at 850 rpm, room temperature. The supernatants were removed and the protein-bound beads were washed sequentially with 100 μ L 80 % EtOH in water for three times and 100 μ L acetonitrile once. An incubation for 1 minute at 850 rpm, room temperature after each addition was carried out. After removal of the last washing solution, the beads were supplemented with 100 μ L 100 mM ammonium bicarbonate buffer and then digested overnight with 1 μ L sequencing-grade trypsin (0.5 mg / mL, Promega) at 37 °C, 600 rpm. The caps for each well were all carefully encapsulated to prevent the evaporation of aqueous medium. On the next day, the peptide mixtures were all transferred to the new Eppendorf tubes and the beads were washed with 50 μ L and 30 μ L 1% FA in water. The beads with washing solutions were incubated at 40 °C, 850 rpm for 5 minutes. The washing solutions were collected all together with digested peptide mixtures, and placed within a MS-vial. For LC-MS/MS measurement, 5 μ L solutions were injected for each sample.

Cell lysates clean-up and protein enrichment with SP2E

After incubation, the mixtures were added with 200 μL 8M urea to a total volume of 400 μL . 100 μL of a 1:1 mixture of hydrophobic and hydrophilic carboxylate-coated magnetic beads (Cytiva) was washed thrice with 500 μL MS-water and the reaction mixture was transferred onto the pre-washed beads followed by the addition of 600 μL absolute ethanol, mixing and incubation at room temperature, 950 rpm for 5 minutes. The beads were washed thrice with 500 μL 80 % EtOH in MS-water. Further, proteins were eluted twice with 500 μL 0.2 % SDS in PBS at room temperature, 950 rpm for 5 minutes from the carboxylate-coated beads and transferred onto 50 μL pre-equilibrated streptavidin-coated magnetic beads (New England Biolabs) by three times washing with 500 μL 0.2 % SDS in PBS. For streptavidin-biotin complex formation, beads were incubated at room temperature, 950 rpm for 1 hour with eluted proteins. Subsequently, beads were washed three times with 500 μL 0.1 % NP-40 in PBS, two times with 500 μL 6 M urea (in MS-water) and two times with 500 μL MS-water by vortexing and short-spin between each wash step. After the last wash step, 80 μL ammonium bicarbonate buffer (125 mM in MS-water), 10 μL TCEP (100 mM in MS-water) and 10 μL chloracetamide (400 mM in MS-water) was added and incubated at 95 $^{\circ}\text{C}$ for 5 min. Then, proteins were digested overnight with 1.5 μL sequencing-grade trypsin (0.5 mg / mL, Promega) at 37 $^{\circ}\text{C}$, 600 rpm.

On the next day, the supernatants were transferred into new Eppendorf tubes, and beads were washed thrice with 100 μL ammonium bicarbonate buffer (100 mM in MS-water) to collect the peptides. The combined fractions were acidified with 2 μL MS-grade formic acid (FA) and peptides were desalted with 50 mg of SepPak C18 cartridges on a vacuum manifold. First, cartridges were equilibrated with 1 mL MS-grade ACN, 1 mL elution buffer (80 % ACN with 0.5 % FA in MS-water) and 3 mL washing buffer (0.5 % FA in MS-water). Second, combined peptide fractions were loaded onto the cartridges and washed with 3 mL of wash buffer. Finally, peptides were eluted twice with 250 μL elution buffer. Desalted peptide eluates were vacuum-dried with a SpeedVac at 35 $^{\circ}\text{C}$, reconstituted in 30 μL 1 % FA in MS-water (vortex and 15 minutes sonication bath), transferred to MS vials and subjected to LC-MS/MS analysis.

7.2.9 Mass Spectrometry Measurements

Direct injection - top-down analysis with high-resolution MS1 and then MS2 fragmentation

High-resolution MS measurements were performed on an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher Scientific) coupled to an UltiMate 3000 Nano-HPLC (Thermo Fisher Scientific) via an EASY-Spray source (Thermo Fisher Scientific) and high field asymmetric waveform ion mobility spectrometry (FAIMS) interface (Thermo Fisher Scientific).

The fully lyophilized reaction mixtures from oxidized or S-arylated glutathione, KEHSPC, RWTPCD, and Ac₄GalNAz-thalidomide conjugate were dissolved within 1 mL 0.5 % FA in acetonitrile / water mixture (1:1, v/v) to a final concentration at around 10 μM , respectively. The samples were then collected by a syringe (500 μL , Hamilton) and loaded onto the pump (chemyx). To start, the flow rate for injection was set at 10 $\mu\text{L}/\text{min}$; the resolution of Orbitrap was selected between 30,000 to 240,000; the Automatic Gain Control (AGC) was 100 %, while the maximal injection time was 100 ms and the microscan was varied between 1 - 5 depending on the measurement situations. The mass window was selected according to the compounds of interest. When the signal of compounds of interest appears, the flow rate was immediately decreased to 4 $\mu\text{L}/\text{min}$. The compensation voltage (CV) was scanned and chosen to achieve the highest

intensity. The MS2 fragmentation was done by inputting the mass of target precursor ions from MS1 and adjusting the HCD for a suitable fragmentation pattern. The measurements and fragmentation were recorded over 5 microscans.

HPLC separation - bottom-up analysis with high-resolution MS1 and then MS2 fragmentation

MS measurements were performed on Orbitrap Eclipse Tribrid Mass Spectrometer coupled to an UltiMate 3000 Nano-HPLC via a Nanospray Flex and FAIMS interface (all from Thermo Fisher Scientific). First, peptides were loaded on an Acclaim PepMap 100 μ -precolumn cartridge (5 μ m, 100 Å; 300 μ M ID x 5 mm, Thermo Fisher Scientific). Then, peptides were separated at 40 °C on a PicoTip emitter (noncoated, 15 cm, 75 μ m ID, 8 μ m tip, New Objective) that was in-house packed with ReprosilPur 120 C18-AQ material (1.9 μ m, 150 Å, Dr. A. Maisch GmbH). The LC buffers consisted of MS-grade water (A) and acetonitrile (B) both supplemented with 0.1% FA. The gradient was run from 4-35.2% B during a 60 min method (0-5 min 4%, 5-6 min to 7%, 7-36 min to 24.8%, 37-41 min to 35.2%, 42-46 min 80%, 47-60 min 4%) at a flow rate of 300 nL/min.

7.2.10 Data Acquisition and Calculation

Data-dependent acquisition (DDA)

FAIMS was performed with two alternating CVs including -50 V and -70 V. The mass spectrometer was operated in dd-MS2 mode with following settings: Polarity: positive; MS1 Orbitrap resolution: 240k; MS1 AGC target: standard; MS1 maximum injection time: 50 ms; MS1 scan range: m/z 375- 1500; RF Lens: 30%; MS2 Orbitrap resolution: 15k; MS2 AGC target: standard; MS2 maximum injection time: 35; HCD collision energy: 30%; RF Lens: 30%; MS2 cycle time: 1.7 s; intensity threshold: 1.0e4 counts; included charge states: 2-6; dynamic exclusion: 60 s.

Data-independent acquisition (DIA)

FAIMS was performed with one CV at -45 V. One DIA cycle comprised one MS1 scan followed by 30 MS2 scans. The mass spectrometer was operated in DIA mode with following settings: Polarity: positive; MS1 Orbitrap resolution: 60k; MS1 AGC target: standard; MS1 maximum injection time: 50 ms; MS1 scan range: m/z 200-1800; RF Lens: 30 %; Precursor Mass Range: m/z 500-740; isolation window: m/z 4; window overlap: m/z 2; MS2 Orbitrap resolution: 30 k; MS2 AGC target: 200 %; MS2 maximum injection time: auto; HCD collision energy: 35 %; RF Lens: 30 %; MS2 scan range: auto.

MaxQuant (closed search)

MS *.raw files were split by FAIMS-MzXML-generator and subsequently analyzed with MaxQuant software 2.0.3.1. Peptides were searched against the *in silico* digested Uniprot database for Homo Sapiens. FDR was determined by using a decoy database and set to 1 % as thresholds. Oxidation (M), acetylation (N-term), and carbamidomethylation (C) were set as variable modifications.

Computation of measured mass spectra using DIA-NN

First, measured *.raw files were converted to *.mzML format with "MSConvert" from the "ProteoWizard" software package (<http://www.proteowizard.org/download.html>) with following settings: "peakPicking" filter with "vendor msLevel = 1"; "Demultiplex" filter with parameters "Overlap Only" and "mass error" set to 10 ppm. Subsequently, they were analyzed with DIA-NN 1.8.1 and peptides were searched against Uniprot database for Homo Sapiens with included contaminants and decoys. The DIA-NN settings were as follows: FASTA digest for library-free search/library generation: enabled; Deep learning-based spectra,

RTs and IMs prediction: enabled; missed cleavages: 2; max number of variable modifications: 3; modifications: *N*-term M excision, oxidation (M) and *N*-term acetylation; precursor charge range: 2-6; precursor range: m/z 500-740; fragment ion range: m/z 200-1800; precursor FDR level: 1 %; match between runs (MBR): enabled; library generation: smart profiling; quantification strategy: Robust LC; mass accuracy: 0; scan windows: 0.

7.2.11 Data Visualization

Figure 36B, S34A, S35A, S36A: Heatmap

For statistical analysis, the generated “report.gg_matrix.tsv” table was loaded in Perseus 1.6.14.0. Then, quantified values were assigned with conditions during GlyTAC treatment. Each biological replicate had the same name to constitute one group. Samples not treated with any of the active components or treated with only one of the two components were always applied as controls. Subsequently, the groups were filtered for at least two valid values out of three columns in each group. The matrix was then processed through multiple-sample test with ANOVA-significant. The rest values were log₂-transformed and missing values were replaced by a normal distribution. The rows in matrix were then filtered with positive ANOVA-significant mode and further proceeded with Z-score. The hierarchical clustering were generated with Euclidean distance and average linkage for both row and column trees, while the row clusters were defined to be 8 to classify groups of interests. The first two groups containing most proteins were therefore mainly down-regulated and up-regulated proteins after sequential GlyTAC treatment. The gene and protein lists for these groups from different GlyTAC treatment conditions are recorded separately for Venn diagram and KEGG pathway analysis.

Figure 25, 36D, S34B, S35B, S36B: Volcano plots

AMPylation project (Figure 25): for statistical analysis, the generated “report.gg_matrix.tsv” table was loaded in Perseus 1.6.14.0. Then, quantified values were log₂-transformed and assigned with conditions for probe **1 / 2 / DMSO (Figure 23)** treatment. Each biological replicate had the same name to constitute one group. Subsequently, the groups were filtered for at least three valid values out of four columns in at least one group. Further, missing values were replaced from a normal distribution. -log₁₀(p-values) were obtained by a two-sided one sample Student’s t-test over replicates with the initial significance level $\alpha = 0.05$ adjustment by the multiple testing correction method of Benjamini and Hochberg (FDR = 0.05) using the volcano plot function. Finally, the volcano plot values from Perseus were transferred to GraphPad PRISM 9 and visualized properly.

GlyTAC project (Figure 36D, S34B, S35B, S36B): for statistical analysis, the generated “report.gg_matrix.tsv” table was loaded in Perseus 1.6.14.0.²²⁰ Then, quantified values were assigned with conditions during GlyTAC treatment. Each biological replicate had the same name to constitute one group. Subsequently, the groups were filtered for at least two valid values out of three columns in each group. The rest values were log₂-transformed and missing values were replaced by a normal distribution. -log₁₀(p-values) were obtained by a two-sided one sample Student’s t-test over replicates with the initial significance level $\alpha = 0.05$ adjustment by the multiple testing correction method of Benjamini and Hochberg (FDR = 0.05) using the volcano plot function. Finally, the volcano plot values from Perseus were transferred to GraphPad PRISM 9 and visualized properly.

Figure 36C, 36E: Venn diagrams

For Figure 36C, the down-regulated proteins' lists after hierarchical clustering (5 μ M and 20 μ M DBCO-Thal treatment respectively) were overlapped through the Venn diagram online tool <https://bioinfogp.cnb.csic.es/tools/venny/> to determine proteins numbers of each group. The results were then visualized by Affinity Designer 2.

For Figure 36E, the down-regulated proteins' lists were overlapped with downloaded human O-GlcNAcylation protein database <https://www.oglcnac.mcw.edu/download/> (generated on 16.07.2024) also using Venn diagram online tool <https://bioinfogp.cnb.csic.es/tools/venny/>. The results were then visualized by Affinity Designer 2.

Figure 36F, 36G, S35C: KEGG pathway analysis and GO term analysis

The up- and down-regulated proteins after sequential GlyTAC treatment were transferred to the online tool ShinyGO 0.80 <http://bioinformatics.sdstate.edu/go/> for KEGG pathway analysis and GO term analysis (Biological Function). The parameters were set as default and the results were visualized directly from the website.

8. Appendix

8.1 Supplementary Figures

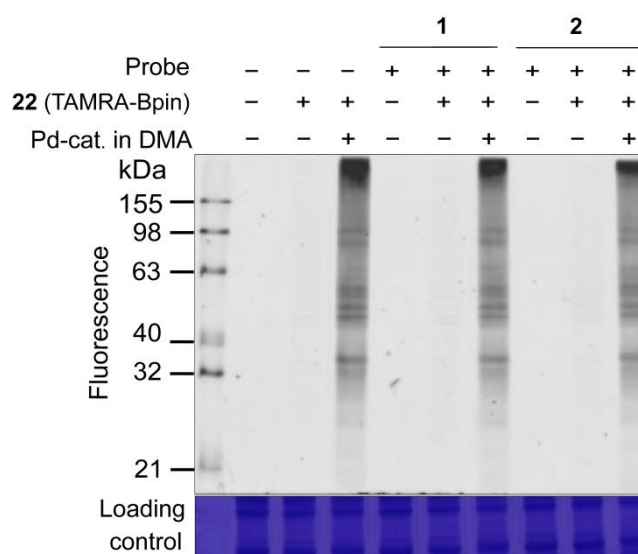


Figure S1 In-gel fluorescence analysis of iodinated-probe dependent labelling for the envisaged Suzuki cross-coupling within HeLa cell lysates. The cell lysates were treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M) in the presence of Pd catalyst in DMA (1 μ L, final conc. 200 μ M) in a total volume of 40 μ L PBS buffer at 50 $^{\circ}$ C, 750 rpm for 2 hours. The labeled proteins were resolved by 10 % SDS-PAGE gel.

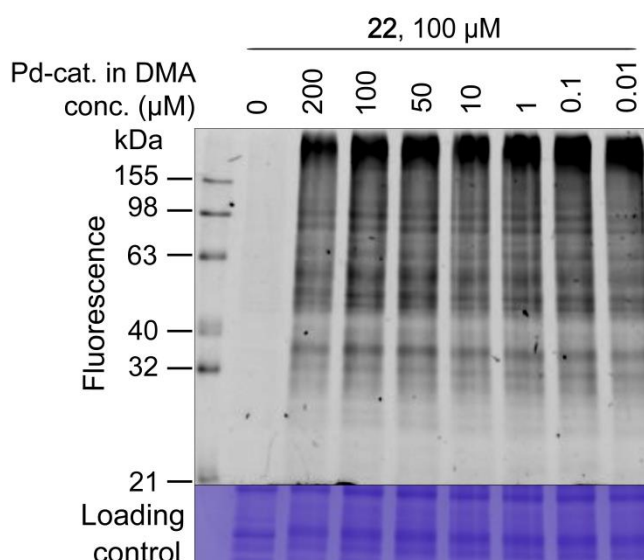


Figure S2 In-gel fluorescence analysis of Pd-cat. concentration-dependent labelling within HeLa cell lysates. The cell lysates were treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M) in the presence of different Pd catalyst concentration in DMA (volume of DMA was kept at 1 μ L in all experiments) in a total volume of 40 μ L PBS buffer at 50 $^{\circ}$ C, 750 rpm for 2 hours. The labeled proteins were resolved by 10 % SDS-PAGE gel.

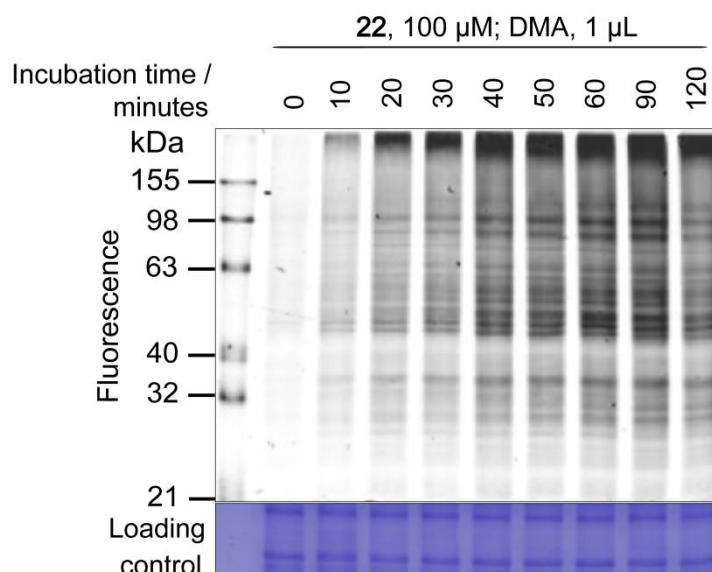


Figure S5 In-gel fluorescence analysis of DMA labelling dynamics within HeLa cell lysates. The cell lysates were treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M) and DMA (1 μ L) in a total volume of 40 μ L PBS buffer at 50 $^{\circ}$ C, 750 rpm for 0, 10, 20, 30, 40, 50, 60, 90 and 120 minutes. The labeled proteins were resolved by 10 % SDS-PAGE gel.

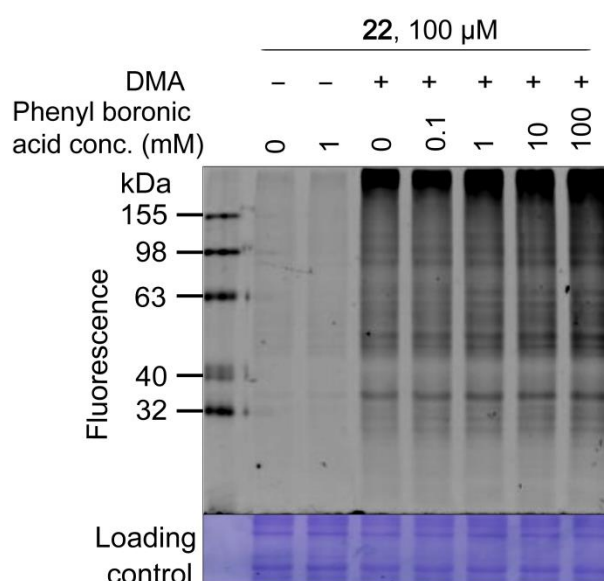


Figure S6 In-gel fluorescence analysis of de-labelling by phenylboronic acid within HeLa cell lysates. The cell lysates were firstly treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M) and DMA (1 μ L) in a total volume of 40 μ L PBS buffer at 50 $^{\circ}$ C for 2 hours. Afterwards phenylboronic acid stock solutions at different concentration (added volume was 1 μ L for all experiments) were added to achieve final concentration at 0.1, 1, 10 and 100 mM. The resulting mixture was further incubated at 50 $^{\circ}$ C, 750 rpm for 2 hours. The labeled proteins were resolved by 10 % SDS-PAGE gel.

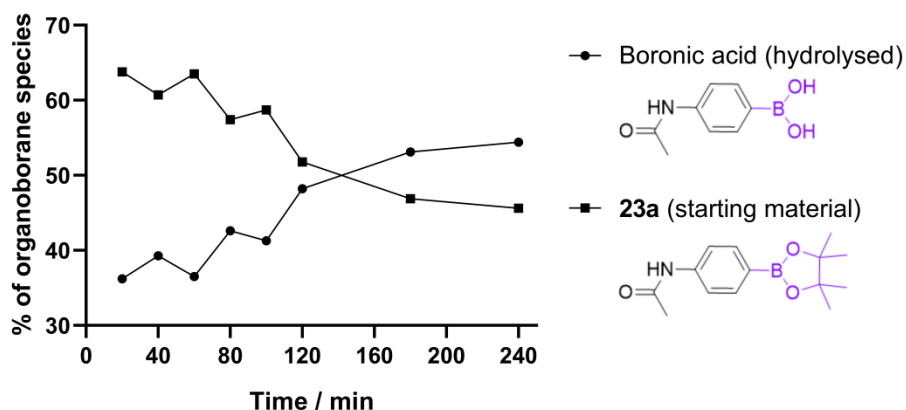


Figure S7 Half-life time measurement of aryl pinacol boronate **23a** in the presence of APS in water. 12.5 mM **23a** and 0.01 % APS (final concentrations) are co-incubated within 40 μ L water at 50 $^{\circ}$ C for 2 hours, and then subjected to lyophilization to remove water. The residuals were reconstituted within 40 μ L 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8 μ L), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 μ L/minute. The conversion was decided based on integral of total ion current (TIC) of each species.

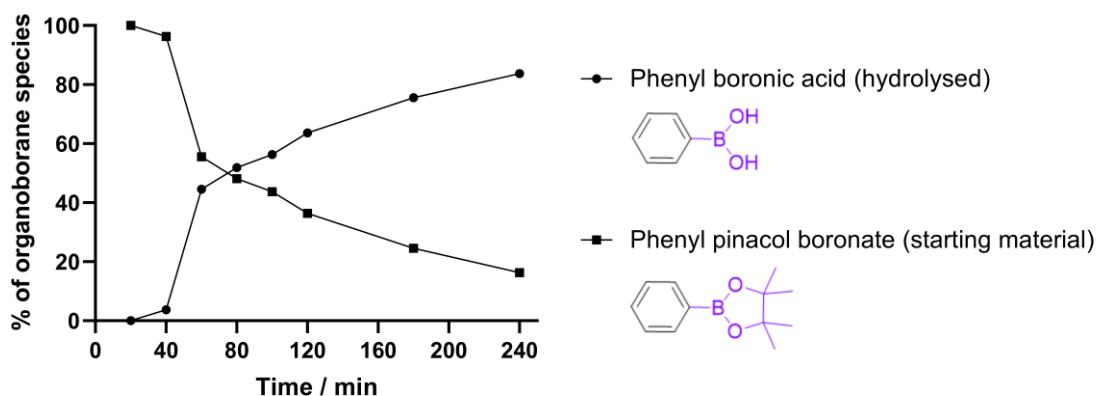


Figure S8 Half-life time measurement of phenyl pinacol boronate in the presence of APS in water. 12.5 mM phenyl pinacol boronate and 0.01 % APS (final concentrations) are co-incubated within 40 μ L water at 50 $^{\circ}$ C for 2 hours, and then subjected to lyophilization to remove water. The residuals were reconstituted within 40 μ L 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8 μ L), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 μ L/minute. The conversion was decided based on integral of total ion current (TIC) of each species.

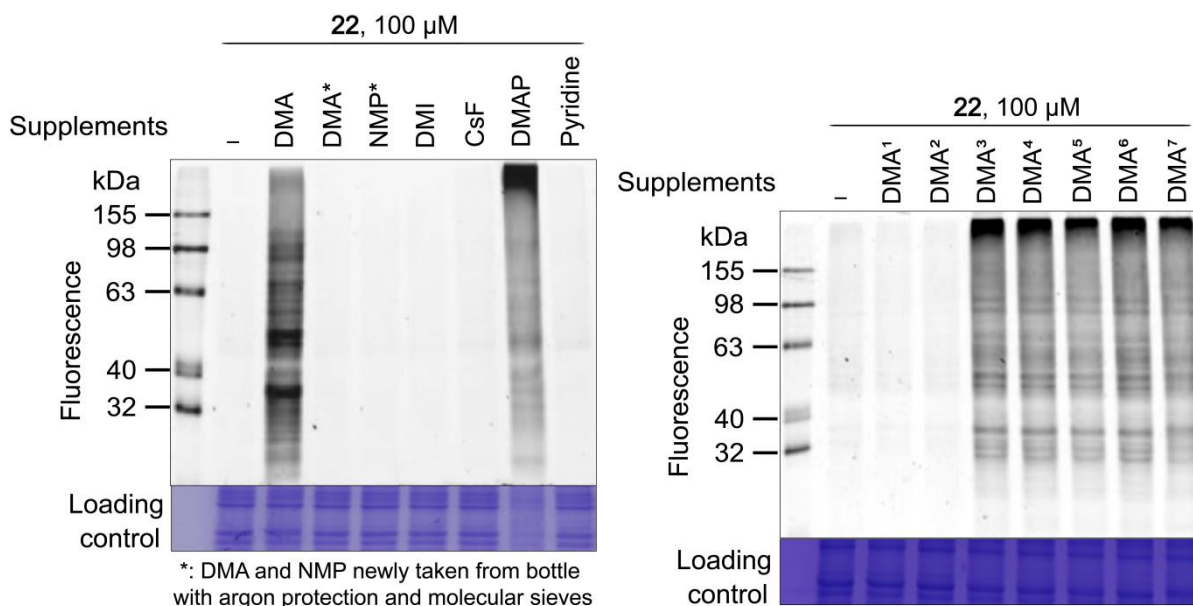


Figure S9 Left: In-gel fluorescence analysis for testing other possible “labelling triggers” within HeLa cell lysates. The cell lysates were treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M) and 1 μ L DMA / newly taken 1 μ L DMA / *N*-methyl-2-pyrrolidone (NMP, 1.06 μ L) / 1,3-dimethyl-2-imidazo-lidinone (DMI, 1.19 μ L) / 1 μ L 10.7 M CsF stock solution / 1.3 mg DMAP / 0.886 μ L pyridine in a total volume of 40 μ L PBS buffer at 50 $^{\circ}$ C, 750 rpm for 2 hours. Right: In-gel fluorescence analysis for protein labelling with different DMA within HeLa cell lysates. The cell lysates were treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M) and DMA (1 μ L, from different sources with different treatments) in a total volume of 40 μ L PBS buffer at 50 $^{\circ}$ C, 750 rpm for 2 hours. DMA 1 and 2 were freshly taken from bottles (1 from Thermo Fisher Scientific, 2 from TCI Deutschland GmbH). DMA 3 - 6 were air-flushed for 2 h, 4 h, 6 h and 1 day, which were further air-exposed for 1 month. DMA 7 was air-exposed for 6 months and used routinely for all other experiments. The labeled proteins were resolved by 10 % SDS-PAGE gel. These results together with **Figure 27A** may reveal the crucial role of oxygen for initiating the reaction.

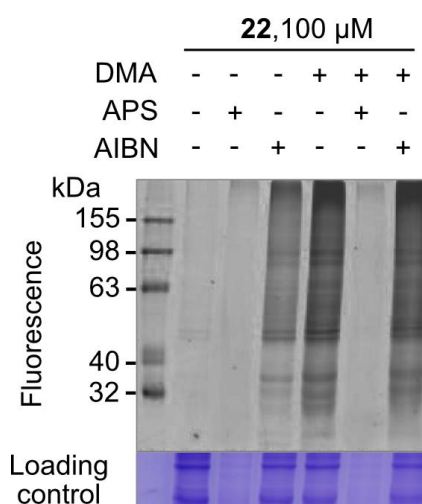


Figure S10 In-gel fluorescence analysis for testing other possible radical initiators within HeLa cell lysates. The cell lysates were treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M), supplemented by 1 μ L APS (10 % m/v in water) or 1 μ L AIBN (5 % m/v in methanol) in the presence or absence of DMA (1 μ L) in a total volume of 40 μ L PBS buffer at 50 $^{\circ}$ C for 2 hours. The labeled proteins were resolved by 10 % SDS-PAGE gel.

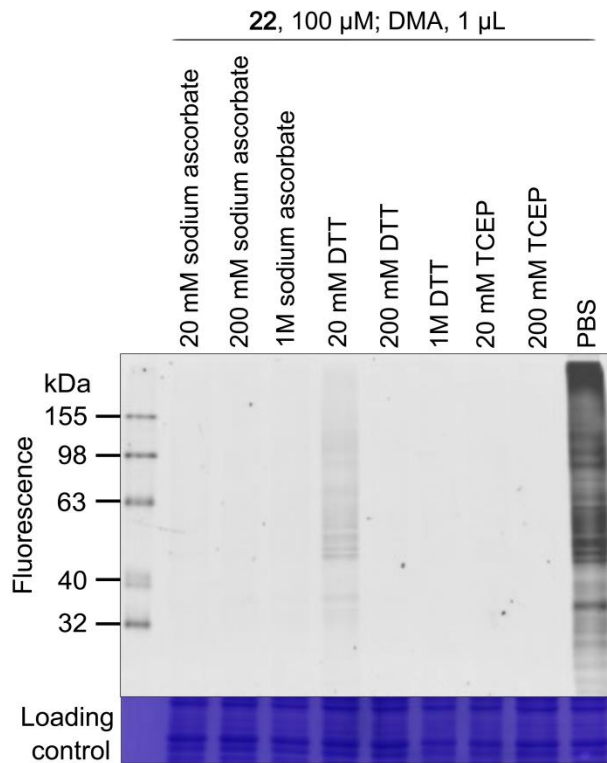


Figure S11 In-gel fluorescence analysis of protein labelling in the presence of reductants within HeLa cell lysates. The cell lysates were firstly added with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M) and DMA (1 μ L) to a total volume of 20 μ L by PBS buffer. PBS buffers with reductants (20 μ L of 2 \times stock) were then added into each sample to the indicated final concentrations. The resulting mixture (40 μ L) was incubated at 50 $^{\circ}$ C, 750 rpm for 2 hours. The labeled proteins were resolved by 10 % SDS-PAGE gel.

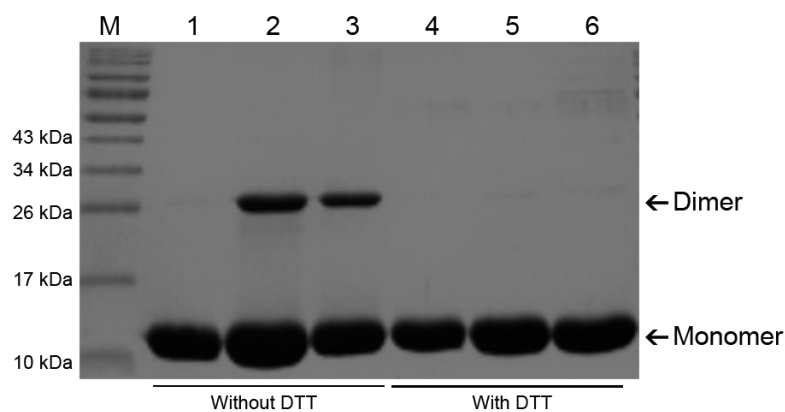
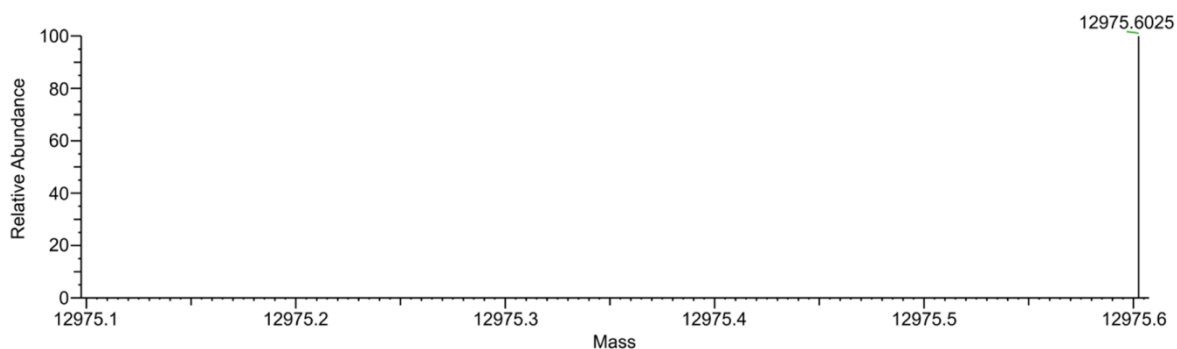
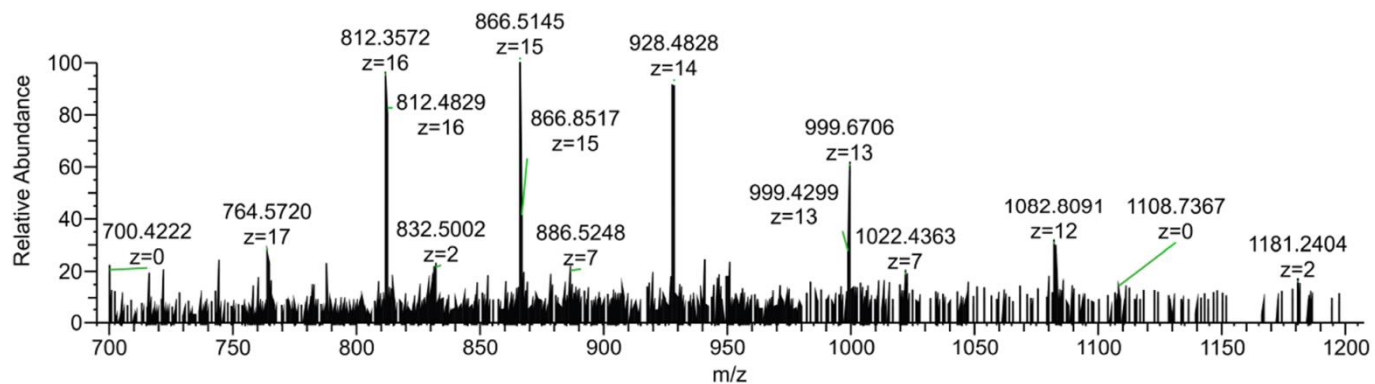
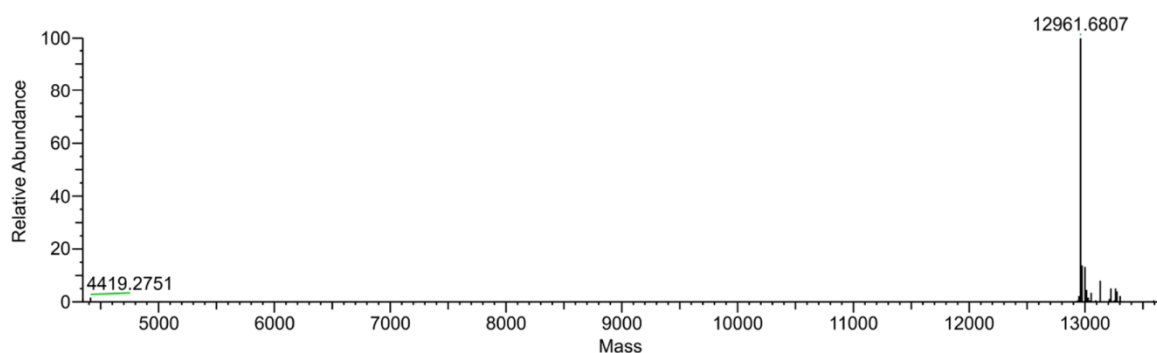
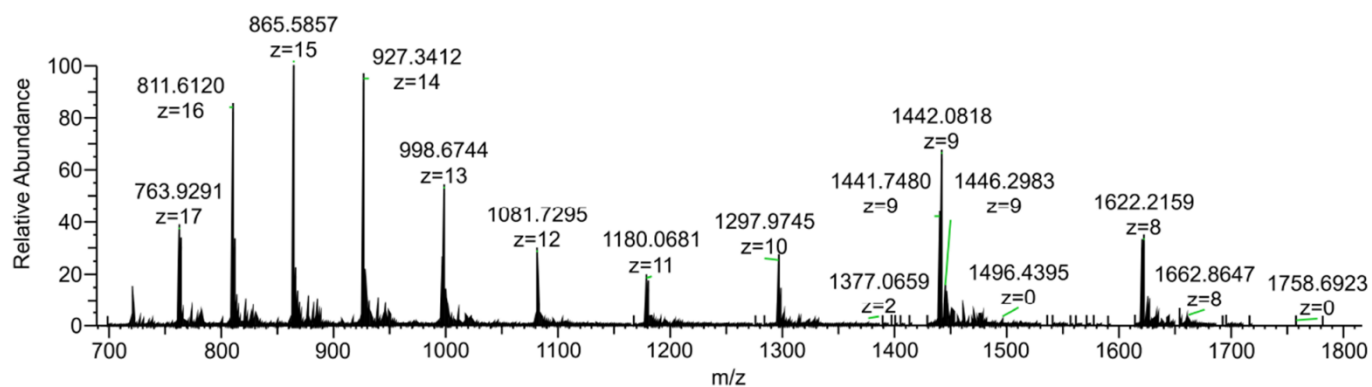


Figure S12 SDS-PAGE gel analysis. M: molecular-weight marker; lanes 1 + 4: His₆-TrxA, lanes 2 + 5: His₆-TrxA^{C33A}, lanes 3 + 6: His₆-TrxA^{C36A}, without (lanes 1 - 3) or with dithiothreitol (DTT) (lanes 4 - 6) in the loading buffer. Lanes 4 - 6 were incubated with 5 mM DTT at room temperature for 10 min, breaking intermolecular disulfide bonds.

ThioredoxinA variants	Molecular formula	Monoisotopic mass (calculated)	Observed mass (deconvoluted)
Wild type	$C_{585}H_{917}N_{155}O_{169}S_5$ (raw)	12977.6529	/
	$C_{585}H_{915}N_{155}O_{169}S_5$ (if disulfide was formed)	12975.6373	12975.6025



ThioredoxinA variants	Molecular formula	Monoisotopic mass (calculated)	Observed mass (deconvoluted)
C33S	$C_{585}H_{917}N_{155}O_{170}S_4$	12961.6758	12961.6807



ThioredoxinA variants	Molecular formula	Monoisotopic mass (calculated)	Observed mass (deconvoluted)
C36S	C ₅₈₅ H ₉₁₇ N ₁₅₅ O ₁₇₀ S ₄	12961.6758	12961.6789

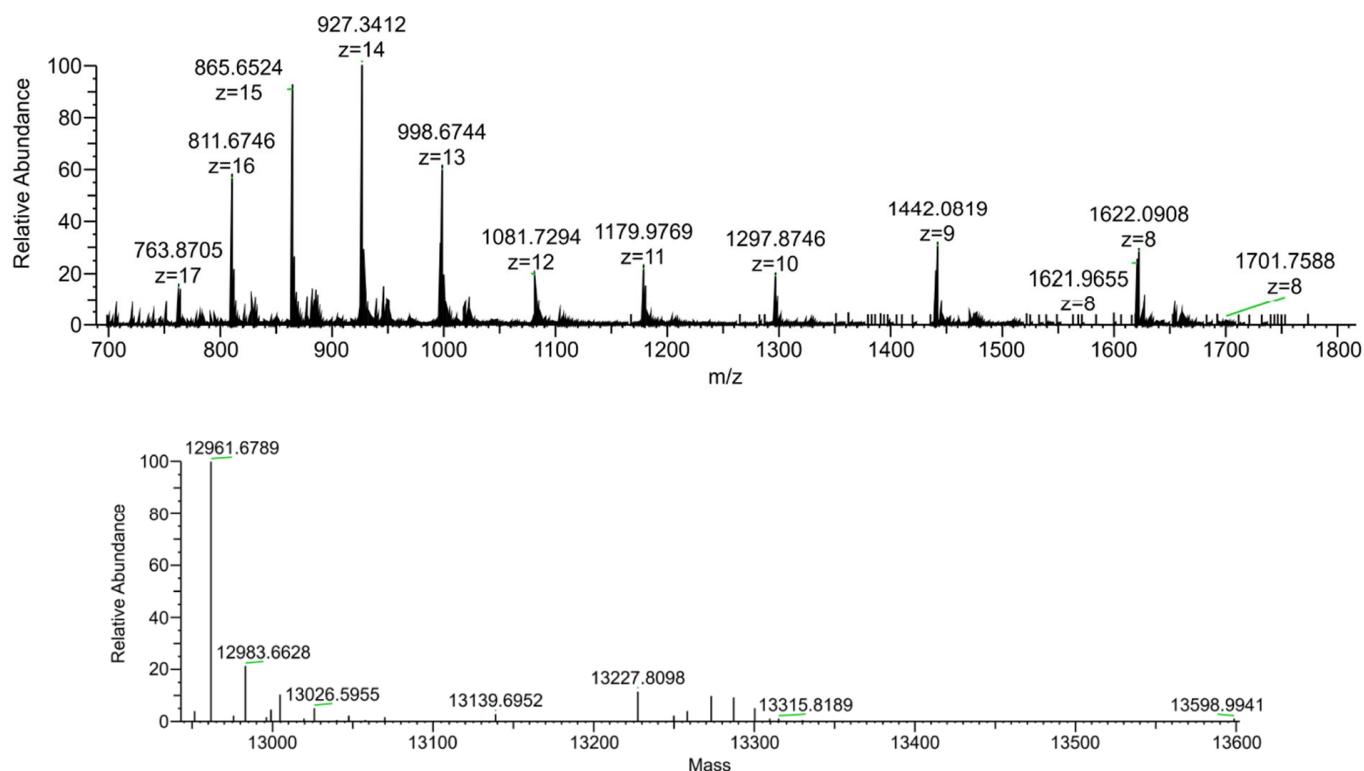


Figure S13 High-resolution MS analysis of ThioredoxinA variants (wild type, C33S, and C36S).

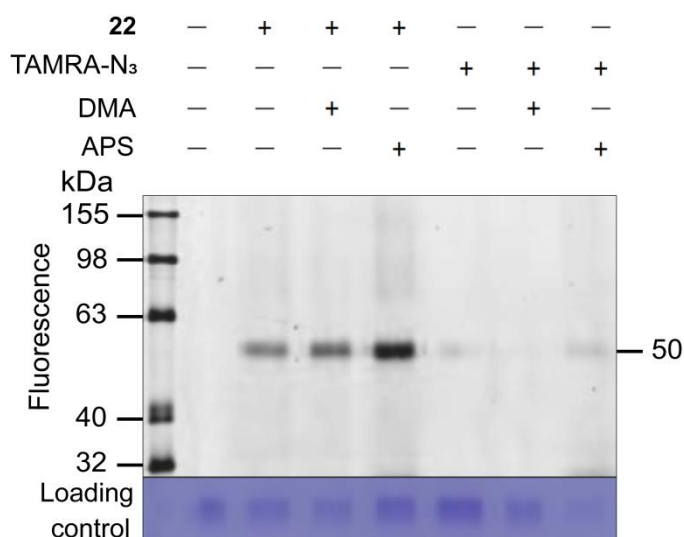


Figure S14 In-gel fluorescence analysis for biotin polyclonal antibody reactivity towards **22** or TAMRA-azide (TAMRA-N₃). 2 µg biotin polyclonal antibody was treated with fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 µL, final conc. 100 µM) or TAMRA-N₃ (1 µL, final conc. 100 µM), with either 1 µL DMA or 1 µL 0.4 % APS (w/v) in a total volume of 40 µL PBS buffer at 50 °C, 750 rpm for 2 hours. The labeled biotin polyclonal antibody (heavy chains) were resolved by 10 % SDS-PAGE gel.

GSSG, calculated mass: 613.1592, $z = 1$; 307.0833, $z = 2$.

Observed mass: 613.1591, $z = 1$; 307.0833, $z = 2$.

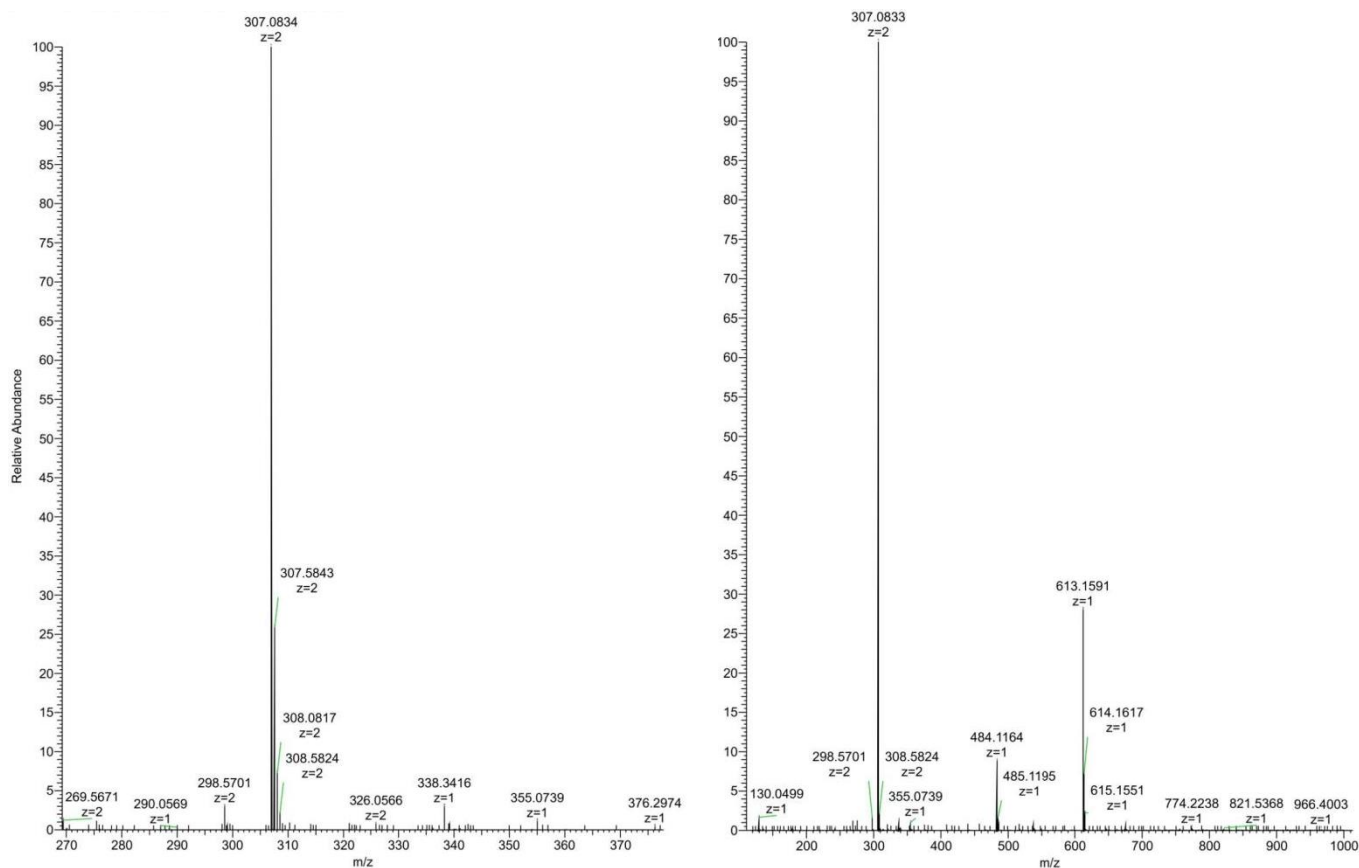


Figure S15 High-resolution MS analysis of oxidized forms of glutathione (GSSG).

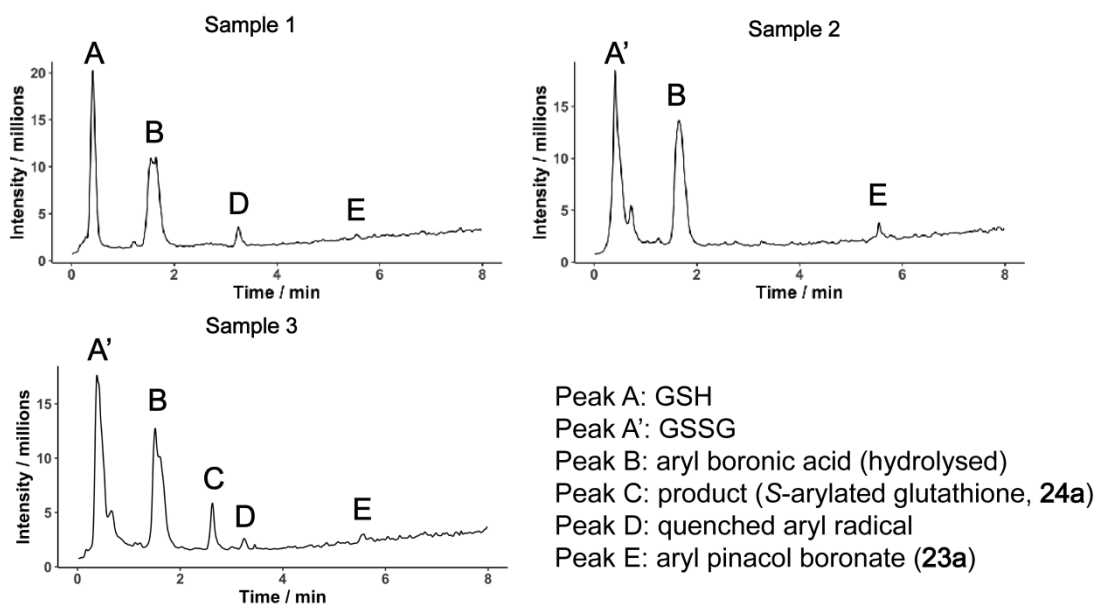


Figure S16 LC-MS chromatogram of GSH / GSSG reactivity in presence of APS. Sample 1: 2.5 mM GSH with 12.5 mM **23a** and 0.01 % APS in 40 μ L water. Sample 2: 2.5 mM GSSG with 12.5 mM **23a** in 40 μ L water. Sample 3: 2.5 mM GSSG with 12.5 mM **23a** and 0.01 % APS in 40 μ L water. These samples were treated at 50 $^{\circ}$ C for 2 hours, and then subjected to lyophilization to remove water. The residual white powders were reconstituted within 40 μ L 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8 μ L), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 μ L/minute.

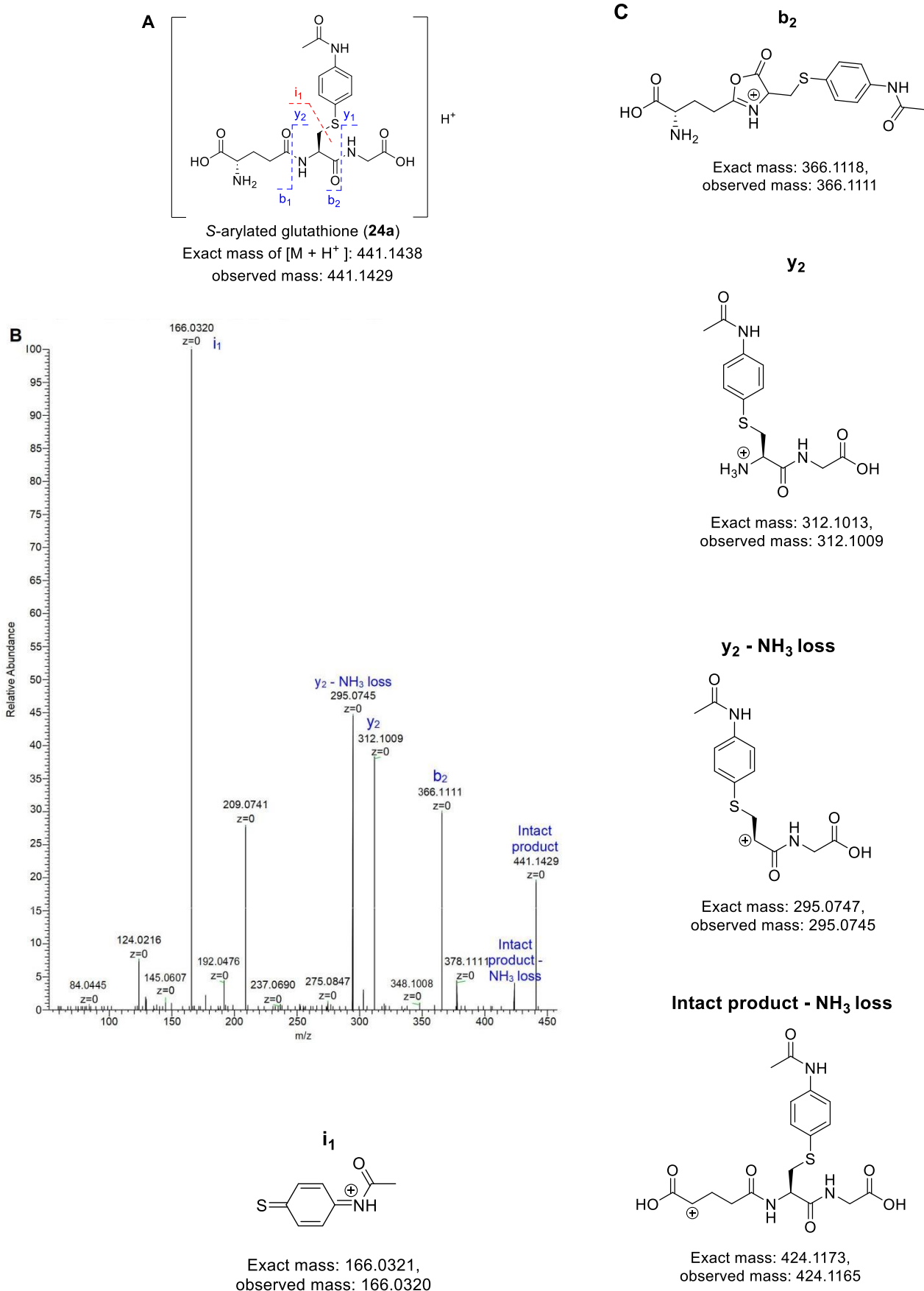


Figure S17 **A** S-arylated glutathione (**24a**) as an intact product with peptide b, y and i fragmentation. **B** High resolution MS2 spectrum of selected MS1 precursor ion ($M = 441$) with HCD fragmentation. NL: $1.28E7$, CV: -20 V, HCD: 20. **C** Structure analysis and assignment of MS2 peaks from S-arylated glutathione (**24a**).

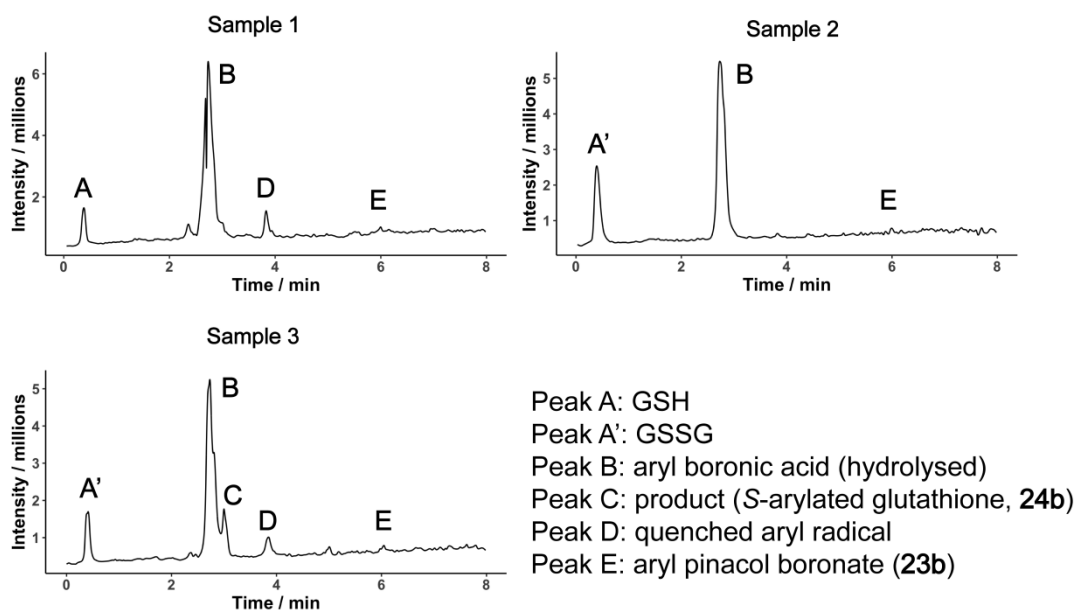


Figure S18 LC-MS chromatogram of GSH / GSSG reactivity in presence of APS. Sample 1: 1.25 mM GSH with 10 mM **23b** and 0.01 % APS in 40 μ L water. Sample 2: 1.25 mM GSSG with 10 mM **23b** in 40 μ L water. Sample 3: 1.25 mM GSSG with 10 mM **23b** and 0.01 % APS in 40 μ L water. These samples were treated at 80 $^{\circ}$ C for 2 hours, and then subjected to lyophilization to remove water. The residual white powders were reconstituted within 40 μ L 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 5 μ L), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 μ L/minute.

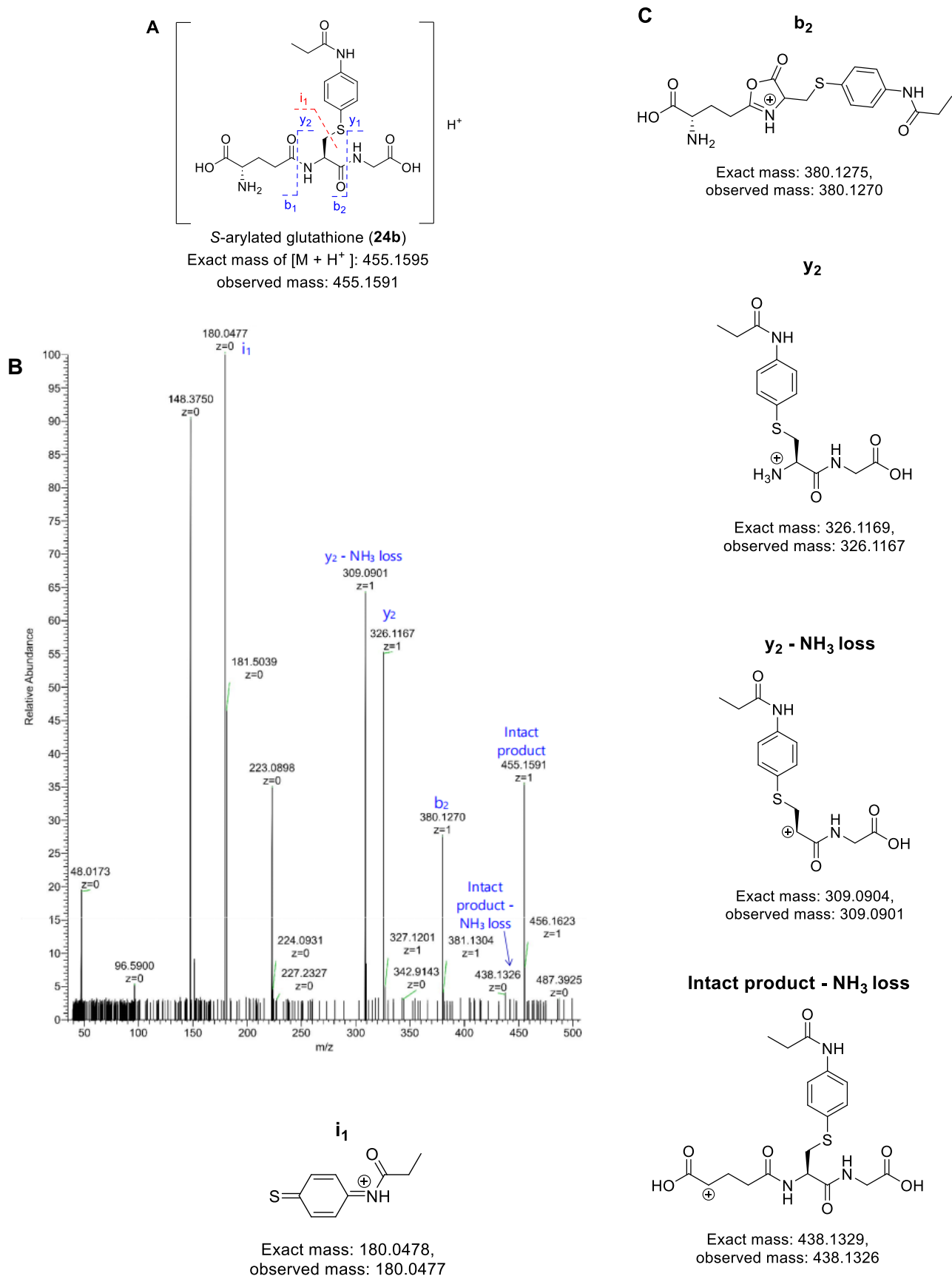


Figure S19 **A** S-arylated glutathione (**24b**) as an intact product with peptide b, y and i fragmentation. **B** High resolution MS2 spectrum of selected MS1 precursor ion ($M = 455$) with HCD fragmentation. NL: 6.69E5, CV: -20 V, HCD: 20. **C** Structure analysis and assignment of MS2 peaks from S-arylated glutathione (**24b**).

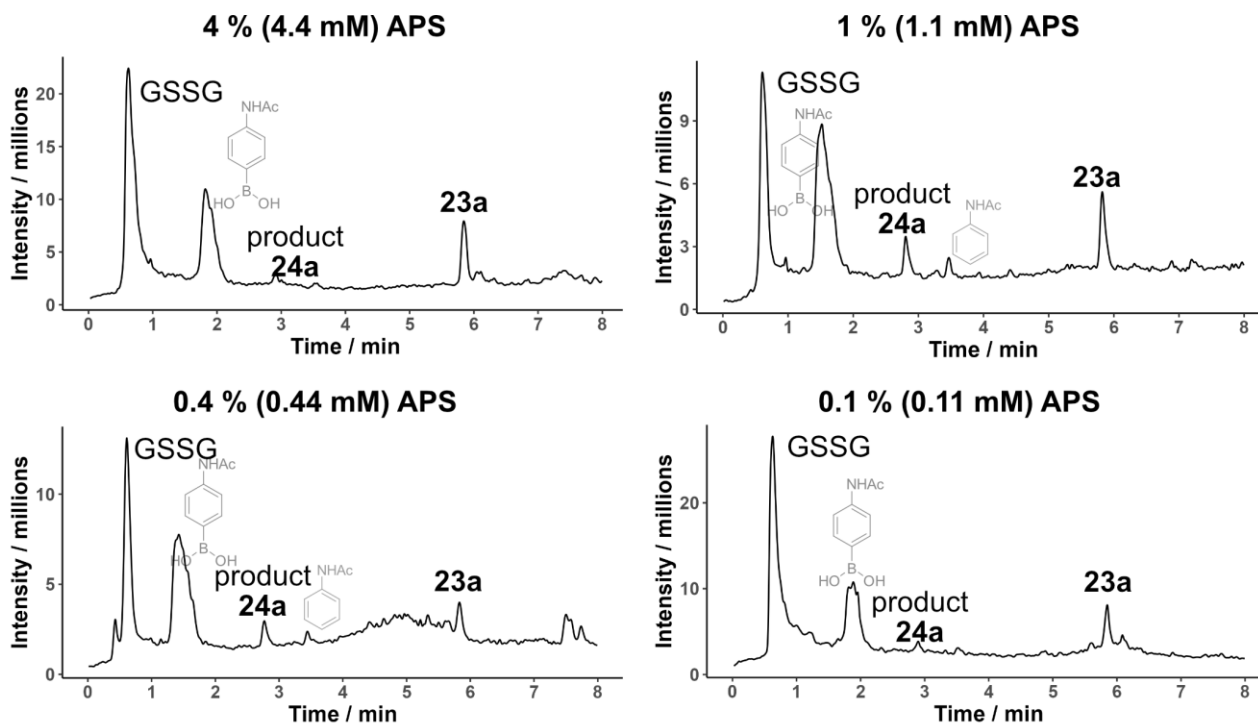


Figure S20 LC-MS chromatogram of varied catalytic amount of APS concentration. 2.5 mM GSSG, 12.5 mM **23a** and listed concentrations of APS are incubated within 40 μ L water at 50 $^{\circ}$ C for 2 hours, and then subjected to lyophilization to remove water. The residuals were reconstituted within 40 μ L 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8 μ L), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 μ L/minute.

KEHSPC disulfide, calculated mass: 1397.5936, $z = 1$; 699.3008, $z = 2$; 350.1543, $z = 4$.

Observed mass: 699.3059, $z = 2$; 350.1569, $z = 4$.

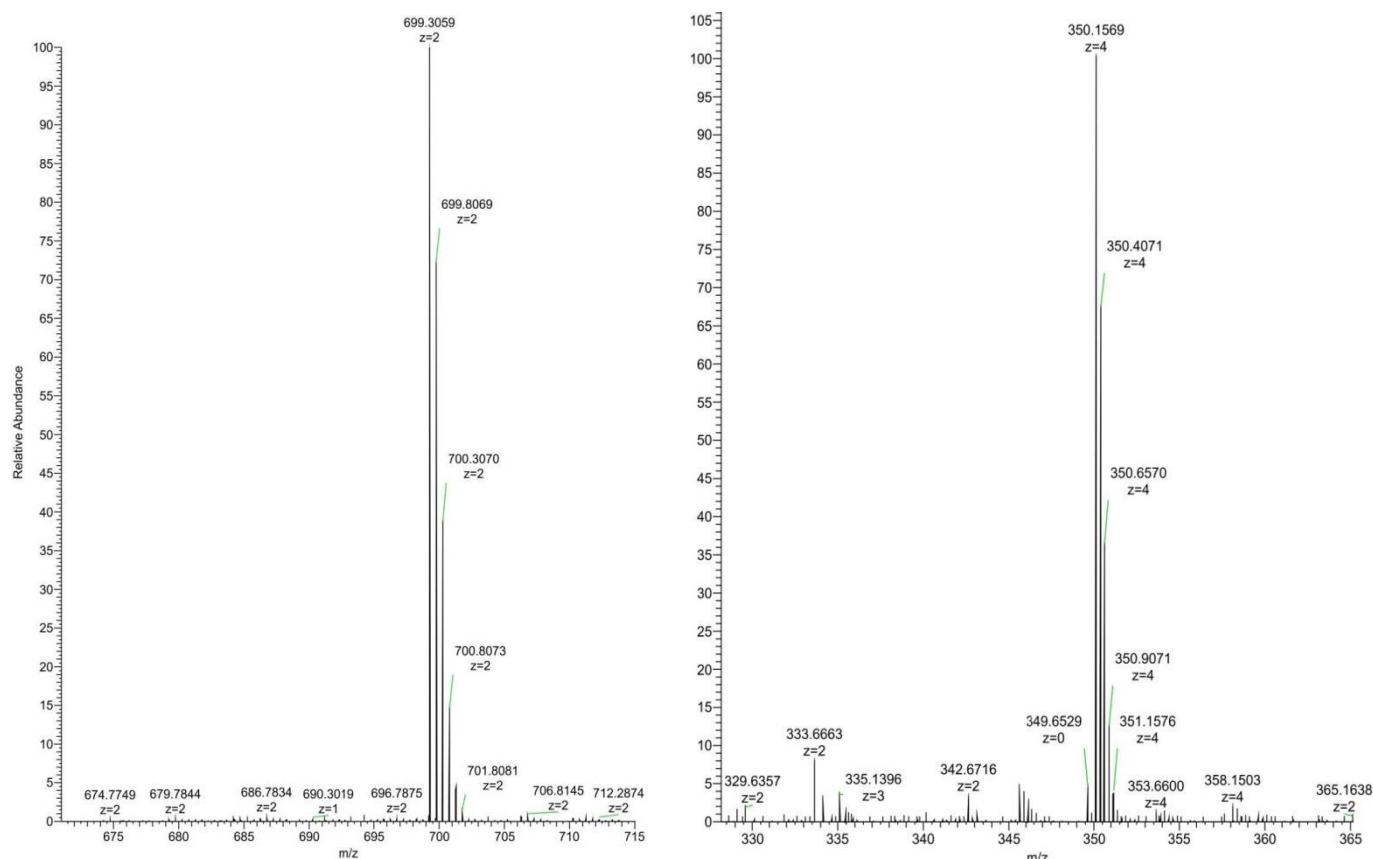


Figure S21 High-resolution MS analysis of oxidized forms of hexapeptide KEHSPC.

RWTPCD disulfide, calculated mass: 1551.6467, $z = 1$; 776.3273, $z = 2$; 517.8875, $z = 3$.

Observed mass: 776.3323, $z = 2$; 517.8914, $z = 3$.

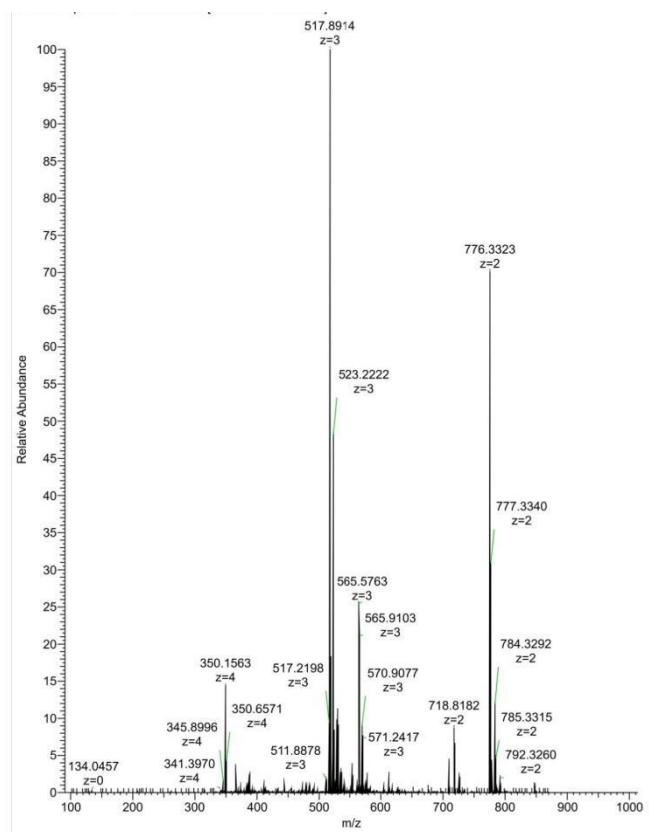


Figure S22 High-resolution MS analysis of oxidized forms of hexapeptide RWTPCD.

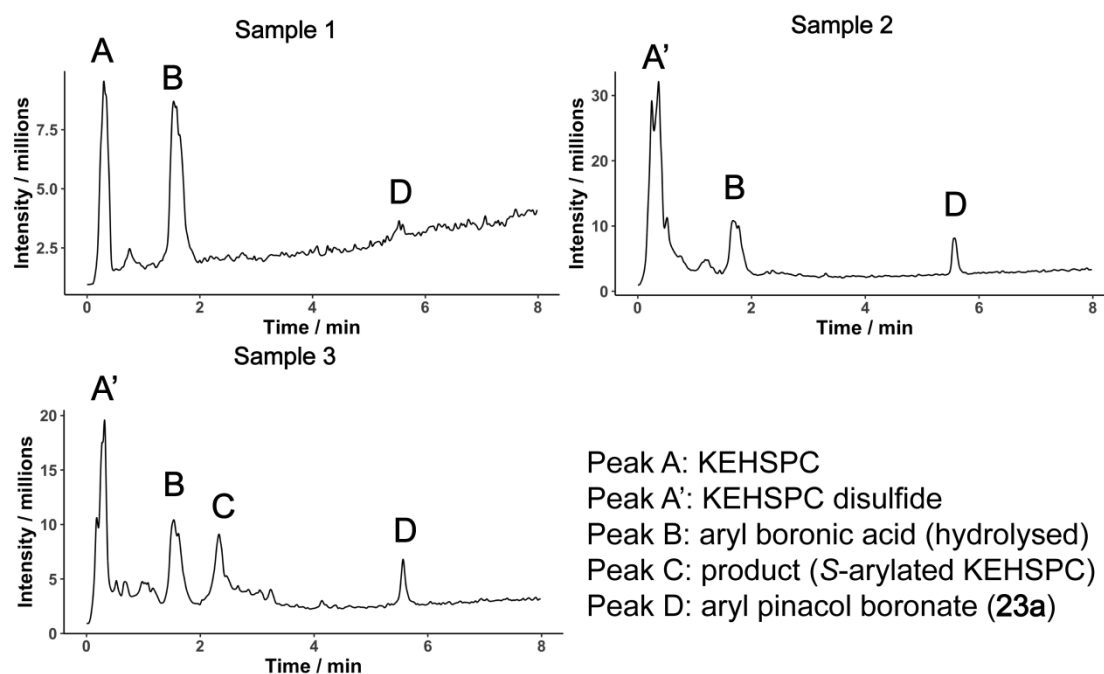


Figure S23 LC-MS chromatogram of KEHSPC / KEHSPC disulfide reactivity in presence of APS. Sample 1: 2.5 mM KEHSPC with 12.5 mM **23a** and 0.01 % APS in 40 μ L water. Sample 2: 2.5 mM KEHSPC disulfide with 12.5 mM **23a** in 40 μ L water. Sample 3: 2.5 mM KEHSPC disulfide with 12.5 mM **23a** and 0.01 % APS in 40 μ L water. These samples were treated at 80 $^{\circ}$ C for 2 hours, and then subjected to lyophilization to remove water. The residual white powders were reconstituted within 40 μ L 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8

μL), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 $\mu\text{L}/\text{minute}$.

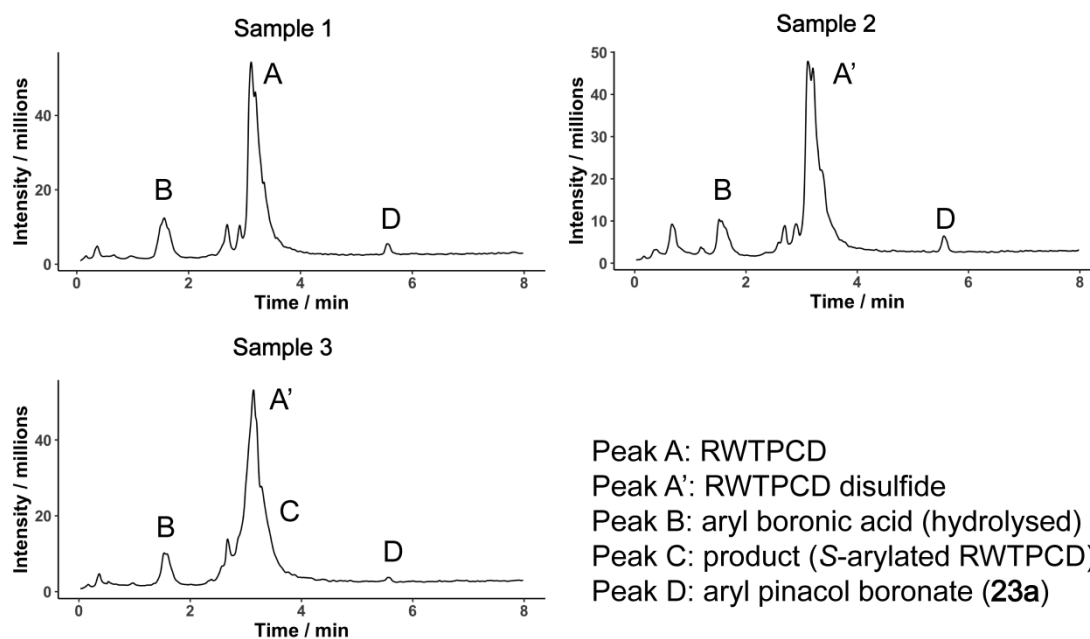


Figure S24 LC-MS chromatogram of RWTPCD / RWTPCD disulfide reactivity in presence of APS. Sample 1: 2.5 mM RWTPCD with 12.5 mM **23a** and 0.01 % APS in 40 μL water. Sample 2: 2.5 mM RWTPCD disulfide with 12.5 mM **23a** in 40 μL water. Sample 3: 2.5 mM RWTPCD disulfide with 12.5 mM **23a** and 0.01 % APS in 40 μL water. These samples were treated at 80 $^{\circ}\text{C}$ for 2 hours, and then subjected to lyophilization to remove water. The residual white powders were reconstituted within 40 μL 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8 μL), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 $\mu\text{L}/\text{minute}$.

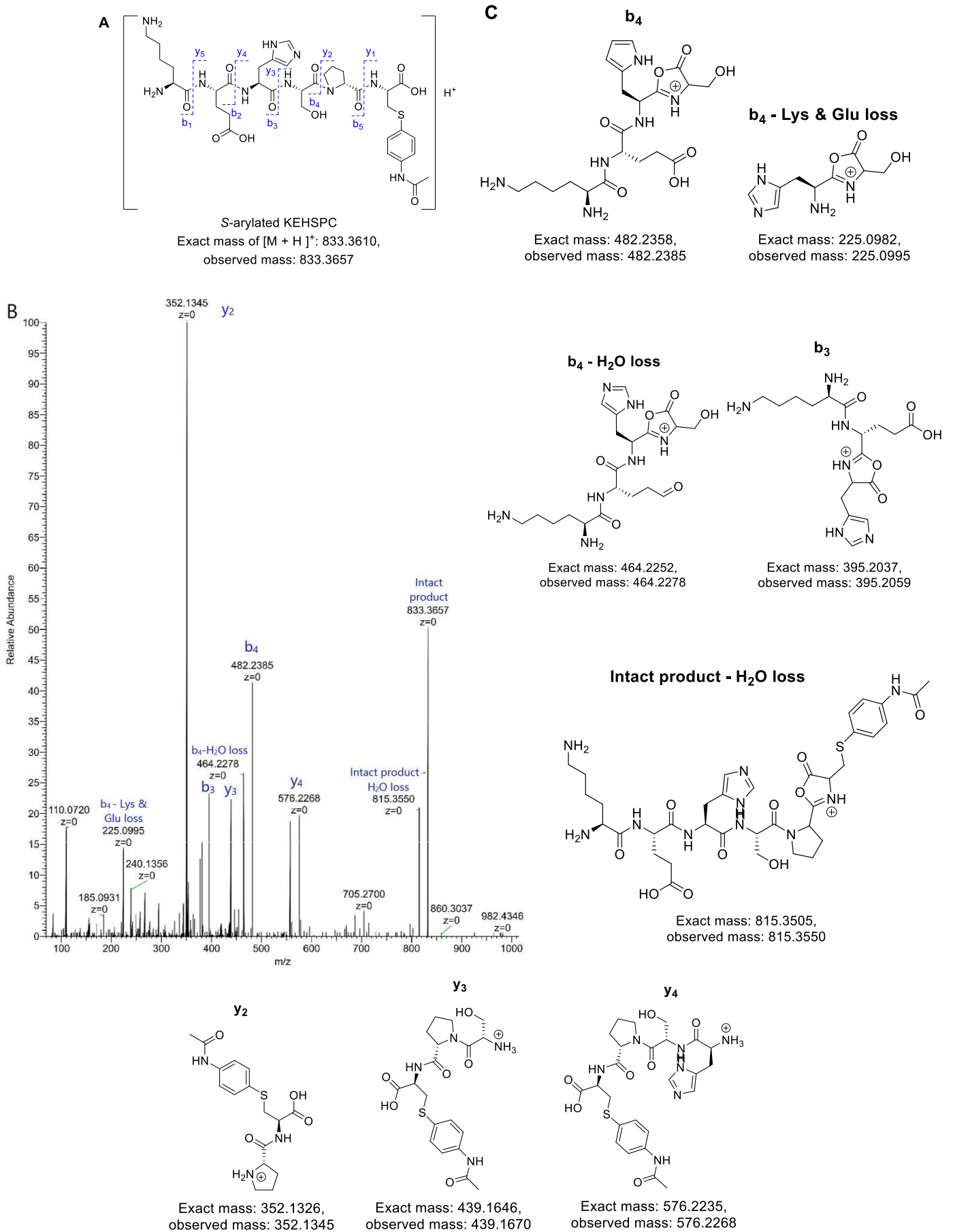


Figure S25 **A** S-arylated KEHSPC as an intact product with peptide b and y fragmentation. **B** High resolution MS2 spectrum of selected MS1 precursor ion ($M = 833$) with HCD fragmentation. NL: 8.93E6, CV: -35 V, HCD: 40. **C** Structure analysis and assignment of MS2 peaks from S-arylated KEHSPC.

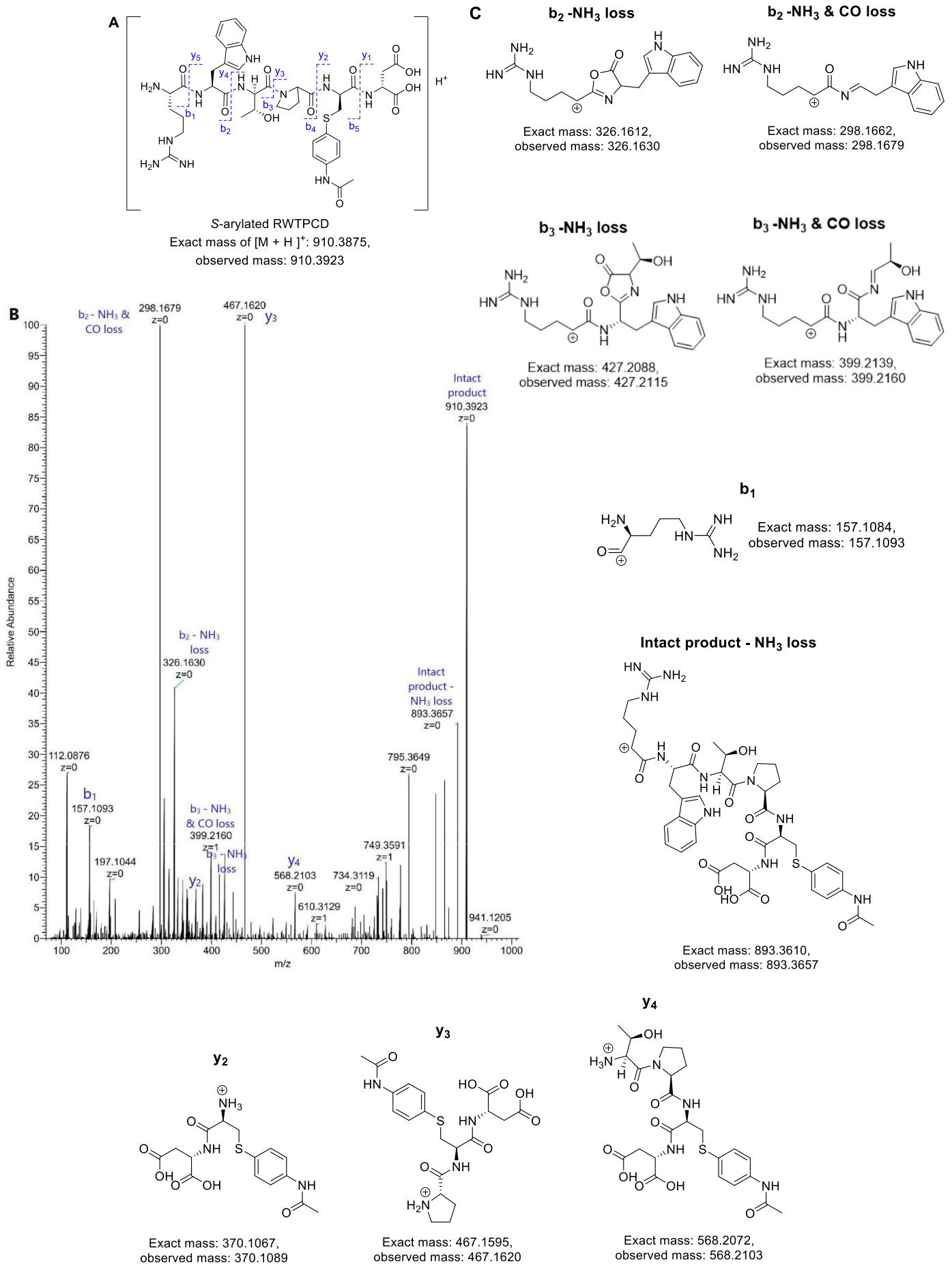
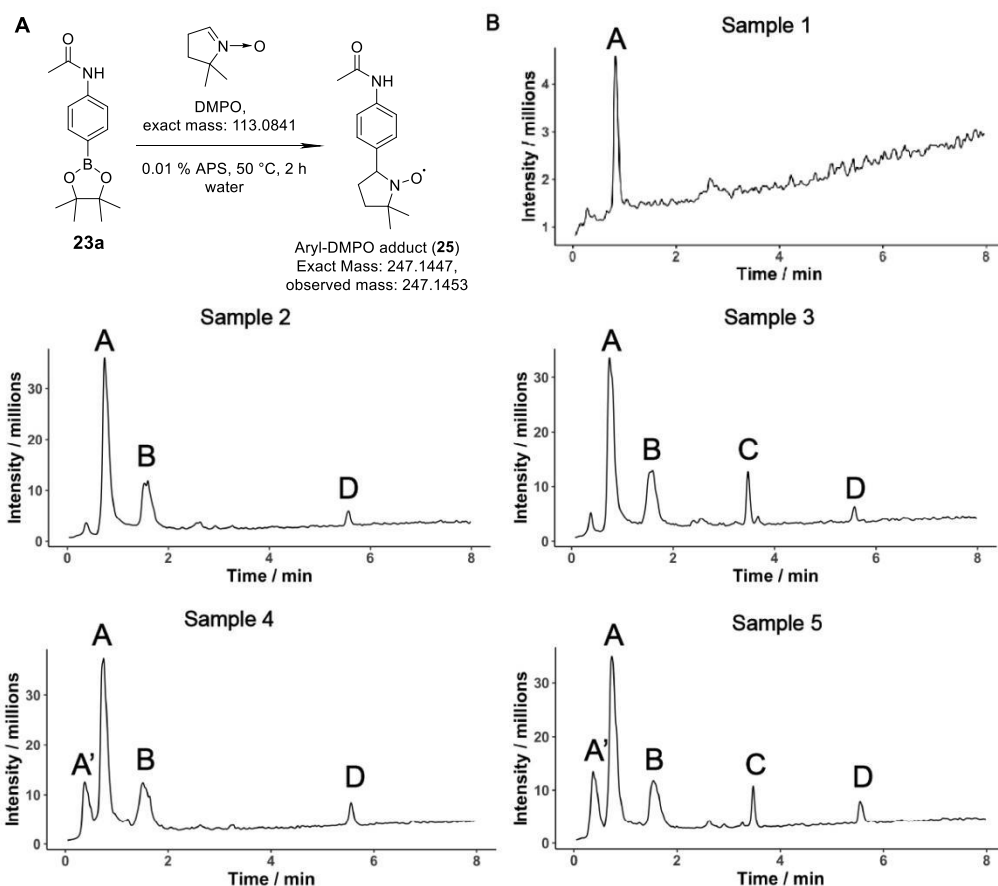


Figure S26 **A** S-arylated RWTPCD as an intact product with peptide b and y fragmentation. **B** High resolution MS2 spectrum of selected MS1 precursor ion ($M = 910$) with HCD fragmentation. NL: 2.03E5, CV: -22 V, HCD: 35. **C** Structure analysis and assignment of MS2 peaks from S-arylated RWTPCD.



Peak A: DMPO; Peak A': GSSG; Peak B: aryl boronic acid (hydrolysed);
Peak C: aryl-DMPO spin adduct (**25**); Peak D: aryl pinacol boronate (**23a**)

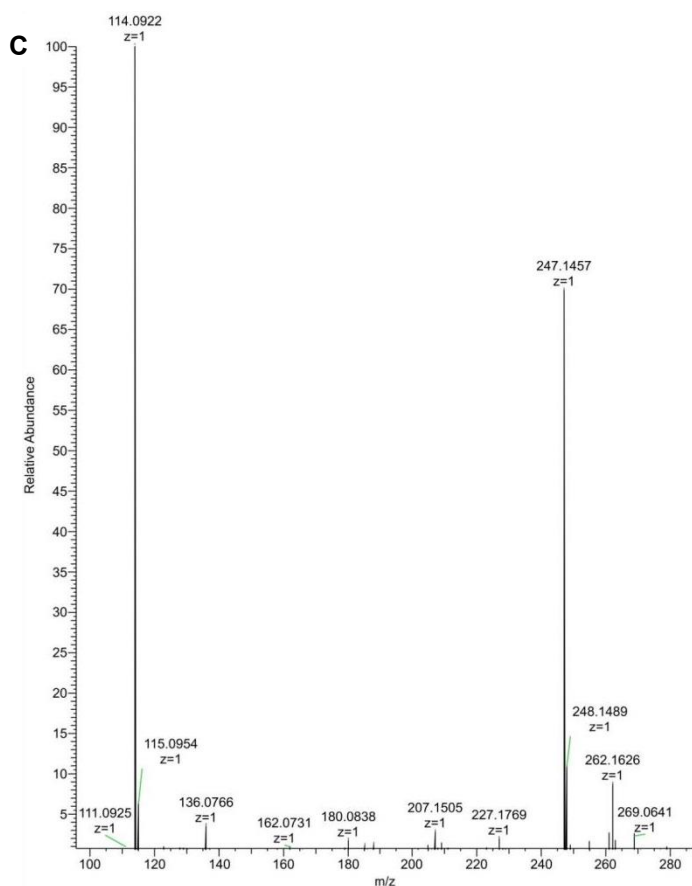


Figure S27 A Spin trap of aryl radicals by 5,5-dimethyl-1-pyrroline *N*-oxide (DMPO). **B** LC-MS chromatograms. Sample 1: 0.01 % APS with 25 mM DMPO in 40 μ L water. Sample 2: 12.5 mM **23a** with 25 mM DMPO in 40 μ L water. Sample 3: 0.01 % APS, 12.5 mM **23a** with 25 mM DMPO in 40 μ L water. Sample 4: 12.5 mM **23a**, 2.5 mM GSSG with 25 mM DMPO in 40 μ L water. Sample 5: 0.01 % APS, 12.5 mM **23a**, 2.5 mM GSSG with 25 mM DMPO in 40 μ L water. These samples were treated at 50 ° C for 2 hours, and then subjected to lyophilization to remove water. The residual white powders were reconstituted within 40 μ L 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8 μ L), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 μ L/minute. **C** DMPO-trapped aryl radical, calculated mass: 247.1447, observed mass: 247.1457. Peak 114.0922: DMPO.

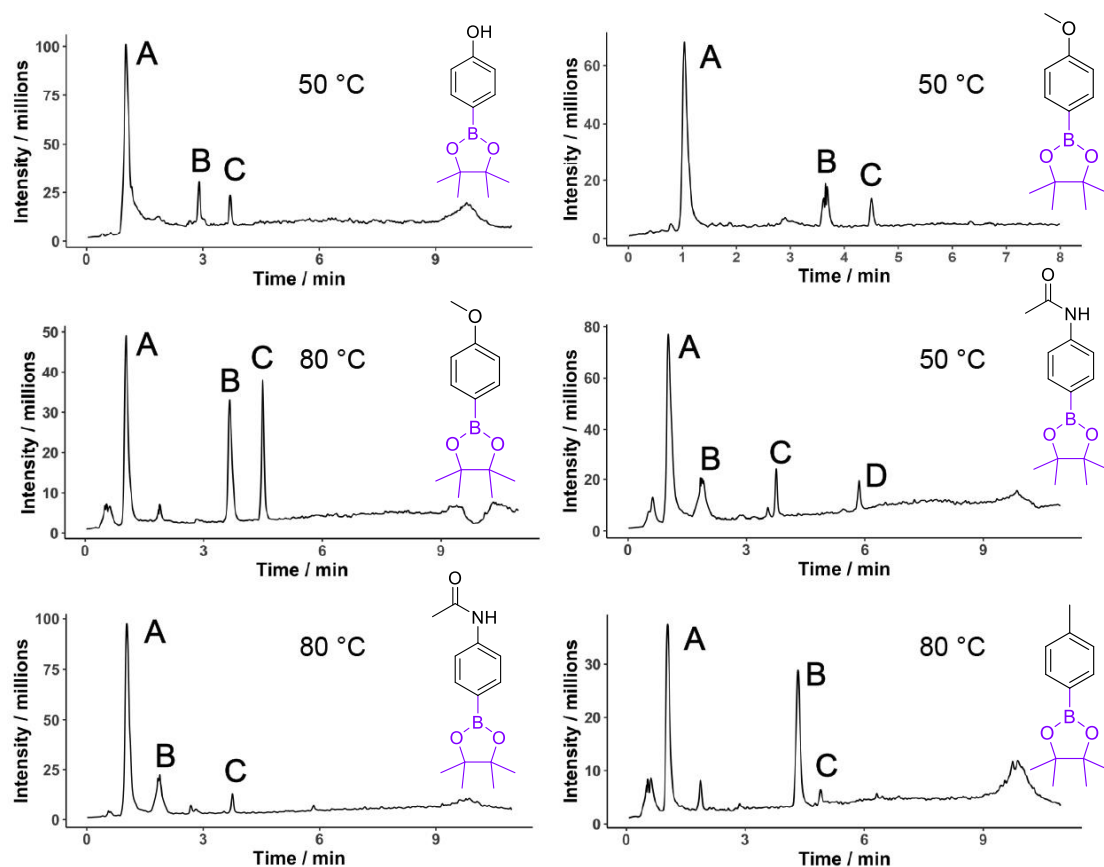
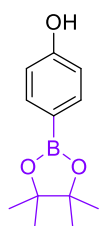
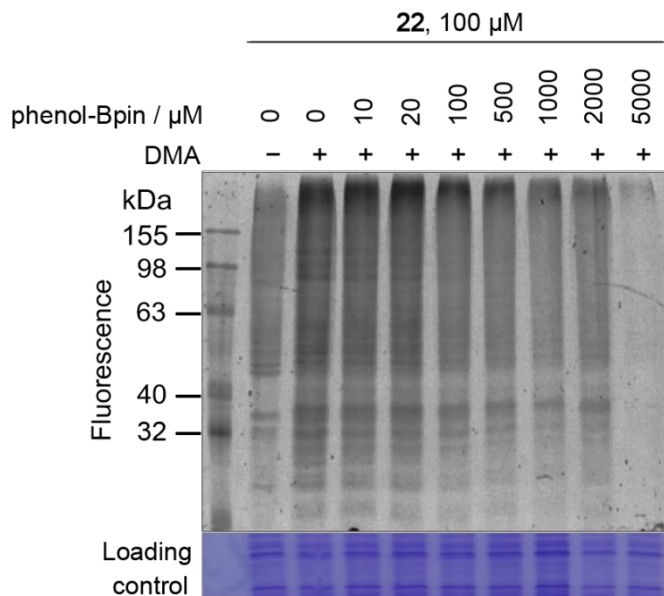
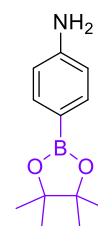
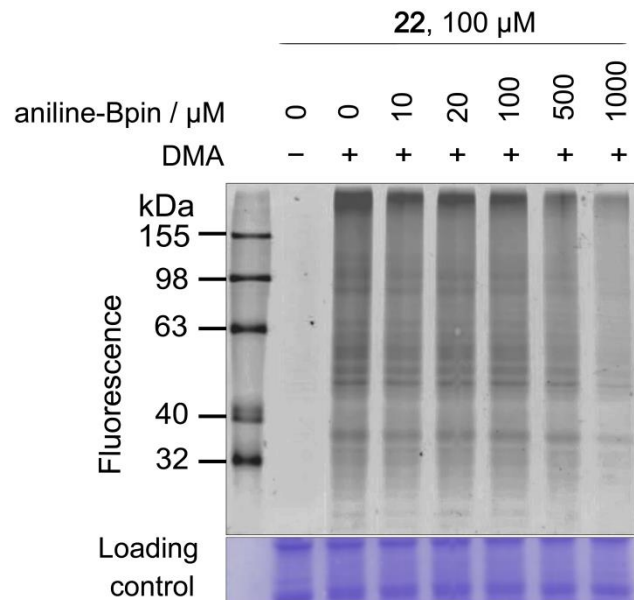


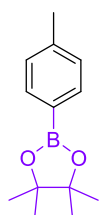
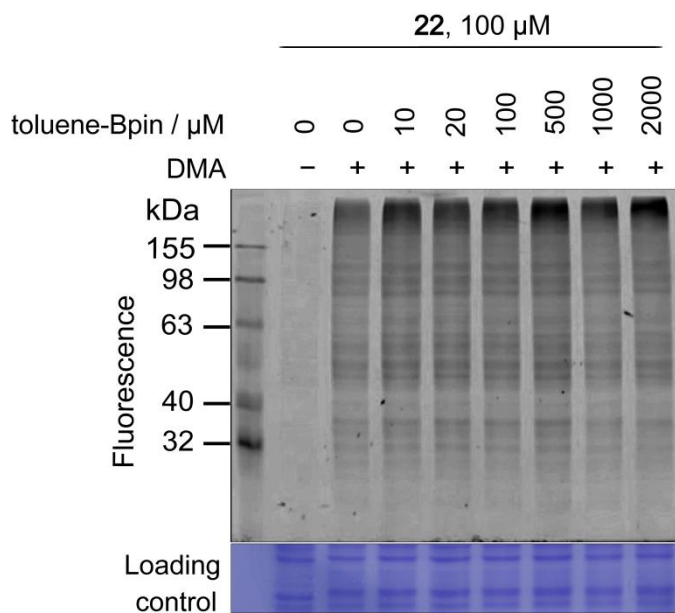
Figure S28 LC-MS chromatograms showing aryl radical formations from aryl pinacol boronates. All samples shown were prepared with 0.01 % APS, 12.5 mM corresponding aryl pinacol boronates, and 25 mM DMPO in 40 μ L water. These samples were treated at temperature shown in the figures for 2 hours, and then subjected to lyophilization to remove water. The residual white powders were reconstituted within 40 μ L 10 % acetonitrile in water and placed in a MS-vial. The samples were measured by LC-MS (injection volume: 8 μ L), with acetonitrile gradient from 5 % to 80 % within 7 minutes and the flow rate was 600 μ L/minute. Peak A: DMPO; Peak B: aryl boronic acid (hydrolysed); Peak C: aryl-DMPO spin adduct; Peak D: aryl pinacol boronate (if applicable).



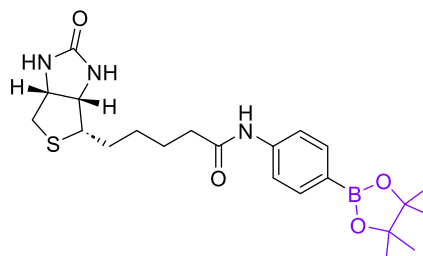
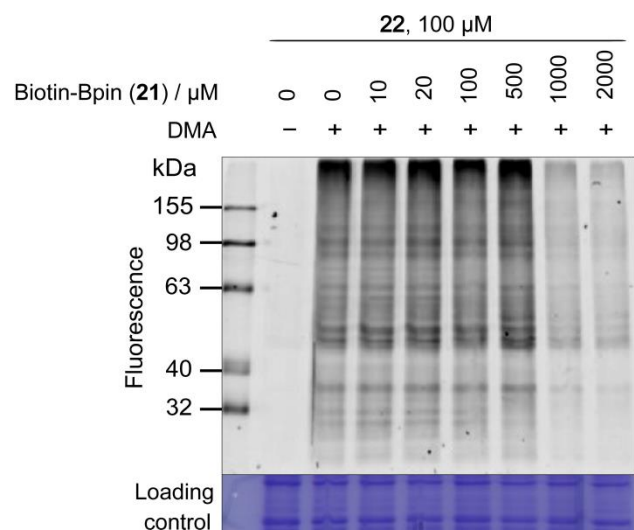
Starts to compete from 1:1 (**22**)



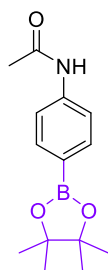
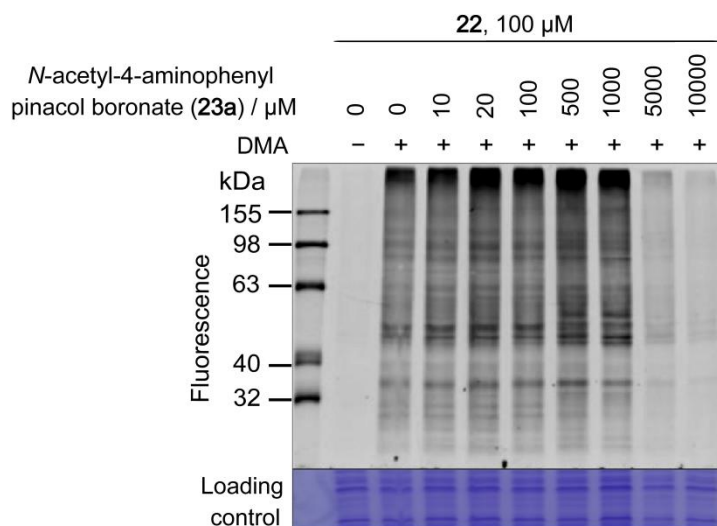
Starts to compete from 5:1 (**22**)



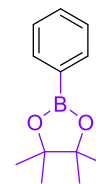
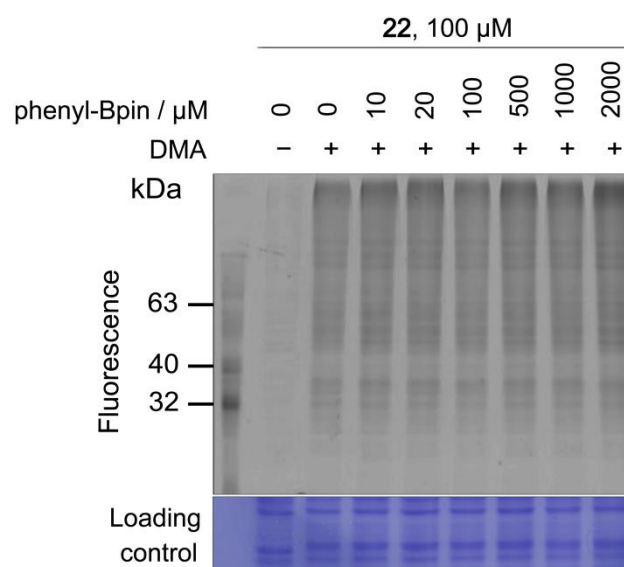
Starts to compete from 10:1 (**22**)



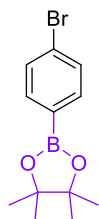
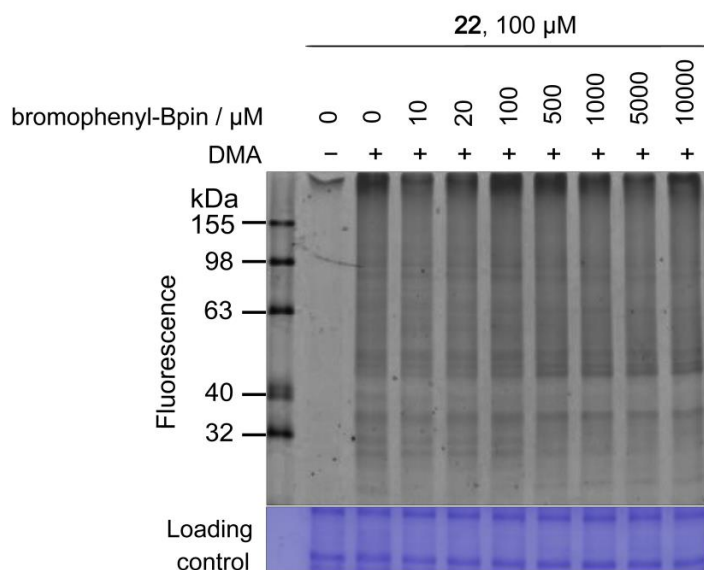
Starts to compete from 10:1 (**22**)



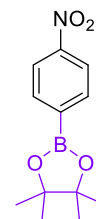
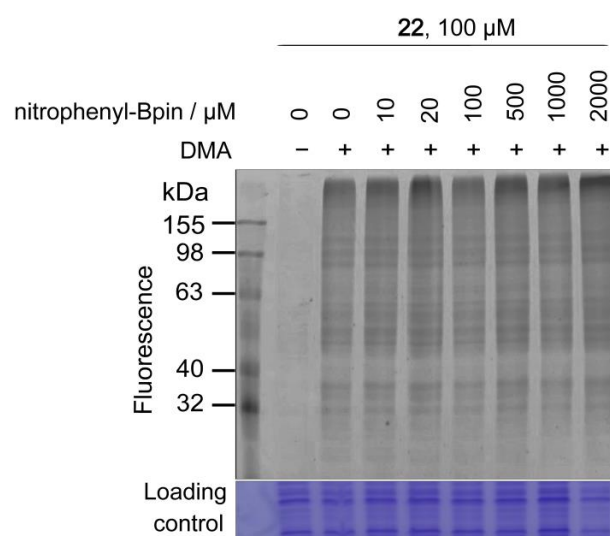
Starts to compete from 50:1 (**22**)



No competition was shown



No competition was shown



No competition was shown

Figure S29 In-gel fluorescence analysis of competition experiments between fluorescent aryl pinacol boronate **22** (TAMRA-Bpin, 1 μ L, final conc. 100 μ M) and other aryl pinacol boronates as competitors (1 μ L, final conc. listed on each figure) within HeLa cell lysates. The reaction mixture (total vol. at 40 μ L) was incubated at 50 $^{\circ}$ C, 750 rpm for 2 hours. The labeled proteins were resolved by 10 % SDS-PAGE gel.

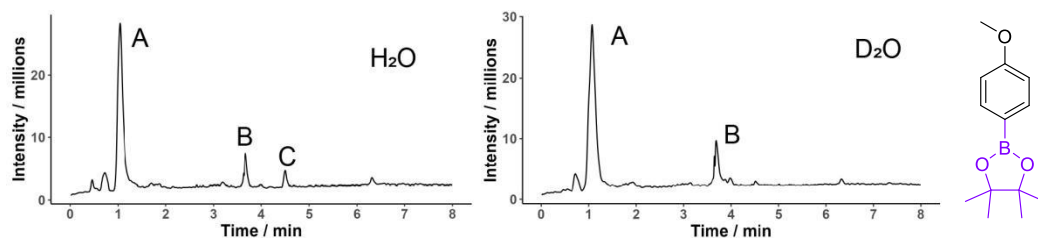


Figure S30 LC-MS chromatogram showing isotopic effect when incubating 12.5 mM anisole-Bpin with 25 mM DMPO and 0.01 % APS in either 40 μ L water (left panel) or 40 μ L deuterated water (right panel). Peak A: DMPO; Peak B: aryl boronic acid (hydrolysed); Peak C: aryl-DMPO spin adduct.

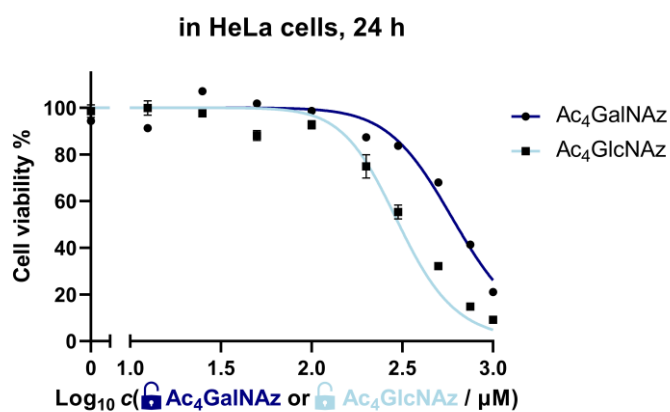


Figure S31 Cell viability assay to evaluate the cytotoxicity of Ac₄GalNAz and Ac₄GlcNAz after 24 h incubation in HeLa cells, $n = 3$.

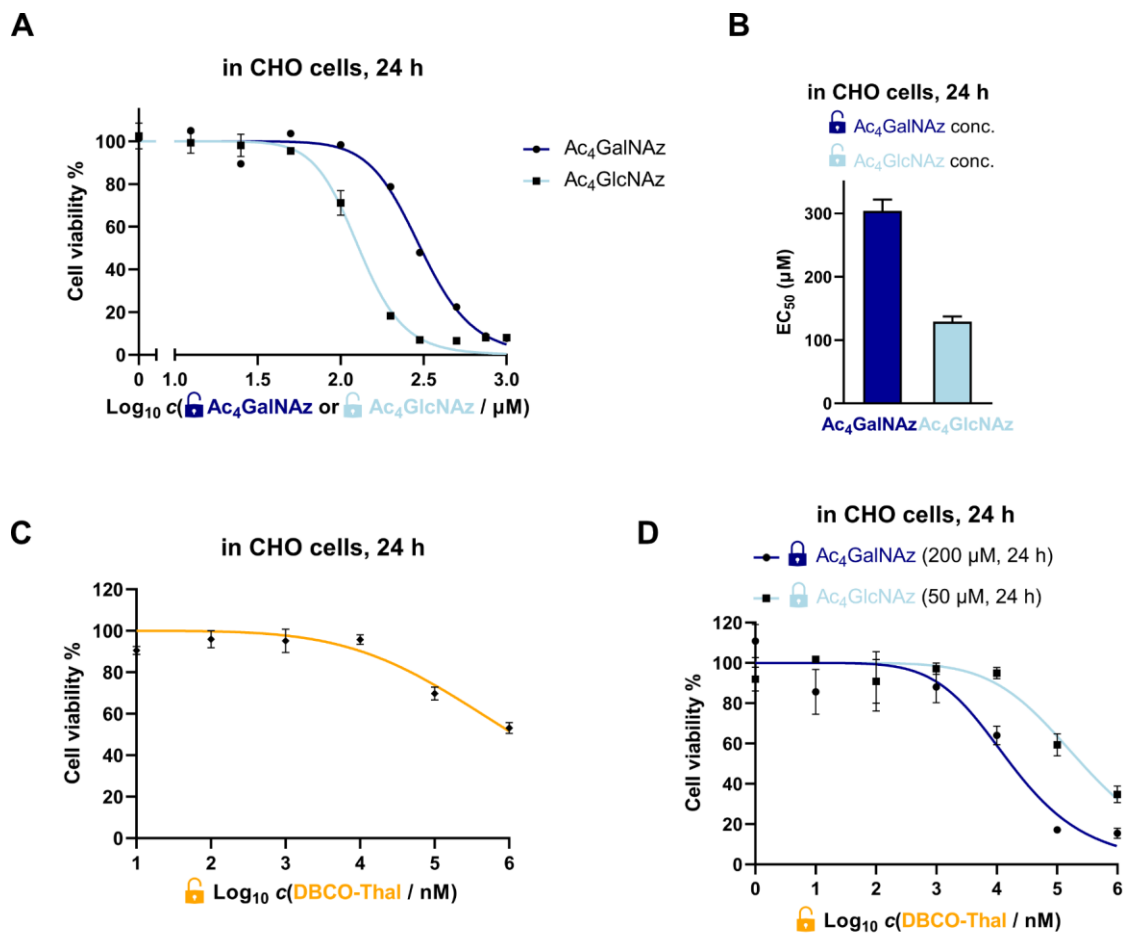


Figure S32 **A** Cell viability assay to evaluate the cytotoxicity of the Ac₄GalNAz and Ac₄GlcNAz in CHO cells, $n = 3$. Ac₄GalNAz gave EC₅₀ value at 303 μM and Ac₄GlcNAz at 129 μM, with no cytotoxicity up to 200 μM and 80 μM, respectively. **B** Bar plot comparing cytotoxicity of Ac₄GalNAz and Ac₄GlcNAz analogues in CHO cells after 24 h incubation, $n = 3$. **C** Concentration-dependent cytotoxicity of DBCO-Thal in CHO cells after 24 h incubation. Concentration range between 10 nM and 1 mM, $n = 3$. **D** DBCO-Thal concentration-dependent cytotoxicity after incubation with Ac₄GalNAz (200 μM, 24 h) or Ac₄GlcNAz (50 μM, 24 h) in CHO cells, $n = 3$. DBCO-Thal gave EC₅₀ value at 17 and 257 μM for Ac₄GalNAz or Ac₄GlcNAz pre-treated CHO cells, respectively.

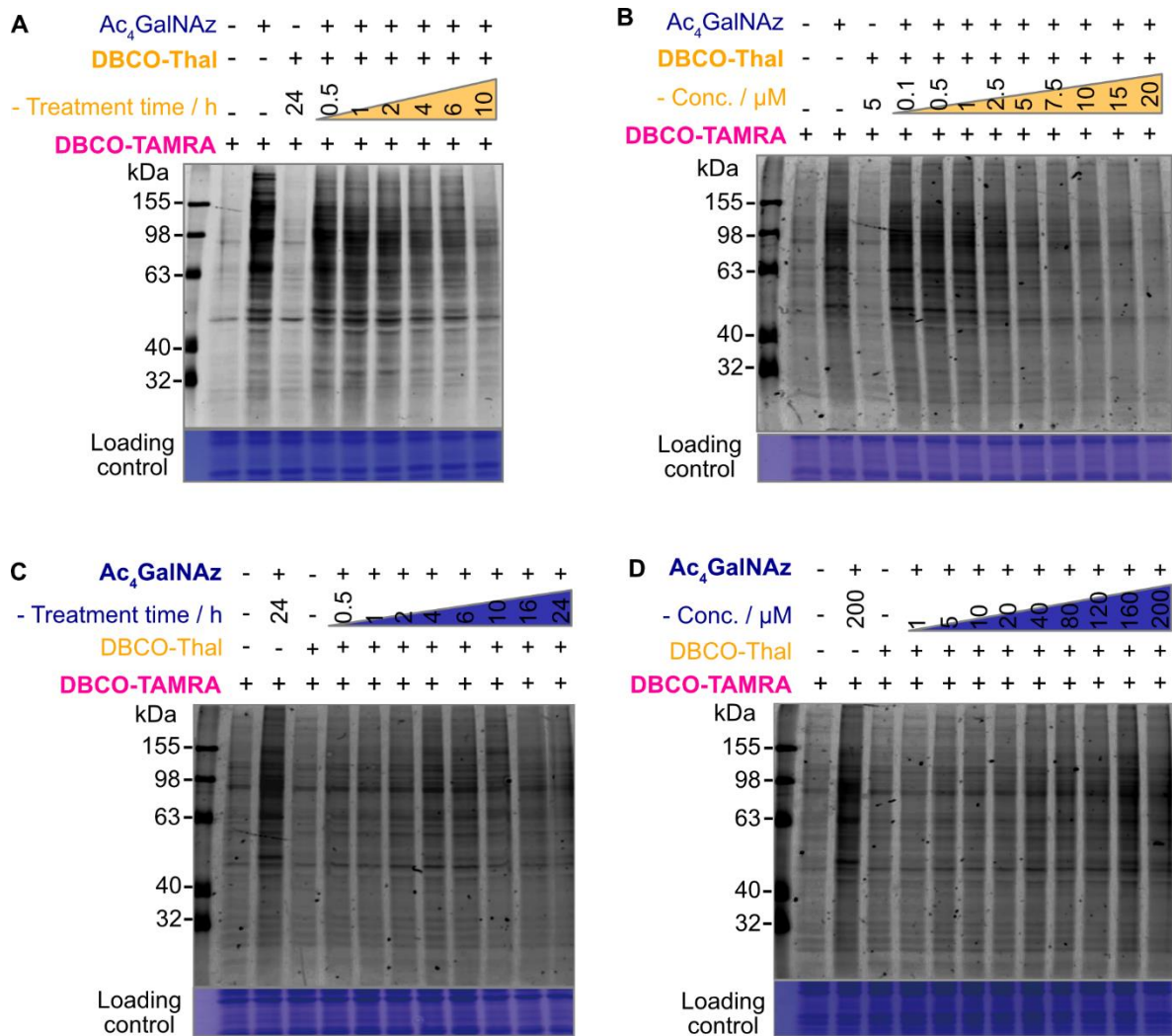


Figure S33 In-gel fluorescence analysis of cell lysates after GlyTAC treatments using DBCO-PEG₄-5-tetramethylrhodamine (DBCO-PEG₄-5-TAMRA, abbreviated as DBCO-TAMRA).

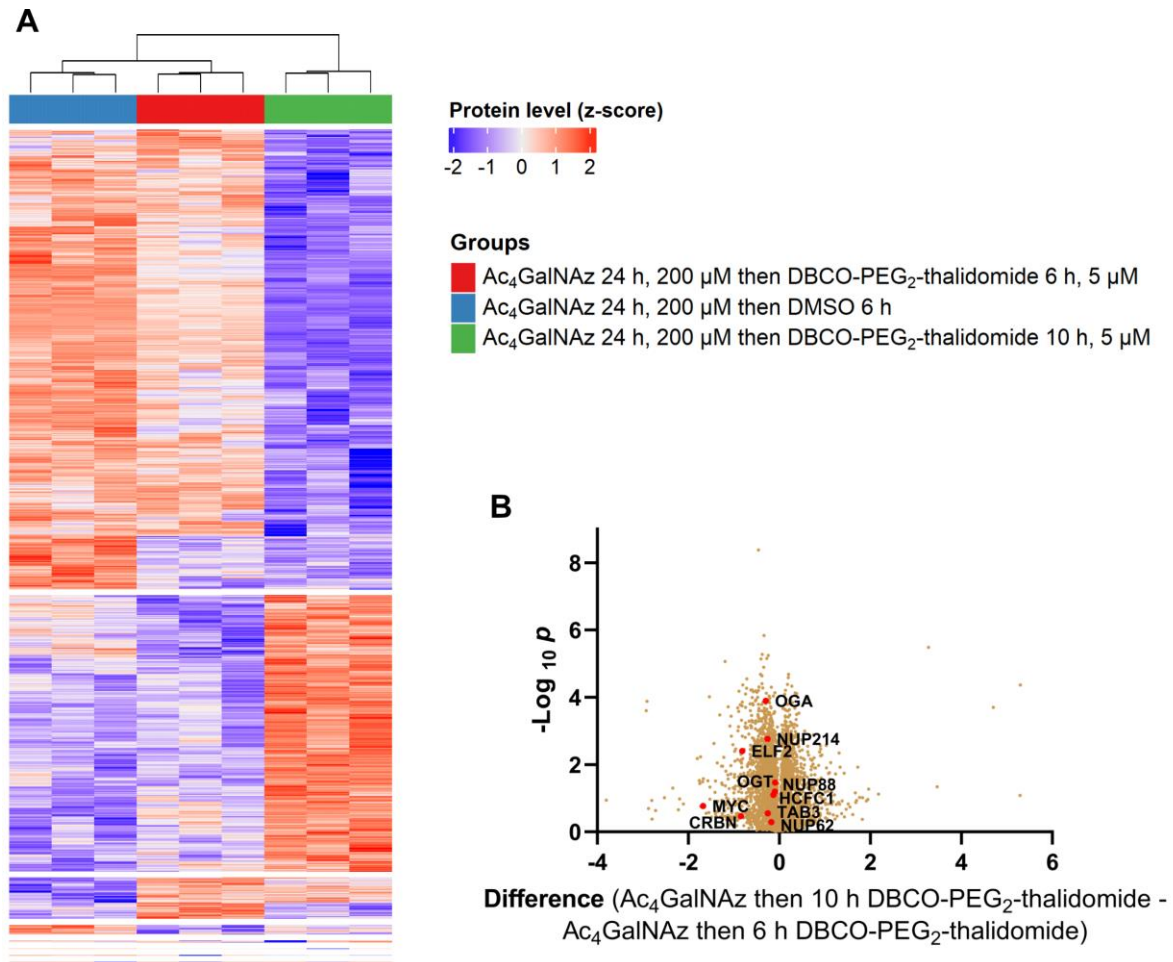


Figure S34 **A** Heatmap visualizing up- and down-regulated proteins in all conditions using ANOVA-significant test. **B** Volcano plot comparing GlyTAC (Ac₄GalNAz then 10 h, 5 μM DBCO-Thal) against GlyTAC (Ac₄GalNAz then 6 h, 5 μM DBCO-Thal) treatment.

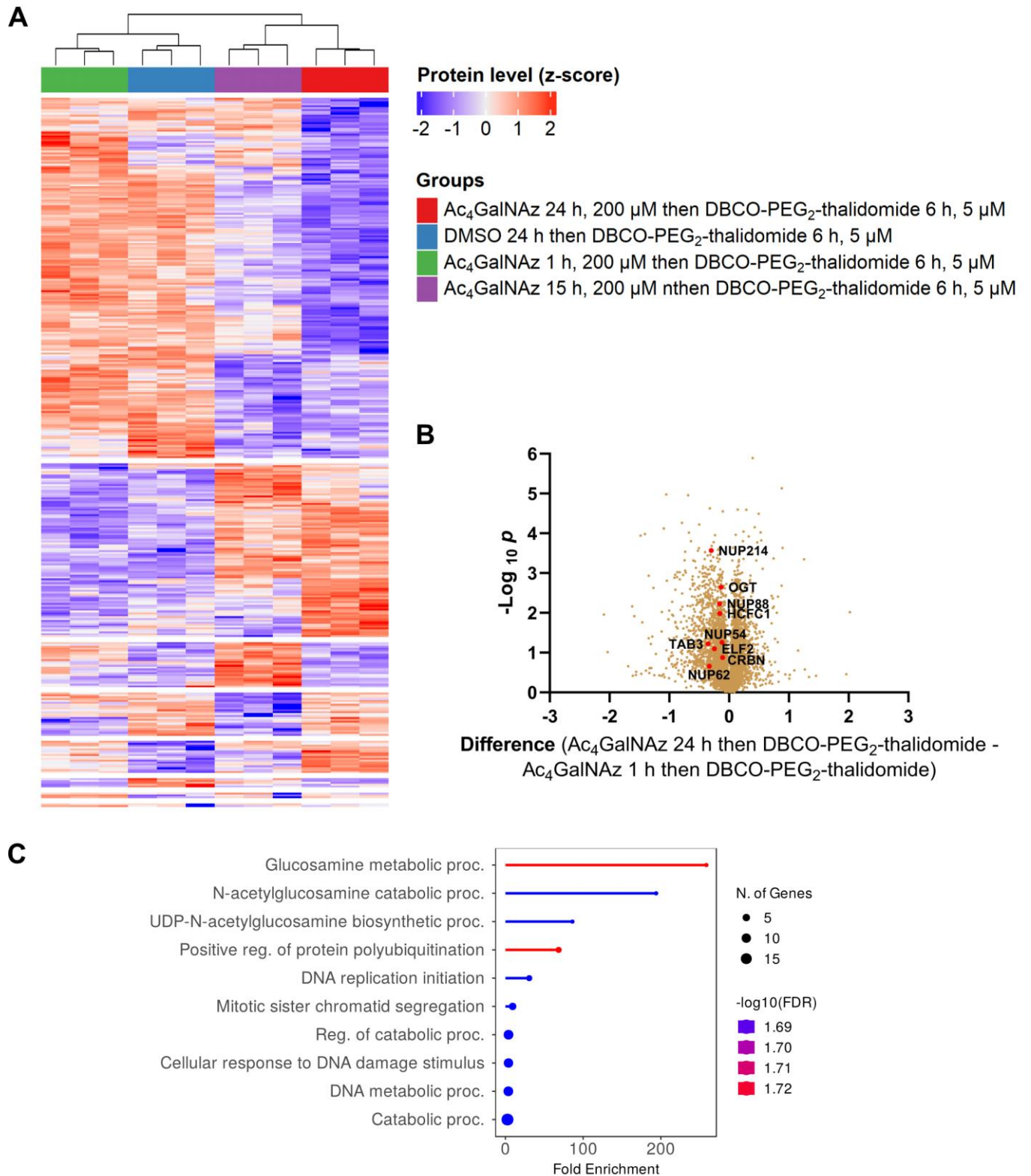


Figure S35 **A** Heatmap visualizing up- and down-regulated proteins in all conditions using ANOVA-significant test. **B** Volcano plot comparing GlyTAC (24 h Ac₄GalNAz then DBCO-Thal) against GlyTAC (1 h Ac₄GalNAz then DBCO-Thal) treatment. **C** GO term (Biological Process) analysis of significantly up-regulated proteins.

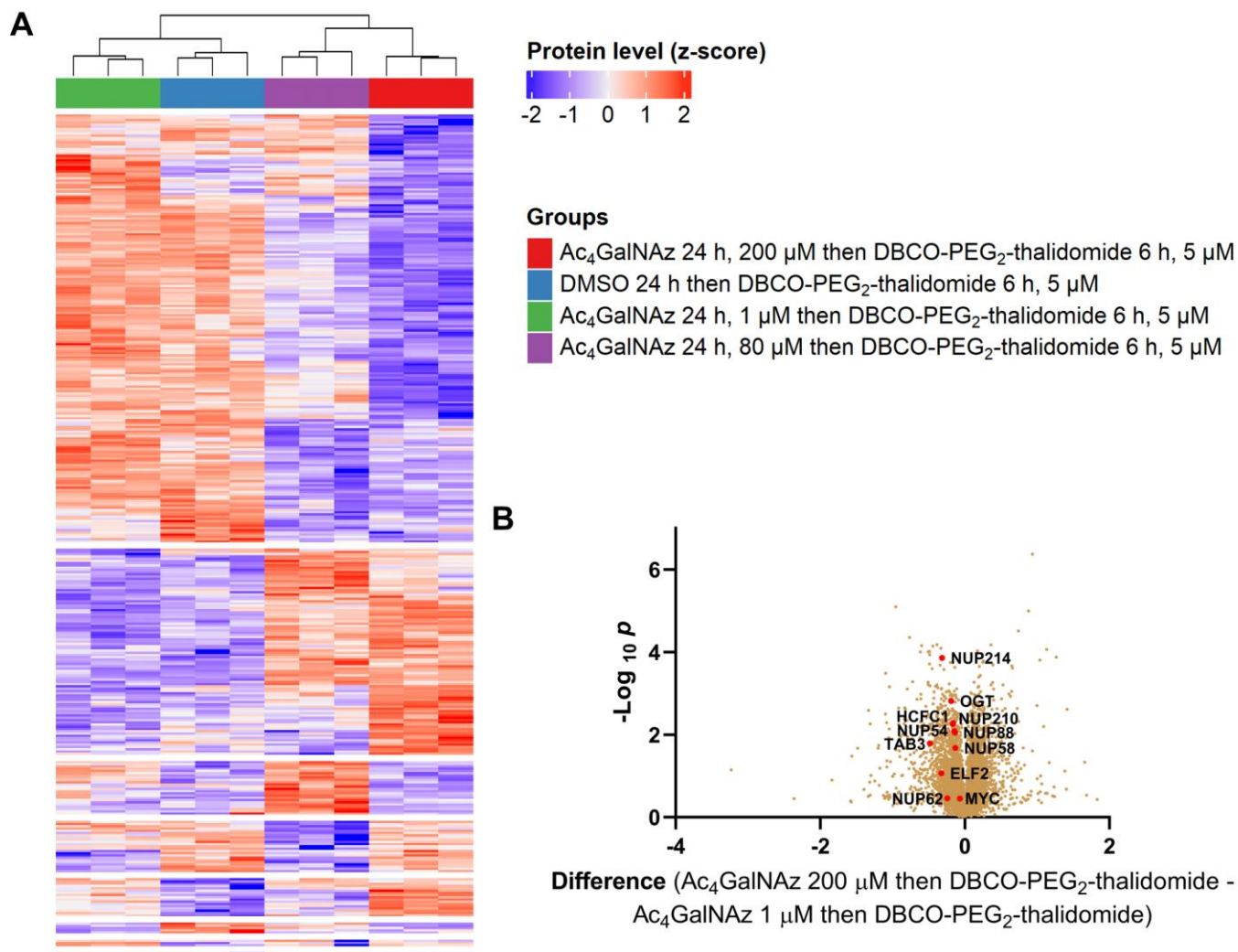


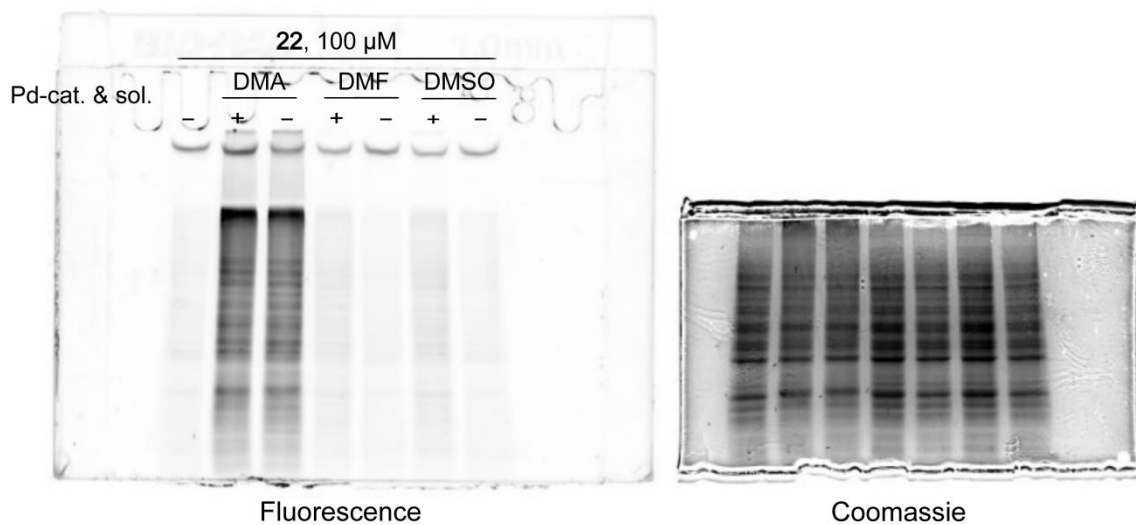
Figure S36 **A** Heatmap visualizing up- and down-regulated proteins in all conditions using ANOVA-significant test. **B** Volcano plot comparing GlyTAC (200 μM Ac₄GalNAz then DBCO-Thal) against GlyTAC (1 μM Ac₄GalNAz then DBCO-Thal) treatment.

8.2 Supplementary Table

Fragment ions	Cal. mass - 24a	Obs. mass - 24a	Cal. mass - 24b	Obs. mass - 24b
Intact product with NH ₃ loss	424.1173	424.1165	438.1329	438.1326
b ₂	366.1118	366.1111	380.1275	380.1270
y ₂	312.1013	312.1009	326.1169	326.1167
y ₂ with NH ₃ loss	295.0747	295.0745	309.0904	309.0901
i ₁	166.0321	166.0320	180.0478	180.0477

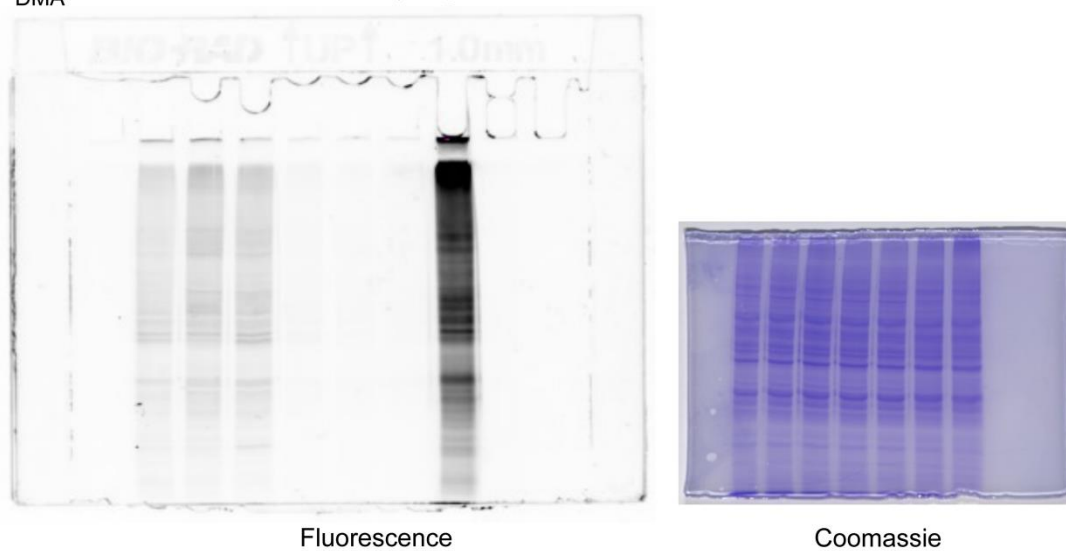
Table S1 The high-resolution MS2 fragments of product **24a** (lane 2 - 3) and **24b** (lane 4 - 5). Fragments structures are visualized in **Figure 28C** and Figure S17, S19. Units for cal. and obs. mass: Da.

8.3 Uncropped Images

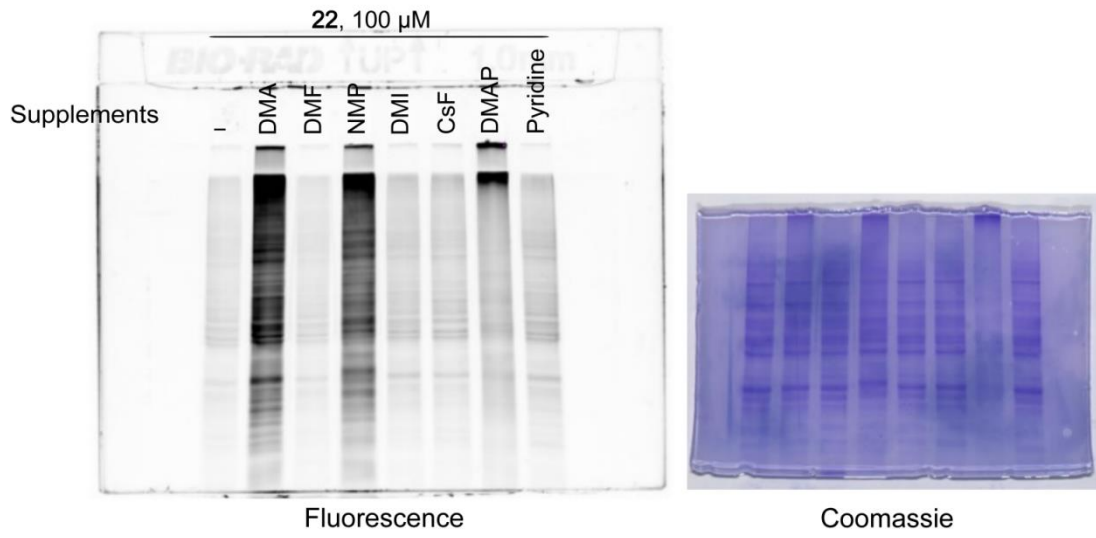


Uncropped fluorescence image and loading control by Coomassie of Figure 26C

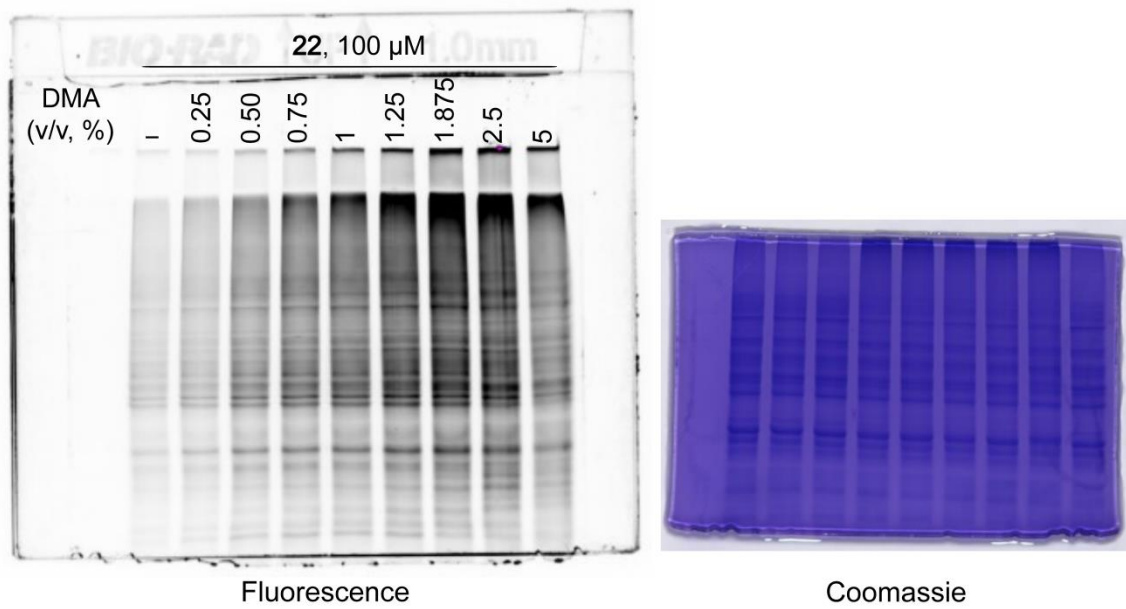
22	+	+	+	-	-	-	+
Pd(OAc) ₂ in acetonitrile	-	+	+	-	+	-	-
TAMRA-N ₃	-	-	-	+	+	+	-
DMA	-	-	-	-	-	+	+



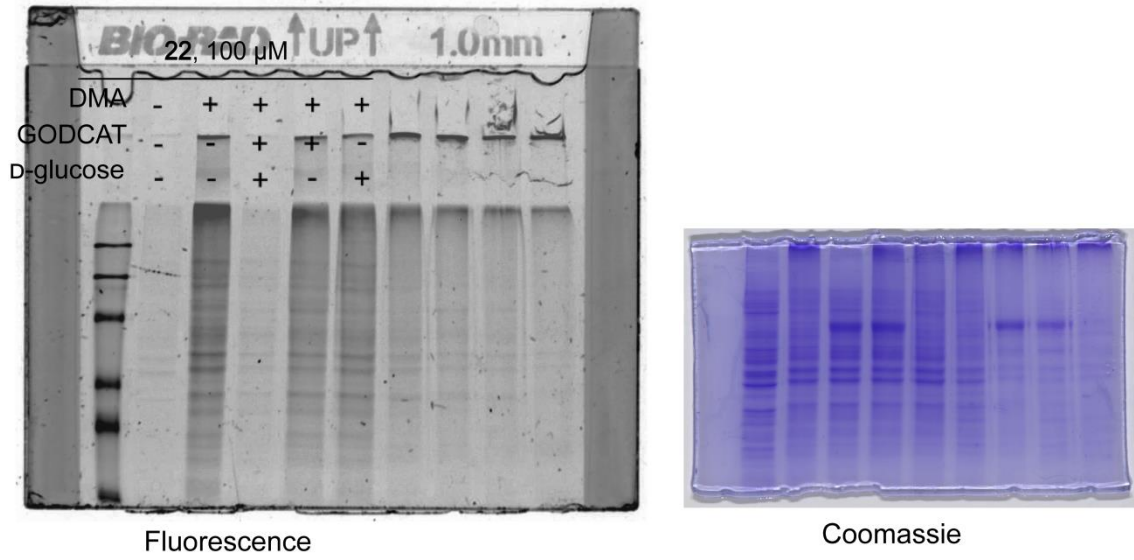
Uncropped fluorescence image and loading control by Coomassie of Figure 26D



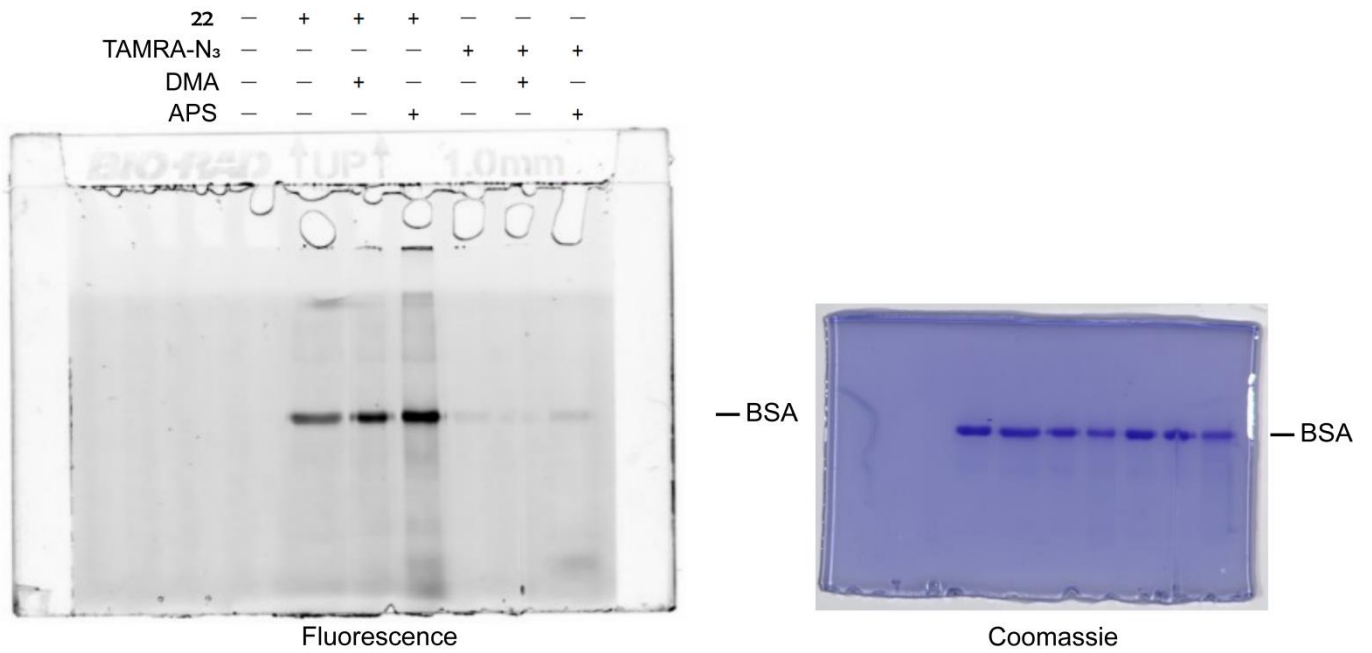
Uncropped fluorescence image and loading control by Coomassie of Figure 26E



Uncropped fluorescence image and loading control by Coomassie of Figure 26F



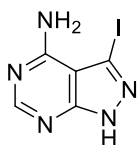
Uncropped fluorescence image and loading control by Coomassie of Figure 27A



Uncropped fluorescence image and loading control by Coomassie of Figure 27C

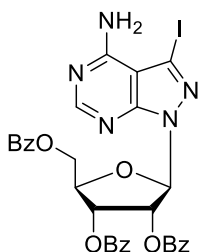
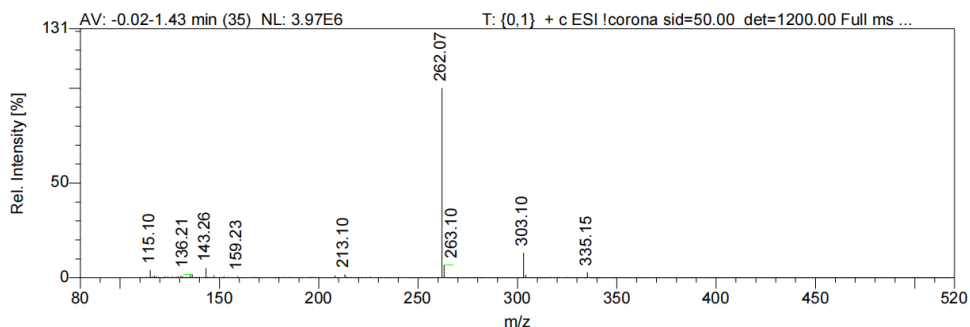
8.4 Mass Analysis

Compounds

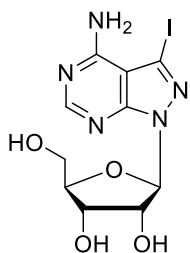
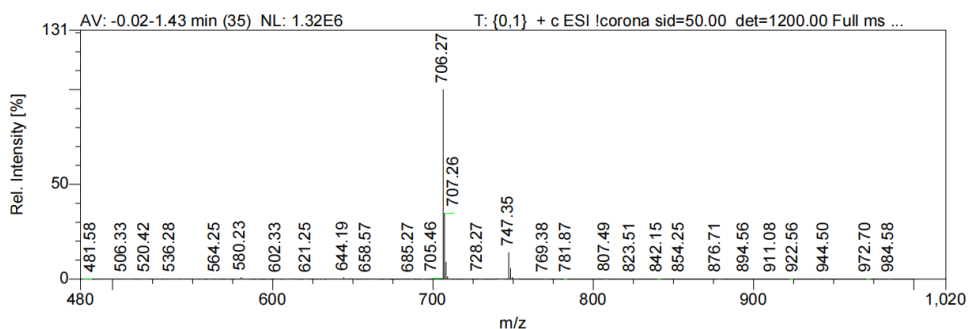


9, C₅H₅N₅I [M+H]⁺:
calculated mass: 261.95;
observed mass: 262.07.

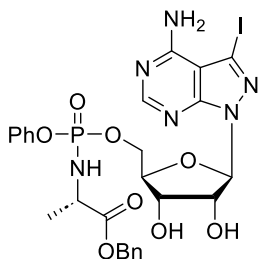
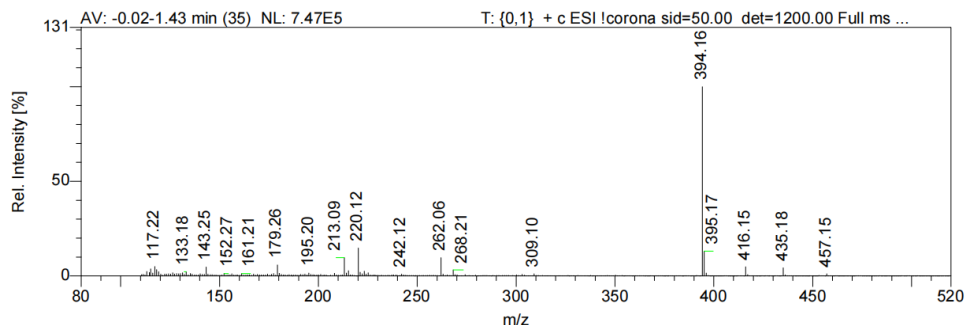
Mass Analysis (with direct injection unless specified)



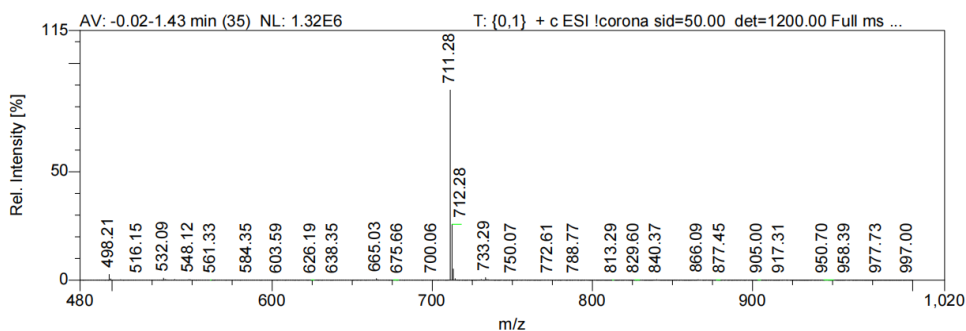
11, C₃₁H₂₅N₅O₇I [M+H]⁺:
calculated mass: 706.07;
observed mass: 706.27.

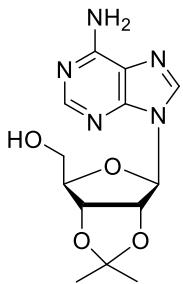


2, C₁₀H₁₃N₅O₄I [M+H]⁺:
calculated mass: 394.03;
observed mass: 394.16.

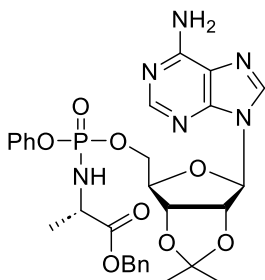
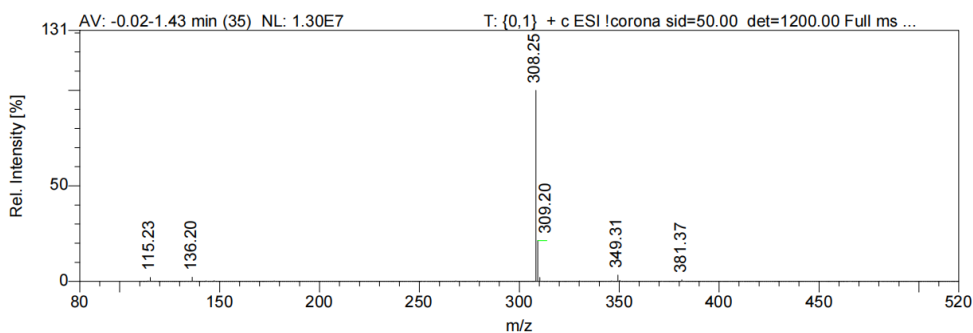


1, C₂₆H₂₉N₆O₈PI [M+H]⁺:
calculated mass: 711.10;
observed mass: 711.28.

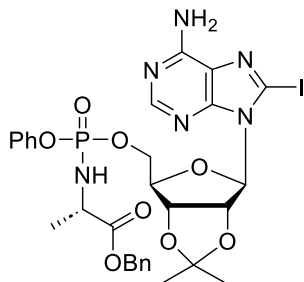
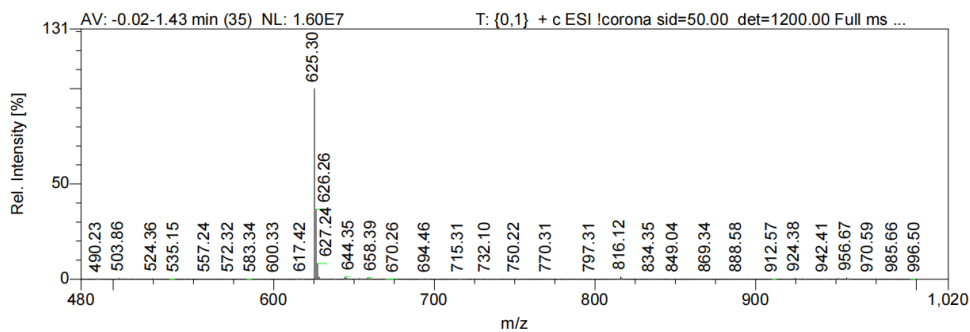




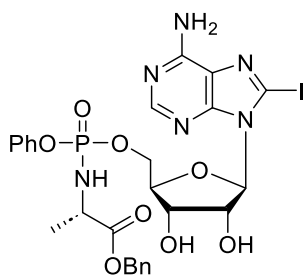
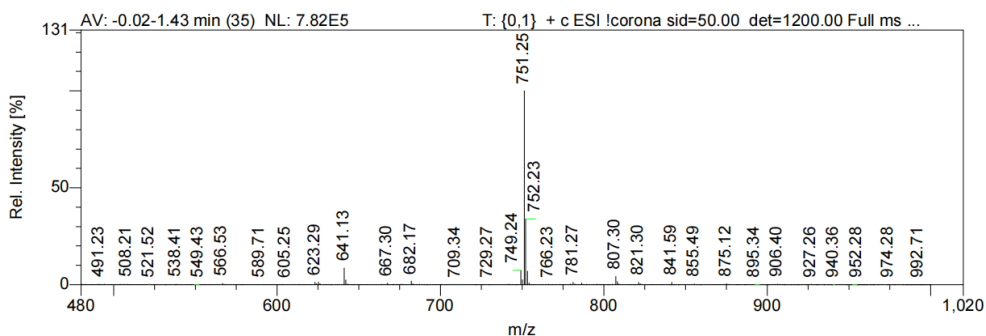
13, C₁₃H₁₈N₅O₄ [M+H]⁺:
calculated mass: 308.15;
observed mass: 308.25.



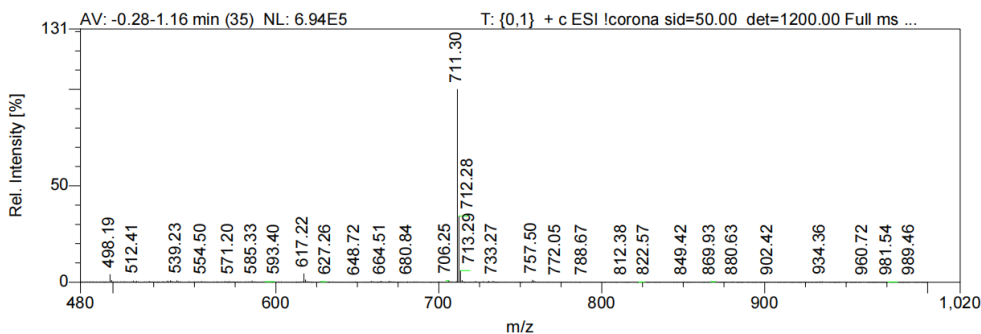
14, C₂₉H₃₄N₆O₈P [M+H]⁺:
calculated mass: 625.21;
observed mass: 625.30.

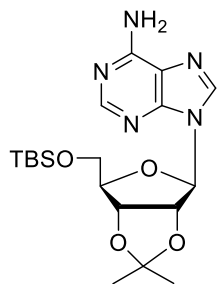


15, C₂₉H₃₃N₆O₈PI [M+H]⁺:
calculated mass: 751.12;
observed mass: 751.25.

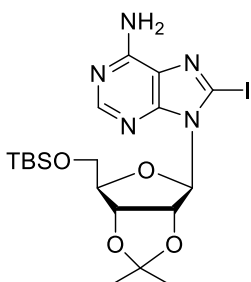
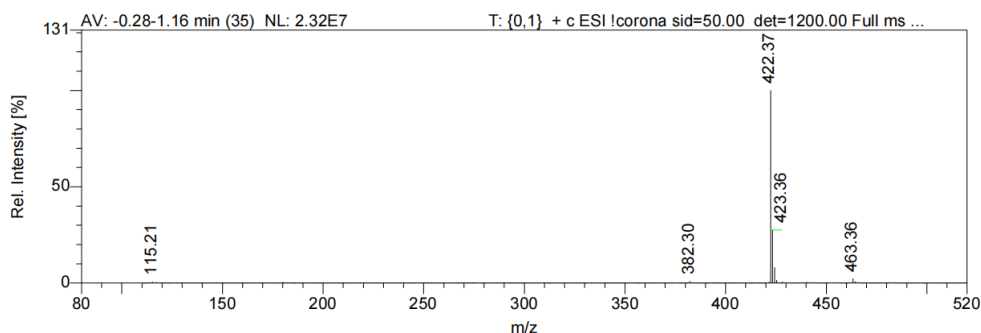


3, C₂₆H₂₉N₆O₈PI [M+H]⁺:
calculated mass: 711.10;
observed mass: 711.30.

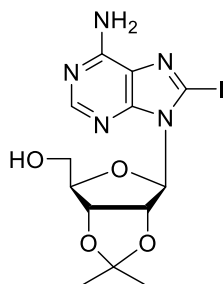
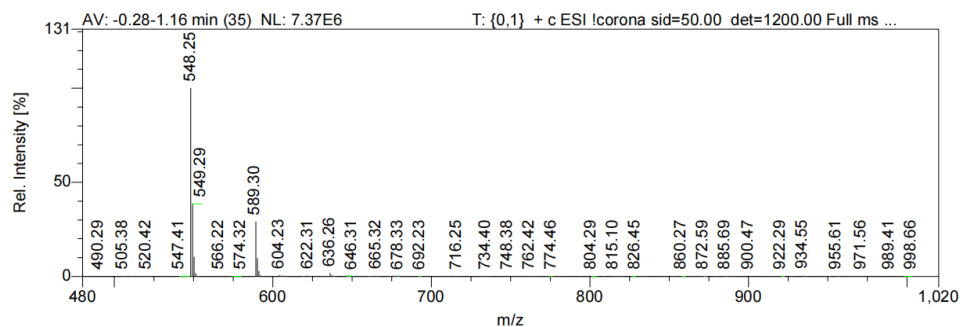




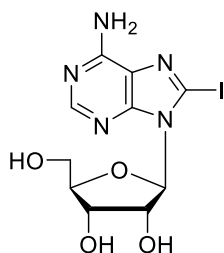
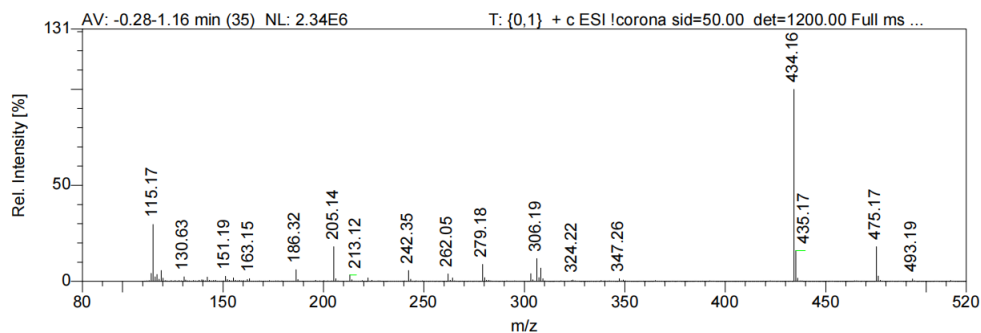
16, C₁₉H₃₂N₅O₄Si [M+H]⁺:
calculated mass: 422.22;
observed mass: 422.37.



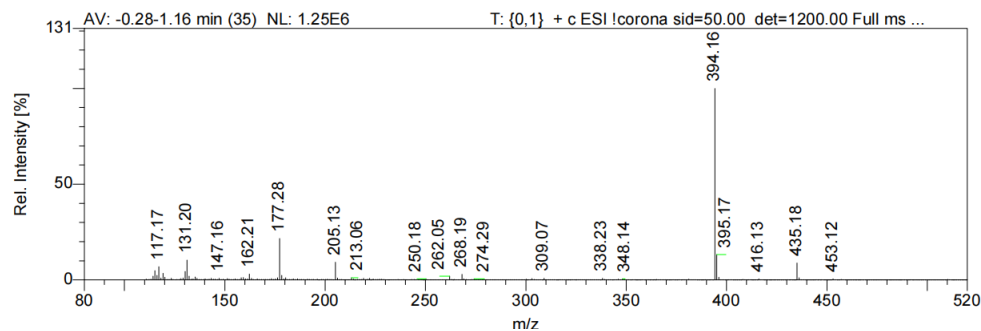
17, C₁₉H₃₁N₅O₄SiI [M+H]⁺:
calculated mass: 548.11;
observed mass: 548.25.

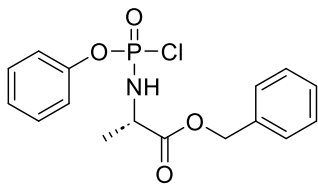


18, C₁₃H₁₇N₅O₄I [M+H]⁺:
calculated mass: 434.04;
observed mass: 434.16.

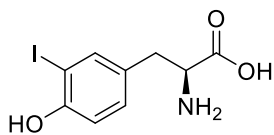
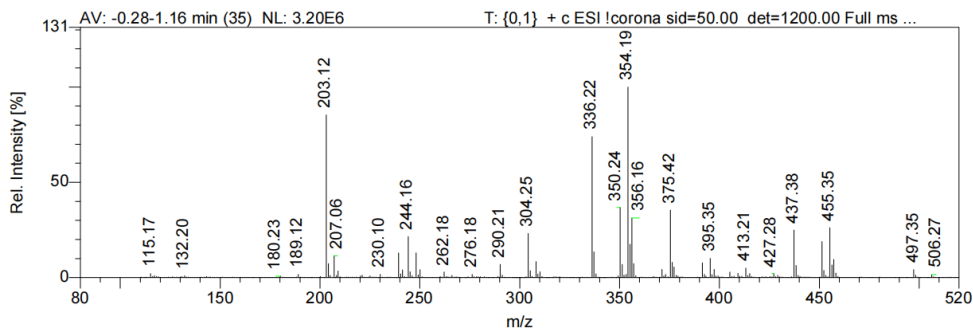


4, C₁₀H₁₃N₅O₄I [M+H]⁺:
calculated mass: 394.06;
observed mass: 394.16.

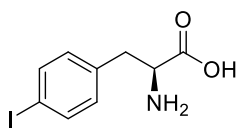
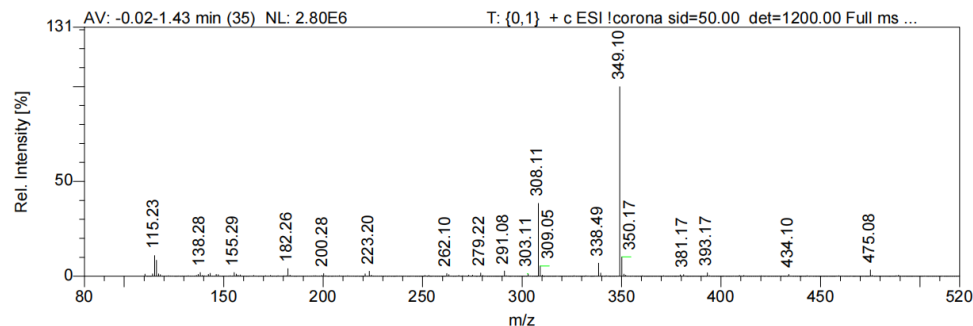




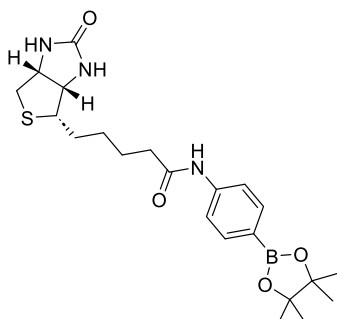
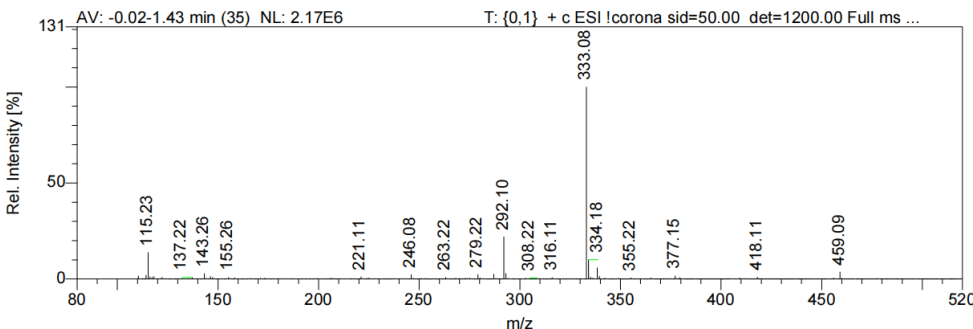
12, C₁₆H₁₈NO₄PCl [M+H]⁺:
calculated mass: 354.08;
observed mass: 354.19.



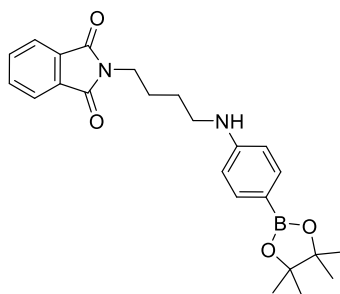
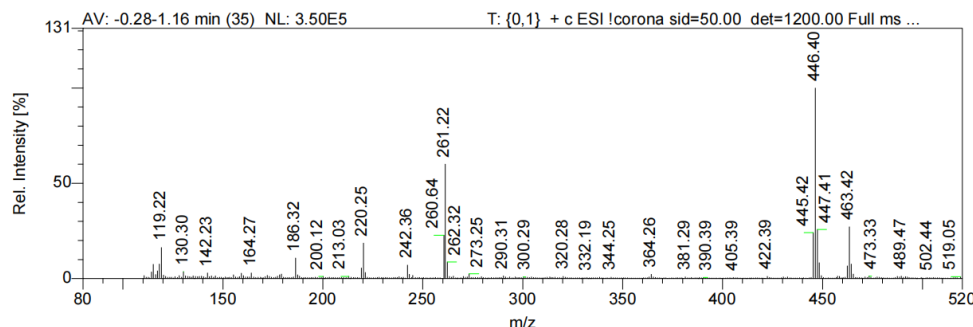
5, C₉H₁₁NO₃ [M+H]⁺:
calculated mass: 307.98;
observed mass: 308.11;
[M+ACN+H]⁺: 349.10.



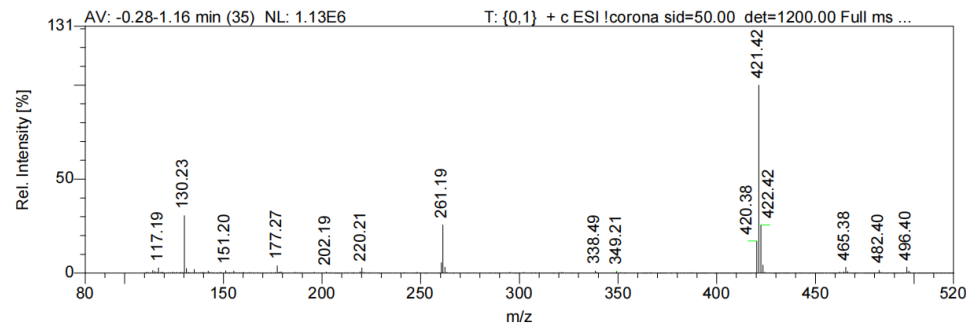
6, C₉H₁₁NO₂I [M+H]⁺:
calculated mass: 291.99;
observed mass: 292.10;
[M+ACN+H]⁺: 333.08.

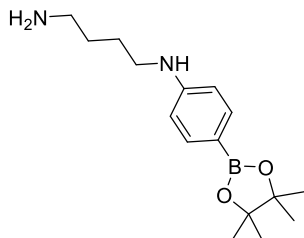


21, C₂₂H₃₃BN₃O₄S [M+H]⁺:
calculated mass: 446.23;
observed mass: 446.40.

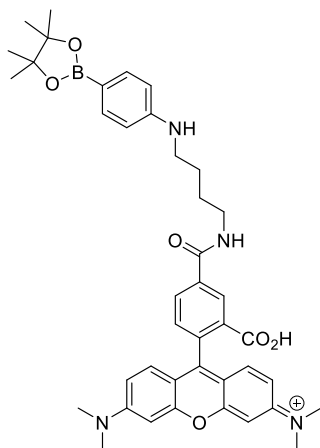
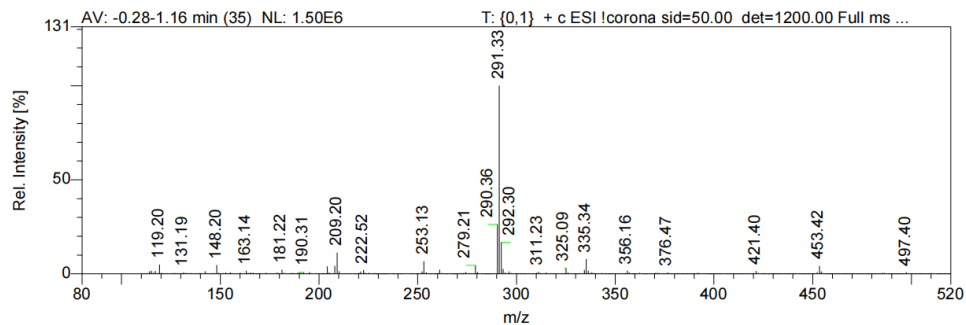


29, C₂₄H₃₀BN₂O₄ [M+H]⁺:
calculated mass: 421.23;
observed mass: 421.42.

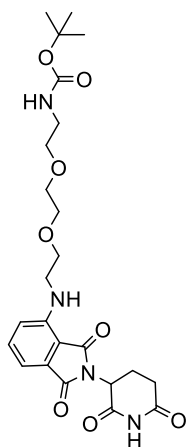
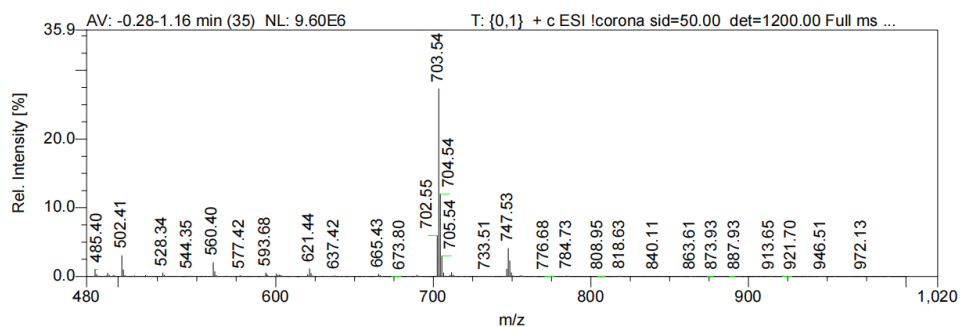




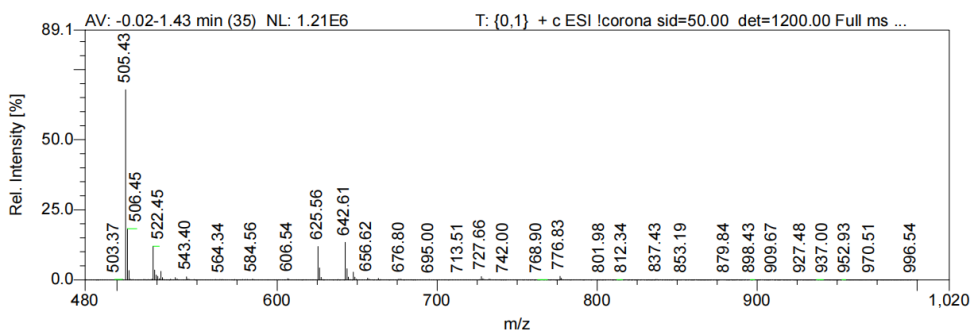
30, C₁₆H₂₈BN₂O₂ [M+H]⁺:
calculated mass: 291.22;
observed mass: 291.33.

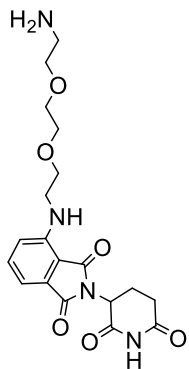


22, C₄₁H₄₈BN₄O₆⁺ [M+H]⁺:
calculated mass: 703.37;
observed mass: 703.54.

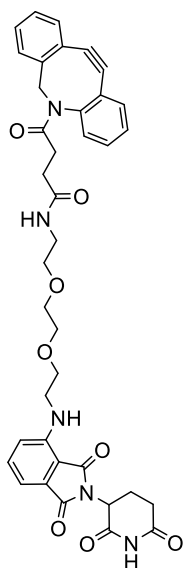
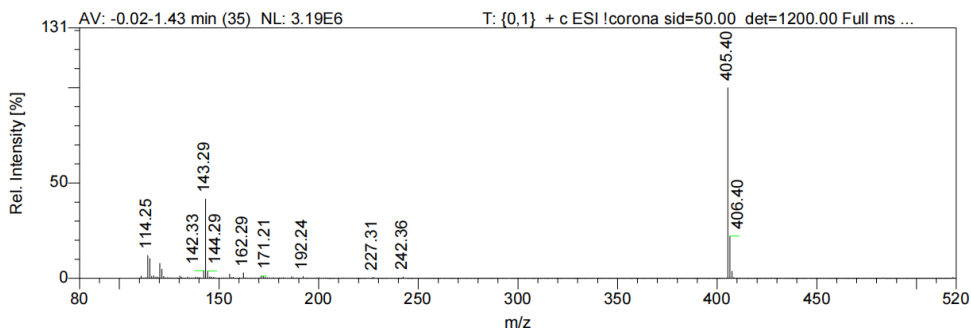


34, C₂₄H₃₃N₄O₈ [M+H]⁺:
calculated mass: 505.23;
observed mass: 505.43.

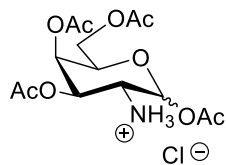
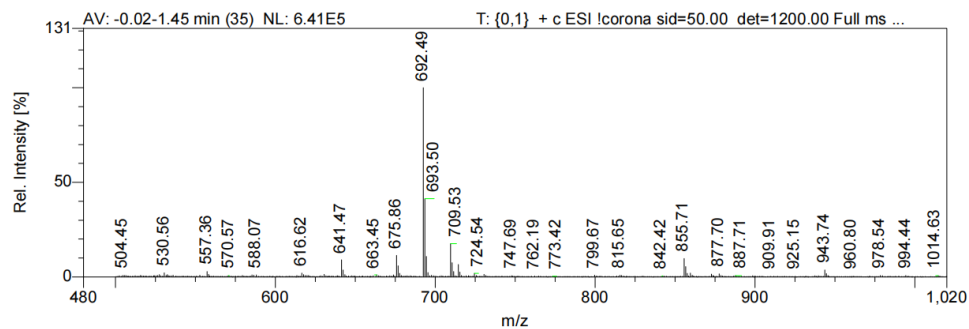




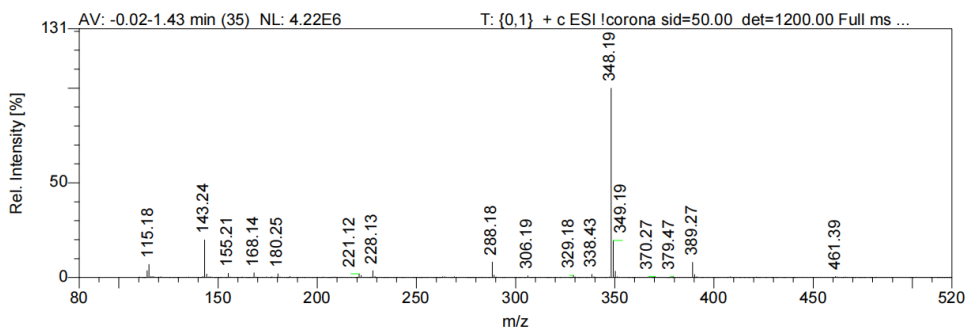
35, $C_{19}H_{25}N_4O_6$ $[M+H]^+$:
 calculated mass: 405.18,
 observed mass: 405.40.



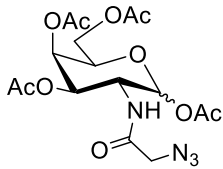
DBC0-Thal,
 $C_{38}H_{38}N_5O_8$ $[M+H]^+$:
 calculated mass: 692.27,
 observed mass: 692.49.



41, $C_{14}H_{22}NO_9$ $[M+H]^+$:
 calculated mass: 348.13,
 observed mass: 348.19.



HCac4GalNAz_20231115043547 #5 RT: 0.20 AV: 1 NL: 4.67E6
T: FTMS + p ESI cv=-21.00 Full ms [400.0000-480.0000]



Ac₄GalNAz,

C₁₆H₂₂N₄O₁₀Na [M+Na]⁺:

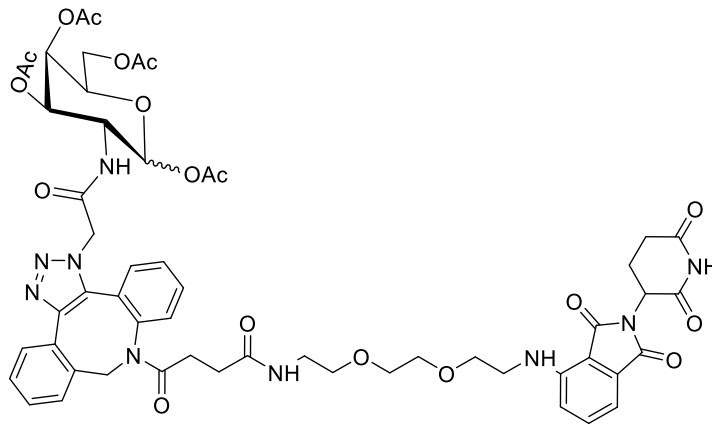
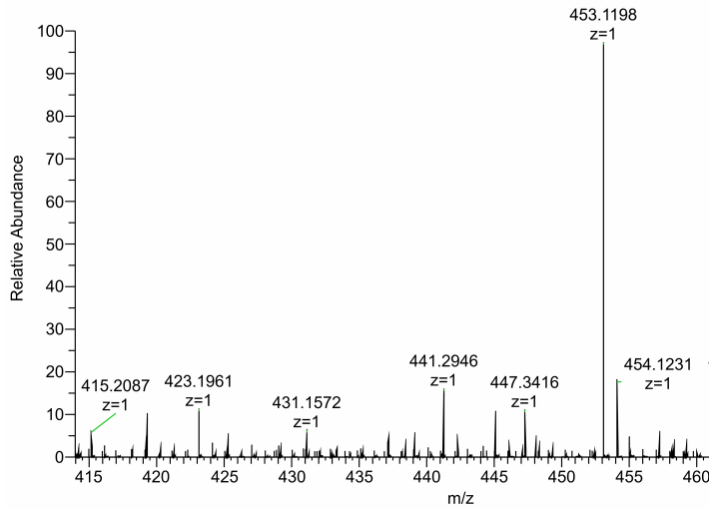
calculated mass:

453.1234;

observed mass: 453.1198

(measured with

high-resolution MS1).



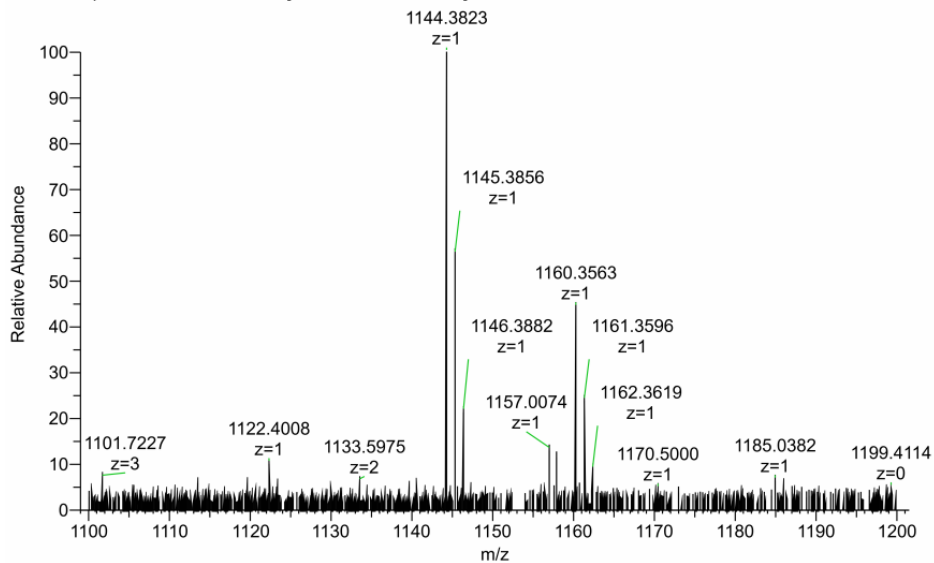
Ac₄GalNAz-DBCO-Thal conjugate,

C₅₄H₅₉N₉O₁₈Na [M+Na]⁺: calculated mass:1144.3870, observed mass: 1144.3823;

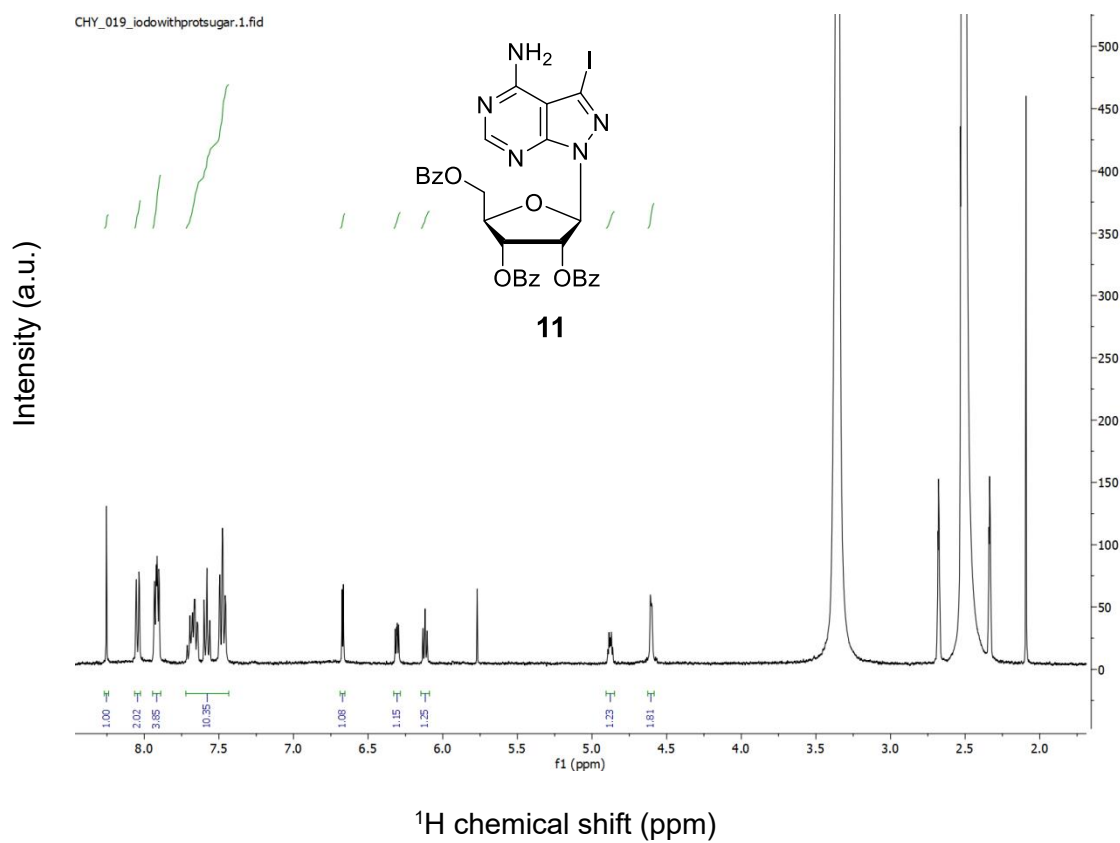
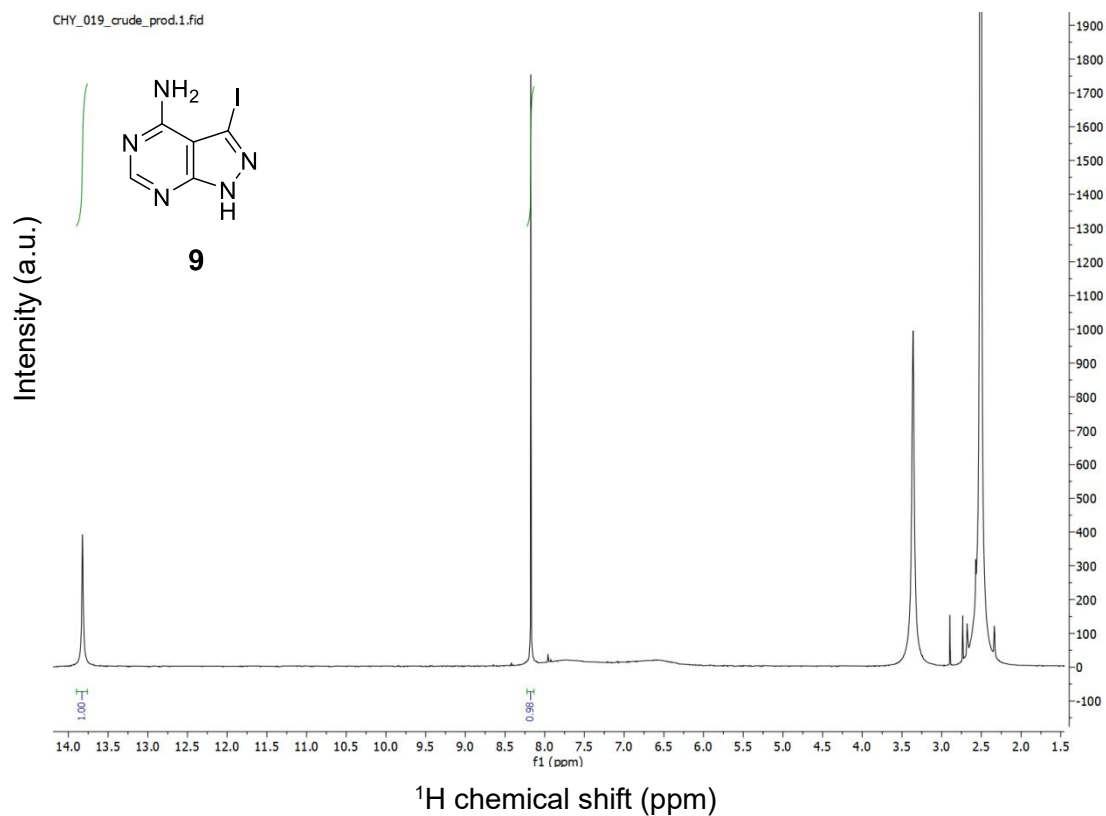
C₅₄H₅₉N₉O₁₈K [M+K]⁺: calculated mass:1160.3610, observed mass: 1144.3563;

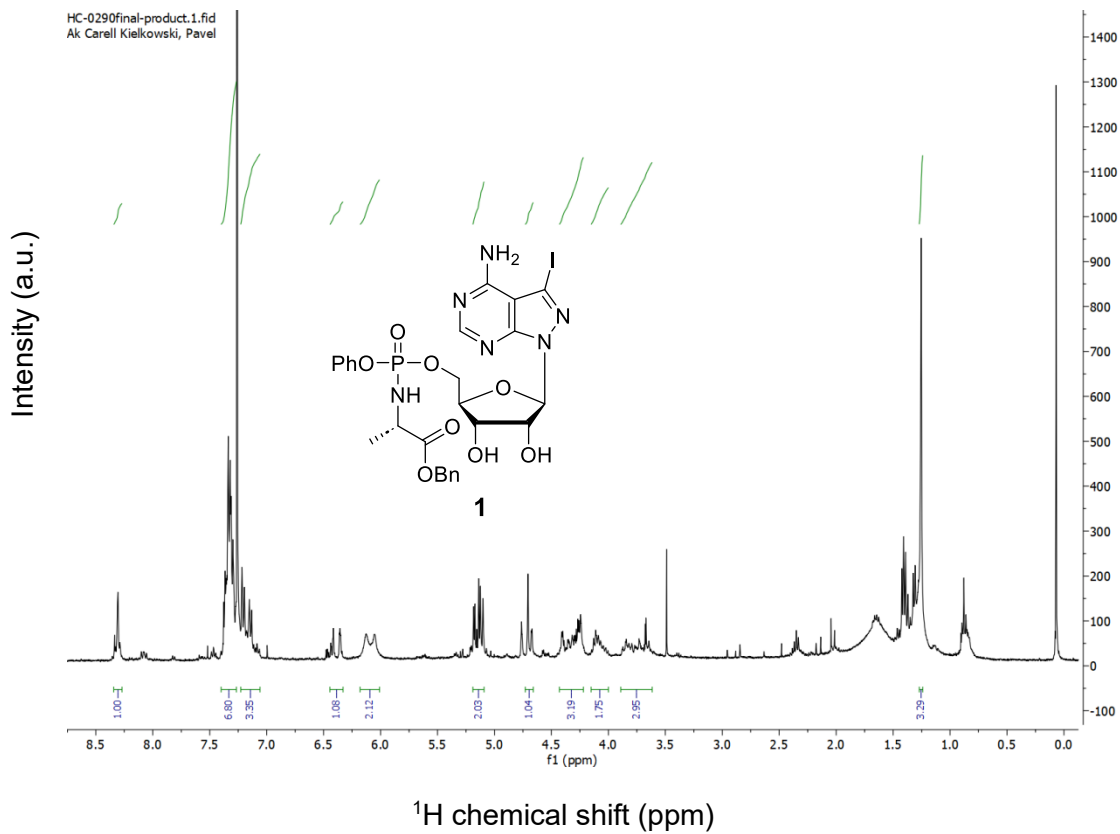
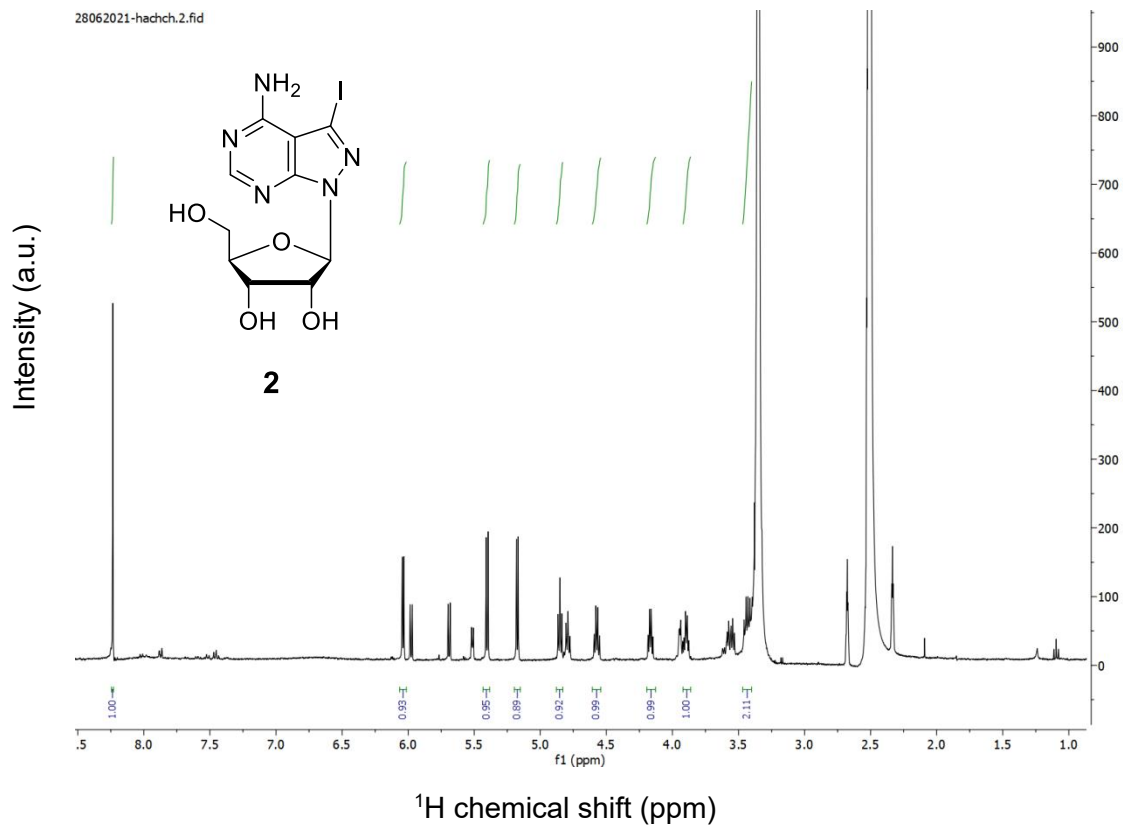
(measured with high-resolution MS1).

HCPROTAC_1140 #3 RT: 0.11 AV: 1 NL: 1.37E5
T: FTMS + p ESI cv=-30.00 Full ms [1100.0000-1200.0000]

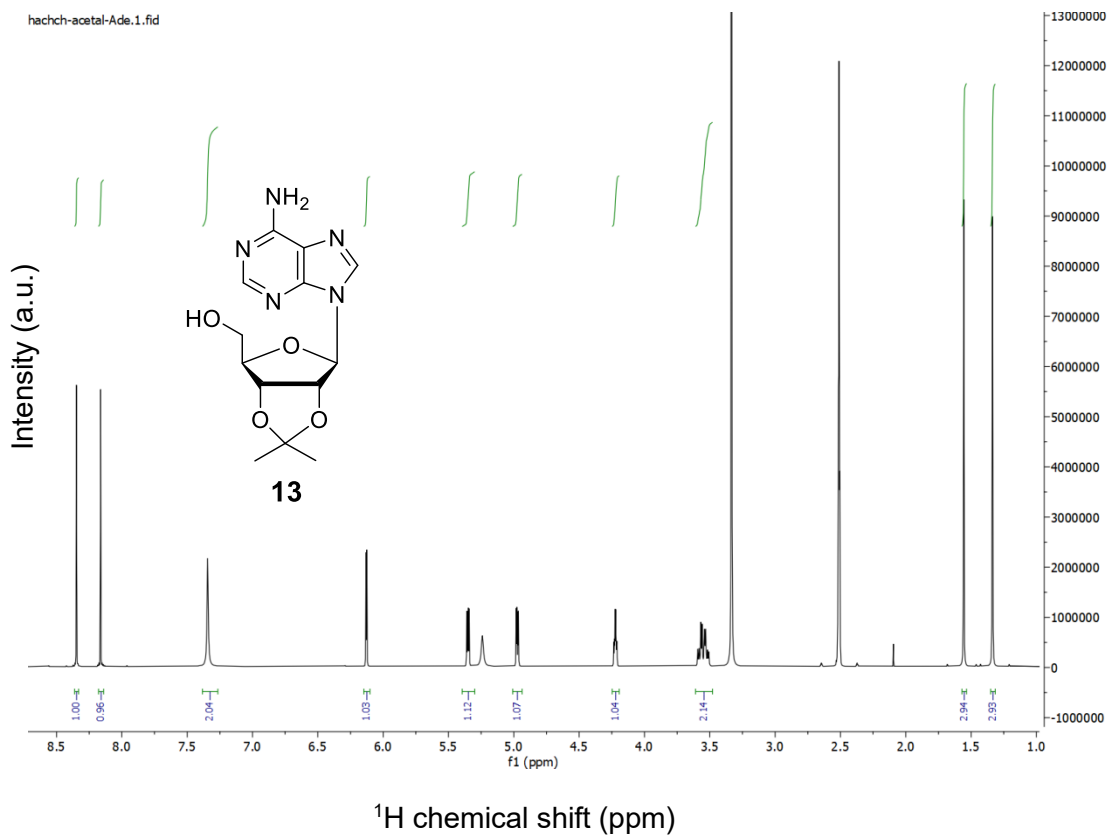
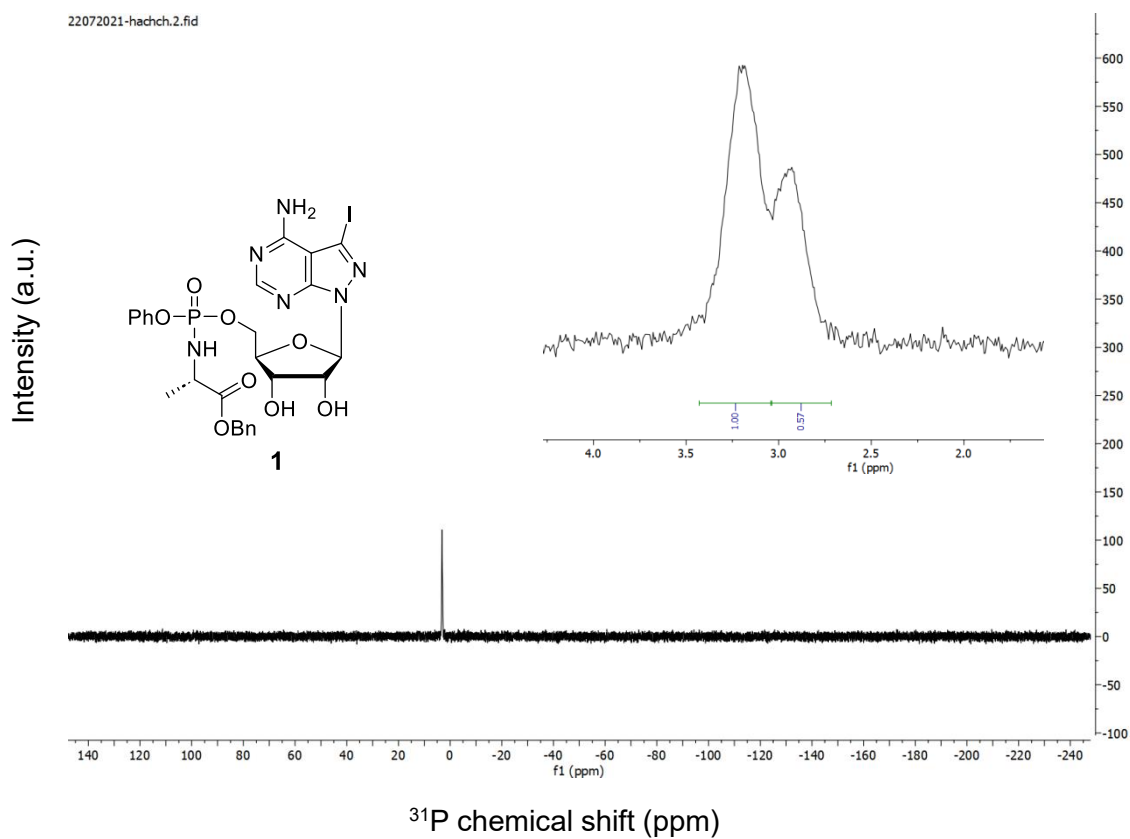


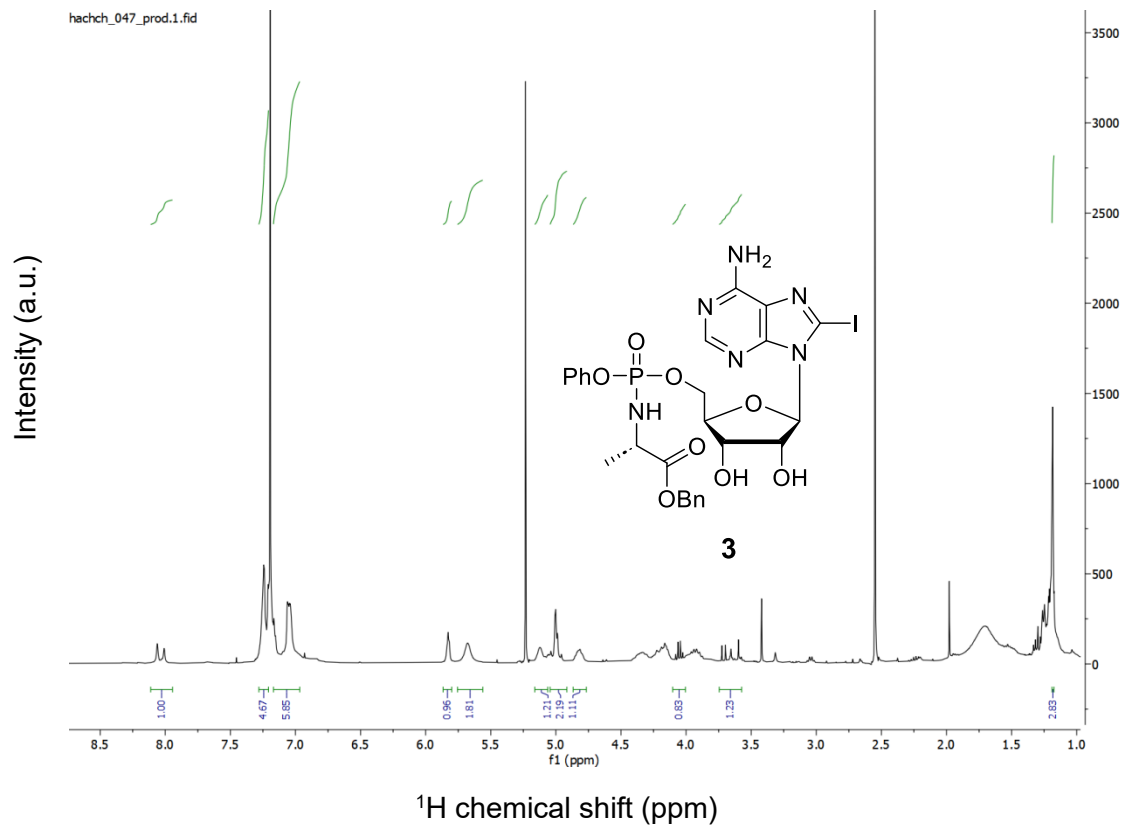
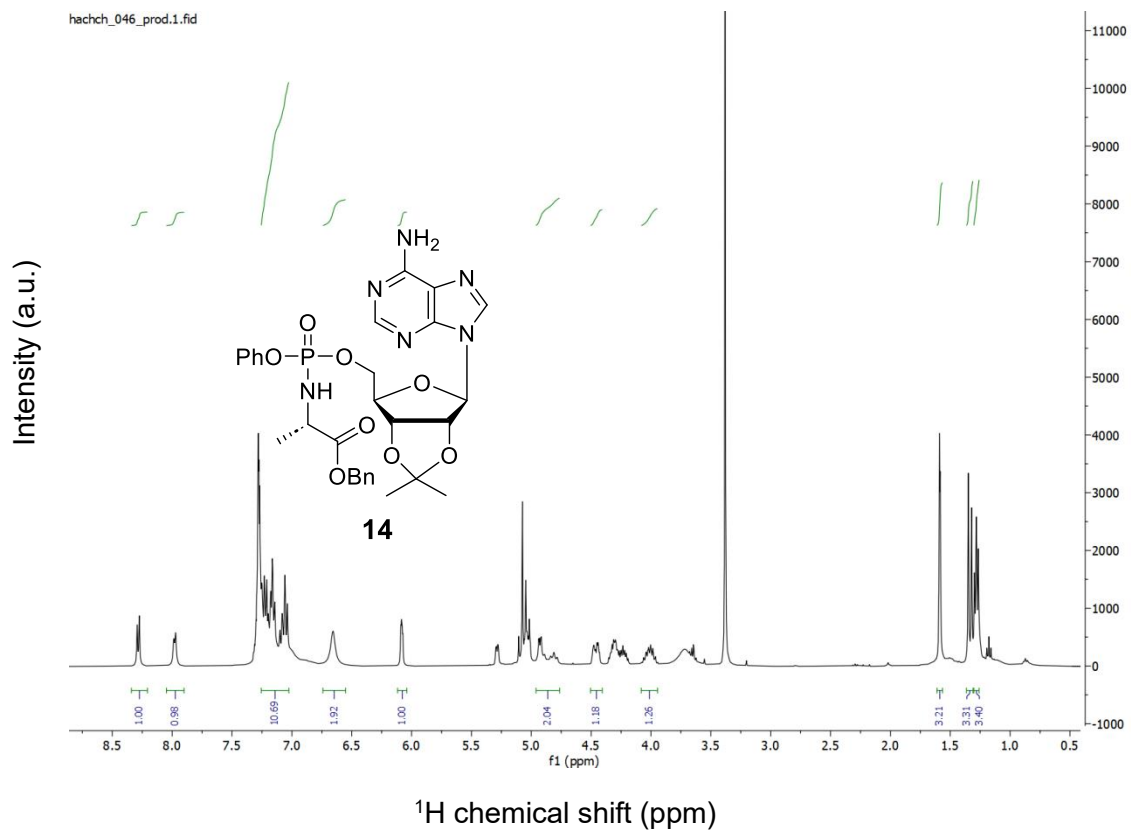
8.5 NMR Spectrums

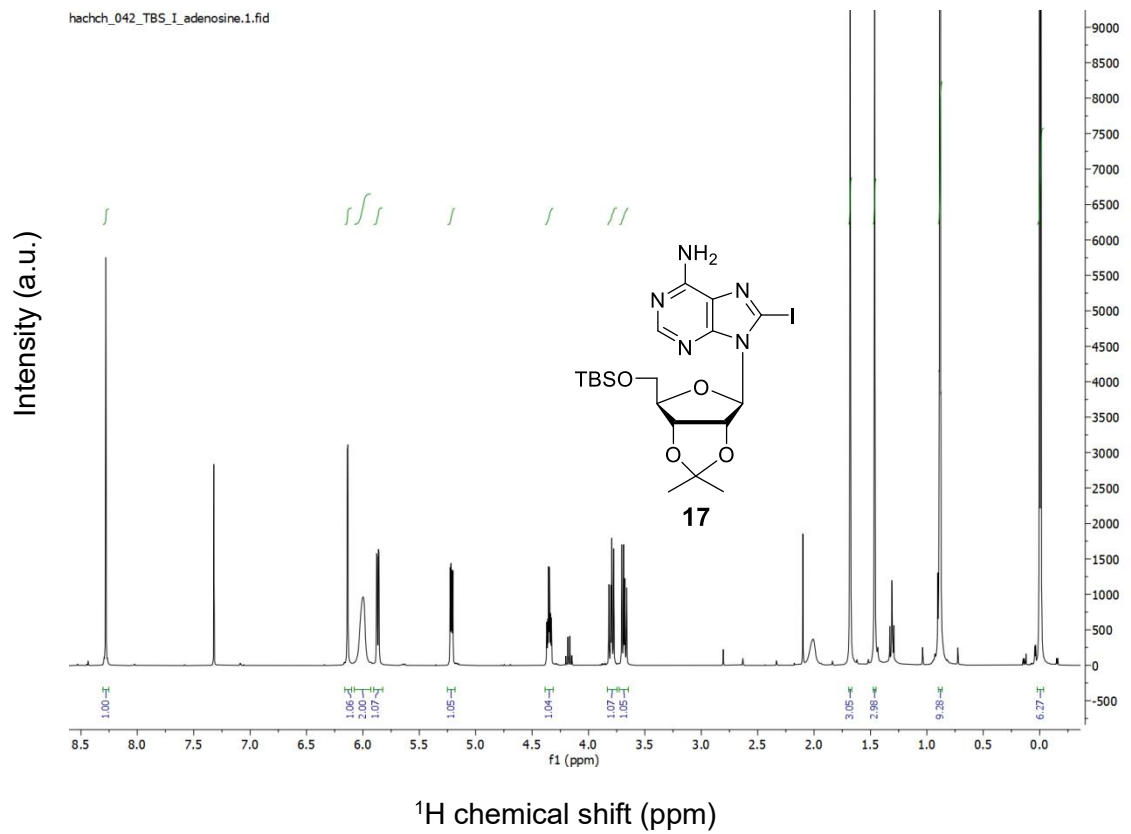
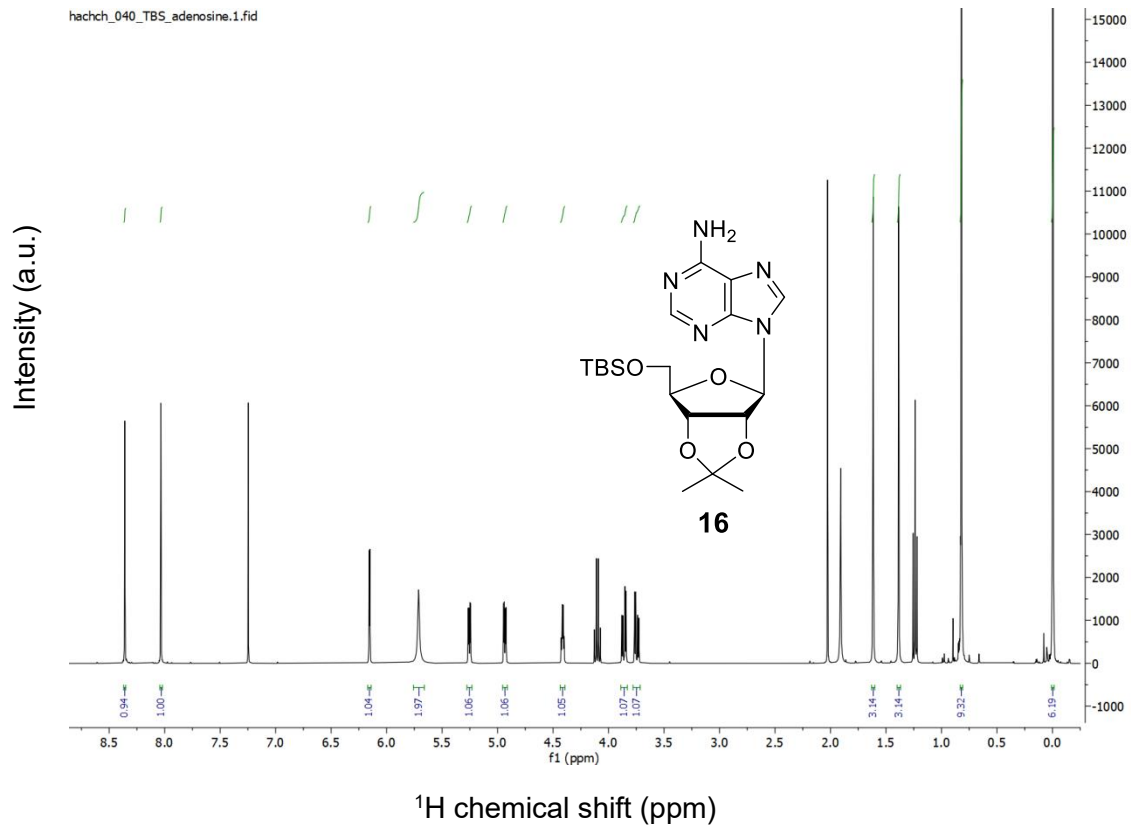


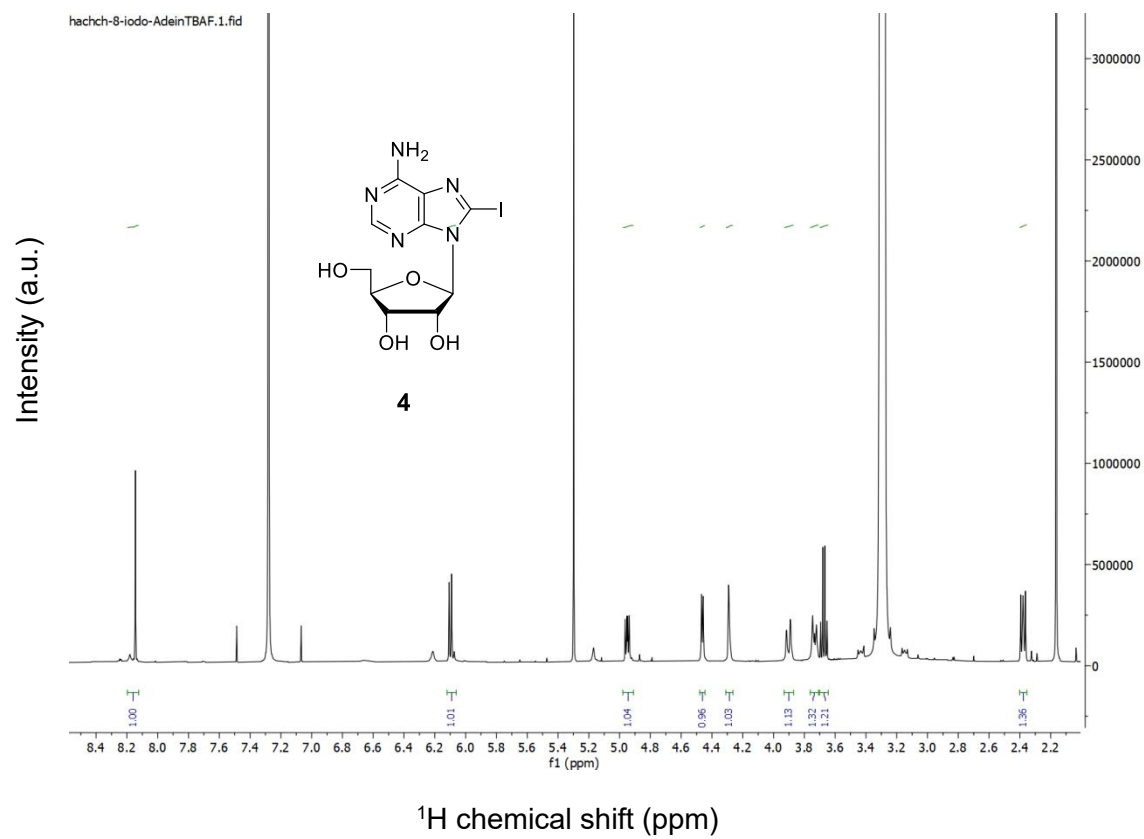
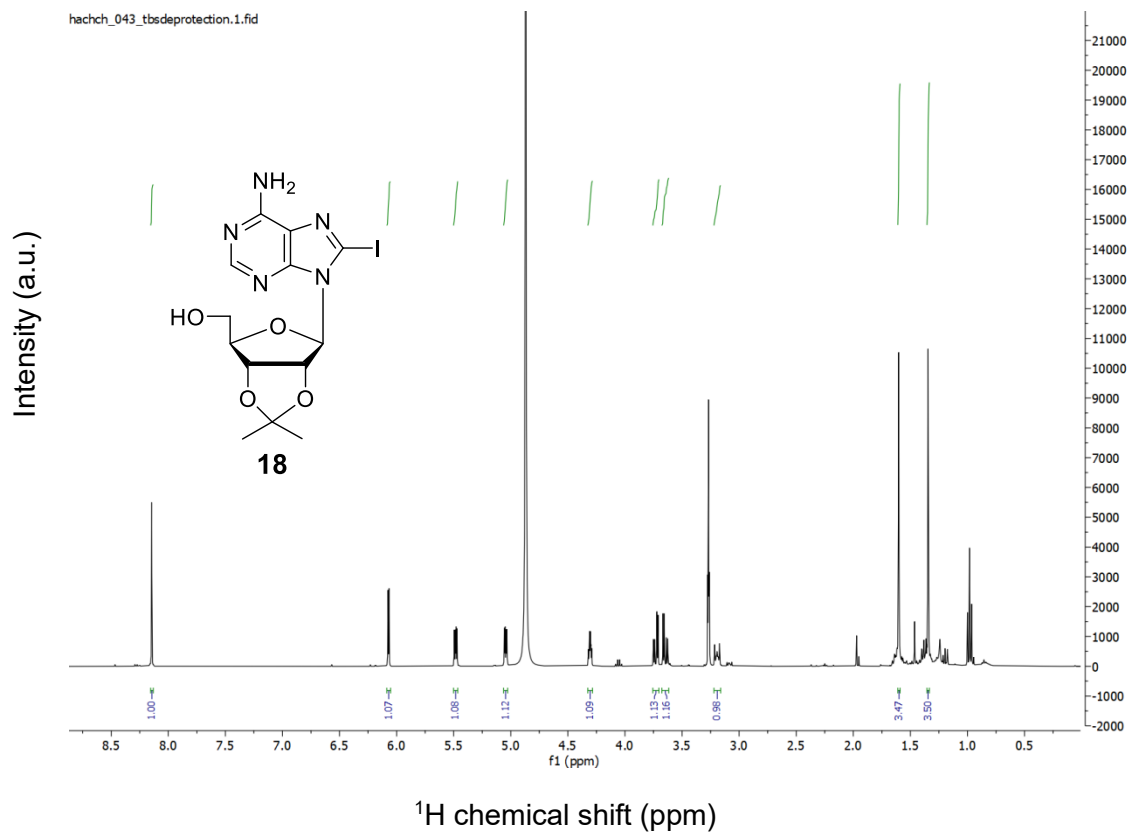


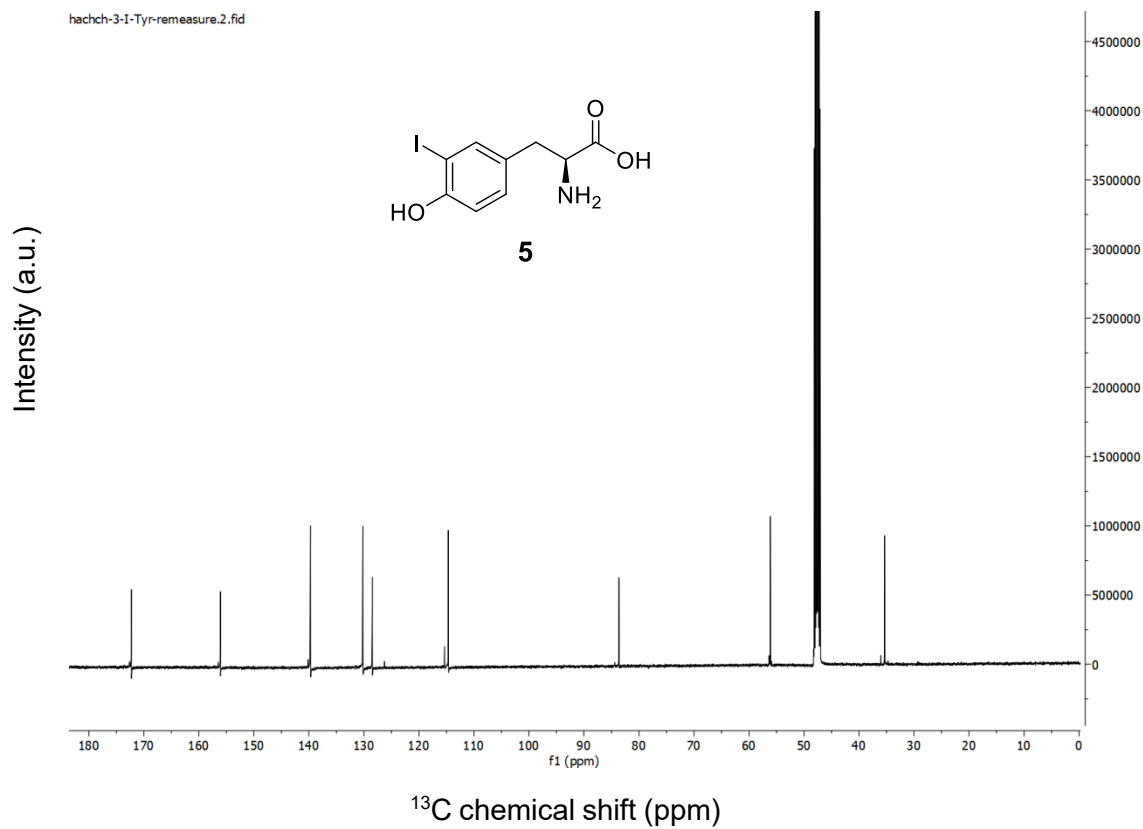
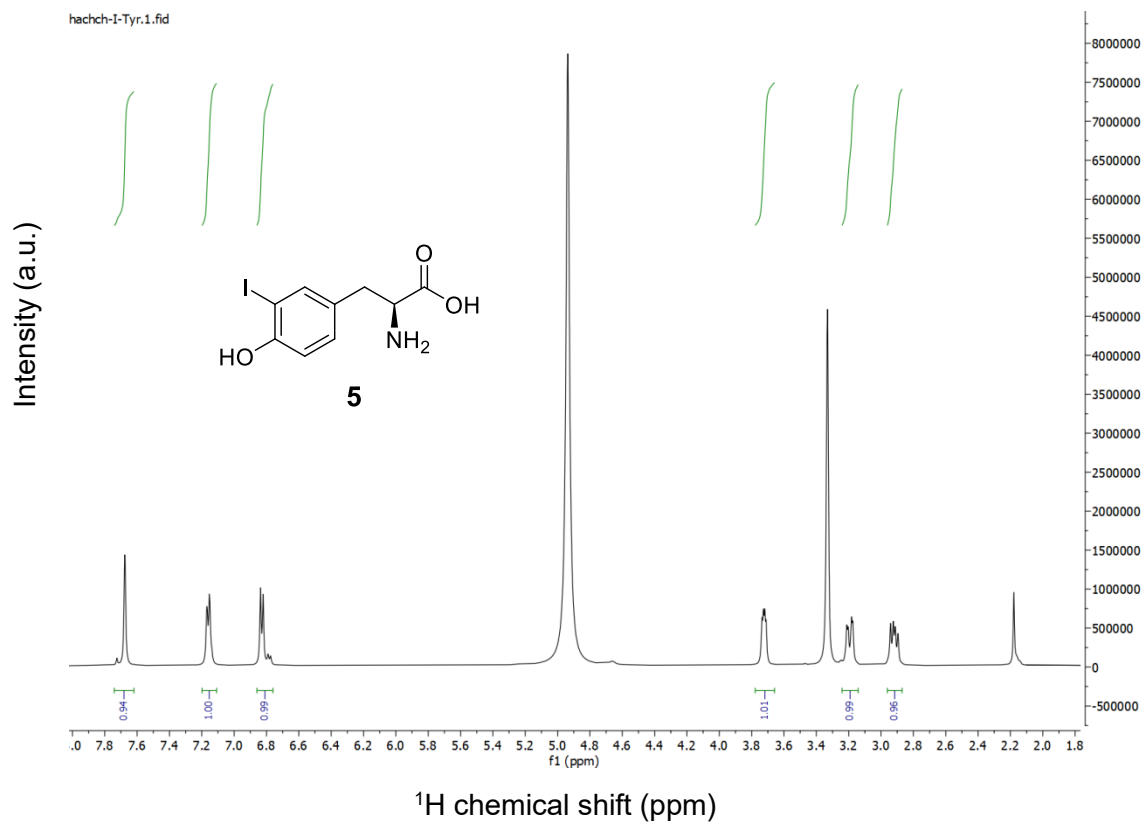
22072021-hachch.2.fid

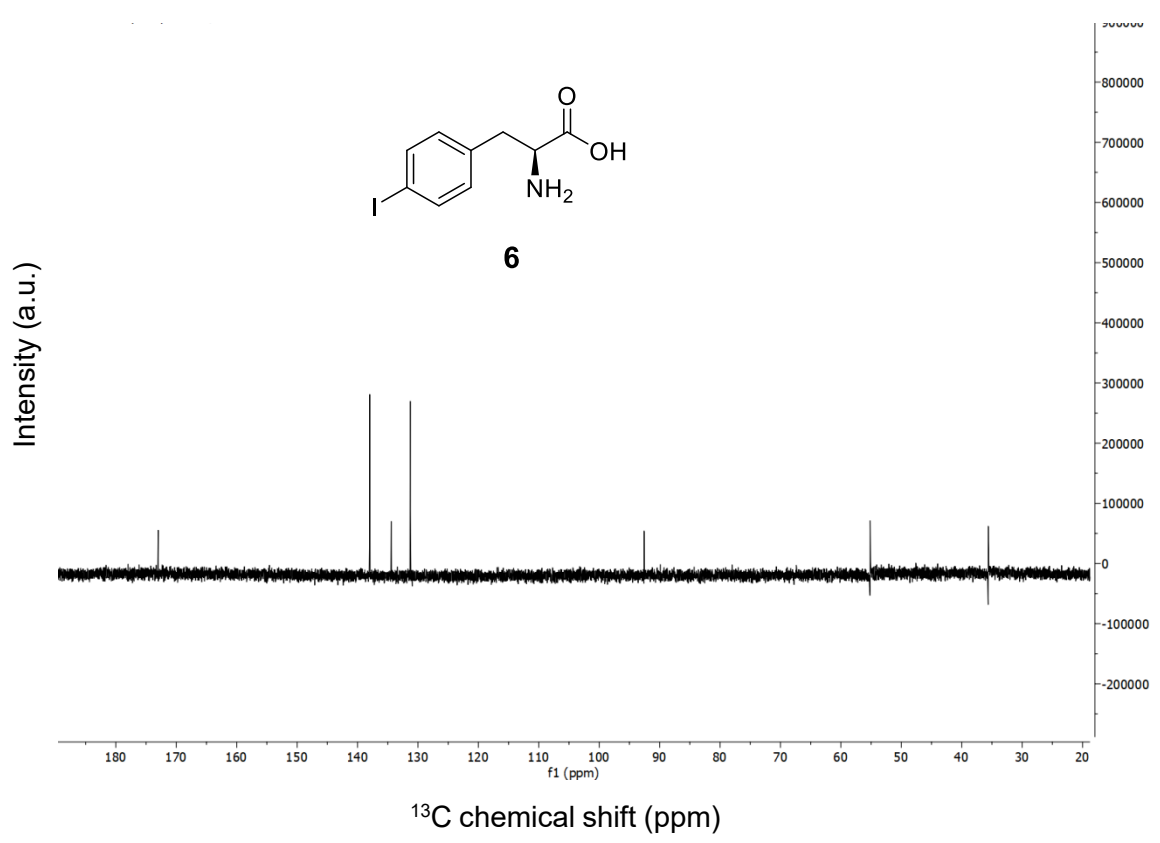
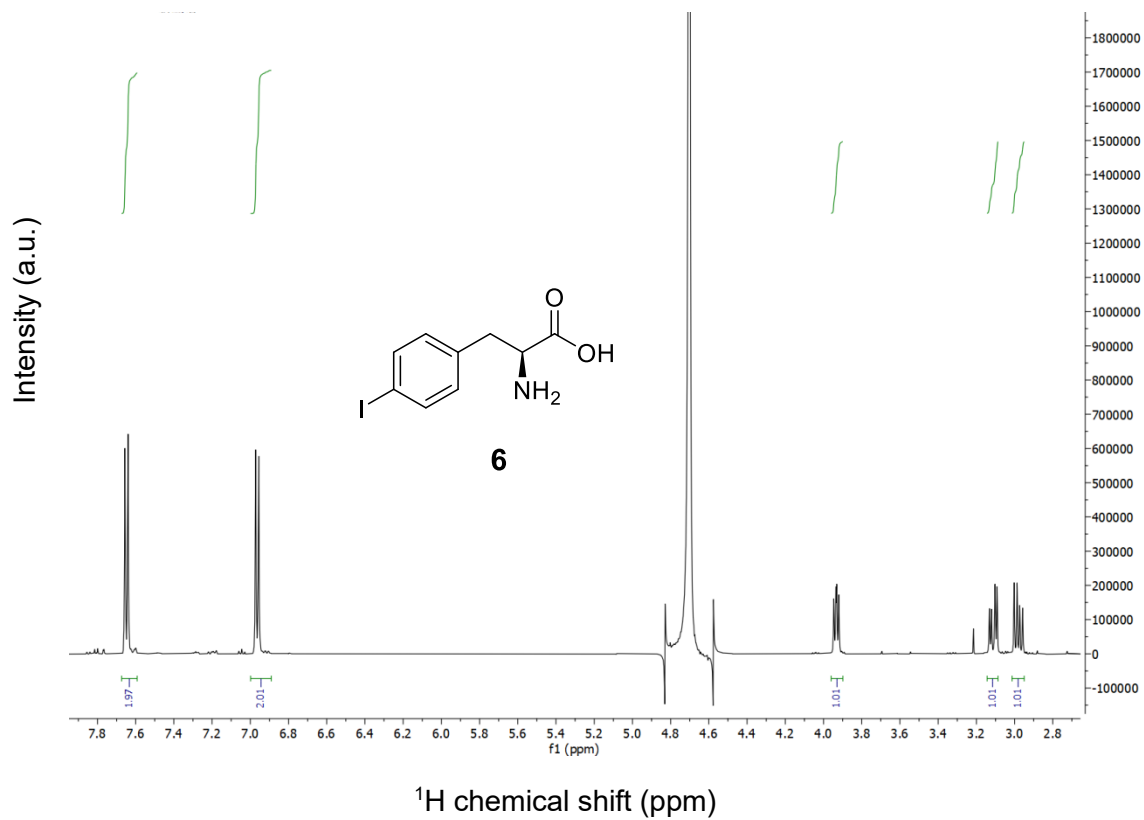


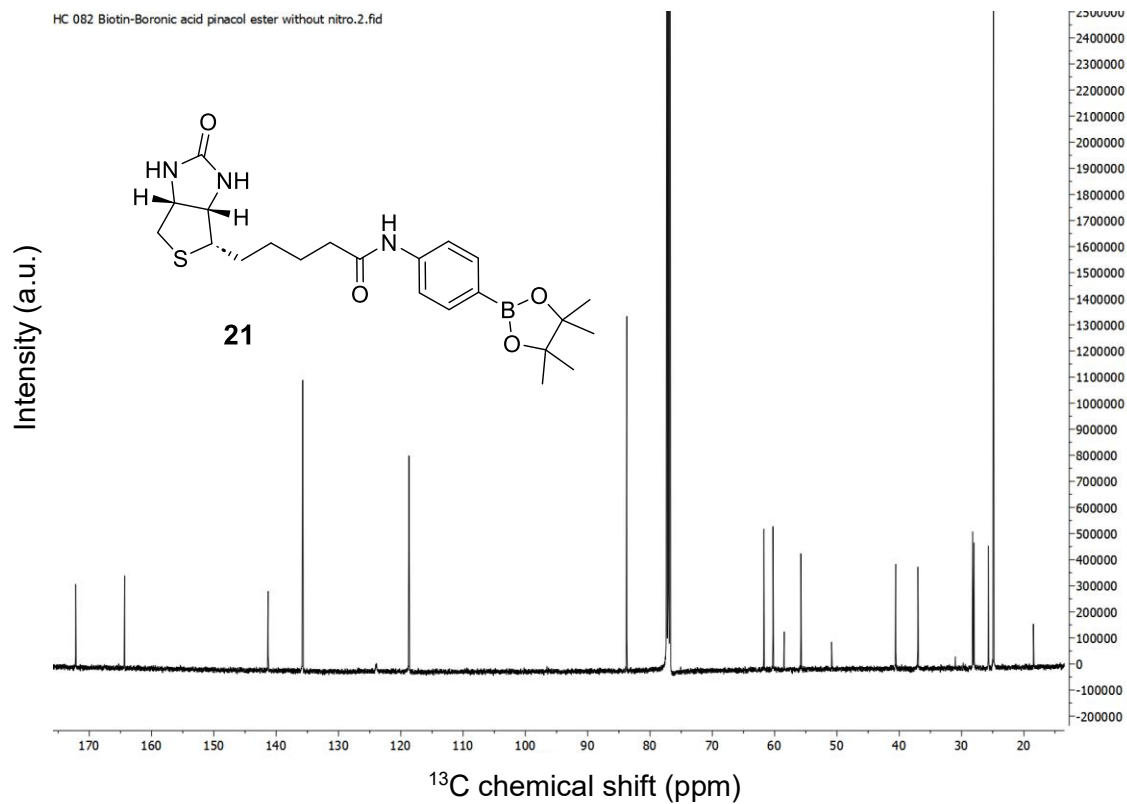
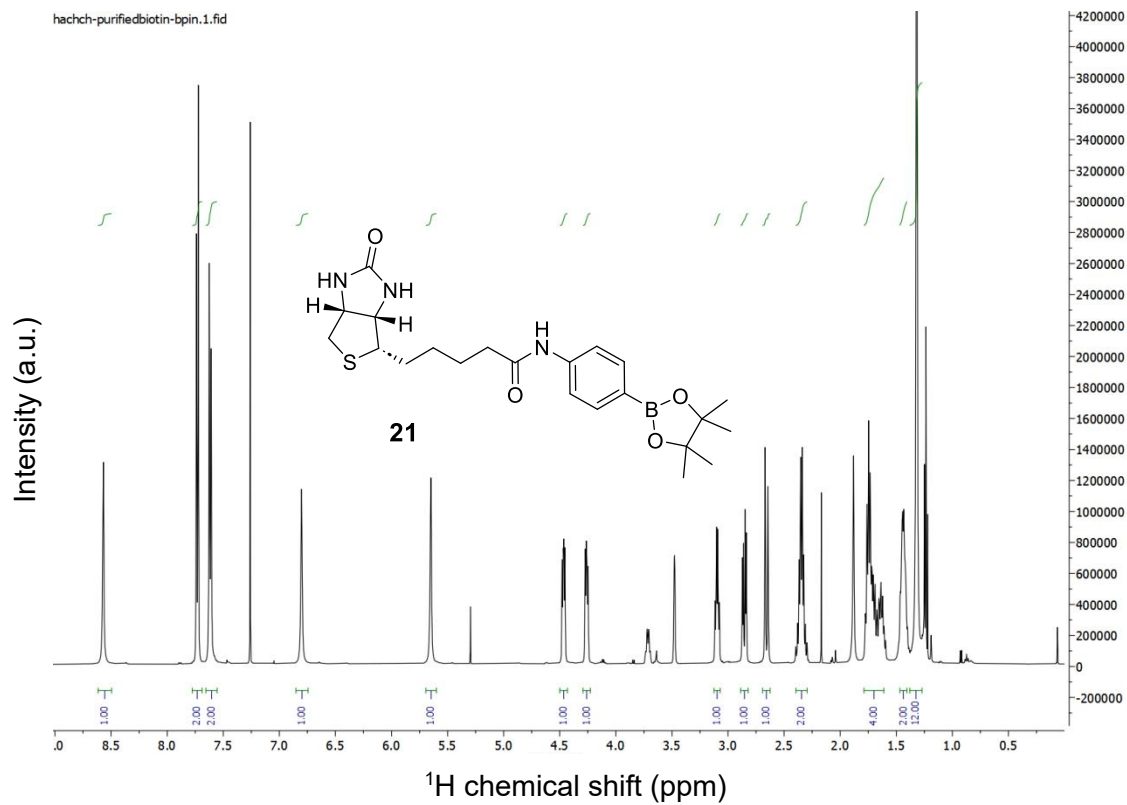


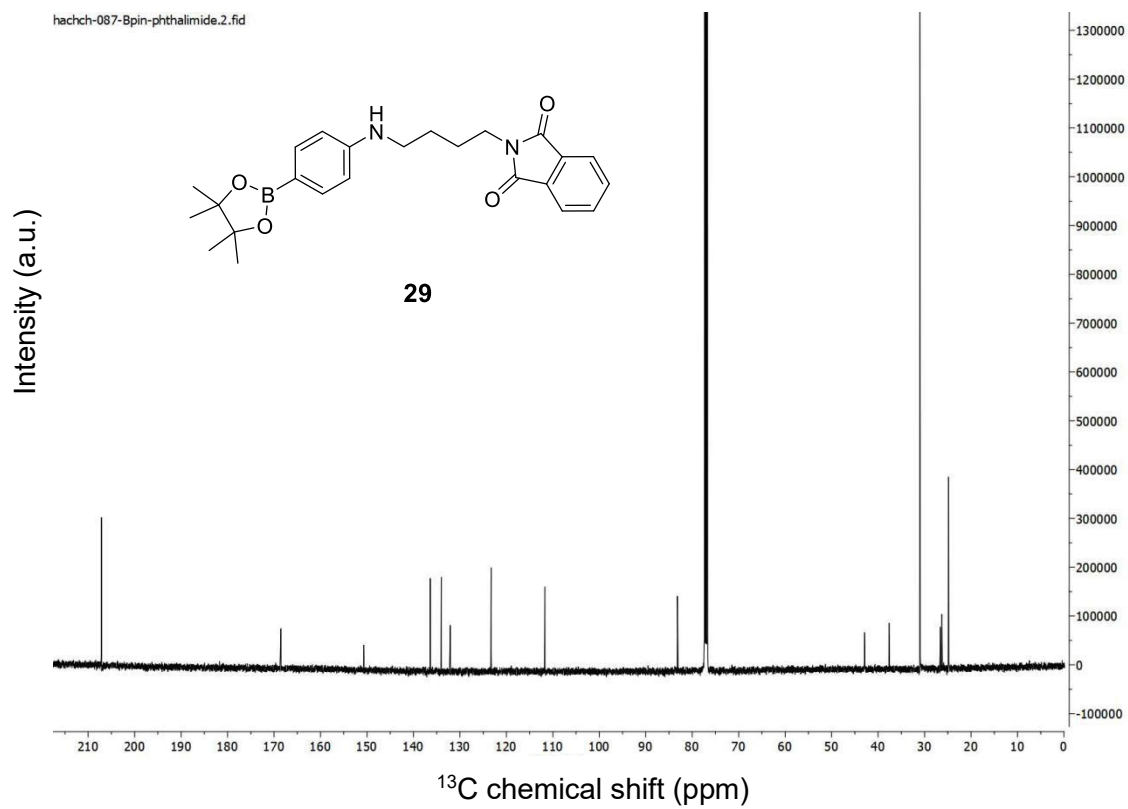
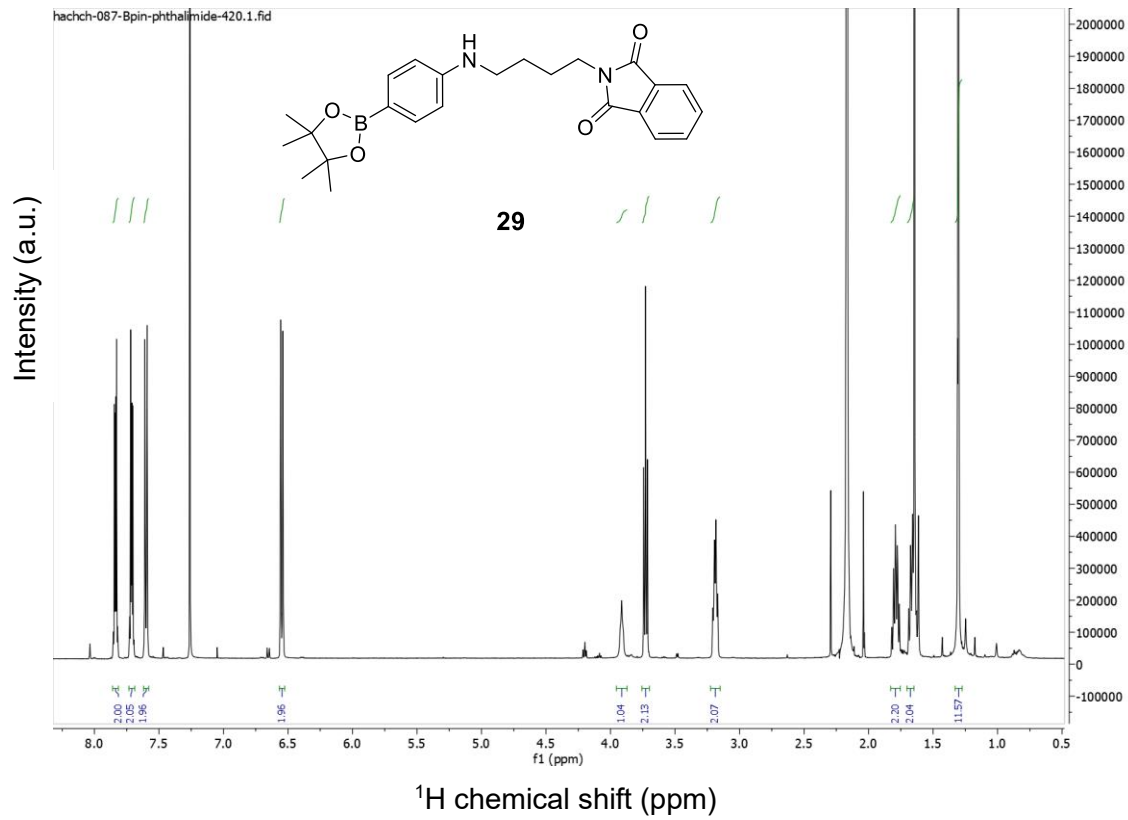


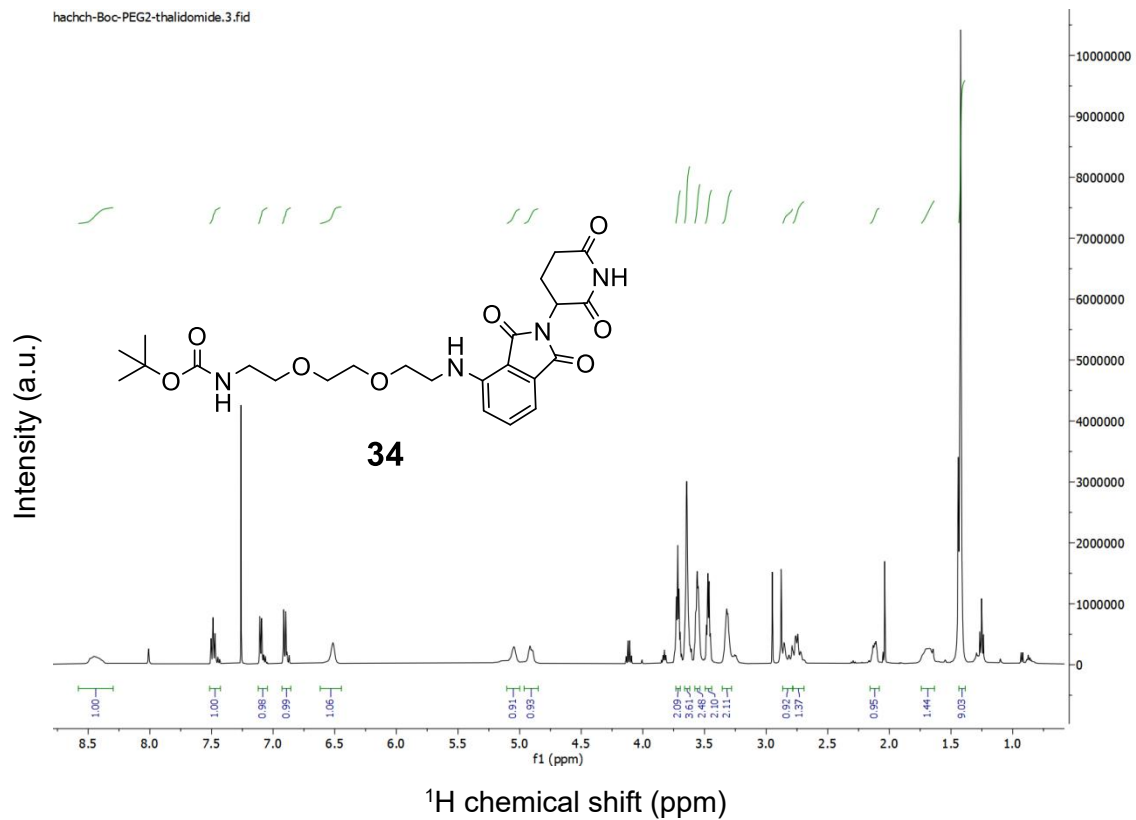
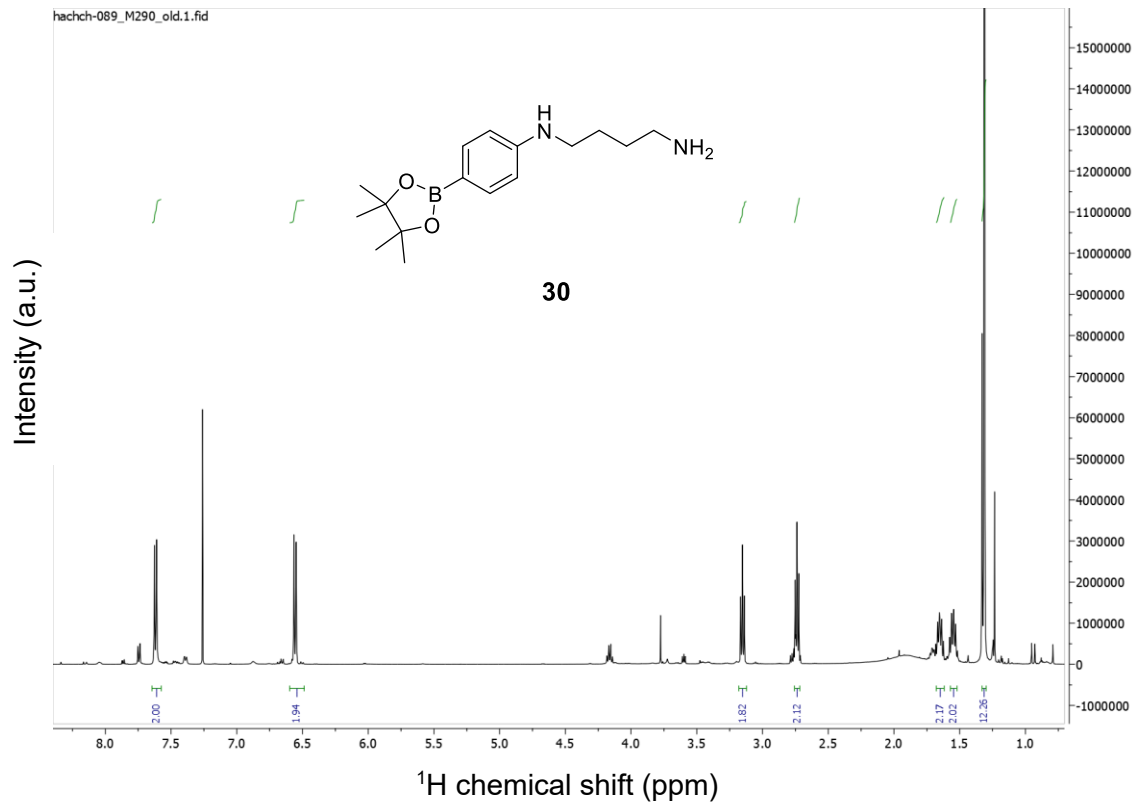


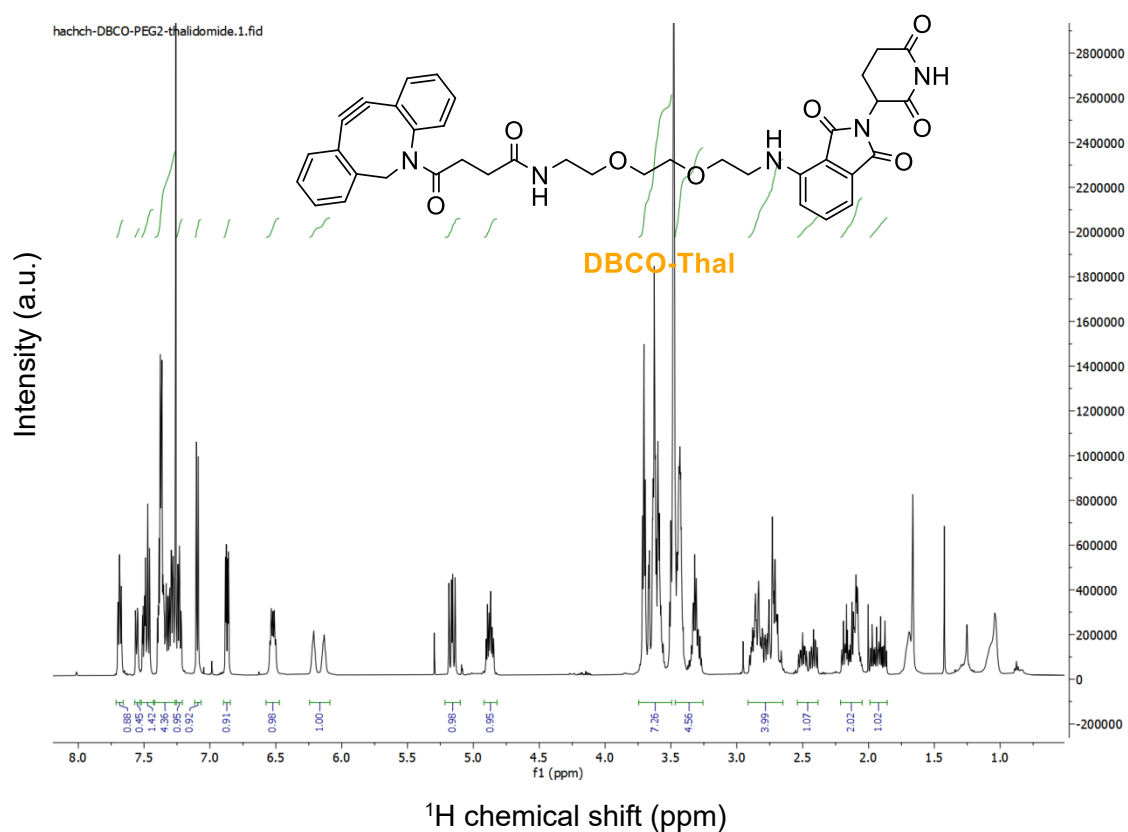
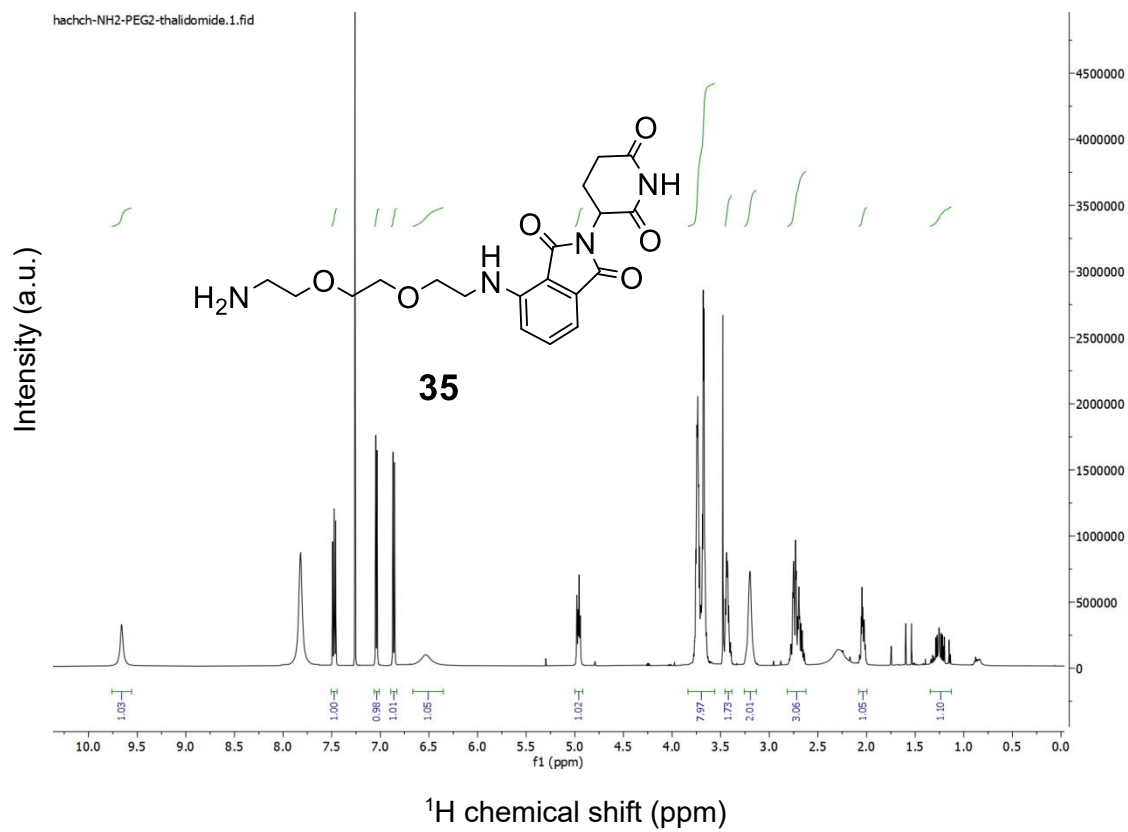


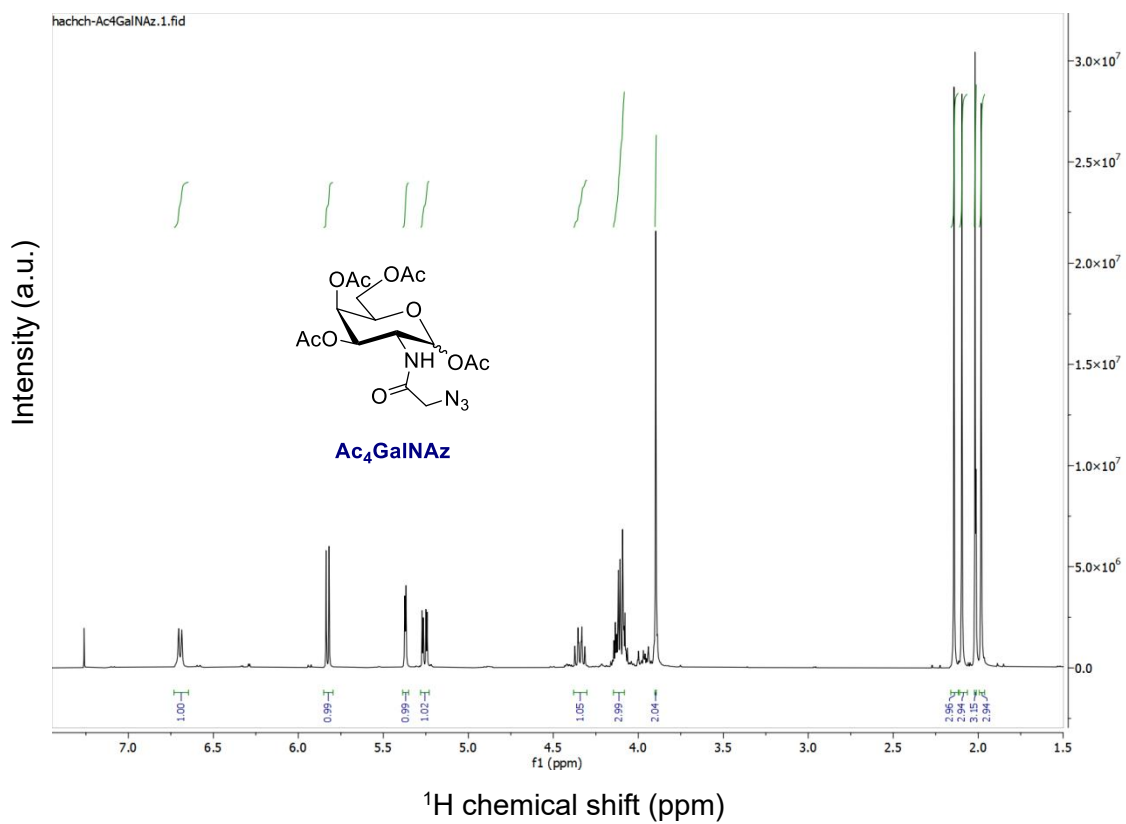
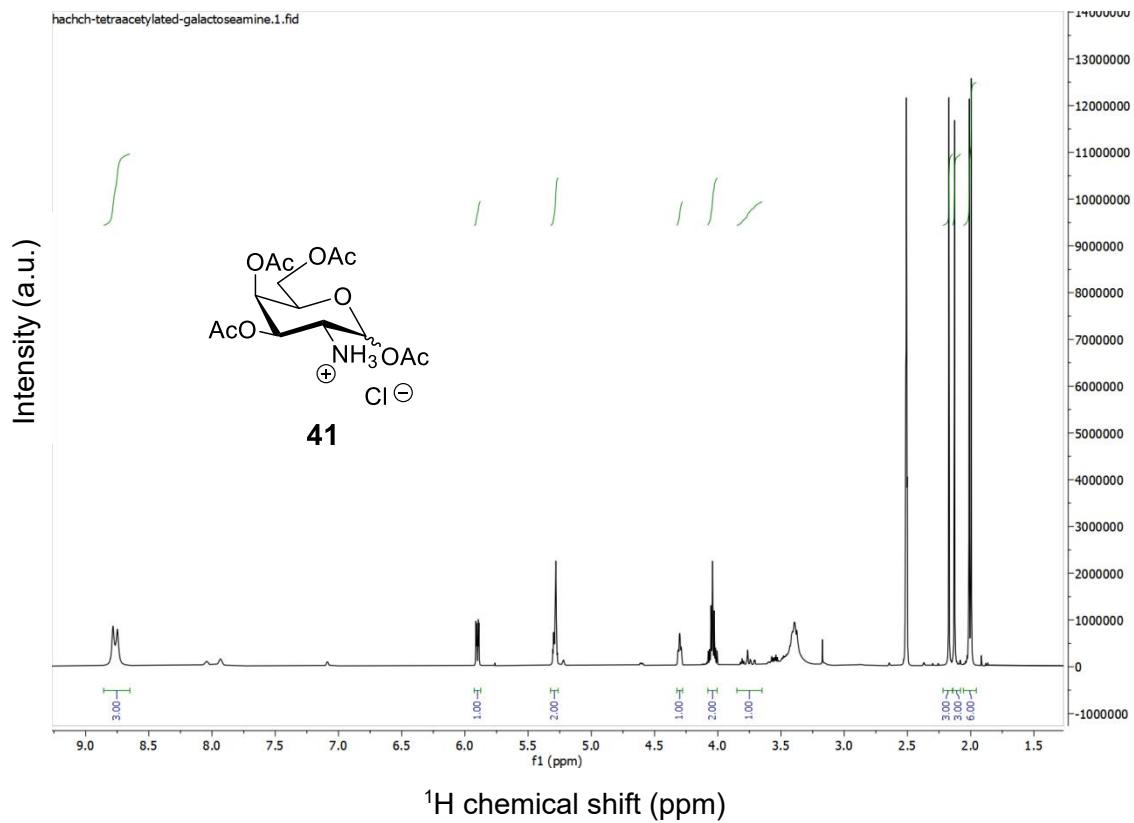












hachchAc4GalNAz.2.fid

